SYMMETRIC UPWIND SCHEME FOR DISCRETE WEIGHTED TOTAL VARIATION

Sonia Tabti, Julien Rabin, Abderrahim Elmoataz

SYMMETRIC UPWIND SCHEME FOR DISCRETE WEIGHTED TOTAL VARIATION

Sonia Tabti , Julien Rabin, Abderrahim Elmoataz

Normandie Univ, ENSICAEN, CNRS, GREYC, 6 Bd du Marchal Juin, Caen, France

ABSTRACT

This paper is devoted to the study of the discrete formulations of the weighted Total Variation (TV) based on upwind schemes that have been proposed for imaging problems in a local setting in [1] and in a non-local setting for graphs and point-clouds in [2]. We focus on two new symmetric formulations based on the ℓ^2 and ℓ^∞ norms respectively and propose a dedicated optimization algorithm to solve convex problems based on such TV penalties. We demonstrate the theoretical and practical interest of such formulations for image processing tasks.

***Index Terms*—** Weighted Total Variation, Upwind scheme, image denoising, isotropy preservation, co-area formula.

1. INTRODUCTION

Context In this paper, we are interested in the discrete formulation of Total Variation (TV) in image processing and computer vision as a prior for piecewise constant images. This topic has been widely studied since it has been proposed by Rudin-Osher-Fatemi (ROF) [3] and remains an active research field; as an example, non-local formulations have been proposed in [4], [5] and a combination of TV with Non-Local means has been proposed in [6], all in order to preserve textures and thin structures. The Total-Generalized-Variation (TGV) studied in [7] and the method in [8] are variants that reduce the staircasing effect observed with TV.

Total Variation is applied in many imaging and computer vision problems such as segmentation [9], inpainting, deconvolution and motion estimation [10], optical flow [11], filtering [12], point cloud processing [13] and classification [14]. Most studies are based on continuous formulations that are then approximated to solve the discrete problem. In this work, we consider directly the discrete setting.

Motivation In [1], a new discrete TV formulation is proposed, referred to in the following as TV_u, and inspired by upwind finite difference schemes used in the numerical resolution of Partial Differential Equations. Despite the upwind scheme [1] ability to better preserve isotropy in comparison with the usual ones, Condat shows in [15] that it is not invariant to contrast inversion. Consequently, some artifacts tends to appear with TV_u such as light isolated dots, which are not suitable especially for filtering purpose. Condat also proposes a new dual formulation of the Total Variation, denoted here by TV_c, which avoids this issue. However, the price to pay is that this formulation is not explicit. Note in addition that both schemes do not check the discrete co-area formula. This formula, verified by the usual anisotropic TV definition, allows to solve some non-convex problems by convex relaxation such as segmentation, see for example [9].

Contributions In this paper, we study two non-local TV formulations on graphs. This type of graph-based formulation was initially proposed in [2]. The proposed formulations rely on a symmetric definition combining an upwind scheme with a downwind scheme expressed in an ℓ^∞ and ℓ^2 norms respectively. We propose an efficient way to express these norms allowing us to solve convex optimization problems with these regularization terms and illustrate their properties on some numerical experiments. These properties are: isotropy preservation, bias reduction in comparison with TV_u, better filtering results in comparison with variational approaches of the same type and the versatility of the non-local ℓ^∞ formulation that can mimic other formulations.

Outline In Section 2, we first recall the local, non-local and upwind formulations of the Total Variation, then, we introduce the proposed symmetric formulations. In Section 3, the optimization algorithm to solve convex variational problems based on these regularization terms is described. Some comparative experiments in Section 4 illustrate the interest of the proposed framework.

2. DISCRETE UPWIND FORMULATIONS OF TOTAL VARIATION

2.1. Notations

A graph $\mathcal{G}(\mathcal{V}, \mathcal{E}, \omega)$ is defined by a set of vertices $\mathcal{V}$ connected by edges $\mathcal{E} \subseteq \mathcal{V} \times \mathcal{V}$ and a weight function $\omega : \mathcal{E} \mapsto \mathbb{R}_+$. A graph is symmetric, or undirected if $\omega(v,u) = \omega(u,v)$ $\forall (u,v) \in \mathcal{E}$. In practice, this property is usually satisfied for applications in imagery except in the domain borders because of special conditions such as Neumann or Dirichlet conditions. In the sequel, we consider a scalar image f defined on a regular cartesian grid $\Omega \subset \mathbb{Z}^2$, $u : x \in \Omega \mapsto f(x) \in \mathbb{R}$. The border of the domain is denoted by $\partial\Omega$. The associated n-

The authors thank ANR-14-CE27-0001 GRAPHSIP and FEDER/FSE 2014/2020 (GRAPHSIP project) for funding.

connexity graph is $\mathcal{G}(\Omega, \mathcal{E}, \omega)$, where: $\mathcal{E} = \{(x, x+t),\ \forall\, x \in \Omega \backslash \partial\Omega, t \in \mathcal{N}_n\}$. In practice, we use: $\mathcal{N}_2 = \{(1,0),(0,1)\}$ or:

$$\mathcal{N}_4 = \{(1,0),(-1,0),(0,1),(0,-1)\} \text{ with } \omega = 1,$$

or: $\mathcal{N}_8 = \{u = (i,j) \text{ s.t. } i,j \in \{0,1,-1\}\}$ with either $\omega(i,j) = 1/\|f(i) - f(j)\|^2$ or $\omega = 1$.
For a field $p : \mathcal{E} \mapsto \mathbb{R}$, the inner product is defined by: $\langle p,\, q\rangle = \sum_{(v,u)\in\mathcal{E}} p(v,u)q(v,u)$ and the canonical norm by: $\|p\| = \sqrt{\langle p,\, p\rangle}$. We also define the composed norm $\|.\|_{1,k}$:

$$\|p\|_{1,k} := \begin{cases} \displaystyle\sum_{v\in\mathcal{V}} \left(\sum_{u \text{ s.t } (u,v)\in\mathcal{E}} |p(v,u)|^k \right)^{\frac{1}{k}}, & \text{if } 1 \leqslant k < \infty \\ \displaystyle\sum_{v\in\mathcal{V}} \max_{u \text{ s.t } (u,v)\in\mathcal{E}} |p(v,u)|, & \text{if } k = \infty \end{cases} \tag{1}$$

which will be useful to unify the different formulations of the total variation studied here. The dual norm is:

$$\|p\|^*_{1,k} = \max_{\|q\|_{1,k}\leqslant 1} \langle p,\, q\rangle = \|p\|_{\infty,k^*} \quad \text{with} \quad \tfrac{1}{k} + \tfrac{1}{k^*} = 1. \tag{2}$$

We note $\mathbf{1}$ (respectively $\mathbf{0}$) the vector fields where each component is equal to 1 (resp. 0). The adjoint of a linear operator K is noted K^*, such that: $\langle Kp,\, q\rangle = \langle p,\, K^*q\rangle$. The operators $(p)_+ = \max\{\mathbf{0}, p\}$ and $(p)_- = \min\{\mathbf{0}, p\}$ apply a thresholding to each component of p. The inequality: $p \geqslant \mathbf{0}$, implies that each component of p is non-negative. The characteristic function of a set $\mathcal{C}$ is denoted by $\chi_{\mathcal{C}}$, and the indicator function $\mathbb{1}_{\mathcal{C}}$. The proximal operator of an application f is: $\mathrm{Prox}_f(x) := \underset{y}{\operatorname{argmin}}\ \frac{1}{2}\|x - y\| + f(y)$.

2.2. Upwind and downwind difference operators

We recall here the definition of difference operators on graphs as proposed in [2]. Without loss of generality and for sake of simplicity, we consider in the sequel fully connected graphs $\mathcal{E} = \mathcal{V} \times \mathcal{V}$, where the lack of an edge is represented by a zero weight ω. The centered difference operator of a vectorial function on $\mathcal{G}$:

$$\nabla_\omega f : v \in \mathcal{V} \mapsto \left(\sqrt{\omega(v,u)}(f(u) - f(v)) \right)_{u\in\mathcal{V}} \in \mathbb{R}^{\mathcal{V}}. \tag{3}$$

The upwind and downwind uncentered difference operators are then respectively defined as follow:

$$\begin{array}{lll} \nabla^-_\omega f : & v \in \mathcal{V} \mapsto (\nabla_\omega f(v))_- & \textit{(upwind)} \\ \nabla^+_\omega f : & v \in \mathcal{V} \mapsto (\nabla_\omega f(v))_+ & \textit{(downwind)} \end{array} \tag{4}$$

The adjoint of the difference operator $\nabla^*_\omega : \mathbb{R}^{\mathcal{E}} \to \mathbb{R}^{\mathcal{V}}$ is defined in the general case by:

$$\nabla^*_\omega p(v) = \sum_{u\in\mathcal{V}} \sqrt{\omega(u,v)}p(u,v) - \sum_{u\in\mathcal{V}} \sqrt{\omega(v,u)}p(v,u). \tag{5}$$

2.3. Usual formulations of TV and properties

Discrete co-area formula As mentioned in the introduction, an important property satisfied by the continuous formulation of the Total Variation is the co-area formula which allows for instance to establish a link between the notion of perimeter and the TV norm. In practice, it is interesting to check that the discrete formulation of TV also statisfies this property which is written as follows in our framework:

$$\mathsf{TV}(f) = \int_{\mathbb{R}} \mathsf{TV}\left(\chi_{\{f\geqslant t\}}\right) dt. \tag{6}$$

TV formulations Most of usual discrete formulations can be written this way:

$$\mathsf{TV}^\diamond_{\omega|k}(.) := \|\nabla^\diamond_\omega .\|_{1,k} \quad \text{with } \diamond \in \{\ , +, -\}. \tag{7}$$

If we consider the local definition of the scalar Total Variation on an n-connexity graph $\mathcal{G}(\Omega, \mathcal{E}, \omega)$ with $\omega = \mathbf{1}$, we get the following definitions: [1]

- $\mathsf{TV}_{|2}$ with $n = 2$: isotropic Total Variation [3],
- $\mathsf{TV}_{|1}$ with $n = 2$: anisotropic Total Variation,
- $\mathsf{TV}_u = \mathsf{TV}^-_{|2}$ with $n = 4$: *upwind* isotropic Total Variation [1].

In the more general case of Non-Local formulations (NL-TV), with any weight function ω, we get:

- $\mathsf{TV}_{\omega|2}$: isotropic Non-Local Total Variation [4, 13],
- $\mathsf{TV}_{\omega|k}$, $\mathsf{TV}^-_{\omega|k}$ and $\mathsf{TV}^+_{\omega|k}$ with $k \in [1,\infty)$ [2] .

Note that, among all previous definitions, only $\mathsf{TV}_{|1}$ and $\mathsf{TV}^\diamond_{\omega|1}$ check the discrete co-area formula (6).
Bias of uncentered upwind (and downwind) schemes Next proposition shows that, unlike usual centered formulations $\mathsf{TV}_{\omega|k}$, uncentered schemes, that is to say upwind and downwind schemes, are biased except in the case of undirected graphs with $k = 1$.

Proposition 1 (Contrast inversion)**.** *We have the contrast-inversion property* $\forall\ k \geqslant 1$ *and* $\forall\, \omega$*:*

$$\mathsf{TV}^-_{\omega|k}(-f) = \mathsf{TV}^+_{\omega|k}(f) \neq \mathsf{TV}^-_{\omega|k}(f). \tag{8}$$

If $k = 1$*, uncentered schemes verify:*

$$\mathsf{TV}_{\omega|1} = \tfrac{1}{2}\mathsf{TV}^-_{\omega|1} + \tfrac{1}{2}\mathsf{TV}^+_{\omega|1} \tag{9}$$

$$\mathsf{TV}^-_{\omega|1}(f) = \tfrac{1}{2}\mathsf{TV}_{\omega|1}(f) - \tfrac{1}{2}\langle f,\, \nabla^*_\omega \mathbf{1}\rangle \tag{10}$$

where the second term is equal to zero in the undirected graph case.

The lack of invariance to contrast inversion is illustrated in [15] where light pixels have a lower value of $\mathsf{TV}^-_{|2}$ than dark pixels, which produces noticeable denoising artifacts.

[1] up to a normalization factor related to the choice of the neighborhood $\mathcal{N}_n$.

2.4. Symmetric uncentered formulation

Inspired by the equation (9) which shows that bias can be removed by defining an uncentered symmetric scheme, we focus on a symmetric definition of Total Variation with the ℓ^∞ and the ℓ^2 norms.

Definition We consider the following formulation:

$$\mathsf{TV}^{\pm}_{\omega|k} := \tfrac{1}{2}\mathsf{TV}^{-}_{\omega|k} + \tfrac{1}{2}\mathsf{TV}^{+}_{\omega|k} \text{ for } k \in \{2, \infty\}. \tag{11}$$

This definition has been proposed in [16] in the case $k = \infty$ but has never been used in practice. Notice also that the case $k = 1$ is not interesting as we have, according to equation (9): $\mathsf{TV}^{\pm}_{\omega|1} = \mathsf{TV}_{\omega|1}$.

Properties of the proposed formulations The symmetric formulation (11) is a semi-norm and verifies the next proposition in the case $k = \infty$.

Proposition 2 (Co-area [16])**.** *A constant weight $\omega(v, u)$ in each vertex $v \in \mathcal{V}$ is a sufficient condition for $\mathsf{TV}^{\pm}_{\omega|\infty}$ to verify the discrete co-area formula* (6)*.*

Using the definitions in section 2.1, and in particular (2), we can show the next proposition:

Proposition 3 (Dual formulation of $\mathsf{TV}^{\pm}_{\omega|k}$ (11))**.**

$$\begin{aligned}\mathsf{TV}^{\pm}_{\omega|k}(f) = \max_{r_1 \geqslant 0, r_2 \geqslant 0} \tfrac{1}{2}\langle \nabla_\omega f,\, r_1 - r_2\rangle \\ - \chi_{\|r_1\|_{\infty,k^*} \leq 1} - \chi_{\|r_2\|_{\infty,k^*} \leq 1}\end{aligned}$$

where k^* verifies $\frac{1}{k} + \frac{1}{k^*} = 1$.

3. APPLICATION AND OPTIMIZATION

3.1. Inverse problems

In order to demonstrate the practical interest of the uncentered symmetric formulations proposed, we focus on the resolution of inverse problems as in the ROF model [3] :

$$\min_f \tfrac{1}{2}\|f - g\|^2 + \lambda \mathsf{TV}^{\pm}_{\omega|k}(f). \tag{12}$$

where g is an image corrupted with Gaussian noise, f is the image to be estimated and λ the regularization parameter.

3.2. Dual formulation

Using proposition 3, we obtain the following primal-dual problem:

$$\min_{f \in \mathbb{R}^{\mathcal{V}}} \max_{r=(r_1,r_2)} \tfrac{1}{2}\|f - g\|^2 + \langle Kf,\, r\rangle - \chi_{r \in \mathcal{S}_{k^*}}, \tag{13}$$

where $K = \frac{\lambda}{2}\left[\nabla_w; -\nabla_w\right]$ so that $\|K\|^2 \leqslant 8\lambda$ with neighborhood $\mathcal{N}_2$ (see [17]) and $\|K\|^2 \leqslant 16\lambda$ with neighborhood $\mathcal{N}_4$. The convex set $\mathcal{S}_{k^*}$ is defined as follows:

$$\mathcal{S}_{k^*} = \{(r_1, r_2) : r_i \in \mathbb{R}_+^{\mathcal{E}}\,, \|r_i\|_{\infty,k^*} \leq 1 \;\forall\, i \in \{1, 2\}\}.$$

3.3. Algorithm

Problem (13) can be solved with the Chambolle-Pock algorithm [10] by following these steps at each iteration t:

$$\begin{cases} r^{(t+1)} & = \mathrm{Proj}_{\mathcal{S}_{k^*}}(r^{(t)} + \gamma^{(t)} K \bar{f}^{(t)}) \\ f^{(t+1)} & = \mathrm{Prox}_{\frac{\tau^{(t)}}{2}\|\cdot - g\|^2}(f^{(t)} - \tau^{(t)} K^* r^{(t+1)}) \\ \bar{f}^{(t+1)} & = f^{(t+1)} + \theta^{(t)}(f^{(t+1)} - f^{(t)}) \\ \text{where} & \quad \theta^{(t)} = 1/\sqrt{1 + 2\rho\tau^{(t)}}, \tau^{(t+1)} = \theta^{(t)}\tau^{(t)}, \gamma^{(t+1)} = \gamma^{(t)}/\theta^{(t)} \end{cases} \tag{14}$$

with these conditions and parameters setting:
$\tau^0, \gamma^0 > 0, \theta^0 \in [0, 1], \bar{f}^0 = g$, $\rho = 0.7/\lambda, \tau^0 = 0.99/\|K\|$, $\tau^{(t)}\gamma^{(t)}\|K\|^2 < 1$.
The proximal operator is (see for example [10]):

$$\mathrm{Prox}_{\frac{\tau}{2}\|\cdot - g\|^2}(f) = \tfrac{1}{1+\tau}(f + \tau g).$$

Next proposition proves that, in order to perform the projection on $\mathcal{S}_{k^*}$, one should perform a projection onto the non-negative portion of the ℓ^{k^*} unit ball. Due to lack of space, the proof of this proposition does not appear in this paper.

Proposition 4 (Projection on $\mathcal{S}_{k^*}$)**.** *For $k^* \in \{1, 2\}$*

$$\mathrm{Proj}_{\mathcal{S}_{k^*}}(r_1, r_2) = (\hat{r}_1, \hat{r}_2) \;\; \textit{where} \;\; \hat{r}_i(v) = \mathrm{Proj}_{\|\cdot\|_{k^*} \leq 1}\left((r_i(v))_+\right),$$

For $i \in \{1, 2\}$. Hence, for $k^* = 2$, we obtain the same result as [1] with: $\mathrm{Proj}_{\|\cdot\| \leqslant 1}(x) = \frac{x}{\max\{1, \|x\|\}}$. For $k^* = 1$, the projection on the ℓ_1-ball can be computed with linear complexity (see for example [20]).

4. EXPERIMENTS AND DISCUSSION

4.1. Study of shape preservation

Figure 2 shows the regularization results of a synthetic image for the various definitions of TV considered previously. Each shape represents an ℓ^k ball and corresponds to the structural element of the adjoint $\mathsf{TV}_{|k^*}$ formulation studied here (see *e.g.* [21] for $\mathsf{TV}_{|2}$). As expected, the proposed isotropic formulation $\mathsf{TV}^{\pm}_{|2}$ provides a satisfying result on the disc shape, like TV_u and TV_c. In contrast, $\mathsf{TV}_{|2}$ produces some blur in a particular direction, as already reported in [1, 15]. This observation stands for the other symmetric formulation considering the diamond for $\mathsf{TV}^{\pm}_{|\infty}$ (in comparison with $\mathsf{TV}_{|\infty}$). Even more interestingly, the result obtained with $\mathsf{TV}^{\pm}_{\omega|\infty}$ for a larger neighborhood ($\mathcal{N}_8$ instead of $\mathcal{N}_4$) and a well chosen weight distribution, $\omega = \frac{1}{\|\cdot\|^2}$, "mimic" the behavior of $\mathsf{TV}^{\pm}_{|2}$. Last, observe that the contrast is better preserved with $\mathsf{TV}^{\pm}_{|\infty}$ in comparison with all these approaches.

4.2. Denoising application: Bias removal

Figure 1 illustrates the interest of the proposed formulations in the context of a simple application: the denoising

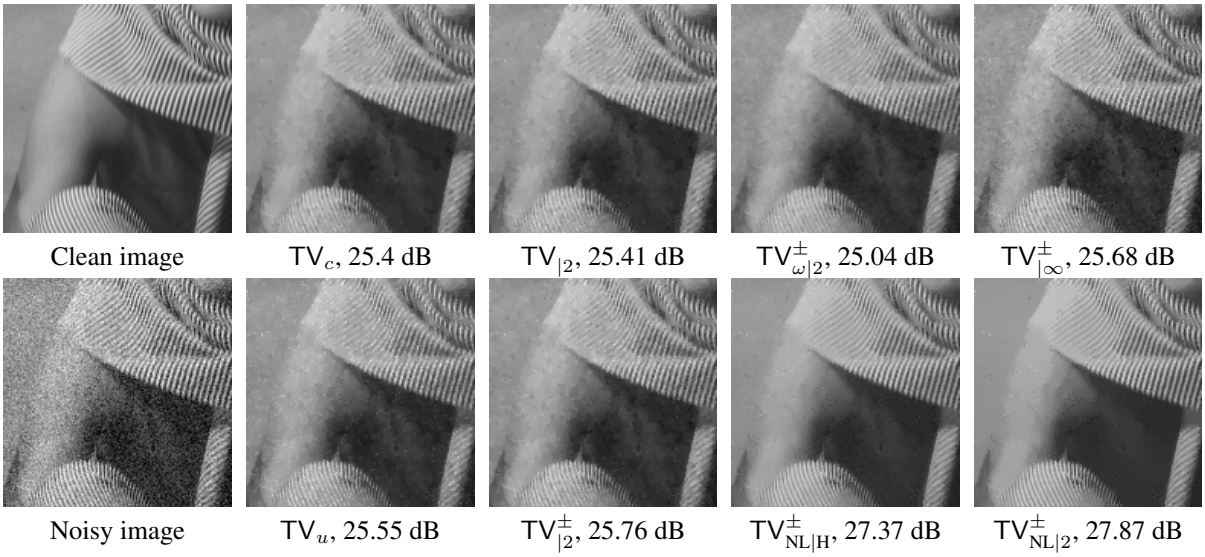

Fig. 1. Denoising results with Gaussian noise ($\sigma = 20$), using 1000 iterations, $\lambda = 0.05$ and $\omega = 1/\|f(u) - f(v)\|^2$. The proposed symmetric formulations do not suffer from artefacts, unlike the upwind scheme TV_u, which results in a better PSNR espcially with the Non-Local (NL) weight computed computed as in [18]. Staircasing effect is reduced by using $\mathsf{TV}_{|\mathrm{H}}$, the Huber TV, see for instance [19]. Note that the results obtained with $\mathsf{TV}^{\pm}_{|2}$ and $\mathsf{TV}^{\pm}_{|\infty}$ can be improved with a higher value of λ but the PSNR obtained with the NL weight will still be the highest.

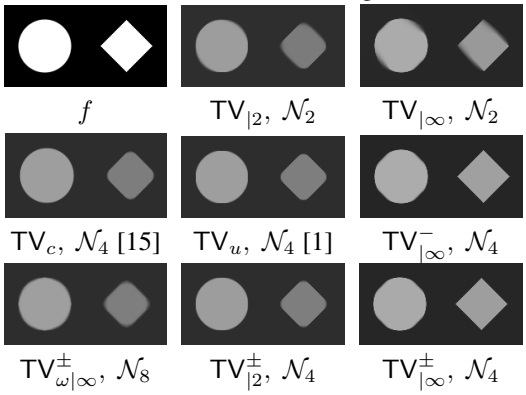

Fig. 2. Regularizations of the image f with 4000 iterations and $\lambda = 6$.

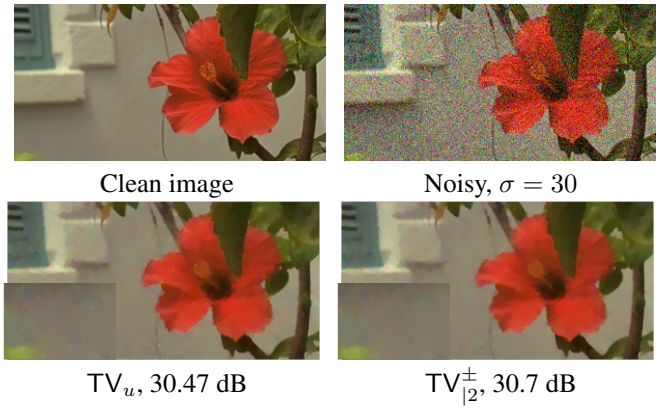

Fig. 3. Color image denoising results with zooms at the bottom left parts with 1000 iterations and $\lambda = 0.16$.

problem (12). In fact, as mentioned in Section 2.2, we observe that TV_u produces a bias represented by light isolated pixels. The proposed regularization terms as well as TV_c, do not suffer from such bias and manage to provide a better PSNR in comparison with the other schemes in most cases. We can also observe that the result obtained with the Non-local formulation is really improved visually and textures are well preserved. Moreover, with Huber TV, the staircasing effect is well reduced. The same benefits are also observed for color images (Figure 3). The vectorial TV definition used is described in [22].

5. CONCLUSION

In this paper, we proposed two uncentered and symmetric discrete TV schemes on weighted graphs with an ℓ_2 and an ℓ_∞ norm respectively. We also developed an efficient optimization algorithm to adress generic problems that may benefit from such new formulations. These terms are semi-norms with an explicit formula unlike the framework proposed in [15]. We illustrated in a denoising application that both proposed formulations remove the bias observed with TV_u [1]. An extension to color image filtering has also been proposed in the case of the ℓ_2 norm with the same advantages as the grayscale version. Our formulations are generic in the sense that the neighborhood $\mathcal{N}_n$ and the weights ω can be defined arbitrarily and adapted to specific problems. This property allows the ℓ^∞ formulation to be able to mimic the properties of other formulations. We also recall that $\mathsf{TV}^{\pm}_{\mathbf{1}|\infty}$ verifies the co-area formula which allows to solve some non-convex optimization problems.

6. REFERENCES

[1] Antonin Chambolle, Stacey E. Levine, and Bradley J. Lucier, "An upwind finite-difference method for total variation-based image smoothing," *SIAM Journal on Imaging Sciences*, vol. 4, no. 1, pp. 277–299, 2011.

[2] Abderrahim Elmoataz, Olivier Lezoray, and Sébastien Bougleux, "Nonlocal discrete regularization on weighted graphs: a framework for image and manifold processing," *IEEE transactions on Image Processing*, vol. 17, no. 7, pp. 1047–1060, 2008.

[3] Leonid I Rudin, Stanley Osher, and Emad Fatemi, "Nonlinear total variation based noise removal algorithms," *Physica D: Nonlinear Phenomena*, vol. 60, no. 1-4, pp. 259–268, 1992.

[4] Guy Gilboa and Stanley Osher, "Nonlocal operators with applications to image processing," *Multiscale Modeling & Simulation*, vol. 7, no. 3, pp. 1005–1028, 2008.

[5] Zhi Li, Francois Malgouyres, and Tieyong Zeng, "Smoothing the finite differences defining the non-local total variation and application in image restoration," *Journal of Mathematical Imaging and Vision*, vol. 59, pp. 296–317, 2017.

[6] Camille Sutour, Charles-Alban Deledalle, and Jean-François Aujol, "Adaptive regularization of the NL-means: Application to image and video denoising," *IEEE Transactions on image processing*, vol. 23, no. 8, pp. 3506–3521, 2014.

[7] Kristian Bredies, Karl Kunisch, and Thomas Pock, "Total generalized variation," *SIAM Journal on Imaging Sciences*, vol. 3, no. 3, pp. 492–526, 2010.

[8] Cécile Louchet and Lionel Moisan, "Total variation denoising using iterated conditional expectation," in *Proceedings of the 22nd European Signal Processing Conference (EUSIPCO)*. IEEE, 2014, pp. 1592–1596.

[9] Tony F Chan, Selim Esedoglu, and Mila Nikolova, "Algorithms for finding global minimizers of image segmentation and denoising models," *SIAM journal on applied mathematics*, vol. 66, no. 5, pp. 1632–1648, 2006.

[10] Antonin Chambolle and Thomas Pock, "A first-order primal-dual algorithm for convex problems with applications to imaging," *Journal of Mathematical Imaging and Vision*, vol. 40, no. 1, pp. 120–145, 2011.

[11] René Ranftl, Kristian Bredies, and Thomas Pock, "Non-local total generalized variation for optical flow estimation," in *Proceedings of ECCV'14*, Cham, 2014, pp. 439–454, Springer International Publishing.

[12] Guy Gilboa, "A total variation spectral framework for scale and texture analysis," *SIAM Journal on Imaging Sciences*, vol. 7, no. 4, pp. 1937–1961, 2014.

[13] Moncef Hidane, Olivier Lézoray, and Abderrahim Elmoataz, "Nonlinear multilayered representation of graph-signals," *Journal of Mathematical Imaging and Vision*, vol. 45, no. 2, pp. 114–137, 2013.

[14] Xavier Bresson, Thomas Laurent, David Uminsky, and James Von Brecht, "Multiclass total variation clustering," in *Advances in Neural Information Processing Systems*, 2013, pp. 1421–1429.

[15] Laurent Condat, "Discrete total variation: New definition and minimization," *preprint hal-01309685*, 2016.

[16] Daniel Tenbrinck, François Lozes, and Abderrahim Elmoataz, "Solving minimal surface problems on surfaces and point clouds," in *International Conference on Scale Space and Variational Methods in Computer Vision*. Springer, 2015, pp. 601–612.

[17] Antonin Chambolle, "An algorithm for total variation minimization and applications," *Journal of Mathematical imaging and vision*, vol. 20, no. 1, pp. 89–97, 2004.

[18] Antoni Buades, Bartomeu Coll, and J-M Morel, "A non-local algorithm for image denoising," in *IEEE Computer Society Conference on CVPR*. IEEE, 2005, vol. 2, pp. 60–65.

[19] Mila Nikolova, "Local strong homogeneity of a regularized estimator," *SIAM Journal on Applied Mathematics*, vol. 61, no. 2, pp. 633–658, 2000.

[20] Laurent Condat, "Fast projection onto the simplex and the l1 ball," *Mathematical Programming*, vol. 158, no. 1, pp. 575–585, 2016.

[21] Tony Chan, Selim Esedoglu, Frederick Park, and A Yip, "Recent developments in total variation image restoration," *Mathematical Models of Computer Vision*, vol. 17, no. 2, 2005.

[22] Xavier Bresson and Tony F Chan, "Fast dual minimization of the vectorial total variation norm and applications to color image processing," *Inverse problems and imaging*, vol. 2, no. 4, pp. 455–484, 2008.