Total Variation Meets Topological Persistence: A First Encounter

Ulrich Bauer, Carola-Bibiane Schönlieb[1] and Max Wardetzky

Institut für Numerische und Angewandte Mathematik, Lotzestraße 16–18, D-37083 Göttingen, Germany

Abstract. We present first insights into the relation between two popular yet apparently dissimilar approaches to denoising of one dimensional signals, based on (i) total variation (TV) minimization and (ii) ideas from topological persistence. While a close relation between (i) and (ii) might phenomenologically not be unexpected, our work appears to be the first to make this connection precise for one dimensional signals. We provide a link between (i) and (ii) that builds on the equivalence between TV-L^2 regularization and taut strings and leads to a novel and efficient denoising algorithm that is contrast preserving and operates in $O(n\log n)$ time, where n is the size of the input.

Keywords: Total variation, topological persistence, denoising
PACS: 02.30.Jr, 02.30.Xx, 02.40.Pc, 02.60.Pn.

INTRODUCTION

Our motivation for writing this note is to provide insights into the relationship between two popular yet apparently dissimilar approaches for *classifying* and *removing* noise from measured data, based on (i) minimization of total variation regularizing functionals and (ii) ideas from topological persistence. Whereas (i) relies on variational principles and the various analytical properties of PDEs, (ii) is grounded on purely topological and combinatorial concepts.

Variational methods, such as (i), commonly classify noise by working with energy functionals that depend on the input function and its (weak) derivatives; they decrease noise by either letting the input evolve for some time along the flow of the energy (sub)gradient or by directly solving for the minimizer under appropriate boundary conditions. The variational approach to denoising typically changes the input in a *global manner*, i.e., it generally alters the input even in regions that would not a priori be classified as noise. In contrast, methods grounded in topological persistence alter the input only *locally* in precisely those regions that are classified as noise.

The notions of persistent homology and persistence pairs were introduced in [1] in order to investigate the change of the homology groups in a filtration of a topological space (a nested sequence of subspaces). Persistence can be thought of as a weight that quantifies the topological significance of pairs of critical points. Denoising is performed *locally* by canceling certain pairs of critical points with low significance while leaving the rest of the function unaltered. Albeit its preferable locality, persistence simplification suffers from its lack of robustness to noise with unbounded support and outliers. TV-L^2 regularization, on the other hand, can handle this situation well.

In order to overcome this deficiency of persistence simplification, while remaining conceptually close to the idea of persistence, we offer a novel significance weight, which we call *filling significance*, together with a corresponding denoising algorithm. We outline the relation of our approach to both traditional persistence simplification and TV-L^2 regularization, pointing out the similarities and dissimilarities between these approaches.

TOTAL VARIATION MINIMIZATION

We briefly outline the method of minimization of total variation regularizing functionals for signal denoising in one dimension. In the sequel we work with an open and bounded domain $\Omega = (a,b) \subset \mathbb{R}$ and piecewise constant functions $f : \overline{\Omega} \to \mathbb{R}$. For a given noisy function f the denoised function $u \in BV(\Omega)$ shall be computed as the minimizer of the

[1] Author supported by the project WWTF Five senses-Call 2006 project No. CI06 003 and the DFG Graduiertenkolleg 1023. Publication partially based on work supported by Award No. KUK-I1-007-43, made by King Abdullah University of Science and Technology (KAUST).

CP1281, *ICNAAM, Numerical Analysis and Applied Mathematics, International Conference 2010*
edited by T. E. Simos, G. Psihoyios, and Ch. Tsitouras

functional

$$\mathscr{J}(u) := \int_\Omega (u-f)^2\,dx + 2\alpha\,|D(u)|\,(\Omega), \tag{1}$$

where $BV(\Omega)$ is the space of functions of bounded variation and $|Du|(\Omega)$ denotes the total variation of u on Ω, see [2]. Here, the parameter α controls the strength of the regularizing effect of the total variation term in (1) on the given function f.

The minimization of total variation regularizing functionals like in (1) traces back to the first uses of such a functional model in noise removal in digital images as proposed by Rudin, Osher, and Fatemi [3]. Extensions to more general situations as appearing in inverse problems and numerical methods for the minimization of the functional have been proposed later in several important contributions, e.g., in [4]. From these pioneering and very successful results, the scientific output related to total variation minimization and its applications in signal and image processing increased dramatically in the last decade. In this short note, it is impossible to do justice to the bulk of previous work.

Characterization of minimizers. A well-known fact is the equivalence of TV-L^2 minimization with the taut string algorithm in one dimension, see [5] and references therein. In the following, we concentrate on the continuous setting discussed in [5]. In the taut string problem one considers the integral of the datum f, i.e.,

$$F(t) := \int_a^t f(x)\,dx, \quad a \le t \le b.$$

Then one seeks for a $U : \overline{\Omega} \to \mathbb{R}$ which minimizes

$$\tilde{\mathscr{J}}(U) := \int_a^b \sqrt{1+U'(t)^2}\,dt \tag{2}$$

subject to

$$U(a) = F(a),\; U(b) = F(b) \quad \text{and} \quad \max_{a\le t\le b} |U(t) - F(t)| \le \alpha. \tag{3}$$

The denoised function u is defined as the derivative of a minimizer U of (2) subject to (3). In [5] the author shows that TV-L^2 minimization (1) is equivalent to the taut string problem (2)-(3) in the sense that the minimizers of both problems are unique and the derivative u of the taut string minimizer U coincides with the minimizer of (1).

By condition (3) a minimizer U of (2)-(3) is contained in a tube of radius α around F, i.e., the residual $u-f$ integrated over an arbitrary interval in $[a,b]$ is uniformly bounded by 2α. In this sense condition (3) can be interpreted as a stable control for the deviation of a minimizer u of (1) from the given function f depending on the strength of regularization α. In the construction of our topological denoising algorithm we will preserve property (3), while providing a method that yields an output closer to the input f than the one provided by TV-L^2 regularization.

TOPOLOGICAL DENOISING

For our purposes, we consider a partition $P = (a = t_0 < t_1 < \cdots < t_n = b)$ of $\overline{\Omega} = [a,b]$ and we assume that $f : \overline{\Omega} \to \mathbb{R}$ is given as a piecewise constant function on this subdivision such that $f(t_i)$ is equal to the minimum value of f on the (at most two) incident open subintervals to t_i, i.e., f becomes lower semi-continuous. Without loss of generality we may assume that f takes on n distinct values $f_1 < f_2 < \cdots < f_n$, and we let $\Omega_i = f^{-1}(-\infty, f_i]$. Note carefully that we do not assume $f_i = f(t_i)$ here. The inclusions $\Omega_i \subset \Omega_j$, for $i \le j$, induce group homomorphisms $\iota_0^{i,j} : H_0(\Omega_i;\mathbb{Z}^2) \to H_0(\Omega_j;\mathbb{Z}^2)$ on the homology groups of degree zero. All higher homology groups vanish for our 1D setting. A homology class $h \in H_0(\Omega_i;\mathbb{Z}^2)$ is said to be *born* at Ω_i if h is not contained in the image of $\iota_0^{i-1,i}$. In this case i is called a *local minimum*. A class h that is born at Ω_i *dies* entering Ω_j if j is the smallest number such that $\iota_0^{i,j}(h)$ lies in the image of $\iota_0^{i-1,j}$ in $H_0(\Omega_j;\mathbb{Z}^2)$, see [1]. In this case j is called a *local maximum* and the pair (i,j) is called a persistence pair with *persistence* $(f_j - f_i) > 0$. The set of critical points of f is the set of all local minima and maxima.

Persistence-based denoising. A given input function f may be denoised by removing all those persistence pairs of critical points that have persistence below a user-specified threshold. We briefly describe the corresponding inductive construction of a finite sequence $(u^{(0)}, u^{(1)}, \dots)$ of piecewise constant functions, starting with $u^{(0)} = f$. In this construction, we assume that each $u^{(k)}$ takes on distinct values $u_1^{(k)} < \dots < u_{n_k}^{(k)}$. For each $\delta > 0$ and for each local minimum i of $u^{(k)}$, let $\Omega_i^{(k)}(\delta)$ be the connected component of ${u^{(k)}}^{-1}[u_i^{(k)}, u_i^{(k)} + \delta]$ that contains ${u^{(k)}}^{-1}(u_i^{(k)})$; similarly, for each local maximum j of $u^{(k)}$, let $\Omega_j^{(k)}(\delta)$ be the connected component of ${u^{(k)}}^{-1}[u_j^{(k)} - \delta, u_j^{(k)}]$ that contains ${u^{(k)}}^{-1}(u_j^{(k)})$. We call δ the *filling level*. Increasing the filling level continuously, starting with $\delta = \delta^{(0)} = 0$, there exists a smallest $\delta^{(k+1)} > 0$ and a corresponding local minimum i and a local maximum j of $u^{(k)}$ such that (i) $u_i^{(k)} + \delta^{(k+1)} = u_j^{(k)} - \delta^{(k+1)}$ and (ii) the intervals $\Omega_i^{(k)}(\delta^{(k+1)})$ and $\Omega_j^{(k)}(\delta^{(k+1)})$ merge. In this case, (i,j) is a persistence pair of $u^{(k)}$ with persistence $2\delta^{(k+1)}$. Once a (minimal) persistence pair has formed in this way, we replace $u^{(k)}$ by a function $u^{(k+1)}$ that is constant on the interval $\Omega_i^{(k)}(\delta^{(k+1)}) \cup \Omega_j^{(k)}(\delta^{(k+1)})$, taking the value $u_i^{(k)} + \delta^{(k+1)} = u_j^{(k)} - \delta^{(k+1)}$ there, and that agrees with $u^{(k)}$ everywhere else. We say that $u^{(k+1)}$ is constructed from $u^{(k)}$ by *canceling* the persistence pair (i,j). Notice that each persistence pair of $u^{(k)}$ corresponds to a persistence pair of f. Moreover, given a user-specified threshold $\delta \geq 0$, there exists a function $u^{(k)}$ that has no persistence pairs with persistence $\leq 2\delta$ and whose persistence pairs correspond to the persistence pairs of f with persistence $> 2\delta$. We call $u^{(k)}$ a δ-*persistence simplification* of the input f. We note that a δ-persistence simplification can be constructed in $O(n \log n)$ time, see [6].

Filling-based denoising. For f given as above, we consider the following process that mimics the construction of canceling persistence pairs. Differently from the above though, we steadily increase the filling *volume* instead of the filling *level*, and we sequentially remove *filling* pairs of critical points instead of *persistence* pairs. Let $u^{(0)} = f$. Given a critical point i of $u^{(k)}$, we call $\alpha_i^{(k)}(\delta) = \int_{\Omega_i^{(k)}(\delta)} (\delta - |u_i^{(k)} - u^{(k)}(x)|)\, dx$ the *filling volume* of i at filling level δ. Vice-versa, given a filling volume α, let $\delta_i^{(k)}(\alpha)$ denote the corresponding filling level at i. (Note that for sufficiently small α there is indeed a bijection between filling volume and filling level at each critical point.) Abusing the above notation, we let $\Omega_i^{(k)}(\alpha) := \Omega_i^{(k)}(\delta_i^{(k)}(\alpha))$. Increasing the filling volume α continuously, starting with $\alpha = \alpha^{(0)} = 0$, let $\alpha^{(k+1)}$ be the smallest filling volume such that there exists a local minimum i of $u^{(k)}$ and a nearby local maximum j such that (i) $u_i^{(k)} + \delta_i^{(k)}(\alpha^{(k+1)}) = u_j^{(k)} - \delta_j^{(k)}(\alpha^{(k+1)})$ and (ii) the intervals $\Omega_i^{(k)}(\alpha^{(k+1)})$ and $\Omega_j^{(k)}(\alpha^{(k+1)})$ merge. In this case, we call (i,j) a *filling pair* of $u^{(k)}$ with *filling significance* $\alpha^{(k+1)}$. Once a filling pair has formed in this way, we replace $u^{(k)}$ by a function $u^{(k+1)}$ that is constant on $\Omega_i^{(k)}(\alpha^{(k+1)}) \cup \Omega_j^{(k)}(\alpha^{(k+1)})$, taking the value $u_i^{(k)} + \delta_i^{(k)}(\alpha^{(k+1)}) = u_j^{(k)} - \delta_j^{(k)}(\alpha^{(k+1)})$ there, and that agrees with $u^{(k)}$ everywhere else. We say that $u^{(k+1)}$ is constructed from $u^{(k)}$ by *canceling* the filling pair (i,j). Given a user-specified threshold $\alpha \geq 0$, there exists a function $u^{(k)}$ that has no filling pairs with filling significance less or equal to α. We call $u^{(k)}$ an α-*filling simplification* of the input f. Similarly to δ-persistence simplifications, an α-filling simplification can be constructed in $O(n \log n)$ time.

COMPARISON

In the following statements, we consider an α-filling simplification sequence $(f = u^{(0)}, u^{(1)}, \dots)$, for increasing α, with corresponding filling volumes $0 = \alpha^{(0)} < \alpha^{(1)} \leq \alpha^{(2)} \leq \dots$. We obtain the following central result.

Theorem. *Each critical point i of $u^{(k)}$ corresponds to a unique critical point $c_k(i)$ of f and $u_i^{(k)} = f_{c_k(i)}$. Moreover, let $\Omega^{(k)} := \Omega_i^{(k)}(\alpha^{(k+1)}) \cup \Omega_j^{(k)}(\alpha^{(k+1)})$ and $\mu^{(k)} := u_i^{(k)} + \delta_i^{(k)}(\alpha^{(k+1)}) = u_j^{(k)} - \delta_i^{(k)}(\alpha^{(k+1)})$, where (i,j) is the filling pair of $u^{(k)}$. Then:*

1. *In analogy to the taut string solution* (3)*, we have*

$$U^{(k)}(a) = F(a),\, U^{(k)}(b) = F(b) \quad \text{and} \quad \max_{a \leq t \leq b} |U^{(k)}(t) - F(t)| \leq \alpha^{(k)}\,,$$

where $U^{(k)}$ is the antiderivative of $u^{(k)}$. More specifically, $u^{(k)}$ satisfies

$$u^{(k)}(x) = \begin{cases} \mu^{(m)} & \text{if } x \in \Omega^{(m)} \text{ for } m = \max_{l<k} l : x \in \Omega^{(l)}, \\ f(x) & \text{otherwise.} \end{cases}$$

2. *Let $\tilde{u}^{(\alpha)}$ be the minimizer of* (1) *and let k be the largest integer such that $\alpha \geq \frac{\alpha^{(k)}}{2}$. Furthermore, we let $\Omega^{(k)}(\alpha) = \{x \in \overline{\Omega} \,|\, x \in \Omega_i^{(k)}(\alpha) \text{ for some critical point } i \text{ of } u^{(k)}\}$. Then*

$$\tilde{u}^{(\alpha)}(x) = \begin{cases} f_{c_k(i)} \pm \delta_i^{(k)}(\alpha) & \text{if } x \in \Omega_i^{(k)}(\alpha) \text{ for some critical point } i \text{ of } u^{(k)}, \\ \mu^{(m)} & \text{if } x \notin \Omega^{(k)}(\alpha) \text{ and } x \in \Omega^{(m)} \text{ for } m = \max_{l<k} l : x \in \Omega^{(l)}, \\ f(x) & \text{otherwise,} \end{cases}$$

where the sign in the first case is positive if i is a local minimum and negative otherwise.

In particular, if $\alpha = \frac{\alpha^{(k)}}{2}$, then $u^{(k)}$ agrees with f whenever $\tilde{u}^{(\alpha)}$ does. However, $u^{(k)}$ is a *local* modification of f, whereas $\tilde{u}^{(\alpha)}$ is a *global modification*. Figure 1 depicts an example of this effect – while total variation denoising changes the signal in a neighborhood of *every* local extremum of f, our filling-based approach only changes the function where such a change actually leads to a *cancelation* of local extrema. Finally, Figure 2 gives a glance at a first attempt to apply ideas of filling-based denoising to two-dimensional image data.

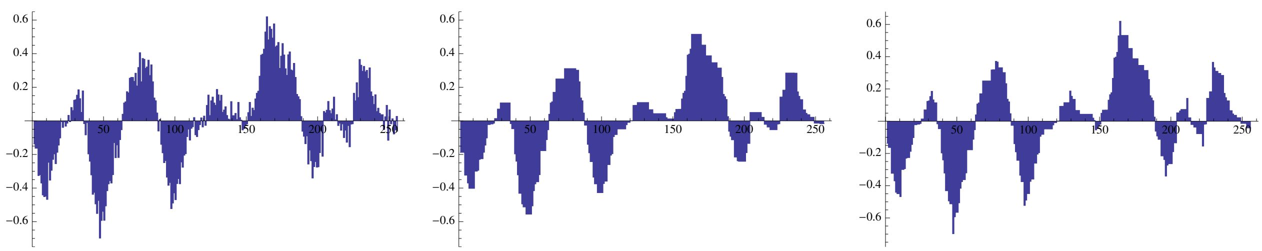

FIGURE 1. Signal denoising (f.l.t.r.): given noisy signal f; TV-L^2 denoising with $\alpha = 0.1$; solution of proposed filling-based denoising with $\alpha = 0.2$.

FIGURE 2. Image denoising (f.l.t.r.): given noisy f; TV-L^2 denoising; solution of proposed filling-based denoising.

REFERENCES

1. H. Edelsbrunner, D. Letscher, and A. Zomorodian, *Discrete and Computational Geometry* **28**, 511–533 (2002), URL `http://dx.doi.org/10.1007/s00454-002-2885-2`.
2. L. Ambrosio, N. Fusco, and D. Pallara, *Functions of Bounded Variation and Free Discontinuity Problems (Oxford Mathematical Monographs)*, Oxford University Press, USA, 2000, ISBN 0198502451, URL `http://www.amazon.com/exec/obidos/redirect?tag=citeulike07-20\&path=ASIN/0198502451`.
3. L. I. Rudin, S. Osher, and E. Fatemi, *Physica D Nonlinear Phenomena* **60**, 259–268 (1992), ISSN 0167-2789, URL `http://dx.doi.org/10.1016/0167-2789(92)90242-F`.

4. L. Vese, *Applied Mathematics and Optimization* **44**, 131–161 (2001), ISSN 0095-4616, URL `http://dx.doi.org/10.1007/s00245-001-0017-7`.
5. M. Grasmair, *Journal of Mathematical Imaging and Vision* **27**, 59–66 (2007), ISSN 0924-9907, URL `http://dx.doi.org/10.1007/s10851-006-9796-4`.
6. U. Bauer, C. Lange, and M. Wardetzky, Optimal topological simplification of discrete functions on surfaces, arXiv preprint (2010), URL `http://arxiv.org/abs/1001.1269`, `1001.1269`.