

COLOR IMAGE PROCESSING BY VECTORIAL TOTAL VARIATION WITH GRADIENT CHANNELS COUPLING

JUAN C. MORENO

Department of Computer Science
University of Beira Interior
6201-001 Covilhã, Portugal

V. B. SURYA PRASATH

Computational Imaging and VisAnalysis (CIVA) Lab
Department of Computer Science
University of Missouri-Columbia
Columbia MO 65211, USA

JOÃO C. NEVES

Department of Computer Science
University of Beira Interior
6201-001 Covilhã, Portugal

(Communicated by Sung Ha Kang)

ABSTRACT. We study a regularization method for color images based on the vectorial total variation approach along with channel coupling for color image processing, which facilitates the modeling of inter channel relations in multi-dimensional image data. We focus on penalizing channel gradient magnitude similarities by using L^2 differences, which allow us to explicitly couple all the channels along with a vectorial total variation regularization for edge preserving smoothing of multichannel images. By using matched gradients to align edges from different channels we obtain multichannel edge preserving smoothing and decomposition. A detailed mathematical analysis of the vectorial total variation with penalized gradient channels coupling is provided. We characterize some important properties of the minimizers of the model as well as provide geometrical results regarding the regularization parameter. We are interested in applying our model to color image processing and in particular to denoising and decomposition. A fast global minimization based on the dual formulation of the total variation is used and convergence of the iterative scheme is provided. Extensive experiments are given to show that our approach obtains good decomposition and denoising results in natural images. Comparison with previous color image decomposition and denoising methods demonstrate the advantages of our approach.

1. Introduction. Techniques based on variational and partial differential equations (PDEs) are widely used in various image processing tasks. In particular, classical ill-posed problems such as image restoration and denoising can be effectively tackled by variational energy minimization models. One of the famous method is the total variation (TV) regularization which was original proposed for edge preserving image restoration [58]. The well-known Rudin-Osher-Fatemi (ROF) image

2010 *Mathematics Subject Classification.* Primary: 47A52; Secondary: 68U10.

Key words and phrases. Vectorial total variation, image decomposition, coupling channels, color image processing, dual formulation, BV space.

denoising model seeks a minimizer of the sum of a fidelity term measured in the square of L^2 -norm along with the TV regularization,

$$(1) \quad \min_{u \in BV(\Omega, \mathbb{R})} \left\{ \int_{\Omega} |\nabla u| \, dx + \lambda \int_{\Omega} |u - f|^2 \, dx \right\}$$

where $f : \Omega \subset \mathbb{R}^2 \rightarrow \mathbb{R}$ is the given noisy image, $BV(\Omega, \mathbb{R})$ is space of bounded variation functions, and $\lambda \geq 0$ is the regularization parameter. The TV term acts as an important regularizer penalizing the function's gradient [63, 36]:

- Sharp edges are recovered without noise and avoids discriminate smoothing associated with L^2 regularization functionals.
- Edge locations are preserved, and under certain conditions edge locations kept exactly. This is important as dislocations of edges while smoothing is detrimental to further higher level image processing tasks.

Despite its usefulness in image restoration, TV regularization is known to exhibit blocky or staircasing artifacts in solutions [47]. To avoid such artifacts various modifications has been studied in the recent years like higher order regularization [38, 41] and adaptive or weighted TV regularization [64, 62, 56, 52, 57]. For example, consider the weighted TV regularization,

$$(2) \quad \min_{u \in BV(\Omega, \mathbb{R})} \left\{ \int_{\Omega} g(x) |\nabla u| \, dx + \lambda \int_{\Omega} |u - f|^2 \, dx \right\}$$

with function g , an edge indicator function that vanishes at object boundaries, $g(x) = (1 + \beta |\nabla f(x)|^2)^{-1}$ where $\beta > 0$ is a constant parameter. Recently, other non-smooth data fidelity terms [45] have found to be useful in various image processing applications [65, 25, 72, 5, 44, 6, 40, 73, 34, 29, 35, 33, 43] where the properties of corresponding minimization formulations are very well established [42, 46].

Bresson et al [14] have considered a similar regularization model with a L^1 data fidelity term,

$$(3) \quad \min_{u \in BV(\Omega, \mathbb{R})} \left\{ \int_{\Omega} g(x) |\nabla u| \, dx + \lambda \int_{\Omega} |u - f| \, dx \right\}.$$

We refer to [22, 30] for the geometric motivation and properties of L^1 fidelity term. The minimization of the $gTV + L^1$ model given by (3) yields a contrast invariant filter [22, 27, 61] and well preserves contrasted features at different scales. The weighted TV norm allows a better preservation of corners and sharp angles in shape denoising. Moreover the introduction of such as function allows to establish a link between gTV and the geodesic active contours model introduced by [17, 50, 19] as an improvement of the original snakes model [39].

The variational minimization problem (3) involves the non-differentiable total variation regularization. Nevertheless, it can be solved by standard calculus of variations and Euler-Lagrange equations toward a smooth approximation of the L^1 -norm driving the function u towards a minimum [9, 23, 14]. Following the work in [24, 20, 5, 8, 7] a convex regularization of the variational model is implemented in [14] by applying a dual formulation of the TV norm. While most existing work involving TV focuses on scalar valued functions, the generalization to vector-valued images remains an important challenge. The main idea of the vectorial total variation (VTV) approach is to extend the TV definition for vector-valued functions $\mathbf{u} = (u_1, \dots, u_N) : \Omega \rightarrow \mathbb{R}^N$, also referred as multispectral images, with $u_i \in BV(\Omega, \mathbb{R})$ such that in the scalar case, $N = 1$, both definitions coincide. Two important notions of vectorial total variation can be considered:

- **Channel by channel summation:** Total variation for vector (multichannel) functions can simply be channel by channel summation [2]. This simple modelling does not assume any coupling between channels and the multidimensional TV consist of summing up the contributions of separate channels. Therefore, chromatic edges can be lost and color smearing can occur between inter-channel edges, see [13, 31] for discussions. The VTV expression is given by,

$$(4) \quad VTV_1(\mathbf{u}) := \sum_{i=1}^N \int_{\Omega} |\nabla u_i| \, dx.$$

- **Implicit coupling with Euclidean norm:** In this approach, an implicit coupling between channels is considered such that each channel uses information coming from other channels [11, 13]. The VTV definition is given by,

$$(5) \quad VTV_2(\mathbf{u}) := \int_{\Omega} \sqrt{\sum_{i=1}^N |\nabla u_i|^2} \, dx.$$

Coupling channels in a variational/PDE setting for improving denoising and other image processing performance has been done by many researchers in the past. One of the earlier works is by Blomgren and Chan [11] who utilized the implicit coupling via the VTV_2 above. However as noted in [66] color smearing can still appear near edges and the staircasing of flat regions can persist. Bresson and Chan [13] used a dual formulation to decrease the computational expense via TV minimization. A multiscale approach along with Moreau-Yosida based primal-dual approach was considered by Dong and Hintermuller [28]. Brook et al [15] used the inter-channel term $E_c(\mathbf{u}) = \sum_{i=1}^3 |\nabla_x u^i \times \nabla_y u^i|^2$, where the derivatives $(\nabla_x u, \nabla_y u)$ are computed in their corresponding (x, y) axis directions. This helps in aligning the color edges as in our model. Moreover the intrachannel term is chosen to be isotropic, $\phi(|\nabla u^i|) = |\nabla u^i|^2$. The corresponding wellposedness of the minimization problem and its numerical convergence are not proven. Also, the term $|\nabla_x u \times \nabla_y u|^2$ is of order $|\nabla u|^4$ and cannot be bounded below by $c|\nabla u|^4$ for some $c > 0$, thus we cannot prove a similar result such as Theorem 2.4 in our work. More recently Ehrhardt and Arridge [32] use a similar alignment for coupling parallel level sets. A closely related approach denoise color or multichannel is to use anisotropic diffusion [59, 66, 12, 55]. They typically involve similar coupling as in the variational approaches or selectively denoise color images in pure chromaticity spaces. Tschumperle and Deriche [68] unified such vector valued diffusion PDEs into one common framework via tensor diffusion. Goldluecke et al [36] use implicit coupling via a natural vectorial total variation motivated from geometric measure theory. Ono and Yamada [48] decorrelate the different channels along with vectorial total variation.

In this paper, we study VTV with an explicit channel coupling term which uses component gradients of a multispectral image. Our aim here is to demonstrate that it is possible to perform an efficient minimization of the VTV with proposed coupling term. This paper, besides developing new coupling models, also proposes a new numerical scheme based on Chambolle's dual minimization [20] to perform the algorithm evolution in an efficient and fast way. The aims of this paper is to

introduce a regularizing term allowing to generalize the TV term to multispectral images. The main contributions of this paper are summarized as follows:

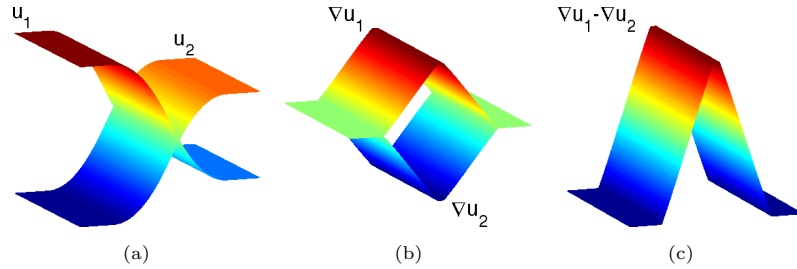
- (1) Introduction of the new vectorial total variation with coupling channels (VTV) approach which penalizes different image gradient components in order to perform edge preservation across channels.
- (2) Presentation of the different properties related to the image processing schemes involving the new VTV norm. In this context, we show the existence of solutions to the decomposition model.
- (3) Extensive numerical implementation to our variational scheme over different image databases show the improvement of our VTV term in relation to other approaches based on functions of bounded variation.

We also consider the L^1 data fidelity term along with VTV regularization with coupling for image decomposition. The coupling terms provide better edge alignment and can also be used in a PDE formulation [55, 51, 53]. Here we utilize weighted TV regularization along with coupling gradient channels for obtaining decomposition and denoising of images. Our aim here is to provide a detailed mathematical analysis of the variational minimization method and provide some applications in decomposition and denoising of noisy images. Experimental results with other related variational/PDE color image denoising schemes indicate that the proposed approach consistently outperforms them in terms of different error metrics. The rest of the paper is organized as follows. Section 2 explains the model in detail using the theory of level sets, and in Section 3 we describe the implementation of the proposed model using the dual minimization. Section 4 provides some experimental results. Section 5 concludes the paper.

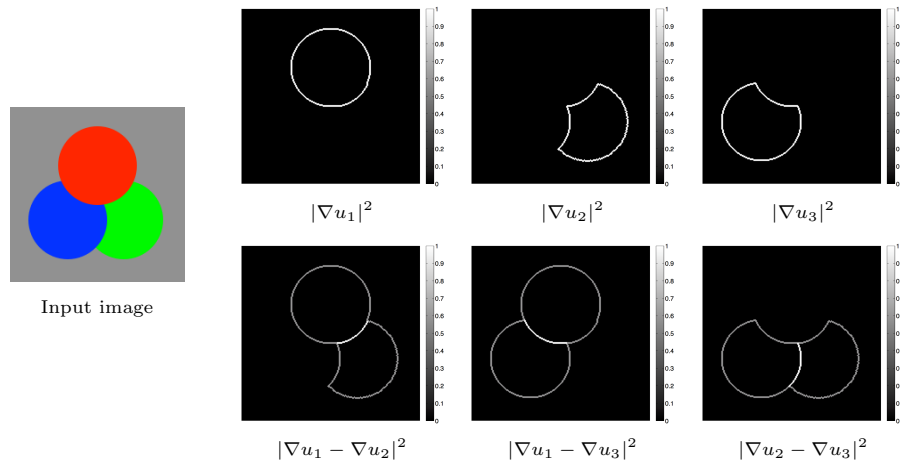
2. Vectorial total variation with channel coupling.

2.1. Coupling channels. For motivating our VTV scheme, we first consider a scenario where two 1-D spatially varying edges meet (Figure 1, First Row). Let us consider the signals u_1 and u_2 (Figure 1(a)) satisfying the following relation in terms of gradient magnitudes: $|\nabla u_1| > |\nabla u_2|$ (Figure 1(b)). In order to define a unique correlated edge we will penalize the expression $\nabla u_1 - \nabla u_2$ (Figure 1(c)) in terms of L^2 norm. Caselles et al [18] have shown that the geometry is constrained by the intensity channel along and the image level lines contain all the salient information. Then, the main idea of the coupling terms is to constrain the regularization near discontinuities by penalizing large derivations in gradients between different channels, therefore making them to look different to the normal Red, Green and Blue image gradient magnitudes. Second row in Figure 1 shows the behavior of the proposed coupling term over a synthetic image defined by the intersection of three different components. This synthetic image is composed of discs determined by pure chromatic colors of red, green, and blue. The gradient differences for each channel clearly differentiates from channel-wise gradients ($|\nabla u_i|^2$) since the coupling term is able to align edges over different channels. This in turn helps to keep the edge contrast during the further vectorial total variation smoothing process. Intersected edges over different colors are penalized by minimizing the terms $J(u_k, u_l)$. From the geometrical point of view, the proposed coupling term is well suited as it does not degrade the edge contrast quality of non-correlated channels, as shown in third row of Figure 1. In this case, a synthetic image where the green color circle does not intersect with the red and blue circles.

Gradient magnitudes and coupling term for two 1-D crossing edges



Behavior of the coupling channels for a color image with crossing edges between different channels



Behavior of the coupling channels for a color image with crossing edges over two color channels

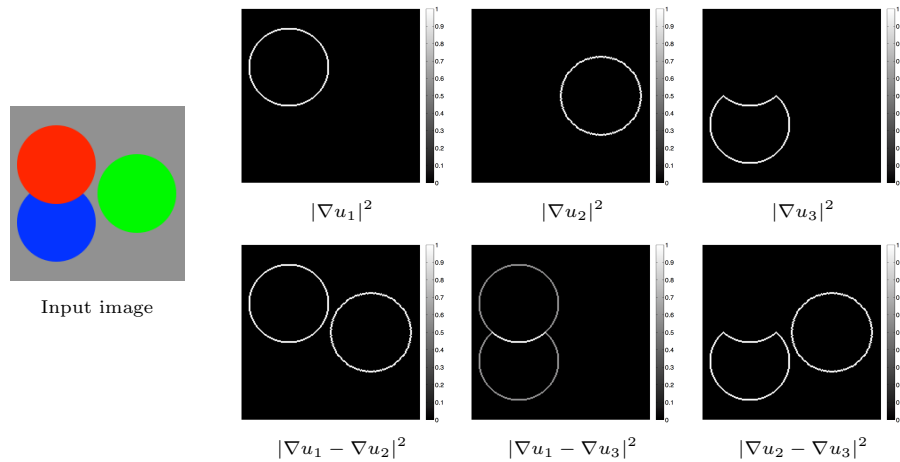


FIGURE 1. First Row: Displays the behaviour of the proposed model for two intersected signals. (a) 1-D edges (u_1, u_2) intersection. (b) Gradients $(\nabla u_1, \nabla u_2)$. (c) Coupling term $\nabla u_1 - \nabla u_2$. Second-Third Rows: Show the edge contrast quality due to the proposed coupling terms over two color images, and the differences with the gradient magnitude components.

In general, we propose the coupling term $J(u_k, u_l) = \int_{\Omega} |\nabla u_k - \nabla u_l|^2 dx$ for two given channels u_k and u_l . Then, the penalizing term extending the TV term for a multispectral image $\mathbf{u} : \Omega \rightarrow \mathbb{R}^N$, with $\Omega \subset \mathbb{R}^M$ a bounded open domain, is given by the following regularization term,

$$(6) \quad \sum_{k=1}^N \sum_{l=1}^N J(u_k, u_l) = \int_{\Omega} \sum_{k=1}^N \sum_{l=1}^N |\nabla u_k - \nabla u_l|^2 dx,$$

where it is expected to penalize large differences between gradient images [55, 51]. We further define the coupling- VTV of $\mathbf{u}$ to be the following,

$$(7) \quad VTV(\mathbf{u}) := \sum_{k=1}^N \int_{\Omega} |Du_k| + \alpha \int_{\Omega} \sum_{k=1}^N \sum_{l=1}^N |\nabla u_k - \nabla u_l|^2 dx,$$

where $\alpha \geq 0$ is a parameter for the cross-coupling term. Thus, the above model unifies the weighted TV regularization with coupling terms for multichannel image processing. We next provide detailed mathematical analysis of the coupled VTV method along with L^1 fidelity case. Note that the L^2 -norm is used in the proposed term to smooth common salient edges within the different channels as well as to distinguish between scaled edge features with nonuniform intensities. A general framework with $J(u_k, u_l) = \int_{\Omega} \psi(|\nabla u_k - \nabla u_l|) dx$ for a convex function ψ can also be accommodated within our framework studied here and we restrict ourselves to the canonical example of L^2 differences.

2.2. Analysis of the coupling VTV model. Under definition (7) and following the original TV theory in [4, 2] we define for $N \geq 1$ the space $BV(\Omega, \mathbb{R}^N)$ of vector valued functions as the set of functions $u \in L^1(\Omega, \mathbb{R}^N)$ such that $VTV(\mathbf{u}) < \infty$ according to equation in (7). The space $BV(\Omega, \mathbb{R}^N)$ endowed with the following norm:

$$(8) \quad \|\mathbf{u}\|_{L^1(\Omega, \mathbb{R}^N)} + \|\mathbf{u}\|_{BV(\Omega, \mathbb{R}^N)},$$

where $\|\mathbf{u}\|_{BV(\Omega, \mathbb{R}^N)} := VTV_1(\mathbf{u})$, is a Banach space. We recall further properties of VTV and a decomposition property which extends to our coupling with VTV .

- **Lower Semicontinuity:** For a given sequence $\{\mathbf{u}_n\}$ of functions in $BV(\Omega, \mathbb{R}^N)$ converging to a function $\mathbf{u} \in L^1(\Omega, \mathbb{R}^N)$

$$(9) \quad \|\mathbf{u}\|_{BV(\Omega, \mathbb{R}^N)} \leq \liminf_{n \rightarrow \infty} \|\mathbf{u}_n\|_{BV(\Omega, \mathbb{R}^N)}.$$

- **Compactness:** Every sequence $\{\mathbf{u}_n\} \in BV(\Omega, \mathbb{R}^N)$ such that $\|\mathbf{u}_n\|_{L^\infty(\Omega, \mathbb{R}^N)} \leq M$, for all $n \in \mathbb{N}$, admits a subsequence $\{\mathbf{u}_{n_k}\}$ converging in $L^1(\Omega, \mathbb{R}^N)$ to a function $\mathbf{u} \in BV(\Omega, \mathbb{R}^N)$.
- **Decomposition of $\|\mathbf{u}\|_{BV(\Omega, \mathbb{R}^N)}$:** The functional VTV can be decomposed as follows

$$(10) \quad \begin{aligned} VTV(\mathbf{u}) &= \sum_{k=1}^N \int_{\Omega} |\nabla u_k| dx + |D^c \mathbf{u}|(\Omega) + \int_{J_{\mathbf{u}} \cap \Omega} |\mathbf{u}^+ - \mathbf{u}^-| dH^{M-1} \\ &+ \alpha \int_{\Omega} \sum_{k=1}^N \sum_{l=1}^N |\nabla u_k - \nabla u_l|^2 dx, \end{aligned}$$

where $D^c \mathbf{u}$ is the Cantor part of $D\mathbf{u}$, $J_{\mathbf{u}}$ is the set of all jump points of $\mathbf{u}$, $\mathbf{u}_k^+$, $\mathbf{u}_k^-$ are the jump functions and H^{M-1} denotes the $(M-1)$ -dimensional Hausdorff measure in $\mathbb{R}^M$ with $\Omega \subset \mathbb{R}^M$.

Given $u \in BV(\Omega, \mathbb{R})$, the gTV term is defined as:

$$(11) \quad gTV(u) = \int_{\Omega} g |Du| \, dx := \sup_{p \in \Phi_g} \int_{\Omega} u(x) \operatorname{div} p(x) \, dx,$$

where $\Phi_g = \{p \in C_c^1(\Omega, \mathbb{R}^2) : |p(x)| \leq g \text{ for all } x \in \mathbb{R}^M\}$. We also adopt the notation [14]

$$(12) \quad \int_{\Omega} g |Du| = \int_{\Omega} g |\nabla u| \, dx.$$

We recall that the perimeter of a Borel measurable set $\Gamma \subset \Omega$ is defined by

$$(13) \quad \operatorname{Per}_g(\Sigma) := gTV(\mathbb{1}_{\Sigma}) = \int_{\Omega} g |\nabla \mathbb{1}_{\Sigma}| \, dx,$$

where $\mathbb{1}_{\Sigma}$ is the characteristic function of Σ . For a given image u and any $\xi \in \mathbb{R}$, we define the upper level set

$$(14) \quad \Gamma_u := \Gamma(u, \xi) = \{x \in \Omega : u(x) > \xi\}.$$

We should also use the co-area formula

$$(15) \quad gTV(u) = \int_{-\infty}^{\infty} \left(\int_{\gamma_{\xi}} g \, ds \right) d\xi = \int_{-\infty}^{\infty} \operatorname{Per}_g(\Gamma(u, \xi)) \, d\xi = \int_{-\infty}^{\infty} \int_{\gamma_{\xi}} g \, ds \, d\xi,$$

where γ_{ξ} is the boundary of the set $\Gamma(u, \xi)$.

The following result shows that our proposed VTV term decouples the level sets of the given multichannel image for each channel, making it a geometry problem for each channel's level set. The proof relies the perimeter (13) and its connection to the TV regularization.

Theorem 2.1. For $\mathbf{u} \in BV(\Omega, \mathbb{R}^N)$

$$(16) \quad \begin{aligned} VTV(\mathbf{u}) = & \int_{-\infty}^{+\infty} \sum_{k=1}^N \operatorname{Per}(\Gamma_{u_k}) \, d\xi_k \\ & + \alpha \int_{-\infty}^{+\infty} \int_0^1 \left(\sum_{i=1}^M \sum_{k=1}^N \sum_{l=1}^N \left(\int_{\Gamma_{\nabla^{x_i} u_k} \setminus \Gamma_{\nabla^{x_i} u_l}} (\xi_k - \mathbb{1}_{[0, \nabla^{x_i} u_l]}) \, dx \right. \right. \\ & + \left. \int_{\Gamma_{\nabla^{x_i} u_l} \setminus \Gamma_{\nabla^{x_i} u_k}} (\mathbb{1}_{[0, \nabla^{x_i} u_l]} - \xi_k) \, dx \right) d\mu_l \, d\xi_k \\ & + \left(\int_{\Gamma_{\nabla^{x_i} u_l} \setminus \Gamma_{\nabla^{x_i} u_k}} (\xi_l - \mathbb{1}_{[0, \nabla^{x_i} u_k]}) \, dx \right. \\ & \left. \left. + \int_{\Gamma_{\nabla^{x_i} u_k} \setminus \Gamma_{\nabla^{x_i} u_l}} (\mathbb{1}_{[0, \nabla^{x_i} u_k]} - \xi_l) \, dx \right) d\mu_k \, d\xi_l. \end{aligned}$$

Proof. From the standard total variation based co-area formula (15),

$$(17) \quad \int_{\Omega} |\nabla u_k| \, dx = \int_{-\infty}^{+\infty} \operatorname{Per}(\Gamma_{u_k}) \, d\xi_k.$$

Now, let $f_k = \nabla^{x_i} u_k$ and $f_l = \nabla^{x_i} u_l$. We have that

$$\begin{aligned}
 (18) \quad & \int_{\Omega} (\nabla^{x_i} u_k - \nabla^{x_i} u_l)^2 dx = \frac{1}{2} \int_{\Omega} (f_k - f_l)^2 dx + \frac{1}{2} \int_{\Omega} (f_l - f_k)^2 dx \\
 &= \frac{1}{2} \left(\int_{\{f_k > f_l\}} (f_k - f_l)^2 dx + \int_{\{f_l > f_k\}} (f_l - f_k)^2 dx \right) \\
 &+ \frac{1}{2} \left(\int_{\{f_l > f_k\}} (f_l - f_k)^2 dx + \int_{\{f_k > f_l\}} (f_k - f_l)^2 dx \right) \\
 &= \left(\int_{\{f_k > f_l\}} \int_{f_l}^{f_k} (\xi_k - f_l) d\xi_k dx + \int_{\{f_l > f_k\}} \int_{f_k}^{f_l} (f_l - \xi_k) d\xi_k dx \right) \\
 &+ \left(\int_{\{f_l > f_k\}} \int_{f_k}^{f_l} (\xi_l - f_k) d\xi_l dx + \int_{\{f_k > f_l\}} \int_{f_l}^{f_k} (f_k - \xi_l) d\xi_l dx \right) \\
 &= \left(\int_{\{f_k > f_l\}} \int_{-\infty}^{+\infty} (\xi_k - f_l) \mathbb{1}_{[f_l, f_k]} d\xi_k dx + \int_{\{f_l > f_k\}} \int_{-\infty}^{+\infty} (f_l - \xi_k) \mathbb{1}_{[f_k, f_l]} d\xi_k dx \right) \\
 &+ \left(\int_{\{f_l > f_k\}} \int_{-\infty}^{+\infty} (\xi_l - f_k) \mathbb{1}_{[f_k, f_l]} d\xi_l dx + \int_{\{f_k > f_l\}} \int_{-\infty}^{+\infty} (f_k - \xi_l) \mathbb{1}_{[f_l, f_k]} d\xi_l dx \right) \\
 &= \left(\int_{-\infty}^{+\infty} \int_{\Omega} (\xi_k - f_l) \mathbb{1}_{[f_l, f_k]} \mathbb{1}_{\Gamma_{f_k} \setminus \Gamma_{f_l}} dx d\xi_k + \int_{-\infty}^{+\infty} \int_{\Omega} (f_l - \xi_k) \mathbb{1}_{[f_k, f_l]} \mathbb{1}_{\Gamma_{f_l} \setminus \Gamma_{f_k}} dx d\xi_k \right) \\
 &+ \left(\int_{-\infty}^{+\infty} \int_{\Omega} (\xi_l - f_k) \mathbb{1}_{[f_k, f_l]} \mathbb{1}_{\Gamma_{f_l} \setminus \Gamma_{f_k}} dx d\xi_l + \int_{-\infty}^{+\infty} \int_{\Omega} (f_k - \xi_l) \mathbb{1}_{[f_l, f_k]} \mathbb{1}_{\Gamma_{f_k} \setminus \Gamma_{f_l}} dx d\xi_l \right).
 \end{aligned}$$

Let us observe that

$$(19) \quad \int_{\Omega} \mathbb{1}_{[\nabla^{x_i} u_k, \nabla^{x_i} u_l]} \mathbb{1}_{\Gamma_{\nabla^{x_i} u_k} \setminus \Gamma_{\nabla^{x_i} u_l}} dx = \int_{\Gamma_{\nabla^{x_i} u_k} \setminus \Gamma_{\nabla^{x_i} u_l}} dx,$$

and similarly

$$(20) \quad \int_{\Omega} \mathbb{1}_{[\nabla^{x_i} u_l, \nabla^{x_i} u_k]} \mathbb{1}_{\Gamma_{\nabla^{x_i} u_l} \setminus \Gamma_{\nabla^{x_i} u_k}} dx = \int_{\Gamma_{\nabla^{x_i} u_l} \setminus \Gamma_{\nabla^{x_i} u_k}} dx.$$

We also set

$$(21) \quad \nabla^{x_i} u_l = \int_0^1 \mathbb{1}_{[0, \nabla^{x_i} u_l]} d\mu_l, \quad \nabla^{x_i} u_k = \int_0^1 \mathbb{1}_{[0, \nabla^{x_i} u_k]} d\mu_k.$$

By using the identities in (18), (19), (20) and (21), we get

$$\begin{aligned}
 (22) \quad & \int_{\Omega} (\nabla^{x_i} u_k - \nabla^{x_i} u_l)^2 dx = \int_{-\infty}^{+\infty} \int_0^1 \left(\int_{\Gamma_{\nabla^{x_i} u_k} \setminus \Gamma_{\nabla^{x_i} u_l}} (\xi_k - \mathbb{1}_{[0, \nabla^{x_i} u_l]}) dx \right. \\
 & \quad \left. + \int_{\Gamma_{\nabla^{x_i} u_l} \setminus \Gamma_{\nabla^{x_i} u_k}} (\mathbb{1}_{[0, \nabla^{x_i} u_l]} - \xi_k) dx \right) d\mu_l d\xi_k \\
 & + \int_{-\infty}^{+\infty} \int_0^1 \left(\int_{\Gamma_{\nabla^{x_i} u_l} \setminus \Gamma_{\nabla^{x_i} u_k}} (\xi_l - \mathbb{1}_{[0, \nabla^{x_i} u_k]}) dx \right. \\
 & \quad \left. + \int_{\Gamma_{\nabla^{x_i} u_k} \setminus \Gamma_{\nabla^{x_i} u_l}} (\mathbb{1}_{[0, \nabla^{x_i} u_k]} - \xi_l) dx \right) d\mu_k d\xi_l.
 \end{aligned}$$

A combination of (17) and (22) leads to the desired result. $\square$

2.3. Vectorial TV with L^1 fidelity term. The $gTV + L^1$ model is one of the most influential variational PDE-based image decomposition models in image processing [22, 14]. This decomposition model separates a grayscale image in cartoon and texture components while preserving main features such as edges. For multi-spectral images, $\mathbf{f} : \Omega \rightarrow \mathbb{R}^N$, we extend the $gTV + L^1$ model to the vectorial case called $gVTV + L^1$ by considering the following variational problem:

$$(23) \quad \min_{\mathbf{u} \in BV(\Omega, \mathbb{R}^N)} \left\{ E_{gVTV}^1(\mathbf{u}) = \sum_{k=1}^N \left(\int_{\Omega} g_k(x) |\nabla u_k| dx + \lambda \int_{\Omega} |u_k - f_k| dx \right) + \alpha \int_{\Omega} \sum_{k=1}^N \sum_{l=1}^N |\nabla u_k - \nabla u_l|^2 dx \right\},$$

where $g_k(x) = (1 + |\nabla f_k(x)|^2)^{-1}$ is an edge indicator function. Note that in the edge indicator function the noisy image f is smoothed with a Gaussian kernel¹. When $g \equiv 1$ we refer to our model as $VTV + L^1$.

Remark 1. Let us note that (23) with L^2 image fidelity term is straightforward [54, 55],

$$(24) \quad \min_{\mathbf{u} \in BV(\Omega, \mathbb{R}^N)} \left\{ E_{gVTV}^2(\mathbf{u}) = \sum_{k=1}^N \left(\int_{\Omega} g_k(x) |\nabla u_k|^2 dx + \lambda \int_{\Omega} |u_k - f_k|^2 dx \right) + \alpha \int_{\Omega} \sum_{k=1}^N \sum_{l=1}^N |\nabla u_k - \nabla u_l|^2 dx \right\}.$$

By following [22] and the result from Theorem 2.1, we show that the $gVTV + L^1$ based functional (23) can also be expressed in terms of level sets. That is, the following theorem proves the equivalence of the energy functional in (23) based on VTV with L^1 data fidelity term and each channel level sets via the co-area formula.

Theorem 2.2. *Given a noise image $\mathbf{f}$ and a function $\mathbf{u} \in BV(\Omega, \mathbb{R}^N)$ we have*

$$(25) \quad \begin{aligned} E_{gVTV}^1(\mathbf{u}) &= \int_{-\infty}^{+\infty} \left(\sum_{k=1}^N \left(\text{Per}_{g_k}(\Gamma_{u_k}) + \lambda \left(\int_{\Gamma_{u_k} \setminus \Gamma_{f_k}} dx + \int_{\Gamma_{f_k} \setminus \Gamma_{u_k}} dx \right) \right) \right) d\xi_k \\ &+ \alpha \int_{-\infty}^{+\infty} \int_0^1 \left(\sum_{i=1}^M \sum_{k=1}^N \sum_{l=1}^N \left(\int_{\Gamma_{\nabla^{x_i} u_k} \setminus \Gamma_{\nabla^{x_i} u_l}} (\xi_k - \mathbb{1}_{[0, \nabla^{x_i} u_l]}) dx \right. \right. \\ &+ \left. \int_{\Gamma_{\nabla^{x_i} u_l} \setminus \Gamma_{\nabla^{x_i} u_k}} (\mathbb{1}_{[0, \nabla^{x_i} u_l]} - \xi_k) dx \right) d\mu_l d\xi_k \\ &+ \left(\int_{\Gamma_{\nabla^{x_i} u_l} \setminus \Gamma_{\nabla^{x_i} u_k}} (\xi_l - \mathbb{1}_{[0, \nabla^{x_i} u_k]}) dx \right. \\ &+ \left. \left. \int_{\Gamma_{\nabla^{x_i} u_k} \setminus \Gamma_{\nabla^{x_i} u_l}} (\mathbb{1}_{[0, \nabla^{x_i} u_k]} - \xi_l) dx \right) d\mu_k d\xi_l. \end{aligned}$$

¹We used a small amount of pre-smoothing to avoid smearing out possible edges, that is we used $\sigma = 1$ for the Gaussian kernel variance.

Proof. The result follows from the weighted co-area formula (15), equation (22) and by noticing that,

$$(26) \quad \int_{\Omega} |u_k - f_k| dx = \int_{-\infty}^{+\infty} \left(\int_{\Gamma_{u_k} \setminus \Gamma_{f_k}} dx + \int_{\Gamma_{f_k} \setminus \Gamma_{u_k}} dx \right) d\xi_k.$$

□

Following the level set formulation of the functional E_{gTV}^1 given in Theorem 2.2, and the convexity of the minimization problem (23), we can consider the following geometric sub-problems

$$(27) \quad \min_{\Gamma_k \subset \Omega} \left\{ \mathcal{G}(\Gamma_k, \Gamma_{f_k}) = \mathcal{G}_1(\Gamma_k) + \mathcal{G}_2(\Gamma_k, \Gamma_{f_k}) + \sum_{i=1}^M \sum_{l=1}^N \alpha \mathcal{G}_{3,l}(\nabla^{x_i}, \Gamma_k) \right\},$$

where

$$(28) \quad \mathcal{G}_1(\Gamma_{u_k}) = \text{Per}_{g_k}(\Gamma_{u_k})$$

$$(29) \quad \mathcal{G}_2(\Gamma_{u_k}, \Gamma_{f_k}) = \lambda \left(\int_{\Gamma_{u_k} \setminus \Gamma_{f_k}} dx + \int_{\Gamma_{f_k} \setminus \Gamma_{u_k}} dx \right)$$

and

$$(30) \quad \begin{aligned} \mathcal{G}_{3,l}(\nabla^{x_i}, \Gamma_{u_k}) &= \int_0^1 \left(\int_{\Gamma_{\nabla^{x_i} u_k} \setminus \Gamma_{\nabla^{x_i} u_l}} (\xi_k - \mathbb{1}_{[0, \nabla^{x_i} u_l]}) dx \right. \\ &\quad \left. + \int_{\Gamma_{\nabla^{x_i} u_l} \setminus \Gamma_{\nabla^{x_i} u_k}} (\mathbb{1}_{[0, \nabla^{x_i} u_l]} - \xi_k) dx \right) d\mu_l \\ &\quad + \int_0^1 \left(\int_{\Gamma_{\nabla^{x_i} u_l} \setminus \Gamma_{\nabla^{x_i} u_k}} (\xi_l - \mathbb{1}_{[0, \nabla^{x_i} u_k]}) dx \right. \\ &\quad \left. + \int_{\Gamma_{\nabla^{x_i} u_k} \setminus \Gamma_{\nabla^{x_i} u_l}} (\mathbb{1}_{[0, \nabla^{x_i} u_k]} - \xi_l) dx \right) d\mu_k. \end{aligned}$$

With the geometrical minimization formulation (27), we can prove the monotonicity property, which is in turn a generalization of the property called the sub-modularity of *Per* functional [1].

Theorem 2.3. (*Monotonicity Property*) Let $\xi_1, \xi_2 \in \mathbb{R}$ with $\xi_1 < \xi_2$. Given an image f_k and two minimizers Γ^1 and Γ^2 with finite parameter of the minimization problem (27) corresponding to $\Gamma_{f_k}^1 = \Gamma(f_k, \xi_1)$ and $\Gamma_{f_k}^2 = \Gamma(f_k, \xi_2)$ respectively, the following identities are satisfied

$$(31) \quad \mathcal{G}(\Gamma^2, \Gamma_{f_k}^1) - \mathcal{G}(\Gamma^1 \cap \Gamma^2, \Gamma_{f_k}^1) \geq \mathcal{G}(\Gamma^1 \cup \Gamma^2, \Gamma_{f_k}^1) - \mathcal{G}(\Gamma^1, \Gamma_{f_k}^1),$$

and

$$(32) \quad 0 \geq \mathcal{G}(\Gamma^2, \Gamma_{f_k}^2) - \mathcal{G}(\Gamma^1 \cap \Gamma^2, \Gamma_{f_k}^2) \geq \mathcal{G}(\Gamma^2, \Gamma_{f_k}^1) - \mathcal{G}(\Gamma^1 \cap \Gamma^2, \Gamma_{f_k}^1).$$

Proof. In order to prove (31) we observe the following identities,

$$(33) \quad (\Gamma_k^2 \setminus \Gamma_{f_k}^1) \setminus ((\Gamma_k^1 \cap \Gamma_k^2) \setminus \Gamma_{f_k}^1) = (\Gamma_k^2 \setminus \Gamma_k^1) \setminus \Gamma_{f_k}^1,$$

and

$$(34) \quad ((\Gamma_k^2 \cup \Gamma_k^1) \setminus \Gamma_{f_k}^1) \setminus (\Gamma_k^1 \setminus \Gamma_{f_k}^1) = (\Gamma_k^2 \setminus \Gamma_k^1) \setminus \Gamma_{f_k}^1.$$

As $(\Gamma_k^1 \cap \Gamma_k^2) \subset \Gamma_k^2$, we have

$$(35) \quad (\Gamma_{f_k}^1 \setminus (\Gamma_k^1 \cap \Gamma_k^2)) \setminus (\Gamma_{f_k}^1 \setminus \Gamma_k^2) = (\Gamma_{f_k}^1 \cap \Gamma_k^2) \setminus \Gamma_k^1,$$

and

$$(36) \quad (\Gamma_{f_k}^1 \setminus \Gamma_k^1) \setminus (\Gamma_{f_k}^1 \setminus (\Gamma_k^1 \cup \Gamma_k^2)) = (\Gamma_{f_k}^1 \cap \Gamma_k^2) \setminus \Gamma_k^1.$$

From the identities (33), (34), (35) and (36) combined all together, it follows

$$(37) \quad \mathcal{G}_2(\Gamma^2, \Gamma_{f_k}^1) - \mathcal{G}_2(\Gamma^1 \cap \Gamma^2, \Gamma_{f_k}^1) - \mathcal{G}_2(\Gamma^1 \cup \Gamma^2, \Gamma_{f_k}^1) + \mathcal{G}_2(\Gamma^1, \Gamma_{f_k}^1) = 0.$$

Similarly, the same previous reasoning can be applied to coupling term $\mathcal{G}_3$, and we get

$$(38) \quad \mathcal{G}_3(\nabla^{x_i}, \Gamma^2) - \mathcal{G}_3(\nabla^{x_i}, \Gamma^1 \cap \Gamma^2) - \mathcal{G}_3(\nabla^{x_i}, \Gamma^1 \cup \Gamma^2) + \mathcal{G}_3(\nabla^{x_i}, \Gamma^1) = 0.$$

By considering the sub modularity property of *Per* functional and the identities (37) and (38) that

$$(39) \quad \begin{aligned} & \mathcal{G}(\Gamma^2, \Gamma_{f_k}^1) - \mathcal{G}(\Gamma^1 \cap \Gamma^2, \Gamma_{f_k}^1) - \mathcal{G}(\Gamma^1 \cup \Gamma^2, \Gamma_{f_k}^1) + \mathcal{G}(\Gamma^1, \Gamma_{f_k}^1) = \\ & = \mathcal{G}_1(\Gamma^2) - \mathcal{G}_1(\Gamma^1 \cap \Gamma^2) - \mathcal{G}_1(\Gamma^1 \cup \Gamma^2) + \mathcal{G}_1(\Gamma^1) \geq 0. \end{aligned}$$

To prove the identity (32), we apply a similar reasoning as before to the term

$$(40) \quad \mathcal{G}(\Gamma^2, \Gamma_{f_k}^2) - \mathcal{G}(\Gamma^1 \cap \Gamma^2, \Gamma_{f_k}^2) - \mathcal{G}(\Gamma^2, \Gamma_{f_k}^1) + \mathcal{G}(\Gamma^1 \cap \Gamma^2, \Gamma_{f_k}^1).$$

□

The following theorem proves that the our $gVTV + L^1$ energy minimization problem (23) is well-posed. The proof involves the direct methods of calculus of variations.

Theorem 2.4. *Given an image $\mathbf{f} \in L^\infty(\Omega, \mathbb{R}^N)$, there exists a minimizer of (23) in $BV(\Omega, \mathbb{R}^N)$.*

Proof. For the proof of this theorem see Appendix A. □

The parameter λ in our minimization functional (23) is crucial in experimental results (see Section 4.2 which shows the importance of λ with respect to image denoising and different error metrics). The value of λ determines how much of important geometrical features are retained. We extend the work of Scherzer et al [60] to the vectorial case and study $G(\Omega)$ color space á la Y. Meyer [42] for obtaining important quantitative guidelines, see Appendix B.

3. Dual minimization based implementation. Implementation of VTV with coupling in (23) can be done in a variety of ways. The dual minimization technique for total variation introduced by Chambolle [20] is an attractive one as it is proven to be very efficient. We adapt dual minimization to our model and provide a convergence result for the alternative minimizations. Following [14] we use a splitting scheme by introducing an auxiliary variable $\mathbf{v}$,

$$(41) \quad \min_{\mathbf{u}, \mathbf{v}} \left\{ \tilde{E}_{gVTV}^1(\mathbf{u}, \mathbf{v}) \right. \\ \left. = \sum_{k=1}^N \left(\int_{\Omega} g_k(x) |\nabla u_k| dx + \frac{1}{2\theta_k} \int_{\Omega} (u_k + v_k - f_k)^2 dx + \lambda \int_{\Omega} |v_k| dx \right) \right\}$$

$$+ \alpha \int_{\Omega} \sum_{k=1}^N \sum_{l=1}^N |\nabla u_k - \nabla u_l|^2 dx \Big\},$$

or equivalently the subproblems

$$(42) \quad \min_{u_k, v_k} \left\{ \tilde{E}_k^1(u_k, v_k) = \int_{\Omega} g_k(x) |\nabla u_k| dx + \frac{1}{2\theta_k} \int_{\Omega} (u_k + v_k - f_k)^2 dx + \lambda \int_{\Omega} |v_k| dx \right. \\ \left. + \alpha \int_{\Omega} \sum_{l=1}^N |\nabla u_k - \nabla u_l|^2 dx \right\},$$

where θ is chosen to be small so that $f_k \sim u_k + v_k$, where (u_k, v_k) represents the cartoon and texture components respectively. Since the functional $\tilde{E}_k^1$ is convex, (42) can be split into two alternating minimization problems and a fast numerical scheme based on Chambolle's dual minimization [20] can be used.

The solution of the minimization problem (42) is given by the following alternating minimization steps:

1. Fixing v_k , the minimization problem in u_k is:

$$(43) \quad \min_{u_k} \left\{ \int_{\Omega} g_k(x) |\nabla u_k| dx + \frac{1}{2\theta_k} \|u_k + v_k - f_k\|_{L^2(\Omega, \mathbb{R})}^2 \right\}.$$

The solution of (43) is given by

$$(44) \quad u_k = f_k - v_k - \theta_k \operatorname{div} \mathbf{p}_k$$

where $\mathbf{p}_k = (p_{k1}, p_{k2})$ satisfies $g_k(x) \nabla(\theta_k \operatorname{div} \mathbf{p}_k - (f_k - v_k)) - |\nabla(\theta_k \operatorname{div} \mathbf{p}_k - (f_k - v_k))| \mathbf{p}_k = 0$, which is solved using a fixed point method: $\mathbf{p}_k^0 = 0$ and

$$\mathbf{p}_k^{n+1} = \frac{\mathbf{p}_k^n + \delta t \nabla(\operatorname{div}(\mathbf{p}_k^n) - (f_k - v_k)/\theta_k)}{1 + \frac{\delta t}{g_k(x)} |\nabla(\operatorname{div}(\mathbf{p}_k^n) - (f_k - v_k)/\theta_k)|}.$$

We remark that (43) can also be implemented using algorithms which solve TV minimization problems of the form $\min_u \{gTV(u) + \lambda \|u - f\|^2\}$, such as the split Bregman [37], fast gradient based (FISTA) [10], and the first-order primal-dual (due to Chambolle and Pock) [21] algorithms, by considering $u_k = u$ and $f = f_k - v_k$.

2. Fixing u_k , the minimization problem in v_k is:

$$(45) \quad \min_{v_k} \left\{ \frac{1}{2\theta} \|u_k + v_k - f_k\|_{L^2(\Omega, \mathbb{R})}^2 + \lambda \|v_k\|_{L^1(\Omega, \mathbb{R})} + \alpha \sum_{l=1}^N \|\nabla v_l - \nabla v_k\|_{L^2(\Omega, \mathbb{R})}^2 \right\},$$

and the solution is found as

$$v_k = 2\alpha \sum_{l=1}^N (\Delta v_l - \Delta v_k) + \begin{cases} f_k - u_k - \theta_k \lambda_k & \text{if } f_k - u_k \geq \theta_k \lambda_k, \\ f_k - u_k + \theta_k \lambda_k & \text{if } f_k - u_k \leq -\theta_k \lambda_k, \\ 0 & \text{if } |f_k - u_k| \leq \theta_k \lambda_k. \end{cases}$$

By implementing the algorithm, we get the sequences,

$$(\mathbf{u}_0, \mathbf{v}_0, \mathbf{u}_1, \mathbf{v}_1, \dots, \mathbf{u}_k, \mathbf{v}_k, \dots).$$

The following theorem provides a proof of convergence of these sequences based on the coerciveness of the functionals in (42). This result is supported by experimental results, see for example Figure 7 energy plots.

Theorem 3.1. *The sequence $(\mathbf{u}_k, \mathbf{v}_k)$ given by the minimization problem (42) converges to coordinatwise minimum on $BV(\Omega, \mathbb{R}^N) \times BV(\Omega, \mathbb{R}^N)$.*

Proof. For the proof of this theorem see Appendix A. $\square$

4. Experimental results. All the images are normalized to the range $[0, 1]$ and we implement the proposed $gVTV + L^1$ model (23) using the dual minimization implementation described in section (3). The algorithm is implemented in MATLAB2012a on a Mac laptop with Intel Core i7 CPU 2.3GHz, 8GB RAM CPU. The edge indicator function $g = 1/(1 + |\nabla I|^2)$ is used for all the results reported here. The $\alpha = 1/2$, $\delta t = 1/8$ and $\theta = 1$ are fixed, and the best results according to the $\max(|u^{n+1} - u^n|, |v^{n+1} - v^n|) \leq \epsilon$ are shown. A quantitative evaluation of the proposed decomposition model is performed for both our approach and other related methods in color images². The comparative performance analysis is done over the following databases: Mosaic art images³, Kodak Lossless True Color Image Suite⁴, Color Test Images Database⁵, USC-SIPI Image Database⁶ and Berkeley segmentation dataset (BSDS500)⁷. Our model takes < 10 sec for 50 iterations for 3 channels image of size 256×256 . To reduce the computational cost one can utilize faster alternative algorithmic implementations such as split Bregman [37], FISTA [10] or Chambolle-Pock first order primal-dual [21]. We consider here color image processing, that is, $N = 3$, for RGB channels and extension of the cross-channel term (6) to cases $N > 3$ is straightforward [51].

4.1. Decomposition and multiscale decomposition results. Image decomposition provides a separation of a given image into cartoon (piecewise constant) and texture components. For obtaining decomposition results we choose $\lambda = 1/2$ in our scheme, since with this value we are able to obtain nice cartoon (smooth) regions along with edge preservation. A better justification for a practical choice of the λ parameter is displayed in Tables 1-3 for the denoising results.

In Figure 2 we use two images from the Mosaic art database to first show the cartoon component result using the proposed VTV model with L^1 (23) and L^2 (24) fidelity terms. We also display the result incorporating the weight function g in the proposed model. By comparing the cartoon component of the proposed $gVTV + L^1$ against the $gVTV + L^2$ scheme for the input image Figure 2(a) we see that they behave qualitatively differently. For example, different shapes of the input image are preserved well in our scheme using the L^1 fidelity terms whereas the result using the implementation with the L^2 fidelity terms looks blurred in the final result. Canny edge map computed with all the three channels using the MATLAB's command `edge(Input, 'canny')` for both different results (implementing $gVTV + L^1$ and $gVTV + L^2$ schemes) indicates that chromatic edges are preserved coherently and the spatial location is intact. Figure 2(b) shows the difference of implementing the

²More results are available online: <http://sites.google.com/site/suryaiit/research/spectral/vtv>

³<http://www.cse.cuhk.edu.hk/leojia/projects/texturesep/>

⁴<http://r0k.us/graphics/kodak>

⁵<http://www.hlevkin.com/TestImages/classic.htm>

⁶<http://sipi.usc.edu>

⁷<http://www.eecs.berkeley.edu/Research/Projects/CS/vision/bsds/>

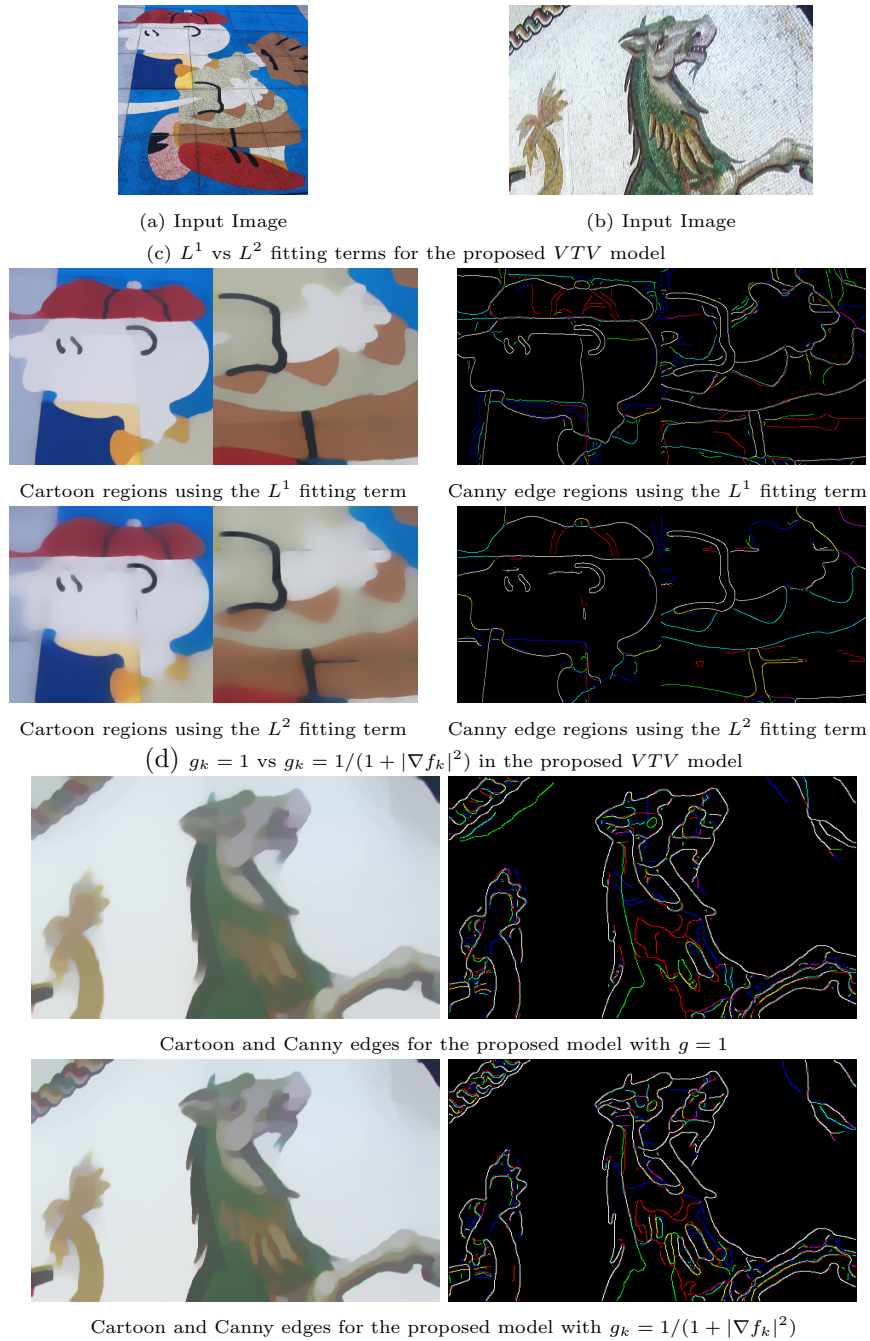


FIGURE 2. Cartoon component results showing the differences between using the L^1 and L^2 norms in the fitting term, as well as the differences of using a weight for the proposed regularized term (stopping parameter $\epsilon = 10^{-2}$). Results show better edge preservation using L^1 norm in the fitting term and $g_k = 1/(1 + |\nabla f_k|^2)$ with our proposed VTV method.

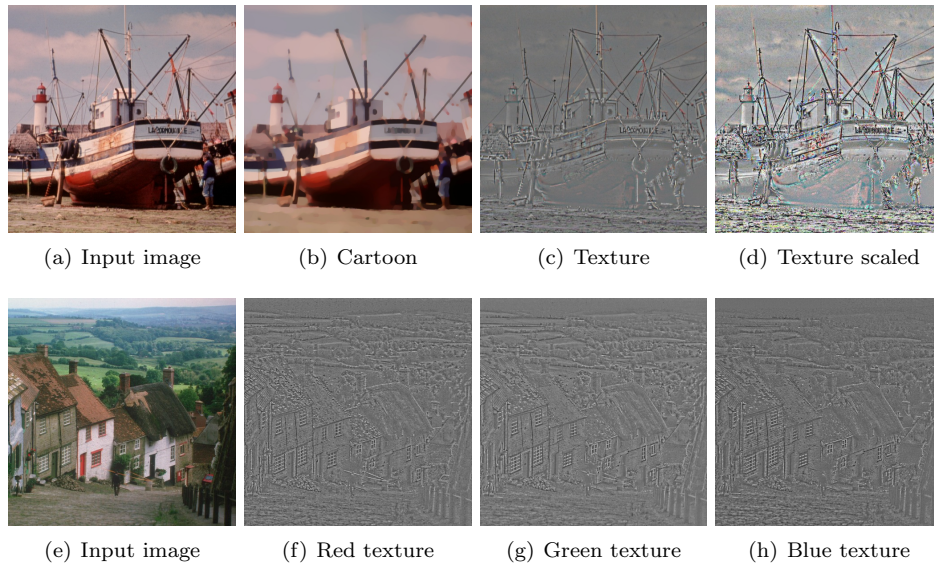


FIGURE 3. Different texture components for two color images (stopping time parameter $\epsilon = 10^{-2}$). First Row: For *Boats* image we show the (a) input image (b) cartoon (smooth), (c) texture component, and (d) scaled texture component. Second Row: For *Goldhill* image as input (e), we show the texture components from (f) Red (g) Green, and (h) Blue texture channels.

proposed VTV with coupling channels approach by using edge indicator functions $g_k = (1 + |\nabla f_k|)^{-1}$ and $g_k = 1$. As can be seen from the corresponding cartoon and edge map components, the result using edge indicator function contains smaller scale chromatic edges compared with setting $g_k = 1$ in our implementation.

Figure 3 shows different texture components for *Boats* and *Goldhill* images. Figure 3(c-d) shows the texture component given by the proposed scheme and its scaled version. The scaled texture component image is obtained by linearly transforming its range to $[0, 255]$ and making differences outside this range to saturate at 0 and 255. Figure 3(f-h) shows the texture component results in RGB channels, proving that the texture available in each channel is captured with our scheme.

Figure 4 shows the proposed color coupling channels for the cartoon components of close-up from *Motorcycle* & *Lighthouse* images. As can be seen, different channel edges are captured by the coupling terms (bottom row) and the final cartoon (smooth) components are devoid of small scale texture details. Moreover, the smoothing is edge preserving due to the weighted TV regularization. Figure 5 provides cartoon and texture components for a variety of natural color images using our scheme as stopping parameter increases from $\epsilon = 10^{-4}$ to $\epsilon = 10^{-2}$. We notice that when the stopping time decreases texture details are gradually lost in the cartoon components and they are added to the texture component.

We next show in Figure 6 a comparison of the cartoon and texture components for the *Lighthouse* image using different vectorial total variation terms along with L^1 data fidelity: $gVTV_1 + L^1$ and $gVTV_2 + L^1$ with our coupling VTV method. We can observe that cartoon component results look visually similar for all three

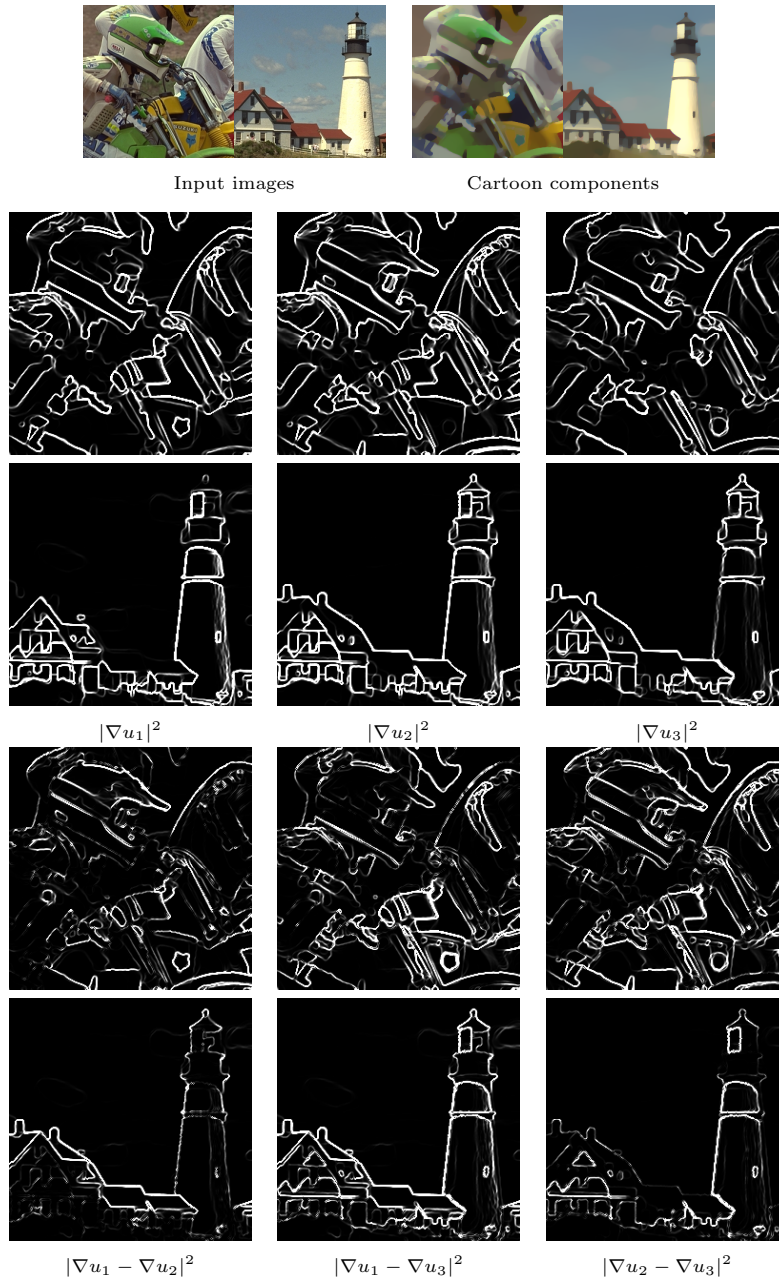
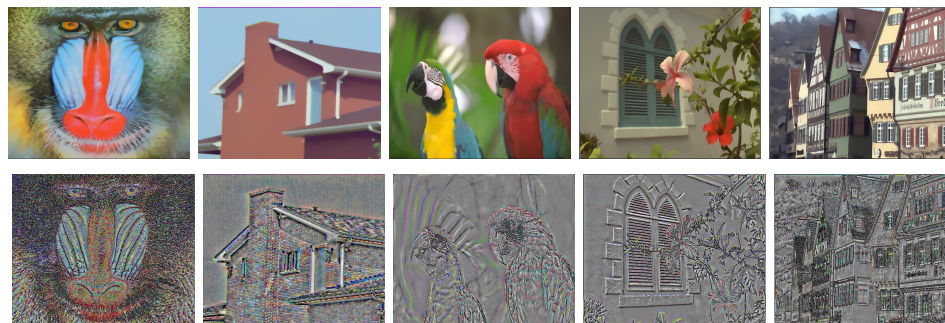


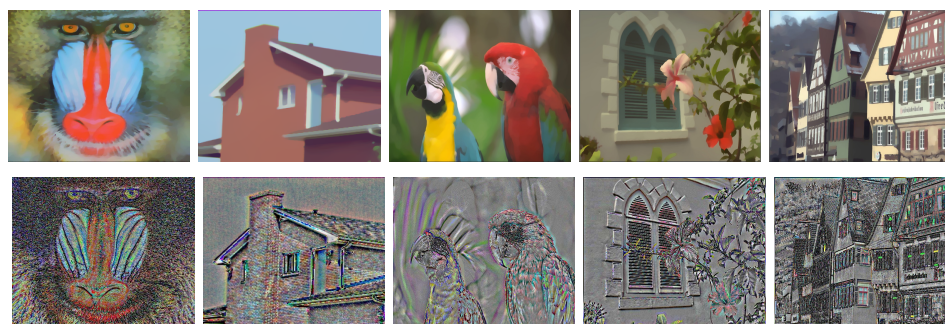
FIGURE 4. Cartoon component results using the proposed $gVTV + L^1$ model together with their Red, Green and Blue gradient magnitudes and the coupling terms (stopping parameter $\epsilon = 10^{-2}$). We show original inputs and their cartoon components, gradient magnitudes for Red (u_1), Green (u_2), Blue (u_3), and corresponding coupling terms.



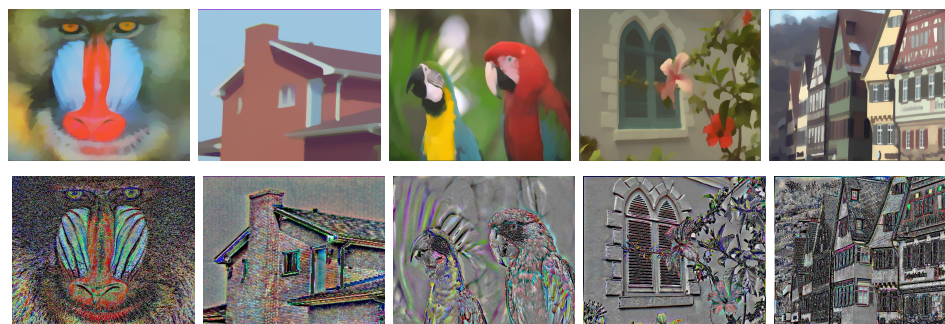
(a) Original (noise-free) images



(b) $\epsilon = 10^{-2}$



(c) $\epsilon = 10^{-3}$



(d) $\epsilon = 10^{-4}$

FIGURE 5. Cartoon and texture decomposition for different color images by implementing the proposed model and different choices of the stopping parameter ϵ . (a) Original color RGB images. (b)-(d) Decomposition results with $\epsilon = 10^{-2}$, $\epsilon = 10^{-3}$ and $\epsilon = 10^{-4}$, respectively.

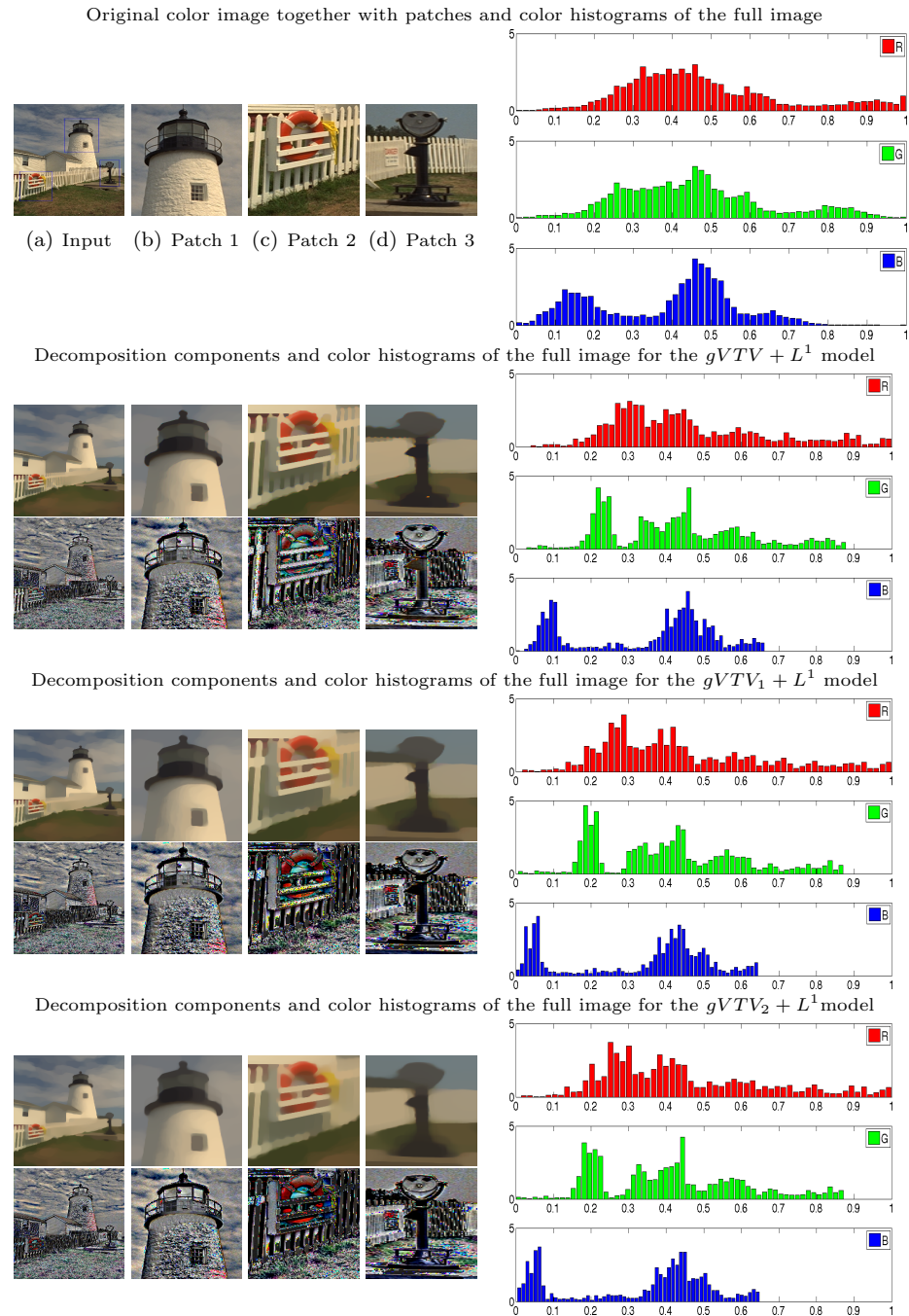


FIGURE 6. Cartoon and texture decomposition of a natural image together with the corresponding R, G, B histograms (stopping parameter $\epsilon = 10^{-4}$). Decomposition results of small patches are also displayed to highlight the cartoon and texture separation of different methods.

models, although our scheme retains the structures better. This can be observed in the color distribution of cartoon components using the corresponding histograms. It can also be noted from scaled texture components that edges are smoother for our proposed $gVTV + L^1$ scheme.

Figure 7 depicts cartoon-texture decomposition experiments using different natural images from Berkeley Segmentation data set (BSDS500), using our $gVTV + L^1$ model, compared with $gVTV_1 + L^1$ and $gVTV_2 + L^1$ schemes. A close look at the cartoon components reveals that our scheme obtains better preservation of small scale geometrical structures. For both the *Butterfly*, *Lady* images we show corresponding energy decrease against number of iterations of the dual minimization implementation for different VTV schemes (see Theorem 3.1).

Following Yin et al [73] and Tang et al [67] we can make the $gVTV + L^1$ (23) model with multiscale parameter $\lambda_k = \lambda_0/2^k$ in the L^1 -fidelity terms with $k = 1, 2, \dots$ and λ_0 a given value. Figure 8 displays 5 scales of our decomposition for the close-up of *Barbara* color image. Cartoon components (Figure 8(b-e)) retain salient edges during increasing steps and corresponding (scaled) texture components (Figure 8(g-j)) start to show progressive capture from small scale textures to big scales.

4.2. Denoising results. Note that the decomposition provides cartoon - piecewise cartoon component which is obtained using a weighted TV ($gVTV$) regularization in an edge preserving way. This allows us to apply our model to color image denoising as well, with u being the denoised image and v noise/texture part. To compare the schemes quantitatively in denoising we report our results using three standard error metrics: mean squared error (MSE), peak signal to noise ratio (PSNR) [4], and mean structural similarity measure (MSSIM) [71]. We also fix the stopping parameter $\epsilon = 10^{-2}$ for all the results.

1. MSE measures the average of the squares of the errors, *i.e.*, the difference between the estimator and what is estimated. Lower MSE value indicates better denoising result. For an estimated RGB image $\mathbf{u}$:

$$(46) \quad \text{MSE}(\mathbf{u}) := \sum_{x \in \Omega} (\mathbf{u} - \mathbf{u}_0)^2 / |\Omega|,$$

where $|\Omega|$ is the size of the image domain and $\mathbf{u}_0$ is the original noise-free color image.

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

$$(47) \quad \text{PSNR}(\mathbf{u}) := 10 * \log_{10} \left(\frac{1}{\sqrt{\text{MSE}}} \right) dB.$$

3. MSSIM index is in the range $[0, 1]$. MSSIM value near one implies the optimal denoising capability of the scheme [71] and is mean value of the SSIM metric. SSIM uses a combined product of luminance, contrast and structure comparison measures. The SSIM is calculated between two windows ω_1 and ω_2 of common size $N \times N$,

$$(48) \quad \text{SSIM}(\omega_1, \omega_2) = \frac{(2\mu_{\omega_1}\mu_{\omega_2} + c_1)(2\sigma_{\omega_1\omega_2} + c_2)}{(\mu_{\omega_1}^2 + \mu_{\omega_2}^2 + c_1)(\sigma_{\omega_1}^2 + \sigma_{\omega_2}^2 + c_2)},$$

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 [71] for more details.

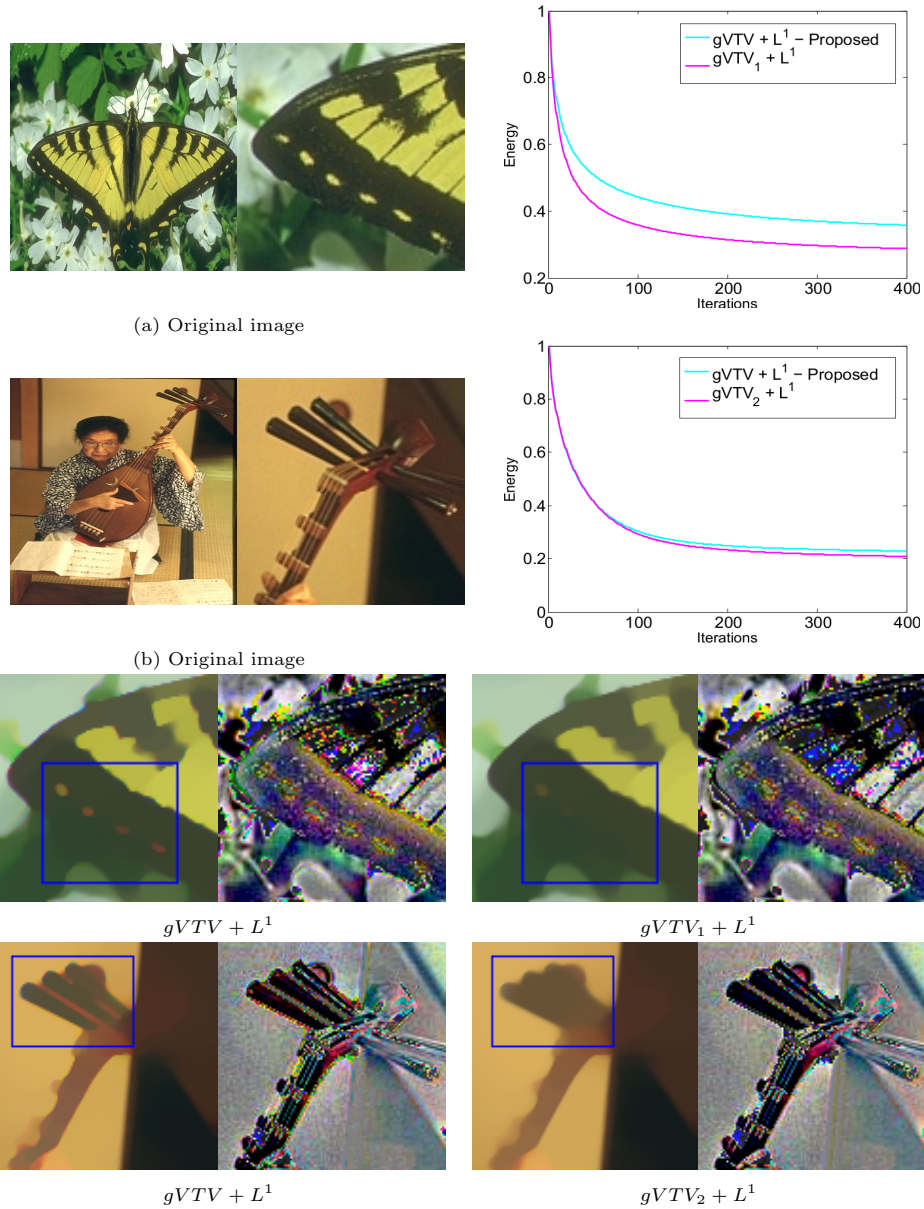


FIGURE 7. Cartoon + Texture decomposition results with corresponding energy versus iterations (stopping parameter $\epsilon = 10^{-4}$). Third Row: $gVTV + L^1$ proposed model and $gVTV_1 + L^1$ approach. Fourth Row: $gVTV + L^1$ proposed model and $gVTV_2 + L^1$ approach.

Since the main application of our proposed model is to study denoising based on chromatic edges through the different channels, we determine the MSE, PSNR and MSSIM error metrics on each channel and use the average as a final value for the

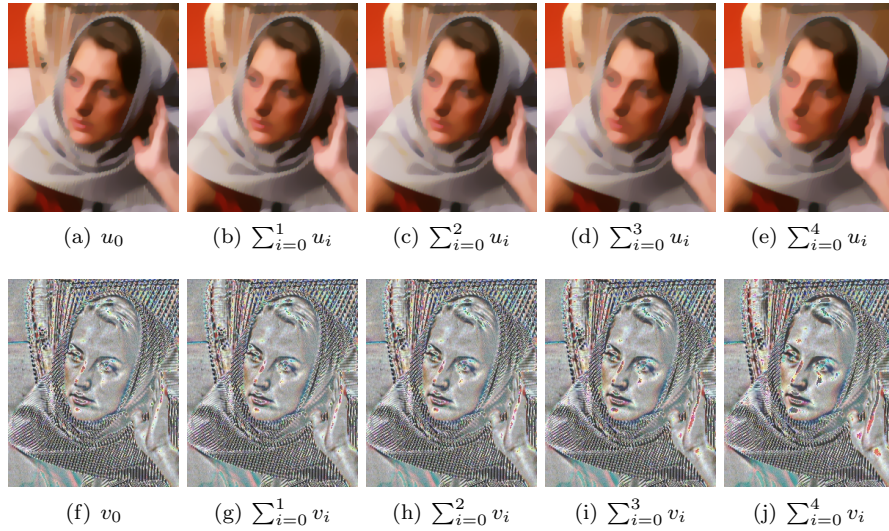


FIGURE 8. Multiscale cartoon and texture decomposition results for the close-up *Barbara* with 5 steps by implementing our proposed approach (stopping parameter $\epsilon = 10^{-4}$).

full Berkeley segmentation dataset (BSDS500) where different noise levels (with Gaussian noise of standard deviation $\sigma = 20$ and $\sigma = 30$) have been added.

Tables 1, 2 and 3 compare the MSE, PSNR and MSSIM performance measures for the $gVTV_1 + L^1$, $gVTV_2 + L^1$ and the proposed $gVTV + L^1$ schemes by using different λ values weighting the L^1 fidelity term averaged over the complete BSDS500, respectively. The idea behind choosing different λ values is to determine which value is able to capture important geometrical features (see Theorem B.5 in Appendix B related to the choice of λ). The proposed vectorial total variation approach with coupling channels scheme performed better than other vectorial total variation approaches over different λ values. We also see that MSE, PSNR and MSSIM value change over different λ values, with decreasing MSE values and increasing PSNR and MSSIM values. We note that $\lambda = 1$ is the best choice for multichannel edge preservation property followed by $\lambda = 1/2^k$ with $k \neq 0$ which tends to over-smooth the images. Note that the MSSIM is a better error metric than MSE and PSNR as it provides a quantitative way of measuring the structural similarity of denoised image against the original noise-free image. Different images for visualizing the improvements of the proposed $gVTV + L^1$ scheme are shown in Figure 9, and Figure 10 where we can see the effect of different VTV regularizations corresponding to λ choices.

Next we use the digitized color images taken from the USC-SIPI standard image dataset to compare the performance of the proposed scheme (23) (Our) together with the multiscale case (M-Our) along with following schemes: variational regularization (BKS, gradient descent with finite differences) [15], vector valued regularization with PDEs of (TD, spatial discretization with centered finite differences) [68], regularization of vector signals based on vectorial TV norm (BC, Chambolle's dual approach [20]) [13], Perona and Malik with a coupling term [55] (PS, finite differences), multiscale vectorial L^r - TV scheme for color images (DHC, using a unified

TABLE 1. Average MSE values for the $gVTV + L^1$ proposed model and the $gVTV_1 + L^1$, $gVTV_2 + L^1$ schemes (see equations (4) and (5)) over different Gaussian noise levels and λ parameters on the complete Berkeley segmentation database (BSDS500). Lower value indicates better result.

Method	Noise		MSE (Average)				
	20%	30%	$\lambda = 1/2^0$	$\lambda = 1/2^1$	$\lambda = 1/2^2$	$\lambda = 1/2^3$	$\lambda = 1/2^4$
$gVTV + L^1$			101.4588	109.3552	116.6560	122.3441	126.6188
$gVTV_1 + L^1$	✓		119.0799	135.8886	149.9856	160.6470	168.5642
$gVTV_2 + L^1$			117.7041	132.0977	145.2602	155.4844	163.1516
$gVTV + L^1$			156.6409	164.6322	173.2940	180.2798	185.5986
$gVTV_1 + L^1$		✓	173.2375	188.5963	203.7099	215.6114	224.5914
$gVTV_2 + L^1$			174.1293	185.4584	198.8797	109.9324	218.4031

TABLE 2. Average PSNR values for the $gVTV + L^1$ proposed model and the $gVTV_1 + L^1$, $gVTV_2 + L^1$ schemes (see equations (4) and (5)) over different Gaussian noise levels and λ parameters on the complete Berkeley segmentation database (BSDS500). Higher value indicates better result.

Method	Noise		PSNR (dB) (Average)				
	20%	30%	$\lambda = 1/2^0$	$\lambda = 1/2^1$	$\lambda = 1/2^2$	$\lambda = 1/2^3$	$\lambda = 1/2^4$
$gVTV + L^1$			29.0622	28.7418	28.4856	28.3011	28.1685
$gVTV_1 + L^1$	✓		28.3252	27.7356	27.2851	26.9635	26.7322
$gVTV_2 + L^1$			28.3311	27.8423	27.4229	27.1128	26.8868
$gVTV + L^1$			27.6266	27.4016	27.2122	27.0823	26.9892
$gVTV_1 + L^1$		✓	27.0860	26.5964	26.1756	25.8667	25.6422
$gVTV_2 + L^1$			27.0151	26.6478	26.2724	25.9839	25.7706

TABLE 3. Average MSSIM values for the $gVTV + L^1$ proposed model and the $gVTV_1 + L^1$, $gVTV_2 + L^1$ schemes (see equations (4) and (5)) over different Gaussian noise levels and λ parameters on the complete Berkeley segmentation database (BSDS500). Values close to 1 indicates better result.

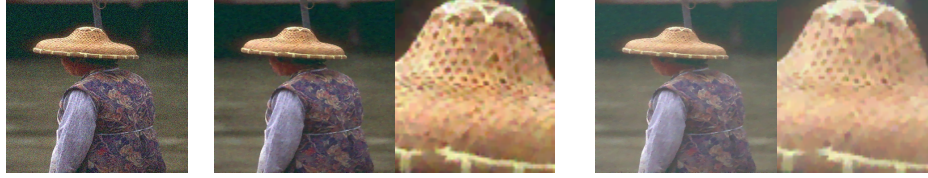
Method	Noise		MSSIM (Average)				
	20%	30%	$\lambda = 1/2^0$	$\lambda = 1/2^1$	$\lambda = 1/2^2$	$\lambda = 1/2^3$	$\lambda = 1/2^4$
$gVTV + L^1$			0.7991	0.7858	0.7740	0.7653	0.7589
$gVTV_1 + L^1$	✓		0.7721	0.7475	0.7286	0.7151	0.7053
$gVTV_2 + L^1$			0.7682	0.7527	0.7363	0.7237	0.7143
$gVTV + L^1$			0.7302	0.7206	0.7100	0.7019	0.6959
$gVTV_1 + L^1$		✓	0.7068	0.6904	0.6749	0.6634	0.6550
$gVTV_2 + L^1$			0.6954	0.6909	0.6790	0.6688	0.6610

Moreau-Yosida based primal-dual approach) [28] which uses the VTV_2 regularization, vectorial total variation derived from the generalized Jacobians from geometric measure (GSC, primal-dual method [21]) [36], multichannel gradients combined by parallel level sets (EA, central differencing scheme) [32], and separate gradient measurement of the luminance component and that of the chrominance (OY, primal-dual method [21]) [48]. In these results, we do not average the PSNR and the MSSIM error metrics over the RGB channels. Instead, we apply directly the PSNR error metric for multichannel images given by equation (46) and (47), respectively. In case of the MSSIM error metric we adapt it to color images by converting the color image into a grayscale (for example, using MATLAB's command `rgb2gray`) and then compute MSSIM for the converted grayscale image. Results are displayed in Table 4 for seventeen standard test images of size 256×256 and 512×512 . Our proposed vectorial total variation with coupling tends to perform better than

Noisy MSSIM	BKS [15]	TD [68]	BC [13]	PS [55]	DHC [28]	GSC [36]	EA [32]	OY [48]	Our	M-Our	BCM [16]	DFKE [26]
 0.520 [†]	27.24 0.724	29.56 0.781	29.42 0.832	27.99 0.784	30.66 0.862	29.75 0.840	30.03 0.860	31.10 0.875	31.05 0.879	31.37 0.887	31.04 0.881	32.23 0.916
 0.524 [†]	26.97 0.682	29.69 0.758	29.02 0.793	29.21 0.774	30.35 0.814	29.41 0.805	30.24 0.824	23.36 0.530	30.20 0.856	31.29 0.864	31.51 0.857	33.18 0.914
 0.381 [†]	29.55 0.865	32.59 0.896	31.75 0.917	23.25 0.890	33.51 0.924	31.97 0.919	33.52 0.923	34.83 0.934	33.99 0.916	34.25 0.926	34.69 0.932	36.62 0.952
 0.505 [†]	27.68 0.811	31.04 0.860	30.87 0.879	24.49 0.855	31.75 0.876	31.14 0.883	30.15 0.896	28.80 0.751	33.68 0.903	32.85 0.905	32.32 0.900	34.10 0.927
 0.487 [†]	26.85 0.730	30.18 0.793	27.43 0.836	24.81 0.784	31.53 0.821	27.67 0.839	29.85 0.854	31.46 0.856	31.78 0.858	32.53 0.858	32.32 0.877	33.78 0.906
 0.663 [†]	23.10 0.638	26.64 0.763	25.09 0.767	23.67 0.729	28.54 0.811	25.39 0.778	27.41 0.850	28.77 0.872	28.72 0.864	29.15 0.861	29.37 0.883	29.70 0.890
 0.393 [†]	29.13 0.903	32.72 0.935	31.22 0.951	27.97 0.949	32.84 0.967	31.64 0.951	31.93 0.948	33.56 0.925	32.51 0.964	33.40 0.969	34.09 0.963	35.52 0.980
 0.465 [†]	27.19 0.872	31.13 0.925	29.74 0.936	28.34 0.938	31.56 0.955	30.03 0.937	31.01 0.925	32.09 0.931	31.59 0.945	32.21 0.963	32.98 0.957	33.59 0.972
 0.764 [‡]	31.11 0.902	33.52 0.911	33.07 0.917	24.90 0.923	34.28 0.922	33.42 0.922	31.36 0.921	22.51 0.748	33.78 0.924	34.41 0.934	34.27 0.926	35.62 0.941
 0.798 [‡]	28.91 0.860	30.88 0.893	30.00 0.887	17.78 0.872	32.62 0.897	30.24 0.895	29.24 0.917	22.91 0.799	32.79 0.924	32.85 0.939	32.20 0.928	33.08 0.951
 0.910 [‡]	20.53 0.604	23.53 0.825	20.82 0.665	19.51 0.681	26.02 0.911	21.09 0.700	24.96 0.914	25.67 0.934	26.19 0.928	26.41 0.934	26.75 0.931	26.97 0.948
 0.808 [‡]	27.39 0.842	29.84 0.893	29.77 0.900	22.52 0.874	31.69 0.905	30.12 0.909	31.01 0.933	31.15 0.933	31.74 0.935	32.27 0.940	31.86 0.937	33.01 0.961
 0.785 [‡]	24.65 0.746	28.50 0.842	27.48 0.911	26.78 0.833	31.66 0.906	27.85 0.921	29.71 0.943	30.35 0.907	31.53 0.938	32.13 0.947	32.16 0.944	33.75 0.962
 0.845 [‡]	23.34 0.678	26.85 0.797	26.30 0.880	23.90 0.767	29.24 0.913	26.72 0.894	28.07 0.933	28.43 0.928	29.35 0.940	29.22 0.941	29.23 0.939	29.48 0.957
 0.809 [‡]	27.81 0.876	30.07 0.908	28.84 0.930	29.09 0.921	30.90 0.912	29.09 0.933	28.93 0.944	29.48 0.912	30.52 0.940	31.14 0.951	30.84 0.948	31.63 0.959
 0.869 [‡]	24.94 0.796	27.27 0.870	25.28 0.836	24.47 0.850	28.68 0.932	25.56 0.849	28.25 0.935	29.54 0.941	28.55 0.945	29.78 0.953	30.93 0.955	33.03 0.974

TABLE 4. PSNR (dB) & MSSIM values for results of various schemes with noise level $\sigma = 20$ for standard images of size $\dagger = 256 \times 256$ (Noisy PSNR = 25.27) and $\ddagger = 512 \times 512$ (Noisy PSNR = 25.83). Higher PSNR & MSSIM values indicate better denoising results. The best results for each image is indicated with **boldface** (within variational/PDE based methods only). The last two columns are nonlocal means [16] and color-BM3D [26] results.

$\lambda = 1/2^0$: Noisy Image (Left), Proposed (Center - PSNR=30.17), $gVTV_1 + L^1$ (Right - PSNR=29.23)



$\lambda = 1/2^1$: Noisy Image (Left), Proposed (Center - PSNR=28.51), $gVTV_1 + L^1$ (Right - PSNR=27.66)



$\lambda = 1/2^2$: Noisy Image (Left), Proposed (Center - PSNR=27.05), $gVTV_2 + L^1$ (Right - PSNR=25.23)



$\lambda = 1/2^3$: Noisy Image (Left), Proposed (Center - PSNR=29.12), $gVTV_1 + L^1$ (Right - PSNR=28.03)



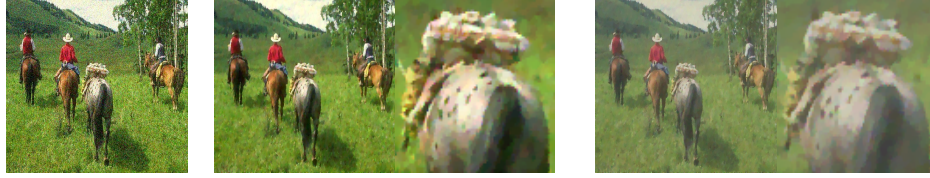
$\lambda = 1/2^4$: Noisy Image (Left), Proposed (Center - PSNR=31.46), $gVTV_2 + L^1$ (Right - PSNR=29.22)



FIGURE 9. Denoising results for different images corrupted with 20% of Gaussian noise. The values of $\lambda = 1/2^0, \lambda = 1/2^1, \lambda = 1/2^2, \lambda = 1/2^3, \lambda = 1/2^4$ are used for comparing the proposed $gVTV + L^1$ scheme with $gVTV_1 + L^1$ and $gVTV_2 + L^1$ models.

other variational/PDEs base diffusion methods by improving color enhancement properties along with edge preservation. Although, we are not comparing with other state-of-the art schemes (non variational-PDE methods) in terms of PSNR and MSSIM, the last two columns in Table 4 show the results for non-local means (BCM [16]) and block matching 3D filtering (DFKE [26]). These results evidence that the proposed model is able to remove noise while preserving edges (according to PSNR (dB) & MSSIM values, which are relatively close to these state of the art schemes).

$\lambda = 1/2^0$: Noisy Image (Left), Proposed (Center - PSNR=26.31), $gVTV_1 + L^1$ (Right - PSNR=25.90)



$\lambda = 1/2^1$: Noisy Image (Left), Proposed (Center - PSNR=26.39), $gVTV_2 + L^1$ (Right - PSNR=25.54)



$\lambda = 1/2^2$: Noisy Image (Left), Proposed (Center - PSNR=29.68), $gVTV_2 + L^1$ (Right - PSNR=28.03)



$\lambda = 1/2^3$: Noisy Image (Left), Proposed (Center - PSNR=26.66), $gVTV_1 + L^1$ (Right - PSNR=24.80)



$\lambda = 1/2^4$: Noisy Image (Left), Proposed (Center - PSNR=25.13), $gVTV_2 + L^1$ (Right - PSNR=23.64)

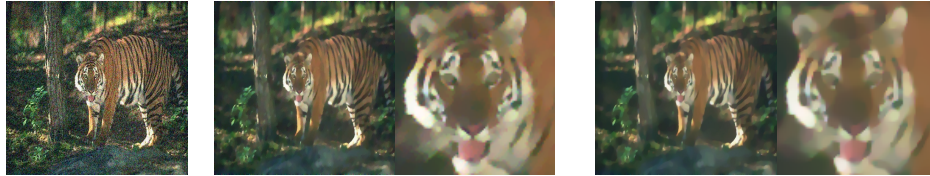


FIGURE 10. Denoising results for different images corrupted with 30% of Gaussian noise. The values of $\lambda = 1/2^0, \lambda = 1/2^1, \lambda = 1/2^2, \lambda = 1/2^3, \lambda = 1/2^4$ are used for comparing the proposed $gVTV + L^1$ scheme with $gVTV_1 + L^1$ and $gVTV_2 + L^1$ models.

Finally, Figure 11 shows a visual comparison of different methods on sub-images (size 330×360) of the noisy *Lena* test image with $\sigma = 20$ for the *TV* based schemes OY [48], GSC [36] and DHC [28]. Visual comparison of various regions show that color smearing artifacts occur in chromatic edges in other schemes due to channel mixing and diffusion transfer and also some staircasing artifacts in flat regions are visible. The method noise comparison shows that our multiscale method leaves no visible structure in the residue in comparison with other schemes. The parameters

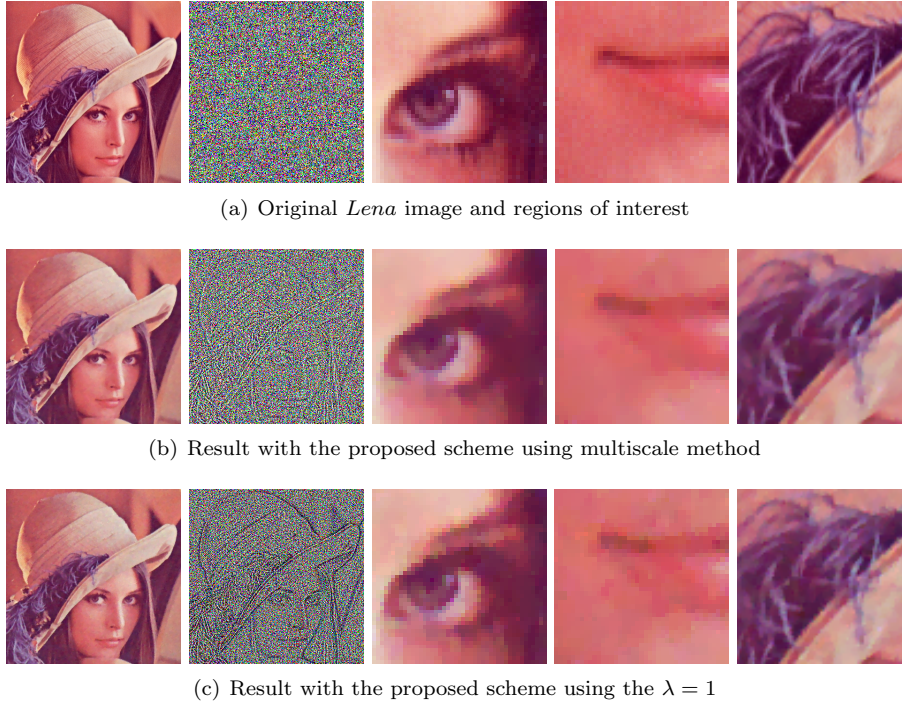


FIGURE 11. Denoising results for the close-up *Lena* image corrupted with Gaussian noise ($\sigma = 20$) together with neighbour regions. (a) Original image, input noise and different sub-regions of the original image. (b)-(c) results of our $gVTV + L^1$ (with multiscale) and $gVTV + L^1$ (with $\lambda = 1$) schemes respectively, displaying denoised images, residual noise/method noise components ($u - u_0$) and different sub-regions of the denoised images.

involved in all the comparison schemes are chosen based on maximum peak signal-to-noise ratio (PSNR).

5. Conclusion. In this paper, we study a multichannel image restoration scheme with total variation regularization and L^1 fidelity. By using matched gradients along with cross channels to align the edges our scheme performs multichannel smoothing and restoration. We provided a complete mathematical characterization of our vectorial total variation along with channel coupling model using its geometrical formulation, which provides some insights for the original model. Utilizing the notion of G -sets and G -values, we give a characterization of a solution of the proposed minimization problem in terms of the λ parameter, as well as the ability of such a parameter to decouple image features at different level sets. A fast dual minimization based implementation was used and experimental results indicate the proposed scheme provides better smoothing than other related models. Experimental results with other related color image denoising schemes indicate that the proposed approach consistently outperforms other color based restoration methods in terms of the perceptually guided PSNR (dB), MSSIM and MSE error metrics.



FIGURE 11. (Continued) Denoised image, method noise $u_0 - u$, and different sub-regions of the denoised image for schemes (d) OY [48], (e) GSC [36] and (f) DHC [28].

Appendix A. Existence of solution and convergence of the dual minimization based scheme.

Proof of Theorem 2.4. Let $C := \inf E_{gTV}^1(\mathbf{u})$, and let $\{\mathbf{u}^n\}_{n=1}^\infty \subseteq BV(\Omega, \mathbb{R}^N)$ be a minimizer sequence for the energy E_{gTV}^1 , i.e.,

$$E_{gTV}^1(\mathbf{u}^n) \xrightarrow{n \rightarrow \infty} C.$$

Then, for a given $k = 1, \dots, N$, it follows

$$\int_{\Omega} |\nabla u_k^n| dx \leq E_{gTV}^1(\mathbf{u}^n),$$

and consequently

$$\sup_{n \in \mathbb{N}} \int_{\Omega} |\nabla u_k^n| dx < \infty.$$

Moreover,

$$\|u_k^n\|_{L^1(\Omega)} \leq \int_{\Omega} dx < \infty.$$

Then, we have that $\{\mathbf{u}^n\}_{n=1}^\infty$ is bounded in $BV(\Omega, \mathbb{R}^N)$. By compactness property of the $BV(\Omega, \mathbb{R}^N)$ space there is a subsequence also denoted by $\{\mathbf{u}^n\}_{n=1}^\infty$, strongly convergent to an element $\mathbf{u}^* = (u_1^*, \dots, u_N^*) \in L^1(\Omega, \mathbb{R}^N)$. Therefore, it follows that

$\mathbf{u}^* \in BV(\Omega, \mathbb{R}^N)$ and

$$(49) \quad \int_{\Omega} |\nabla u_k^*| dx \leq \liminf_{n \rightarrow \infty} \int_{\Omega} |\nabla u_k^n| dx.$$

We also notice that the fidelity term and the coupling term in E_{gVTV}^1 are convex in $\mathbf{u}$, then both term are lower semicontinuous with respect to the $BV(\Omega, \mathbb{R}^N)$ weak*-topology and therefore

$$(50) \quad \int_{\Omega} |\nabla u_k^* - \nabla u_l^*|^2 dx \leq \liminf_{n \rightarrow \infty} \int_{\Omega} |\nabla u_k^n - \nabla u_l^n|^2 dx.$$

and

$$(51) \quad \int_{\Omega} |u_k^* - f_k| dx \leq \liminf_{n \rightarrow \infty} \int_{\Omega} |u_k^n - f_k| dx.$$

Finally, combining inequalities (49), (50) and (51) on a suitable sequence $\{\mathbf{u}^n\}_{n=1}^{\infty}$, we get

$$E_{gVTV}^1(\mathbf{u}^*) \leq \liminf_{n \rightarrow \infty} E_{gVTV}^1(\mathbf{u}^n) = C,$$

i.e., $\mathbf{u}^*$ is a minimizer of E_{gVTV}^1 . $\square$

Proof of Theorem 3.1. Since we are solving a splitting scheme defined by the sub-problems (43) and (45), the following inequalities hold

$$(52) \quad \tilde{E}_{gVTV}^1(\mathbf{u}^{k+1}, \mathbf{v}^{k+1}) \leq \tilde{E}_{gVTV}^1(\mathbf{u}^k, \mathbf{v}^{k+1}) \leq \tilde{E}_{gVTV}^1(\mathbf{u}^k, \mathbf{v}^k).$$

We also have that the sequence $\tilde{E}_{gVTV}^1(\mathbf{u}^k, \mathbf{v}^k)$ is non increasing and bounded by zero. Then, there exist $n \in \mathbb{R}$ such that $n = \lim_{k \rightarrow \infty} \tilde{E}_{gVTV}^1(\mathbf{u}^k, \mathbf{v}^k)$. Since $\tilde{E}_{gVTV}^1$ is coercive, we have that the sequence $(\mathbf{u}^k, \mathbf{v}^k)$ is bounded in $BV(\Omega, \mathbb{R}^N) \times BV(\Omega, \mathbb{R}^N)$ allowing us to extract a subsequence $(\mathbf{u}^{k_r}, \mathbf{v}^{k_r})$ converging to $(\mathbf{u}^*, \mathbf{v}^*)$, i.e., $(\mathbf{u}^{k_r}, \mathbf{v}^{k_r}) \xrightarrow{k_r \rightarrow \infty} (\mathbf{u}^*, \mathbf{v}^*)$ and by (44) we have

$$(53) \quad \mathbf{u}^{k_r} = \mathbf{v}^{k_r-1} - \theta \operatorname{div} \mathbf{p} \xrightarrow{k_r \rightarrow \infty} \mathbf{u}^* = \mathbf{v}^* - \theta \operatorname{div} \mathbf{p}.$$

Considering that

1. $\tilde{E}_{gVTV}^1(\mathbf{u}^{k_r+1}, \mathbf{v}^{k_r}) \leq \tilde{E}_{gVTV}^1(\mathbf{u}, \mathbf{v}^{k_r})$ for all $\mathbf{u} \in BV(\Omega, \mathbb{R}^N)$, and
2. $\tilde{E}_{gVTV}^1(\mathbf{u}^{k_r}, \mathbf{v}^{k_r}) \leq \tilde{E}_{gVTV}^1(\mathbf{u}^{k_r}, \mathbf{v})$ for all $\mathbf{v} \in BV(\Omega, \mathbb{R}^N)$,

it follows from (53) that for all $\mathbf{u} \in BV(\Omega, \mathbb{R}^N)$

$$(54) \quad \tilde{E}_{gVTV}^1(\mathbf{u}^*, \mathbf{v}^*) \leq \tilde{E}_{gVTV}^1(\mathbf{u}, \mathbf{v}^*)$$

and for all $\mathbf{v} \in BV(\Omega, \mathbb{R}^N)$

$$(55) \quad \tilde{E}_{gVTV}^1(\mathbf{u}^*, \mathbf{v}^*) \leq \tilde{E}_{gVTV}^1(\mathbf{u}^*, \mathbf{v})$$

allowing us to conclude that $(\mathbf{u}^*, \mathbf{v}^*)$ is a minimum with $\tilde{E}_{gVTV}^1(\mathbf{u}^*, \mathbf{v}^*) = n$. $\square$

Appendix B. The color $G(\Omega)$ space. Characterizing the minimizers of non-differentiable regularization functional has been analyzed by many authors. For grayscale images, Y. Meyer [42] defined the space $G(\mathbb{R}^2)$ to model texture using the ROF scheme [58]. Here, we extend the notion of G -sets and G -values studied in [60] to the vectorial case. For an observed color image $\mathbf{f}$, we show a relation between the solutions of the minimization problem (23) and the parameter λ which balances the $gVTV$ regularization term and L^1 fitting term. This will help us to correctly model the cartoon and texture decomposition, i.e., $\mathbf{f} \approx \mathbf{u} + \mathbf{v}$. A quantitative study using

MSE, PSNR (db) and MSSIM error metrics is given on different selections for the λ in denoising, see Section 4.2 and Tables 1, 2 and 3. See also [3, 70, 69, 49, 73, 31, 61] for related efforts to study TV models in this context.

The following definition is for vectorial G -values using measurable functions.

Definition B.1. Let $\vec{\Psi} = (\Psi_1, \dots, \Psi_N)$ where $\Psi_k : \Omega \rightarrow 2^{\mathbb{R}}$ and $k = 1, \dots, N$ be a set valued function and let

$$(56) \quad \vec{\Psi} := \{\psi : \Omega \rightarrow \mathbb{R}^N \text{ is a measurable and } \psi(x) \in \vec{\Psi}(x) \text{ almost everywhere}\} \neq \emptyset.$$

We define the G -value of $\vec{\Psi}$ as follows:

$$(57) \quad \begin{aligned} G(\vec{\Psi}) &:= \sup_{\{\mathbf{h} \in C_0^\infty(\Omega, \mathbb{R}^N) : gVTV(\mathbf{h})=1\}} - \sup_{\psi \in \vec{\Psi}} \sum_{k=1}^N \int_{\Omega} \psi_k h_k dx \\ &= \sup_{\{\mathbf{h} \in C_0^\infty(\Omega, \mathbb{R}^N) : gVTV(\mathbf{h})=1\}} \inf_{\psi \in \vec{\Psi}} \sum_{k=1}^N \int_{\Omega} \psi_k h_k dx. \end{aligned}$$

Remark 2. There is no distinction between the set $\vec{\Psi}$ and the function $\vec{\Psi}$. In case $\vec{\Psi}$ is a single value and measurable then $G(\vec{\Psi})$ is the G -norm of $\vec{\Psi}$ as defined in [31].

Definition B.2. The G_λ -values of Ψ are defined as

$$(58) \quad G_\lambda(\vec{\Psi}) := \sup_{\{\mathbf{h} \in C_0^\infty(\Omega, \mathbb{R}^N) : \|\mathbf{h}\|_\lambda=1\}} - \sup_{\psi \in \vec{\Psi}} \int_{\Omega} \sum_{k=1}^N \psi_k h_k dx.$$

We next derive the definition of G -sets for the vectorial case. Let $\partial|g|$ denote the set-valued subdifferential of $\|g\|_1$, i.e.,

$$(59) \quad \partial|g_i| = \begin{cases} \text{sgn}(g_i) & \text{if } g_i \neq 0, \\ [-1, 1] & \text{if } g_i = 0. \end{cases}$$

An important G -value property of $\partial|g|$ is given by the following lemma which relates the parameter λ and our $gVTV$ energy.

Lemma B.3. For $g \in L^1(\Omega, \mathbb{R}^N)$, $G(\partial|g|) \leq \lambda$ if and only if

$$(60) \quad \lambda \sum_{k=1}^N \left(\left| \int_{\{g_k \neq 0\}} \text{sgn}(g_k) h_k dx \right| - \int_{\{g_k=0\}} |h_k| dx \right) \leq gVTV(\mathbf{h})$$

for all $\mathbf{h} \in BV(\Omega, \mathbb{R}^N)$. Moreover

$$(61) \quad G(\partial|g|) = \sup_{\{\mathbf{h} \in BV(\Omega, \mathbb{R}^N) : gVTV(\mathbf{h})=1\}} - \sup_{\psi \in \vec{\Psi}} \int_{\Omega} \sum_{k=1}^N \psi_k h_k dx.$$

Definition B.4. Let $\mathbf{f} \in L^1(\Omega, \mathbb{R}^N)$ and $\mathbf{u} \in BV(\Omega, \mathbb{R}^N)$. Assume that for every $\mathbf{h} \in BV(\Omega, \mathbb{R}^N)$

$$(62) \quad \begin{aligned} \lambda \sum_{k=1}^N \left(- \int_{\{u_k \neq f_k\}} \text{sgn}(u_k - f_k) h_k dx - \int_{\{u_k=f_k\}} |h_k| dx \right) \\ \leq gVTV(\mathbf{u} + \mathbf{h}) - gVTV(\mathbf{u}). \end{aligned}$$

We define the G -set for the vectorial case as

$$(63) \quad \mathcal{G}_{\mathbf{u}}(\partial|\mathbf{u} - \mathbf{f}|) := \{\lambda \in [0, \infty] : \lambda \text{ satisfies (62)}\}.$$

Note that for $\lambda \in \mathcal{G}_u(\partial|\mathbf{u} - \mathbf{f}|)$ it follows that for every $\mathbf{h} \in BV(\Omega, \mathbb{R}^N)$

$$(64) \quad \lambda \sum_{k=1}^N \left(\left| \int_{\{u_k \neq f_k\}} \operatorname{sgn}(u_k - f_k) h_k \, dx \right| - \int_{\{u_k = f_k\}} |h_k| \, dx \right) \leq gVTV(\mathbf{u} + \mathbf{h}) - gVTV(\mathbf{u}),$$

and thus

$$(65) \quad G(\partial(\mathbf{u} - \mathbf{f})) \leq \frac{1}{\lambda}.$$

An equivalence relation between finding a solution of the minimization problem (23) and choosing an appropriate parameter λ for the $gVTV + L^1$ model is stated in the next theorem.

Theorem B.5. *For $\lambda > 0$, $\mathbf{u}_\lambda$ is a minimizer of the vectorial $gVTV + L^1$ model (23), if only if λ satisfies*

$$(66) \quad \lambda \sum_{k=1}^N \left(- \int_{\Omega} \operatorname{sgn}(u_k^\lambda - f_k) h_k \, dx - \int_{\Omega} |h_k| \, dx \right) \leq gVTV(\mathbf{u}^\lambda + \mathbf{h}) - gVTV(\mathbf{u}^\lambda),$$

for all $\mathbf{h} \in BV(\Omega, \mathbb{R}^N)$, i.e., $\lambda \in \mathcal{G}_u(\partial|\mathbf{u} - \mathbf{f}|)$.

Proof. If $\mathbf{u}^\lambda = (u_1^\lambda, \dots, u_N^\lambda)$ is a minimizer of (23), we have that for any $\mathbf{h} \in BV(\Omega, \mathbb{R}^N)$ and $\epsilon = (\epsilon_1, \dots, \epsilon_N)$ with $\epsilon_k > 0$ for every $k = 1, \dots, N$ the following inequalities

$$(67) \quad \begin{aligned} & gTV(u_k^\lambda) + \alpha \sum_{l=1}^N J_{k,l}(u_k^\lambda, u_l^\lambda) + \lambda \int_{\Omega} |u_k^\lambda - f_k| \, dx \\ & \leq gTV(u_k^\lambda + \epsilon_k h_k) + \alpha \sum_{l=1}^N J_{k,l}(u_k^\lambda + \epsilon_k h_k, u_l^\lambda + \epsilon_l h_l) + \lambda \int_{\Omega} |u_k^\lambda + \epsilon_k h_k - f_k| \, dx \\ & \leq gTV(u_k^\lambda + \epsilon_k h_k) + \alpha \sum_{l=1}^N J_{k,l}(u_k^\lambda + \epsilon_k h_k, u_l^\lambda + \epsilon_l h_l) + \lambda \int_{\Omega} |u_k^\lambda - f_k| \, dx \\ & \quad + \epsilon_k \lambda \int_{\{u_k^\lambda \neq f_k\}} \operatorname{sgn}(u_k^\lambda - f_k) \, dx + \lambda \eta(\epsilon_k, u_k^\lambda, h_k), \end{aligned}$$

where $\eta(\epsilon_k, u_k^\lambda, h_k) = o(\epsilon_k)$. This implies

$$(68) \quad \begin{aligned} & \lambda \left(- \int_{\{u_k^\lambda \neq f_k\}} \operatorname{sgn}(u_k^\lambda - f_k) h_k \, dx - \int_{\{u_k^\lambda = f_k\}} |h_k| \, dx \right) \\ & \leq \frac{gTV(u_k^\lambda + \epsilon_k h_k) + \alpha \sum_{l=1}^N J_{k,l}(u_k^\lambda + \epsilon_k h_k, u_l^\lambda + \epsilon_l h_l) - gTV(u_k^\lambda) - \sum_{l=1}^N J_{k,j}(u_k^\lambda, u_l^\lambda)}{\epsilon_k} \\ & \quad + \lambda \left(\eta(\epsilon_k, u_k^\lambda, h_k) - \int_{\{u_k^\lambda = f_k\}} |h_k| \, dx \right). \end{aligned}$$

The convexity of $gTV(u_k^\lambda) + \alpha \sum_{l=1}^N J_{k,l}(u_k^\lambda, u_l^\lambda)$ implies

$$(69) \quad \frac{gTV(u_k^\lambda + \epsilon_k h_k) + \alpha \sum_{l=1}^N J_{k,l}(u_k^\lambda + \epsilon_k h_k, u_l^\lambda + \epsilon_l h_l) - (gTV(u_k^\lambda) + \alpha \sum_{l=1}^N J_{k,l}(u_k^\lambda, u_l^\lambda))}{\epsilon_k} \\ \leq gTV(u_k^\lambda + h_k) + \alpha \sum_{l=1}^N J_{k,l}(u_k^\lambda + h_k, u_l^\lambda + \epsilon_l h_l) - gTV(u_k^\lambda) + \alpha \sum_{l=1}^N J_{k,l}(u_k^\lambda, u_l^\lambda)$$

for all $0 < \epsilon_k < 1$. Since $\eta(\epsilon_k, u_k^\lambda, h_k) \rightarrow \int_{\{u_k=f_k\}} |h_k| dx$, we found by using (69) and letting $\epsilon \rightarrow 0$ that

$$(70) \quad \lambda \sum_{k=1}^N \left(- \int_{\Omega} \operatorname{sgn}(u_k^\lambda - f_k) h_k dx - \int_{\Omega} |h_k| dx \right) \leq gVTV(\mathbf{u}^\lambda + \mathbf{h}) - gVTV(\mathbf{u}^\lambda),$$

Conversely by (66) and the convexity of $|\cdot|$ it follows that

$$(71) \quad \lambda \sum_{k=1}^N \int_{\Omega} |u_k^\lambda + h_k - f_k| dx + gVTV(\mathbf{u}^\lambda + \mathbf{h}) \\ \geq gVTV(\mathbf{u}^\lambda) \\ + \lambda \sum_{k=1}^N \left(\int_{\Omega} |u_k^\lambda - f_k| dx + \int_{\{u_k^\lambda \neq f_k\}} \operatorname{sgn}(u_k^\lambda - f_k) h_k dx + \int_{\{u_k^\lambda = f_k\}} |h_k| dx \right) \\ + gVTV(\mathbf{u}^\lambda + \mathbf{h}) - gVTV(\mathbf{u}^\lambda) \\ \geq \lambda \sum_{k=1}^N \int_{\Omega} |u_k^\lambda - f_k| dx + gVTV(\mathbf{u}^\lambda).$$

□

The proposed $gVTV + L^1$ model is expected to discriminate between different features in an image during regularizing while preserving strong geometrical features such as edges and corners. For this purpose, it is essential to get bounds on the values for the parameter λ that can provide a guideline in experiments. In the next theorem, it is shown that for some particular selections of the parameter λ , we can separate different level sets of an image at multiple scales.

Theorem B.6. *Let $\mathcal{N} = \mathcal{N}_1 \cup \mathcal{N}_2 \subset \Omega$ with $\mathcal{N}_1$ and $\mathcal{N}_2$ disjoint, and let $\mathbf{f} : \Omega \rightarrow \mathbb{R}^N$ such that $\mathbf{f}_{\mathcal{N}} = \mathbf{f}_{\mathcal{N}_1} \mathbb{1}_{\mathcal{N}_1} + \mathbf{f}_{\mathcal{N}_2} \mathbb{1}_{\mathcal{N}_2}$ (being $\mathbf{f}_{\mathcal{N}_i}$ the restriction of $\mathbf{f}$ in $\mathcal{N}_i$). Given a parameter λ satisfying*

$$(72) \quad \frac{1}{G(\partial|\mathbf{f}_{\mathcal{N}_2}|)} < \lambda < \frac{1}{G(\partial|\mathbf{f}_{\mathcal{N}_1}|)}$$

with $G(\partial|\mathbf{f}_{\mathcal{N}_2}|) > G(\partial|\mathbf{f}_{\mathcal{N}_1}|)$, if $\mathbf{u}$ is a minimizer of the $gVTV + L^1$ model (23), it follows $\mathbf{u}_{\mathcal{N}_1} = 0$ and $\mathbf{u}_{\mathcal{N}_2} \neq 0$.

Proof. Let $\epsilon = (\epsilon_1, \dots, \epsilon_N)$ with $0 < \epsilon_k < 1$ for every $k = 1, \dots, N$, and $\mathbf{h} \in BV(\mathcal{N}, \mathbb{R}^N)$. By contradiction, we prove that $\mathbf{u}_{\mathcal{N}_2} \neq 0$ i.e., we suppose $\mathbf{u}_{\mathcal{N}_2} = 0$. From the convexity of $gTV(\cdot)$ and $|\cdot|$, the following inequalities hold

$$(73) \quad \epsilon_k \left(gTV(h_k) + \sum_{l=1}^N J_{k,l}(h_k, h_l) \right) \geq gTV((u_k)_{\mathcal{N} \setminus \mathcal{N}_2} + \epsilon_k h_k)$$

$$\begin{aligned}
& + \sum_{l=1}^N J_{k,l}((u_k)_{\mathcal{N} \setminus \mathcal{N}_2} + \epsilon_k h_k, (u_l)_{\mathcal{N} \setminus \mathcal{N}_2} + \epsilon_l h_l) \\
& - \left(gTV((u_k)_{\mathcal{N} \setminus \mathcal{N}_2}) + \sum_{l=1}^N J_{k,l}((u_k)_{\mathcal{N} \setminus \mathcal{N}_2} + u_l) \right)
\end{aligned}$$

and

$$\begin{aligned}
(74) \quad & \int_{\mathcal{N}_2} |(f_k)_{\mathcal{N}_2}| dx - \int_{\mathcal{N}_2} |(f_k)_{\mathcal{N}_2} - \epsilon_k h_k| dx \\
& \geq \epsilon_k \left(\left| \int_{\{(f_k)_{\mathcal{N}_2} \neq \epsilon_k h_k\}} \operatorname{sgn}((f_k)_{\mathcal{N}_2} - \epsilon_k h_k) h_k dx \right| - \int_{\{(f_k)_{\mathcal{N}_2} = \epsilon_k h_k\}} h_k dx \right).
\end{aligned}$$

We also have from the minimizer $\mathbf{u}$ that

$$\begin{aligned}
(75) \quad & gVTV(\mathbf{u}_{\mathcal{N} \setminus \mathcal{N}_2}) + \lambda \sum_{k=1}^N \left(\int_{\mathcal{N}_2} |(f_k)_{\mathcal{N}_2}| dx + \int_{\mathcal{N} \setminus \mathcal{N}_2} |(u_k)_{\mathcal{N} \setminus \mathcal{N}_2} - (f_k)_{\mathcal{N} \setminus \mathcal{N}_2}| dx \right) \\
& \leq gVTV(\mathbf{u}_{\mathcal{N} \setminus \mathcal{N}_2} + \epsilon \mathbf{h}) + \lambda \sum_{k=1}^N \left(\int_{\mathcal{N}_2} |(f_k)_{\mathcal{N}_2} - \epsilon_k h_k| dx + \int_{\mathcal{N} \setminus \mathcal{N}_2} |(u_k)_{\mathcal{N} \setminus \mathcal{N}_2} - (f_k)_{\mathcal{N} \setminus \mathcal{N}_2}| dx \right).
\end{aligned}$$

By properly combining identities (73), (74) and (75), we get

$$\begin{aligned}
(76) \quad & \sum_{k=1}^N \epsilon_k \left(gTV(h_k) + \sum_{i=1}^N J_{k,i}(h_k, h_i) \right) \\
& \geq gVTV(\mathbf{u}_{\mathcal{N} \setminus \mathcal{N}_2} + \epsilon \mathbf{h}) - gVTV(\mathbf{u}_{\mathcal{N} \setminus \mathcal{N}_2}) \\
& \geq \lambda \sum_{k=1}^N \left(\int_{\mathcal{N}_2} |(f_k)_{\mathcal{N}_2}| dx - \int_{\mathcal{N}_2} |(f_k)_{\mathcal{N}_2} - \epsilon_k h_k| dx \right) \\
& \geq \lambda \sum_{k=1}^N \epsilon_k \left(\left| \int_{\{(f_k)_{\mathcal{N}_2} \neq \epsilon_k h_k\}} \operatorname{sgn}((f_k)_{\mathcal{N}_2} - \epsilon_k h_k) h_k dx \right| \right. \\
& \quad \left. - \int_{\{(f_k)_{\mathcal{N}_2} = \epsilon_k h_k\}} h_k dx \right),
\end{aligned}$$

and with $\epsilon \rightarrow 0^+$ in (76), it happens

$$(77) \quad gVTV(\mathbf{h}) \geq \lambda \sum_{k=1}^N \left(\int_{\{(f_k)_{\mathcal{N}_2} \neq 0\}} \operatorname{sgn}((f_k)_{\mathcal{N}_2}) h_k dx - \int_{\{(f_k)_{\mathcal{N}_2} = 0\}} h_k dx \right).$$

On the other hand, from the assumed inequalities in (72) we conclude from Lemma B.3 that there exists $\tilde{\mathbf{h}} \in BV(\mathcal{N}, \mathbb{R}^N)$ satisfying

$$(78) \quad \lambda \sum_{k=1}^N \left(\left| \int_{\{(f_k)_{\mathcal{N}_2} \neq 0\}} \operatorname{sgn}((f_k)_{\mathcal{N}_2}) \tilde{h}_k dx \right| - \int_{\{(f_k)_{\mathcal{N}_2} = 0\}} \tilde{h}_k dx \right) \geq gVTV(\tilde{\mathbf{h}})$$

contradicting (77), and we conclude that $\mathbf{u}_{\mathcal{N}_2} \neq 0$.

In order to prove that $\mathbf{u}_{\mathcal{N}_1} = 0$, we proceed by contradiction by assuming that $\mathbf{u}_{\mathcal{N}_1} \neq 0$. Given $\mathcal{N}_3 = \mathcal{N} \setminus (\mathcal{N}_1 \cup \mathcal{N}_2)$ with $\mathbf{u}_{\mathcal{N}_3} \neq 0$, we notice that

$$\begin{aligned}
 & gTV((u_k)_{\mathcal{N}_3} + (u_k)_{\mathcal{N}_3^c}) + \sum_{l=1}^N J_{k,l}((u_k)_{\mathcal{N}_3} + (u_k)_{\mathcal{N}_3^c}, (u_l)_{\mathcal{N}_3} + (u_l)_{\mathcal{N}_3^c}) \\
 &= \sup_{p \in \Phi_g} \left\{ \int_{\Omega} (u_k)_{\mathcal{N}_3} \operatorname{div} \mathbf{p} \, dx + \int_{\Omega} (u_k)_{\mathcal{N}_3^c} \operatorname{div} \mathbf{p} \, dx \right\} \\
 (79) \quad &+ \int_{\Omega} \sum_{l=1}^N |(\nabla(u_k)_{\mathcal{N}_3} - \nabla(u_l)_{\mathcal{N}_3}) + (\nabla(u_k)_{\mathcal{N}_3^c} - \nabla(u_l)_{\mathcal{N}_3^c})|^2 \, dx \\
 &\geq \sup_{p \in \Phi_g} \left\{ \int_{\Omega} (u_k)_{\mathcal{N}_3} \operatorname{div} \mathbf{p} \, dx \right\} + \int_{\Omega} \sum_{l=1}^N |(\nabla(u_k)_{\mathcal{N}_3} - \nabla(u_l)_{\mathcal{N}_3})|^2 \, dx \\
 &= gTV((u_k)_{\mathcal{N}_3^c}) + \sum_{l=1}^N J_{k,l}((u_k)_{\mathcal{N}_3^c}, (u_l)_{\mathcal{N}_3^c}).
 \end{aligned}$$

Therefore,

$$(80) \quad gVTV(\mathbf{u}_{\mathcal{N}_3} + \mathbf{u}_{\mathcal{N}_3^c}) + \lambda \sum_{k=1}^N \int_{\Omega} |(u_k)_{\mathcal{N}_3}| \, dx \geq gVTV(\mathbf{u}_{\mathcal{N}_3^c}),$$

On the other hand, since $\mathbf{u}$ is a minimizer

$$\begin{aligned}
 & gVTV(\mathbf{u}_{\mathcal{N}_3} + \mathbf{u}_{\mathcal{N}_3^c}) + \lambda \sum_{k=1}^N \left(\int_{\Omega} |(u_k)_{\mathcal{N}_3} - (f_k)_{\mathcal{N}_3}| \, dx + \int_{\Omega} |(u_k)_{\mathcal{N}_3^c} - (f_k)_{\mathcal{N}_3^c}| \, dx \right) \\
 (81) \quad &\leq gVTV(\mathbf{u}_{\mathcal{N}_3^c}) + \lambda \sum_{k=1}^N \left(\int_{\Omega} |(f_k)_{\mathcal{N}_3}| \, dx + \int_{\Omega} |(u_k)_{\mathcal{N}_3^c} - (f_k)_{\mathcal{N}_3^c}| \, dx \right),
 \end{aligned}$$

and due to the fact that $\mathbf{f}_{\mathcal{N}_3} = 0$, we get the inequality

$$(82) \quad gVTV(\mathbf{u}_{\mathcal{N}_3} + \mathbf{u}_{\mathcal{N}_3^c}) + \lambda \sum_{k=1}^N \int_{\Omega} |(u_k)_{\mathcal{N}_3}| \, dx \leq gVTV(\mathbf{u}_{\mathcal{N}_3^c}),$$

which is different from the identity in (80). Then, considering the definition of the subset $\mathcal{N}_3$, it necessarily happens that $\mathbf{u}_{\mathcal{N}_3} = 0$, and we get

$$\begin{aligned}
 & E_{gTV}^1(\mathbf{u}_{\mathcal{N} \setminus \mathcal{N}_1}) \\
 &\geq E_{gTV}^1(\mathbf{u}_{\mathcal{N} \setminus \mathcal{N}_1} + \mathbf{u}_{\mathcal{N}}) \\
 &= \sum_{k=1}^N \int_{\mathcal{N} \setminus \mathcal{N}_1} g_k(x) |\nabla(u_k)_{\mathcal{N} \setminus \mathcal{N}_1}| \, dx \\
 (83) \quad &+ \alpha \int_{\mathcal{N} \setminus \mathcal{N}_1} \sum_{k=1}^N \sum_{l=1}^N |\nabla(u_k)_{\mathcal{N} \setminus \mathcal{N}_1} - \nabla(u_l)_{\mathcal{N} \setminus \mathcal{N}_1}|^2 \, dx \\
 &+ \sum_{k=1}^N \int_{\mathcal{N}_1} g_k(x) |\nabla(u_k)_{\mathcal{N}_1}| \, dx + \alpha \int_{\mathcal{N}_1} \sum_{k=1}^N \sum_{l=1}^N |\nabla(u_k)_{\mathcal{N}_1} - \nabla(u_l)_{\mathcal{N}_1}|^2 \, dx
 \end{aligned}$$

$$\begin{aligned}
& + \sum_{k=1}^N \left(\lambda \int_{\mathcal{N} \setminus \mathcal{N}_1} |(u_k)_{\mathcal{N} \setminus \mathcal{N}_1} - (f_k)_{\mathcal{N} \setminus \mathcal{N}_1}| dx + \lambda \int_{\mathcal{N}_1} |(f_k)_{\mathcal{N}_1} - (u_k)_{\mathcal{N}_1}| dx \right) \\
& = E_{gVTV}^1(\mathbf{u}_{\mathcal{N} \setminus \mathcal{N}_1}) + \sum_{k=1}^N \int_{\mathcal{N}_1} g_k(x) |\nabla(u_k)_{\mathcal{N}_1}| dx \\
& + \alpha \int_{\mathcal{N}_1} \sum_{k=1}^N \sum_{l=1}^N |\nabla(u_k)_{\mathcal{N}_1} - \nabla(u_l)_{\mathcal{N}_1}|^2 dx \\
& + \lambda \sum_{k=1}^N \left(\int_{\mathcal{N}_1} |(f_k)_{\mathcal{N}_1} - (u_k)_{\mathcal{N}_1}| dx - \int_{\mathcal{N}_1} |(f_k)_{\mathcal{N}_1}| dx \right)
\end{aligned}$$

and therefore

$$\begin{aligned}
(84) \quad & \sum_{k=1}^N \int_{\mathcal{N}_1} g_k(x) |\nabla(u_k)_{\mathcal{N}_1}| dx + \alpha \int_{\mathcal{N}_1} \sum_{k=1}^N \sum_{l=1}^N |\nabla(u_k)_{\mathcal{N}_1} - \nabla(u_l)_{\mathcal{N}_1}|^2 dx \\
& \leq \lambda \sum_{k=1}^N \left(\int_{\mathcal{N}_1} |(f_k)_{\mathcal{N}_1}| dx - \int_{\mathcal{N}_1} |(f_k)_{\mathcal{N}_1} - (u_k)_{\mathcal{N}_1}| dx \right).
\end{aligned}$$

We also have from the convexity of $|\cdot|$ that

$$\begin{aligned}
(85) \quad & \int_{\mathcal{N}_1} |(f_k)_{\mathcal{N}_1}| dx - \int_{\mathcal{N}_1} |(f_k)_{\mathcal{N}_1} - u_k| dx \\
& \leq \left(\int_{\{(f_k)_{\mathcal{N}_1} \neq 0\}} \operatorname{sgn}((f_k)_{\mathcal{N}_1}) u_k dx \right) - \int_{\{(f_k)_{\mathcal{N}_1} = 0\}} u_k dx.
\end{aligned}$$

From inequalities in (84) and (85) we get

$$(86) \quad gVTV(\mathbf{u}_{\mathcal{N}_1}) \leq \lambda \sum_{k=1}^N \left(\left| \int_{\{(f_k)_{\mathcal{N}_1} \neq 0\}} \operatorname{sgn}((f_k)_{\mathcal{N}_1}) h_k dx \right| - \int_{\{(f_k)_{\mathcal{N}_1} = 0\}} h_k dx \right),$$

and at the same time from Lemma B.3, it happens

$$(87) \quad gVTV(\mathbf{u}_{\mathcal{N}_1}) \geq \lambda \sum_{k=1}^N \left(\left| \int_{\{(f_k)_{\mathcal{N}_1} \neq 0\}} \operatorname{sgn}((f_k)_{\mathcal{N}_1}) h_k dx \right| - \int_{\{(f_k)_{\mathcal{N}_1} = 0\}} h_k dx \right),$$

which is a contradiction with (86). Then, we have $\mathbf{u}_{\mathcal{N}_1} = 0$. $\square$

REFERENCES

- [1] L. Ambrosio, N. Fusco and D. Pallara, *Functions of Bounded Variation and Free Discontinuity Problems*, Oxford University Press, Cambridge, UK, 2000.
- [2] H. Attouch, G. Buttazzo and G. Michaille, *Variational Analysis in Sobolev and BV Spaces: Applications to PDEs and Optimization*, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, USA, 2006.
- [3] G. Aubert and J. F. Aujol, [Modelling very oscillating signals. Application to image processing](#), *Applied Mathematics and Optimization*, **51** (2005), 163–182.
- [4] G. Aubert and P. Kornprobst, *Mathematical Problems in Image Processing: Partial Differential Equation and Calculus of Variations*, Springer-Verlag, New York, USA, 2006.
- [5] J.-F. Aujol and A. Chambolle, [Dual norms and image decomposition models](#), *International Journal of Computer Vision*, **63** (2005), 85–104.
- [6] J.-F. Aujol and T. F. Chan, [Combining geometrical and textured information to perform image classification](#), *Journal of Visual Communication and Image Representation*, **17** (2006), 1004–1023.

- [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** (2006), 111–136.
- [8] J.-F. Aujol and S. H. Kang, [Color image decomposition and restoration](#), *Journal of Visual Communication and Image Representation*, **17** (2006), 916–928.
- [9] L. Bar, N. Sochen and N. Kiryati, [Image deblurring in the presence of salt-and-pepper noise](#), in *Lecture Notes in Computer Science*, Germany, **3459** (2005), 107–118.
- [10] A. Beck and M. Teboulle, [Fast gradient-based algorithms for constrained total variation image denoising and deblurring problems](#), *IEEE Transactions on Image Processing*, **18** (2009), 2419–2434.
- [11] P. Blomgren and T. F. Chan, [Color TV: Total variation methods for restoration of vector valued images](#), *IEEE Transactions on Image Processing*, **7** (1998), 304–309.
- [12] G. Boccignone, M. Ferraro and T. Caelli, [Generalized spatio-chromatic diffusion](#), *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **24** (2002), 1298–1309.
- [13] 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** (2008), 455–484.
- [14] X. Bresson, S. Esedoglu, P. Vanderghenst, J. Thiran and S. Osher, [Fast global minimization of the active contour/snake model](#), *Journal of Mathematical Imaging and Vision*, **28** (2007), 151–167.
- [15] A. Brook, R. Kimmel and N. Sochen, [Variational restoration and edge detection for color images](#), *Journal of Mathematical Imaging and Vision*, **18** (2003), 247–268.
- [16] A. Buades, B. Coll and J. M. Morel, [A review of image denoising methods, with a new one](#), *Multiscale Modeling and Simulation*, **4** (2005), 490–530.
- [17] V. Caselles, F. Catté, T. Coll and F. Dibos, [A geometric model for active contours in image processing](#), *Numerische Mathematik*, **66** (1993), 1–31.
- [18] V. Caselles, B. Coll and J.-M. Morel, [Geometry and color in natural images](#), *Journal of Mathematical Imaging and Vision*, **16** (2002), 89–105.
- [19] V. Caselles, R. Kimmel and G. Sapiro, [Geodesic active contours](#), *International Journal of Computer Vision*, (1995), 694–699.
- [20] A. Chambolle, [An algorithm for total variation minimization and applications](#), *Journal of Mathematical Imaging and Vision*, **20** (2004), 89–97.
- [21] A. Chambolle and T. Pock, [A first-order primal-dual algorithm for convex problems with applications to imaging](#), *Journal of Mathematical Imaging and Vision*, **40** (2011), 120–145.
- [22] T. F. Chan and S. Esedoglu, [Aspects of total variation regularized \$L^1\$ function approximation](#), *SIAM Journal on Applied Mathematics*, **65** (2005), 1817–1837.
- [23] T. F. Chan, S. Esedoglu and M. Nikolova, [Algorithms for finding global minimizers of image segmentation and denoising models](#), *SIAM Journal on Applied Mathematics*, **66** (2006), 1632–1648.
- [24] T. F. Chan, G. Golub and P. Mulet, [A nonlinear primal-dual method for total variation-based image restoration](#), *SIAM Journal on Scientific Computing*, **20** (1999), 1964–1967.
- [25] T. F. Chan and J. Shen, [Variational image inpainting](#), *Communications on Pure and Applied Mathematics*, **58** (2005), 579–619.
- [26] K. Dabov, A. Foi, V. Katkovnik and K. Egiazarian, [Color image denoising via sparse 3d collaborative filtering with grouping constraint in luminance-chrominance space](#), in *IEEE International Conference on Image Processing*, San Antonio, TX, USA, **1** (2007), 313–316.
- [27] J. Darbon, [Total variation minimization with \$L^1\$ data fidelity as a contrast invariant filter](#), in *4th Symposium on Image and Signal Processing and Analysis (ISPA)*, Zagreb, Croatia, 2005, 221–226.
- [28] Y. Dong, M. Hintermuller and M. M. Rincon-Camacho, [A multi-scale vectorial \$L^7\$ -TV framework for color image restoration](#), *International Journal of Computer Vision*, **92** (2011), 296–307.
- [29] S. Durand, J. Fadili and M. Nikolova, [Multiplicative noise removal using \$L^1\$ fidelity on frame coefficients](#), *Journal of Mathematical Imaging and Vision*, **36** (2010), 201–226.
- [30] V. Duval, J.-F. Aujol and Y. Gousseau, [The TVL1 model: A geometrical point of view](#), *Multiscale Modelling & Simulation*, **8** (2009), 154–189.
- [31] V. Duval, J.-F. Aujol and L. Vese, [Mathematical modelling of textures: Application to color image decomposition with a projected gradient algorithm](#), *Journal of Mathematical Imaging and Vision*, **37** (2010), 232–248.

- [32] M. J. Ehrhardt and S. R. Arridge, [Vector-valued image processing by parallel level sets](#), *IEEE Transactions on Image Processing*, **23** (2014), 9–18.
- [33] G. Gilboa, [A total variation spectral framework for scale and texture analysis](#), *Multiscale Modelling and Simulation*, **7** (2014), 1937–1961.
- [34] J. Gilles, [Noisy image decomposition: A new structure, texture and noise model based on local adaptivity](#), *Journal of Mathematical Imaging and Vision*, **28** (2007), 285–295.
- [35] J. Gilles, [Multiscale texture separation](#), *Multiscale Modeling & Simulation*, **10** (2012), 1409–1427.
- [36] B. Goldluecke, E. Strekalovskiy and D. Cremers, [The natural vectorial total variation which arises from geometric measure theory](#), *SIAM Journal on Imaging Sciences*, **5** (2012), 537–564.
- [37] T. Goldstein and S. Osher, [The split Bregman algorithm for L1 regularized problems](#), *SIAM Journal on Imaging Sciences*, **2** (2009), 323–343.
- [38] J. B. Greer and A. L. Bertozzi, [Traveling wave solutions of fourth order PDEs for image processing](#), *SIAM Journal on Mathematical Analysis*, **36** (2004), 38–68.
- [39] M. Kass, A. Witkin and D. Terzopoulos, [Snakes: Active contour models](#), *International Journal of Computer Vision*, **1** (1988), 321–331.
- [40] T. M. Le and L. A. Vese, [Image decomposition using total variation and div\(BMO\)](#), *Multiscale Modelling & Simulation*, **4** (2005), 390–423.
- [41] X. Liu, L. Huang and Z. Guo, [Adaptive fourth-order partial differential equation filter for image denoising](#), *Applied Mathematics Letters*, **24** (2011), 1282–1288.
- [42] 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.
- [43] J. C. Moreno, V. B. S. Prasath, D. Vorotnikov, H. Proenca and K. Palaniappan, [Adaptive Diffusion Constrained Total Variation Scheme with Application to Cartoon + Texture + Edge Image Decomposition](#), Technical Report 1505.00866, ArXiv, 2015, URL <http://arxiv.org/abs/1505.00866>.
- [44] Y. Nesterov, [Smooth minimization of non-smooth functions](#), *Mathematical Programming*, **103** (2005), 127–152.
- [45] M. Nikolova, [Minimizers of cost-function involving nonsmooth data-fidelity terms](#), *SIAM Journal on Numerical Analysis*, **40** (2002), 965–994.
- [46] M. Nikolova, [A variational approach to remove outliers and impulse noise](#), *Journal of Mathematical Imaging and Vision*, **20** (2004), 99–120, Special Issue on Mathematics and Image Analysis.
- [47] M. Nikolova, [Weakly constrained minimization: application to the estimation of images and signals involving constant regions](#), *Journal of Mathematical Imaging and Vision*, **21** (2004), 155–175.
- [48] S. Ono and I. Yamada, [Decorrelated vectorial total variation](#), in *IEEE Conference on Computer Vision and Pattern Recognition*, Columbus, OH, USA, 2014, 4090–4097.
- [49] S. Osher and O. Scherzer, [G-norm properties of bounded variation regularization](#), *Communications in Mathematical Sciences*, **2** (2004), 237–254.
- [50] N. Paragios, O. Mellina-Gottardo and V. Ramesh, [Gradient vector flow fast geodesic active contours](#), *Computer Vision, 2001. ICCV 2001. Proceedings. Eighth IEEE International Conference on*, **1** (2001), 67–73.
- [51] V. B. S. Prasath, [Weighted Laplacian differences based multispectral anisotropic diffusion](#), in *IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, Vancouver BC, Canada, 2011, 4042–4045.
- [52] V. B. S. Prasath, [A well-posed multiscale regularization scheme for digital image denoising](#), *International Journal of Applied Mathematics and Computer Science*, **21** (2011), 769–777.
- [53] V. B. S. Prasath, J. C. Moreno and K. Palaniappan, [Color image denoising by chromatic edges based vector valued diffusion](#), Technical Report 1304.5587, ArXiv, 2013, URL <http://arxiv.org/abs/1304.5587>.
- [54] V. B. S. Prasath and A. Singh, [Multichannel image restoration using combined channel information and robust M-estimator approach](#), *International Journal Tomography and Statistics*, **12** (2010), 9–22.
- [55] V. B. S. Prasath and A. Singh, [Multispectral image denoising by well-posed anisotropic diffusion scheme with channel coupling](#), *International Journal of Remote Sensing*, **31** (2010), 2091–2099.

- [56] V. B. S. Prasath and A. Singh, Well-posed inhomogeneous nonlinear diffusion scheme for digital image denoising, *Journal of Applied Mathematics*, **2010** (2010), 14pp, Article ID 763847.
- [57] 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** (2012), 1250003, 18 pp.
- [58] L. Rudin, S. Osher and E. Fatemi, Nonlinear total variation based noise removal algorithms, *Physica D*, **60** (1992), 259–268.
- [59] G. Sapiro and D. L. Ringach, Anisotropic diffusion of multivalued images with application to color filtering, *IEEE Transactions on Image Processing*, **5** (1996), 1582–1586.
- [60] O. Scherzer, W. Yin and S. Osher, Slope and G -set characterization of set-valued functions and applications to non-differentiable optimization problems, *Communications in Mathematical Sciences*, **3** (2005), 479–492.
- [61] Y. Shi, L.-L. Wang and X.-C. Tai, Geometry of total variation regularized L^p -model, *Journal of Computational and Applied Mathematics*, **236** (2012), 2223–2234.
- [62] D. Strong, *Adaptive Total Variation Minimizing Image Restoration*, PhD thesis, UCLA Mathematics Department, USA, 1997.
- [63] D. Strong and T. Chan, Edge-preserving and scale-dependent properties of total variation regularization, *Inverse Problems*, **19** (2003), 165–187.
- [64] 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.
- [65] E. Tadmor, S. Nezzar and L. Vese, A multiscale image representation using hierarchical (BV, L^2) decomposition, *Multiscale Modeling and Simulation*, **2** (2004), 554–579.
- [66] B. Tang, G. Sapiro and V. Caselles, Color image enhancement via chromaticity diffusion, *IEEE Transactions on Pattern Analysis Machine Intelligence*, **10** (2001), 701–707.
- [67] L. Tang and C. He, Multiscale texture extraction with hierarchical (BV, G_p, L^2) , *Journal of Mathematical Imaging and Vision*, **45** (2013), 148–163.
- [68] D. Tschumperle and R. Deriche, Vector-valued image regularization with PDE's: A common framework for different applications, *Computer Vision and Pattern Recognition, 2003. Proceedings. 2003 IEEE Computer Society Conference on*, **1** (2003), 651–656.
- [69] L. A. Vese and S. J. Osher, Modeling textures with total variation minimization and oscillating patterns in image processing, *Journal of Scientific Computing*, **19** (2003), 553–572.
- [70] L. A. Vese and S. J. Osher, Color texture modeling and color image decomposition in variational PDE approach, in *Proceedings of the Eighth International Symposium on Symbolic and Numeric Algorithms for Scientific Computing (SYNASC '06)*, New York, USA, 2006, 103–110.
- [71] 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** (2004), 600–612, URL <http://ece.uwaterloo.ca/~z70wang/research/ssim/>.
- [72] W. Yin, D. Goldfarb and S. Osher, Image cartoon-texture decomposition and feature selection using total variation regularized L^1 functional, in *Variational, Geometric, and Level Set Methods in Computer Vision*, **3752** (2005), 73–84.
- [73] W. Yin, D. Goldfarb and S. Osher, The total variation regularized L^1 model for multiscale decomposition, *Multiscale Modeling & Simulation*, **6** (2007), 190–211.

Received December 2014; 1st revision June 2015; 2nd revision August 2015.

E-mail address: jcmb@ubi.pt

E-mail address: prasaths@missouri.edu

E-mail address: jcneves@ubi.pt