TOTAL GENERALIZED VARIATION FOR GRAPH SIGNALS

Shunsuke Ono[†], *Isao Yamada*[††], *and Itsuo Kumazawa*[†]

[†]Tokyo Institute of Technology, Imaging Science & Engineering Laboratory, Kanagawa, Japan
[††]Tokyo Institute of Technology, Dept. Communications & Computer Engineering, Tokyo, Japan

ABSTRACT

This paper proposes a second-order discrete total generalized variation (TGV) for arbitrary graph signals, which we call the graph TGV (G-TGV). The original TGV was introduced as a natural higher-order extension of the well-known total variation (TV) and is an effective prior for piecewise smooth signals. Similarly, the proposed G-TGV is an extension of the TV for graph signals (G-TV) and inherits the capability of the TGV, such as avoiding staircasing effect. Thus the G-TGV is expected to be a fundamental building block for graph signal processing. We provide its applications to piecewise-smooth graph signal inpainting and 3D mesh smoothing with illustrative experimental results.

***Index Terms*—** Graph signal processing, total generalized variation (TGV), proximal splitting.

1. INTRODUCTION

The concept of *graph signal* explicitly models the connection or relation among signal samples by assigning each of them to each vertex of a graph. Since various types of data, such as images and videos, traffic and sensor network data, mesh data, and biomedical data, can be represented as graph signals, the framework of graph signal processing can provide fundamental tools to analyze, compress, and process a broad class of data in a unified way, see, e.g., [1, 2, 3, 4] and references therein.

In many applications, one can only access noisy and/or incomplete observations of a true graph signal, e.g., cortical activation [5] and mesh surfaces obtained through a scanning process [6], so that one needs to leverage some a priori knowledge to estimate the true graph signal. This approach is usually realized through optimization involving *priors*, functions reflecting the knowledge.

The *total variation* for graph signals (we call it the graph TV (G-TV)) is defined as an absolute sum of the discrete difference of a graph signal and is a prior suitable for *piecewise-smooth* graph signals. It was first introduced in [7] as an extension of the original total variation (TV) [8], and its weighted generalizations were successively proposed in [9, 10]. In [11], a new G-TV was developed for directed graphs. Recently, a very general version of the G-TV was introduced in [6], where the authors define their G-TV in the dual domain and establish an efficient optimization framework for it.

This work is supported in part by JSPS Grants-in-Aid for JSPS fellows (24·2522) and (B-25280073).

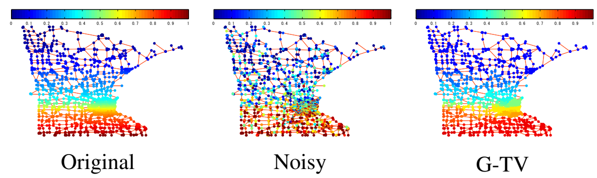


Fig. 1. Example of the staircasing effect in a graph signal: The denoised graph signal by the G-TV (right) is too flattened and has artifact boundaries compared with the original one.

When we use the TV, the so-called *staircasing effect*, the undesirable appearance of edges, often occurs. This is also the case with the G-TV (see Fig. 1). To overcome this limitation, the *total generalized variation* (TGV) was proposed in [12] as a reasonable higher-order generalization of the TV, which has been successfully applied to various image processing applications [13, 14, 15, 16, 17, 18].[1]

The main purpose of the paper is to develop a second-order TGV on graphs (we call it the graph TGV (G-TGV)) and apply it to graph signal restoration and smoothing. The G-TGV is defined by incorporating several mathematical elements from graph theory (see, e.g., [19, 20]) into the TGV. We also provide an efficient algorithm based on a primal-dual splitting method [21, 22] for solving optimization problems involving the G-TGV. Since the G-TGV can be used for restoration and smoothing of any undirected graph signals, there are many potential applications. Specifically, we propose two applications: piecewise-smooth graph signal inpainting and 3D mesh smoothing. Experimental results show its superiority over the G-TV in the applications.

The rest of the paper is organized as follows. Sec. 2 is devoted to establish the G-TGV and its utilization in inverse problems of graph signals via optimization. We present applications of the G-TGV in Sec. 3 with illustrative experimental results, and conclude the paper in Sec. 4.

[1]The TGV of second-order is mainly used.

2. GRAPH TOTAL GENERALIZED VARIATION

2.1. Definition

We consider a weighted graph $G = (V, E, \mathbf{W})$ with vertices $v_i \in V$ and edges $e_{i,j} \in E \subset V \times V$, where the numbers of vertices and edges are denoted by $|V|$ and $|E|$, respectively, and $1 \leq i, j \leq |V|$ are the indices of vertices such that $i < j$. Each weight assigned to each edge $e_{i,j}$ is denoted by $w_{e_{i,j}} > 0$ (the number of $w_{e_{i,j}}$ equals to $|E|$), and $\mathbf{W} \in \mathbb{R}^{|E| \times |E|}$ is the diagonal matrix containing all $w_{e_{i,j}}$.

We also introduce the *incidence matrix* $\mathbf{D} \in \mathbb{R}^{|E| \times |V|}$ of a graph G, of which the entries are defined by

$$D_{e_{i,j}, v_k} := \begin{cases} -1, & \text{if } i = k, \\ 1, & \text{if } j = k, \\ 0, & \text{otherwise.} \end{cases}$$

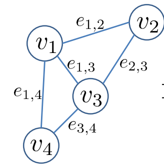
For readers' convenience, we give a simple graph and its incidence matrix in Fig. 2. It is known that $\mathbf{D}$ can be seen as the discrete gradient operator for graph signals on G and satisfies $\mathbf{D}^\top \mathbf{W}^2 \mathbf{D} = \mathbf{L}$, where $\mathbf{L} \in \mathbb{R}^{|V| \times |V|}$ is the so-called *graph Laplacian matrix* of a weighted graph G (see, e.g., [23]).

Let $\mathbf{u} = (u_1, \ldots, u_{|V|})^\top \in \mathbb{R}^{|V|}$ be a graph signal vector where u_i resides on v_i. Based on the original TGV [12], we newly formulate a second-order graph total generalized variation (G-TGV) for an arbitrary weighted graph G as follows:

$$\begin{aligned} \mathrm{TGV}_G^\alpha(\mathbf{u}) &:= \min_{\mathbf{p}, \mathbf{q} \in \mathbb{R}^{|E|}} \alpha \|\mathbf{p}\|_1 + (1 - \alpha) \|\mathbf{D}^\top \mathbf{W} \mathbf{q}\|_1 \\ &\qquad\qquad \text{s.t. } \mathbf{W}\mathbf{D}\mathbf{u} = \mathbf{p} + \mathbf{q}, \\ &= \min_{\mathbf{q} \in \mathbb{R}^{|E|}} \alpha \|\mathbf{W}\mathbf{D}\mathbf{u} - \mathbf{q}\|_1 + (1 - \alpha) \|\mathbf{D}^\top \mathbf{W} \mathbf{q}\|_1, \end{aligned}$$

where $0 < \alpha < 1$, and $\| \cdot \|_1$ stands for the ℓ_1-norm, i.e., the sum of absolute values of all the entries of $(\cdot)$. In this definition, $\mathbf{W}\mathbf{D}\mathbf{u}$ corresponds to the (weighted) first-order difference vector of the graph signal $\mathbf{u}$, and $\mathbf{p}$ and $\mathbf{q}$ are its portions such that $\mathbf{W}\mathbf{D}\mathbf{u} = \mathbf{p} + \mathbf{q}$. This implies that $\mathbf{D}^\top \mathbf{W} \mathbf{q}$ in the right term is the (weighted) second-order difference vector w.r.t. the portion $\mathbf{q}$. As a result, the left term measures the total magnitude of a part of the first-order difference of $\mathbf{u}$, and the right term evaluates the second-order difference of $\mathbf{u}$ w.r.t. the residual part of the first-order difference of $\mathbf{u}$. Since $\mathbf{D}^\top \mathbf{W} \mathbf{q}$ is a part of the Laplacian of $\mathbf{u}$ due to $\mathbf{L}\mathbf{u} = \mathbf{D}^\top \mathbf{W}^2 \mathbf{D} \mathbf{u} = \mathbf{D}^\top \mathbf{W}(\mathbf{p} + \mathbf{q})$, we believe that the above definition is a natural incorporation of elements of graph theory into the TGV.

As the original TGV, the G-TGV is obviously a proper lower semicontinuous convex function[2] over $\mathbb{R}^{|V|}$, and thus we will benefit from convex optimization techniques for solving problems involving the G-TGV in the next section.



$$\mathbf{D} = \begin{array}{c} \begin{array}{cccc} v_1 & v_2 & v_3 & v_4 \end{array} \\ \begin{pmatrix} -1 & 1 & 0 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & 0 & -1 & 1 \\ -1 & 0 & 1 & 0 \\ -1 & 0 & 0 & 1 \end{pmatrix} \end{array} \begin{array}{c} \\ e_{1,2} \\ e_{2,3} \\ e_{3,4} \\ e_{1,3} \\ e_{1,4} \end{array}$$

Fig. 2. A graph and its incidence matrix.

2.2. Problem Statement

Consider the following graph signal observation model:

$$\mathbf{b} = \mathbf{\Phi} \bar{\mathbf{u}} + \mathbf{n},$$

where $\mathbf{\Phi} \in \mathbb{R}^{K \times |V|}$ is a linear degradation operator (e.g., decimation), $\bar{\mathbf{u}} \in \mathbb{R}^{|V|}$ is an unknown original graph signal on a given graph G, and $\mathbf{n} \in \mathbb{R}^K$ is an additive noise. Here we assume that $\bar{\mathbf{u}}$ is piecewise-smooth.

To estimate $\bar{\mathbf{u}}$, we propose to solve the following convex optimization problem:

$$\min_{\mathbf{u} \in \mathbb{R}^{|V|}} \mathrm{TGV}_G^\alpha(\mathbf{u}) \text{ s.t. } \mathbf{u} \in R_{\underline{\mu}, \overline{\mu}} \text{ and } \mathbf{\Phi}\mathbf{u} \in B_{\mathbf{b}, \varepsilon}, \tag{1}$$

where

$$\begin{aligned} R_{\underline{\mu}, \overline{\mu}} &:= \{\mathbf{x} \in \mathbb{R}^{|V|} |\ \underline{\mu} \leq x_i \leq \overline{\mu} \ \ (i = 1, \ldots, |V|)\}, \\ B_{\mathbf{b}, \varepsilon} &:= \{\mathbf{x} \in \mathbb{R}^K |\ \|\mathbf{x} - \mathbf{b}\|_2 \leq \varepsilon\}. \end{aligned}$$

The set $R_{\underline{\mu}, \overline{\mu}}$ is a box constraint with $\underline{\mu} < \overline{\mu}$ which represents some known numerical range of $\bar{\mathbf{u}}$ (if such information is unavailable, set $\underline{\mu}$ ($\overline{\mu}$) to a sufficiently small (large) value). Such a box constraint also guarantees the existence of a minimizer of (1). The set $B_{\mathbf{b}, \varepsilon}$ is a $\mathbf{b}$-centered ℓ_2-norm ball with the radius $\varepsilon > 0$, which serves as a fidelity constraint w.r.t. the observation $\mathbf{b}$. Using a fidelity constraint instead of an additive fidelity term facilitates the parameter setting since ε is directly related to a statistical parameter of noise (e.g., variance) and can be determined independent of what a prior function is employed. Hence, solving the problem would bring a good estimate of the original piecewise-smooth graph signal.

2.3. Optimization

We use a primal-dual splitting method [21, 22] to solve (1). The method can solve optimization problems of the form:

$$\min_{\mathbf{x} \in \mathbb{R}^N} f_1(\mathbf{x}) + f_2(\mathbf{x}) + f_3(\mathbf{A}\mathbf{x}), \tag{2}$$

where f_1 is a differentiable convex function with the β-Lipschitzian gradient ∇f_1 for some $\beta > 0$, $f_2 \in \Gamma_0(\mathbb{R}^N)$ and $f_3 \in \Gamma_0(\mathbb{R}^M)$ are *proximable*[3], and $\mathbf{A} : \mathbb{R}^{M \times N}$ is a matrix.

[2] A function $f : \mathbb{R}^N \to \mathbb{R} \cup \{\infty\}$ is called *proper lower semicontinuous convex* if $\mathrm{dom} f := \{\mathbf{x} \in \mathbb{R}^N |\ f(\mathbf{x}) < \infty\} \neq \emptyset$, $\mathrm{lev}_{\leq \alpha}(f) := \{\mathbf{x} \in \mathbb{R}^N |\ f(\mathbf{x}) \leq \alpha\}$ is closed for every $\alpha \in \mathbb{R}$, and $f(\lambda \mathbf{x} + (1 - \lambda)\mathbf{y}) \leq \lambda f(\mathbf{x}) + (1 - \lambda) f(\mathbf{y})$ for every $\mathbf{x}, \mathbf{y} \in \mathbb{R}^N$ and $\lambda \in (0, 1)$, respectively. The set of all proper lower semicontinuous convex functions over $\mathbb{R}^N$ is denoted by $\Gamma_0(\mathbb{R}^N)$.

[3] The *proximity operator* [24] of a function $f \in \Gamma_0(\mathbb{R}^N)$ of an index $\gamma > 0$ is defined by $\mathrm{prox}_{\gamma f}(\mathbf{x}) := \arg\min_{\mathbf{y} \in \mathcal{X}} f(\mathbf{y}) + \frac{1}{2\gamma} \|\mathbf{x} - \mathbf{y}\|_2^2$. If an efficient computation of $\mathrm{prox}_{\gamma f}$ is available, we call f proximable.

The algorithm is given by

$$\left\lfloor\begin{array}{l}\mathbf{x}^{(n+1)} = \mathrm{prox}_{\gamma_1 f_2}[\mathbf{x}^{(n)} - \gamma_1(\nabla f_1(\mathbf{x}^{(n)}) + \mathbf{A}^\top \mathbf{y}^{(n)})], \\ \mathbf{y}^{(n+1)} = \mathrm{prox}_{\gamma_2 f_3^*}[\mathbf{y}^{(n)} + \gamma_2 \mathbf{A}(2\mathbf{x}^{(n+1)} - \mathbf{x}^{(n)})],\end{array}\right.$$

where f_3^* the *Fenchel-Rockafellar conjugate function*[4] of f_3, and $\gamma_1, \gamma_2 > 0$ satisfy $\frac{1}{\gamma_1} - \gamma_2 \lambda_1(\mathbf{A}^\top \mathbf{A}) \geq \frac{\beta}{2}$ ($\lambda_1(\cdot)$ stands for the maximum eigen value of $\cdot$). Under some mild conditions on f_2, f_3, and $\mathbf{A}$, the sequence $(\mathbf{x}^{(n)})_{n\in\mathbb{N}}$ converges to a solution to (2).

Using the indicator functions[5] of $R_{\underline{\mu},\overline{\mu}}$ and $B_{\mathbf{b},\varepsilon}$, we can rewrite (1) as

$$\min_{\mathbf{u}\in\mathbb{R}^{|V|},\mathbf{q}\in\mathbb{R}^{|E|}} \alpha\|\mathbf{WDu} - \mathbf{q}\|_1 + (1-\alpha)\|\mathbf{D}^\top \mathbf{Wq}\|_1 + \iota_{R_{\underline{\mu},\overline{\mu}}}(\mathbf{u}) + \iota_{B_{\mathbf{b},\varepsilon}}(\mathbf{\Phi u}), \tag{3}$$

Now, let $\mathbf{x} := (\mathbf{u}^\top \ \mathbf{q}^\top)^\top$, and $\mathbf{y} := (\mathbf{y}_1^\top \ \mathbf{y}_2^\top \ \mathbf{y}_3^\top)$ with $\mathbf{y}_1 \in \mathbb{R}^{|E|}$, $\mathbf{y}_2 \in \mathbb{R}^{|V|}$, and $\mathbf{y}_3 \in \mathbb{R}^K$. Then, by defining $f_1(\mathbf{x}) := 0$, $f_2(\mathbf{x}) := \iota_{R_{\underline{\mu},\overline{\mu}}}(\mathbf{u})$,

$$f_3(\mathbf{y}) := \alpha\|\mathbf{y}_1\|_1 + (1-\alpha)\|\mathbf{y}_2\|_1 + \iota_{B_{\mathbf{b},\varepsilon}}(\mathbf{y}_3),$$
$$\mathbf{A} := \begin{pmatrix} \mathbf{WD} & -\mathbf{I} \\ \mathbf{O} & \mathbf{D}^\top \mathbf{W} \\ \mathbf{\Phi} & \mathbf{O} \end{pmatrix},$$

problem (2) is reduced to (3), i.e., (1).

The proximity operator of f_2 is the metric projection onto $R_{\underline{\mu},\overline{\mu}}$ (see footnote 5), i.e.,

$$[\mathrm{prox}_{\gamma\iota_{R_{\underline{\mu},\overline{\mu}}}}(\mathbf{z})]_i = [P_{R_{\underline{\mu},\overline{\mu}}}(\mathbf{z})]_i = \begin{cases} \underline{\mu}, & \text{if } z_i < \underline{\mu} \\ \overline{\mu} & \text{if } z_i > \overline{\mu} \\ z_i & \text{otherwise.} \end{cases}$$

Meanwhile, the proximity operator of f_3 can be decomposed into that of each term. The proximity operator of $\|\cdot\|_1$ is reduced to the so-called soft-thresholding operation:

$$[\mathrm{prox}_{\gamma\|\cdot\|_1}(\mathbf{z})]_i = [\mathrm{ST}(\mathbf{z},\gamma)]_i = \mathrm{sgn}(z_i)\max\{0, |z_i| - \gamma\},$$

where sgn denotes the signum function. As in the case of $\iota_{R_{\underline{\mu},\overline{\mu}}}$, the proximity operator of $\iota_{B_{\mathbf{b},\varepsilon}}$ is given by

$$\mathrm{prox}_{\gamma\iota_{B_{\mathbf{b},\varepsilon}}}(\mathbf{z}) = P_{B_{\mathbf{b},\varepsilon}}(\mathbf{z}) = \begin{cases} \mathbf{z}, & \text{if } \mathbf{z} \in B_{\mathbf{b},\varepsilon}, \\ \mathbf{b} + \frac{\varepsilon(\mathbf{z}-\mathbf{b})}{\|\mathbf{z}-\mathbf{b}\|_2}, & \text{otherwise.} \end{cases}$$

Hence, the proximity operators of f_2 and f_3 are computable with $\mathcal{O}(|V|)$ and $\mathcal{O}(|E|+|V|+K)$, respectively, implying we can efficiently solve (1) by the primal-dual splitting method. Finally, we show the detailed algorithm in Algorithm 1.

[4]The *Fenchel-Rockafellar conjugate function* of $f \in \Gamma_0(\mathbb{R}^N)$ is defined by $f^*(\boldsymbol{\xi}) := \sup_{\mathbf{x}\in\mathbb{R}^N}\{\langle\mathbf{x},\boldsymbol{\xi}\rangle - f(\mathbf{x})\}$. The proximity operator of f^* can be expressed as $\mathrm{prox}_{\gamma f^*}(\mathbf{x}) = \mathbf{x} - \gamma\mathrm{prox}_{\gamma^{-1}f}(\gamma^{-1}\mathbf{x})$.

[5] For a given nonempty closed convex set $C \in \mathbb{R}^N$, 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., $\mathrm{prox}_{\gamma\iota_C}(\mathbf{x}) = \arg\min_{\mathbf{y}\in C}\|\mathbf{x}-\mathbf{y}\|_2 =: P_C(\mathbf{x})$ ($\forall\gamma > 0$).

Algorithm 1: Primal-dual splitting method for (1)

input : $\mathbf{u}^{(0)}, \mathbf{q}^{(0)}, \mathbf{y}_i^{(0)} (i = 1,2,3)$
output : $\mathbf{u}^{(n)}$
while *A stopping criterion is not satisfied* **do**
- $\mathbf{u}^{(n+1)} = P_{R_{\underline{\mu},\overline{\mu}}}(\mathbf{u}^{(n)} - \gamma_1(\mathbf{D}^\top \mathbf{W}\mathbf{y}_1^{(n)} + \mathbf{\Phi}^\top \mathbf{y}_3^{(n)}))$;
- $\mathbf{q}^{(n+1)} = \mathbf{q}^{(n)} - \gamma_1(-\mathbf{y}_1^{(n)} + \mathbf{WD}\mathbf{y}_2^{(n)})$;
- $\mathbf{y}_1^{(n)} \leftarrow \mathbf{y}_1^{(n)} + \gamma_2(\mathbf{WD}(2\mathbf{u}^{(n+1)} - \mathbf{u}^{(n)}) - (2\mathbf{q}^{(n+1)} - \mathbf{q}^{(n)}))$;
- $\mathbf{y}_2^{(n)} \leftarrow \mathbf{y}_2^{(n)} + \gamma_2\mathbf{D}^\top \mathbf{W}(2\mathbf{q}^{(n+1)} - \mathbf{q}^{(n)})$;
- $\mathbf{y}_3^{(n)} \leftarrow \mathbf{y}_3^{(n)} + \gamma_2\mathbf{\Phi}(2\mathbf{u}^{(n+1)} - \mathbf{u}^{(n)})$;
- $\mathbf{y}_1^{(n+1)} = \mathbf{y}_1^{(n)} - \gamma_2\mathrm{ST}(\frac{1}{\gamma_2}\mathbf{y}_1^{(n)}, \frac{\alpha}{\gamma_2})$;
- $\mathbf{y}_2^{(n+1)} = \mathbf{y}_2^{(n)} - \gamma_2\mathrm{ST}(\frac{1}{\gamma_2}\mathbf{y}_2^{(n)}, \frac{1-\alpha}{\gamma_2})$;
- $\mathbf{y}_3^{(n+1)} = \mathbf{y}_3^{(n)} - \gamma_2 P_{B_{\mathbf{b},\varepsilon}}(\frac{1}{\gamma_2}\mathbf{y}_3^{(n)})$;
- $n \leftarrow n + 1$;

3. APPLICATIONS

We present applications of the G-TGV to graph signal inpainting and 3D mesh smoothing with illustrative experimental results. We fixed γ_1 and γ_2 as 0.1 and $\frac{1}{30\gamma_1}$ in Algorithm 1, which satisfies $\frac{1}{\gamma_1} - \gamma_2\lambda_1(\mathbf{A}^\top \mathbf{A}) \geq \frac{\beta}{2}$ in the following applications. The stopping criteria of Algorithm 1 is set to $\|\mathbf{u}^{(n+1)} - \mathbf{u}^{(n)}\|_2 < 1.0 \times 10^{-5}$. In all the experiments, the parameter α of the G-TGV was simply chosen as 0.5, and the radius of $B_{\mathbf{b},\varepsilon}$ was fixed to the oracle value, i.e., $\varepsilon = \|\mathbf{\Phi}\bar{\mathbf{u}} - \mathbf{b}\|_2$. We compared the G-TGV with the G-TV, where optimization problems involving the G-TV were also solved by the primal-dual splitting method. For the objective evaluation, we use the (normalized) root mean square error RMSE $= \frac{\|\mathbf{u}^{(n)} - \bar{\mathbf{u}}\|_2}{\|\bar{\mathbf{u}}\|_2}$ ($\mathbf{u}^{(n)}$ and $\bar{\mathbf{u}}$ stand for an estimated and the original graph signals, respectively).

3.1. Graph Signal Inpainting

For the first application, we present graph singal inpainting with denoising using the G-TGV. We corrupted a piece-wise smooth signal on the Minnesota road network graph [25] with an additive white Gaussian noise (standard deviation $\sigma = 0.25$), and then randomly masked it (the numerical range of the original graph signal is from $\underline{\mu} = 0$ to $\overline{\mu} = 1$). Specifically, we consider the two cases: (i) 0% missing (only noise corruption) and (ii) 50% missing. All the weights of the graph are set to 1, i.e., $\mathbf{W} = \mathbf{I}$.

The results are shown in Fig. 3. The graph signals restored by the G-TV exhibit the staircasing effect (see the third column from left). In contrast, minimizing the G-TGV results in a good smoothing and their RMSEs are better than those of the G-TV's results (see the right end column), which demonstrates the effectiveness of the G-TGV.

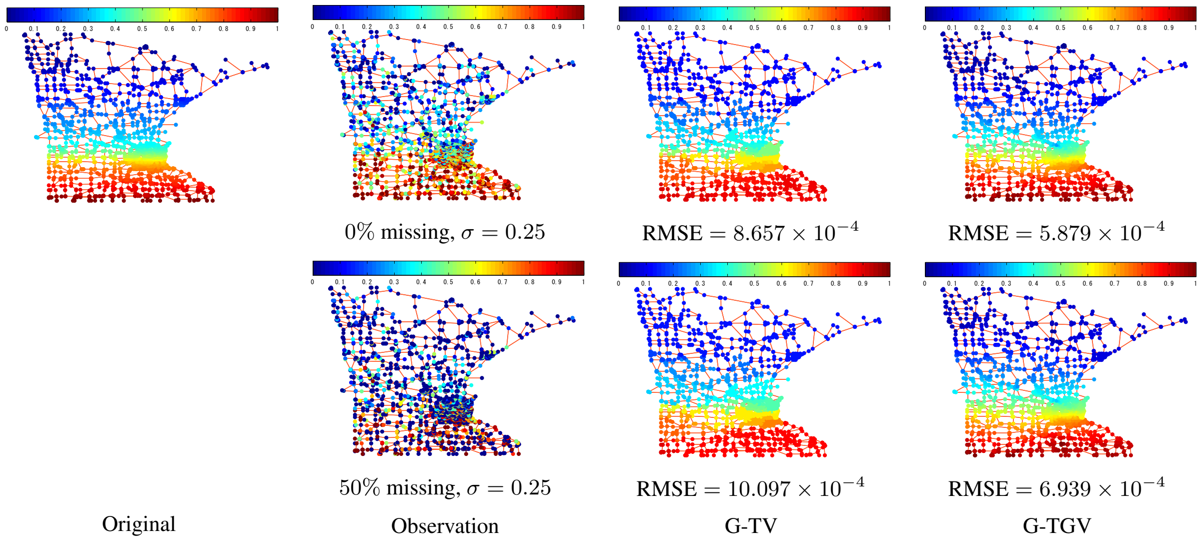


Fig. 3. Graph signal inpainting results: The G-TGV well reconstructs the original graph signal without the staircasing effect.

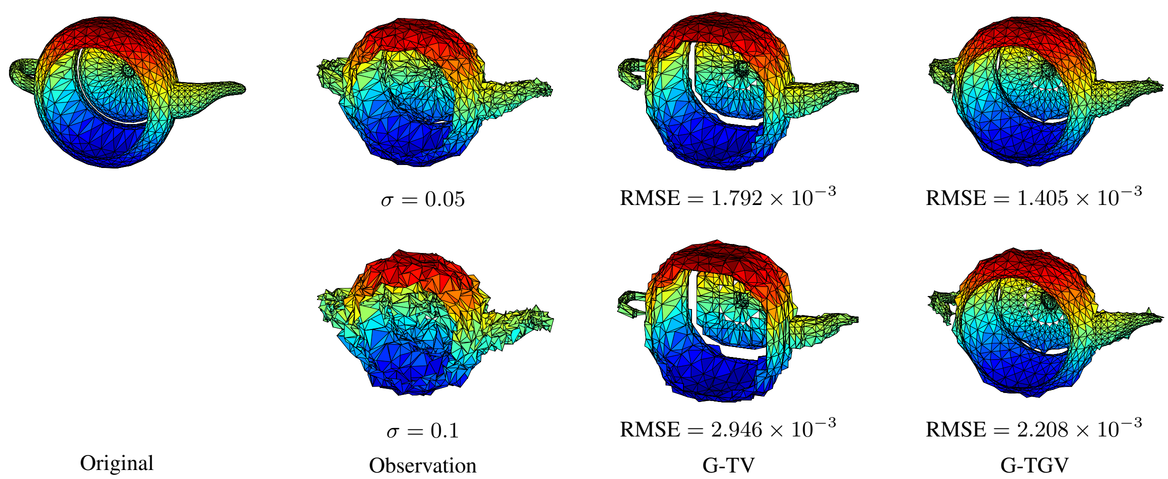


Fig. 4. 3D mesh smoothing results: The G-TGV produces smooth surfaces which are visually more pleasing than the G-TV's results.

3.2. 3D Mesh Smoothing

We also apply the G-TGV to 3D mesh smoothing. In this case, there are totally three graph signals $(\bar{\mathbf{u}}_x, \bar{\mathbf{u}}_y, \bar{\mathbf{u}}_z)$ that represent the spatial coordinates of a triangle mesh node. We added a randomly oriented white Gaussian noise ($\sigma = 0.05$ or 0.1) to the original coordinates of the teapot triangle mesh [6], i.e., the observations are given by $(\mathbf{b}_x, \mathbf{b}_y, \mathbf{b}_z) = (\bar{\mathbf{u}}_x + \mathbf{n}_x, \bar{\mathbf{u}}_y + \mathbf{n}_y, \bar{\mathbf{u}}_z + \mathbf{n}_z)$, where $\mathbf{n}_x$,$\mathbf{n}_y$, and $\mathbf{n}_z$ are uncorrelated vectors of additive noises. The original coordinates are in the range from $\underline{\mu} = -3$ to $\overline{\mu} = 3$. As the case of graph signa inpainting in Sec. 3.1, all the weights of the graph are set to 1.

We show the results in Fig. 4. The resulting meshes by minimizing the G-TV are too square-shaped compared with the original one (see the third column from left). This is because the G-TV only measures the first-order difference of the coordinates. On the other hand, the G-TGV well models higher-order smoothness underlying mesh surfaces, leading to much better smoothing results with lower RMSEs than the G-TV's results (see the right end column).

4. CONCLUDING REMARKS

We have proposed an extension of the TGV for graph signals (G-TGV). The G-TGV well evaluates the piecewise-smoothness of signals on arbitrary undirected graphs, and its design is suitable for inverse problems of graph signals and enables efficient optimization. We have illustrated the G-TGV over several applications, where it is superior to the G-TV. The G-TGV can also be used as a constraint via the technique proposed in [26], which would enhance the potential utility of the G-TGV.

5. REFERENCES

[1] D. K. Hammond, P. Vandergheynst, and R. Gribonval, "Wavelets on graphs via spectral graph theory," *Appl. Comput. Harmon. Anal.*, vol. 30, no. 2, pp. 129150, 2011.

[2] D. I. Shuman, S. K. Narang, P. Frossard, A. Ortega, and P. Vandergheynst, "The emerging field of signal processing on graphs: Extending high-dimensional data analysis to networks and other irregular domains," *IEEE Signal Process. Mag.*., vol. 30, no. 3, pp. 83–98, 2013.

[3] A. Sandryhaila and J. M. F. Moura, "Discrete signal processing on graphs," *IEEE Trans. Signal Process.*, vol. 61, no. 7, pp. 1644–1656, 2013.

[4] S. K. Narang and A Ortega, "Compact support biorthogonal wavelet filterbanks for arbitrary undirected graphs," *IEEE Trans. Signal Process.*, vol. 61, no. 19, pp. 4673–4685, 2013.

[5] D. K. Hammond, B. Scherrer, and S. K. Warfield, "Cortical graph smoothing: A novel method for exploiting DWI-derived anatomical brain connectivity to improve EEG source estimation," *IEEE Trans. Med. Imag.*, vol. 32, no. 10, pp. 1952–1963, 2013.

[6] C. Couprie, L. Grady, L. Najman, J.-C. Pesquet, and H. Talbot, "Dual constrained tv-based regularization on graphs," *SIAM J. Imag. Sci.*, vol. 6, no. 3, pp. 1246–1273, 2013.

[7] T.F. Chan, S. Osher, and J. Shen, "The digital TV filter and nonlinear denoising," *IEEE Trans. Image Process.*, vol. 10, no. 2, pp. 231–241, 2001.

[8] L. I. Rudin, S. Osher, and E. Fatemi, "Nonlinear total variation based noise removal algorithms," *Phys. D*, vol. 60, no. 1-4, pp. 259–268, 1992.

[9] G. Gilboa and S. Osher, "Nonlocal linear image regularization and supervised segmentation," *Multi. Model. Simul.*, vol. 6, no. 2, pp. 595–630, 2007.

[10] A. Elmoataz, O. Lezoray, and S. Bougleux, "Nonlocal discrete regularization on weighted graphs: A framework for image and manifold processing," *IEEE Trans. Image Process.*, vol. 17, no. 7, pp. 1047–1060, 2008.

[11] S. Chen., A. Sandryhaila, G. Lederman, Z. Wang., J. M. F. Moura, P. Rizzo, J. Bielak, J. H. Garrett, and J. Kovacević, "Signal inpainting on graphs via total variation minimization," in *Proc. IEEE ICASSP*, 2014.

[12] K. Bredies, K. Kunisch, and T. Pock, "Total generalized variation," *SIAM J. Imag. Sci.*, vol. 3, no. 3, pp. 92–526, 2010.

[13] K. Bredies, "Recovering piecewise smooth multichannel images by minimization of convex functionals with total generalized variation penalty," in *Efficient Algorithms for Global Optimization Methods in Computer Vision*, Lecture Notes in Computer Science, pp. 44–77. Springer Berlin Heidelberg, 2014.

[14] T. Valkonen, K. Bredies, and F. Knoll, "Total generalized variation in diffusion tensor imaging," *SIAM J. Imag. Sci.*, vol. 6, no. 1, pp. 487–525, 2013.

[15] K. Bredies and M. Holler, "A TGV regularized wavelet based zooming model," in *Scale Space and Variational Methods in Computer Vision*, Lecture Notes in Computer Science, pp. 149–160. Springer Berlin Heidelberg, 2013.

[16] T. Miyata, "L infinity total generalized variation for color image recovery," in *Proc. IEEE Int. Conf. Image Process. (ICIP)*, 2013, pp. 449–453.

[17] D. Ferstl, C. Reinbacher, R. Ranftl, Matthias Rüther, and H. Bischof, "Image guided depth upsampling using anisotropic total generalized variation," in *Proc. IEEE ICCV*, 2013.

[18] S. Ono and I. Yamada, "Decorrelated vectorial total variation," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, 2014.

[19] C. Godsil and G. F. Royle, *Algebraic graph theory*, Springer New York, 2001.

[20] F. K. Chung, *Spectral graph theory*, American Mathematical Soc., 1997.

[21] L. Condat, "A primal-dual splitting method for convex optimization involving Lipschitzian, proximable and linear composite terms," *J. Optim. Theory Appl.*, vol. 158, no. 2, pp. 460–479, 2013.

[22] B. C. Vu, "A splitting algorithm for dual monotone inclusions involving cocoercive operators," *Adv. Comput. Math.*, vol. 38, no. 3, pp. 667–681, 2013.

[23] L. J. Grady and J. R. Polimeni, *Discrete calculus: Applied analysis on graphs for computational science*, Springer, 2010.

[24] J. J. Moreau, "Fonctions convexes duales et points proximaux dans un espace hilbertien," *C. R. Acad. Sci. Paris Ser. A Math.*, vol. 255, pp. 2897–2899, 1962.

[25] N. Perraudin, J. Paratte, D. Shuman, V. Kalofolias, P. Vandergheynst, and D. K. Hammond, "GSPBOX: A toolbox for signal processing on graphs," *ArXiv e-prints*, 2014.

[26] S. Ono and I. Yamada, "Second-order total generalized variation constraint," in *Proc. IEEE ICASSP*, 2014.