A Nonlinear Structure Tensor with Diffusivity Matrix based on Image

by

Jooyoung Hahn and Chang-Ock Lee

Applied Mathematics
Research Report
08 - 01
January 3, 2008

DEPARTMENT OF MATHEMATICAL SCIENCES

KAIST 한국과학기술원
Korea Advanced Institute of Science and Technology

A nonlinear structure tensor with diffusivity matrix based on image *

Jooyoung Hahn† and Chang-Ock Lee‡

Department of Mathematical Sciences
KAIST, Daejeon, 305-701, Korea

Abstract

We propose a nonlinear partial differential equation (PDE) for regularizing a tensor which contains the first derivative information of an image such as strength of edges and a direction of the gradient of the image. Unlike a typical diffusivity matrix which consists of derivatives of a tensor data, we propose a diffusivity matrix which consists of the tensor data itself, *i.e.*, derivatives of an image. This allows directional smoothing for the tensor along edges which are not in the tensor but in the image. That is, a tensor in the proposed PDE is diffused fast along edges of an image but slowly across them. Since we have a regularized tensor which properly represents the first derivative information of an image, the tensor is useful to improve the quality of image denoising, image enhancement, corner detection, and ramp preserving denoising. We also prove the uniqueness and existence of solution to the proposed PDE.

Keywords: nonlinear diffusion, structure tensor, image denoising, image enhancement, corner detection, ramp preserving denoising.

1 Introduction

The image processing based on partial differential equations (PDEs) has been extensively studied for last 20 years and remarkably successful in problems of computer vision. It starts with an assumption that a gray image is a positive real valued function defined on a rectangular domain even though a digital image is represented by integers from 0 as black to 255 as white. In many low-level topics in computer vision such as image denoising, image enhancement, edge and corner detection, and image segmentation, it is crucial to obtain a regularized derivative information of an image. Since a digital image is represented by integral values, it is difficult to obtain a good approximation of derivatives of the image using a standard finite difference scheme. If a given image has noise, it will be a much harder problem. One of simple solutions is to regularize an image and then differentiate the regularized image. However, in this paper, we use the opposite order of operation; we get derivative information of an image and then regularize it even though the image has noise.

*This work was supported by KRF-2006-311-C00015.

†jyhahn76@amath.kaist.ac.kr

‡colee@kaist.edu

In other words, we propose a noble PDE for regularizing a tensor which contains the first derivative information of an image such as strength of edges and a direction of the gradient of the image.

The Perona-Malik (PM) model [1,2] has been a fundamental frame for adaptive smoothing process based on a nonlinear PDE. Let $I_0 \colon \Omega \subset \mathbf{R}^2 \to \mathbf{R}^+$ be an initial noisy image and $h \colon \mathbf{R}^+ \to \mathbf{R}^+$ be an weight function with a regularization factor ϵ:

$$h\big(s^2\big) = \frac{1}{\sqrt{\epsilon^2 + s^2}}. \tag{1.1}$$

As time evolves, the PM model generates regularized images $I(x,t)$ which satisfy the PDE

$$\begin{aligned} \frac{\partial I}{\partial t}(x,t) = \nabla \cdot \Big(h\Big(\big|\nabla I_\sigma\big|^2\Big)\nabla I\Big) \quad &\text{in} \quad \Omega \times (0,T_1], \\ \nabla I(x,t) \cdot \mathbf{n} = 0 \quad &\text{on} \quad \partial\Omega \times (0,T_1], \\ I(x,0) = I_0(x) \quad &\text{on} \quad \Omega, \end{aligned} \tag{1.2}$$

where $\mathbf{n}$ is a normal vector to $\partial\Omega$ and $I_\sigma \equiv G_\sigma * I$ is the convolution of I with the two-dimensional Gaussian kernel with a standard deviation σ. The PM model uses regularized strength of edges as $|\nabla I_\sigma|$ which makes an adaptive smoothing process. Note that I_σ is an isotropically smoothed image since the Gaussian convolution is equivalent to solve an isotropic linear heat equation. If an initial image is highly noisy, we need to use large σ in order to obtain reliable information of strength of edges. However, the regularized strength of edges with large σ is smeared out and it fails to make effective adaptive smoothing in order to keep edges and corners in an original image. Even though we take small σ, large end time T_1 is needed to obtain a regularized image which has visually small amount of noise. It is not guaranteed that edges and corners are preserved with large end time T_1. Moreover, it is hard to take a proper σ in order to obtain good regularized strength of edges because it is hard to measure an amount of noise in practice.

Weickert [3] proposed the coherence-enhancing diffusion based on a diffusivity matrix which explicitly represents directional smoothing. The diffusivity matrix is obtained by a structure tensor [4–6]

$$G_\sigma * \big(\nabla I_\rho \nabla {I_\rho}^{\mathrm{T}}\big), \tag{1.3}$$

where T is the transpose and $(G_\sigma * M)_{ij} = G_\sigma * m_{ij}$ for a matrix $M = (m_{ij})$. The noise scale ρ is determined by an amount of noise in an initial image and the integration scale σ reflects a size of neighborhood for a local structure analysis. The structure tensor has a remarkable feature in obtaining regularized strength of edges when ρ is taken small enough to compute ∇I_ρ. While the regularized strength of edges in the PM model is obtained by regularizing an image and then differentiating the regularized image, it is obtained in a structure tensor by differentiating an image and then regularizing derivatives of the image. Changing the order of two operations, the structure tensor has better representation of flow-like structures in an image [3]. However, it still has the same problem of choosing a value σ as in (1.2) when an initial image is highly noisy.

From two PDE-based image denoising models in the above, we notice that, for the purpose of adaptive smoothing of an image, it is crucial to obtain regularized strength of

edges which properly represents local structures of the image such as edges and corners robust to a change of an amount of noise in the image. Moreover, if we have a regularized direction of the gradient of an image which properly represents the orthogonal orientation to edges in the image, it will improve the quality of image denoising based on an anisotropic diffusion.

In this paper, we propose a nonlinear PDE for regularizing a tensor which contains the first derivative information of an image such as strength of edges and a direction of the gradient of the image. Unlike a typical diffusivity matrix which consists of derivatives of a tensor data [7–13], we propose a diffusivity matrix which consists of the tensor data itself, *i.e.*, derivatives of an image. This allows directional smoothing of a tensor along edges which are not in the tensor but in the image. That is, the tensor is diffused fast along edges of an image but slowly across them. It explains that the proposed PDE generates regularized tensors which adapt to the first derivative information of an image as time evolves. Moreover, the regularized tensor is used to improve the quality of image denoising, image enhancement, corner detection, and ramp preserving denoising.

The rest of this paper is organized as follows. We propose a nonlinear PDE for regularizing a tensor which contains the first derivative information of an image in Section 2. The existence and uniqueness of solution to the proposed PDE are proved in Section 3. Applications which use derivative information of an image are shown with examples in Section 4 and we show how a regularized tensor in the proposed PDE is used to improve the quality of results. The paper is concluded in Section 5.

2 A nonlinear PDE for regularizing a tensor

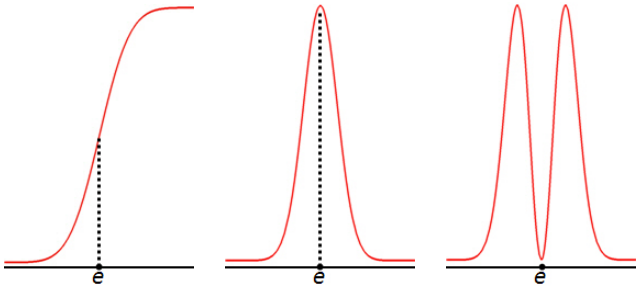
2.1 Modeling of PDE

In this section, we propose a nonlinear PDE for regularizing a tensor which contains the first derivative information of an image. Let us consider a gray image $I\colon \Omega \subset \mathbf{R}^2 \to \mathbf{R}^+$, which is smooth enough to compute a structure tensor [4–6]:

$$\nabla I \nabla I^{\mathrm{T}} = \begin{pmatrix} \left(\dfrac{\partial I}{\partial x}\right)^2 & \dfrac{\partial I}{\partial x}\dfrac{\partial I}{\partial y} \\ \dfrac{\partial I}{\partial x}\dfrac{\partial I}{\partial y} & \left(\dfrac{\partial I}{\partial y}\right)^2 \end{pmatrix}. \tag{2.1}$$

It contains all first derivative information of an image because the maximum eigenvalue and the corresponding eigenvector provide strength of edges and a direction of the gradient of the image, respectively. There have been many different types of diffusion for regularizing a tensor data with applications to optical flow estimation [7–9] and diffusion tensor magnetic resonance imaging [10–13]. In these works, diffusion coefficients for regularizing a tensor data depend on derivatives of the tensor data. If a regularized tensor data with such diffusion coefficients is used in a diffusivity matrix of denoising an image, it is not desirable to preserve edges and corners in a denoised image.

If a regularized tensor is used as diffusion coefficients of image denoising, the desirable features of the tensor are as follows: the maximum eigenvalue has a larger value on edges than on homogeneous regions, it has the local maximum on edges, and the corresponding



(a) (b) (c)

Figure 2.1: (a) is an one-dimensional image data $I(x)$. The point $x = e$ is an edge in (a). (b) is $u(x,0) = \left(\frac{dI}{dx}(x)\right)^2$. (c) is $\left|\frac{\partial u}{\partial x}(x,0)\right|^2$.

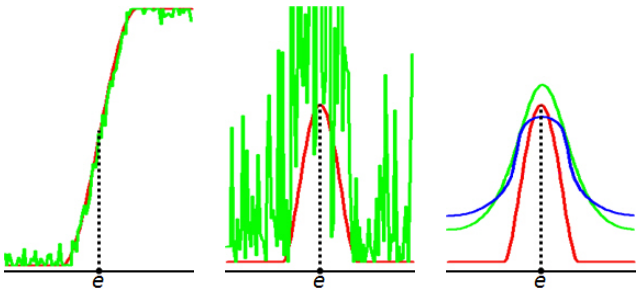
eigenvector near edges is orthogonally aligned to edges of an image. For a clear explanation, we consider an image data I as a function defined on a closed interval $[0,1]$ in Figure 2.1-(a). Let us write the corresponding formulation to typical tensor regularization in [9, 14] (see (B.2) and (B.3) in Appendix B):

$$\begin{aligned} \frac{\partial u}{\partial \tau}(x,\tau) &= \frac{\partial}{\partial x}\left(h\left(\left|\frac{\partial u}{\partial x}\right|^2\right)\frac{\partial u}{\partial x}\right) \quad \text{in} \quad [0,1]\times(0,T_2]\,, \\ \frac{\partial u}{\partial x}(0,\tau) &= \frac{\partial u}{\partial x}(1,\tau) = 0 \quad \text{on} \quad (0,T_2]\,, \\ u(x,0) &= \left(\frac{dI}{dx}(x)\right)^2 \quad \text{on} \quad [0,1]. \end{aligned} \tag{2.2}$$

The point $x = e$ in Figure 2.1-(a) is an edge of the image data I which is characterized by a local maximum of the derivative of the image. In order to keep the edge in a denoised image, the value $u(e,T_2)$ should be as large as possible. However, a regularized u near the edge at $x = e$ is easily smeared out since $|\frac{\partial u}{\partial x}|^2$ has a local minimum at the edge. It explains that the function $|\frac{\partial u}{\partial x}|^2$ in the diffusion coefficient for regularizing u does not help to preserve a large value of u at edges of an image when end time T_2 become large; see Figure 2.1. Instead of using such a function, we propose a diffusion process of u whose diffusion coefficient consists of the diffused quantity u itself:

$$\begin{aligned} \frac{\partial u}{\partial \tau}(x,\tau) &= \frac{\partial}{\partial x}\left(h\big(u^2\big)\frac{\partial u}{\partial x}\right) \quad \text{in} \quad [0,1]\times(0,T_2]\,, \\ \frac{\partial u}{\partial x}(0,\tau) &= \frac{\partial u}{\partial x}(1,\tau) = 0 \quad \text{on} \quad (0,T_2]\,, \\ u(x,0) &= \left(\frac{dI}{dx}(x)\right)^2 \quad \text{on} \quad [0,1]. \end{aligned} \tag{2.3}$$

Then a regularized u near the edge at $x = e$ is not easily smeared out since u^2 has a local maximum at the edge. Even if there is noise in a given image, a solution of the proposed PDE (2.3) outperforms a solution of the PDE (2.2) since it has larger value at $x = e$; see Figure 2.2.



(a) (b) (c)

Figure 2.2: (a) is an one-dimensional noisy image data $I(x)$ (green) and an original data (red). (b) is an initial data of (2.3) (green) and the square of derivative of the original data (red). (c) is regularized derivatives at $T_2 = 600$ using different PDEs. The green curve is a result of the proposed PDE (2.3) and the blue curve is a result of (2.2).

Now, by extending an image data I in the PDE (2.3) to a gray image, we propose a noble nonlinear PDE for regularizing a tensor which contains derivatives of the image:

$$\begin{aligned} \frac{\partial u_{ij}}{\partial \tau}(x,\tau) = \nabla \cdot \big(g\big(U_\sigma\big)\nabla u_{ij}\big) \quad &\text{in} \quad \Omega \times (0,T_2]\,, \\ \big(g\big(U_\sigma\big)\nabla u_{ij}\big)\cdot \mathbf{n} = 0 \quad &\text{on} \quad \partial\Omega \times (0,T_2]\,, \\ u_{ij}(x,0) = \Big(\nabla I(x)\nabla I(x)^{\mathrm{T}}\Big)_{ij} \quad &\text{on} \quad \Omega. \end{aligned} \tag{2.4}$$

The function g is defined on a set $\mathcal{S}$ of real symmetric 2×2 matrices:

$$g(M) \equiv \tilde{g}(\Lambda) v_\Lambda {v_\Lambda}^{\mathrm{T}} + \tilde{g}(\lambda) v_\lambda {v_\lambda}^{\mathrm{T}}, \tag{2.5}$$

where (Λ, v_Λ) and (λ, v_λ) are eigenpairs of $M \in \mathcal{S}$, $\Lambda \geq \lambda$, and $\tilde{g}$ is even symmetric to the function (1.1). Note that it is easily extended to a color image in Section 3. Comparing (2.4) with nonlinear PDEs (B.2) and (B.3) for regularizing a tensor, the only difference is a diffusivity matrix depending upon the tensor itself, not derivatives of the tensor. The key point is that the maximum eigenvalue and the corresponding eigenvector of a tensor in the proposed PDE provides regularized strength of edges of an image and a regularized direction of the gradient of the image. That is, the tensor is diffused fast along edges of an image but slowly across them. Note that the tensor in (B.2) or (B.3) is diffused fast along edges of a tensor data but slowly across them. In the next section, we will show that a regularized tensor from (2.4) helps to preserve edges and corners in a denoised image and it represents derivative information of an image better than regularized tensors from (B.1), (B.2), and (B.3).

2.2 Quality of the nonlinear structure tensor

In this section, we show distinctive features of a regularized tensor from the proposed PDE (2.4). Two examples are presented to show a superiority of the proposed method by comparing it with other methods. The first is image denoising with different end time for

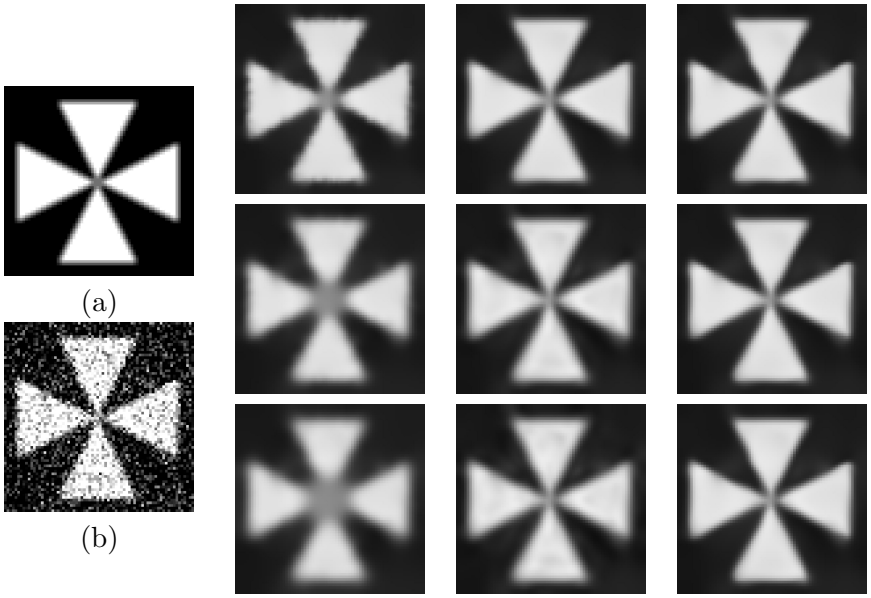


Figure 2.3: (a) is a clean image I_c. In (b), we add the Gaussian white noise with zero mean and the standard deviation 70 (SNR $\simeq$ 7.62). (c), (d), and (e) are obtained by different PDEs for image denoising (A.1), (A.2), and (A.3) at $T_1 = 30$, respectively (see Appendix A). From top to bottom, different end time T_2 for regularizing a tensor is used as 1, 5, and 10. SNR (2.6) and relative H^1 norm error (2.7) are shown in Table 2.1. Note that denoised images in (e) from (A.3) which uses our regularized tensor (2.4) preserve geometric features such as edges and corners in the original image (a) and have steady and high SNR with various end time T_2.

regularizing a tensor and the second is to check how well a regularized tensor represents derivative information of an image. For a comparison, we list PDEs for image denoising using different regularized derivative information of an image in Appendix A and PDEs for regularizing a tensor in Appendix B.

In Figure 2.3, the Gaussian white noise with the zero mean and the standard deviation 70 is added to a clean image I_c. The quality of regularized images is measured by two different methods in Table 2.1. The first is signal-to-noise ratio (SNR):

$$\text{SNR} = 10 \log_{10} \frac{V(I_r)}{V(I_r - I_c)}, \tag{2.6}$$

where $V(\cdot)$ is the variance of an image and I_r is the solution at $t = T_1$ of PDEs in Appendix A. The second is relative H^1 norm error:

$$\text{Rel. } H^1 = \frac{\|I_r - I_c\|_{H^1}}{\|I_c\|_{H^1}}, \qquad \text{where} \quad \|I\|_{H^1}^2 = \int_\Omega I^2 + \int_\Omega |\nabla I|^2. \tag{2.7}$$

End time for image denoising is fixed as $T_1 = 30$ and denoised images are presented by using different end time $T_2 = 1$, 5, and 10 for regularizing a tensor in Figure 2.3. As it is expected,

Comparison of the solution to different PDEs for image denoising

T_2	(A.1)		(A.2)		(A.3)	
	SNR	Rel. H^1	SNR	Rel. H^1	SNR	Rel. H^1
1	16.19	0.260	16.92	0.251	17.54	0.242
5	12.98	0.314	15.56	0.271	17.38	0.244
10	10.89	0.357	14.87	0.281	17.26	0.245

Table 2.1: We use different image denoising methods (A.1), (A.2), and (A.3) of the noisy image in Figure 2.3-b.

the PM model (A.1) gives wiggly image with small T_2 and too smeared image with large T_2. Regularized images from (A.2) look clean, but SNR is smaller and relative H^1 norm error is larger than results from the proposed method. Denoised images from (A.3) which uses the proposed PDE (2.4) for regularizing a tensor have steady and high SNR and robust to various end time T_2. Considering with the necessity of choosing a proper noise scale in (A.1) which depends on an amount of noise, results from (A.3) are practically useful because it is hard to measure an amount of noise in a given image. Denoised images from (A.3) also preserve geometric features such as edges and corners in the original image. It explains that a regularized tensor by our method has distinctive information of derivatives of image which improves the quality of image denoising.

Now, we show how well a regularized tensor in the proposed PDE (2.4) represents derivative information near edges in an image, as time τ evolves. Since we regularize a tensor $U(\cdot, 0)$ whose the maximum eigenvalue and the corresponding eigenvector are $|\nabla I|^2$ and ∇I, respectively, the maximum eigenvalue Λ and the corresponding unit eigenvector v_Λ of a tensor $U(\cdot, \tau)$ are considered as regularized strength of edges and a regularized direction of the gradient of an image, respectively. In order to check that both Λ and v_Λ near edges are properly obtained, we compute a normalized vector field $\mathcal{V}$ which points a local maxima of Λ along a direction of v_Λ:

$$\mathcal{V} \equiv \operatorname{sgn}\left(\nabla\Lambda \cdot v_\Lambda\right) v_\Lambda, \tag{2.8}$$

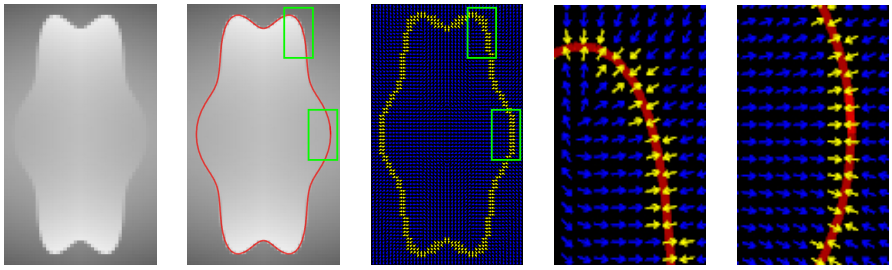
where $\operatorname{sgn}(\cdot)$ is the sign function. If an original image I is smooth and a regularized tensor of the proposed PDE (2.4) preserves an initial condition very well, we have,

$$\begin{aligned}\mathcal{V} &\simeq \operatorname{sgn}\left(\nabla\big(|\nabla I|^2\big) \cdot \frac{\nabla I}{|\nabla I|}\right)\frac{\nabla I}{|\nabla I|} = \operatorname{sgn}\left(\nabla\big(|\nabla I|\big) \cdot \frac{\nabla I}{|\nabla I|}\right)\frac{\nabla I}{|\nabla I|} \\ &= \operatorname{sgn}\left(\mathcal{D}^2_{\frac{\nabla I}{|\nabla I|}} I\right)\frac{\nabla I}{|\nabla I|},\end{aligned}$$

where $\mathcal{D}^2_v(\cdot)$ is the second directional derivative along the direction v. In this case, if we restrict $\mathcal{V}$ on a set of points in the pixel resolution, which is defined by

$$\mathcal{R} \equiv \big\{x \in \Omega \mid \mathcal{V}(x^*) \cdot \mathcal{V}(x) < 0 \text{ and } x^* = x + \mathcal{V}(x)\big\}, \tag{2.9}$$

then vectors in $\mathcal{V}|_{\mathcal{R}}$, the restriction of $\mathcal{V}$ on $\mathcal{R}$, should have two properties. First, they are placed in the closest pixel to edges which are defined by zeros of the second directional

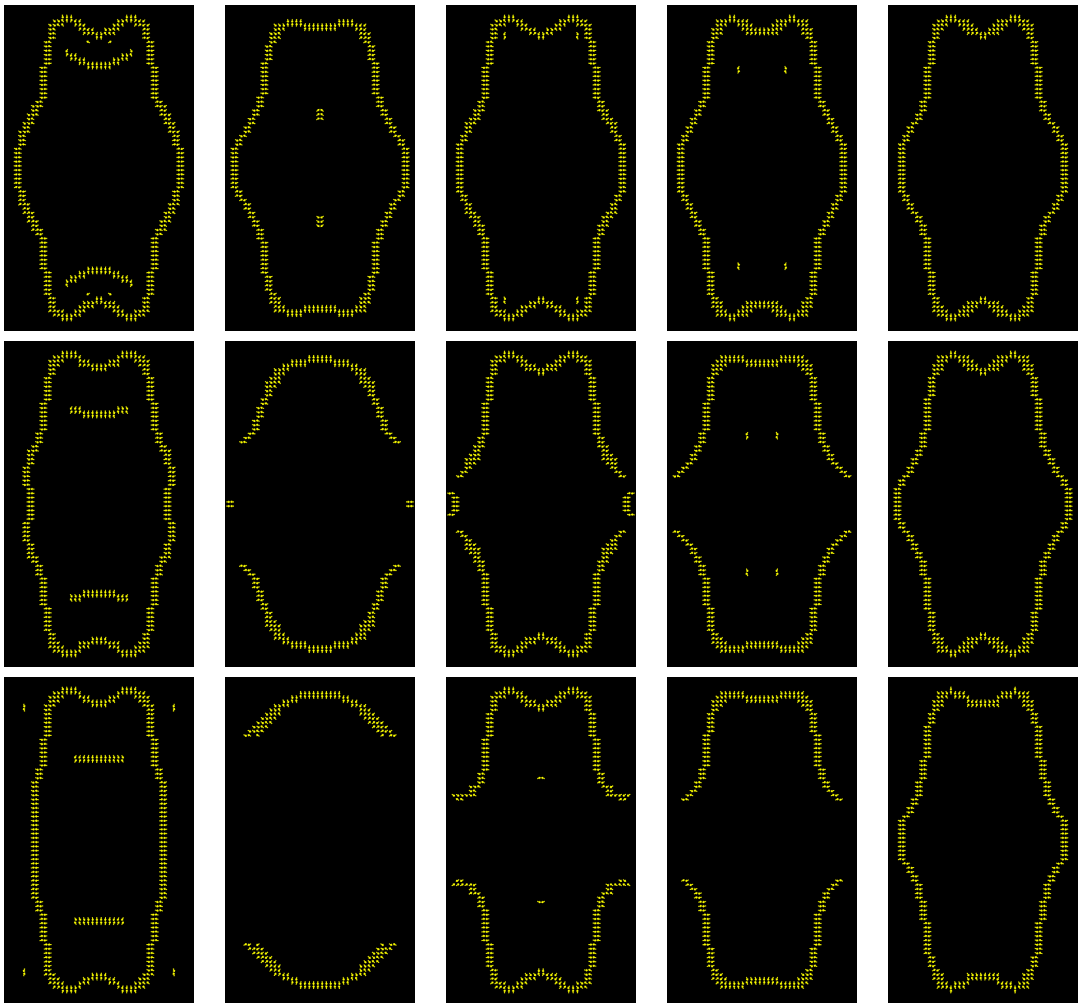


(a) (b) (c) (d) (e)

Figure 2.4: (a) is a given image and the red curve in (b) is the exact location of edges. (c) is the vector field $\mathcal{V}$ in (2.8) obtained by the regularized tensor of the proposed PDE (2.4) at $T_2 = 10$. Vectors on the set $\mathcal{R}$ in (2.9) are highlighted in yellow. (d) and (e) are magnified images from both the red curve in (b) and vectors in (c) on green square regions. Note that vectors in highlighted in yellow are placed in pixels close to edges. They also point to edges and are aligned orthogonally to edges.

derivative of an image along a direction of the gradient of the image. Second, they point to edges and are aligned orthogonally to edges. Therefore, if $\mathcal{V}|_{\mathcal{R}}$ from a regularized tensor does not have these properties, the regularized tensor does not properly represents derivative information near edges in an original image.

In Figure 2.4, we obtain the vector field $\mathcal{V}$ (2.8) from our regularized tensor (2.4) at $T_2 = 10$. Vectors on the set $\mathcal{R}$ in (2.9) are highlighted in yellow. They are placed in pixels close to edges. They also point to edges and are aligned orthogonally to edges. A regularized tensor from the proposed PDE correctly represents derivative information near edges in an image. In Figure 2.5, we use the same image in Figure 2.4-(a) to compare a vector field $\mathcal{V}|_{\mathcal{R}}$ from our regularized tensor (B.4) with vector fields from different regularized tensors obtained by (B.1), (B.2), and (B.3). More interestingly, we also compare a vector field $\mathcal{V}|_{\mathcal{R}}$ from a tensor $\nabla\tilde{I}\nabla\tilde{I}^{\mathrm{T}}$, where a regularized image $\tilde{I}$ is obtained by the PM model (A.1). The last comparison raises a question that which one is better to preserve derivative information of an image between the method of regularizing an image first and taking derivative of the regularized image later and the method of taking derivative of an image first and regularizing it later. Weickert [3] already mentioned that the latter is better in the case of a flow-like image by comparing $\nabla(G_\sigma * I)$ with eigenvector which corresponds to maximum eigenvalue in a structure tensor $G_\sigma * \left(\nabla I \nabla I^{\mathrm{T}}\right)$. We put the problem in more difficult situation. The image in Figure 2.4 has weak edges and we use more advanced regularization of the image (A.1) instead of using an isotropic linear regularization of the image. For the comparison, end time T_1 for an image diffusion in (A.1) and T_2 for tensor diffusions (B.1), (B.2), (B.3), and (B.4) are taken as the same value. A vector field $\mathcal{V}|_{\mathcal{R}}$ from our regularized tensor (B.4) gives the best quality for preserving derivative information near edges in an image, as time τ evolves. It also shows the better performance than a result from regularizing an image first and taking derivative of the regularized image later.



(a) (b) (c) (d) (e)

Figure 2.5: A restricted vector field $\mathcal{V}|_{\mathcal{R}}$ is obtained by (2.8) and (2.9) from a tensor. Vectors in (a) are $\mathcal{V}|_{\mathcal{R}}$ from a tensor $\nabla\tilde{I}\nabla\tilde{I}^{\mathrm{T}}$, where a regularized image $\tilde{I}$ is obtained by the PM model (A.1). Vectors in (b), (c), (d), and (e) are $\mathcal{V}|_{\mathcal{R}}$ from the regularized tensors of (B.1), (B.2), (B.3) and (B.4), respectively. From top to bottom, end time T_1 for regularizing an image and T_2 for regularizing a tensor are taken as same values, $11, 33$, and 55. Apparently, vectors from our regularized tensor (B.4) show the best result for preserving derivative information near edges in an image as time evolves.

3 Existence and uniqueness of the proposed PDE

In this section, we extend the proposed PDE (2.4) for a color image and provide mathematical justification to use the extended PDE. The proof of the existence and uniqueness of solution to the extended PDE follows the main argument in [2] with an extension to a multidimensional data. Since the PDE for image denoising (A.3) has the solution when a diffusivity matrix is positive definite, we will also prove a solution of the proposed PDE is positive definite if an initial data is.

Recall that $\mathcal{S}$ is a set of real symmetric 2×2 matrices. Any matrix $X \in \mathcal{S}$ is diagonalizable, *i.e.*,

$$X = \begin{pmatrix} x_1 & -x_2 \\ x_2 & x_1 \end{pmatrix} \begin{pmatrix} \Lambda^X & 0 \\ 0 & \lambda^X \end{pmatrix} \begin{pmatrix} x_1 & -x_2 \\ x_2 & x_1 \end{pmatrix}^{\mathrm{T}},$$

where $\big(\Lambda^X, (x_1, x_2)^{\mathrm{T}}\big)$ and $\big(\lambda^X, (-x_2, x_1)^{\mathrm{T}}\big)$ are eigenpairs of X and $\Lambda^X \geq \lambda^X$. Note that $x_1^2 + x_2^2 = 1$. The existence and uniqueness of solution to the proposed PDE (2.4) are based on a simple property of a Lipschitz function defined on $\mathcal{S}$.

Lemma 1. *Let f be a function from $\mathcal{S}$ to $\mathcal{S}$ defined by*

$$f(X) \equiv \begin{pmatrix} x_1 & -x_2 \\ x_2 & x_1 \end{pmatrix} \begin{pmatrix} \tilde{f}\big(\Lambda^X\big) & 0 \\ 0 & \tilde{f}\big(\lambda^X\big) \end{pmatrix} \begin{pmatrix} x_1 & -x_2 \\ x_2 & x_1 \end{pmatrix}^{T},$$

where $\tilde{f}$ is a Lipschitz function on $\mathbf{R}$. Then f is also a Lipschitz function on $\mathcal{S}$.

Proof. It is enough to show that f is a Lipschitz function on $\mathcal{S}$ with the Frobenius norm $\|\cdot\|_F$. Let X and Y be two elements in $\mathcal{S}$. Since X and Y are diagonalizable, we have

$$X - Y = \begin{pmatrix} \Lambda^X x_1^2 + \lambda^X x_2^2 - \Lambda^Y y_1^2 - \lambda^Y y_2^2 & \big(\Lambda^X - \lambda^X\big)x_1x_2 - \big(\Lambda^Y - \lambda^Y\big)y_1y_2 \\ \big(\Lambda^X - \lambda^X\big)x_1x_2 - \big(\Lambda^Y - \lambda^Y\big)y_1y_2 & \Lambda^X x_2^2 + \lambda^X x_1^2 - \Lambda^Y y_2^2 - \lambda^Y y_1^2 \end{pmatrix},$$

where $\big(\Lambda^X, (x_1, x_2)\big)$ and $\big(\lambda^X, (-x_2, x_1)\big)$ are eigenpairs of X with $\Lambda^X \geq \lambda^X$ and $\big(\Lambda^Y, (y_1, y_2)\big)$ and $\big(\lambda^Y, (-y_2, y_1)\big)$ are eigenpairs of Y with $\Lambda^Y \geq \lambda^Y$. Note that we have two identities

$$x_1^2 + x_2^2 = 1 \text{ and } y_1^2 + y_2^2 = 1. \tag{3.1}$$

If we denote components of $X - Y$ by a, b, and c, then

$$\|X - Y\|_F^2 = \left\| \begin{pmatrix} a & b \\ b & c \end{pmatrix} \right\|_F^2 = a^2 + 2b^2 + c^2 = \frac{1}{2}\big((a+c)^2 + (a-c)^2 + 4b^2\big). \tag{3.2}$$

For the simplicity of notation, we define

$$\alpha \equiv \Lambda^X - \lambda^X,\ \beta \equiv \Lambda^Y - \lambda^Y,\ \gamma \equiv \Lambda^X - \Lambda^Y, \text{ and } \delta \equiv \lambda^X - \lambda^Y,$$

then we obtain, by (3.1),

$$\begin{aligned} (a+c)^2 &= (\gamma+\delta)^2, \\ (a-c)^2 + 4b^2 &= \alpha^2 + \beta^2 + 2\alpha\beta\big((x_2y_1 - x_1y_2)^2 - (x_1y_1 + x_2y_2)^2\big). \end{aligned} \tag{3.3}$$

Let θ be an angle between (x_1, x_2) and (y_1, y_2). Since (x_1, x_2) and $(-x_2, x_1)$ are orthogonal, we get

$$(a-c)^2 + 4b^2 = \alpha^2 + \beta^2 + 2\alpha\beta\big(2\sin^2\theta - 1\big). \tag{3.4}$$

From (3.2), (3.3) and (3.4), we deduce

$$\|X - Y\|_F^2 = \frac{1}{2}\big((\gamma+\delta)^2 + (\alpha-\beta)^2 + 4\alpha\beta\sin^2\theta\big).$$

By using $\alpha - \beta = \gamma - \delta$, we finally obtain

$$\|X - Y\|_F^2 = \big(\Lambda^X - \Lambda^Y\big)^2 + \big(\lambda^X - \lambda^Y\big)^2 + 2\big(\Lambda^X - \lambda^X\big)\big(\Lambda^Y - \lambda^Y\big)\sin^2\theta. \tag{3.5}$$

Now, we show that f is a Lipschitz function on S. From (3.5), it is obvious that

$$\begin{aligned}\|f(X) - f(Y)\|_F^2 = &\Big(\tilde{f}\big(\Lambda^X\big) - \tilde{f}\big(\Lambda^Y\big)\Big)^2 + \Big(\tilde{f}\big(\lambda^X\big) - \tilde{f}\big(\lambda^Y\big)\Big)^2 \\ &+ 2\Big(\tilde{f}\big(\Lambda^X\big) - \tilde{f}\big(\lambda^X\big)\Big)\Big(\tilde{f}\big(\Lambda^Y\big) - \tilde{f}\big(\lambda^Y\big)\Big)\sin^2\theta.\end{aligned}$$

Since we have $\Lambda^X \geq \lambda^X$ and $\Lambda^Y \geq \lambda^Y$,

$$\|f(X) - f(Y)\|_F \leq C\|X - Y\|_F,$$

where a constant C is a Lipschitz constant of $\tilde{f}$. □

Now, we extend the proposed PDE (2.4) for a color image. Let Ω be $(0,1)\times(0,1) \subset \mathbf{R}^2$ and $h\colon \mathbf{R} \to \mathbf{R}^+$ be smooth, decreasing on $[0,\infty)$, even, $h(0) = 1$, $\lim_{s\to\infty} h(s) = 0$, and $s \longmapsto h(\sqrt{|s|})$ smooth. Note that a typical choice of function h is even symmetric to the function (1.1). For the simplicity of notation, we define

$$u_a \equiv a^{\mathrm{T}} \begin{pmatrix} u_{11} & u_{12} \\ u_{21} & u_{22} \end{pmatrix} a,$$

where a vector $a \in \mathbf{R}^2$ is constant with $|a| = 1$. Let us consider a color image $I\colon \Omega \subset \mathbf{R}^2 \to (\mathbf{R}^+)^3$ as a function whose components I_k are smooth enough to compute

$$M = \begin{pmatrix} m_{11} & m_{12} \\ m_{21} & m_{22} \end{pmatrix} \equiv \sum_{k=1}^{3} \nabla I_k \nabla {I_k}^{\mathrm{T}}. \tag{3.6}$$

An extension of the proposed PDE (2.4) for a color image is a system of nonlinear parabolic partial differential equations

$$\begin{aligned} \frac{\partial u_{ij}}{\partial \tau} &= \nabla\cdot\big(g(U_\sigma)\nabla u_{ij}\big) \quad \text{in} \quad \Omega\times(0,T), \\ \big(g(U_\sigma)\nabla u_{ij}\big)\cdot\mathbf{n} &= 0 \quad \text{on} \quad \partial\Omega\times(0,T), \\ u_{ij}(x,0) &= m_{ij}(x) \quad \text{on} \quad \Omega, \end{aligned} \tag{3.7}$$

where a function g from $\mathcal{S}$ to $\mathcal{S}$ was defined in (2.5). In order to prove existence and uniqueness, we introduce standard function spaces. Let $H^1(\Omega)$ be the Sobolev space of functions $u \in L^2(\Omega)$ with its distributional derivatives $\mathcal{D}^s u$ of order $|s| = \sum_{j=1}^2 s_j \leq 1$ being in $L^2(\Omega)$. It is a Hilbert space with the norm

$$\|u\|_{H^1(\Omega)} = \left(\sum_{|s|\leq 1} \|\mathcal{D}^s u\|^2_{L^2(\Omega)} \right)^{\frac{1}{2}}$$

and we denote by $\big(H^1(\Omega)\big)'$ its dual space. For a given Banach space $\mathcal{B}$ with a norm $\|\cdot\|_{\mathcal{B}}$, we denote by $L^p((0,T);\mathcal{B})$ the set of all strongly measurable functions $u\colon [0,T] \to \mathcal{B}$ with

$$\|u\|_{L^p((0,T);\mathcal{B})} \equiv \left(\int_0^T \|u(\tau)\|^p_{\mathcal{B}} \, d\tau \right)^{\frac{1}{p}} < \infty$$

for $1 \leq p < \infty$, and

$$\|u\|_{L^\infty((0,T);\mathcal{B})} \equiv \operatorname*{ess\,sup}_{0\leq t\leq T} \|u(\tau)\|_{\mathcal{B}} < \infty.$$

We also denote by $C([0,T];\mathcal{B})$ the set of all continuous functions $u\colon [0,T] \to \mathcal{B}$ with

$$\|u\|_{C([0,T];\mathcal{B})} \equiv \max_{[0,T]} \|u(\tau)\|_{\mathcal{B}} < \infty.$$

In a similar way, we denote by $L^\infty((0,T);C^\infty(\Omega))$ the set of all functions such that $u(\tau) \in C^\infty(\Omega)$ for almost every $\tau \in (0,T)$ with

$$\|u\|_{L^\infty((0,T);C^\infty(\Omega))} \equiv \inf_\alpha \Big\{ \|u(\tau)\|_{C^\infty(\Omega)} \leq \alpha, \text{ for almost every } \tau \in (0,T) \Big\} < \infty.$$

See [15] for more details of function spaces. Now, we prove following statements for the proposed PDE (3.7).

Theorem 1. *If $m_{ij} \in L^2(\Omega)$, then there exist unique solutions*

$$u_{ij} \in C([0,T];L^2(\Omega)) \cap L^2((0,T);H^1(\Omega)) \ \textit{with} \ \frac{\partial u_{ij}}{\partial \tau} \in L^2((0,T);\big(H^1(\Omega)\big)')$$

satisfying (3.7) *in the distributional sense. Moreover,*

$$u_{ij} \in C^\infty((0,T)\times\bar{\Omega}) \ \textit{and} \ \|u_{ij}\|_{L^\infty((0,T);L^2(\Omega))} \leq \|m_{ij}\|_{L^2(\Omega)}.$$

Proof. Part **1**: Uniqueness. Let u_{ij} and v_{ij} be two solutions of (3.7). That is, we have

$$\begin{aligned}
\frac{\partial u_{ij}}{\partial \tau} &= \nabla \cdot \big(g(U_\sigma)\nabla u_{ij}\big), \quad \big(g(U_\sigma)\nabla u_{ij}\big)\cdot \mathbf{n} = 0, \quad u_{ij}(x,0) = m_{ij}(x), \\
\frac{\partial v_{ij}}{\partial \tau} &= \nabla \cdot \big(g(V_\sigma)\nabla v_{ij}\big), \quad \big(g(V_\sigma)\nabla v_{ij}\big)\cdot \mathbf{n} = 0, \quad v_{ij}(x,0) = m_{ij}(x).
\end{aligned}$$

Denoting $w_{ij} \equiv u_{ij} - v_{ij}$, we obtain

$$\frac{\partial w_{ij}}{\partial \tau} - \nabla \cdot \big(g(U_\sigma)\nabla w_{ij}\big) = \nabla \cdot \big((g(U_\sigma) - g(V_\sigma))\nabla v_{ij}\big).$$

For a constant vector $a \in \mathbf{R}^2$ with $|a| = 1$, from above equations we have

$$\frac{\partial w_a}{\partial \tau} - \nabla \cdot \big(g(U_\sigma)\nabla w_a\big) = \nabla \cdot \big((g(U_\sigma) - g(V_\sigma))\nabla v_a\big).$$

Multiplying it by w_a and integrating over Ω, we get

$$\begin{aligned} &\frac{1}{2}\frac{d}{d\tau}\int_\Omega |w_a(x,\tau)|^2 dx + \int_\Omega \big(\nabla g(U_\sigma)(x,\tau)\nabla w_a(x,\tau)\big)\cdot w_a(x,\tau)dx = \\ &\quad -\int_\Omega \big(\big(g(U_\sigma) - g(V_\sigma)\big)(x,\tau)\nabla v_a(x,\tau)\big)\cdot \nabla w_a(x,\tau)dx, \quad \text{a.e} \quad \tau \in (0,T). \end{aligned} \tag{3.8}$$

Now, we claim following statements that will be proved at the end of Part 1,

L-1. The diffusivity matrix $g(U_\sigma)$ is uniformly positive definite almost everywhere, *i.e.*, there exists a constant ν independent of (x,τ) such that $a^{\mathrm{T}} g(U_\sigma) a \geq \nu > 0$ almost everywhere in $\Omega \times (0,T)$ for all $a \in \mathbf{R}^2$ with $|a| = 1$.

L-2. There exits a constant K, depending only on G_σ such that

$$\begin{aligned} &\int_\Omega \big(\big(g(U_\sigma) - g(V_\sigma)\big)(x,\tau)\nabla v_a(x,\tau)\big)\cdot \nabla w_a(x,\tau)dx \\ &\qquad\qquad \leq KF(\tau)\|\nabla v_a(\tau)\|_{L^2(\Omega)}\|\nabla w_a(\tau)\|_{L^2(\Omega)}, \end{aligned}$$

where $F(\tau) = \max\limits_{\substack{a\in\mathbf{R}^2 \\ |a|=1}} \big\|a^{\mathrm{T}} W(\tau) a\big\|_{L^2(\Omega)}$ and $W \equiv U - V$.

From (3.8), we use L-1 to obtain

$$\begin{aligned} &\frac{1}{2}\frac{d}{d\tau}\|w_a(\tau)\|^2_{L^2(\Omega)} + \nu\|\nabla w_a(\tau)\|^2_{L^2(\Omega)} \\ &\qquad \leq \int_\Omega \big(\big(g(U_\sigma) - g(V_\sigma)\big)(x,\tau)\nabla v_a(x,\tau)\big)\cdot \nabla w_a(x,\tau)dx, \quad \text{a.e} \quad \tau \in (0,T). \end{aligned}$$

Using L-2 and Young's inequality, we deduce

$$\begin{aligned} &\frac{1}{2}\frac{d}{d\tau}\|w_a(\tau)\|^2_{L^2(\Omega)} + \nu\|\nabla w_a(\tau)\|^2_{L^2(\Omega)} \\ &\qquad \leq \frac{K^2}{2\nu}\{F(\tau)\}^2\|\nabla v_a(\tau)\|^2_{L^2(\Omega)} + \frac{\nu}{2}\|\nabla w_a(\tau)\|^2_{L^2(\Omega)}, \quad \text{a.e} \quad \tau \in (0,T). \end{aligned}$$

Then, we obtain

$$\frac{d}{d\tau}\|w_a(\tau)\|^2_{L^2(\Omega)} \leq \frac{K^2}{\nu}\{F(\tau)\}^2\|\nabla v_a(\tau)\|^2_{L^2(\Omega)}, \quad \text{a.e} \quad \tau \in (0,T). \tag{3.9}$$

Since the inequality (3.9) holds for an arbitrarily given constant vector $a \in \mathbf{R}^2$ with $|a| = 1$, we deduce

$$\frac{d}{d\tau}\{F(\tau)\}^2 \leq \phi(\tau)\{F(\tau)\}^2, \quad \text{a.e} \quad \tau \in (0,T), \tag{3.10}$$

where $\phi(\tau) \equiv \dfrac{K^2}{\nu} \max\limits_{\substack{a\in\mathbf{R}^2\\|a|=1}} \|\nabla v_a(\tau)\|^2_{L^2(\Omega)}$.

Now, we apply Gronwall's inequality to (3.10),

$$\{F(\tau)\}^2 \le \{F(0)\}^2 \exp \int_0^\tau \phi(s)ds,$$

By using $F(0) = 0$, we obtain

$$F(\tau) = \max_{\substack{a\in\mathbf{R}^2\\|a|=1}} \left\|a^{\mathrm{T}} W(\tau) a\right\|^2_{L^2(\Omega)} = 0, \quad \text{a.e. } \tau \in (0,T).$$

Finally, we conclude

$$a^{\mathrm{T}} W(x,\tau) a = 0, \quad \text{a.e. } (x,\tau) \in \Omega \times (0,T),$$

for all constant vectors $a \in \mathbf{R}^2$ with $|a| = 1$. It implies that

$$u_{ij}(x,\tau) = v_{ij}(x,\tau), \quad \text{a.e. } (x,\tau) \in \Omega \times (0,T), \quad \forall i,j \in \{1,2\}.$$

Proof of L-1. Since the function h is decreasing, it is enough to show that all eigenvalues of U_σ are bounded almost everywhere in $\Omega \times (0,T)$. Since $u_{ij} \in L^\infty((0,T);L^2(\Omega))$ from the result of existence, we get $G_\sigma * u_{ij} \in L^\infty((0,T);C^\infty(\Omega))$ for all i,j. Then we have $a^{\mathrm{T}} U_\sigma a \in L^\infty((0,T);C^\infty(\Omega))$ for any constant vectors $a \in \mathbf{R}^2$ with $|a| = 1$. That is, there exits a constant $\tilde{K}$, depending only on G_σ and $\|m_{ij}\|_{L^2(\Omega)}$, such that $\left|a^{\mathrm{T}} U_\sigma a\right| = \left|G_\sigma * a^{\mathrm{T}} U a\right| \le \tilde{K}$, almost everywhere in $\Omega \times (0,T)$.

Proof of L-2. First, we simply obtain

$$\begin{aligned}\int_\Omega \big(\big(g(U_\sigma) - g(V_\sigma)\big)(x,\tau)\nabla v_a(x,\tau)\big) \cdot \nabla w_a(x,\tau)dx \\ \le \max_{x\in\Omega} \|(g(V_\sigma) - g(U_\sigma))(x,\tau)\|_2 \|\nabla v_a(\tau)\|_{L^2(\Omega)} \|\nabla w_a(\tau)\|_{L^2(\Omega)},\end{aligned}$$

where $\|\cdot\|_2$ is a matrix norm as the spectral radius for symmetric matrices. We have a norm equivalence between $\|\cdot\|_2$ and the Frobenius norm $\|\cdot\|_F$ such that

$$\frac{1}{\sqrt{2}}\|A\|_F \le \|A\|_2 \le \|A\|_F. \tag{3.11}$$

By Lemma 1 and (3.11), there exits a constant $C(h)$ such that

$$\|g(V_\sigma) - g(U_\sigma)\|_2 \le C(h)\|V_\sigma - U_\sigma\|_2.$$

Then there is a constant K depending on G_σ and h such that

$$\|g(V_\sigma) - g(U_\sigma)\|_2 \le K(h, G_\sigma) \max_{\substack{a\in\mathbf{R}^2\\|a|=1}} \left\|a^{\mathrm{T}}(U - V)(\tau)a\right\|_{L^2(\Omega)}.$$

Note that the constant K is independent of $(x,\tau)\in\Omega\times(0,T)$. Therefore, we obtain

$$\int_\Omega \big(\big(g(U_\sigma)-g(V_\sigma)\big)(x,\tau)\nabla v_a(x,\tau)\big)\cdot\nabla w_a(x,\tau)dx \\ \le K(h,G_\sigma)\max_{\substack{a\in\mathbf{R}^2\\|a|=1}}\big\|a^{\mathrm{T}}W(\tau)a\big\|_{L^2(\Omega)}\|\nabla v_a(\tau)\|_{L^2(\Omega)}\|\nabla w_a(\tau)\|_{L^2(\Omega)},$$

where $W\equiv U-V$.

Part **2**: Existence. Let us define a Hilbert space

$$S\equiv\left\{w\in L^2((0,T);H^1(\Omega))\;\middle|\;\frac{\partial w}{\partial\tau}\in L^2((0,T);\big(H^1(\Omega)\big)')\right\}$$

with a norm

$$\|w\|_S\equiv\left(\|w\|^2_{L^2((0,T);H^1(\Omega))}+\left\|\frac{\partial w}{\partial\tau}\right\|^2_{L^2((0,T);(H^1(\Omega))')}\right)^{\frac12}.$$

We choose a fixed matrix $W=\begin{pmatrix}w_{11}&w_{12}\\w_{21}&w_{22}\end{pmatrix}$ which satisfies

$$w_{ij}\in S\cap L^\infty((0,T);L^2(\Omega))\ \text{ and }\ \|w_{ij}\|_{L^\infty((0,T);L^2(\Omega))}\le\|m_{ij}\|_{L^2(\Omega)}.$$

Consider a linear system of variational problems

$$\left\langle\frac{\partial u_{ij}}{\partial\tau}(\tau),v\right\rangle_{(H^1(\Omega))'\times H^1(\Omega)}+\int_\Omega\big(g(W_\sigma(x,\tau))\nabla u_{ij}(x,\tau)\big)\cdot\nabla v dx=0,\tag{3.12}$$

for all $v\in H^1(\Omega)$ and almost every $\tau\in(0,T)$. From the proof of L-1 in Part 1, there exists a constant ν such that $a^{\mathrm{T}}g(W_\sigma)a\ge\nu>0$ almost everywhere in $\Omega\times(0,T)$ for all $a\in\mathbf{R}^2$ with $|a|=1$. Therefore, by classical results on parabolic partial differential equations in [15], equations (3.12) have unique solutions u^W_{ij} of satisfying following estimates

$$\begin{aligned}&\big\|u^W_{ij}\big\|_{L^2((0,T);H^1(\Omega))}\le C,\\&\left\|\frac{\partial u^W_{ij}}{\partial\tau}\right\|_{L^2((0,T);(H^1(\Omega))')}\le C,\\&\big\|u^W_{ij}\big\|_{L^\infty((0,T);L^2(\Omega))}\le\|m_{ij}\|_{L^2(\Omega)},\end{aligned}$$

where C is a constant depending only on h, G_σ, and $\|m_{ij}\|_{L^2(\Omega)}$. From these estimates, we define a subspace S^{ij}_0 of S

$$S^{ij}_0\equiv\left\{w_{ij}\in S\;\middle|\;\begin{array}{l}w_{ij}(0)=m_{ij},\|w_{ij}\|_{L^2((0,T);H^1(\Omega))}\le C,\\\left\|\dfrac{\partial w_{ij}}{\partial\tau}\right\|_{L^2((0,T);(H^1(\Omega))')}\le C,\\\|w_{ij}\|_{L^\infty((0,T);L^2(\Omega))}\le\|m_{ij}\|_{L^2(\Omega)}\end{array}\right\}.$$

We denote $S_0 \equiv \prod_{(i,j)\in\mathbb{I}} S_0^{ij} \subset S^4$, where $\mathbb{I} \equiv \mathbb{Z}_2 \times \mathbb{Z}_2$. By the construction of S_0^{ij}, we have a well-defined mapping

$$(w_{11}, w_{12}, w_{21}, w_{22}) \mapsto Q(w_{11}, w_{12}, w_{21}, w_{22}) \equiv \left(u_{11}^W, u_{12}^W, u_{21}^W, u_{22}^W\right)$$

from S_0 to S_0. Following the argument in [2], S_0 is a nonempty, convex, weakly compact in S^4 and the mapping Q is weakly continuous. By the Schauder-Tychonoff's fixed-point theorem, there exits $(w_{11}, w_{12}, w_{21}, w_{22}) \in S_0$ such that $(w_{11}, w_{12}, w_{21}, w_{22}) = Q(w_{11}, w_{12}, w_{21}, w_{22})$. Since each component of $Q(w_{11}, w_{12}, w_{21}, w_{22})$ solves (3.12), it proves the existence of a solution of (3.7) in the distributional sense. The regularity is also proved from a general theory of parabolic partial differential equations. □

At last, we prove that the positive definiteness of a regularized tensor $U(x,\tau)$ in (3.7) is guaranteed if an initial tensor M in (3.6) is positive definite. It is crucial condition to solve the PDE for image denoising (A.3) numerically.

Theorem 2. *If $m_{ij} \in L^\infty(\Omega)$, then the solutions u_{ij} of (3.7) satisfy*

$$\operatorname*{ess\,inf}_{x\in\Omega} \lambda(x) \le u_a(x,\tau) \le \operatorname*{ess\,sup}_{x\in\Omega} \Lambda(x) \quad on \quad \Omega \times (0,T],$$

where a is any constant vector in $\mathbf{R}^2$ with $|a| = 1$ and $\Lambda(x)$ and $\lambda(x)$ are maximum and minimum eigenvalues of $M(x)$ given by (3.6), respectively.

Proof. This proof follows Stampacchia's truncation method in [16] and is done by the same method in [3]. We denote a constant

$$C \equiv \operatorname*{ess\,sup}_{x\in\Omega} \Lambda(x).$$

and show the maximum principle first. Let $\psi \in C^1(\mathbf{R})$ be a function with $\psi(s) = 0$ on $(-\infty, 0]$ and $0 < \psi'(s) \le K$ on $(0,\infty)$ for some constant K. For a constant vector $a \in \mathbf{R}^2$ with $|a| = 1$, we define

$$\varphi(s) \equiv \int_0^s \psi(r)dr \quad \text{and} \quad \phi(\tau) \equiv \int_\Omega \varphi(u_a(x,\tau) - C)dx, \quad \tau \in [0,T].$$

Since $\psi(u_a(\tau) - C) \in H^1(\Omega)$ and $\dfrac{\partial u_a}{\partial \tau}(\tau) \in \left(H^1(\Omega)\right)'$, we have

$$\left|\int_\Omega \psi(u_a(x,\tau) - C)\frac{\partial u_a}{\partial \tau}(x,\tau)dx\right| \le \|\psi(u_a(\tau) - C)\|_{H^1(\Omega)} \left\|\frac{\partial u_a}{\partial \tau}(\tau)\right\|_{(H^1(\Omega))'} < \infty.$$

So, the function $\phi(\tau)$ is differentiable for $\tau \in (0,T)$. Now, we obtain the sign of $\phi'(\tau)$ using the statement L-1:

$$\begin{aligned}
\frac{d}{d\tau}\phi(\tau) &= \int_\Omega \psi(u_a(x,\tau) - C)\frac{\partial u_a}{\partial \tau}(x,\tau)dx \\
&= \int_\Omega \psi(u_a(x,\tau) - C)\nabla \cdot \left(g(U_\sigma)(x,\tau)\nabla u_a(x,\tau)\right)dx \\
&= -\int_\Omega \psi'(u_a(x,\tau) - C)\left(g(U_\sigma)(x,\tau)\nabla u_a(x,\tau)\right) \cdot \nabla u_a(x,\tau)dx \le 0.
\end{aligned}$$

By means of $\varphi(s) \le \frac{K}{2}s^2$, we have

$$0 \le \phi(\tau) \le \int_\Omega \varphi(u_a(x,\tau) - m_a)dx \le \frac{K}{2}\|u_a(\tau) - m_a\|^2_{L^2(\Omega)}.$$

Since $u_a \in C([0,T]; L^2(\Omega))$, we get

$$\|u_a(\tau) - m_a\|_{L^2(\Omega)} \to 0 \quad \text{as} \quad \tau \to 0^+.$$

It proves the continuity of $\phi(\tau)$ at $\tau = 0$.

Now, from $\phi \in C[0,T]$, $\phi(0) = 0$, $\phi(\tau) \ge 0$ on $[0,T]$, and $\phi'(\tau) \le 0$, we deduce that $\phi(\tau) = 0$ on $[0,T]$. Hence, for all $\tau \in [0,T]$, we obtain $u_a(x,\tau) - C \le 0$ almost everywhere on Ω. Due to the smoothness of u_a for $\tau > 0$, we finally obtain $u_a(x,\tau) \le C$ on $\Omega \times (0,T]$. Similarly, the minimum principle is proved from the maximum principle when we apply the initial tensor $-M$. □

4 Applications

A regularized tensor from proposed PDEs (2.4) and (3.7) is useful to low-level topics in computer vision such as corner detection, image denoising, image enhancement, and ramp preserving denoising process. Since a regularized tensor by our method adapts to the first derivative information of an image as time evolves, it is interesting to apply our regularized tensor to applications where the first derivative information of the image has been used in a PDE-based method so far. In Section 2.2, we already showed that our regularized tensor improves the quality of image denoising. In this section, we show more examples in order to see how a regularized tensor by the proposed PDE works in applications.

4.1 Corner detection

To find out locations of corners in an image is one of primitive tasks as a preprocess of high-level topics in computer vision such as motion tracking, object recognition, and registration. Many methods have been studied for finding locations of corners in an image and we use a method based on a structure tensor of the image. Since an initial tensor in proposed PDEs (2.4) and (3.7) is regularized along edges of an image, the minimum eigenvalue of a regularized tensor at a corner which is a junction of more than two edges is increased at the corner. In [17], corners are detected where a local maximum of minimum eigenvalues in a structure tensor occurs.

In Figure 4.1, we show that a regularized tensor by our method has more effective directional smoothing than regularized tensors by other methods. From (c) to (f), they are profiles of minimum eigenvalues from different regularized tensors in (B.1), (B.2), (B.3), and (B.4) at end time $T_2 = 70$, respectively. The profile (c) loses the feature which identifies positions of corners. The profile (d) does not show any information yet. Note that, in a numerical experiment, end time in (B.2) should be taken roughly 100 times larger than end time in (B.4) in order to obtain similar information in (f). However, a problem is that such large end time T_2 cannot be used in Figure 2.5 which has very delicate weak edges in an image. The profile (e) has peaks at corners but too bumpy shape (e-1) which gives false information for finding positions of corners. The profile (f-1) shows the best result which has four sharp peaks at corners and flatter shape on homogeneous regions in the image.

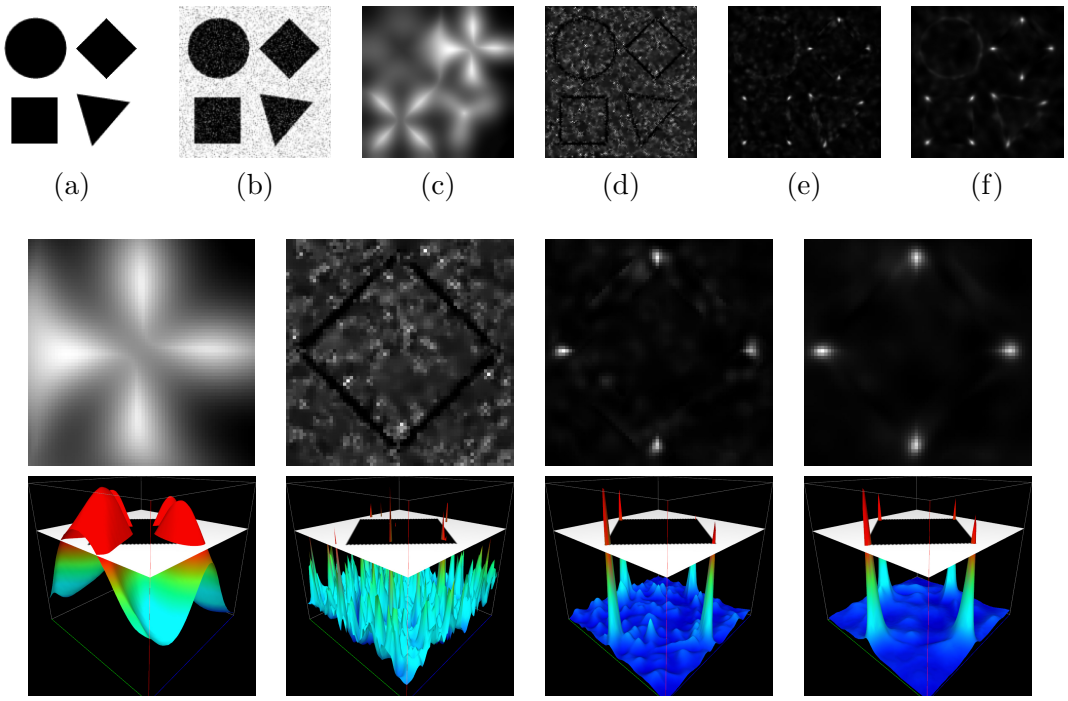
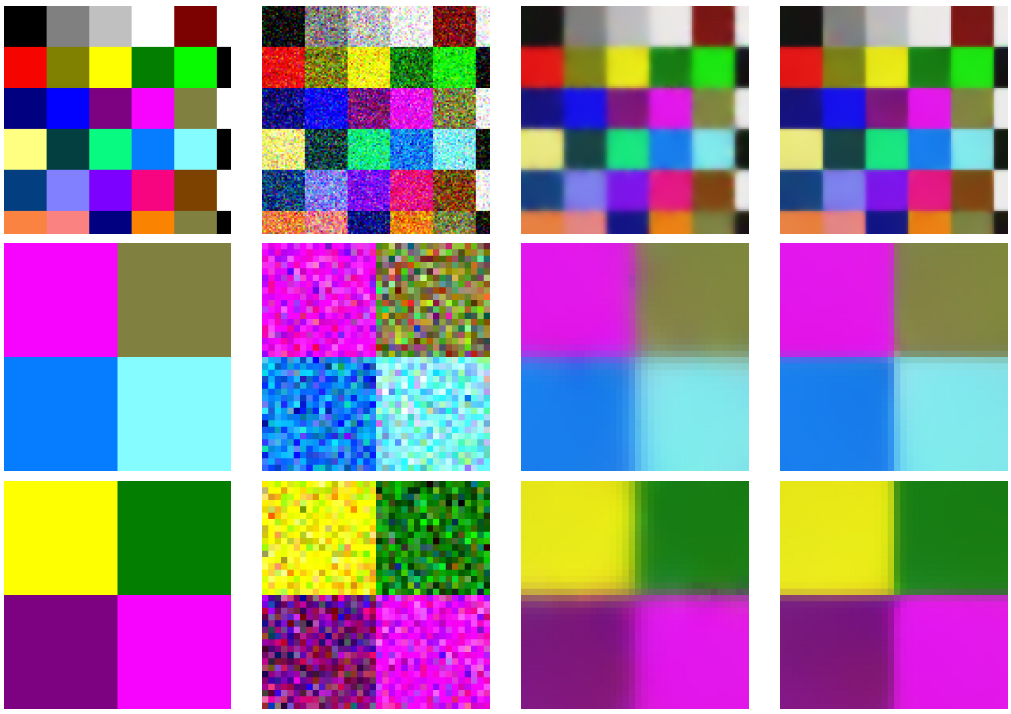


Figure 4.1: (a) is an original image. In (b), we add the Gaussian white noise with zero mean and the standard deviation 50 (SNR $\simeq$ 14.49). From (c) to (f), they are profiles of minimum eigenvalues from different regularized tensors in (B.1), (B.2), (B.3), and (B.4) at end time $T_2 = 70$, respectively. Images from (c-1) to (f-1) are magnified from a part on the top right rectangle from (c) to (f) and we plot intensity graphs of the minimum eigenvalue with the part of the image. The profile (f-1) shows the best result which has four sharp peaks at corners and flatter shape on homogeneous regions in the image.



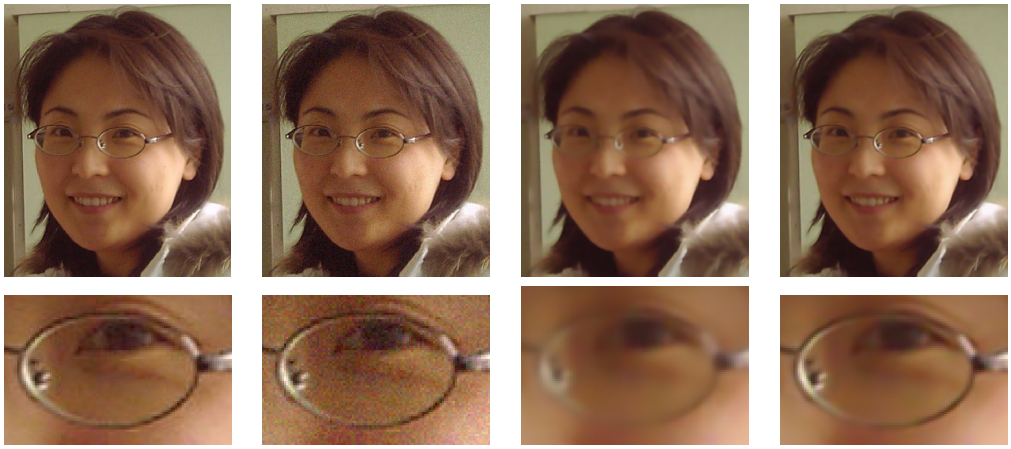
(a) (b) (c) (d)

Figure 4.2: (a) is an original clean image. In (b), we add the Gaussian white noise with zero mean and the standard deviation 50 (SNR $\simeq$ 11.97). We use end time $T_2 = 5$ for obtaining a diffusivity matrix in (4.2). The result (c) is a combination of the Perona-Malik model for a color image with a fidelity term and a shock filter. (d) is the result of proposed method (4.2) which preserves corners and edges very well. In the second and the third row, we magnify two parts in the first row.

4.2 Image denoising and enhancement

In Section 3, we extended the proposed PDE (2.4) for a color image by changing an initial condition. The maximum eigenvalue of a regularized tensor from the extended PDE (3.7) is the rate of maximum change and the corresponding eigenvector provides a direction of maximal change, which is not in the tensor but in the color image. In this section, we extend the image denoising PDE (A.3) for a color image and we combine it with a fidelity term in [18] and a shock filter [19] to enhance and denoise the color image simultaneously. In a research of image enhancement, this kind of combination was introduced in [20] to overcome a problem that a PDE-based shock filter in [21] enhances some noise in an image. Alvarez and Mazorra [22] combined a shock filter and an anisotropic diffusion to solve the same problem. Weickert [19] used a structure tensor to find out more accurate gradient information of an image and used it to enhance the image coherently based on a shock filter. In these formulations, the first derivative information of an image is crucial to enhance and denoise an image simultaneously.

Now, we extend the image denoising PDE (A.3) for a color image with a given noisy



(a) (b) (c) (d)

Figure 4.3: (a) is an original image which has jpeg artifacts. In (b), we add the Gaussian white noise with zero mean and the standard deviation 10. We use the end time $T_2 = 10$ for obtaining a diffusivity matrix in (4.2). (c) and (d) are obtained in the same way as in Figure 4.2-(c) and 4.2-(d), respectively. The result in (d) from the proposed method preserves significant features around the frame of glasses. In the second row, we magnify a region around the right eye in the first row.

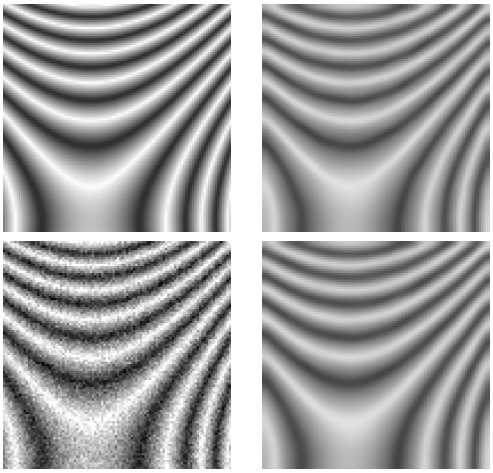
image $I^0 \colon \Omega \subset \mathbf{R}^2 \to \left(\mathbf{R}^+\right)^3$:

$$\begin{aligned} \frac{\partial I_k}{\partial t}(x,t) = \nabla \cdot \left(g\big(U(x,T_2)\big)\nabla I_k\right) \quad &\text{in} \quad \Omega \times (0,T_1], \\ I_k(x,0) = I_k^0(x) \quad &\text{on} \quad \Omega, \\ \frac{\partial u_{ij}}{\partial \tau}(x,\tau) = \nabla \cdot \left(g\big(U_\sigma\big)\nabla u_{ij}\right) \quad &\text{in} \quad \Omega \times (0,T_2], \\ u_{ij}(x,0) = \left(\sum_{k=1}^{3} w_k(x,t) w_k(x,t)^{\mathrm{T}}\right)_{ij} \quad &\text{on} \quad \Omega, \end{aligned} \tag{4.1}$$

where $w_k(x,t) = \nabla(G_\rho * I_k(x,t))$ for $k \in \{1,2,3\}$. We combine the first equation with a fidelity term in [18] and a shock filter [19] to enhance and denoise a color image simultaneously:

$$\frac{\partial I_k}{\partial t}(x,t) = \nabla \cdot \left(g\big(U(x,T_2)\big)\nabla I_k\right) + C_f\left(I_k - I_k^0\right) - C_s \operatorname{sgn}\left(\mathcal{D}^2_{v_\Lambda} I_k\right)|\nabla I_k|, \tag{4.2}$$

where C_s and C_f are constants and v_Λ is the eigenvector corresponding to the maximum eigenvalue of $U(x,T_2)$. Note that the PDE in [20] used an weighted constant C_s depending on a position in an image. In Figures 4.2 and 4.3, we show the quality of image enhancement using our regularized tensor. Since an initial image has large noise, we need large end time T_2 for regularizing a tensor in (4.1). $C_s = 0.02$ and $C_f = 255 \times 10^{-6}$ are chosen in all



(a) (b)

Figure 4.4: Images in (a) are an original clean image at the top row and an image at the bottom row with the Gaussian white noise with zero mean and the standard deviation 20 (SNR $\simeq$ 14.56). (b) is a result of the proposed model (4.3) at $T_1 = 17$. We use different end time $T_2 = 1$ at the top row and $T_2 = 5$ at the bottom row. The result preserves ramp structure of the original image.

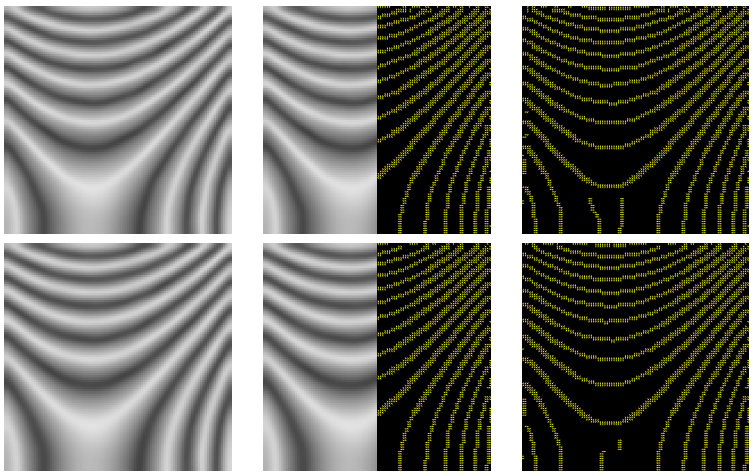
computations. Instead of the diffusivity matrix $g(U(x, T_2))$ in (4.2), if we use

$$h\left(\sum_{k=1}^{3} \left|\nabla G_{\sqrt{2T_2}} * I_k\right|^2\right),$$

then it becomes a combination of the PM model for a color image with a fidelity term and a shock filter. Note that our regularized tensor improves the quality of image enhancement because it properly represents derivative information of a color image.

4.3 Ramp preserving denoising

In medical image processing, regularizing a data $I_0 \colon \Omega \subset \mathbf{R}^2 \to \mathbf{R}^+$, that preserves ramp structure of the data is useful [23]. To obtain ramp preserving denoising, we take a different initial condition for tensor regularization in (A.3). Let $w = (w_1, w_2) = \nabla I$ and regard w as an image with two channels. Then a jump of slope in the data I is considered as an edge of w. Ramp preserving denoising means that a data I is smoothed fast along edges of w and slowly across them. Therefore, we need a tensor whose maximum eigenvalue has a local maximum on edges of w along a direction of the corresponding eigenvector. Now, we



(a) (b) (c)

Figure 4.5: (a) is same images in Figure 4.4-(b). (c) is a vector field obtained by (2.8) and (2.9) from a tensor in (4.3). The position of yellow vector field indicates where the maximum eigenvalue of a tensor has a local maximum. The noisy data in Figure 4.4-(a) is diffused slowly across yellow regions in the proposed model (4.3) and ramp structure is preserved. In (b), we show both the left of (a) and the right of (c).

propose the following denoising scheme:

$$\begin{aligned} \frac{\partial I}{\partial t}(x,t) &= \nabla \cdot \big(g\big(U(x,T_2)\big)\nabla I\big) \quad \text{in} \quad \Omega \times (0,T_1], \\ I(x,0) &= I_0(x) \quad \text{on} \quad \Omega, \\ \frac{\partial u_{ij}}{\partial \tau}(x,\tau) &= \nabla \cdot \big(g\big(U_\sigma\big)\nabla u_{ij}\big) \quad \text{in} \quad \Omega \times (0,T_2]\,, \\ u_{ij}(x,0) &= \left(\sum_{k=1}^{2} \nabla w_k(x,t) \nabla w_k(x,t)^{\mathrm{T}}\right)_{ij} \quad \text{on} \quad \Omega, \end{aligned} \tag{4.3}$$

where $\big(w_1(x,t), w_2(x,t)\big) = \nabla I(x,t)$. A regularized tensor has an advantage of a directional smoothing with respect to structure of ∇I while a regularized tensor in (2.4) is smoothed with respect to structure of I. The maximum eigenvalue of the tenor in (4.3) has a local maximum on edges of w, *i.e.*, a jump of slope in a data, and it regularizes the data while preserving structure of ramps. On the contrary, the maximum eigenvalue of the tensor in (2.4) has a local maximum on edges of I and it regularizes image while preserving structure of edges.

In Figure 4.4, we show a result of the proposed model (4.3). A result of the proposed model preserves ramp structure in an original data even though we use different end time $T_2 = 1$ and $T_2 = 5$ for regularizing a tensor. In Figure 4.5, we show the vector field which is obtained by (2.8) and (2.9) from a tensor in (4.3). The position of yellow vector field indicates where the maximum eigenvalue of a tensor has a local maximum. Therefore, the noisy data in Figure 4.4-(a) is diffused slowly across yellow regions and ramp structure is preserved in the proposed model.

5 Conclusions

We proposed a nonlinear PDE for regularizing a tensor which contains the first derivative information of an image such as strength of edges and a direction of the gradient of the image. Unlike a typical diffusivity matrix which consists of derivatives of a tensor data, we proposed a diffusivity matrix which consists of the tensor data itself. This allows directional smoothing for a tensor along edges which are not in the tensor but in the image. That is, the tensor data is diffused fast along edges of image but slowly across them. The proposed PDE provides a regularized tensor which adapts to the first derivative information of an image as time evolves. We also proved the uniqueness and existence of solution to the proposed PDE. Since we obtained a regularized tensor which properly represents the first derivative information of an image, our regularized tensor improves the quality of image denoising, image enhancement, and corner detection. Moreover, we obtained a ramp preserving denoising process by simply changing an initial condition for regularizing a tensor.

Appendix

A PDEs for image denoising

We list PDEs for image denoising using different regularized derivative information of an image with an initial noisy image $I_0 \colon \Omega \subset \mathbf{R}^2 \to \mathbf{R}^+$.

1. The Perona-Malik model for image denoising (1.2):

$$\begin{aligned} \frac{\partial I}{\partial t}(x,t) = \nabla \cdot \Big(h\Big(\big| \nabla G_{\sqrt{2T_2}} * I \big|^2 \Big) \nabla I \Big) \quad &\text{in} \quad \Omega \times (0, T_1], \\ I(x,0) = I_0(x) \quad &\text{on} \quad \Omega. \end{aligned} \tag{A.1}$$

2. The nonlinear PDE model for image denoising with a structure tensor (1.3):

$$\begin{aligned} \frac{\partial I}{\partial t}(x,t) = \nabla \cdot (g(U(x,T_2)) \nabla I) \quad &\text{in} \quad \Omega \times (0, T_1], \\ I(x,0) = I_0(x) \quad &\text{on} \quad \Omega, \\ \frac{\partial u_{ij}}{\partial \tau}(x,\tau) = \triangle u_{ij} \quad &\text{in} \quad \Omega \times (0, T_2], \\ u_{ij}(x,0) = \Big(\nabla I_\rho(x,t) \nabla I_\rho(x,t)^{\mathrm{T}} \Big)_{ij} \quad &\text{on} \quad \Omega. \end{aligned} \tag{A.2}$$

   Note that $U(x,T_2) = G_{\sqrt{2T_2}} * \big(\nabla I_\rho \nabla {I_\rho}^{\mathrm{T}} \big)$. This model is one of variant of models in [3, 9, 14, 24].

3. The nonlinear PDE model for image denoising with our regularized tensor (2.4):

$$\begin{aligned} \frac{\partial I}{\partial t}(x,t) = \nabla \cdot (g(U(x,T_2)) \nabla I) \quad &\text{in} \quad \Omega \times (0, T_1], \\ I(x,0) = I_0(x) \quad &\text{on} \quad \Omega, \\ \frac{\partial u_{ij}}{\partial \tau}(x,\tau) = \nabla \cdot \big(g\big(U_\sigma\big) \nabla u_{ij} \big) \quad &\text{in} \quad \Omega \times (0, T_2], \\ u_{ij}(x,0) = \Big(\nabla I_\rho(x,t) \nabla I_\rho(x,t)^{\mathrm{T}} \Big)_{ij} \quad &\text{on} \quad \Omega. \end{aligned} \tag{A.3}$$

A boundary condition of diffusion processes for an image and a tensor is the free flux condition. The value ρ is taken small enough to compute ∇I_ρ.

B PDEs for tensor regularization

We summarize PDEs for tensor regularization with a given image $I\colon \Omega \subset \mathbf{R}^2 \to \mathbf{R}^+$. A boundary condition is the free flux condition and an initial condition is a tensor which contains derivative information of an image:

$$U(x,0) = \nabla I(x) \nabla I(x)^{\mathrm{T}} \quad \text{on} \quad \Omega.$$

1. The linear PDE model [3]:

$$\frac{\partial u_{ij}}{\partial \tau}(x,\tau) = \triangle u_{ij} \quad \text{in} \quad \Omega \times (0, T_2]\,. \tag{B.1}$$

Note that $U(x,T_2)$ is exactly same as the structure tensor $G_{\sqrt{2T_2}} * \left(\nabla I(x) \nabla I(x)^{\mathrm{T}}\right)$.

2. The nonlinear PDE model [9]:

$$\frac{\partial u_{ij}}{\partial \tau}(x,\tau) = \nabla \cdot \left(h\left(\sum_{i,j} |\nabla u_{ij}|^2 \right) \nabla u_{ij} \right) \quad \text{in} \quad \Omega \times (0, T_2]\,. \tag{B.2}$$

3. The nonlinear PDE model with a structure tensor [9]:

$$\frac{\partial u_{ij}}{\partial \tau}(x,\tau) = \nabla \cdot \left(g\left(\sum_{i,j} \nabla u_{ij} \nabla u_{ij}{}^{\mathrm{T}} \right) \nabla u_{ij} \right) \quad \text{in} \quad \Omega \times (0, T_2]\,. \tag{B.3}$$

4. The proposed PDE model (2.4):

$$\frac{\partial u_{ij}}{\partial \tau}(x,\tau) = \nabla \cdot \big(g\big(U_\sigma\big) \nabla u_{ij} \big) \quad \text{in} \quad \Omega \times (0, T_2]\,. \tag{B.4}$$

References

[1] P. Perona and J. Malik, "Scale space and edge detection using anisotropic diffusion," IEEE Trans. Pattern Anal. Machine Intell., Vol. 12, pp. 629–639, 1990.

[2] F. Catté, P.L. Lions, J.M. Morel, and T. Coll, "Image selective smoothing and edge detection by nonlinear diffusion," SIAM J. Numer. Anal., Vol. 29, pp. 182–193, 1992.

[3] J. Weickert, "Coherence-enhancing diffusion filtering," Int. J. Comput. Vis., Vol. 31, pp. 111–127, 1999.

[4] J. Bigün and G. H. Granlund, "Optimal orientation detection of linear symmetry," In First International Conference on Computer Vision, London, 1987, pp. 433–438.

[5] M. Kass and A. Witkin, "Analyzing oriented patterns," Comput. Vis. Graph. Image Process., Vol. 37, pp. 362–385, 1987.

[6] B. Jähne, Digital image processing, Springer-Velarg: Berlin, Heidelberg, 2005.

[7] J. Bigün, G. H. Granlund, and J. Wiklund, "Multidimensional orientation estimation with applications to texture analysis and optical flow," IEEE Trans. Pattern Anal. Machine Intell., Vol. 13, pp. 1349–1356, 1991.

[8] J. Weickert and C. Schnörr, "A theoretical framework for convex regularizers in PDE-based computation of image motion," Int. J. Comput. Vis., Vol. 45, pp. 245–264, 2001.

[9] T. Brox, J. Weickert, B. Burgeth, and P. Mrázek, "Nonlinear structure tensors," Image Vis. Comput., Vol. 24, pp. 41–55, 2006.

[10] G. Gerig, O. Kübler, R. Kikinis, and F. A. Jolesz, "Nonlinear anisotropic filtering of MRI data," IEEE Trans. Med. Imaging, Vol. 1, pp. 221–232, 1992.

[11] G. J. M. Parker, J. A. Schnabel, M. R. Symms, D. J. Werring, and G. J. Barker, "Nonlinear smoothing for reduction systematic and random errors in diffusion tensor imaging," J. Mag. Reson. Imaging., Vol. 11, pp. 702–710, 2000.

[12] D. Tschumperlé and R. Deriche, "Diffusion tensor regularization with constraints preservation," In IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Kauai Marriott, Hawaii, 2001, pp. 948–953.

[13] D. Tschumperlé and R. Deriche, "Orthonormal vector sets regularization with PDE's and applications," Int. J. Comput. Vis., Vol. 50, pp. 237–252, 2002.

[14] J. Weickert and T. Brox, "Diffusion and regularization of vector- and matrix-valued images," In M. Z. Nashed, O. Scherzer (Eds.): Inverse Problems, Image Analysis, and Medical Imaging. Contemporary Mathematics, vol. 313, AMS, Providence, RI, 2002, pp. 251–268.

[15] C.L. Evans, Partial Differential Equations, Graduate Studies in Mathematics, Vol. 19, American Mathematical Society: Providence, Role Island, 1998.

[16] H. Brezis, Analyse fonctionnelle, Dunod: Masson, Paris, 1992.

[17] U. Köthe, "Edge and junction detection with an improved structure tensor," In B. Michaelis, G. Krell (Eds.): Pattern Recognition, Proc. of 25th DAGM Symposium, Magdeburg, Lecture Notes in Computer Science 2781, 2003, pp. 25–32.

[18] L. I. Rudin, S. Osher, and E. Fatemi, "Nonlinear total variation based noise removal algorithms," Physica D, Vol. 60, pp. 259–268, 1992.

[19] J. Weickert, "Coherence-enhancing shock filters," In B. Michaelis, G. Krell (Eds.): Pattern Recognition, Proc. of 25th DAGM Symposium, Magdeburg, Lecture Notes in Computer Science 2781, 2003, pp. 1–8.

[20] P. Kornprobst, R. Deriche, and G. Aubert. "Image coupling, restoration and enhacement via PDE's." In Proc. of the International Conference on Image Processing, vol. 2, Santa Barbara, California, 1997, pp. 458–261.

[21] S. Osher and L. I. Rudin, "Feature-oriented image enhancement using shock filters," SIAM J. Numer. Anal., Vol. 27, pp. 919–940, 1990.

[22] L. Alvarez and L. Mazorra, "Signal and image restoration using shock filters and anisotropic diffusion." SIAM J. Numer. Anal., Vol. 31, pp. 590–605, 1994.

[23] S. Lee, J.K. Seo, C. Park, B.I. Lee, E.J. Woo, S.Y. Lee, O. Kwon, and J. Hahn, "Conductivity image reconstruction from defective data in MREIT: Numerical simulation and animal experiment," IEEE Trans. Med. Imaging, Vol. 25, pp. 168–176, 2006.

[24] J. Weickert, "Scale-space properties of nonlinear diffsusion filtering with a diffusion tensor," Tech. Rep. 110, Laboratory of Technomathematics, University of Kaiserslautern, Germany, Oct. 1994.