

ADAPTIVE DIFFUSION CONSTRAINED TOTAL VARIATION SCHEME WITH APPLICATION TO CARTOON + TEXTURE + EDGE IMAGE DECOMPOSITION

JUAN C. MORENO, V. B. SURYA PRASATH, DMITRY VOROTNIKOV AND HUGO PROENÇA

ABSTRACT: We consider an image decomposition model involving a variational (minimization) problem and an evolutionary PDE. We utilize a linear inhomogeneous diffusion constrained weighted total variation (TV) scheme for image adaptive decomposition. An adaptive weight along with TV regularization splits a given image into three components representing the geometrical (cartoon), textural (small scale-microtextures), and edges (big scale-macrottextures). We study the wellposedness of the coupled variational-PDE scheme along with an efficient numerical scheme based on Chambolle's dual minimization. We provide extensive experimental results in cartoon-texture-edges decomposition as well compare with other related variational, anisotropic diffusion PDE based models.

KEYWORDS: Image decomposition, total variation, linear diffusion, adaptive weight, multi-scale.

AMS SUBJECT CLASSIFICATION (2010): 68U10, 35Q68, 65F22.

1. Introduction

Decomposing an image into meaningful components is an important and challenging inverse problem in image processing. Image denoising is a very well known example of image decomposition. In such a decomposition, the given image is assumed to be under the influence of noise, and the main purpose is to remove noise without destroying edges. This denoising task can be regarded as a decomposition of the image into noise-free signal and noise part. There exist various methodologies for image restoration, where variational minimization and partial differential equation (PDE) are two of the most popular ones [3].

Received December 10, 2013.

The third author was partially supported by CMUC funded by the European Regional Development Fund through the program COMPETE and by the Portuguese Government through FCT under the project PEst-C/MAT/UI0324/2013, and by the projects PTDC/MAT/098060/2008, UTA-CMU/MAT/0007/2009, PTDC/MAT-CAL/0749/2012.

Another important example of image decomposition is based on its smooth and texture components using the total variation (TV) regularization method. This was first studied by Rudin et al [41] for image restoration. The TV regularization can be written as an unconstrained minimization,

$$\min_u E_{TV}(u) = \int_{\Omega} |\nabla u| \, dx + \frac{1}{2\alpha} \int_{\Omega} |u - f|^2 \, dx. \quad (1)$$

The parameter $\alpha > 0$ balances the fidelity term with respect to the TV regularization. Let the given image be written as $f = u + v$, where we assume that the function u models well structured homogeneous regions (cartoon) and v defines oscillating patterns such as noise and texture. Meyer [28] established the scale connection property of TV for image decomposition, namely that the α parameter is related to the scale of objects in the image. In particular, Meyer proved that if f is a characteristic function and is sufficiently small with respect to a norm ($|f|_* \leq 1/2\alpha$), then the minimization of the TV regularization (1) gives $u = 0$ and $f = 0$, which is counter-intuitive since one expects $u = f$ and $v = 0$. Thus, Meyer proposed to use dual of the closure in the bounded variation (BV) space of the Schwartz class for non-trivial solutions, we refer to [28] for more details. This crucial fact has been exploited by Vese and Osher [46] to obtain numerical approximations to the Meyer decomposition model, see also [4, 6]. While the literature on cartoon and texture decomposition is extensive by now, three lines of inquiry are closely related to the work contained here.

- **Different function spaces for modeling the textures and discrete approximations.** Following Meyer's seminal work [28], various authors have considered different function spaces to model textures accurately [5, 25, 7, 33, 18, 21, 26, 23, 45].
- **L^1 fidelity based TV regularization models.** In another related direction the fidelity term can be made L^1 and is proven to provide contrast preserving restorations. We refer to Chan and Esedoglu [13] for the geometric motivation and fine properties of L^1 fidelity term, see also [15]. Applications of L^1 -TV models for cartoon plus texture decomposition are considered as well [31, 48, 49, 21, 16, 2].
- **Different Regularizers instead of TV which is known to provide blocky solutions.** The well-known staircasing property of the TV regularization has been analyzed by many in the past [32, 10, 11] and various modifications have also been studied. To avoid staircasing

and blocky restoration results from the classical TV model there have been a variety of methods studied previously. Weighted and adaptive [43, 42, 38, 37, 39], nonlocal-TV [24], and higher order [14, 22, 34]. Out of other related decomposition models we mention multi-scale parameter based models [44, 20, 45]. Extension to multichannel images in general, RGB images in particular, is also an important area of research [8]. In this paper, we propose a new image decomposition scheme which splits a given image into its geometrical and textural part along with an edge capturing diffusion constraint. Following the success of weighted and adaptive TV models, our scheme is based on a weighted TV regularization where the edge-variable-based weight is computed in a data-adaptive way. The scheme is implemented using the splitting method of [9] along with Chambolle’s dual minimization scheme for the weighted TV regularization [12]. As a by-product of the implementation via dual minimization we obtain an auxiliary variable which is akin to textural component of the given image. Thus, the scheme studied here provides a cartoon, texture, edge (CTE) decomposition for digital images, see Figure 1 for some examples. We consider the color image decomposition using the dual minimization based color TV model. Multi-scale decomposition following the recent work of Tang and He [45] is also given. Moreover, we provide theoretical analysis of the model with a priori estimates and prove its solvability. Extensive experimental results on synthetic, textured real images are given. Further, illustrative comparison with other state-of-the-art models are undertaken and the results indicate superior performance of our scheme with respect to cartoon, texture separation and edge preservation.

The rest of the paper is organized as follows. Section 2 introduces the adaptive TV regularization coupled with a diffusion PDE. Section 3 provides wellposedness results for our coupled model. In Section 4 we provide the experimental results conducted on real and synthetic images and comparison with related schemes. Finally, Section 5 concludes the paper.

2. Diffusion constrained regularization

2.1. Weighted total variation minimization. The total variation based regularization functional [41] given in Eqn. (1) is well-known in edge preserving image restoration, we rewrite it as follows,

$$\min_u E_{TV}^\mu(u) = \int_{\Omega} |\nabla u| \, dx + \mu \int_{\Omega} |u - f|^2 \, dx \quad (2)$$

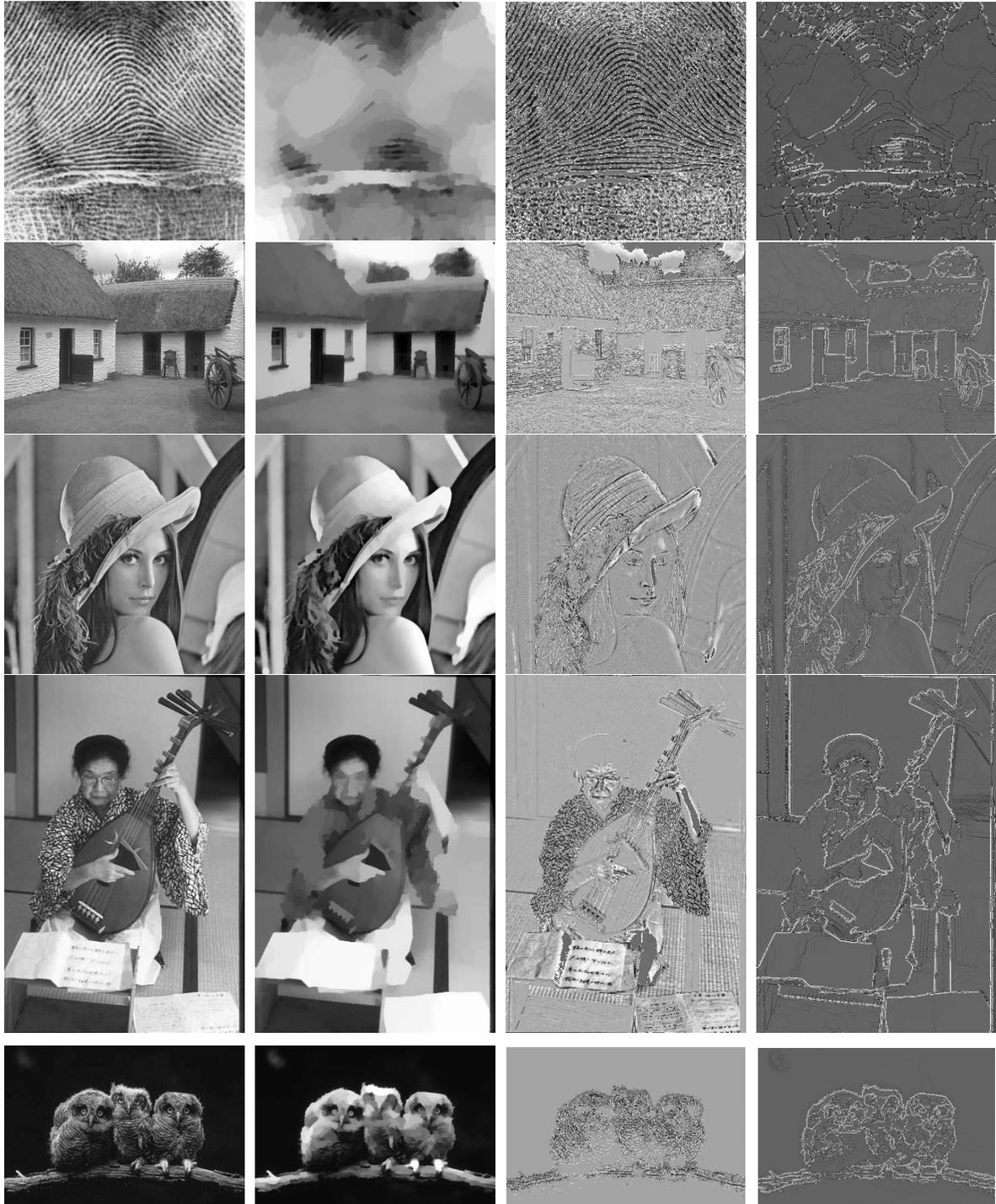
(a) Input (f)(b) Cartoon (u)(c) Texture (v)(d) (Pseudo) Edges (w)

FIGURE 1. The proposed coupled TV regularization with linear diffusion PDE model provides cartoon, texture and edge decomposition of images.

where now $\mu > 0$ is the image fidelity parameter which is important in obtaining results in denoising and cartoon+texture decomposition. A related approach is to consider a weighted total variation,

$$\min_u E_{gTV}^\mu(u) = \int_{\Omega} g(x, u, \nabla u) |\nabla u| dx + \mu \int_{\Omega} |u - f|^2 dx \quad (3)$$

where $g(x, u, \nabla u)$ represents the generalized weight function. For example, Bresson et al [9] have considered a convex regularization of the variational model

$$\min_u \left\{ \int_{\Omega} g(x) |\nabla u| dx + \mu \int_{\Omega} |u - f| dx \right\}$$

using a fast minimization based on a dual formulation to get a partition in the geometric and texture information. Note that the image fidelity is changed to L^1 norm, we refer to [13] for more details. The convex regularization version is considered in [9],

$$\min_{u,v} \left\{ \int_{\Omega} g(x) |\nabla u| dx + \frac{1}{2\theta} \int_{\Omega} (u + v - f)^2 dx + \mu \int_{\Omega} |v| dx \right\}, \quad (4)$$

where the parameter $\theta > 0$ is chosen to be small so that f almost satisfies $f \sim u + v$, with the function u representing geometric information, *i.e.* the piecewise-smooth regions, and function v captures texture information lying in the given image. The function g is an edge indicator function that vanishes at object boundaries, for example,

$$g(x) := \frac{1}{1 + \beta |\nabla f(x)|^2},$$

where f is the original image and β is an arbitrary positive constant. Thus, we see that TV based minimization models naturally lead to cartoon and texture decomposition of digital images. The image fidelity parameter μ can be made data adaptive to obtain texture preserving restorations, see [19].

2.2. Description of the model. In our work, we consider the following regularization model which was motivated by a successful coupled PDE modeling done in [40] for image restoration,

$$\min_u \left\{ \int_{\Omega} \phi(x, u, |\nabla u|) dx + \mu \int_{\Omega} |u - f| dx \right\}, \quad (5)$$

$$\frac{\partial w}{\partial t} = \lambda \operatorname{div}(\nabla w) + (1 - \lambda)(|\nabla u| - w). \quad (6)$$

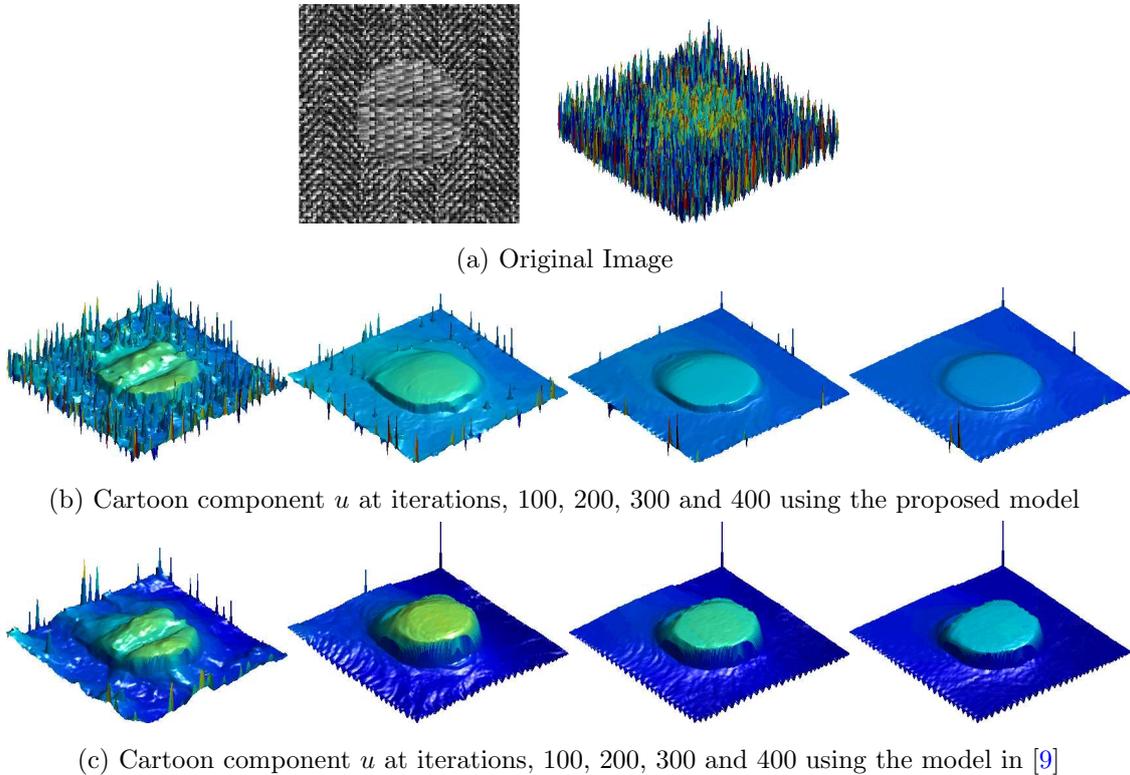


FIGURE 2. (Color online) Our adaptive diffusion constrained total variation scheme (see Eqn. (9)) with constant μ, λ (second row) provides better edge preserving image decomposition when compared to the traditional TV regularization model (Eqn. (4)) of [9] (third row) as number of iterations increase from 100 to 400. The proposed scheme keeps the structure without diffusing its boundary with the background.

The choice of regularizer ϕ depends on an application area and among a plethora of convex and non-convex functions available with the classical TV [41] and the non-local TV [24]. Motivated from the above discussions in Section 2.1, and success enjoyed by the weighted L^1 -TV regularization model in image denoising and segmentation, we use L^1 -TV regularizer model as a prime example to illustrate our model here. The proposed CTE model thus consists of a minimization along with a non-homogeneous diffusion equation,

$$\min_u \left\{ \int_{\Omega} g(w) |\nabla u| dx + \int_{\Omega} \mu(x) |u - f| dx \right\}, \quad (7)$$

$$\frac{\partial w}{\partial t} = \lambda(x) \operatorname{div}(\nabla w) + (1 - \lambda(x)) (|\nabla u| - w), \quad (8)$$

where $g(w) = \frac{1}{1+w^2}$, or $g(w) = \exp(-w^2)$ (Perona-Malik type diffusion functions [35]), or $g(w) = |w|^{-1}$, or $g(w) = \frac{1}{\sqrt{\epsilon^2+|w|^2}}$ (total variation diffusion function [41]). That is we solve adaptive data fidelity based weighted total variation minimization for the smooth part u using Eqn. (7) along with a linear non-homogenous diffusion constraint on w by solving Eqn. (8). Note that the balancing parameter λ and image fidelity μ taking values in $[0, 1]$ are important in our experimental results. Adaptive ways of choosing these parameters are explained below in Section 2.3. Following [9] we use a splitting with an auxiliary variable v to obtain

$$\min_{u,v} \left\{ \int_{\Omega} g(w)|\nabla u| dx + \frac{1}{2\theta} \int_{\Omega} (u+v-f)^2 dx + \int_{\Omega} \mu(x)|v| dx \right\}, \quad (9)$$

$$\frac{\partial w}{\partial t} = \lambda(x)\text{div}(\nabla w) + (1 - \lambda(x))(|\nabla u| - w).$$

Thus, the computed solution of these equations provides a representation $f \sim u+v+w$, where the function u represents the geometric information, the function v captures the texture information, and the function w represents the edges lying in the given image. Figure 2 shows a comparison of our scheme (μ and λ were taken as 1) and Bresson et al [9] scheme Eqn. (4) for a synthetic texture image which contains two different texture patterns. As can be seen, our scheme (Figure 2(b)) retains the cartoon edges better without diffusing the boundary and the shape is preserved in contrast to Bresson et al's result (Figure 2(c)).

The above coupled system is solved in an alternating iterative way for all the variables (u, v, w) involved and Chambolle's dual minimization scheme [12] is used for the weighted TV minimization step. We start with the initial conditions $(u, v, w)|_{n=0} = (f, \mathbf{0}, \mathbf{1})$ and use the following steps to compute CTE components:

- (1) Solving the linear diffusion PDE (8) for w with (u, v) fixed:

$$w^{n+1} = w^n + \frac{\delta t}{(\delta x)^2} (\lambda(x)\tilde{\Delta}w^n + (1 - \lambda(x))(|\nabla u| - w^n)), \quad (10)$$

where δx is spatial discretization step (natural pixel grid), $\tilde{\Delta}$ is the standard finite difference discretization for the Laplacian and δt is the step size.

(2) Solving for the cartoon component u with (v, w) fixed:

The minimization problem in u is given by (see Eqn. (9)),

$$\min_u \left\{ \int_{\Omega} g(w) |\nabla u| dx + \frac{1}{2\theta} \int_{\Omega} (u + v - f)^2 dx \right\}. \quad (11)$$

The solution of (11) is given by

$$u = f - v - \theta \operatorname{div} \mathbf{p},$$

where $\mathbf{p} = (p_1, p_2)$ satisfies $g(w) \nabla(\theta \operatorname{div} \mathbf{p} - (f - v)) - |\nabla(\theta \operatorname{div} \mathbf{p} - (f - v))| \mathbf{p} = 0$, which is solved using a fixed point method: $\mathbf{p}^0 = 0$ and

$$\mathbf{p}^{n+1} = \frac{\mathbf{p}^n + \delta t \nabla(\theta \operatorname{div}(\mathbf{p}^n) - (f - v)/\theta)}{1 + \frac{\delta t}{g(w)} |\nabla(\theta \operatorname{div}(\mathbf{p}^n) - (f - v)/\theta)|}.$$

(3) Solving for the texture component v with (u, w) fixed:

$$\min_v \left\{ \frac{1}{2\theta} \int_{\Omega} (u + v - f)^2 dx + \int_{\Omega} \mu(x) |v| dx \right\}, \quad (12)$$

and the solution is found as

$$v = \begin{cases} f - u - \theta \mu(x) & \text{if } f - u \geq \theta \mu(x), \\ f - u + \theta \mu(x) & \text{if } f - u \leq -\theta \mu(x), \\ 0 & \text{if } |f - u| \leq \theta \mu(x). \end{cases}$$

Next we describe a data adaptive way for choosing the fidelity parameter μ using the cartoon component at a previous iteration u^n .

Remark 1. *We interchangeably use edges and pseudo-edges as the w component provides an edge like features from a given image. The definition of edges in a digital image depends on the context and many traditional definitions depend on the magnitude of gradients (i.e., $|\nabla I|$), hence a solution of the PDE (8) provides a pseudo-edge map, see Figure 1(d).*

2.3. Adaptive fidelity parameter. Here, we consider the data adaptive parameters selection strategies which can provide a balanced approach in obtaining better CTE decomposition results. For the image fidelity parameter μ in Eqn. (7) we utilize a local histogram measure which provides a better texture separation [30]. For a given gray-scale image $I : \bar{\Omega} \rightarrow [0, L]$, let $\mathcal{N}_{x,r}$ be the local region centered at x with radius r . We compute the local

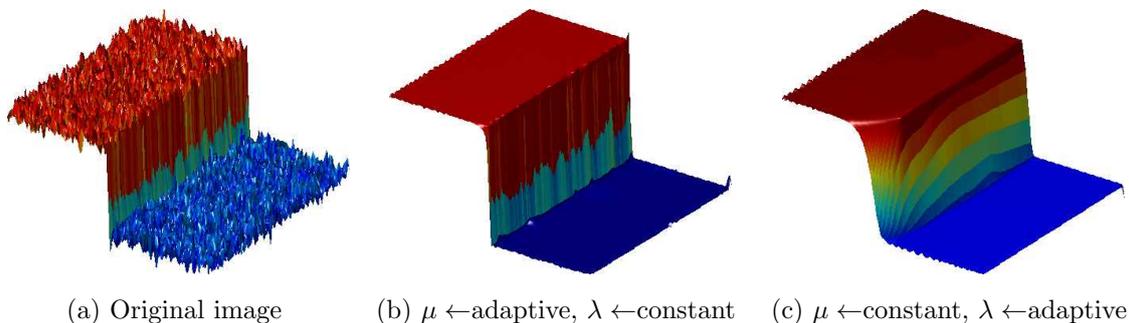


FIGURE 3. (Color Online) The proposed model with adaptive μ_1 (see Eqn. (14)) and constant $\lambda (= 1)$ provides better edge preservation in the cartoon component and captures small scale oscillations in the texture component against constant $\mu (= 1)$ and adaptive λ (using the definition of μ_1 from Eqn. (14)).

histogram of the pixel $x \in \Omega$ and its corresponding cumulative distribution function

$$P_x(y) = \frac{|\{z \in \mathcal{N}_{x,r} \cap \Omega \mid I(z) = y\}|}{\mathcal{N}_{x,r} \cap \Omega} \quad F_x(y) = \frac{|\{z \in \mathcal{N}_{x,r} \cap \Omega \mid I(z) \leq y\}|}{\mathcal{N}_{x,r} \cap \Omega} \quad (13)$$

for $0 \leq y \leq L$, respectively. This allows us to define the following measurable function $\mu : \Omega \rightarrow \mathbb{R}$, such that for each $x \in \Omega$,

$$\mu(x) = \mu_1(x) = \frac{\int_0^L F_x(y) dy}{\max_{x \in \Omega} \int_0^L F_x(y) dy}, \quad (14)$$

allowing us to get a weight of how much nonhomogeneous intensity is present in a local region $\mathcal{N}_{x,r}$ of a given pixel x . This new feature of the image does not depend on the pixel properties instead provides regional properties, see [30] for more details. Thus, we see that the μ is chosen according to local histogram information and is computed in an image adaptive way using the cartoon u^n in the iterative scheme (12). We compare our approach with two related adaptive functions:

- (1) The adaptive formulation of [40] which uses a summation of cartoon components up-to iteration n .

$$\mu_2(x) = \sum_{i=0}^n G_{\rho_i} \star u^n(x) \quad (15)$$

with $\rho_i = 1/i^2$ and at $n = 0$ the $\lambda_2(x) = 0.05$.

- (2) Relative reduction rate based parameter proposed in local TV based scheme [24].

$$\mu_3(x) = \frac{G_\rho \star |\nabla f(x)| - G_\rho \star |L_\sigma \star \nabla f(x)|}{G_\rho \star |\nabla f(x)|}, \quad (16)$$

where L_σ is a low pass filter. Note that this adaptive parameter uses only the initial input image f whereas the previous choices use u computed at a previous (Eqn. (14)) or every (Eqn. (15)) iteration.

Remark 2. *Figure 3 explains the choice of adaptiveness in our coupled model (7-8) for a synthetic image with different texture patterns. The first case with μ adaptive, λ constant provides persisting cartoon component whereas the second case with μ constant, λ adaptive (same local histogram based measure, Eqn. (14) is used for defining $\lambda(x)$) blurs the boundaries in the final result. Thus, in what follows, we use only μ adaptive parameter to illustrate our decomposition results.*

Figure 4 shows a comparison of different adaptive μ functions for a synthetic texture image. We see that the local histogram based μ_1 captures the texture components from all the quadrants. Moreover, Figure 4(e) shows that the energy value decreases similarly for different μ functions as the iteration increases. Figure 5 shows the usage of different μ function when we apply our model (7-8) for the same synthetic image to obtain cartoon (u) + texture (v) + pseudo-edges (w) decomposition. Note that the texture image v is obtained by linearly transforming its range to $[0, 255]$ for visualization purposes. Differences outside this range are saturated to 0 and 255 respectively. A similar transformation is applied to the edges (w) component as well. Next, we study the wellposedness of the model (7-8) using weak solutions concept and prove some a priori estimates and solvability of the proposed adaptive coupled model.

3. Wellposedness

3.1. Preliminaries. 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 [1]. Obviously, rectangular domains have this property. Fix also a time interval $[0, T]$, $T > 0$.

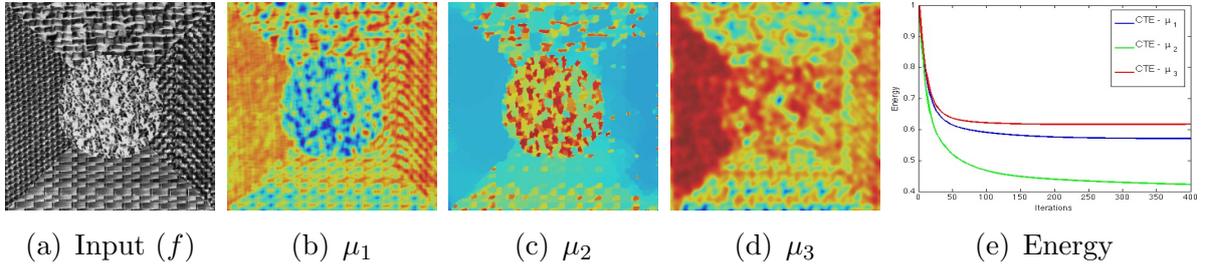


FIGURE 4. (Color Online) Comparison of different μ functions computed using the given input image. (a) Original image. (b) μ_1 based on local histograms Eqn. (14). (c) μ_2 based on the work of [40] Eqn. (15). (d) μ_3 base on the work of [24] Eqn. (16). (e) Energy versus iteration for different adaptive μ functions based energy minimization scheme (7-8).

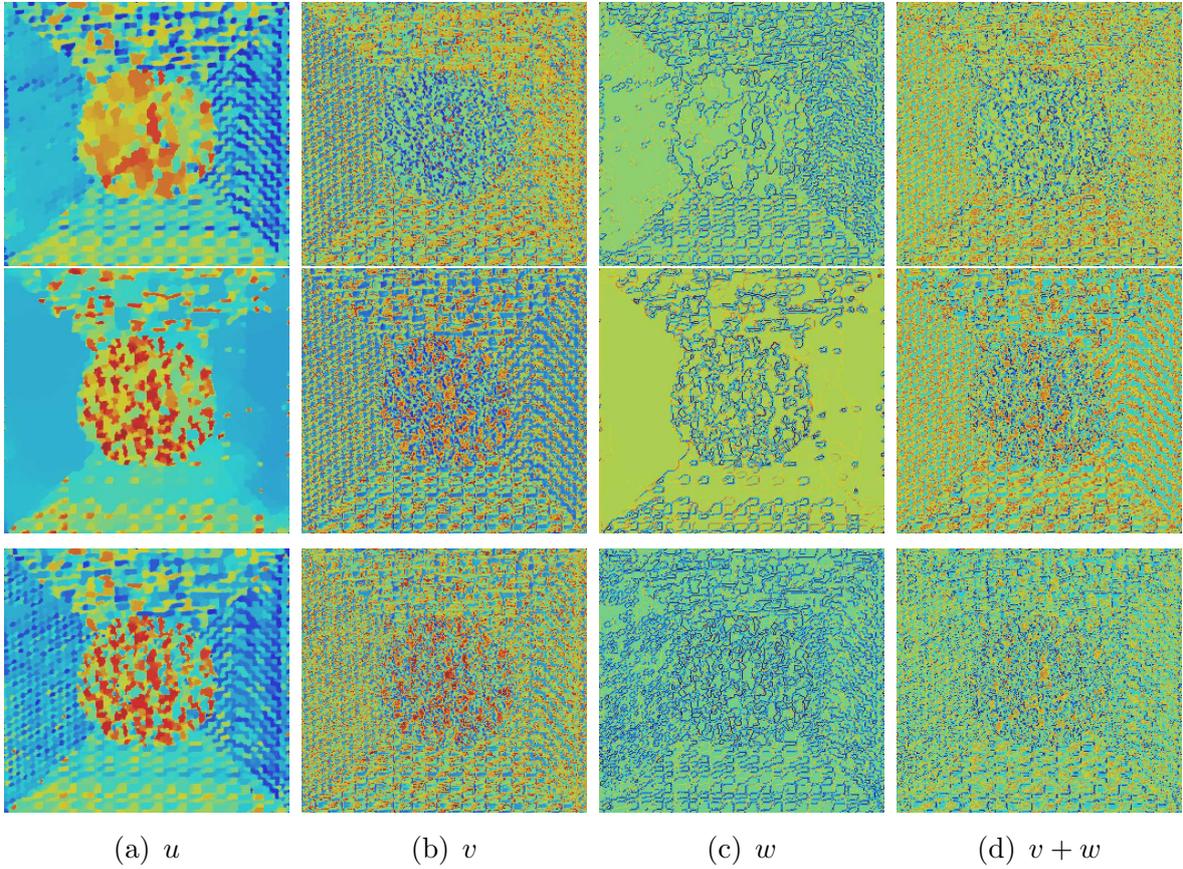


FIGURE 5. (Color Online) Different μ functions based CTE scheme (with λ constant) results. Top: μ_1 (local histogram) based result. Middle: μ_2 result. Bottom: μ_3 result. (a) Cartoon u . (b) Texture v . (c) Pseudo-edges w . (d) $v + w$. Best viewed electronically, zoomed in.

We study wellposedness of the weighted TV

$$u(t, x) = \mathbf{u}(x) : \min_{\mathbf{u}: \Omega \rightarrow \mathbb{R}} \left\{ \int_{\Omega} g(w(t, x)) |\nabla \mathbf{u}(x)| dx + \int_{\Omega} \mu(x) |\mathbf{u}(x) - f(x)| dx \right\} \quad (17)$$

with the diffusion constraint

$$\frac{\partial w(t, x)}{\partial t} = \Delta_{p, \lambda} w(t, x) + (1 - \lambda(x)) (|\nabla u(t, x)| - w(t, x)), \quad (18)$$

$$w(t, x) = 0, \quad x \in \partial\Omega, \quad (19)$$

$$w(0, x) = F(x), \quad (20)$$

where $p \geq 2$, and $f : \Omega \rightarrow \mathbb{R}$, $F : \Omega \rightarrow [0, +\infty)$, $\lambda : \Omega \rightarrow (0, 1]$, $\mu : \Omega \rightarrow (0, +\infty)$, $g : [0, +\infty) \rightarrow (0, +\infty)$ are prescribed functions. The operator $\Delta_{p, \lambda}$ is a weighted p -Laplacian:

$$\Delta_{p, \lambda} v = \lambda \operatorname{div}(|\nabla v|^{p-2} \nabla v) - (1 - |\nabla v|^{p-2}) \nabla v \cdot \nabla \lambda. \quad (21)$$

In particular, for $p = 2$ we recover the linear diffusion case. In this section, for the sake of generality, we admit adaptive λ and generic g .

Note that w is non-negative by the maximum principle.

We consider the Dirichlet boundary condition for w , but other boundary conditions can also be handled.

We use the brief notations L_q ($q \geq 1$), W_q^m ($m \in \mathbb{R}$), $W_0^{m, q}$ ($m > 0$) for the Lebesgue and Sobolev spaces on Ω with values in \mathbb{R} or \mathbb{R}^2 . Parentheses denote the bilinear form

$$(u, v) = \int_{\Omega} u(x) \cdot v(x) dx.$$

The norm in L_2 is $\|u\| = \sqrt{(u, u)}$.

The symbols $S(\mathcal{J}; E)$, $C(\mathcal{J}; E)$, $L_1(\mathcal{J}; E)$, etc., denote the spaces of Bochner measurable, continuous, Bochner integrable, etc., functions on an interval $\mathcal{J} \subset \mathbb{R}$ with values in a Banach space E .

Let \mathcal{M} be the Banach space of finite Radon measures on Ω . It is the dual of the space $C_0(\Omega)$ (the space of continuous functions on Ω that vanish at $\partial\Omega$, see e.g. [17]).

Let BV be the space of functions of bounded variation on Ω . For $v \in BV$, and $\phi \in C(\overline{\Omega})$, $\phi \geq 0$, the *weighted total variation* of v is

$$TV_{\phi}(v) = \sup_{\psi \in C_0^{\infty}(\Omega): |\psi| \leq \phi} (v, \operatorname{div} \psi). \quad (22)$$

In particular, the *total variation* of u is

$$TV(v) = TV_1(v). \quad (23)$$

Due to lower semicontinuity of suprema, for every non-negative $\phi \in C(\overline{\Omega})$ and a weakly- $*$ converging sequence $\{v_m\} \subset BV$, we have

$$TV_\phi(v) \leq \liminf_{m \rightarrow +\infty} TV_\phi(v_m). \quad (24)$$

A more refined argument of the same nature proves

Lemma 1. *For any $\varphi \in S(0, T; C(\overline{\Omega}))$, $\varphi \geq 0$ for a.a. $t \in (0, T)$, and a weakly- $*$ converging sequence $\{v_m\} \subset L_q(0, T; BV)$, $q > 1$, one has*

$$TV_{\varphi(t)}(v(t)) \leq \liminf_{m \rightarrow +\infty} TV_{\varphi(t)}(v_m(t)) \quad (25)$$

for a.a. $t \in (0, T)$.

For $v \in BV$, $|\nabla v|$ will denote the corresponding total variation measure. The operator

$$|\nabla(\cdot)| : BV \rightarrow \mathcal{M} \quad (26)$$

is bounded. We recall the duality relation

$$TV_\phi(v) = \langle |\nabla v|, \phi \rangle_{\mathcal{M} \times C_0(\Omega)}. \quad (27)$$

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 will use the embeddings

$$BV \subset L_q, \quad q \leq 2, \quad (28)$$

$$W_2^1 \subset L_q, \quad q < +\infty, \quad (29)$$

$$W_p^1 \subset C(\overline{\Omega}), \quad p > 2, \quad (30)$$

and

$$\mathcal{M} \subset W_q^{-1}, \quad q < 2, \quad (31)$$

and the Poincaré inequality

$$\|v\|_{W_p^1} \leq C \|\nabla v\|_{L_p}, \quad p \geq 1, \quad v|_{\partial\Omega} = 0. \quad (32)$$

Embeddings (28)–(31) are compact (except for (28) with $q = 2$).

We assume that λ , $\nabla\lambda$ and g are Lipschitz functions,

$$\lambda_0 = \inf_{x \in \Omega} \lambda(x) > 0, \quad g_0 = \sup_{y \geq 0} g(y) < +\infty,$$

and there exists a constant C_g so that

$$\left| \frac{d(\log g(y))}{dy} \right| \leq \frac{C_g}{1+y} \text{ for a.a. } y \geq 0. \quad (33)$$

The last condition means that g can have at most polynomial decay at infinity.

We assume that $\mu \in L_\infty(\Omega)$, and

$$0 < \mu_1 = \operatorname{ess\,inf}_{x \in \Omega} \mu(x) \leq \mu_2 = \operatorname{ess\,sup}_{x \in \Omega} \mu(x) < +\infty.$$

Finally, we assume that

$$F \in L_2, \quad (34)$$

and **at least one** of the following three conditions holds:

$$p > 2, f \in BV, \quad (35)$$

$$p = 2, \exists q > 1 : f \in W_q^1, \quad (36)$$

$$p = 2, f \in BV \cap L_\infty, \exists c_g > 0 : \frac{1}{g(y)} \leq c_g(1+y), y \geq 0. \quad (37)$$

3.2. A priori estimates. Before specifying the underlying function spaces and defining the notion of solution, let us derive a formal a priori estimate for problem (17)–(20).

The Euler-Lagrange equation for (17) is

$$-\operatorname{div} \left(g(w) \frac{\nabla u}{|\nabla u|} \right) + \mu \frac{u-f}{|u-f|} = 0, \quad \frac{\partial u / \partial \nu}{|\nabla u|} \Big|_{\partial \Omega} = 0. \quad (38)$$

For each $t \in [0, T]$, multiplying (38) by $\frac{w(u-f)}{g(w)}$, and integrating over Ω , we get

$$\left(g(w) \frac{\nabla u}{|\nabla u|}, \nabla \left(\frac{w(u-f)}{g(w)} \right) \right) + \left(\frac{\mu w}{g(w)}, |u-f| \right) = 0. \quad (39)$$

Thus,

$$\begin{aligned} & \left(\frac{\nabla u}{|\nabla u|}, \nabla w(u-f) \right) - \left(\frac{\nabla u}{|\nabla u|}, \frac{g'(w)w}{g(w)} \nabla w(u-f) \right) \\ & + (w, |\nabla u|) - \left(w \frac{\nabla u}{|\nabla u|}, \nabla f \right) + \left(\frac{\mu w}{g(w)}, |u-f| \right) = 0. \end{aligned} \quad (40)$$

Multiplying (38) by $\frac{u-f}{g(w)}$, and integrating over Ω , we find

$$- \left(\frac{\nabla u}{|\nabla u|}, \frac{g'(w)}{g(w)} \nabla w(u-f) \right) + (1, |\nabla u|) - \left(\frac{\nabla u}{|\nabla u|}, \nabla f \right) + \left(\frac{\mu}{g(w)}, |u-f| \right) = 0. \quad (41)$$

Since the last term is non-negative, we conclude that

$$TV(u) \leq TV(f) + C_g \|\nabla w\| \|u-f\|. \quad (42)$$

Multiplying (38) by $u-f$, and integrating over Ω , we derive

$$(g(w), |\nabla u|) + \|\mu(u-f)\|_{L_1} \leq g_0 TV(f) \quad (43)$$

Multiplying (38) by $u|u|$, and integrating over Ω , we deduce

$$2(g(w)|u|, |\nabla u|) + \left(\frac{u-f}{|u-f|}, \mu u|u| \right). \quad (44)$$

It is not difficult to obtain the following scalar inequality

$$|a-b|^3 \leq 2(a-b)a|a| + 2b^2|a-b|, \quad a, b \in \mathbb{R}, \quad (45)$$

which enables to conclude from (44) that

$$\|\sqrt{\mu}(u-f)\|^2 + 4(g(w)|u|, |\nabla u|) \leq 2\|\sqrt{\mu}f\|^2. \quad (46)$$

Hence, due to (28),

$$\|u-f\| \leq C(\|f\|) \leq C(\|f\|_{BV}). \quad (47)$$

By (42) and (47),

$$TV(u) \leq C(1 + \|\nabla w\|). \quad (48)$$

Multiplying (18) by w , and integrating over Ω , we get

$$\frac{1}{2} \frac{d\|w\|^2}{dt} + (\lambda, |\nabla w|^p) + ((1-\lambda)w, w) = ((1-\lambda)w, |\nabla u|) - (w \nabla \lambda, \nabla w). \quad (49)$$

Using Hölder's inequality and (33), we deduce from (40) that

$$(w, |\nabla u|) \leq (1 + C_g) \|\nabla w\|_{L_p} \|u-f\|_{L_{p/p-1}} + (w, |\nabla f|). \quad (50)$$

From (49), (50) and Young's inequality we infer

$$\frac{1}{2} \frac{d\|w\|^2}{dt} + \frac{\lambda_0}{2} \|\nabla w\|_{L_p}^p \leq C(g, \lambda_0, p) \|u - f\|_{L_{p/p-1}}^{p/p-1} + C(\lambda, p) \|w\|_{L_{p/p-1}}^{p/p-1} + (w, |\nabla f|). \quad (51)$$

Provided (35) or (36), estimate (47), and embeddings (30) or (29), resp., imply

$$\frac{d\|w\|^2}{dt} + \lambda_0 \|\nabla w\|_{L_p}^p \leq C(1 + \|w\|^{p/p-1} + \|w\|_{W_p^1}). \quad (52)$$

By (32), (34) and usual arguments, (52) yields

$$\|w\|_{L_\infty(0,T;L_2)} + \|w\|_{L_p(0,T;W_0^{1,p})} \leq C. \quad (53)$$

In case (37), we multiply (38) by $\frac{wu}{g(w)}$, and integrate over Ω , arriving at

$$\begin{aligned} \left(\frac{\nabla u}{|\nabla u|}, u \nabla w \right) - \left(\frac{\nabla u}{|\nabla u|}, u \frac{g'(w)w}{g(w)} \nabla w \right) + (w, |\nabla u|) \\ + \left(\frac{\mu w}{g(w)}, |u - f| \right) + \left(\frac{u - f}{|u - f|}, \frac{\mu w f}{g(w)} \right) = 0. \end{aligned} \quad (54)$$

Then, since the penultimate term is non-negative,

$$(w, |\nabla u|) \leq (1 + C_g) \|\nabla w\| \|u\| + \mu_2 c_g (|f|w, 1 + w) \leq C(1 + \|\nabla w\| + \|w\|^2). \quad (55)$$

Now, (49) and (55) yield

$$\frac{d\|w\|^2}{dt} + \lambda_0 \|\nabla w\|_{L_p}^p \leq C(1 + \|w\|^2), \quad (56)$$

which implies (53). In all the three cases, (43), (48) and (53) imply

$$\|u\|_{L_\infty(0,T;L_1)} + \|u\|_{L_p(0,T;BV)} \leq C. \quad (57)$$

The operator

$$\Lambda : BV \rightarrow \mathcal{M}, \quad \Lambda(v) = (1 - \lambda)|\nabla v|, \quad \langle \Lambda(v), \phi \rangle_{\mathcal{M} \times C_0(\Omega)} = \langle |\nabla v|, (1 - \lambda)\phi \rangle_{\mathcal{M} \times C_0(\Omega)},$$

is bounded, so

$$\|(1 - \lambda)|\nabla u|\|_{L_p(0,T;\mathcal{M})} \leq C. \quad (58)$$

The weighted p -Laplacian operator

$$\Delta_{p,\lambda} : W_0^{1,p} \rightarrow W_{p/p-1}^{-1}$$

is also bounded. Hence, (18),(31),(53) and (58) yield an estimate for the time derivative of w :

$$\|w'\|_{L_p(0,T;W_q^{-1})} \leq C, \quad q < 2, \quad \frac{q}{q-1} \geq p. \quad (59)$$

3.3. Solvability.

Definition 1. Assume (34) and (35). A pair of functions (u, w) from the class

$$\begin{aligned} u &\in L_\infty(0, T; L_1) \cap L_p(0, T; BV), \\ w &\in L_\infty(0, T; L_2) \cap L_p(0, T; W_0^{1,p}) \cap W_p^1(0, T; W_{p/p-1}^{-1}), \end{aligned}$$

is called a weak solution to problem (17)–(20) if

$$TV_{g(w(t))}(u(t)) + \|\mu(u(t) - f)\|_{L_1} \leq TV_{g(w(t))}(\mathbf{u}(t)) + \|\mu(\mathbf{u}(t) - f)\|_{L_1}, \quad (60)$$

for any $\mathbf{u} \in S(0, T; BV)$ and a.a. $t \in (0, T)$,

$$w' - \Delta_{p,\lambda} + (1 - \lambda)w = (1 - \lambda)|\nabla u| \quad (61)$$

in the space $W_{p/p-1}^{-1}$ for a.a. $t \in (0, T)$, and

$$w(0) = F \quad (62)$$

in $W_{p/p-1}^{-1}$.

Remark 3. This definition is correct since all members of (61) belong to $W_{p/p-1}^{-1}$ for a.a. $t \in (0, T)$ (cf. the end of Subsection 3.2), and $w \in W_p^1(0, T; W_{p/p-1}^{-1}) \subset C([0, T]; W_{p/p-1}^{-1})$.

Theorem 1. Assume (34) and (35). Then there exists a weak solution to (17)–(20).

Proof: We can prove the existence of weak solutions via approximation of (17)–(20) by a more regular problem, and consequent passage to the limit (cf. [40, 50]). Let (u_m, w_m) be a sequence of “approximate” solutions (possibly with “approximate” data f_m and F_m). It essentially suffices to show that (60)–(62) is the limiting case of (17)–(20), i.e., that it is possible to pass to the limit in all the members.

Due to estimates (53), (57), (59), without loss of generality we may suppose that

$$u_m \rightarrow u \text{ weakly } - * \text{ in } L_\infty(0, T; \mathcal{M}), \quad (63)$$

$$u_m \rightarrow u \text{ weakly } - * \text{ in } L_p(0, T; BV), \quad (64)$$

$$w_m \rightarrow w \text{ weakly } - * \text{ in } L_\infty(0, T; L_2), \quad (65)$$

$$w_m \rightarrow w \text{ weakly in } L_p(0, T; W_0^{1,p}), \quad (66)$$

$$w'_m \rightarrow w' \text{ weakly in } L_p(0, T; W_{p/p-1}^{-1}). \quad (67)$$

Note that

$$u \in L_\infty(0, T; L_1) \subset L_\infty(0, T; \mathcal{M}) \cap L_p(0, T; BV). \quad (68)$$

By (30), (31) and the Aubin-Lions-Simon theorem [50],

$$w_m \rightarrow w \text{ strongly in } L_p(0, T; C(\bar{\Omega})), \quad (69)$$

$$w_m \rightarrow w \text{ strongly in } C([0, T]; W_{p/p-1}^{-1}), \quad (70)$$

so

$$w_m(0) \rightarrow w(0) \text{ in } W_{p/p-1}^{-1}, \quad (71)$$

and we can pass to the limit in (62).

Using the representation

$$\|v\|_{L_1} = \sup_{\varphi \in L_\infty, \|\varphi\|_{L_\infty} \leq 1} (\varphi, v), \quad (72)$$

and lower semicontinuity of suprema, we can check that

$$\|\mu(u(t) - f)\|_{L_1} \leq \liminf_{m \rightarrow +\infty} \|\mu(u_m(t) - f)\|_{L_1} = \liminf_{m \rightarrow +\infty} \|\mu(u_m(t) - f_m)\|_{L_1} \quad (73)$$

for a.a. $t \in (0, T)$. By Lemma 1,

$$TV_{g(w(t))}(u(t)) \leq \liminf_{m \rightarrow +\infty} TV_{g(w(t))}(u_m(t)). \quad (74)$$

But

$$\begin{aligned} |TV_{g(w_m(t))}(u_m(t)) - TV_{g(w(t))}(u_m(t))| &\leq \|g(w_m(t)) - g(w(t))\|_{L_\infty} TV(u_m(t)) \\ &\leq C(g) \|w_m(t) - w(t)\|_{L_\infty} TV(u_m(t)), \end{aligned} \quad (75)$$

so

$$\|TV_{g(w_m)}(u_m) - TV_{g(w)}(u_m)\|_{L_{p/2}(0, T)} \leq C \|w_m - w\|_{L_p(0, T; L_\infty)} \|u_m\|_{L_p(0, T; BV)} \rightarrow 0. \quad (76)$$

Therefore, without loss of generality,

$$TV_{g(w_m(t))}(u_m(t)) - TV_{g(w(t))}(u_m(t)) \rightarrow 0 \quad (77)$$

for a.a. $t \in (0, T)$. Due to (73), (74), (77), we can pass to the limit in (60).

On the other hand, (60) with $u = u_m$, $f = f_m$, $w = w_m$, $\mathbf{u} = u$ gives

$$TV_{g(w_m(t))}(u_m(t)) + \|\mu(u_m(t) - f_m)\|_{L_1} \leq TV_{g(w_m(t))}(u(t)) + \|\mu(u(t) - f_m)\|_{L_1} \quad (78)$$

Similarly to (75)–(77), we can check that

$$TV_{g(w_m(t))}(u(t)) - TV_{g(w(t))}(u(t)) \rightarrow 0. \quad (79)$$

From (73), (74), (77)–(79) we conclude that

$$TV_{g(w(t))}(u_m(t)) \rightarrow TV_{g(w(t))}(u(t)) \quad (80)$$

for a.a. $t \in (0, T)$.

Fix any non-negative function $\phi \in C_0(\Omega)$. Let

$$\kappa(t) = \left\| \frac{\phi}{g(w(t))} \right\|_{L^\infty} \quad (81)$$

and

$$\varphi(t) = \kappa(t)g(w(t)) - \phi. \quad (82)$$

For a.a. $t \in (0, T)$, $\varphi(t)$ is a non-negative continuous function on $\bar{\Omega}$. By Lemma 1 and (80), we infer that

$$\begin{aligned} TV_\phi(u(t)) &= \kappa(t)TV_{g(w(t))}(u(t)) - TV_{\varphi(t)}(u(t)) \\ &\geq \limsup_{m \rightarrow +\infty} (\kappa(t)TV_{g(w(t))}(u_m(t)) - TV_{\varphi(t)}(u_m(t))) = \limsup_{m \rightarrow +\infty} TV_\phi(u_m(t)). \end{aligned} \quad (83)$$

But, due to (25),

$$TV_\phi(u(t)) \leq \liminf_{m \rightarrow +\infty} TV_\phi(u_m(t)). \quad (84)$$

Thus,

$$TV_\phi(u(t)) = \lim_{m \rightarrow +\infty} TV_\phi(u_m(t)), \quad (85)$$

for every non-negative $\phi \in C_0(\Omega)$, which yields

$$|\nabla u_m(t)| \rightarrow |\nabla u(t)| \quad (86)$$

weakly-* in \mathcal{M} for a.a. $t \in (0, T)$. Then (31) implies

$$(1 - \lambda)|\nabla u_m(t)| \rightarrow (1 - \lambda)|\nabla u(t)| \quad (87)$$

strongly in $W_{p/p-1}^{-1}$ for a.a. $t \in (0, T)$. Due to (64) and (31),

$$\|(1 - \lambda)|\nabla u_m|\|_{L_p(0, T; W_{p/p-1}^{-1})} \leq C. \quad (88)$$

By (87), (88) and [29, Proposition 2.8, Remark 2.10],

$$(1 - \lambda)|\nabla u_m| \rightarrow (1 - \lambda)|\nabla u| \text{ strongly in } L_q(0, T; W_{p/p-1}^{-1}), \quad \forall q < p. \quad (89)$$

Rewrite (61) as

$$w' + Aw = K(u, w), \quad (90)$$

where

$$A(w) = -\operatorname{div}(\lambda|\nabla w|^{p-2}\nabla w) + (1 - \lambda)w, \quad K(u, w) = -\nabla w \cdot \nabla \lambda + (1 - \lambda)|\nabla u|. \quad (91)$$

It is easy to see that the operator $A : W_0^{1,p} \rightarrow W_{p/p-1}^{-1}$ is monotone, coercive and hemi-continuous (cf. [27]). By (69) and (89),

$$K(u_m, w_m) \rightarrow K(u, w) \text{ strongly in } L_{p/p-1}(0, T; W_{p/p-1}^{-1}).$$

Hence, we can successfully pass to the limit in (90) via Minty-Browder monotonicity technique (cf. [27]). \blacksquare

Definition 2. Assume (34) and (36) or (37). A pair of functions (u, w) from the class

$$u \in L_\infty(0, T; L_1) \cap L_2(0, T; BV), \quad (92)$$

$$w \in L_\infty(0, T; L_2) \cap L_2(0, T; W_0^{1,2}) \cap W_2^1(0, T; W_q^{-1}), \quad \forall q < 2, \quad (93)$$

is called a pseudosolution to problem (17)–(20) if there is a sequence (u_m, w_m, p_m) such that each pair (u_m, w_m) is a weak solution to (17)–(20) with $p = p_m$,

$$\begin{aligned} u_m &\rightarrow u \text{ weakly } - * \text{ in } L_\infty(0, T; \mathcal{M}), \\ u_m &\rightarrow u \text{ weakly } - * \text{ in } L_2(0, T; BV), \\ w_m &\rightarrow w \text{ weakly } - * \text{ in } L_\infty(0, T; L_2), \\ w_m &\rightarrow w \text{ weakly in } L_2(0, T; W_0^{1,2}), \\ w_m &\rightarrow w \text{ strongly in } L_2(0, T; L_q), \quad \forall q < +\infty, \\ w_m &\rightarrow w \text{ strongly in } C([0, T]; W_q^{-1}), \quad \forall q < 2, \\ w'_m &\rightarrow w' \text{ weakly in } L_2(0, T; W_q^{-1}), \quad \forall q < 2, \\ p_m &\rightarrow 2. \end{aligned}$$

Theorem 2. *Assume (34) and (36) or (37). Then there exists a pseudosolution to (17)–(20).*

The proof is based on estimates (53), (57), (59) and the proof of Theorem 1.

4. Experimental Results

4.1. Implementation details. The proposed scheme is implemented using the dual minimization [12] for the weighted TV (Eqn. (7)) and explicit Euler finite difference scheme for the non-homogenous linear diffusion (Eqn. (8)). The edge indicator function $g(w) = 1/(1 + w^2)$ is used for all the results reported here. We obtained similar results for other g functions. The adaptive μ_1 based results are reported here unless otherwise stated explicitly and μ_2 provided similar results whereas μ_3 provided blurred cartoon components, see Section 2.3 for details. The parameters $\delta x = 1$, $\delta t = 1/8$ and $\theta = 10^{-2}$ are fixed, and the best results according to the $\max(|u^{n+1} - u^n|, |v^{n+1} - v^n|) \leq \epsilon$ are shown. By constant λ or μ in the results we mean they are taken as constant value 1. Currently there are no quantitative ways to evaluate different decomposition algorithms. In particular which smooth, texture and edge separation model are better is an open research question in image quality assessment. The algorithm is visualized in MATLAB 7.8(R2009a) on a 64-bit Windows 7 laptop with 3GB RAM, 2.20GHz CPU. It takes < 10 sec for 50 iterations for 3 channels image of size 481×321 .

4.2. Image decomposition results.

4.2.1. Gray-scale images. We first show decomposition results of Bresson et al [9] with our model in Figure 6 for a synthetic image which consists of two different texture regions. Comparing the cartoon - texture decomposition of our scheme (Figure 6(b)) with the results of Bresson et al (Figure 6(c)), we see that they behave different visually. For example, the shape of the diamond at the center is preserved well in our scheme whereas the Bresson et al [9] scheme blurs it in the final result. Figure 6(c) shows the energy value against number of iterations for the same synthetic image, which indicates that our adaptive CTE scheme decreases the energy values comparable to Bresson et al [9] model. More grayscale image decomposition results are given in Figure 1. We see that the cartoon component obtained are piecewise constant segments indicating the potential for image segmentation [36]. The

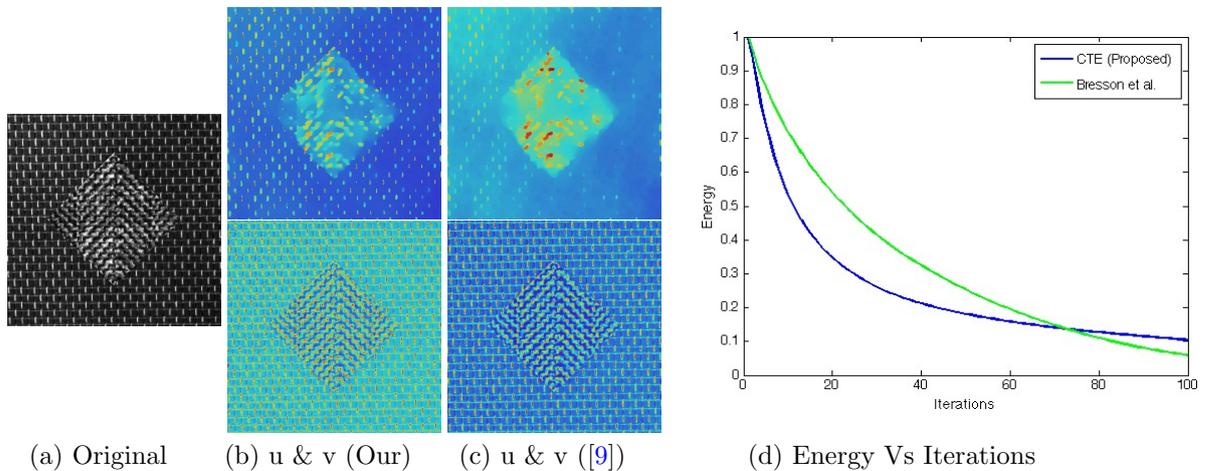


FIGURE 6. (Color online) Comparison of our constant μ , λ proposed scheme (second column) with Bresson et al [9] (third column), shows that our scheme preserves large-scale textures and shape boundaries. Energy value comparison between our scheme CTE and Bresson et al. [9] shows similar convergence property. Best viewed electronically, zoomed in.

texture and edges component are complementary and it is clear that edges are based on the cartoon subregions, see for example, Figure 1(d) top row.

4.2.2. Color images. We further provide image decomposition for color images [8, 16] by using vectorial TV version in Eqn. (7) summing up the contributions of the separate channels. In these sense, we consider the following vectorial non-homogenous diffusion equation:

$$\min_{\mathbf{u}=(u_1, u_2, u_3)} \left\{ \sum_{i=1}^3 \int_{\Omega} g(w_i) |\nabla u_i| dx + \sum_{i=1}^3 \int_{\Omega} \mu_i(x) |u_i - f_i| dx \right\},$$

$$\frac{\partial w_i}{\partial t} = \lambda_i(x) \operatorname{div}(\nabla w_i) + (1 - \lambda_i(x)) (|\nabla u_i| - w_i),$$

were each scalar function $u_i : \Omega \rightarrow \mathbb{R}$, $1 \leq i \leq 3$ represent one component of the RGB color system. Figure 7 shows the cartoon components of our CTE scheme with constant and adaptive μ_1 against the traditional TV based scheme of [9] for two standard RGB test images. As can be seen, our scheme obtains better edge preserving cartoon (u) components. The close-up shots indicate that our scheme also avoids blurring of edges, see for example *Barbara* face. See also Figure 2 where the proposed scheme with adaptive μ_1 provides better shape preservation as the number of iterations increase

(b) Our constant μ , λ based scheme

(c) Bresson et al [9]

(b) Our adaptive μ_1 , constant λ based scheme

(c) Bresson et al [9]

FIGURE 7. (Color online) Our diffusion constrained total variation scheme provides better edge preserving cartoon component u when compared to the traditional TV regularization model [9]. Even with constant μ , λ the proposed scheme provides better results (see top row (b)). The crop regions highlight that the proposed scheme provides better preservation of large scale textures compared to [9] model. Best viewed electronically, zoomed in.

from 100 to 400. Next, Figures 8-10 shows decomposition for a variety of RGB images for two different iteration values in our proposed CTE scheme with constant μ against adaptive μ_1 based results. As can be seen in Figure 8, increasing the number of iterations removes more texture details and provides piecewise constant (smoother) cartoon images. Our adaptive μ_1 based scheme results (last row) on the other hand keep most of the salient edges. This can be seen further in Figure 9 (last row) where the adaptive μ_1 based scheme captures only small scale oscillations corresponding to texture components whereas the constant μ based results remove strong edges as well. Figure 10 show the corresponding edge functions, and it can be seen that adaptive scheme has more information than in constant parameter case. Thus, we conclude that, using adaptive μ provides an image adaptive way

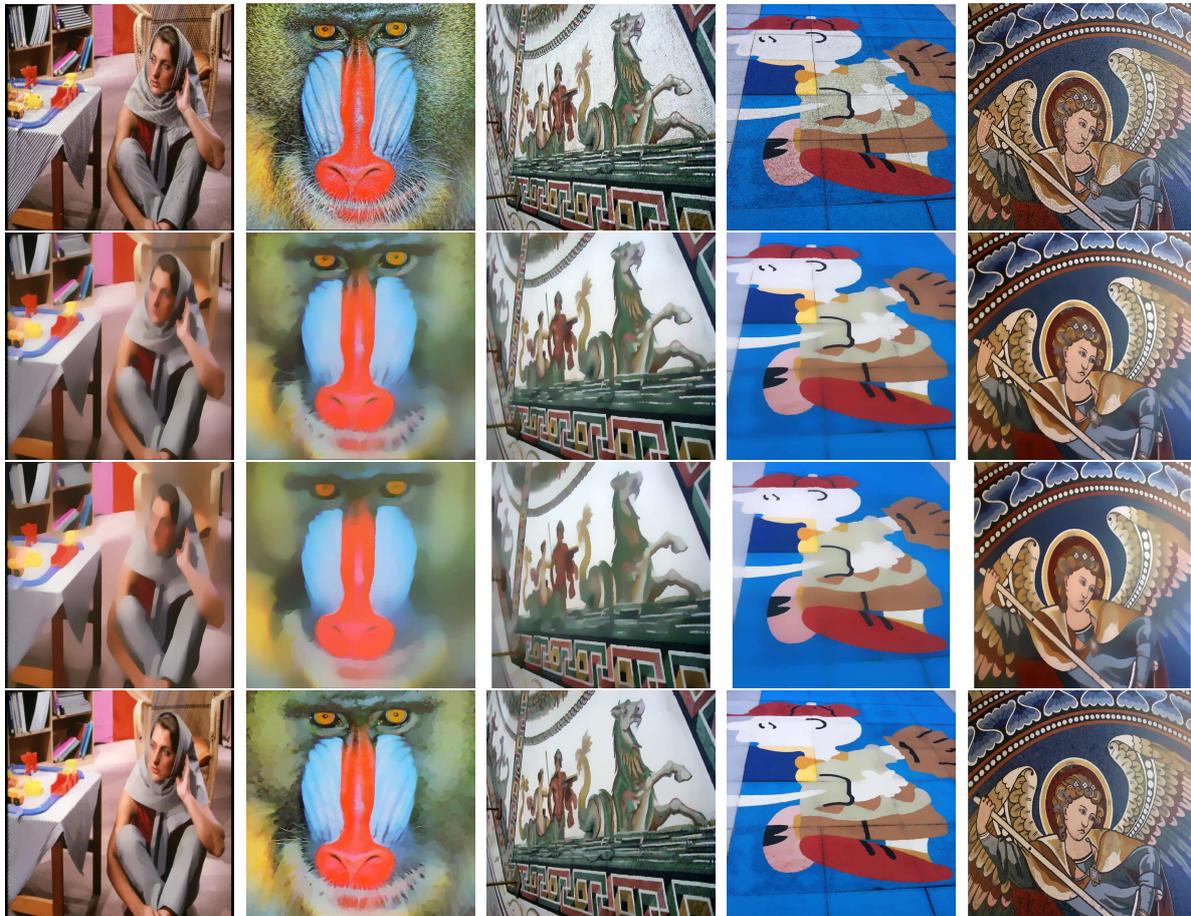


FIGURE 8. (Color online) Effect of constant μ in color image decomposition using our coupled scheme on the cartoon (u) component. First row: Original color RGB images. Cartoon component result for constant μ , λ with total number of iterations, Second row: 100, and Third row: 200. As can be seen increasing the number of iterations removes more texture details and provides piecewise constant cartoon image. Last row: Shows the proposed scheme results with adaptive μ_1 , see Eqn. (14) and constant λ . Best viewed electronically, zoomed in.

of obtaining edge preserving cartoon components without sacrificing overall decomposition properties.

4.3. Image denoising results. Note that the decomposition provides piecewise cartoon component which is obtained using a weighted TV regularization in an edge preserving way, see Figure 8 (last row). Hence, as a byproduct we obtain image denoising, with u the denoised image and $v + w$ the ‘noise’ part. To compare the schemes quantitatively for denoising we utilize two commonly used error metrics in the image denoising literature, one is the

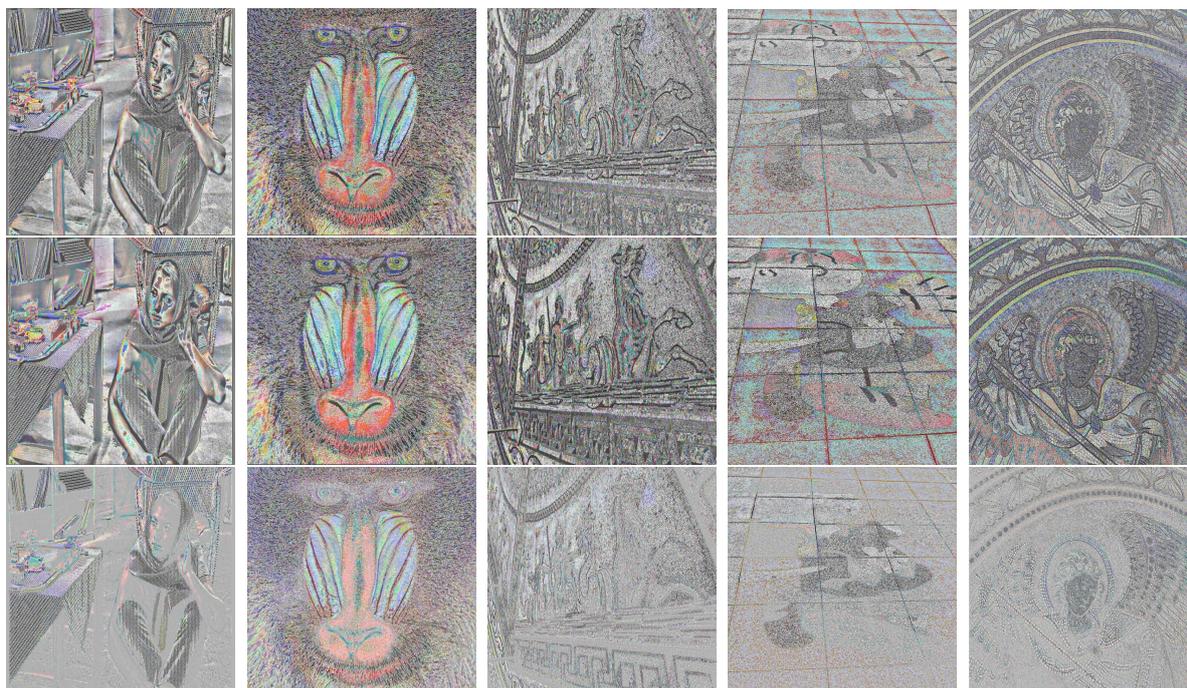


FIGURE 9. (Color online) Corresponding texture components v . Arrangement is as in Figure 8. Best viewed electronically, zoomed in.



FIGURE 10. (Color online) Corresponding edge functions w . Arrangement is as in Figure 8. Best viewed electronically, zoomed in.

classical peak signal to noise ratio (PSNR) [3], and the other one is the mean structural similarity measure (MSSIM) [47]:

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

$$\text{PSNR}(u) := 20 * \log_{10} \left(\frac{u_{max}}{\sqrt{MSE}} \right) dB, \quad (94)$$

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 $u_{max} = 255$.

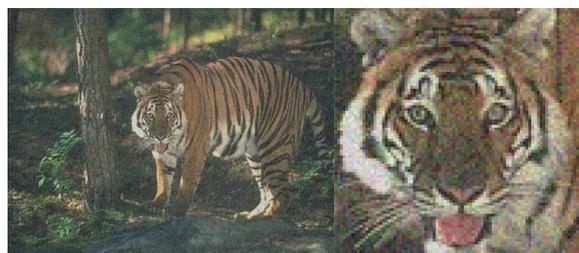
- (2) MSSIM index is in the range $[0, 1]$. The MSSIM value near one implies the optimal denoising capability of the scheme [47] and is mean value of the SSIM metric. The SSIM is calculated between two windows ω_1 and ω_2 of common size $N \times N$,

$$\text{SSIM}(\omega_1, \omega_2) = \frac{(2\mu_{\omega_1}\mu_{\omega_2} + c_1)(2\sigma_{\omega_1\omega_2} + c_2)}{(\mu_{\omega_1}^2 + \mu_{\omega_2}^2 + c_1)(\sigma_{\omega_1}^2 + \sigma_{\omega_2}^2 + c_2)}, \quad (95)$$

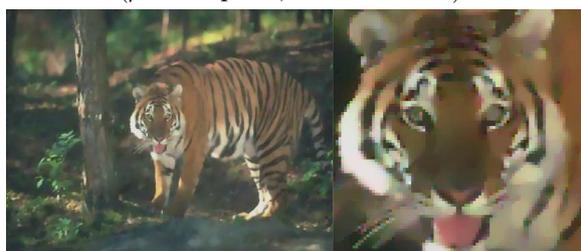
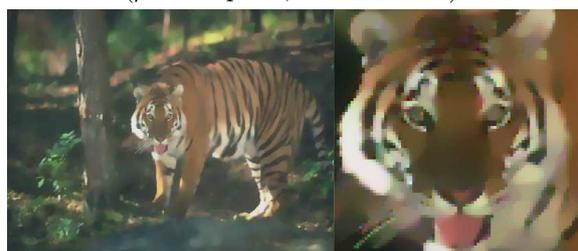
where μ_{ω_i} the average of ω_i , $\sigma_{\omega_i}^2$ the variance of ω_i , $\sigma_{\omega_1\omega_2}$ the covariance, c_1, c_2 stabilization parameters, see [47] for more details.

Remark that the SSIM is a better error metric than PSNR as it provides a quantitative way of measuring the structural similarity of denoised image against the original noise-free image. Table 1 compares the proposed scheme with that of [40] using the both PSNR (dB) and MSSIM error metrics average for Berkeley segmentation dataset (BSDS) images. We implemented both the schemes on the full Berkeley segmentation dataset of 500 noisy images for two different noise levels and obtained similar improvements. Figures 11-12 show some example images corresponding to Table 1. As can be seen from the zoomed in versions, the proposed CTE scheme provides cleaner cartoon components (denoised images, see Figures 11(b-e)-12(b,c)) in contrast to original coupled PDE model [40] which either excessively blurs out details (Figure 11(f,g)) or keeps noisy regions (Figure 12(d)) in final results.

4.4. Multi-scale decomposition. Following Tang and He [45] we can make the weighted TV model with multi-scale parameter λ . Note that this is slightly different from the original multi scale usage in [44], here we use it in the constraint Eqn. (8). Let us briefly recall the model proposed in [45] where the texture component v is modeled using G_p norm. That is, the



(a) Input noise 20%

(b) Proposed 30 iterations
($\mu \leftarrow$ adaptive, $\lambda \leftarrow$ constant)(c) Proposed 100 iterations
($\mu \leftarrow$ adaptive, $\lambda \leftarrow$ constant)(d) Proposed 30 iterations
($\mu \leftarrow$ constant, $\lambda \leftarrow$ adaptive)(e) Proposed 40 iterations
($\mu \leftarrow$ constant, $\lambda \leftarrow$ adaptive)

(f) [40] 20 iterations



(g) [40] 60 iterations

FIGURE 11. (Color online) Better edge preserving image restoration results were obtained using our scheme in comparison with the original coupled PDE model [40]. The iteration numbers were chosen according to maximum MSSIM values, see Table 1 for the corresponding values.

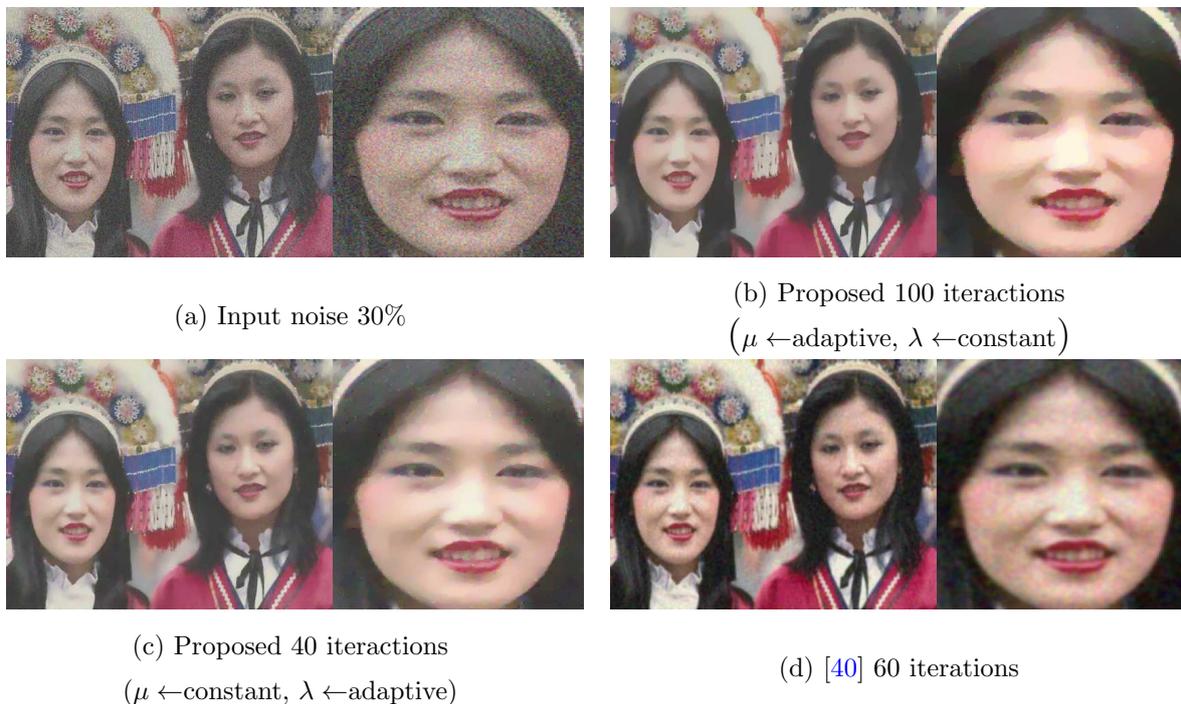


FIGURE 12. (Color online) Better edge preserving image restoration results were obtained using our scheme in comparison with the original coupled PDE model [40]. The iteration numbers were chosen according to maximum MSSIM values, see Table 1 for the corresponding values.

minimization is carried out for both u, v ,

$$\inf_{u,v} \left\{ E(f, \lambda; u, v) = |u|_{BV(\Omega)} + \mu \|f - u - v\|_{L^2(\Omega)}^2 + \lambda \|v\|_{G_p(\Omega)} \right\},$$

with G_p consisting of all distributions which can be written as,

$$v = \partial_x g_1 + \partial_y g_2 = \operatorname{div}(\vec{g}), \quad \vec{g} \in L^p(\Omega, \mathbb{R}^2).$$

The G_p norm is defined as,

$$\|v\|_{G_p(\Omega)} = \inf \left\{ \|\vec{g}\|_{L^p(\Omega)} \mid v = \operatorname{div}(\vec{g}), \vec{g} \in L^p(\Omega, \mathbb{R}^2) \right\}.$$

We utilize the same modeling for the texture component v in our splitting step of the proposed weighted TV model (see Eqn. (4)),

$$\min_{u,v} \left\{ \int_{\Omega} |\nabla u| dx + \mu \|f - u - v\|_{L^2(\Omega)}^2 + \lambda \|v\|_{G_p(\Omega)} \right\} \quad (96)$$

Finally, we compare our multi-scale version of the scheme (96) with the multi-scale TV decomposition results of [45]. Figure 13 shows the comparison result

| Method | $\mu \leftarrow$ constant $\lambda \leftarrow$ adaptive | $\mu \leftarrow$ adaptive $\lambda \leftarrow$ constant | Figure | Noise | Iterations | Convergence Time (Seconds) | PSNR (dB) | MSSIM |
|----------------|------------------------------------------------------------|------------------------------------------------------------|--------|-------|------------|-------------------------------|-----------|--------|
| CTE (Proposed) | | ✓ | 11 (b) | 20% | 30 | 30.47 | 28.7938 | 0.7723 |
| | | | 11 (c) | 20% | 100 | 38.71 | 27.3510 | 0.7147 |
| | | | 12 (b) | 30% | 100 | 38.61 | 26.5238 | 0.6774 |
| CTE (Proposed) | ✓ | | 11 (d) | 20% | 30 | 30.96 | 28.2464 | 0.7237 |
| | | | 11 (e) | 20% | 40 | 32.20 | 26.1983 | 0.6796 |
| | | | 12 (c) | 30% | 40 | 32.33 | 26.7451 | 0.6792 |
| [40] | ✓ | | 11 (f) | 20% | 20 | 5.07 | 27.3646 | 0.7178 |
| | | | 11 (g) | 20% | 60 | 40.57 | 25.4476 | 0.6970 |
| | | | 12 (d) | 30% | 60 | 40.30 | 25.0559 | 0.6543 |

TABLE 1. Image denoising error metrics (average) comparison using original coupled PDE scheme [40] for different noise levels and parameters on the Berkeley segmentation dataset (BSDS) 500. Some examples corresponding to the entries are shown above in Figure 11 and Figure 12.

on a synthetic image for 5 steps and our scheme retains the cartoon component clearly than the Tang and He [45]. Moreover, the texture components show a progressive capture of small scale oscillations.

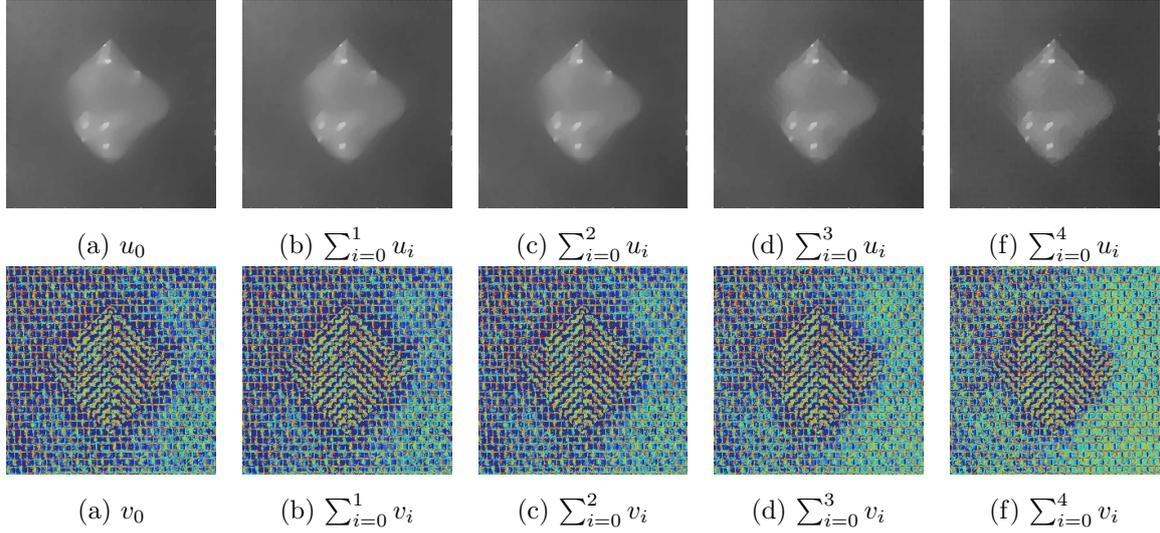
5. Conclusion

We have presented a new image decomposition model coupling a variational and PDE problem via a weighted total variation regularization algorithm. Our main contribution is twofold: 1) the proposed decomposition model gets an image decomposition into its cartoon, texture and edge components with fixed and adaptive parameters by utilizing a local histogram along with a diffusion equation. Extensive experiments using a fast dual minimization based splitting implementation indicates that the proposed scheme is useful for edge preserving image decomposition on real and noisy images and comparative results indicate the proposed scheme is useful in denoising natural images as well in multi scale image decomposition. 2) We fashioned a new well posed scheme to transform the non-linear problem to a set of subproblems that are much easier to solve quickly via an optimization technique and a linear diffusion PDE solution.

References

- [1] R. Adams. *Sobolev spaces*. Academic Press, New York, NY, USA, 1975. Pure and Applied Mathematics, Vol. 65.
- [2] P. Athavale and E. Tadmor. Integro-differential equations based on (BV, L^1) image decomposition. *SIAM Journal on Imaging Sciences*, 4(1):300–312, 2011.
- [3] G. Aubert and P. Kornprobst. *Mathematical problems in image processing: Partial differential equation and calculus of variations*. Springer-Verlag, New York, USA, 2006.

Multiscale cartoon and texture extraction using our proposed CTE method



Multiscale cartoon and texture extraction using the hierarchical decomposition of Tang and He [45]

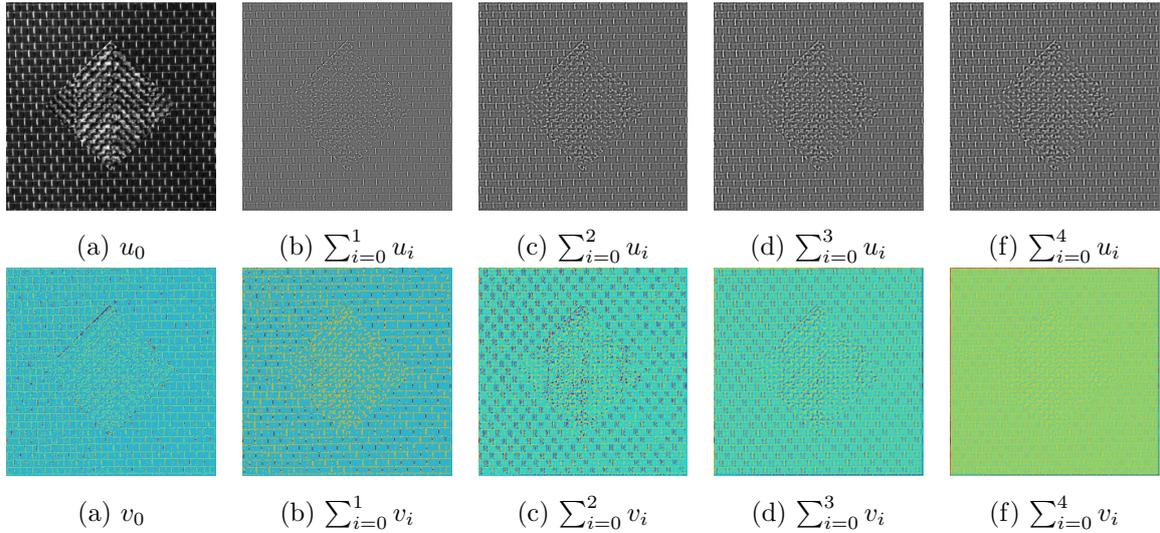


FIGURE 13. (Color online) Multiscale cartoon and texture decomposition of a synthetic image for 5 steps by implementing our proposed approach (top two rows) and the hierarchical method from [45] (bottom two rows).

- [4] J.-F. Aujol, G. Aubert, L. Blanc-Féraud, and A. Chambolle. Image decomposition into a bounded variation component and a oscillating component. *Journal of Mathematical Imaging and Vision*, 22(1):71–88, 2005.
- [5] J.-F. Aujol and A. Chambolle. Dual norms and image decomposition models. *International Journal of Computer Vision*, 63(1):85–104, 2005.

- [6] J.-F. Aujol, G. Gilboa, T. Chan, and S. Osher. Structure-texture image decomposition – modeling, algorithms and parameter selection. *International Journal of Computer Vision*, 67(1):111–136, 2006.
- [7] J.-F. Aujol, G. Gilboa, T. Chan, and S. Osher. Structure-texture image decomposition – modeling, algorithms and parameter selection. *International Journal of Computer Vision*, 67(1):111–136, 2006.
- [8] X. Bresson and T. F. Chan. Fast dual minimization of the vectorial total variation norm and applications to color image processing. *Inverse Problem and Imaging*, 2(4):455–484, 2008.
- [9] X. Bresson, S. Esedoglu, P. Vandergheynst, J. Thiran, and S. Osher. Fast global minimization of the active contour/snake model. *Journal of Mathematical Imaging and Vision*, 28(2):151–167, 2007.
- [10] V. Caselles, A. Chambolle, D. Cremers, M. Novaga, and T. Pock. An introduction to total variation in image analysis. In *Theoretical Foundations and Numerical Methods for Sparse Recovery (Ed. M. Fornasier)*, volume 9 of *Radon Series on Computational and Applied Mathematics*, pages 263–340. De Gruyter, 2010.
- [11] V. Caselles, A. Chambolle, and M. Novaga. Total variation in imaging. In *Handbook of Mathematical Methods in Imaging (Ed. O. Scherzer)*, pages 1016–1057. Springer, 2011.
- [12] A. Chambolle. An algorithm for total variation minimization and applications. *Journal of Mathematical Imaging and Vision*, 20(1–2):89–97, 2004.
- [13] T. F. Chan and S. Esedoglu. Aspects of total variation regularized L1 function approximation. *SIAM Journal on Applied Mathematics*, 65(5):1817–1837, 2005.
- [14] T. F. Chan, A. Marquina, and P. Mulet. High-order total variation-based image restoration. *SIAM Journal on Scientific Computing*, 22(2):503–516, 2000.
- [15] V. Duval, J.-F. Aujol, and Y. Gousseau. The TVL1 model: A geometric point of view. *Multiscale Modeling and Simulation*, 8(1):154–189, 2009.
- [16] V. Duval, J.-F. Aujol, and L. Vese. Mathematical modeling of textures: application to color image decomposition with a projected gradient algorithm. *Journal of Mathematical Imaging and Vision*, 37(3):232–248, 2010.
- [17] Irene Fonseca and Giovanni Leoni. *Modern methods in the calculus of variations: L^p spaces*. Springer Monographs in Mathematics. Springer, New York, 2007.
- [18] J. Garnett, T. Le, Y. Meyer, and L. Vese. Image decomposition using bounded variation minimization and generalized homogeneous Besov spaces. *Applied and Computational Harmonic Analysis*, 23(1):25–56, 2007.
- [19] G. Gilboa, N. Sochen, and Y. Y. Zeevi. Variational denoising of partly textured images by spatially varying constraints. *IEEE Transactions on Image Processing*, 15(8):2281–2289, 2006.
- [20] J. Gilles. Multiscale texture separation. *Multiscale Modelling and Simulation*, 10(4):1409–1427, 2012.
- [21] A. Haddad. Texture separation $BV - G$ and $BV - L^1$ models. *Multiscale Modelling and Simulation*, 6(1):273–286, 2007.
- [22] Y. Hu and M. Jacob. Higher degree total variation (hdtv) regularization for image recovery. *IEEE Transactions on Image Processing*, 21(5):2559–2571, 2012.
- [23] T. Le, L. Lieu, and L. Vese. (Φ, Φ^*) image decomposition and minimization algorithms. *Journal of Mathematical Imaging and Vision*, 33(2):135–148, 2009.
- [24] T. Le, J.-M. Morel, and L. Vese. Fast cartoon + texture image filters. *IEEE Transactions on Image Processing*, 19(8):1978–1986, 2010.
- [25] T. Le and L. Vese. Image decomposition using total variation and $\text{div}(\text{BMO})$. *Multiscale Modeling and Simulation*, 4(2):390–423, 2005.

- [26] L. Lieu and L. Vese. Image restoration and decomposition via bounded total variation and negative Hilbert Sobolev spaces. *Applied Mathematics and Optimization*, 58(2):167–193, 2008.
- [27] J.-L. Lions. *Quelques méthodes de résolution des problèmes aux limites non linéaires*. Dunod, 1969.
- [28] Y. Meyer. *Oscillating Patterns in Image Processing and Nonlinear Evolution Equations*. American Mathematical Society, Boston, MA, USA, 2001. The Fifteenth Dean Jacqueline B. Lewis Memorial Lectures, Vol. 22 of University Lecture Series.
- [29] Diego R. Moreira and Eduardo V. Teixeira. On the behavior of weak convergence under nonlinearities and applications. *Proc. Amer. Math. Soc.*, 133(6):1647–1656 (electronic), 2005.
- [30] K. Ni, X. Bresson, T. Chan, and S. Esedoglu. Local histogram based segmentation using the Wasserstein distance. *International Journal of Computer Vision*, 84(1):97–111, 2009.
- [31] M. Nikolova. Minimizers of cost-function involving nonsmooth data-fidelity terms. *SIAM Journal on Numerical Analysis*, 40(3):965–994, 2004.
- [32] M. Nikolova. Weakly constrained minimization: application to the estimation of images and signals involving constant regions. *Journal of Mathematical Imaging and Vision*, 21(2):155–175, 2004.
- [33] S. Osher, A. Solé, and L. Vese. Image decomposition and restoration using total variation minimization and the h^1 . *Multiscale Modelling and Simulation*, 1(3):349–370, 2006.
- [34] K. Papafitsoros and C. Schonlieb. A combined first and second order variational approach for image reconstruction. *Journal of Mathematical Imaging and Vision*, 2013.
- [35] P. Perona and J. Malik. Scale-space and edge detection using anisotropic diffusion. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 12(7):629–639, 1990.
- [36] V. B. S. Prasath. Color image segmentation based on vectorial multiscale diffusion with inter-scale linking. In *Third International Conference on Pattern Recognition and Machine Intelligence (PReMI-09)*, pages 339–344, Delhi, India, December 2009. Springer LNCS 5909 Eds.: S. Chaudhury, S. Mitra, C. A. Murthy, P. S. Sastry, Sankar K. Pal.
- [37] V. B. S. Prasath and A. Singh. A hybrid convex variational model for image restoration. *Applied Mathematics and Computation*, 215(10):3655–3664, 2010.
- [38] V. B. S. Prasath and A. Singh. Well-posed inhomogeneous nonlinear diffusion scheme for digital image denoising. *Journal of Applied Mathematics*, 2010:14pp, 2010. Article ID 763847.
- [39] V. B. S. Prasath and A. Singh. An adaptive anisotropic diffusion scheme for image restoration and selective smoothing. *International Journal of Image and Graphics*, 12(1):18pp, 2012.
- [40] V. B. S. Prasath and D. Vorotnikov. On a system of adaptive coupled pdes for image restoration. *Journal of Mathematical Imaging and Vision*, Online First, 2012. Available at arXiv:1112.2904.
- [41] L. Rudin, S. Osher, and E. Fatemi. Nonlinear total variation based noise removal algorithms. *Physica D*, 60(1–4):259–268, 1992.
- [42] D. Strong. *Adaptive total variation minimizing image restoration*. PhD thesis, UCLA Mathematics Department, USA, August 1997.
- [43] D. M. Strong and T. F. Chan. Spatially and scale adaptive total variation based regularization and anisotropic diffusion in image processing. Technical Report 96–46, UCLA CAM, 1996.
- [44] E. Tadmor, S. Nezzar, and L. Vese. A multiscale image representation using hierarchical (BV,L2) decomposition. *Multiscale Modeling and Simulation*, 2(4):554–579, 2004.
- [45] L. Tang and C. He. Multiscale texture extraction with hierarchical (BV, G_p, L^2) . *Journal of Mathematical Imaging and Vision*, 45(2):148–163, 2013.
- [46] L. Vese and S. Osher. Modeling textures with total variation minimization and oscillating patterns in image processing. *Journal of Scientific Computing*, 19(1–3):553–572, 2003.

- [47] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli. Image quality assessment: from error visibility to structural similarity. *IEEE Transactions on Image Processing*, 13(4):600–612, 2004.
- [48] W. Yin, D. Goldfarb, and S. Osher. Image cartoon-texture decomposition and feature selection using the total variation regularized L^1 functional. In *Variational, Geometric, and Level Set Methods in Computer Vision*, pages 73–84, Beijing, China, 2005. Lecture Notes in Computer Science Vol. 3752 (Eds. N. Paragios, O. Faugeras, T. Chan and C. Schnörr).
- [49] W. Yin, D. Goldfarb, and S. Osher. The total variation regularized L^1 model for multi scale decomposition. *Multiscale Modeling and Simulation*, 6(1):190–211, 2006.
- [50] V. G. Zvyagin and D. A. Vorotnikov. *Topological approximation methods for evolutionary problems of nonlinear hydrodynamics*, volume 12 of *de Gruyter Series in Nonlinear Analysis and Applications*. Walter de Gruyter & Co., Berlin, Germany, 2008.

JUAN C. MORENO

IT, DEPARTMENT OF COMPUTER SCIENCE, UNIVERSITY OF BEIRA INTERIOR, 6201–001, PORTUGAL.

E-mail address: `jmoreno@ubi.pt`

V. B. SURYA PRASATH

DEPARTMENT OF COMPUTER SCIENCE, UNIVERSITY OF MISSOURI-COLUMBIA, MO 65211 USA.

E-mail address: `prasaths@missouri.edu`

DMITRY VOROTNIKOV

CMUC, DEPARTMENT OF MATHEMATICS, UNIVERSITY OF COIMBRA, 3001–501 COIMBRA, PORTUGAL.

E-mail address: `mitvorot@mat.uc.pt`

HUGO PROENÇA

IT, DEPARTMENT OF COMPUTER SCIENCE, UNIVERSITY OF BEIRA INTERIOR, 6201–001, PORTUGAL.

E-mail address: `hugomcp@ubi.pt`