On a system of adaptive coupled PDEs for image restoration

V. B. Surya Prasath* and D. Vorotnikov†

Abstract

In this paper, we consider a coupled system of partial differential equations (PDEs) based model for image restoration. Both the image and the edge variables are incorporated by coupling them into two different PDEs. It is shown that the initial-boundary value problem has global in time dissipative solutions (in a sense going back to P.-L. Lions), and several properties of these solutions are established. Some numerical examples are given to highlight the denoising nature of the proposed model along with some comparison results.

Keywords: Image restoration, Coupled PDE, Nonlinear diffusion, Edge variable, Wellposedness, Dissipative solutions.

1 Introduction

Partial differential equation (PDE) based image restoration is now a well-researched area within the image processing community [?, ?, ?]. Starting with the parabolic paradigm of Perona and Malik [?] a wide variety of PDEs have been studied for the past past two decades. Among a wealth of PDE based schemes available for image restoration we mention total variation [?, ?, ?, ?], Shock filters [?, ?] and fourth order PDEs [?, ?, ?, ?, ?, ?, ?, ?, ?] based approaches. Other approaches include combining different type of PDEs [?, ?, ?, ?], integro-differential equations [?], fractional anisotropic diffusion [?, ?, ?, ?] etc.

Most of these schemes use the absolute value of the gradient image as a guiding road map in the diffusion process to restore the noisy images. It is well–known that under noisy conditions gradient map can give spurious oscillations [?] in the restoration process. There have been numerous efforts to improve/built upon the successful restoration results obtained with the classical PDEs and to avoid gradient based artifacts. Based on the approach they take, we can classify such improvements into two broad categories: (a) adaptive schemes [?, ?, ?, ?, ?, ?] - a single PDE with some kind of adaptive edge map estimation included and (b) coupled PDEs [?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?] - a separate PDE for

*Department of Computer Science, University of Missouri-Columbia, MO 65211, USA. Email: prasaths@missouri.edu
†CMUC, Department of Mathematics, University of Coimbra, 3001-454 Coimbra, Portugal. Email:

estimating a better edge map. Separate estimation of the edge map for restoring noisy images can be considered as solving for an *edge variable* along with image variable.

Starting with the pioneering work of Geman and Geman [?] various researchers have studied the concept of a separate edge variable. For example, half–quadratic method studied by Charbonnier et al [?] compute the edge variable separately using an alternative minimization scheme. This type of coupled edge variable computation has connections to the famous Mumford–Shah functional [?] in image segmentation, for example phase field method [?] utilizes a sort of inverse edge variable, see also [?]. Another approach is to statistically model the edges present in an image and treat them in Markov random field theory [?, ?]. In this case, the edge variable is known as edge prior and can be utilized in finding the contours of objects present.

In this paper, we study a coupled PDE which combines the Gaussian smoothing based regularization approach of Catté et al [?] with that of the Perona–Malik anisotropic PDE [?]. The PDE for the edge variable is devised using a balanced approach which interpolates between the spatial smoothing approach with that of the anisotropic diffusion. It is shown that the corresponding Dirichlet initial-boundary value problem possesses global in time *dissipative solutions*; uniqueness, regularity and some other properties of these solutions are studied. The concept of dissipative solution was suggested in [?] for the Euler equations of ideal fluid flow. Later, existence of dissipative solutions was established for Boltzmann's equation [?, ?], the ideal MHD equations [?], Navier-Stokes-Maxwell equations [?], Euler-α and Maxwell-α models [?] and viscoelastic diffusion equations [?].

The features of our problem (8)–(11) which oppose strong and classical weak wellposedness are the presence of a nonlinear function (modulus) of the gradient of u in the right-hand side of (9) and the Perona-Malik-like form of g. The inequality (19) in the definition of dissipative solutions turns out to contain the absolute value function as well. Therefore, unlike in the previous works on dissipative solutions, it is impossible to pass to the limit in this inequality via weak and weak-* compactness argument. Nevertheless, we manage to do it via strong compactness, although it is not sufficiently strong to obtain classical (i.e. not dissipative) weak solutions. Numerical comparison of the results with anisotropic diffusion PDEs and coupled PDEs is undertaken on noisy synthetic and real images, highlighting the advantages of the proposed model.

Rest of the paper is organized as follows. Section 2 introduces diffusion PDE models in image restoration and the coupled PDE studied in this paper. Section 3 presents the wellposedness theory for the model. Section 4 gives some numerical examples to illustrate the effect of the proposed approach against some well–known PDE based schemes. Finally, Section 5 concludes the paper.

2 Diffusion for image restoration

2.1 Anisotropic diffusion

Perona–Malik [**?**] considered the following anisotropic diffusion PDE to improve the denoising capabilities of the linear diffusion

$$\frac{\partial u}{\partial t} = div\left(g\left(|\nabla u|\right)\nabla u\right) \tag{1}$$

with $u(0) = u_0$, i.e. the input noisy image is the initial datum, and the above PDE is run for a finite time $T > 0$ to obtain denoised image $u(\cdot, T)$. The choice of the diffusion function $g : [0, \infty) \to [0, \infty)$ is important in controlling the smoothing and even enhancement of edges. In [**?**] the following two diffusion functions are considered

$$g_{pm1}(s) = \frac{1}{1 + (s/K)^2}, \qquad g_{pm2}(s) = \exp\left(-(s/K)^2\right) \tag{2}$$

where $K > 0$ is the contrast parameter. Separating and finding edges from a digital image is a well studied problem. Due to the usage of edge maps (via the diffusion coefficient function $g(\nabla u)$) in the restoration process a well-defined edge modelling can give better denoising results. Catté et al [**?**] in their pioneering work to make the Perona–Malik type PDE work better as well as to prove wellposedness introduced the following modification

$$\frac{\partial u}{\partial t} = div\left(g(|G_\sigma \star \nabla u|)\nabla u\right) \tag{3}$$

where $G_\sigma(x) = (2\pi\sigma)^{-1} \exp -\left(|x|^2 / 2\sigma\right)$ is the Gaussian kernel and the operation $\star$ means convolution. This introduction of spatial pre-smoothing not only made the gradient computation robust to outliers it also provided a smooth edge map for the diffusion to operate upon. Following Koenderink [**?**] one can observe that such a Gaussian smoothing is equivalent to solving the following linear diffusion equation up to time $T = \sigma/2$

$$v' = \Delta v$$

with initial datum $v(0, x) = \nabla u(t, x)$, and consequently substitute the Catté et al.'s modification with the following coupled PDE

$$u' = div\left(g(v)\nabla u\right), \quad v' = \Delta v = div(\nabla v).$$

The models following the above idea of using a separate PDE to create better edge maps, which rely not only on the absolute value of gradient, have been studied by some researchers in the past [**?**, **?**, **?**, **?**].

2.2 Proposed coupled PDE model

In this paper we consider the following coupled PDE which combines both the Perona–Malik PDE (1) and Catté et al's spatially regularization framework (3),

$$\frac{\partial u}{\partial t} = div\,(g(v)\nabla u) \tag{4}$$

$$\frac{\partial v}{\partial t} = \lambda\, div(\nabla v) + (1-\lambda)\,(|\nabla u| - v) \tag{5}$$

where $g(s) = \frac{1}{1+(s/K)^2}$ (Perona-Malik type diffusion function) or $g(s) = |v|^{-1}$ (total variation diffusion function). The balancing parameter $0 \leq \lambda \leq 1$ is an important parameter, see Section 2.3 below.

The first PDE is the usual Perona–Malik type PDE. Here it is modified and instead of using a gradient based diffusion function $g = g(|\nabla u|)$, we separate it into another variable v and incorporate into that function $g = g(v)$. Note that the gradient $|\nabla u|$ acts like an edge map computed from the image u and is prone to noise and can lead to staircasing artifacts. So this separation will give better restoration as we can control the edge map better by using a separate PDE. The second term in Eqn. (5) is important as it constrains the variable v to be like $|\nabla u|$, i.e $v \sim |\nabla u|$. The parameter λ which appears in the second PDE (5) balances between the PM model (1) and the Catté et al's model (3). Hence it is important in localizing denoising effects of the diffusion based scheme. That is, Catté et al's model can lead to poor edge localization if the pre-smoothing is higher whereas the PM model can lead to staircasing artifacts in flat regions of the image. A balanced model can avoid both these drawbacks and can give better results. A related model to the proposed coupled system is that of Nitzberg and Shoita [?] who considered the following relaxation model:

$$\frac{\partial u}{\partial t} = div\,(g(v)\nabla u)$$

$$\frac{\partial v}{\partial t} = wG_\sigma \star |\nabla u|^2 - wv$$

where $w > 0$ relaxation parameter. We give a brief overview of the known mathematical treatments of the systems close to (4)–(5) in Section 3, Remark 4.

2.3 Role of balancing parameter

As λ approaches unity, the proposed model Eqns. (4-5) behaves like the Catté et al's spatial smoothing based PDE model Eqn. (3), and hence the denoised image lack edge localization. Figure 1 shows that as $\lambda \to 1$ restored images tends to become smoothed along the edges as well as around them. This effect can be better explained by looking at the corresponding edge variable v in Figure 1.

To fix the parameter in an automatic way here we consider an adaptive approach based on probable image edges found at multiple scales. Note that the edge variable $v(t, x)$ gives an estimate of edges present at scale t and for a fixed λ certain scale edges are retained, compare for example Figure 1(a) and

(abc)

λλλ=

00240

(def)

λλλ=

00680

Figure 1: Influence of balancing parameter λ in the restoration process by the proposed coupled PDE Eqns. (4-5) as the λ value increases from 0 to 1. Noisy *Peppers* image is used as the initial image u_0 with noise level $\sigma_n = 30$. In each sub-figure top shows the denoised image u and the bottom is the corresponding edge variable v. We refer to the electronic version for better visualization of the fine scale details in the edge variable images.

(abc)

Figure 2: Examples to highlight adaptive λ in denoising using (a) synthetic piecewise constant *Shapes* image, (b) piecewise smooth *Peppers* image, and (c) strongly textured *Barbara* image. (Top row) denoised images u, (middle rows) some regions taken from each image showing flat regions with no staircasing artifacts, other edge, texture details are well preserved under the coupled PDE model (bottom row) corresponding adaptive λ computed from the edge variable v using Eqn. (6).

(f). When $\lambda = 0$ small scale edges as well as some staircasing artifacts are visible in flat regions of the middle pepper (Figure 1(a) bottom) whereas when $\lambda = 1$, except some big scale edges other features are washed away. A simple way to combine probable edges found by the edge variable is to sum them up

$$\lambda = \lambda(x) = \sum_{\tau=0}^{t-1} G_{\sigma_\tau} \star v(\tau, x) \tag{6}$$

where G_{σ_τ} represent Gaussian kernels with half-width $\sigma_\tau > 0$. At $t = 0$ we fix $\lambda = 0.05$ uniformly and further iterations follow Eqn. (6) with $\sigma_\tau = 1/\tau^2$. The multiscale Gaussian pre-smoothing is done to avoid outliers in the edge variable causing oscillations in the restoration process. Moreover, as the iteration t increases, due to the smoothing property of the diffusion PDE noise is reduced and hence Gaussian filter width is reduced accordingly to avoid losing fine scale edges. Note that Eqn. (6) sums edge maps found at all the previous iterations from t at zero to $t-1$. Figure 2 shows three different standard test images and their denoised version using the coupled PDE Eqns. (4-5) with adaptive λ using formula in Eqn. (6). Note the near perfect recovery of piecewise constant *Shapes* image in Figure 2(a). The scheme does preserve piecewise smooth *Peppers* image in Figure 2(b) without any staircasing artifacts usually associated with Perona and Malik type PDE based schemes. In the textured *Barbara* image, Figure 2(c), the scheme does preserve textures but small scale textures are removed due to the Gaussian smoothing utilized in the adaptive parameter term λ.

Remark 1. *The parameter λ in the proposed coupled PDE is related to the regularization parameter selection problem from variational minimization. Gilboa et al [**?**] used the relation to propose an adaptive parameter for denoising partially textured images.*

Remark 2. *Further adaptation of the balancing parameter λ is also possible, for example, $\lambda = \lambda(x, u(t, x))$. Such consideration can lead to a more general restoration model and will be studied elsewhere.*

Remark 3. *Nordstörm [**?**] proposed a biased version following the relation between the PDE and variational minimization methods*

$$\frac{\partial u}{\partial t} = div\left(g\left(|\nabla u|\right)\nabla u\right) - \lambda(u - u_0) \tag{7}$$

The term on the right hand side of the above equation comes from the data fidelity and is added to keep the restored image diverging far away from the input image u_0. Here we do not consider this term in the restoration step (PDE for u) and instead utilize it in the edge variable step (PDE for v).

3 Wellposedness of the problem

The objective of this section is to prove Theorem 1 concerning existence, uniqueness, regularity and some other properties of dissipative solutions to the problem

$$\frac{\partial u(t,x)}{\partial t} = \mathrm{div}(g(v(t,x))\nabla u(t,x)), \tag{8}$$

$$\frac{\partial v(t,x)}{\partial t} - \lambda(x)\Delta v(t,x) = (1-\lambda(x))(|\nabla u(t,x)| - v(t,x)), \tag{9}$$

$$u\Big|_{\partial\Omega} = 0, v\Big|_{\partial\Omega} = 0, \tag{10}$$

$$u|_{t=0} = u_0,\ v|_{t=0} = v_0. \tag{11}$$

We consider the simplest Dirichlet boundary condition (10), but other boundary conditions can also be handled.

In the section, Ω is considered to be a bounded *domain* (i.e. an open set in $\mathbb{R}^2$) possessing the *cone property*. We recall [?] that this means that each point $x \in \Omega$ is a vertex of a *finite cone* C_x contained in Ω, and all these cones C_x are congruent. A finite cone is a set of the form

$$C_x = B_1 \cap \{x + \xi(y-x) | y \in B_2, \xi > 0\}$$

where B_1 and B_2 are open balls in $\mathbb{R}^2$, B_1 is centered at x, and B_2 does not contain x. Obviously, rectangular domains have this property.

The symbol C will stand for a generic positive constant that can take different values in different lines. We sometimes write $C(\dots)$ to specify that the constant depends on a certain parameter or value.

We assume that $g : \mathbb{R} \to \mathbb{R}$, $\frac{1}{\sqrt{g}}$ and $\lambda : \Omega \to \mathbb{R}$ are Lipschitz functions having positive values, g is bounded, $\lambda \le 1$,

$$\lambda_0 = \inf_{x\in\Omega} \lambda(x) > 0.$$

The assumptions on g hold, for instance, if

$$g(s) = \frac{a}{b + c|s|^d}, \tag{12}$$

where a, b, c, d are positive numbers, and $1 \le d \le 2$.

Note that

$$\frac{1}{\sqrt{g(s)}} \le \Big|\frac{1}{\sqrt{g(s)}} - \frac{1}{\sqrt{g(0)}}\Big| + \frac{1}{\sqrt{g(0)}} \le C(g)(1+|s|). \tag{13}$$

Remark 4. *In [?, ?], equation (8) is considered to be coupled with*

$$\frac{\partial v}{\partial t} + v = F(|\nabla u|^2), \tag{14}$$

where F is a smooth function (instead of coupling with (9)). The resulting model coincides with the Nitzberg–Shiota one [?] if $F(\xi) = \xi$, and with our model provided $\lambda \equiv 0$ and $F(\xi) = \sqrt{\xi}$ (non-smooth at zero). Existence and uniqueness of local in time strong solutions is proved in [?]. Global in time weak solution is shown to exist in [?] provided F is uniformly bounded (thus excluding the Nitzberg–Shiota model). Another time averaging model, with (14) replaced by

$$v(t,x) = \int_{-\infty}^{+\infty} |\nabla u(s,x)|^2 \theta(t-s)\, ds, \tag{15}$$

with fixed function θ, is studied in [?]. Global in time strong wellposedness is established when the support of θ is bounded, lies in the positive semi-axis and is separated from 0 (if it approaches 0, the local wellposedness takes place). The Nitzberg–Shiota model corresponds to the case $\theta(s) = 0, s < 0$; $\theta(s) = e^{-s}, s \geq 0$, where the support is unbounded and includes 0. Global in time solvability (in any sense) for both Nitzberg–Shiota model and our model with $\lambda \equiv 0$ remains an open problem.

We use the standard notations $L_p(\Omega)$, $W^m_p(\Omega)$, $H^m(\Omega) = W^m_2(\Omega)$ for the Lebesgue and Sobolev spaces. We will often keep the function space symbol and omit Ω.

The Euclidean norm in finite-dimensional spaces is denoted by $|\cdot|$. The symbol $\|\cdot\|$ will stand for the Euclidean norm in $L_2(\Omega)$. The corresponding scalar products is denoted by a dot $\cdot$ and parentheses $(\cdot,\cdot)$.

Let $H^1_0(\Omega)$ be the closure of the set of smooth, compactly supported in Ω, functions in $H^1(\Omega)$. By virtue of Friedrichs' inequality, the Euclidean norm $\|\cdot\|_1$ corresponding to the scalar product

$$(u,v)_1 = (\nabla u, \nabla v)$$

is a norm in H^1_0.

The set $V_2 = H^1_0(\Omega) \cap H^2(\Omega)$ is a Hilbert space with the scalar product

$$(u,v)_2 = (u,v)_1 + \sum_{|\alpha|=2} (D^\alpha u, D^\alpha v).$$

Denote the corresponding Euclidean norm by $\|\cdot\|_2$.

Let V_r, $1 < r < 2$, be the closure of V_2 in W^1_r.

We recall the following abstract observation [?, ?]. Assume that we have two Hilbert spaces, $X \subset Y$, with continuous embedding operator $i : X \to Y$, and $i(X)$ is dense in Y. The adjoint operator $i^* : Y^* \to X^*$ is continuous and, since $i(X)$ is dense in Y, one-to-one. Since i is one-to-one, $i^*(Y^*)$ is dense in X^*, and one may identify Y^* with a dense subspace of X^*. Due to the Riesz representation theorem, one may also identify Y with Y^*. We arrive at the chain of inclusions:

$$X \subset Y \equiv Y^* \subset X^*.$$

Both embeddings here are dense and continuous. Observe that in this situation, for $f \in Y, u \in X$, their scalar product in Y coincides with the value of the functional f from X^* on the element $u \in X$:

$$(f, u)_Y = \langle f, u \rangle. \tag{16}$$

Such triples (X, Y, X^*) are called Lions triples. We use the Lions triples (V_2, L_2, V_2^*) and (H_0^1, L_2, H^{-1}).

The symbols $C(\mathcal{J}; E)$, $C_w(\mathcal{J}; E)$, $L_2(\mathcal{J}; E)$ etc. denote the spaces of continuous, weakly continuous, quadratically integrable etc. functions on an interval $\mathcal{J} \subset \mathbb{R}$ with values in a Banach space E. We recall that a function $u : \mathcal{J} \to E$ is *weakly continuous* if for any linear continuous functional g on E the function $g(u(\cdot)) : \mathcal{J} \to \mathbb{R}$ is continuous.

We require the following spaces

$$W_1 = W_1(\Omega, T) = \{\tau \in L_2(0, T; V_2),\ \tau' \in L_2(0, T; V_2^*)\},$$

$$\|\tau\|_{W_1} = \|\tau\|_{L_2(0,T;V_2)} + \|\tau'\|_{L_2(0,T;V_2^*)},$$

$$W_2 = W_2(\Omega, T) = \{\tau \in L_2(0, T; H_0^1),\ \tau' \in L_2(0, T; H^{-1})\},$$

$$\|\tau\|_{W_2} = \|\tau\|_{L_2(0,T;H_0^1)} + \|\tau'\|_{L_2(0,T;H^{-1})}.$$

Let us introduce the operator

$$A : V_2 \to V_2^*,\ \langle Au, \varphi \rangle = (u, \varphi)_2,$$

where φ is an arbitrary element of V_2.

Denote by $\mathcal{R}$ the following class of pairs of functions:

$$\mathcal{R} = L_{4,loc}(0, \infty; V_2) \cap L_\infty(0, \infty; W_\infty^1) \cap W_{4,loc}^1(0, \infty; L_2)$$

$$\times L_{2,loc}(0, \infty; V_2) \cap L_\infty(0, \infty; L_\infty) \cap W_{2,loc}^1(0, \infty; L_2).$$

Observe that the following expressions, where δ is a positive number, are well-defined for $(w, \tau) \in \mathcal{R}$, and their values are in $L_{2,loc}(0, \infty; L_2)$:

$$E_1(w, \tau, \delta) = -\frac{\partial w}{\partial t} + \delta \operatorname{div}(g(\tau)\nabla w),$$

$$E_2(w, \tau, \delta) = -\frac{\partial \tau}{\partial t} + \lambda \Delta \tau + \delta(1 - \lambda)(|\nabla w| - \tau) + (1 - \delta)(\nabla \tau \cdot \nabla \lambda),$$

$$E_1(w, \tau) = E_1(w, \tau, 1),$$

$$E_2(w, \tau) = E_2(w, \tau, 1).$$

Let us recall the Sobolev inequality

$$\|u\|_{L_\infty} \le C(\Omega)\|u\|_2,\ u \in V_2, \tag{17}$$

and the Ladyzhenskaya inequality [?]

$$\|u^2\| \le \sqrt{2}\|u\|\,\|\nabla u\|,\ u \in H_0^1. \tag{18}$$

The following Gronwall-like lemma will be useful.

Lemma 1. *([?, Lemma 3.1]) Let $f, \chi, L, M : [0,T] \to \mathbb{R}$ be scalar functions, $\chi, L, M \in L_1(0,T)$, and $f \in W_1^1(0,T)$ (i.e. f is absolutely continuous). If*

$$\chi(t) \ge 0, L(t) \ge 0$$

and

$$f'(t) + \chi(t) \le L(t)f(t) + M(t)$$

for almost all $t \in (0,T)$, then

$$f(t) + \int_0^t \chi(s)\,ds \le \exp\left(\int_0^t L(s)ds\right)\left[f(0) + \int_0^t \exp\left(\int_s^0 L(\xi)d\xi\right) M(s)\,ds\right]$$

for all $t \in [0,T]$.

We can now give

Definition 1. *Let $u_0, v_0 \in L_2(\Omega)$. A pair of functions (u,v) from the class*

$$u, v \in C_w([0,\infty); L_2),$$

is called a dissipative solution to problem (8) – (11) if, for all test functions $(\zeta,\theta) \in \mathcal{R}$ and all non-negative moments of time t, one has

$$\gamma^{\|u(t)\|^2}\left[\|u(t)-\zeta(t)\|^2 + \|v(t)-\theta(t)\|^2\right] \le \gamma^{2t+\|u_0\|^2}\Big\{\|u_0-\zeta(0)\|^2 + \|v_0-\theta(0)\|^2 + \int_0^t 2\gamma^{-s}\Big|\big(E_1(\zeta,\theta)(s), u(s)-\zeta(s)\big) + \big(E_2(\zeta,\theta)(s), v(s)-\theta(s)\big)\Big|\Big\} \tag{19}$$

where $\gamma = \gamma(\Omega, g, \lambda, \zeta, \theta) > 1$ is a certain function of $\Omega, g, \lambda, \zeta$ and θ.

Theorem 1. *a) Given $u_0, v_0 \in L_2$, there is a dissipative solution to problem (8) – (11).*

b) This solution (u,v) belongs to $L_{4/3,loc}(0,\infty; V_{-\epsilon+4/3}) \times L_{2,loc}(0,\infty; H_0^1)$, $0 < \epsilon < \frac{1}{3}$.

c) If, for some $u_0, v_0 \in L_2$, there exist $T > 0$ and a strong solution (u_T, v_T) to problem (8) – (11), which is a restriction of a function from $\mathcal{R}$ to $(0,T)$. Then the restriction of any dissipative solution (with the same initial data) to $(0,T)$ coincides with (u_T, v_T).

d) Every strong solution $(u,v) \in \mathcal{R}$ is a (unique) dissipative solution.

e) The dissipative solutions satisfy the initial condition (11).

To prove Theorem 1, we consider the following auxiliary problem:

$$\frac{\partial u}{\partial t} + \varepsilon A u = \delta \operatorname{div}(g(v)\nabla u), \tag{20}$$

$$\frac{\partial v}{\partial t} - \lambda \Delta v = \delta(1-\lambda)(|\nabla u| - v) + (1-\delta)(\nabla v \cdot \nabla \lambda), \tag{21}$$

$$u\Big|_{\partial\Omega} = 0, v\Big|_{\partial\Omega} = 0, \tag{22}$$

$$u|_{t=0} = \delta u_0,\ \ v|_{t=0} = \delta v_0. \tag{23}$$

Here, $\varepsilon > 0$ and $0 \le \delta \le 1$ are parameters. The weak formulation of (20) – (23) is as follows.

Definition 2. *A pair of functions (u,v) from the class*

$$u \in W_1,\ \ v \in W_2$$

is a weak solution to problem (20) – (23) if the equalities

$$\frac{d}{dt}(u,\varphi) + \varepsilon(u,\varphi)_2 + \delta(g(v)\nabla u, \nabla\varphi) = 0, \tag{24}$$

and

$$\frac{d}{dt}(v,\phi) + (\lambda\nabla v, \nabla\phi) + \delta(\nabla v, \phi\nabla\lambda) - \delta\left((1-\lambda)(|\nabla u| - v), \phi\right) = 0 \tag{25}$$

are satisfied for all $\varphi \in V_2,\ \phi \in H_0^1$ almost everywhere in $(0,T)$, and (22) and (23) hold.

Lemma 2. *Let (u,v) be a weak solution to problem (20) – (23). Then, for all test functions $(\zeta,\theta) \in \mathcal{R}$*

and $0 \le t \le T$*, one has*

$$\gamma^{\|u(t)\|^2}\{\|u(t)-\zeta(t)\|^2+\|v(t)-\theta(t)\|^2 + 2\varepsilon\int_0^t \|u(s)-\zeta(s)\|_2^2\,ds+\lambda_0\int_0^t \|v(s)-\theta(s)\|_1^2\,ds\} \le \gamma^{2t+\delta\|u_0\|^2}\Big\{\|\delta u_0-\zeta(0)\|^2+\|\delta v_0-\theta(0)\|^2 + \int_0^t 2\gamma^{-s}\Big|\big(E_1(\zeta,\theta,\delta)(s),u(s)-\zeta(s)\big) + \big(E_2(\zeta,\theta,\delta)(s),v(s)-\theta(s)\big)-\varepsilon(\zeta(s),u(s)-\zeta(s))_2\Big|\,ds\Big\} \quad (26)$$

where $\gamma=\gamma(\Omega,g,\lambda,\zeta,\theta)>1$ *is a certain function of* Ω,g,λ,ζ *and* θ*.*

Proof. Let us first derive the straightforward energy estimate. For almost all $t\in(0,T)$, let $\varphi=u(t)$ in (24). Then[1]

$$\frac{1}{2}\frac{d}{dt}(u,u)+\delta(g(v)\nabla u,\nabla u)+\varepsilon(u,u)_2=0. \quad (27)$$

Integration in time gives

$$\frac{1}{2}\|u(t)\|^2+\int_0^t(\delta g(v(s))\nabla u(s),\nabla u(s))\,ds\le\frac{\delta}{2}\|u_0\|^2. \quad (28)$$

Observe now that

$$\frac{d}{dt}(\zeta,\varphi)+\delta(g(\theta)\nabla\zeta,\nabla\varphi)+(E_1(\zeta,\theta,\delta),\varphi)+\varepsilon(\zeta,\varphi)_2=\varepsilon(\zeta,\varphi)_2, \quad (29)$$

and

$$\frac{d}{dt}(\theta,\phi)+(\lambda\nabla\theta,\nabla\phi)+\delta(\nabla\theta,\phi\nabla\lambda)-\delta\left((1-\lambda)(|\nabla\zeta|-\theta),\phi\right)+(E_2(\zeta,\theta,\delta),\phi)=0. \quad (30)$$

for $\varphi\in V_2,\ \phi\in H_0^1$. Denote $w=u-\zeta$ and $\varsigma=v-\theta$. For almost all $t\in(0,T)$, put $\varphi=w(t)$ and $\phi=\varsigma(t)$. Add the difference between (24) and (29) with the difference between (25) and (30), arriving at

$$\frac{1}{2}\frac{d}{dt}(w,w)+\frac{1}{2}\frac{d}{dt}(\varsigma,\varsigma)+\delta(g(v)\nabla w,\nabla w) +\varepsilon(w,w)_2+(\lambda\nabla\varsigma,\nabla\varsigma)+\delta\left((1-\lambda)\varsigma,\varsigma\right) =-\delta([g(v)-g(\theta)]\nabla\zeta,\nabla w)+\delta\left((1-\lambda)(|\nabla u|-|\nabla\zeta|),\varsigma\right)-\delta(\nabla\varsigma,\varsigma\nabla\lambda) +(E_1(\zeta,\theta,\delta),w)+(E_2(\zeta,\theta,\delta),\varsigma)-\varepsilon(\zeta,w)_2. \quad (31)$$

[1] See e.g. [?, p. 153] on how $\frac{1}{2}$ appears in (27).

Let us estimate the first three terms in the right-hand side.

$$
\begin{aligned}
&-\delta([g(v)-g(\theta)]\nabla\zeta,\nabla w)+\delta\left((1-\lambda)(|\nabla u|-|\nabla\zeta|),\varsigma\right)\\
&\qquad\le C(\zeta,g)\delta(|v-\theta|,|\nabla w|)\\
&\qquad\le C(\zeta,g)\left(\frac{|\varsigma|}{\sqrt{g(v)}},\sqrt{\delta g(v)}|\nabla w|\right)\\
&= C(\zeta,g)\left[\left(\frac{|\varsigma|}{\sqrt{g(0)}},\sqrt{\delta g(v)}|\nabla w|\right)+\left(|\varsigma|\left(\frac{1}{\sqrt{g(\theta)}}-\frac{1}{\sqrt{g(0)}}\right),\sqrt{\delta g(v)}|\nabla w|\right)\right]\\
&\qquad+C(\zeta,g)\left(|\varsigma|\left(\frac{1}{\sqrt{g(v)}}-\frac{1}{\sqrt{g(\theta)}}\right),\sqrt{\delta g(v)}|\nabla w|\right)\\
&\le C(\zeta,\theta,g)\left(|\varsigma|,\sqrt{\delta g(v)}|\nabla w|\right)+C(\zeta,g)\left(\varsigma^2,\sqrt{\delta g(v)}(|\nabla\zeta|+|\nabla u|)\right)\\
&\qquad\le\|\sqrt{\delta g(v)}\nabla w\|^2+C(\zeta,\theta,g)\|\varsigma\|^2+C(\zeta,g)\left(\varsigma^2,\sqrt{\delta g(v)}|\nabla u|\right), \qquad (33)
\end{aligned}
$$

and

$$
-\delta(\nabla\varsigma,\varsigma\nabla\lambda)\le C(\lambda)(\varsigma,\nabla\varsigma)\le\frac{\lambda_0}{4}\|\varsigma\|_1^2+C(\lambda)\|\varsigma\|^2
$$

Now, (31) implies

$$
\begin{aligned}
&\frac{1}{2}\frac{d}{dt}(w,w)+\frac{1}{2}\frac{d}{dt}(\varsigma,\varsigma)+\varepsilon(w,w)_2+\frac{3\lambda_0}{4}\|\varsigma\|_1^2\\
&\qquad\le C(\zeta,\theta,\lambda,g)\left(\varsigma^2,1+\sqrt{\delta g(v)}|\nabla u|\right)+(E_1(\zeta,\theta,\delta),w)+(E_2(\zeta,\theta,\delta),\varsigma)-\varepsilon(\zeta,w)_2.
\end{aligned}
$$

Denote $\Phi(t)=\left\|1+\sqrt{\delta g(v(t))}|\nabla u(t)|\right\|$. Due to (18),

$$
\begin{aligned}
&\frac{d}{dt}(w,w)+\frac{d}{dt}(\varsigma,\varsigma)+2\varepsilon(w,w)_2+\frac{3\lambda_0}{2}\|\nabla\varsigma\|^2\\
&\qquad\le C(\zeta,\theta,\lambda,g)\Phi\|\varsigma\|\|\nabla\varsigma\|+2(E_1(\zeta,\theta,\delta),w)+2(E_2(\zeta,\theta,\delta),\varsigma)-2\varepsilon(\zeta,w)_2.
\end{aligned}
$$

Thus,

$$
\begin{aligned}
&\frac{d}{dt}\|w\|^2+\frac{d}{dt}\|\varsigma\|^2+2\varepsilon\|w\|_2^2+\lambda_0\|\nabla\varsigma\|^2\\
&\qquad\le C(\zeta,\theta,\lambda,g)\Phi^2\|\varsigma\|^2\\
&\qquad\quad+2(E_1(\zeta,\theta,\delta),w)+2(E_2(\zeta,\theta,\delta),\varsigma)-2\varepsilon(\zeta,w)_2. \qquad (35)
\end{aligned}
$$

We now require two estimates for Φ,

$$\int_0^t \Phi^2(s)ds = \int_0^t \int_\Omega [1 + \sqrt{\delta g(v(s))}|\nabla u(s)|]^2 \, dx \, ds$$
$$\leq 2\int_0^t \int_\Omega \, dx \, ds + 2\int_0^t \int_\Omega \delta g(v(s))|\nabla u(s)|^2 \, dx \, ds$$
$$\leq 2t|\Omega| + \delta\|u_0\|^2 - \|u(t)\|^2, \quad (36)$$

by virtue of (28), and

$$\int_0^t \Phi^2(s)ds \geq \int_0^t \int_\Omega \, dx \, ds = t|\Omega|. \quad (37)$$

With the help of Lemma 1, we derive from (35)– (37) that

$$\|w(t)\|^2 + \|\varsigma(t)\|^2 + 2\varepsilon \int_0^t \|w(s)\|_2^2 \, ds + \lambda_0 \int_0^t \|\nabla\varsigma(s)\|^2 \, ds$$
$$\leq \exp\left(C(\zeta,\theta,\lambda,g)\int_0^t \Phi^2(s)ds\right)\Big\{\|w(0)\|^2 + \|\varsigma(0)\|^2 +$$
$$\int_0^t \exp\left(C(\zeta,\theta,\lambda,g)\int_s^0 \Phi^2(\xi)d\xi\right)[2(E_1(\zeta,\theta,\delta)(s), w(s))$$
$$+ 2(E_2(\zeta,\theta,\delta)(s), \varsigma(s)) - 2\varepsilon(\zeta(s), w(s))_2] \, ds\Big\}$$
$$\leq \exp\left(C(\zeta,\theta,\lambda,g)(2t|\Omega| + \delta\|u_0\|^2 - \|u(t)\|^2)\right)\Big\{\|w(0)\|^2 + \|\varsigma(0)\|^2 +$$
$$\int_0^t \exp\left(-C(\zeta,\theta,\lambda,g)s|\Omega|\right)\big|2(E_1(\zeta,\theta,\delta)(s), w(s))$$
$$+ 2(E_2(\zeta,\theta,\delta)(s), \varsigma(s)) - 2\varepsilon(\zeta(s), w(s))_2\big| \, ds\Big\}$$
$$\leq \exp\left(C(\zeta,\theta,\lambda,g)(|\Omega|+1)(2t + \delta\|u_0\|^2 - \|u(t)\|^2)\right)\Big\{\|w(0)\|^2 + \|\varsigma(0)\|^2 +$$
$$\int_0^t \exp\left(-C(\zeta,\theta,\lambda,g)s(|\Omega|+1)\right)\big|2(E_1(\zeta,\theta,\delta)(s), w(s))$$
$$+ 2(E_2(\zeta,\theta,\delta)(s), \varsigma(s)) - 2\varepsilon(\zeta(s), w(s))_2\big| \, ds\Big\}, \quad (38)$$

since $s \leq 2t$. Now (38) yields (26) with

$$\gamma = \exp\{C(\zeta,\theta,\lambda,g)(|\Omega|+1)\}.$$

□

Lemma 3. *Let (u, v) be a weak solution to problem (20) – (23). The following estimates are valid:*

$$\|u\|_{L_\infty(0,T;L_2)} + \|v\|_{L_\infty(0,T;L_2)} + \|v\|_{L_2(0,T;H_0^1)} \le C, \tag{39}$$

$$\|u\|_{L_2(0,T;V_2)} \le \frac{C}{\sqrt{\varepsilon}}, \tag{40}$$

$$\|\nabla u\|_{L_2(0,T;L_1)} + \|\nabla u\|_{L_1(0,T;L_r)} + \|\nabla u\|_{L_{4/3}(0,T;L_{-\epsilon+4/3})} \le C, \tag{41}$$
$$1 < r < 2, 0 < \epsilon < \frac{1}{3},$$

$$\|u'\|_{L_2(0,T;V_2^*)} + \|v'\|_{L_2(0,T;H^{-2})} \le (1+\sqrt{\varepsilon})C, \tag{42}$$

$$\|v'\|_{L_2(0,T;H^{-1})} \le (1+1/\sqrt{\varepsilon})C. \tag{43}$$

The constants $C = C(T, \|u_0\|, \|v_0\|, \lambda, g, \Omega)$ are independent of ε and δ.

Proof. The estimates (39) and (40) are direct consequences of (26) with $\zeta \equiv \theta \equiv 0$.

Then, using (13) and (28), we have

$$\begin{aligned}\|\nabla u\|_{L_2(0,T;L_1)} &\le \|\sqrt{\delta g(v)}\nabla u\|_{L_2(0,T;L_2)}\|1/\sqrt{g(v)}\|_{L_\infty(0,T;L_2)}\\ &\le C\|1+|v|\|_{L_\infty(0,T;L_2)} \le C,\end{aligned}$$

and, since $H_0^1 \subset L_p$ for any $p < \infty$ by Sobolev embedding,

$$\begin{aligned}\|\nabla u\|_{L_1(0,T;L_r)} &\le \|\sqrt{\delta g(v)}\nabla u\|_{L_2(0,T;L_2)}\|1+|v|\|_{L_2(0,T;L_{2r/(2-r)})}\\ &\le C(1+\|v\|_{L_2(0,T;H_0^1)}) \le C.\end{aligned}$$

By the time-space Hölder inequality [**?**, Lemma 2.2.1(b)],

$$\begin{aligned}\|\nabla u\|_{L_{4/3}(0,T;L_{-\epsilon+4/3})} &\le \||\nabla u|^{1/2}\|_{L_4(0,T;L_2)}\||\nabla u|^{1/2}\|_{L_2(0,T;L_{\frac{8-6\epsilon}{2+3\epsilon}})}\\ &\le \sqrt{\|\nabla u\|_{L_2(0,T;L_1)}\|\nabla u\|_{L_1(0,T;L_{\frac{4-3\epsilon}{2+3\epsilon}})}} \le C.\end{aligned}$$

It remains to estimate the time derivatives, expressing them from (24) and (25). Utilizing (28), we get

$$\begin{aligned}\|\langle u', \varphi\rangle\|_{L_2(0,T)} &\le \delta\|(g(v)\nabla u, \nabla\varphi)\|_{L_2(0,T)} + \varepsilon\|(u,\varphi)_2\|_{L_2(0,T)}\\ &\le \|\sqrt{\delta g(v)}\|_{L_\infty(0,T;L_\infty)}\|\sqrt{\delta g(v)}\nabla u\|_{L_2(0,T;L_2)}\|\nabla\varphi\| + \sqrt{\varepsilon}\sqrt{\varepsilon}\|u\|_{L_2(0,T;V_2)}\|\varphi\|_2\\ &\le C(1+\sqrt{\varepsilon})\|\varphi\|_2,\end{aligned}$$

and

$$\begin{aligned}\|\langle v',\phi\rangle\|_{L_2(0,T)} \le \|(\lambda\nabla v,\nabla\phi)\|_{L_2(0,T)} + \delta\|(\nabla v,\phi\nabla\lambda)\|_{L_2(0,T)} \\ + \delta\|\left((1-\lambda)v,\phi\right)\|_{L_2(0,T)} + \delta\|\left((1-\lambda)|\nabla u|,\phi\right)\|_{L_2(0,T)} \\ \le \|v\|_{L_2(0,T;H_0^1)}\|\phi\|_1 + C(\lambda)\|v\|_{L_2(0,T;H_0^1)}\|\phi\| \\ + \|\nabla u\|_{L_2(0,T;L_1)}\|\phi\|_{L_\infty} \le C\|\phi\|_2.\end{aligned}$$

In order to get (43), it suffices to observe that

$$\delta\|\left((1-\lambda)|\nabla u|,\phi\right)\|_{L_2(0,T)} \le \|\nabla u\|_{L_2(0,T;L_2)}\|\phi\| \le C\|u\|_{L_2(0,T;V_2)}\|\phi\|_1 \le \frac{C}{\sqrt{\varepsilon}}\|\phi\|_1.$$

□

Lemma 4. *Given $T > 0$ and $u_0, v_0 \in L_2$, there exists a weak solution to problem (20) – (23) with $\delta = 1$.*

Proof. Let us rewrite the weak statement of (20) – (23) in the suitable operator form

$$\tilde{A}(u,v) = \delta Q(u,v). \tag{44}$$

The operators $\tilde{A}, Q : W_1 \times W_2 \to L_2(0,T;V_2^*) \times L_2(0,T;H^{-1}) \times L_2 \times L_2$ are determined by the formulas

$$\langle \tilde{A}(u,v),(\varphi,\phi)\rangle = \Big(\frac{d}{dt}(u,\varphi) + \varepsilon(u,\varphi)_2, \frac{d}{dt}(v,\phi) + (\lambda\nabla v,\nabla\phi), u|_{t=0}, v|_{t=0}\Big),$$

$$\langle Q(u,v),(\varphi,\phi)\rangle = \Big(-(g(v)\nabla u,\nabla\varphi), -(\nabla v,\phi\nabla\lambda) + \left((1-\lambda)(|\nabla u| - v),\phi\right), u_0, v_0\Big).$$

Here $\varphi \in V_2$ and $\phi \in H_0^1$ are test functions.

The operator Q is continuous and compact. Here we only explain this claim for its first component, and for the others the proof is more straightforward. We observe first that the embedding $W_1 \subset L_p(0,T;W_p^1)$ is compact for some $p > 2$. This can be shown using [?, Corollary 8]. The embedding $W_2 \subset L_2(0,T;L_2)$ is compact by [?, Corollary 4]. Let $(u_m,v_m) \rightharpoonup (u_0,v_0)$ be a weakly converging sequence in $W_1 \times W_2$. Then (u_m,v_m) is strongly converging in $L_p(0,T;W_p^1) \times L_2(0,T;L_2)$. By Krasnoselskii's theorem [?, Theorem 2.1], $g(v_m) \to g(v_0)$ in $L_q(0,T;L_q)$ for any $q < +\infty$. Thus, $g(v_m)\nabla u_m \to g(v_0)\nabla u_0$ in $L_2(0,T;L_2)$, and the claim follows.

The linear operator $\tilde{A}$ is continuous by [?, Corollary 2.2.3] and invertible by [?, Lemma 3.1.3]. Thus, (44) can be rewritten as

$$(u,v) = \delta\tilde{A}^{-1}Q(u,v)$$

in the space $W_1 \times W_2$.

Lemma 3 yields the a priori estimate

$$\|u\|_{W_1} + \|v\|_{W_2} \le C,$$

where C may depend on ε but does not depend on δ. By Schaeffer's theorem [**?**, p. 539], there exists a fixed point of the map $\tilde{A}^{-1}Q$, which is the required solution. $\square$

We will also need the following simple fact.

Proposition 1. *Let G be a measurable set in a finite-dimensional space, $\chi : \mathbb{R} \to \mathbb{R}$ be a continuous function, and let $y_m : G \to \mathbb{R}$ be a sequence of functions. Assume that $\{y_m\}$ is uniformly bounded in $L_\infty(G)$, and $y_m \to y_0$ in $L_q(G)$, $q \geq 1$. Then $\chi(y_m) \to \chi(y_0)$ in $L_p(G)$ for any $p < \infty$.*

Proof. Due to the uniform boundedness of $\{y_m\}$, without loss of generality we may assume that χ is also bounded, and then it suffices to apply [**?**, Theorem 2.1]. $\square$

Based on the obtained lemmas, we can proceed with the sketch of the **proof of Theorem 1**. We refer to [**?**] for the details of the technique, and mainly focus on the new issues. To prove a) and b), one passes to the limit in (26) with $\delta = 1$ as $\varepsilon = \varepsilon_m \to 0$ on every interval $(0, T)$, $T > 0$. However, unlike in [**?**, **?**, **?**, **?**], in view of the presence of the absolute value in the right-hand member of (26), it is not possible to do it via weak and weak-* compactness.

Let (u_m, v_m) be the weak solution to problem (20) – (23) with $\varepsilon = \varepsilon_m$. Lemma 3, [**?**, Corollary 4] and the compact Sobolev embedding $W^1_{-\epsilon+4/3} \subset L_2$ imply that without loss of generality $u_m \to u$ in $L_{4/3}(0, T; L_2)$, $v_m \to v$ in $L_2(0, T; L_2)$. Then, by (39) and Proposition 1,

$$\gamma^{\|u_m(t)\|^2} \to \gamma^{\|u(t)\|^2}$$

in $L_2(0, T)$. Furthermore, by the same proposition, $\|u_m(t) - \zeta(t)\|^2 \to \|u(t) - \zeta(t)\|^2$, $\|v_m(t) - \theta(t)\|^2 \to \|v(t) - \theta(t)\|^2$ in $L_2(0, T)$. Therefore

$$\gamma^{\|u_m(t)\|^2} \left\{ \|u_m(t) - \zeta(t)\|^2 + \|v_m(t) - \theta(t)\|^2 \right\}$$
$$\to \gamma^{\|u(t)\|^2} \left\{ \|u(t) - \zeta(t)\|^2 + \|v(t) - \theta(t)\|^2 \right\}$$

in $L_1(0, T)$. Note that

$$\theta \in L_4(0, T; H^1) \subset L_\infty(0, T; L_2) \cap L_2(0, T; H^2).$$

This yields $E_1(\zeta, \theta) \in L_4(0, T; L_2)$. Remember that $E_2(\zeta, \theta) \in L_2(0, T; L_2)$. Thus, we can pass to the limit in the right-hand side of (26) as well; the last summand (the one with ε) goes to zero due to (40).

To get c), one lets $\zeta = u_T$, $\theta = v_T$ in (19) for $t \in (0, T)$, and then the right-hand member of (19) vanishes there. And e) is obtained by putting $t = 0$ in (19) and applying a density argument. Finally, d) is a consequence of a), e) and c).

Figure 3: Comparison of denoising results on noisy *Lena* image. (a) Perona and Malik [?] (PM) (b) Catté et al [?] (CLMC) (c) Coupled PDE Eqn. (4-5) with $\lambda = 0.5$ (CPDE), and (d) with adaptive λ using Eqn. (6) (ADAP). Top row shows the denoised image and the bottom row shows method noise, i.e., $(|u_0 - u|)$

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

Figure 4: Comparison results with classical diffusion schemes for a circle taken from the *Shapes* test image. (a) Perona and Malik [?] (PM) (b) Catté et al [?] (CLMC) (c) Coupled PDE Eqn. (4-5) with $\lambda = 0.5$ (CPDE), and (d) with adaptive λ using Eqn. (6) (ADAP). Top row shows the surface visualization and the bottom row shows corresponding level lines as contours.

4 Numerical experiments

4.1 Implementation

Implementing the proposed coupled PDE Eqns. (4-5) can be done in a variety of ways [?]. Here we follow a standard finite difference approach and utilize an explicit Euler scheme for both PDEs as a proof of concept. Dirichlet boundary conditions are used and the initial image $u = u_0$ and initial edge map $v = 1$ are fixed. An alternating scheme is used, that is, at each iteration we solve for the image variable u and then for the edge variable v. In this case, the first PDE Eqn. (4) is an inhomogeneous linear PDE in the image variable u which can be solved very efficiently, and the second PDE Eqn. 5 is a time dependent inhomogeneous Poisson problem in the edge variable v and we can adapt fast Poisson solvers for it. Note that the adaptive parameter λ in Eqn. (6) requires storage of the entire scale space of $v(\tau, x)_{t=0}^{t-1}$ at every iteration $t > 1$. To speed up the computational efforts we can utilize down-scaling techniques or other advanced numerical techniques such as operator splitting formulae for solving coupled PDE systems.

4.2 Comparison results and discussion

The proposed system of coupled PDE (we denote CPDE the non-adaptive $\lambda = 0.5$ and ADAP the adaptive case Eqn. (6) respectively) are compared numerically first with the following two classical single

(a)

Figure 5: One dimensional signal taken from the *Shapes* image illustrating the edge preserving and noise removal properties of the proposed coupled PDE scheme. Original signal is given by $(-.-)$ dash-dotted line, noisy by $(\cdots)$ dotted, and the restored signal is in solid line.

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

Figure 6: Comparison of denoising results on noisy ($\sigma_n = 20$) $Montage$ image. (a) Nitzberg and Shoita [?] (NS) (b) Chen and Levine [?] (CL) (c) Belahmidi and Chambolle [?] (BC) (d) Amann [?] (AM) (e) Coupled PDE Eqns. (4-5) with constant $\lambda = 0.5$ (CPDE), and (d) with adaptive λ using Eqn. (6) (ADAP). From top to bottom: the denoised image u, edge indicator based function $g(v)$, method noise ($|u_0 - u|$), surface visualization of the piecewise smooth part, and corresponding level lines shown as contours.

PDE schemes:

(a) Perona and Malik [?]:

$$\frac{\partial u}{\partial t} = div\left(\frac{\nabla u}{1 + |\nabla u|^2 / K^2}\right)$$

(b) Catté et al [?]:

$$\frac{\partial u}{\partial t} = div\left(\frac{\nabla u}{1 + |\nabla G_\sigma \star u|^2 / K^2}\right)$$

Note that, to make a fair comparison we utilize the same diffusion function g_{pm1} Eqn. (2) in all schemes. The contrast parameter $K > 0$ can be chosen in a variety of ways, see for example [?]. For simplicity we utilize the original suggestion given by Perona and Malik [?] uniformly for all the schemes. Further, the proposed coupled PDEs are compared numerically with the following coupled PDE schemes from recent literature:

(a) Nitzberg and Shoita [?]:

$$\begin{aligned}\frac{\partial u}{\partial t} &= div\left(g(v)\nabla u\right)\\ \frac{\partial v}{\partial t} &= wG_\sigma \star |\nabla u|^2 - wv\end{aligned}$$

where $w > 0$ relaxation parameter.

(b) Chen and Levine [?]:

$$\begin{aligned}\frac{\partial u}{\partial t} &= div\left(L(v)\nabla u\right) - \lambda(u - u_0)\\ \tau\frac{\partial v}{\partial t} &= (\nabla G_\sigma \star u - v)\end{aligned}$$

where L is the matrix valued diffusion tensor.

(b) Belahmidi and Chambolle [?]:

$$\begin{aligned}\frac{\partial u}{\partial t} &= div\left(g(v)\nabla u\right) - \lambda(u - u_0)\\ \frac{\partial v}{\partial t} &= F(|\nabla u|^2) - v\end{aligned}$$

where F is a smoothed version of truncation $s \to \min(s, M)$, $M > 0$ large.

(b) Amann [?]:

$$\frac{\partial u}{\partial t} = div\left(\frac{\nabla u}{1 + (\theta \star |\nabla u|^2)/K^2}\right)$$

where $\theta\star|\nabla u|^2(t) = \int_{t-\delta}^{t} |\nabla u(\tau)|^2\, d\tau$ represents the time-delayed convolution. Note that technically this is not a coupled system although it can be written as a relaxation similar to our model Eqns. (4-5).

The parameters[2] in all these schemes were tuned to obtain the best possible PSNR values (see Eqn. (45) below).

Figure 3 shows a comparison results for the noisy ($\sigma_n = 20$) *Lena* gray scale image with the classical diffusion PDEs. As can be seen from Figure 3, the coupled PDE model performs well in general and avoids the staircasing artifacts associated with the classical PDEs of Perona and Malik [?] and Rudin et al [?]. Moreover compared to Catté et al [?] the proposed method preserves fine scale structures better. To highlight the smoothing property of the proposed scheme, in Figure 4 we show the surface and level lines of a circle taken from the synthetic *Shapes* image for different schemes.

Figure 5 shows a line of 80 pixel width taken across the noisy *Shapes* image (at pixel position $x = 250$ and $y = 140$ to 220, corresponds to the circle and the spiral at the right end of the image) and the corresponding restored version of it using our scheme with adaptive choice for the parameter. As can be seen, the jumps seen at pixel ranges 50-60 and 70 are well-preserved, whereas the noisy perturbations at pixel range 10-40 are smoothed out. By comparing with the original signal one can see clearly the strong smoothing effects of the proposed coupled PDE scheme in flat regions. The sharp corners are slightly blurred due to the Laplacian involved in Eqn. (5).

Figure 6 shows a comparison of systems of coupled PDEs for the noisy ($\sigma_n = 20$) *Montage* gray scale image. As can be seen by comparing the piecewise constant circle and the ramp slope part the proposed system of coupled PDEs preserve them while removing noise effectively. To compare the schemes quantitatively two commonly used error metrics from the image processing literature are utilized:

1. PSNR is given in decibels (dB). A difference of $0.5\,dB$ can be identified visually. Higher PSNR value indicates optimum denoising capability.

$$\text{PSNR}(u) := 20 * \log 10\left(\frac{u_{max}}{\sqrt{MSE}}\right) dB \tag{45}$$

[2] Unfortunately there is no universal guideline for choosing parameters in diffusion based schemes and maximum PSNR based selection is done by sweeping the parameter set thoroughly. The important parameter σ in smoothing kernel G_σ is set $\sigma = 2$ for all the schemes and experiments reported here. This parameter needs to be increased if the noise level σ_n is higher.

(a)(b)(c)

(d)(e)

Figure 7: Application of denoising bio-medical images using the proposed scheme. (a) Input image (b) Output image u (c) Edge variable image v. Surface visualization of $BrainMRI$ image: (d) Input image (e) Output image.

where MSE $= (mn)^{-1}\sum\sum(u-u_0)$, $m \times n$ denotes the image size, u_{max} denotes the maximum value, for example in 8-bit images $u_{max} = 255$.

2. MSSIM index is in the range $[0, 1]$. The MSSIM value near one implies the optimal denoising capability of a scheme and is mean value of the SSIM metric. The SSIM is calculated between two windows ω_1 and ω_2 of common size $N \times N$

$$\mathrm{SSIM}(\omega_1, \omega_2) = \frac{(2\mu_{\omega_1}\mu_{\omega_2} + c_1)(2\sigma_{\omega_1\omega_2} + c_2)}{(\mu^2_{\omega_1} + \mu^2_{\omega_2} + c_1)(\sigma^2_{\omega_1} + \sigma^2_{\omega_2} + c_2)}$$

where μ_{ω_i} the average of ω_i, $\sigma^2_{\omega_i}$ the variance of ω_i, $\sigma_{\omega_1\omega_2}$ the covariance, c_1, c_2 stabilization parameters, see [?] for more details[3].

Table 1 shows the comparison results using these three metrics for different test images. As can be seen, the proposed scheme performs well for a variety of images (*Barbara*[4], *Cameraman* [5], *Montage*, and standard test images taken from USC-SIPI miscellaneous database[6]). Even with the global parameter $\lambda = 0.5$, the coupled PDE outperforms the standard diffusion PDEs of Perona and Malik [?] and Catte et al [?]. Further test results and images used here are available online[7]. Moreover, for textured images (*Mandrill*, *Barbara* etc) the non adaptive coupled PDE system seems to perform better than the adaptive case. We stress however that this work, the system of coupled PDE, does not aim to give state-of-the-art results for image denoising, and instead concentrates on demonstrating how a coupled PDE combined with an adaptive parameter choice can be harnessed directly for noise removal and edge detection. For instance, denoising will give similar or even better results as with total variation regularization through the classical ROF model [?] if one is able to identify an appropriate regularization parameters involved in the model [?]. Our examples are again a proof-of-concept that uses the coupled system and we do not claim it outperforms state of the art TV regularization based schemes.

As an application of the proposed system we consider denoising medical images. Figure 7 shows input *Ultrasound* (481×403), *Bacteria* (391×380), *BrainMRI* (210×210) images and its corresponding

[3] Code available at http://ece.uwaterloo.ca/~z70wang/research/ssim/

[4] Image courtesy of J. Portilla and available online at http://decsai.ugr.es/~javier/denoise/barbara.png

[5] Image courtesy of MIT

[6] Available at http://sipi.usc.edu/database/

[7] http://sites.google.com/site/suryaiit/research/aniso

(a) (b) (c)

(d) (e)

(f) (g)

Figure 8: Top row: Noisy $Barbara$ image decomposition using the adaptive coupled PDE system (a) smoothed image u (b) edge variable v (c) noise residue $w = u_0 - (u + v)$ Bottom rows: Edges detected from noise-free $Aircraft$ 659×409 image using the adaptive coupled PDE system with reaction terms ($\epsilon_1 = \epsilon_2 = 0.0015$) (d) Canny detector [?] with $\sigma = 1$ (e) Canny detector with $\sigma = 2$ (f) Synchronization coupled PDE scheme [?] (g) Modified proposed system of coupled PDEs.

(u, v) functions. Figure 7(d,e) shows both input u_0 and the result u in surface format which highlights the selective smoothing property of the scheme.

We can further modify the scheme to obtain meaningful decomposition of a digital image. For example, Figure 8 (top row) shows the decomposition of the $Barbara$ image into three different components, i.e, $u0 = u + v + w$ where w component is computed simply by $w = u0 - (u + v)$. Note that such a three part decomposition model is originally devised to obtain smooth + edges + texture part. In our case, we obtain texture as part of the edge variable v itself and the w component includes mainly random noise present in the image. Thus, we naturally obtain image decomposition as part of the proposed system of coupled PDEs [?]. Moreover, following a similar idea in [?] we can obtain edge detection as part of the image decomposition using the common initial condition, namely the input image, for both the PDEs. A weak coupling is utilized with the addition of reaction terms of the form $\epsilon_1(u - v)$, $\epsilon_2(v - u)$ to the coupling PDEs Eqn. (4-5). Finally, the difference (residual) $u(x, T) - v(x, T)$ is advocated as synchronization of the two dynamical systems which can facilitate better edge detection, we refer to [?] for more details. Figure 8(bottom rows) illustrate this for $Aircraft$[8] image and compares it with the scheme in [?]. As can be seen we obtain similar results but with much smoother output as we use different diffusion terms in the system. Compared with Canny edge detector [?] with two different parameters[9] $\sigma = 1, 2$ the proposed scheme provides better edge map as well.

Note that, adding the usual fidelity $(u - u_0)$ (a reaction term) such as the Nordstörm's bias PDE version Eqn. (7) does not modify the proofs presented in Section 3. Currently, we are studying a model

[8]Image courtesy of UCF CVPR Group and available online at http://marathon.csee.usf.edu/edge/edge_detection.html

[9]Implemented using the MATLAB command `edge(`u_0`,'canny',` σ`)`.

which involves a L^1 fidelity as well as adaptive fidelity parameter for better texture preserving denoising,

$$\begin{aligned} \frac{\partial u}{\partial t} &= div\,(g(v)\nabla u) - \mu(x)\frac{u-u_0}{|u-u_0|} && (46) \\ \frac{\partial v}{\partial t} &= \lambda(x)\,div(\nabla v) + (1-\lambda(x))\,(|\nabla u| - v) && (47) \end{aligned}$$

Further, the edge variable PDE can be generalized as well

$$\frac{\partial v}{\partial t} = \lambda(x)\,div(\tilde{g}(u)\nabla v) + (1-\lambda(x))\,(F(|\nabla u|) - v) \tag{48}$$

where $\tilde{g}, F \in C^1([0,+\infty))$, $F(0) = 0$, $g(0) = 1$, $\lim_{s\to\infty} g(s) = 0$. Extension of the results presented in Section 3 for these generalized system of coupled PDEs is the subject of our ongoing work.

5 Conclusions

A novel coupled PDE based scheme is studied for image restoration. By utilizing a separate PDE for the edge variable our proposed model improves the denoising results significantly. A combination of edge preserving Perona–Malik and Catté et al's smoothing PDEs is considered for image restoration. Adaptive choice for choosing the balancing parameter involved in the edge variable PDE has been studied. Existence and uniqueness result for the coupled PDE model is proved using the theory of dissipative solutions due to P.-L. Lions. Further, numerical experiments conducted on a variety of noisy images indicate that the model gives artifact free restoration results than other related schemes from the past.

Image	PM [?]	CLMC [?]	NS [?]	CL [?]	BC [?]	AM [?]	CPDE	ADAP
*Girl*1	16.17/0.7965	16.27/0.8465	19.21/0.8904	19.21/0.8884	19.21/0.9120	19.37/0.8824	**21.31**/0.9500	21.10/**0.9562**
*Couple*1	16.18/0.7965	16.22/0.7865	19.27/0.8984	18.72/0.9102	19.20/0.9091	21.20/0.9280	19.10/0.9081	**22.45/0.9421**
*Girl*2	16.09/0.8210	17.43/0.8303	19.51/0.8885	19.54/0.8650	19.21/0.8205	19.54/0.9010	20.32/0.9150	**20.82/0.9231**
*Girl*3	15.97/0.8192	15.80/0.8548	18.50/0.8311	18.22/0.8872	18.97/0.8900	19.22/0.8985	20.86/**0.8995**	**21.01**/0.8945
*House*1	15.86/0.7966	15.72/0.7949	18.29/0.8219	19.00/0.8985	18.75/0.8282	19.19/0.9099	21.31/0.9085	**22.17/0.9455**
Tree	18.15/0.8287	18.45/0.8116	18.52/0.8110	17.95/0.8018	17.48/0.8256	17.93/0.8452	19.15/0.8401	**19.88/0.8483**
*Jelly*1	16.17/0.7696	16.91/0.7555	19.69/0.7657	19.53/0.7586	19.39/0.7683	19.21/0.7885	21.01/0.7908	**21.56/0.7998**
*Jelly*2	16.00/0.7968	15.75/0.8219	18.72/0.8562	18.28/0.8231	19.04/0.8143	19.28/0.8184	19.23/0.8765	**19.85/0.8804**
Splash	15.96/0.7966	15.46/0.7898	18.89/0.8248	19.17/0.8720	18.94/0.9164	18.73/0.9105	**19.80**/0.9215	19.57/**0.9316**
Tiffany	16.24/0.7889	16.50/0.8108	17.00/0.8115	18.25/0.8018	18.73/0.8229	18.24/0.8049	18.69/0.8522	**18.80/0.8904**
Mandrill	15.35/0.8231	15.87/0.8484	16.27/0.8349	16.82/0.8146	16.84/0.8390	17.53/0.8727	**17.84/0.8970**	17.56/0.8851
Lena	15.62/0.7960	16.03/0.8187	17.12/0.8450	18.29/0.8384	17.56/0.8900	18.87/0.9454	19.22/0.9667	**19.85/0.9874**
Barbara	15.65/0.7965	15.45/0.7982	17.48/0.8994	17.59/0.9210	17.23/0.8945	18.00/0.7868	**18.72/0.9498**	17.81/0.8996
Cameraman	15.71/0.8025	16.82/0.8091	17.90/0.8703	18.19/0.8451	18.94/0.7918	17.57/0.7887	**18.96/0.9118**	17.97/0.8862
Montage	15.32/0.7965	15.32/0.8465	17.45/0.8982	17.45/0.8982	17.40/0.8982	17.75/0.8983	18.84/0.9499	**18.86/0.9799**

Table 1: PSNR and MSSIM comparison of various schemes for standard test images from the USC-SIPI database. In each case noisy image (PSNR = 15.21 dB) is obtained by adding random Gaussian noise of strength $\sigma_n = 30$ to the original gray-scale image of size 256×256. Each row indicates PSNR/MSSIM values for different test images. The proposed coupled PDE with $\lambda = 0.5$ and with adaptive choice for choosing λ are given as CPDE and ADAP (last two columns) respectively. Best results are indicated in boldface.