

Decorrelated Vectorial Total Variation

Shunsuke Ono

Isao Yamada

Department of Communications and Computer Engineering, Tokyo Institute of Technology, Japan

{ono/isao}@sp.ce.titech.ac.jp

Abstract

This paper proposes a new vectorial total variation prior (VTV) for color images. Different from existing VTVs, our VTV, named the decorrelated vectorial total variation prior (D-VTV), measures the discrete gradients of the luminance component and that of the chrominance one in a separated manner, which significantly reduces undesirable uneven color effects. Moreover, a higher-order generalization of the D-VTV, which we call the decorrelated vectorial total generalized variation prior (D-VTGV), is also developed for avoiding the staircasing effect that accompanies the use of VTVs. A noteworthy property of the D-VT(G)V is that it enables us to efficiently minimize objective functions involving it by a primal-dual splitting method. Experimental results illustrate their utility.

1. Introduction

Priors characterizing reasonable estimates play a prominent role in computer vision and image processing because numerous problems including denoising, deconvolution, inpainting, super-resolution and compressed sensing, are ill-posed or ill-conditioned. Existing priors for images can be roughly classified into two categories, namely, *local* and *nonlocal* ones. Nonlocal priors, such as the *nonlocal means* [6], the *nonlocal total variation* [15, 37], and a number of patch-based priors (e.g., [11, 21, 8]), are often superior to local ones owing to their full-activation of information in a given data.

On the other hand, improving local priors is still an important issue due to, for example, the following reasons. First, local priors can be well defined as a “convex function” in many cases, which enables us to leverage them in optimization problems designed for a variety of scenarios and to efficiently solve the problems by convex optimization techniques. By contrast, many nonlocal priors can be described only as “procedure” for a particular scenario (typically only for Gaussian denoising), which precludes their direct application to other scenarios. Second, the implementation of local priors is usually easier than that of non-

local priors. Specifically, local priors are free from complicated (and often chicken-and-egg) self-similarity evaluation such as block matching, which is necessary in the use of nonlocal priors. Third, initial estimation required for nonlocal priors (especially in the case where a given data involves severe degradation such as blur and missing components) is usually executed by local priors. Indeed, the quality of the initial estimation affects that of the final one.

With this background, the development of a class of *vectorial total variation priors* (VTV) [30, 2, 1, 5, 12], which are extensions of the well-known *total variation prior* (TV) [29] and are successful local priors for color images, is indeed a very active research topic, with new techniques still emerging [16, 17, 23, 19]. Existing VTVs are defined as some norm of the discrete gradients of neighboring pixels of (all channels of) a color image, where they are different mainly in terms of what norm is employed. Almost all of them, however, do not fully exploit *inter-channel correlation*, so that the use of them sometimes produces undesirable *uneven color effects*, as observed in the results obtained by existing VTVs (see Section 4). Although the VTV proposed in [23] explicitly takes the correlation into account and succeeds in reducing (but not sufficiently) uneven color effects, it also has several drawbacks: i) it is *anisotropic*, i.e., the vertical and horizontal gradients are decoupled, resulting in the generation of blocky artifacts around contours; ii) it requires the calculation of the projection onto the ℓ_1 -norm ball in optimization, which increases the computational cost.

This paper proposes a new isotropic VTV, named the *decorrelated vectorial total variation* (D-VTV), with its theoretical properties and applications. The idea is twofold, that is, i) to incorporate a color transform into the definition of VTV for decorrelating RGB channels and ii) to separately measure luminance and chrominance variations with a weight controlling the balance between them. Idea i) appears similar to but is actually different from such a two-step method that first applies some color transform to a given observation and then performs existing VTV regularization. Such method cannot be applied to inverse problems with missing components and cannot guarantee the

optimality in the whole process. Moreover, most existing VTVs are designed to be invariant with respect to orthogonal color transforms, implying that incorporating such transforms into the existing VTVs is meaningless. By contrast, the D-VTV is color-transform-variant because of Idea ii) and can suppress chrominance variation in preference to luminance one, so that it can drastically reduce uneven color effects. In addition, the structure of the D-VTV makes the minimization of objective functions involving it very efficient. We will elaborate on the essential differences between existing VTVs and the D-VTV in Remark 2.1.

Furthermore, we extend the D-VTV to a higher-order version, which we call the *decorrelated vectorial total generalized variation* (D-VTGV) inspired by the prior works [4, 31, 3, 24]. This is introduced for reducing the *staircase effect* that accompanies the use of VTVs.

The D-VT(G)V is applied to imaging inverse problems, where we provide an efficient algorithmic solution to the associated convex optimization problems involving the D-VT(G)V via a primal-dual splitting method [7, 10, 35]. Experimental comparisons of existing VT(G)V and the D-VT(G)V are given in Section 4, which demonstrates the superiority of our proposed prior.

2. Proposed Prior

In the following, \mathbb{N} , \mathbb{R} , \mathbb{R}_+ , and \mathbb{R}_{++} denote the sets of positive integers, all, nonnegative, and positive real numbers, respectively, and let $\overline{\mathbb{R}} := \mathbb{R} \cup \{\infty\}$. We adopt the vector notation for color images. That is, the channel components on a color image of size $N_v \times N_h \times 3$ are stacked into a vector $\mathbf{u} := [\mathbf{u}_R^\top \mathbf{u}_G^\top \mathbf{u}_B^\top]^\top \in \mathbb{R}^{3N}$ in lexicographic order, where $N = N_v N_h$ is the number of the pixels, $\mathbf{u}_R, \mathbf{u}_G, \mathbf{u}_B \in \mathbb{R}^N$ are the RGB channels (also expressed as vectors), and \cdot^\top stands for the transposition.

2.1. Definition

By letting $\mathbf{D}_v, \mathbf{D}_h \in \mathbb{R}^{N \times N}$ be the vertical and horizontal gradient operators with Neumann boundary, the first-order gradient operator for a single channel is defined by $\mathbf{D}_1 := [\mathbf{D}_v^\top \mathbf{D}_h^\top]^\top \in \mathbb{R}^{2N \times N}$, and that for a color image by $\mathbf{D} := \text{diag}(\mathbf{D}_1, \mathbf{D}_1, \mathbf{D}_1) \in \mathbb{R}^{6N \times 3N}$. We also introduce an orthonormal color transform $\mathbf{C} : \mathbb{R}^{3N} \rightarrow \mathbb{R}^{3N} : \mathbf{u} \mapsto [\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3]$ for decorrelating color channels, where $\mathbf{u}_1 := \frac{1}{\sqrt{3}}(\mathbf{u}_R + \mathbf{u}_G + \mathbf{u}_B)$ is the luminance channel, and $\mathbf{u}_2 := \frac{1}{\sqrt{2}}(\mathbf{u}_R - \mathbf{u}_B)$ and $\mathbf{u}_3 := \frac{1}{\sqrt{6}}(\mathbf{u}_R - 2\mathbf{u}_G + \mathbf{u}_B)$ are the chrominance channels. This transform is the so-called opponent transform [28] and provides effective reduction of the correlation among RGB channels, as employed in color image denoising methods [14, 36]. Moreover, for $k_1, k_2 \in \mathbb{N}$ and $w \in (0, 1)$, we define the function:

$$\|\cdot\|_{1,2}^{(w,k_1,k_2)} : \mathbb{R}^{(k_1+k_2)N} \rightarrow \mathbb{R}_+ : \mathbf{x} \mapsto w\|\mathbf{x}_1\|_{1,2}^{(k_1)} + \|\mathbf{x}_2\|_{1,2}^{(k_2)},$$

where $\|\cdot\|_{1,2}^{(k)} : \mathbb{R}^{kN} \rightarrow \mathbb{R}_+ : \mathbf{x} \mapsto \sum_{i=1}^N \sqrt{\sum_{j=0}^{k-1} x_{i+j}^2}$ is the mixed $\ell_{1,2}$ norm ($k \in \mathbb{N}$ and x_i denotes the i th entry of \mathbf{x}), and $\mathbf{x} = [\mathbf{x}_1^\top \mathbf{x}_2^\top]^\top$ with $\mathbf{x}_1 \in \mathbb{R}^{k_1 N}$ and $\mathbf{x}_2 \in \mathbb{R}^{k_2 N}$. The function $\|\cdot\|_{1,2}^{(w,k_1,k_2)}$ is obviously a norm on $\mathbb{R}^{(k_1+k_2)N}$. Using $\|\cdot\|_{1,2}^{(w,k_1,k_2)}$, the *Decorrelated Vectorial Total Variation prior* (D-VTV) is then defined by

$$J_{\text{VTV}}^w : \mathbb{R}^{3N} \rightarrow \mathbb{R}_+ : \mathbf{u} \mapsto \|\mathbf{DCu}\|_{1,2}^{(w,2,4)}.$$

Proposition 2.1 *The D-VTV is continuous and convex.*

Proof: This follows at once from the continuity and convexity of $\|\cdot\|_{1,2}^{(k)}$, and the linearity of \mathbf{D} and \mathbf{C} . \square

Remark 2.1 (The D-VTV versus existing VTVs)

- The design of the weight w being smaller than one is intended to suppress chrominance variation (the second term) in preference to luminance one (the first term), which prevents uneven color effects. As will be demonstrated in Section 4, the value of the D-VTV increases rapidly based on the level of noise contamination compared to existing VTVs, implying that noise magnifies chrominance variation rather than luminance one and thus should be suppressed first. This property also tends to keep meaningful details (e.g., edges and texture) in color images because such details often have large luminance variation. These things suggest that the D-VTV is a very suitable prior for color images.
- Since the D-VTV adopts the mixed $\ell_{1,2}$ norm, the minimization of objective functions involving the D-VTV becomes computationally efficient. This is because, in such minimization, we require the computation of the so-called *proximity operator*¹ of the norm employed in the corresponding VTV, and in the D-VTV case it is the mixed $\ell_{1,2}$ norm with the associated proximity operator given by a simple soft-thresholding type operation: for $\gamma \in \mathbb{R}_{++}$ and for $i = 1, \dots, kN$,

$$[\text{prox}_{\gamma\|\cdot\|_{1,2}^{(k)}}(\mathbf{x})]_i = \max\{1 - \gamma(\sum_{j=0}^{k-1} x_{i+j}^2)^{-\frac{1}{2}}, 0\}x_i.$$

By contrast, the proximity operators of the norms employed in state-of-the-art VTVs require more expensive operations (see also the next item), so that the D-VTV is also preferable in terms of computational cost.

- We summarize the comparison of VTVs in Tab. 1, where we refer to the prior in [5, 12] as the *basic VTV* (B-VTV), in [16, 17] as the *spectral VTV* (S-VTV), in [19] as the *nuclear VTV* (N-VTV), and in [23] as the *L-infinity VTV* (I-VTV). Note that the S-VTV and the N-VTV require the singular value decomposition, and the I-VTV the weighted ℓ_1 -norm ball projection, in their optimization, respectively. Indeed, the B-VTV, which also employs the

¹ The *proximity operator* of a proper lower semicontinuous convex function $f : \mathcal{X} \rightarrow \overline{\mathbb{R}}$ (\mathcal{X} is a Euclidean space) of index $\gamma \in \mathbb{R}_{++}$ is given by $\text{prox}_{\gamma f}(\mathbf{x}) := \arg \min_{\mathbf{y} \in \mathcal{X}} \{f(\mathbf{y}) + \frac{1}{2\gamma}\|\mathbf{x} - \mathbf{y}\|_2^2\}$ ($\|\cdot\|_2$ is the ℓ_2 norm).

Table 1: Comparison of VTVs

	norm	isotropy	uneven color	cost
B-VTV [5, 12]	$\ell_{1,2}$	○	△	⊙
S-VTV [16, 17]	spectral	○	△	△
I-VTV [23]	$\ell_{1,\infty}$	×	○	△
N-VTV [19]	nuclear	○	△	△
D-VTV (ours)	$\ell_{1,2}$	○	⊙	⊙

mixed $\ell_{1,2}$ norm, is the most computationally-efficient among the said existing VTVs, and the additional computation required for the D-VTV compared to the B-VTV is just the color transform \mathbf{C} , the cost of which is negligible.

2.2. Higher-order generalization

The *staircasing effect*, i.e., the undesirable appearance of edges, often accompanies the use of VTVs. The *total generalized variation* (TGV) [4] was introduced as a well-established higher-order generalization of the TV in order to overcome this limitation, and is recently utilized in a number of applications, e.g., [34, 13]. The notion of the TGV has also been incorporated into the B-VTV [3] and the I-VTV [24]. The D-VTV is no exception to the said limitation but it is expected to be more effective than existing VTVs, which motivates us to extend the TGV to a new vectorial version in the view of the D-VTV.

To proceed, let $\mathbf{G}_1 := \begin{bmatrix} -\mathbf{D}_v^\top & \mathbf{O} \\ -\mathbf{D}_h^\top & -\mathbf{D}_v^\top \\ \mathbf{O} & -\mathbf{D}_h^\top \end{bmatrix} \in \mathbb{R}^{3N \times 2N}$ and $\mathbf{G} := \text{diag}(\mathbf{G}_1, \mathbf{G}_1, \mathbf{G}_1) \in \mathbb{R}^{9N \times 6N}$ (\mathbf{O} denotes zero matrices of appropriate size). The *Decorrelated Vectorial Total Generalized Variation* (D-VTGV) is defined as follows: $J_{\text{VTGV}}^{\mathbf{w},\alpha} : \mathbb{R}^{3N} \rightarrow \mathbb{R}_+$:

$$\mathbf{u} \mapsto \min_{\mathbf{p} \in \mathbb{R}^{6N}} \alpha \|\mathbf{DCu} - \mathbf{p}\|_{1,2}^{(w_1,2,4)} + (1 - \alpha) \|\mathbf{Gp}\|_{1,2}^{(w_2,3,6)},$$

where $\mathbf{w} := [w_1 w_2] \in (0, 1) \times (0, 1)$ and $\alpha \in (0, 1)$. Since \mathbf{GD} becomes the second-order gradient operator, the D-VTGV can be seen as the *infimal convolution* of the first- and second-order terms with the parameter α controlling their balance.

Proposition 2.2 *The D-VTGV is continuous and convex.*

Proof: Let $g_1 := \alpha \|\cdot\|_{1,2}^{(w_1,2,4)}$ and $g_2 := (1 - \alpha) \|\mathbf{G} \cdot\|_{1,2}^{(w_2,3,6)}$ be functions on \mathbb{R}^{3N} , and let $g := \min_{(\cdot)=\mathbf{x}+\mathbf{y}} g_1(\mathbf{x}) + g_2(\mathbf{y})$ be the infimal convolution of them. Then g is continuous and convex from [31, Theorem 2.1] with the fact that $\|\cdot\|_{1,2}^{(w,k_1,k_2)}$ is a norm and that \mathbf{G} is a linear operator. Finally, the statement follows from $J_{\text{VTGV}}^{\mathbf{w},\alpha} = g \circ (\mathbf{DC})$ and the linearity of \mathbf{DC} . \square

3. Application

Since the D-VT(G)V is explicitly defined as a convex prior, it has a potential to be used as a building block for various applications in the computer vision and image processing fields. We focus on imaging inverse problems.

3.1. Problem formulation

Consider the observation model $\mathbf{v} = \mathcal{D}(\Phi \mathbf{u}_{\text{org}})$, where $\mathbf{u}_{\text{org}} \in \mathbb{R}^{3N}$ is an original color image we wish to estimate, $\Phi \in \mathbb{R}^{M \times 3N}$ a linear operator representing some acquisition process (e.g., blur), $\mathcal{D} : \mathbb{R}^M \rightarrow \mathbb{R}^M$ a noise contamination process being not necessarily additive, and $\mathbf{v} \in \mathbb{R}^M$ (M and $3N$ may be different) an observation. Based on this model, we propose a general form of convex optimization problems involving the D-VT(G)V as follows: find \mathbf{u}^* in

$$\arg \min_{\mathbf{u} \in [0, 255]^{3N}} J(\mathbf{u}) + F_{\mathbf{v}}(\Phi \mathbf{u}), \quad (1)$$

where $J \in \{J_{\text{VTV}}^{\mathbf{w}}, J_{\text{VTGV}}^{\mathbf{w},\alpha}\}$, $F_{\mathbf{v}} : \mathbb{R}^M \rightarrow \overline{\mathbb{R}}$ is a proper lower semicontinuous convex fidelity function regarding \mathbf{v} with its proximity operator computable (required in optimization), and $[0, 255]^{3N} \subset \mathbb{R}^{3N}$ is an intensity range constraint for eight-bit color images.

Proposition 3.1 (1) *has at least one solution.*

Proof: From Proposition 2.1 and 2.2, $J(\mathbf{u}) + F_{\mathbf{v}}(\Phi \mathbf{u})$ is proper lower semicontinuous convex. The compactness of the intensity range constraint then proves the statement.

Remark 3.1 (Examples of $F_{\mathbf{v}}$)

- (Gaussian noise) A reasonable choice is the ℓ_2 -norm ball: $\mathcal{B}_{\mathbf{v},\varepsilon}^2 := \{\mathbf{x} \in \mathbb{R}^M \mid \|\mathbf{x} - \mathbf{v}\|_2 \leq \varepsilon\}$, where $\varepsilon \in \mathbb{R}_+$. We express this constraint-type fidelity by the *indicator function*² of $\mathcal{B}_{\mathbf{v},\varepsilon}^2$, i.e., $F_{\mathbf{v}} := \iota_{\mathcal{B}_{\mathbf{v},\varepsilon}^2}$ with its proximity operator given by $\text{prox}_{\gamma F_{\mathbf{v}}}(\mathbf{x}) = \mathbf{x}$, if $\mathbf{x} \in \mathcal{B}_{\mathbf{v},\varepsilon}^2$; $\mathbf{v} + \frac{\varepsilon(\mathbf{x} - \mathbf{v})}{\|\mathbf{x} - \mathbf{v}\|_2}$, otherwise. The resulting formulation corresponds to minimizing the D-VTV while keeping the Gaussian likelihood at a certain level determined by ε . Compared to the standard additive fidelity, i.e., $J_{\text{VTV}}^{\mathbf{w}}(\mathbf{u}) + \frac{\lambda}{2} \|\Phi \mathbf{u} - \mathbf{v}\|_2^2$, the constraint-type fidelity would facilitate parameter setting because ε has a clearer meaning than λ and thus can be easily adjusted based on noise standard deviation.
- (Impulsive noise) The ℓ^1 norm has been adopted for impulsive noise contamination as a suitably robust fidelity term, which is defined by $F_{\mathbf{v}} := \lambda \|\cdot - \mathbf{v}\|_1$ ($\|\cdot\|_1$ denotes the ℓ^1 norm). The proximity operator of $F_{\mathbf{v}}$ is given, for $\gamma \in \mathbb{R}_{++}$ and for $i = 1, \dots, M$, by $[\text{prox}_{\gamma F_{\mathbf{v}}}(\mathbf{x})]_i = v_i + \text{sgn}(x_i - v_i) \max\{|x_i - v_i| - \gamma\lambda, 0\}$, where sgn denotes the signum function.
- (Poisson noise) It has been shown that under Poisson noise contamination, the so-called the *generalized Kulback-Leibler divergence* is suitable for $F_{\mathbf{v}}$. Its definition and the computation of the associated proximity operator can be found in [9].

² For a given nonempty closed convex set $C \in \mathcal{X}$, the indicator function of C is defined by $\iota_C(\mathbf{x}) := 0$, if $\mathbf{x} \in C$; ∞ , otherwise. Using the indicator function, we can express a convex constraint as an additive term. The proximity operator of ι_C is equivalent to the metric projection onto C , i.e., $\text{prox}_{\gamma \iota_C}(\mathbf{x}) = \arg \min_{\mathbf{y} \in C} \|\mathbf{x} - \mathbf{y}\| =: P_C(\mathbf{x})$ ($\forall \gamma \in \mathbb{R}_{++}$).

Algorithm 1: Solver for (1) with J_{VTV}^w

```

input :  $\mathbf{u}^{(0)}, \mathbf{z}_1^{(0)} := [\mathbf{z}_{1,1}^{(0)\top} \ \mathbf{z}_{1,(2,3)}^{(0)\top}]^\top, \mathbf{z}_2^{(0)}$ 
output :  $\mathbf{u}^{(n)}$ 
1 while A stopping criterion is not satisfied do
2    $\mathbf{u}^{(n+1)} = P_{[0,255]^{3N}}(\mathbf{u}^{(n)} - \gamma_1(\mathbf{C}^\top \mathbf{D}^\top \mathbf{z}_1^{(n)} + \Phi^\top \mathbf{z}_2^{(n)}));$ 
3    $[\mathbf{z}_{1,1}^{(n)\top} \ \mathbf{z}_{1,(2,3)}^{(n)\top}]^\top \leftarrow \mathbf{z}_1^{(n)} + \gamma_2 \mathbf{DC}(2\mathbf{u}^{(n+1)} - \mathbf{u}^{(n)});$ 
4    $\mathbf{z}_2^{(n)} \leftarrow \mathbf{z}_2^{(n)} + \gamma_2 \Phi(2\mathbf{u}^{(n+1)} - \mathbf{u}^{(n)});$ 
5    $\mathbf{z}_{1,1}^{(n+1)} = \mathbf{z}_{1,1}^{(n)} - \gamma_2 \text{prox}_{\frac{w}{\gamma_2} \|\cdot\|_{1,2}^{(2)}}(\frac{1}{\gamma_2} \mathbf{z}_{1,1}^{(n)});$ 
6    $\mathbf{z}_{1,(2,3)}^{(n+1)} = \mathbf{z}_{1,(2,3)}^{(n)} - \gamma_2 \text{prox}_{\frac{1}{\gamma_2} \|\cdot\|_{1,2}^{(4)}}(\frac{1}{\gamma_2} \mathbf{z}_{1,(2,3)}^{(n)});$ 
7    $\mathbf{z}_2^{(n+1)} = \mathbf{z}_2^{(n)} - \gamma_2 \text{prox}_{\frac{1}{\gamma_2} F_v}(\frac{1}{\gamma_2} \mathbf{z}_2^{(n)});$ 
8    $n \leftarrow n + 1;$ 

```

3.2. Optimization

To solve (1), we utilize the primal-dual splitting method [7, 10, 35] which brings an algorithmic solution to the following convex optimization problem: find \mathbf{x}^* in

$$\arg \min_{\mathbf{x} \in \mathcal{X}} f_1(\mathbf{x}) + f_2(\mathbf{L}\mathbf{x}), \quad (2)$$

where $f_1 : \mathcal{X} \rightarrow \overline{\mathbb{R}}$ and $f_2 : \mathcal{Y} \rightarrow \overline{\mathbb{R}}$ (\mathcal{X} and \mathcal{Y} are Euclidean spaces) are proper lower semicontinuous convex functions, the proximity operators of which are available, and $\mathbf{L} : \mathcal{X} \rightarrow \mathcal{Y}$ is a linear operator.

3.2.1 D-VTV case

In (2), let $\mathbf{L} := [(\mathbf{DC})^\top \ \Phi^\top]^\top$, $f_1 : \mathbb{R}^{3N} \rightarrow \overline{\mathbb{R}} : \mathbf{x} \mapsto \iota_{[0,255]^{3N}}(\mathbf{x})$, and $f_2 : \mathbb{R}^{6N+M} \rightarrow \overline{\mathbb{R}} : \mathbf{y} \mapsto \|\mathbf{y}_1\|_{1,2}^{(w,2,4)} + F_v(\mathbf{y}_2)$, where $\mathbf{y} = [\mathbf{y}_1^\top \ \mathbf{y}_2^\top]^\top$ ($\mathbf{y}_1 \in \mathbb{R}^{6N}$, $\mathbf{y}_2 \in \mathbb{R}^M$). Then (2) is reduced to (1) with J_{VTV}^w , and the proximity operators of f_1 and f_2 are computable using those of $\iota_{[0,255]^{3N}}$, $\|\cdot\|_{1,2}^{(k)}$ and F_v . The proximity operator of $\iota_{[0,255]^{3N}}$ is equal to the metric projection (see footnote 2) onto $[0, 255]^{3N}$, which can be calculated by pushing the entries into $[0, 255]$. For the computation of the proximity operators of $\|\cdot\|_{1,2}^{(k)}$ and F_v , see Remark 2.1 and 3.1. Consequently, applying the primal-dual splitting method to (2) with this setting produces a solver for (1) with J_{VTV}^w , as shown in Algorithm 1.

3.2.2 D-VTGV case

By noting that F_v in (1) with $J_{\text{VTGV}}^{w,\alpha}$ does not involve \mathbf{p} , we can remove the nested structure in (1) with $J_{\text{VTGV}}^{w,\alpha}$ (due to the infimal convolution), leading to finding $(\mathbf{u}^*, \mathbf{p}^*)$ in

$$\arg \min_{\mathbf{u} \in [0,255]^{3N}, \mathbf{p} \in \mathbb{R}^{6N}} \alpha \|\mathbf{DC}\mathbf{u} - \mathbf{p}\|_{1,2}^{(w_1,2,4)} + (1 - \alpha) \|\mathbf{G}\mathbf{p}\|_{1,2}^{(w_2,3,6)} + F_v(\Phi\mathbf{u}). \quad (3)$$

Let $\mathbf{L} := \begin{bmatrix} \mathbf{DC} & -\mathbf{I} \\ \mathbf{O} & \mathbf{G} \\ \Phi & \mathbf{O} \end{bmatrix}$, $f_1 : \mathbb{R}^{9N} \rightarrow \overline{\mathbb{R}} : \mathbf{x} \mapsto \iota_{[0,255]^{3N}}(\mathbf{x}_1)$, and $f_2 : \mathbb{R}^{15N+M} \rightarrow \overline{\mathbb{R}} : \mathbf{y} \mapsto \alpha \|\mathbf{y}_1\|_{1,2}^{(w_1,2,4)} + (1 -$

Algorithm 2: Solver for (1) with $J_{\text{VTGV}}^{w,\alpha}$

```

input :  $\mathbf{u}^{(0)}, \mathbf{p}^{(0)}, \mathbf{z}_i^{(0)} := [\mathbf{z}_{i,1}^{(0)\top} \ \mathbf{z}_{i,(2,3)}^{(0)\top}]^\top (i = 1, 2), \mathbf{z}_3^{(0)}$ 
output :  $\mathbf{u}^{(n)}$ 
1 while A stopping criterion is not satisfied do
2    $\mathbf{u}^{(n+1)} = P_{[0,255]^{3N}}(\mathbf{u}^{(n)} - \gamma_1(\mathbf{C}^\top \mathbf{D}^\top \mathbf{z}_1^{(n)} + \Phi^\top \mathbf{z}_3^{(n)}));$ 
3    $\mathbf{p}^{(n+1)} = \mathbf{p}^{(n)} - \gamma_1(-\mathbf{z}_1^{(n)} + \mathbf{G}^\top \mathbf{z}_2^{(n)});$ 
4    $[\mathbf{z}_{1,1}^{(n)\top} \ \mathbf{z}_{1,(2,3)}^{(n)\top}]^\top \leftarrow$ 
    $\mathbf{z}_1^{(n)} + \gamma_2(\mathbf{DC}(2\mathbf{u}^{(n+1)} - \mathbf{u}^{(n)}) - (2\mathbf{p}^{(n+1)} - \mathbf{p}^{(n)}));$ 
5    $[\mathbf{z}_{2,1}^{(n)\top} \ \mathbf{z}_{2,(2,3)}^{(n)\top}]^\top \leftarrow \mathbf{z}_2^{(n)} + \gamma_2 \mathbf{G}(2\mathbf{p}^{(n+1)} - \mathbf{p}^{(n)});$ 
6    $\mathbf{z}_3^{(n)} \leftarrow \mathbf{z}_3^{(n)} + \gamma_2 \Phi(2\mathbf{u}^{(n+1)} - \mathbf{u}^{(n)});$ 
7    $\mathbf{z}_{1,1}^{(n+1)} = \mathbf{z}_{1,1}^{(n)} - \gamma_2 \text{prox}_{\frac{\alpha w_1}{\gamma_2} \|\cdot\|_{1,2}^{(2)}}(\frac{1}{\gamma_2} \mathbf{z}_{1,1}^{(n)});$ 
8    $\mathbf{z}_{1,(2,3)}^{(n+1)} = \mathbf{z}_{1,(2,3)}^{(n)} - \gamma_2 \text{prox}_{\frac{\alpha}{\gamma_2} \|\cdot\|_{1,2}^{(4)}}(\frac{1}{\gamma_2} \mathbf{z}_{1,(2,3)}^{(n)});$ 
9    $\mathbf{z}_{2,1}^{(n+1)} = \mathbf{z}_{2,1}^{(n)} - \gamma_2 \text{prox}_{\frac{(1-\alpha)w_2}{\gamma_2} \|\cdot\|_{1,2}^{(3)}}(\frac{1}{\gamma_2} \mathbf{z}_{2,1}^{(n)});$ 
10   $\mathbf{z}_{2,(2,3)}^{(n+1)} = \mathbf{z}_{2,(2,3)}^{(n)} - \gamma_2 \text{prox}_{\frac{1-\alpha}{\gamma_2} \|\cdot\|_{1,2}^{(6)}}(\frac{1}{\gamma_2} \mathbf{z}_{2,(2,3)}^{(n)});$ 
11   $\mathbf{z}_3^{(n+1)} = \mathbf{z}_3^{(n+1)} - \gamma_2 \text{prox}_{\frac{1}{\gamma_2} F_v}(\frac{1}{\gamma_2} \mathbf{z}_3^{(n)});$ 
12   $n \leftarrow n + 1;$ 

```

$\alpha) \|\mathbf{y}_2\|_{1,2}^{(w_2,3,6)} + F_v(\mathbf{y}_3)$, where \mathbf{I} denotes an identity matrix of appropriate size, $\mathbf{x} = [\mathbf{x}_1^\top \ \mathbf{x}_2^\top]^\top$ ($\mathbf{x}_1 \in \mathbb{R}^{3N}$, $\mathbf{x}_2 \in \mathbb{R}^{6N}$) and $\mathbf{y} = [\mathbf{y}_1^\top \ \mathbf{y}_2^\top \ \mathbf{y}_3^\top]^\top$ ($\mathbf{y}_1 \in \mathbb{R}^{6N}$, $\mathbf{y}_2 \in \mathbb{R}^{9N}$, $\mathbf{y}_3 \in \mathbb{R}^M$). As the D-VTV case, (2) becomes equivalent to (3), i.e., (1) with $J_{\text{VTGV}}^{w,\alpha}$, and the proximity operators of f_1 and f_2 are also computable using those of $\iota_{[0,255]^{3N}}$, $\|\cdot\|_{1,2}^{(k)}$ and F_v . Thus we can solve (1) with $J_{\text{VTGV}}^{w,\alpha}$ by the primal-dual splitting method as Algorithm 2.

Proposition 3.2 (Convergence of Algorithm 1 and 2)

Suppose that $\gamma_i > 0$ ($i = 1, 2$) satisfy $\gamma_1 \gamma_2 \sigma_1(\mathbf{L}) \leq 1$ ($\sigma_1(\cdot)$ is the largest singular value of a matrix \cdot). Then the sequence generated by Algorithm 1 or 2 converges to a solution of (1) with J_{VTV}^w or $J_{\text{VTGV}}^{w,\alpha}$.

Proof: A direct consequence of [10, Theorem 3.3].

4. Experiments

We examine the utility of the D-VT(G)V by comparing with the existing VTVs listed in Tab. 1 on various scenarios. All experiments were performed using MATLAB (R2013a), on a Windows 7 (64bit) desktop computer with an Intel Core i7 2.8 GHz processor and 8.0 GB of RAM.³ We fixed γ_1 and γ_2 as 0.01 and $1/(12\gamma_1)$ in Algorithm 1 and 2, and set their stopping criteria as $\frac{\|\mathbf{u}^{(n+1)} - \mathbf{u}^{(n)}\|_2}{255} \leq 0.01$.

4.1. Denoising

We first consider simple Gaussian noise removal experiments. Clean test images are contaminated by an additive white Gaussian noise with standard deviation σ . Following the discussion in Remark 3.1, the ℓ_2 -norm ball is adopted

³ The MATLAB source code is available at <https://sites.google.com/site/thunsukeono/>.

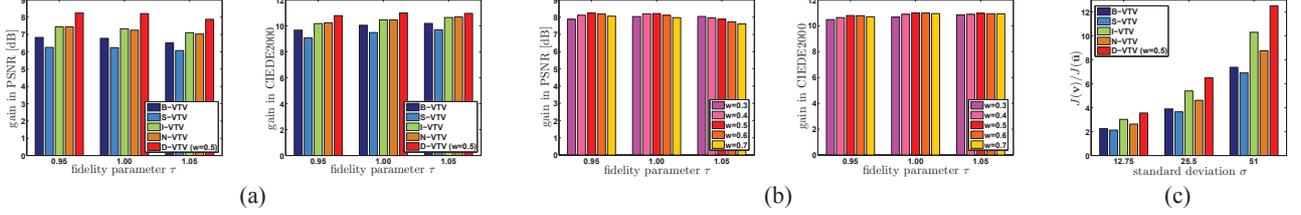


Figure 1: Quantitative comparisons on Gaussian denoising experiments: (a) Comparison of the existing VTVs and the D-VTV ($w = 0.5$) in terms of the average gains in PSNR [dB] and CIEDE2000 ($\sigma = 25.5$); (b) Comparison of the D-VTV with various w in terms of the average gains in PSNR [dB] and CIEDE2000 ($\sigma = 25.5$); and (c) Comparison of the existing VTVs and the D-VTV in terms of $J(\mathbf{v})/J(\mathbf{u}_{\text{org}})$ ($\sigma = 12.75, 25.5, \text{ and } 51$).

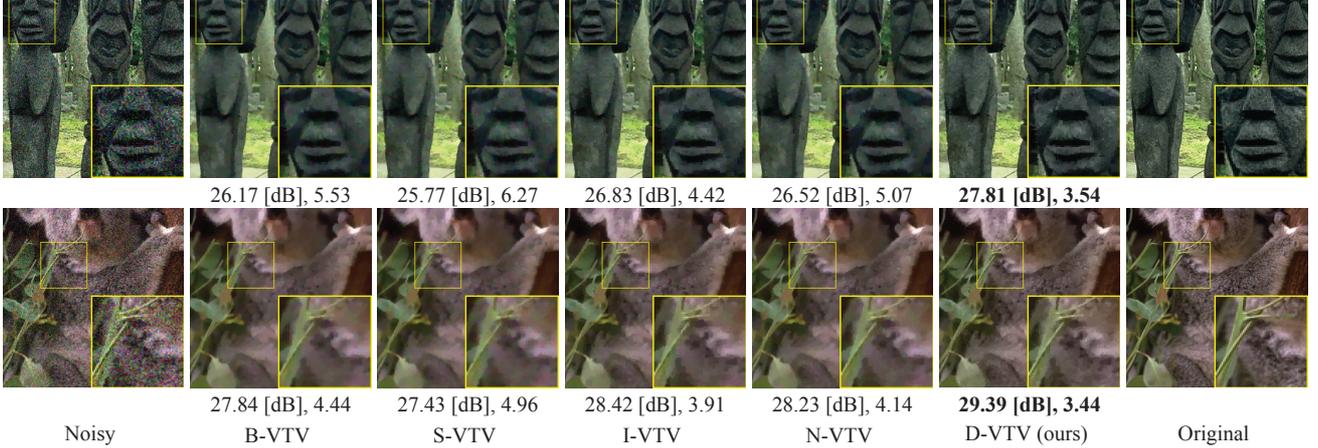


Figure 2: Resulting images obtained by using the existing VTVs and the D-VTV on Gaussian denoising experiments.

for the fidelity function, i.e., denoised images are obtained by solving the following problem: find \mathbf{u}^* in

$$\arg \min_{\mathbf{u} \in [0, 255]^{3N}} J(\mathbf{u}) \quad \text{s.t.} \quad \|\mathbf{u} - \mathbf{v}\|_2 \leq \varepsilon, \quad (4)$$

where J is the D-VT(G)V or one of the existing VT(G)V, and \mathbf{v} is a given noisy image. This constraint-type fidelity is convenient for giving a fair comparison of the VTVs, since a reasonable ε can be determined based only on a given σ (no matter what VT(G)V we use).

Quantitative comparisons of the existing VTVs and the D-VTV ($w = 0.5$) in terms of the average gains in PSNR (left) and CIEDE2000 [32] (right)⁴ are given in Fig. 1(a), which were conducted using 300 images taken from the *Berkeley Segmentation Database*⁵ [22]. The radius ε of the ℓ_2 -norm ball is set to $\varepsilon = \tau\sqrt{3N}\sigma^2$, where the fidelity parameter τ is selected as 0.95, 1, and 1.05, respectively (the smaller τ corresponds to the tighter fidelity). The use of the D-VTV results in the best average performance for every τ (about 1 [dB] (PSNR) and 0.7 (CIEDE2000) better than the best one among the existing VTVs). We also examined how the performance of the D-VTV varies with respect to the luminance weight w . The results are plotted in Fig. 1(b),

⁴ PSNR is defined by $10 \log_{10}(3N * 255^2 / \|\mathbf{u} - \mathbf{u}_{\text{org}}\|_2^2)$. CIEDE2000 (see [32] for the definition) is known as a good color quality assessment (Note: a smaller value of CIEDE2000 indicates a higher quality).

⁵ For each image, the center region of size 256×256 is cropped.

where w is increased from 0.3 to 0.7 by 0.1. As expected, a change in w affects the performance, but note that, for all the examined values of w , the D-VTV outperforms the existing VTVs (see (a) and (b)).

Some resulting images are shown in Fig. 2 with their corresponding PSNR [dB] (left) and CIEDE2000 (right). Notice that the images restored by using the D-VTV contain much less uneven color effects than those obtained by using the other VTVs and preserve details (such as body hair of the koala). They also indicate the best PSNR and CIEDE2000 values.

To demonstrate the suitability of the D-VTV as a prior for color images, we evaluated the value of each VTV both on clean and noisy images with different values of standard deviation. Specifically, since the scale of the values of VTVs are different from one another, we computed the ratio of each VTV on noisy images and that on clean images, i.e., $J(\mathbf{v})/J(\mathbf{u}_{\text{org}})$ (J stands for one VTV), for measuring how much the value is increased by noise, where \mathbf{v} and \mathbf{u}_{org} are noisy and clean images, respectively. Fig. 1(c) indicates the average of $J(\mathbf{v})/J(\mathbf{u}_{\text{org}})$ for each VTV based on the 300 images. The value of the D-VTV is rapidly increased by noise compared to those of the other VTVs, which implies that the D-VTV well distinguishes clean and noisy images, i.e., it is a very suitable prior for color images.

Average CPU times [sec] of each method are as follows:



Figure 3: Resulting images obtained by several VTVs and the corresponding VTGVs on a Gaussian denoising experiment.

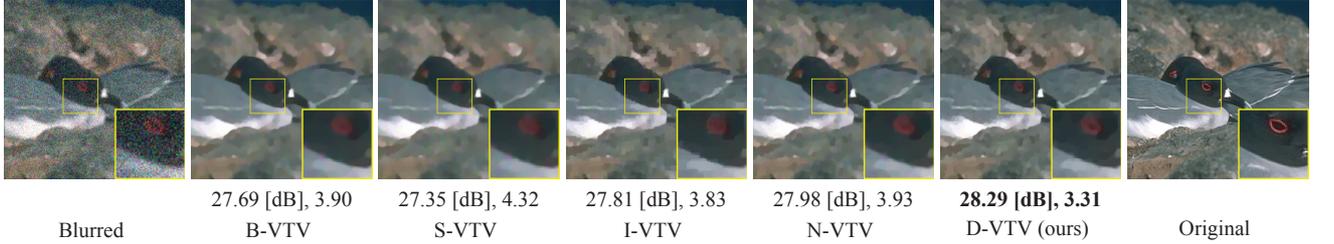


Figure 4: Resulting images obtained by using the existing VTVs and the D-VTV on a deblurring experiment.

the B-VTV, 1.53; the S-VTV, 72.21; the I-VTV, 88.68; the N-VTV, 71.76; and the D-VTV, 1.62. Note that since all the program codes were implemented by MATLAB and are not optimized, they can be accelerated by, e.g., parallel implementations using GPU.

We also compared the D-VTV (and several existing VTVs) with the D-VTGV (and the corresponding existing VTGVs). The parameter α is simply set to 0.5 for all the VTGVs (the B-VTGV [3], the I-VTGV [24], and the D-VTGV). The resulting images⁶ with their PSNR [dB] and CIEDE2000 values are given in Fig. 3, where we observe that the staircasing effect is significantly reduced by using the VTGVs; and that the D-VT(G)V effectively removes uneven color effects. As a result, the image restored by using the D-VTGV indicates the best PSNR and CIEDE2000, and seems to be close to the original image.

4.2. Deblurring

Next, we apply the D-VTV to a deblurring problem. Specifically, we solve the following problem: find \mathbf{u}^* in

$$\arg \min_{\mathbf{u} \in [0, 255]^{3N}} J(\mathbf{u}) \quad \text{s.t.} \quad \|\Phi \mathbf{u} - \mathbf{v}\|_2 \leq \varepsilon,$$

where Φ is a blur operator, and \mathbf{v} is a blurred image. The blur kernel is set to the 5×5 Gaussian blur with standard deviation 2, and then a white Gaussian noise is added ($\sigma = 25.5$), so that a suitable fidelity is the ℓ_2 -ball (ε is fixed to $0.95\sqrt{3N\sigma^2}$). The deblurring results with their PSNR and CIEDE2000 values are shown in Fig. 4 (the test image was taken from the Berkeley Segmentation Database). One sees that there are uneven color effects in the images obtained by using the B-VTV, the S-VTV, and the N-VTV. The deblurred result by using the I-VTV exhibits less uneven color effect but contains blocky artifacts around smooth

contours because of the anisotropy (see, e.g., the eye of the bird). The use of the D-VTV nicely resolves both uneven color effects and smooth contours.

4.3. Missing component recovery

We consider the recovery of a complete image from an image randomly missing 70% of its RGB channel components. The problem to be solved is finding \mathbf{u}^* in

$$\arg \min_{\mathbf{u} \in [0, 255]^{3N}} J(\mathbf{u}) \quad \text{s.t.} \quad \Phi \mathbf{u} = \mathbf{v},$$

where Φ is a random sampling operator, and \mathbf{v} is an incomplete observation (without noise). Here the hard constraint can be dealt with as the ℓ_2 -ball fidelity with the radius $\varepsilon = 0$. We depict the results in Fig. 5 with their PSNR and CIEDE2000 (the test images were taken from the Berkeley Segmentation Database). One can see that the D-VTV successfully reconstructs missing components without producing unnatural color changes that are observed in the resulting images obtained by using the existing VTVs.

4.4. Detail magnification

Aside from restoration, the D-VTV can be utilized for detail magnification. Details are extracted by solving the following problem: find \mathbf{u}^* in

$$\arg \min_{\mathbf{u} \in [0, 255]^{3N}} J(\mathbf{u}) \quad \text{s.t.} \quad \|\mathbf{u} - \mathbf{u}_{\text{org}}\|_2 \leq \varepsilon,$$

where \mathbf{u}^* is a processed image without details, so that the details are obtained as $\mathbf{u}_{\text{org}} - \mathbf{u}^*$. Here ε controls the quantity of details to be extracted. The magnification result is then generated by adding the magnified details $\beta(\mathbf{u}_{\text{org}} - \mathbf{u}^*)$ to \mathbf{u}^* (β denotes the magnification rate). Fig. 6 shows input images and the corresponding magnification results ($w = 0.01$, $\varepsilon = \sqrt{2.55 * 3N}$, and $\beta = 5$). The use of the B-VTV results in a standard high frequency enhancement. On the other hand, since the nature of the D-VTV enables us

⁶cc licensed (BY) flickr photo by ~@~ina (Irina Patrascu): http://flickr.com/photos/angel_ina/3201337190/

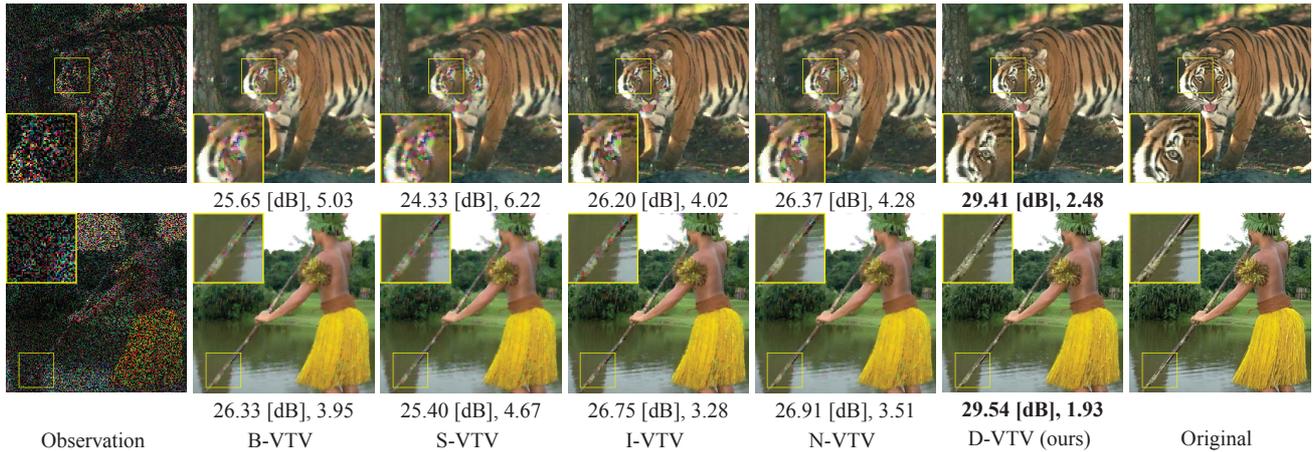


Figure 5: Resulting images obtained by using the existing VTVs and the D-VTV on missing component recovery.

to extract details with chrominance variation in preference to those with luminance one, we can magnify only color details by using the D-VTV, producing vivid color images.

4.5. Examples on a real-world image

We finally apply denoising using the existing VTVs and the D-VTV, i.e., solving (4), to a real-world image (480×480) with an ISO noise, which was taken by a cell phone camera with a high ISO speed. The radius ε was manually selected for each VTV in such a way as to produce a visually pleasing result. Fig. 7 shows the results. Uneven color effects appear in the result obtained by using the existing VTVs. By contrast, such effects are much reduced in the result based on the D-VTV, which seems to be the most preferable processed image among them.

5. Concluding Remarks

We have proposed a novel vectorial total variation prior named the D-VTV. Compared to existing VTVs, the D-VTV enjoys the capability of preventing uneven color effects and preserving meaningful details. It is also suitable for optimization thanks to its convexity and computational efficiency. Moreover, we have extended the D-VTV to a higher-order version called the D-VTVG, which overcomes the staircasing effect that accompanies the use of VTVs. We also presented a general form of convex optimization problems involving the D-VTV(G)V for imaging inverse problems, and provided efficient algorithmic solutions to the problems via a primal-dual splitting method.

We should remark that the color transform and norm involved in the D-VTV are not limited to the current ones, and a systematic performance analysis with various combinations of color transforms and norms is an important future work. Together with a recently proposed prior for color artifact reduction [27], the use of the D-VTV(G)V can lead to more effective color image restoration. Cartoon-texture decomposition using the D-VTV(G)V with various texture

modelings (e.g., [1, 12, 26, 25]) is also an interesting topic. The decorrelation idea of the D-VTV(G)V can be incorporated into other extensions of TV, such as the TV in light fields [18], in manifold [20], and the CS-specific TV [33].

Acknowledgements. We would like to thank the anonymous reviewers for their helpful comments. This work is supported in part by JSPS Grants-in-Aid for JSPS fellows (24-2522) and (B-21300091).

References

- [1] J. F. Aujol and S. H. Kang. Color image decomposition and restoration. *J. Vis. Commun. Image Rep.*, 17(4):916–928, 2006. 1, 7
- [2] P. Blomgren and T. F. Chan. Color TV: Total variation methods for restoration of vector valued images. *IEEE Trans. Image Process.*, 7(3):304–309, 1998. 1
- [3] K. Bredies. Recovering piecewise smooth multichannel images by minimization of convex functionals with total generalized variation penalty. SFB-Report, 2012. 2, 3, 6
- [4] K. Bredies, K. Kunisch, and T. Pock. Total generalized variation. *SIAM J. Imag. Sci.*, 3(3):92–526, 2010. 2, 3
- [5] X. Bresson and T. F. Chan. Fast dual minimization of the vectorial total variation norm and applications to color image processing. *Inverse Probl. Imag.*, 2(4):455–484, 2008. 1, 2, 3
- [6] A. Buades, B. Coll, and J.-M. Morel. A non-local algorithm for image denoising. In *Proc. IEEE CVPR*, 2005. 1
- [7] A. Chambolle and T. Pock. A first-order primal-dual algorithm for convex problems with applications to imaging. *J. Math. Imag. Vis.*, 40(1):120–145, 2010. 2, 4
- [8] X. Chen, S. B. Kang, J. Yang, and J. Yu. Fast patch-based denoising using approximated patch geodesic paths. In *Proc. IEEE CVPR*, 2013. 1
- [9] P. L. Combettes and J.-C. Pesquet. A Douglas-Rachford splitting approach to nonsmooth convex variational signal recovery. *IEEE J. Sel. Top. Signal Process.*, 1:564–574, 2007. 3
- [10] L. Condat. A primal-dual splitting method for convex optimization involving Lipschitzian, proximable and linear composite terms. *J. Optim. Theory Appl.*, 158(2):460–479, 2013. 2, 4
- [11] K. Dabov, A. Foi, V. Katkovnik, and K. Egiazarian. Image denoising by sparse 3-D transform-domain collaborative filtering. *IEEE Trans. Image Process.*, 16(8):2080–2095, 2007. 1
- [12] V. Duval, J. F. Aujol, and L. Vese. Mathematical modeling of textures : application to color image decomposition with a projected gradient algorithm. *J. Math. Imag. Vis.*, 37(3):232–248, 2010. 1, 2, 3, 7



Figure 6: Detail magnified results by using the B-VTV and the D-VTV.



Figure 7: Results on a real-world image.

- [13] D. Ferstl, C. Reinbacher, R. Ranftl, M. R  ther, and H. Bischof. Image guided depth upsampling using anisotropic total generalized variation. In *Proc. IEEE ICCV*, 2013. 3
- [14] A. Foi, V. Katkovich, and K. Egiazarian. Pointwise shape-adaptive dct for high-quality denoising and deblocking of grayscale and color images. *IEEE Trans. Image Process.*, 16(5):1395–1411, 2007. 2
- [15] G. Gilboa and S. Osher. Nonlocal operators with applications to image processing. *Multiscale Model. Simul.*, 7(3):1005–1028, 2009. 1
- [16] B. Goldluecke and D. Cremers. An approach to vectorial total variation based on geometric measure theory. In *Proc. IEEE CVPR*, 2010. 1, 2, 3
- [17] B. Goldluecke, E. Strelakovski, and D. Cremers. The natural vectorial total variation which arises from geometric measure theory. *SIAM J. Imag. Sci.*, 5(2):537–563, 2012. 1, 2, 3
- [18] B. Goldluecke and S. Wanner. The variational structure of disparity and regularization of 4D light fields. In *Proc. IEEE CVPR*, 2013. 7
- [19] S. Lefkimmiatis, A. Roussos, M. Unser, and P. Maragos. Convex generalizations of total variation based on the structure tensor with applications to inverse problems. In *Proc. SSVM*, 2013. 1, 2, 3
- [20] J. Lellmann, E. Strelakovski, S. Koetter, and D. Cremers. Total variation regularization for functions with values in a manifold. In *Proc. IEEE ICCV*, 2013. 7
- [21] J. Mairal, F. Bach, J. Ponce, G. Sapiro, and A. Zisserman. Non-local sparse models for image restoration. In *Proc. IEEE ICCV*, 2009. 1
- [22] D. Martin, C. Fowlkes, D. Tal, and J. Malik. A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics. In *Proc. IEEE ICCV*, 2001. 5
- [23] T. Miyata. Total variation defined by weighted L infinity norm for utilizing inter channel dependency. In *Proc. IEEE ICIP*, 2012. 1, 2, 3
- [24] T. Miyata. L infinity total generalized variation for color image recovery. In *Proc. IEEE ICIP*, 2013. 2, 3, 6
- [25] S. Ono, T. Miyata, and I. Yamada. Cartoon-texture image decomposition using blockwise low-rank texture characterization. *IEEE Trans. Image Process.*, 23(3):1128–1142, 2014. 7
- [26] S. Ono, T. Miyata, I. Yamada, and K. Yamaoka. Image recovery by decomposition with component-wise regularization. *IEICE Trans. Fundam.*, E95-A(12):2470–2478, 2012. 7
- [27] S. Ono and I. Yamada. A convex regularizer for reducing color artifact in color image recovery. In *Proc. IEEE CVPR*, 2013. 7
- [28] K. Plataniotis and A. Venetsanopoulos. *Color Image Processing and Applications*. Springer, 2000. 2
- [29] L. I. Rudin, S. Osher, and E. Fatemi. Nonlinear total variation based noise removal algorithms. *Phys. D*, 60(1-4):259–268, 1992. 1
- [30] G. Sapiro. Vector-valued active contours. In *Proc. IEEE CVPR*, 1996. 1
- [31] S. Setzer, G. Steidl, and T. Teuber. Infimal convolution regularizations with discrete ℓ_1 -type functionals. *Commun. Math. Sci.*, 9(3):797–827, 2011. 2, 3
- [32] G. Sharma, W. Wu, and E. N. Dalal. The CIEDE2000 color-difference formula: Implementation notes, supplementary test data, and mathematical observations. *Color Res. Appl.*, 30(1):21–30, 2005. 5
- [33] X. Shu and N. Ahuja. Hybrid compressive sampling via a new total variation TVL1. In *Proc. ECCV*, 2010. 7
- [34] T. Valkonen, K. Bredies, and F. Knoll. Total generalized variation in diffusion tensor imaging. *SIAM J. Imag. Sci.*, 6(1):487–525, 2013. 3
- [35] B. C. Vu. A splitting algorithm for dual monotone inclusions involving cocoercive operators. *Adv. Comput. Math.*, 38(3):667–681, 2013. 2, 4
- [36] G. Yu and G. Sapiro. DCT image denoising: a simple and effective image denoising algorithm. *Image Processing On Line*, 2011. 2
- [37] X. Zhang, M. Burger, X. Bresson, and S. Osher. Bregmanized nonlocal regularization for deconvolution and sparse reconstruction. *SIAM J. Imag. Sci.*, 3(3):253–276, 2010. 1