Generalized Total Variation: Tying the Knots

Ivan W. Selesnick

***Abstract*—This paper formulates a convex generalized total variation functional for the estimation of discontinuous piecewise linear signals from corrupted data. The method is based on (1) promoting pairwise group sparsity of the second derivative signal and (2) decoupling the principle knot parameters so they can be separately weighted. The proposed method refines the recent approach by Ongie and Jacob.**

I. Introduction

Algorithms for signal estimation (denoising, restoration, and reconstruction) rely on some signal model, even if the model is implicit. Algorithms based on total variation (TV) regularization assume the signal of interest is piecewise constant; i.e., its derivative is sparse [30]. TV regularization is widely used in sparse signal processing. However, the piecewise constant signal model is often unrealistic; hence, several generalizations of TV have been proposed [1], [3], [4], [6], [14]–[17], [21], [25], [33], [34].

Recently, Ongie and Jacob proposed a generalization of TV for the purpose of estimating discontinuous piecewise polynomial signals [28]. In this model, the signal consists of polynomial segments partitioned by points (*knots*) where the signal is discontinuous. This model is of relevance because signals in various domains exhibit discontinuities; for example (1) edges of objects in images may produce discontinuities in pixel values and (2) arrivals of particles in biosensors may produce discontinuities in measured signals [7].

Following Ongie and Jacob, we propose a generalized TV (GTV) functional for the purpose of estimating discontinuous piecewise linear signals. The GTV functional, proposed by Ongie and Jacob, is based on group sparsity of the second derivative [28]. Our approach builds upon this approach, but differs in two ways. First, we apply a linear transform to each group in the group-sparse representation of the second derivative. This linear transform serves to separate the two principle knot parameters so they can be separately weighted. Second, we promote group sparsity using a synthesis formulation rather than an analysis formulation of group sparsity.

Ongie and Jacob describe both convex and non-convex forms of their proposed GTV functional, but emphasize the non-convex form. Experiments show that the convex form of their proposed GTV functional tends to miss signal discontinuities. The two refinements we propose lead to a new convex GTV functional that more accurately recovers discontinuities.

The author is with the Department of Electrical and Computer Engineering, School of Engineering, New York University, 6 Metrotech Center, Brooklyn, NY 11201, USA. Email: selesi@nyu.edu.

MATLAB software available at http://eeweb.poly.edu/iselesni/GTVknots

A. Tying the Knots

Figure 1 illustrates the difficulty in accurately recovering discontinuities of piecewise linear signals by convex regularization. Figures 1(a) and 1(b) illustrate a representative discontinuous signal (with and without noise). The signal obtained by TV denoising, i.e.,

$$\hat{x} = \arg\min_x \left\{ \frac{1}{2} \sum_n (y_n - x_n)^2 + \lambda \sum_n |x_n - x_{n-1}| \right\} \quad (1)$$

where y is the noisy data, accurately recovers the discontinuity but not the linear behavior of the signal (c.f., staircase artifacts, Fig. 1(c)). On the other hand, the signal obtained by second-order TV denoising, i.e., minimizing

$$\begin{aligned} &\frac{1}{2} \sum_n (y_n - x_n)^2 + \lambda \sum_n |x_n - 2x_{n-1} + x_{n-2}| \\ &\qquad = \frac{1}{2} \|y - x\|_2^2 + \lambda \|Dx\|_1 \end{aligned} \quad (2)$$

where D is the second-order difference operator, accurately recovers the linear behavior of the signal but not the discontinuity (Fig. 1(d)). Instead of a single knot at the discontinuity, the denoised signal possesses two knots, identifiable as isolated negative and positive impulses in the second-order difference signal. These two impulses are separated by a gap.

To accurately recover the discontinuity, the denoised signal should possess a single knot at the discontinuity (hence, the second-order difference signal should exhibit a positive-negative impulse pair). The two knots in the second-order TV solution should be brought (tied) together. To achieve this behavior, it is reasonable to use a regularization functional that promotes pairwise group sparsity [28]. The denoised signal obtained using this approach indeed exhibits group sparsity in its second-order difference signal (Fig. 1(e)). However, the denoised signal is very similar to the one obtained using second-order TV (Fig. 1(d)). In particular, the discontinuity is not more accurately recovered.

A more accurate recovery of the discontinuity is possible using a non-convex formulation [28]; but this raises the question: can the discontinuity be more accurately recovered using some other convex functional? A positive answer is demonstrated by the denoised signal, shown in Fig. 1(f), obtained using the convex functional (20) proposed below. It accurately recovers both the discontinuity and linear behavior of the signal.

II. Knot Parameters

To introduce the approach, consider the discrete-time signal

$$x_n = \begin{cases} a_1 n + a_0, & n < 0 \\ b_1 n + b_0, & n \geqslant 0 \end{cases} \quad (3)$$

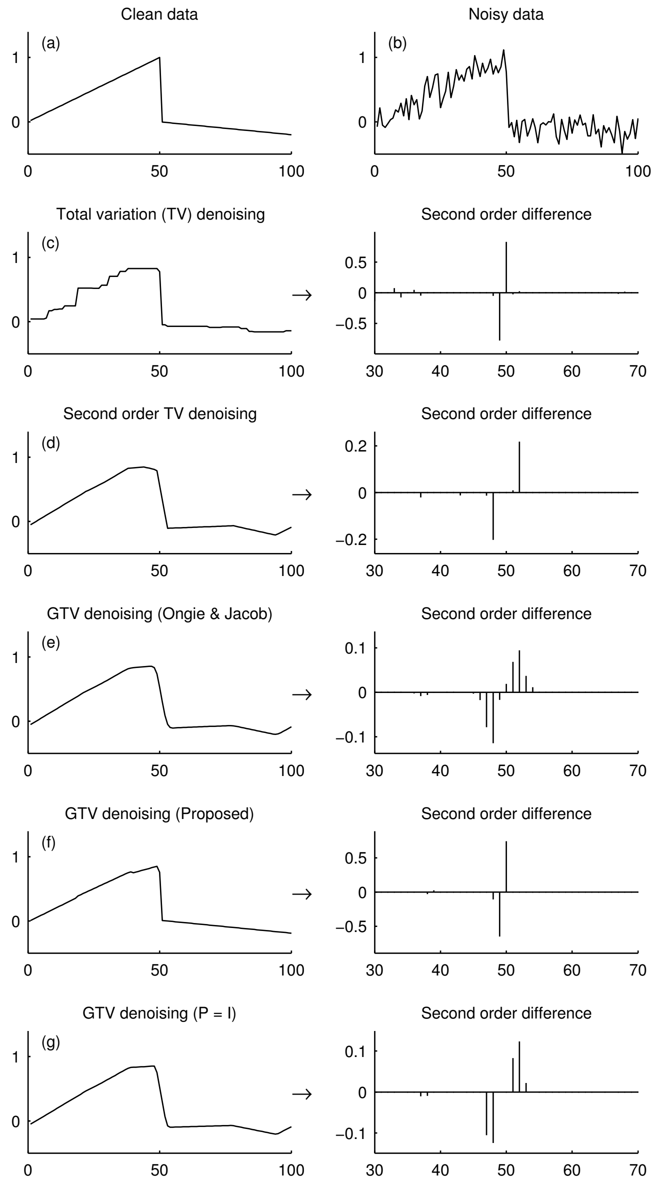

Fig. 1. Comparison of denoising using several convex TV functionals. (a) Noise-free signal. (b) Noisy signal. (c) TV. (d) Second-order TV. (e) GTV [28]. (f) GTV (proposed). (g) GTV ($P = I_2$).

illustrated in Fig. 2. Informally, this signal exhibits a discontinuity in both its value and its slope. Let v be the discrete-time second derivative

$$v_n = x_n - 2x_{n-1} + x_{n-2} \tag{4}$$

illustrated in Fig. 2. Using (3), the signal v is given by

$$v_0 = b_0 - a_0, \tag{5a}$$
$$v_1 = (b_1 - a_1) - (b_0 - a_0), \tag{5b}$$
$$v_n = 0, \; n \notin \{0, 1\}. \tag{5c}$$

Not only is v group-sparse, but its non-zero values are correlated. To clarify, define two functions on the real line,

$$x_{\text{left}}(t) = a_1 t + a_0, \qquad x_{\text{right}}(t) = b_1 t + b_0, \tag{6}$$

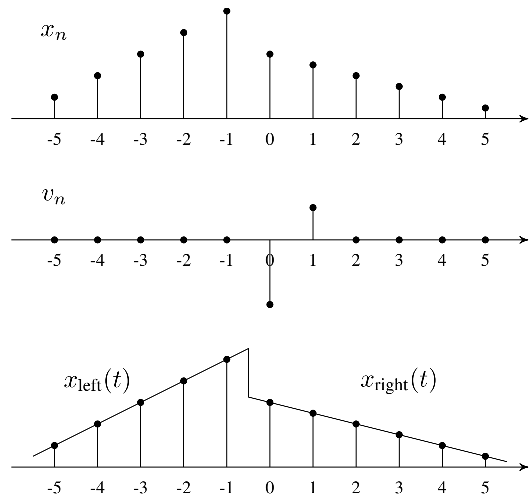

Fig. 2. Discontinuous piecewise linear signal with a knot at $t = -1/2$.

as illustrated in Fig. 2. Both the value and slope of the function are discontinuous at the knot $t = -1/2$. The amplitudes of the discontinuities are given respectively by

$$\Delta_{\text{value}} = x_{\text{right}}(-0.5) - x_{\text{left}}(-0.5) \tag{7}$$
$$= (b_0 - a_0) - (b_1 - a_1)/2 \tag{8}$$

and

$$\Delta_{\text{slope}} = x'_{\text{right}}(-0.5) - x'_{\text{left}}(-0.5) \tag{9}$$
$$= b_1 - a_1. \tag{10}$$

We consider Δ_{value} and Δ_{slope} to be the principle knot parameters. From (5), it can be observed that

$$\begin{aligned} v_0 &= \Delta_{\text{slope}}/2 + \Delta_{\text{value}} \\ v_1 &= \Delta_{\text{slope}}/2 - \Delta_{\text{value}}. \end{aligned} \tag{11}$$

Both v_0 and v_1 depend on both knot parameters. For a given class of signals, it is likely that Δ_{value} and Δ_{slope} exhibit quite distinct amplitude ranges. Hence, they should be weighted differently. From (11), we write

$$\begin{aligned} \Delta_{\text{slope}} &= v_0 + v_1 \\ 2\Delta_{\text{value}} &= v_0 - v_1. \end{aligned} \tag{12}$$

That is, the change in *value* and *slope* can be decoupled and separately weighted.

In light of this, our approach is based on un-mixing the knot parameters, Δ_{slope} and Δ_{value}, and using different weights for them (in addition to promoting group sparsity). This is accomplished with the matrix P in (21) below.

III. Definitions

The elements of the discrete-time signal $x \in \ell^2(\mathbb{Z})$ are denoted x_n or $[x]_n$. We define the second-order difference operator $D \colon \ell^2(\mathbb{Z}) \to \ell^2(\mathbb{Z})$ by

$$[Dx]_n = x_n - 2x_{n-1} + x_{n-2}. \tag{13}$$

The operator D is a discrete approximation of the second derivative. We define $B_n \colon \ell^2(\mathbb{Z}) \to \mathbb{R}^2$ by

$$B_n(r) = [r_{2n}, \; r_{2n+1}]^{\mathsf{T}}. \tag{14}$$

The operator B_n extracts non-overlapping two-point blocks.

To promote group sparsity, we use a synthesis formulation of group sparsity [26]. First, we define $S\colon \ell^2(\mathbb{Z}) \to \ell^2(\mathbb{Z})$ by

$$[Sr]_n = r_{2n-1} + r_{2n}. \tag{15}$$

Next, we define the group-sparse regularizer $R\colon \ell^2(\mathbb{Z}) \to \mathbb{R}$,

$$\begin{aligned} R(v) = \min_r \; & \sum_n \|B_n(r)\|_2 \\ \text{such that } \; & v = Sr. \end{aligned} \tag{16}$$

In the functional (16), the operator S synthesizes the signal v as a sum of overlapping two-point blocks. For a finite-length signal v, we truncate the operator S to a finite matrix. For example, a four-point signal v can be written as a sum of three overlapping blocks (groups) as

$$v = Sr \Leftrightarrow \begin{bmatrix} v_0 \\ v_1 \\ v_2 \\ v_3 \end{bmatrix} = \begin{bmatrix} 1 & & & & & \\ & 1 & 1 & & & \\ & & & 1 & 1 & \\ & & & & & 1 \end{bmatrix} \begin{bmatrix} r_0 \\ r_1 \\ r_2 \\ r_3 \\ r_4 \\ r_5 \end{bmatrix}, \tag{17}$$

where the three groups are

$$B_0(r) = \begin{bmatrix} r_0 \\ r_1 \end{bmatrix}, \quad B_1(r) = \begin{bmatrix} r_2 \\ r_3 \end{bmatrix} \quad B_2(r) = \begin{bmatrix} r_4 \\ r_5 \end{bmatrix}. \tag{18}$$

The vector r is the group-sparse *representation* of the signal v. The cost function to be minimized in (16) is a mixed ℓ_2–ℓ_1 norm [18]; it is geared to minimize the number of non-zero groups. (Previous works on group-sparse TV use an analysis formulation of group sparsity [22], [28], [31].)

To achieve the goals of this work, we generalize the group-sparse regularizer (16) to include an invertible matrix P,

$$\begin{aligned} R(v; P) = \min_r \; & \sum_n \|PB_n(r)\|_2 \\ \text{such that } \; & v = Sr. \end{aligned} \tag{19}$$

Note that $R(v; aP) = aR(v; P)$ for $a > 0$.

Also, if K is a banded matrix, then SKS^{T} is also banded. This fact contributes to the computational efficiency of the algorithm developed in Sec. V.

IV. Problem Formulation

We formulate the denoising of a discontinuous piecewise linear signal as the strictly convex optimization problem,

$$\hat{x} = \arg\min_x \left\{ \frac{1}{2}\|y - x\|_2^2 + \lambda R(Dx; P) \right\} \tag{20}$$

where R is defined by (19), $\lambda > 0$, and

$$P = \begin{bmatrix} \rho_1 & \\ & \rho_2 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} = \begin{bmatrix} \rho_1 & \rho_1 \\ \rho_2 & -\rho_2 \end{bmatrix} \tag{21}$$

where $\rho_i > 0$. The formulation promotes pairwise group sparsity of the second-order difference signal. The parameters ρ_1 and ρ_2 weight the parameters Δ_{slope} and Δ_{value} respectively; hence, they influence the knot behavior.

Three parameters $(\lambda, \rho_1, \rho_2)$ appear in problem (20), but the problem really has only two independent parameters. Without loss of generality, we may set $\lambda = 1$, because $\lambda R(Dx; P) = R(Dx; \lambda P)$ for $\lambda > 0$.

V. Algorithm

In this section, we derive an algorithm to solve (20). The function (19) can be written as

$$\begin{aligned} R(v; P) = \min_{r,u} \; & \sum_n \|B_n(u)\|_2 \\ \text{such that } \; & \begin{cases} v = Sr \\ B_n(u) = PB_n(r) \text{ for all } n. \end{cases} \end{aligned} \tag{22}$$

The constraint $B_n(u) = PB_n(r)$ for all n can be written as $u = (I \otimes P)r$, where $\otimes$ is the Kronecker product. Hence, $r = (I \otimes P^{-1})u$ and we may write (22) as

$$\begin{aligned} R(v; P) = \min_u \; & \sum_n \|B_n(u)\|_2 \\ \text{such that } \; & v = S(I \otimes P^{-1})u. \end{aligned} \tag{23}$$

Therefore, problem (20) can be expressed as

$$\begin{aligned} \min_{x,u} \; & \left\{ F(x,u) = \frac{1}{2}\|y - x\|_2^2 + \lambda \sum_n \|B_n(u)\|_2 \right\} \\ \text{such that } \; & Dx = S(I \otimes P^{-1})u. \end{aligned} \tag{24}$$

To find the solution of problem (24) (and thereby of (20)), we use the majorization-minimization (MM) approach [11], [20]. This leads to the iteration

$$\begin{aligned} \{x^{(i)}, u^{(i)}\} = \arg\min_{x,u} \; & F^{\mathsf{M}}(x, u; u^{(i-1)}) \\ \text{such that } \; & Dx = S(I \otimes P^{-1})u \end{aligned} \tag{25}$$

where i is the iteration index and F^{M} is a majorizer of F, i.e.,

$$F^{\mathsf{M}}(x, u; z) \geqslant F(x, u), \quad F^{\mathsf{M}}(x, u; u) = F(x, u). \tag{26}$$

To obtain a majorizer of F, we use the inequality

$$\frac{1}{2}\frac{\|u\|_2^2}{\|z\|_2} + \frac{1}{2}\|z\|_2 \geqslant \|u\|_2, \quad z \neq 0. \tag{27}$$

A key point is that the left-hand side is quadratic in u. Using (27), a majorizer F^{M} is given by

$$F^{\mathsf{M}}(x, u; z) = \frac{1}{2}\|y - x\|_2^2 + \frac{\lambda}{2}\sum_n \frac{\|B_n(u)\|_2^2}{\|B_n(z)\|_2} + C \tag{28}$$

where C depends on neither x nor u. Using this majorizer, the iteration (25) is given by

$$\begin{aligned} \{x^{(i)}, u^{(i)}\} = \arg\min_{x,u} \; & \left\{ \|y - x\|_2^2 + \lambda u^{\mathsf{T}}(W^{(i-1)} \otimes I_2)u \right\} \\ \text{such that } \; & Dx = S(I \otimes P^{-1})u \end{aligned}$$

where $W^{(i)}$ is a diagonal matrix with elements

$$W^{(i)}_{n,n} = \frac{1}{\|B_n(u^{(i)})\|_2}. \tag{29}$$

This is a least squares problem and its explicit solution is obtained straightforwardly as

$$\Gamma^{(i-1)} = [W^{(i-1)}]^{-1}, \tag{30}$$

$$\mu^{(i)} = \left(\lambda DD^{\mathsf{T}} + A\left(\Gamma^{(i-1)} \otimes I_2\right)A^{\mathsf{T}}\right)^{-1} Dy, \tag{31}$$

$$x^{(i)} = y - \frac{1}{\lambda}D^{\mathsf{T}}\mu^{(i)}, \tag{32}$$

$$u^{(i)} = \left(\Gamma^{(i-1)} \otimes I_2\right)A^{\mathsf{T}}\mu^{(i)}, \tag{33}$$

TABLE I
ALGORITHM FOR GTV-TK DENOISING (20).

Input: $y \in \mathbb{R}^N,\ \lambda > 0,\ \rho_1 > 0,\ \rho_2 > 0$
Initialization: $u = Dy$ or $u = 1$
$P = \left[\begin{smallmatrix} \rho_1 & \rho_1 \\ \rho_2 & -\rho_2 \end{smallmatrix}\right]$
$A = S\,(I \otimes P^{-1})$
Repeat:

$$\Gamma_{n,n} = (u_{2n}^2 + u_{2n+1}^2)^{1/2}$$

$$\mu = \left(\lambda D D^{\mathsf{T}} + A\,(\Gamma \otimes I_2)\,A^{\mathsf{T}}\right)^{-1} D y$$

$$u = (\Gamma \otimes I_2) A^{\mathsf{T}} \mu,$$

Until convergence
Return: $x = y - (1/\lambda) D^{\mathsf{T}} \mu$

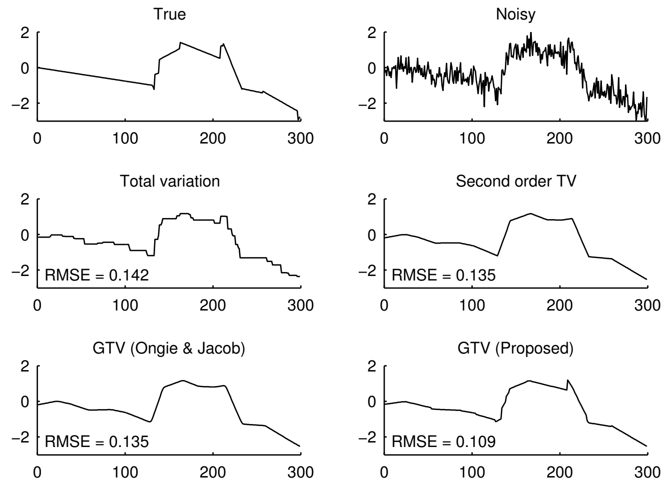

Fig. 3. Example 2. Denoising of a random discrete piecewise linear signal.

where $A = S\,(I \otimes P^{-1})$. Using (29) and (14), $\Gamma^{(i)}$ is a diagonal matrix with elements

$$\Gamma_{n,n}^{(i)} = \|B_n(u^{(i)})\|_2 \tag{34}$$

$$= \big((u_{2n}^{(i)})^2 + (u_{2n+1}^{(i)})^2\big)^{1/2}. \tag{35}$$

The total algorithm to solve (20) is summarized in Table I. Each iteration has cost $O(N)$ as all matrices are banded. In particular, the matrix to be inverted in (31) is banded; hence, μ can be obtained using a fast banded system solver.

By MM principles, the value of the objective function F decreases at each iteration. However, if $B_m(u^{(j)}) = 0$ for some iteration j and index m, then $\Gamma_{m,m}^{(j)} = 0$. Consequently, $B_m(u^{(i)}) = 0$ for all subsequent iterations, $i > j$, by (33). In this case, convergence to the global minima is not assured. In practice, this 'zero-locking' phenomenon is safely avoided by initializing u to non-zero values [11], [27]. An alternate approach to solve (24) is to use proximal methods, e.g., Douglas-Rachford splitting [5].

VI. EXAMPLES

Example 1. The denoised signal obtained by applying the proposed method to the noisy data in Fig. 1(b) is shown in Fig. 1(f). We used the algorithm in Table I to minimize (20) with $\lambda = 1$, $\rho_1 = 10$, and $\rho_2 = 0.5$. As intended, the signal approximates the discontinuity by a single knot. The ratio $\rho_1/\rho_2 = 20$ means that Δ_{slope} is more penalized than Δ_{value}, which promotes discontinuities in the denoised signal.

To illustrate the importance of matrix P in (20), we apply the proposed method with $P = I_2$ (Fig. 1(g)). The second-order difference signal exhibits pairwise group sparsity; however, the signal does not accurately recover the discontinuity. The discontinuity is approximated by two isolated knots, not a single knot. With $P = I_2$, the method fails to tie the knots.

Similarly, we have found that analysis group sparsity with $P \neq I_2$ fails to accurately recover the discontinuity. Hence, we conclude that both synthesis group sparsity and weighting (i.e., $P \neq I_2$) are important in the proposed GTV method.

Example 2. To further evaluate the proposed method, we simulate random discontinuous piecewise linear signals as in [28]. We simulate signals of length 300 each with 10 uniformly distributed discontinuities. The jumps and slopes are distributed as $\mathcal{N}(0,1)$ and $\mathcal{N}(0,0.1)$, respectively. We add white noise distributed as $\mathcal{N}(0,0.4)$. We then denoise each signal using first-order TV, second-order TV, and GTV using the method of [28] and the proposed method. In the proposed method, we set $\rho_1 = 10$ and $\rho_2 = 1$. For each method and each realization, we set λ to minimize the root-mean-square-error (RMSE). Over 50 realizations, the average RMSE is 0.199 for first-order TV, 0.181 for second-order TV, 0.179 for GTV (method of [28]), and 0.151 for GTV (proposed method). An example realization is illustrated in Fig. 3.

Generally, the ratio ρ_1/ρ_2 should depend on the average ratio of the discontinuities in the value and the slope, respectively. If ρ_1/ρ_2 is too small or too large, then the proposed GTV method will be like second-order or first-order TV denoising, respectively. In Example 1, $\rho_1/\rho_2 = 1$ (i.e., $P = I$) is a 'small' value which yields results resembling second-order TV. On the other hand, setting ρ_1/ρ_2 much greater than 20 in Example 1 will yield a result like that of first-order TV.

VII. CONCLUSION

We have proposed a convex generalized total variation (GTV) functional for estimating discontinuous piecewise linear signals from corrupted data. The approach uses the synthesis form of group sparsity to promote pairwise group sparsity of the second derivative. Additionally, a pairwise linear transform is used to decouple the knot parameters.

Extensions to (1) discontinuous higher-order polynomial signals and (2) multidimensional signals will be of interest. It will also be of interest to develop a formulation of convex GTV denoising using a non-convex form of the proposed GTV regularizer, as has been done for standard TV denoising [32].

Finally, we remark that TV regularization is perhaps most effective when used in combination with wavelets or other transforms [2], [8]–[10], [12], [13], [19], [23], [24], [29], [35]. The proposed GTV functional may likewise be most effective when used in conjunction with other regularization methods.

REFERENCES

[1] K. Bredies, K. Kunisch, and T. Pock. Total generalized variation. *SIAM J. Imag. Sci.*, 3(3):492–526, 2010.

[2] E. J. Candés and F. Guo. New multiscale transforms, minimum total variation synthesis: applications to edge-preserving image reconstruction. *Signal Processing*, 82(11):1519–1543, 2002.

[3] A. Chambolle and P.-L. Lions. Image recovery via total variation minimization and related problems. *Numerische Mathematik*, 76:167–188, 1997.

[4] T. Chan, A. Marquina, and P. Mulet. High-order total variation-based image restoration. *SIAM J. Sci. Comput.*, 22(2):503–516, 2000.

[5] P. L. Combettes and J.-C. Pesquet. Proximal splitting methods in signal processing. In H. H. Bauschke et al., editors, *Fixed-Point Algorithms for Inverse Problems in Science and Engineering*, pages 185–212. Springer-Verlag, 2011.

[6] C. Couprie, L. Grady, L. Najman, J.-C. Pesquet, and H. Talbot. Dual constrained TV-based regularization on graphs. *SIAM J. Imag. Sci.*, 6(3):1246–1273, 2013.

[7] V. R. Dantham, S. Holler, V. Kolchenko, Z. Wan, and S. Arnold. Taking whispering gallery-mode single virus detection and sizing to the limit. *Applied Physics Letters*, 101(4):043704, 2012.

[8] Y. Ding and I. W. Selesnick. Artifact-free wavelet denoising: Non-convex sparse regularization, convex optimization. *IEEE Signal Processing Letters*, 22(9):1364–1368, September 2015.

[9] S. Durand and M. Nikolova. Denoising of frame coefficients using ℓ^1 data-fidelity term and edge-preserving regularization. *Multiscale Modeling & Simulation*, 6(2):547–576, 2007.

[10] G. R. Easley, D. Labate, and F. Colonna. Shearlet-based total variation diffusion for denoising. *IEEE Trans. Image Process.*, 18(2):260–268, February 2009.

[11] M. Figueiredo, J. Bioucas-Dias, and R. Nowak. Majorization-minimization algorithms for wavelet-based image restoration. *IEEE Trans. Image Process.*, 16(12):2980–2991, December 2007.

[12] A. Gholami and S. M. Hosseini. A balanced combination of Tikhonov and total variation regularizations for reconstruction of piecewise-smooth signals. *Signal Processing*, 93(7):1945–1960, 2013.

[13] W. Guo, J. Qin, and W. Yin. A new detail-preserving regularization scheme. *SIAM J. Imag. Sci.*, 7(2):1309–1334, 2014.

[14] Y. Hu and M. Jacob. Higher degree total variation (HDTV) regularization for image recovery. *IEEE Trans. Image Process.*, 21(5):2559–2571, May 2012.

[15] Y. Hu, G. Ongie, S. Ramani, and M. Jacob. Generalized higher degree total variation (HDTV) regularization. *IEEE Trans. Image Process.*, 23(6):2423–2435, June 2014.

[16] F. I. Karahanoglu, I. Bayram, and D. Van De Ville. A signal processing approach to generalized 1-d total variation. *IEEE Trans. Signal Process.*, 59(11):5265–5274, November 2011.

[17] S.-J. Kim, K. Koh, S. Boyd, and D. Gorinevsky. l1 trend filtering. *SIAM Review*, 51(2):339–360, 2009.

[18] M. Kowalski. Sparse regression using mixed norms. *J. of Appl. and Comp. Harm. Analysis*, 27(3):303–324, 2009.

[19] J. Krommweh and J. Ma. Tetrolet shrinkage with anisotropic total variation minimization for image approximation. *Signal Processing*, 90(8):2529–2539, 2010.

[20] K. Lange, E. C. Chi, and H. Zhou. A brief survey of modern optimization for statisticians. *Int. Stat. Rev.*, 82(1):46–70, 2014.

[21] S. Lefkimmiatis, A. Bourquard, and M. Unser. Hessian-based norm regularization for image restoration with biomedical applications. *IEEE Trans. Image Process.*, 21(3):983–995, March 2012.

[22] J. Liu, T.-Z. Huang, I. W. Selesnick, X.-G. Lv, and P.-Y. Chen. Image restoration using total variation with overlapping group sparsity. *Information Sciences*, 295:232–246, 2015.

[23] J. Ma. Towards artifact-free characterization of surface topography using complex wavelets and total variation minimization. *Applied Mathematics and Computation*, 170(2):1014–1030, 2005.

[24] F. Malgouyres. Minimizing the total variation under a general convex constraint for image restoration. *IEEE Trans. Image Process.*, 11(12):1450–1456, December 2002.

[25] M. Nikolova. Energy minimization methods. In O. Scherzer, editor, *Handbook of Mathematical Methods in Imaging*, chapter 5, pages 138–186. Springer, 2011.

[26] G. Obozinski, L. Jacob, and J. P. Vert. Group lasso with overlaps: the latent group lasso approach. Technical report, HAL-Inria-00628498, 2011. arXiv:1110.0413.

[27] J. Oliveira, J. Bioucas-Dias, and M. A. T. Figueiredo. Adaptive total variation image deblurring: A majorization-minimization approach. *Signal Processing*, 89(9):1683–1693, September 2009.

[28] G. Ongie and M. Jacob. Recovery of discontinuous signals using group sparse higher degree total variation. *IEEE Signal Processing Letters*, 22(9):1414–1418, September 2015.

[29] N. Pustelnik, C. Chaux, and J. Pesquet. Parallel proximal algorithm for image restoration using hybrid regularization. *IEEE Trans. Image Process.*, 20(9):2450–2462, September 2011.

[30] L. Rudin, S. Osher, and E. Fatemi. Nonlinear total variation based noise removal algorithms. *Physica D*, 60:259–268, 1992.

[31] I. W. Selesnick and P.-Y. Chen. Total variation denoising with overlapping group sparsity. In *Proc. IEEE Int. Conf. Acoust., Speech, Signal Processing (ICASSP)*, pages 5696–5700, May 2013.

[32] I. W. Selesnick, A. Parekh, and I. Bayram. Convex 1-D total variation denoising with non-convex regularization. *IEEE Signal Processing Letters*, 22(2):141–144, February 2015.

[33] S. Setzer, G. Steidl, and T. Teuber. Infimal convolution regularizations with discrete l1-type functionals. *Commun. Math. Sci.*, 9(3):797–827, 2011.

[34] W. Stefan, R. Renaut, and A. Gelb. Improved total variation-type regularization using higher order edge detectors. *SIAM J. Imag. Sci.*, 3(2):232–251, 2010.

[35] L. Wang, L. Xiao, J. Zhang, and Z. Wei. New image restoration method associated with tetrolets shrinkage and weighted anisotropic total variation. *Signal Processing*, 93(4):661–670, 2013.

MATLAB PROGRAM

```
function [x, u, cost] = gtvd(y, rho1, rho2, Nit)
% [x, u, cost] = gtvd(y, rho1, rho2, Nit)
%
% Generalized Total Variation Denoising
%
% The method is based on pairwise group sparse representation of
% second derivative (using synthesis group sparsity).
%
% INPUT
%   y - noisy signal
%   rho1, rho2 -  knot regularization parameters (positive)
%   Nit - number of iterations
%
% OUTPUT
%   x - denoised signal
%   u - sparse representation
%   cost - cost function history

% Ivan Selesnick, selesi@nyu.edu, 2015
% Reference:
% Generalized Total Variation: Tying the Knots
% IEEE Signal Processing Letters

y = y(:);                                  % Convert to column vector
cost = zeros(1, Nit);                      % Cost function history
N = length(y);

e = ones(N, 1);
D = spdiags([e -2*e e], [0 1 2], N-2, N);  % second order difference
DDT = D * D';

S = kron(speye(N-2), [1 1]);
S = S(:, 2:end-1);

P = [rho1 rho1; rho2 -rho2];
L = N - 3;        % Number of groups = L = N-3

A = S * kron(speye(L), inv(P) );
Dy = D*y;
w = ones(L,1);      % Initialization

for k = 1:Nit
    G = spdiags(kron(w, [1; 1]), 0, 2*L, 2*L);
    mu = (DDT + A * G * A') \ Dy;

    x = y - D' * mu;
    u = G *  A' * mu;
    w = sqrt(u(1:2:end).^2 + u(2:2:end).^2);

    %   max(abs(A*u - D*x)) % verify (should be zero)
    cost(k) = 0.5 * sum(abs( x - y ).^2) + sum( w );

end
```