6th International Workshop on New Computational Methods for Inverse Problems IOP Publishing
Journal of Physics: Conference Series **756** (2016) 012007 doi:10.1088/1742-6596/756/1/012007

Total Variation Denoising and Support Localization of the Gradient

A. Chambolle[1], **V. Duval**[2], **G. Peyré**[2,3] **and C. Poon**[2,3,4]

[1] CNRS and CMAP, Ecole Polytechnique 91128 Palaiseau Cedex, FRANCE
[2] INRIA, MOKAPLAN, Domaine de Voluceau Le Chesnay, FRANCE
[3] CNRS and CEREMADE, Université Paris-Dauphine, Place du Marechal De Lattre De Tassigny, 75775 Paris 16, FRANCE
[4] DAMTP, University of Cambridge, Wilberforce Road, Cambridge CB3 0DZ, UK

E-mail: `antonin.chambolle@polytechnique.fr`, `vincent.duval@inria.fr`, `{gabriel.peyre, clarice.poon}@ceremade.dauphine.fr`

Abstract. This paper describes the geometrical properties of the solutions to the total variation denoising method. A folklore statement is that this method is able to restore sharp edges, but at the same time, might introduce some staircasing (i.e. "fake" edges) in flat areas. Quite surprisingly, put aside numerical evidences, almost no theoretical result are available to backup these claims. The first contribution of this paper is a precise mathematical definition of the "extended support" (associated to the noise-free image) of TV denoising. This is intuitively the region which is unstable and will suffer from the staircasing effect. Our main result shows that the TV denoising method indeed restores a piece-wise constant image outside a small tube surrounding the extended support. Furthermore, the radius of this tube shrinks toward zero as the noise level vanishes and in some cases, an upper bound on the convergence rate is given.

1. Introduction

The total variation (TV) denoising method was introduced by Rudin, Osher and Fatemi in [1]. It is one of the first proposed non-linear image restoration method, and has had an enormous impact on shaping modern imaging sciences. Despite being quite old, this method is still routinely used today, and its popularity probably stems from both its simplicity and its ability to restore "cartoon-looking" images.

Total Variation Denoising The total variation of a function $u \in \mathrm{L}^2\left(\mathbb{R}^2\right)$ is defined as

$$J(u) \overset{\text{def.}}{=} \int_{\mathbb{R}^2} |Du| \overset{\text{def.}}{=} \sup \left\{ \int_{\mathbb{R}^2} u \operatorname{div} z \; ; \; z \in \mathscr{C}_c^1(\mathbb{R}^2, \mathbb{R}^2), \|z\|_\infty \leqslant 1 \right\}. \tag{1}$$

Given some noisy input function f, following [1], we are interested in the total variation denoising problem

$$\min_{u \in \mathrm{L}^2(\mathbb{R}^N)} \lambda J(u) + \frac{1}{2}\|u - f\|_{L^2}^2. \tag{$\mathcal{P}_\lambda(f)$}$$

Here, $\lambda > 0$ is the regularization parameter, and it should adapted by the user to the noise level.

In this paper, we present the results of [2], which studies the ability of TV denoising to restore the geometrical structures (in particular the edges) of some (typically unknown) noise-free function f by solving $\mathcal{P}_\lambda(f+w)$. Here w accounts for some additive noise in the image formation process, and is assumed to have a finite L^2 norm $\|w\|_{L^2}$.

Level lines in the low noise regime Although it is not very difficult to see that, as $\lambda \to 0^+$ and $\|w\|_{L^2} \to 0^+$, the solution $u_{\lambda,w}$ to $\mathcal{P}_\lambda(f+w)$ converges towards f in the L^2 topology, our goal is to describe this convergence more precisely: *is it possible to say that the level lines of $u_{\lambda,w}$ converge to those of f? In what sense? Moreover, does the support of $Du_{\lambda,w}$ converge towards the support of Df?*

1.1. Previous works

Jump sets stability A landmark result in the characterization the structural properties of TV denoising solutions is [3], which proves that TV regularization does not introduce jumps, i.e. the "jump set" of the solution of $(\mathcal{P}_\lambda(f))$ is included in that of the input f. A review of this result and extensions can be found in [4].

These results are however of little interest when f is replaced by a noisy function $f+w$, since the noise w, which is only assumed to be in L^2, might introduce jumps everywhere. It is actually the presence of this noise which is responsible for the "staircasing" effect, which creates spurious edges in flat area. The present paper aims to fill this theoretical gap by analyzing the impact of the noise on the gradient support of the solution to $\mathcal{P}_\lambda(f+w)$, when both $\|w\|_{L^2}$ and λ are not too large.

Explicit solutions and calibrable sets Of particular importance for the analysis of TV methods are indicator functions of sets, and their behavior under the regularization. Indicator functions which are invariant (up to a rescaling) under TV denoising define the so-called "calibrable" sets. More precisely, by denoting the perimeter of a set by $P(C) \overset{\text{def.}}{=} J(\mathbb{1}_C)$ and its Lebesgue measure by $|C|$, we say that $C \subset \mathbb{R}^2$ is a calibrable set if $h_C \mathbb{1}_C \in \partial J(\mathbb{1}_C)$, where $h_C = P(C)/|C|$; and if C is a calibrable set, then the solution of $\mathcal{P}_\lambda(\mathbb{1}_C)$ is $u_{\lambda,0} = (1-h_C\lambda)^+\mathbb{1}_C$. A full characterization of convex calibrable sets is given in [5]. It was also proved in [5] that if C is a convex set, then the solution u_λ to $\mathcal{P}_\lambda(\mathbb{1}_C)$ satisfies

$$\{u_\lambda \geqslant t\} = \emptyset \text{ if } t > 1 - \frac{\lambda}{R^*} \quad \text{and} \quad \{u_\lambda \geqslant t\} = C_{\lambda/(1-t)} \text{ if } 0 \leqslant t \leqslant 1 - \frac{\lambda}{R^*}$$

where $C_r = \bigcup \{B(x,r)\ ;\ B(x,r) \subseteq C\}$ and C_{R^*} is the maximal calibrable set in C. Throughout this paper, $B(x,r)$ denotes the ball of radius r centred at x. In this paper, we shall show that in the presence of L^2 additive noise, calibrable sets play the role of "stable" sets and the corresponding edges are well restored by TV denoising.

Integral error estimates The systematic study of noise stability of regularization schemes relies on the so-called *source condition* [6], which in the simple denoising setting, is the requirement that the subdifferential of J, $\partial J(f)$, should be non-empty. In the case of total variation regularization, the subdifferential of $J : \mathrm{L}^2(\mathbb{R}^2) \to \mathbb{R} \cup \{+\infty\}$ can be written as follows.

$$\partial J(0) = \left\{\operatorname{div} z\ ;\ z \in L^\infty(\mathbb{R}^2;\mathbb{R}^2),\ \operatorname{div} z \in L^2(\mathbb{R}^2),\ \|z\|_{L^\infty} \leqslant 1\right\},$$
$$\partial J(u) = \left\{v \in \partial J(0)\ ;\ \int_{\mathbb{R}^2} uv = J(u)\right\}.$$

For non-smooth regularizations over Banach spaces, stability studies with respect to the source condition started with the seminal paper of Burger and Osher [7] who show that the

Journal of Physics: Conference Series **756** (2016) 012007 doi:10.1088/1742-6596/756/1/012007

source condition implies stability of the solution according to the Bregman divergence associated to J. In particular, if the source condition is satisfied for $f \in L^2(\mathbb{R}^2) \cap BV(\mathbb{R}^2)$, so that there exists $v = -\operatorname{div} z \in \partial J(f) \neq \emptyset$, then one has support stability in each region U where $\sup_{x\in U} |z(x)| < 1$. Formally, we say that z is *non-degenerate* on U and smaller values of $\sup_{x\in U} |z(x)|$ will allow for stronger stability bounds. More precisely, it was proved in [7] that for $U \subset \mathbb{R}^2$ and $\delta > 0$ such that $\sup_{x\in U} |z(x)| < 1 - \delta$, the solution $u_{\lambda,w}$ to $\mathcal{P}_\lambda(f + w)$ satisfies

$$(1-\delta)\int_U |Du_{\lambda,w}| \leqslant \frac{\|w\|_{L^2}^2}{2\lambda} + \frac{\lambda\|v\|_{L^2}^2}{2} + \|w\|_{L^2}\|v\|_{L^2}.$$

We remark however that this result does not lead to a precise geometric characterization of the TV regularized solutions. In this paper, we shall describe when $u_{\lambda,w}$ is constant on U and hence, our analysis can be seen as a generalization and refinement of this source condition approach.

2. Duality and the minimal norm certificate

We assume for the remainder of this paper that $f \in L^2(\mathbb{R}^2)$ and that the source condition holds. Instead of studying the properties of any element in $\partial J(f)$ as done in [7], we consider the minimal norm certificate:

$$v_0 \overset{\text{def.}}{=} \operatorname{argmin}\{\|v\|_{L^2} \; ; \; v \in \partial J(f)\}. \tag{2}$$

The study of v_0 will lead to an understanding of the geometric properties of $u_{\lambda,w}$ in the *low noise regime*, where $\lambda_0 > 0$, $\alpha_0 > 0$, $(\lambda, w) \in D_{\lambda_0,\alpha_0}$ and

$$D_{\lambda_0,\alpha_0} \overset{\text{def.}}{=} \left\{(\lambda, w) \in \mathbb{R}_+ \times \mathrm{L}^2\left(\mathbb{R}^2\right) \; ; \; 0 \leqslant \lambda \leqslant \lambda_0 \quad \text{and} \quad \|w\|_{L^2} \leqslant \alpha_0\lambda\right\}. \tag{3}$$

This particular element v_0 naturally arises from the dual problems associated with $(\mathcal{P}_\lambda(f))$, which we now describe. The Fenchel-Rockafellar dual problem of $(\mathcal{P}_\lambda(f))$ is

$$\sup_{v\in\partial J(0)} \langle f, v\rangle - \frac{1}{2\lambda}\|v\|_{L^2}^2, \tag{$\mathcal{D}'_\lambda(f)$}$$

$$\text{or equivalently} \qquad \inf_{v\in\partial J(0)} \left\|\frac{f}{\lambda} - v\right\|_{L^2}^2 \tag{$\mathcal{D}_\lambda(f)$}$$

We remark that since the solution $v_{\lambda,w}$ to $(\mathcal{D}_\lambda(f + w))$ is simply the projection of f/λ onto a convex set, $v_{\lambda,w}$ exists and is unique. Furthermore, there is strong duality.

The limit of $(\mathcal{P}_\lambda(f))$ as $\lambda \to 0^+$ is the trivial problem

$$\min_{u\in L^2(\mathbb{R}^2)\cap BV(\mathbb{R}^2)} J(u) \quad \text{s.t.} \quad u = f, \tag{$\mathcal{P}_0(f)$}$$

having $u = f$ as the solution. The dual is

$$\sup_{v\in\partial J(0)} \langle f, v\rangle, \tag{$\mathcal{D}_0(f)$}$$

with solutions $\partial J(f)$. Again, there is strong duality, although there is no guarantee of uniqueness or existence of solutions to $(\mathcal{D}_0(f))$ as it is possible that $\partial J(f) = \emptyset$. Note that since the dual solution $v_{\lambda,w}$ of $(\mathcal{D}_\lambda(f+w))$ is the projection of $(f+w)/\lambda$ onto a convex set, the non-expansiveness of the projection yields

$$\forall(\lambda, w) \in \mathbb{R}_+^* \times \mathrm{L}^2\left(\mathbb{R}^2\right), \quad \|v_{\lambda,0} - v_{\lambda,w}\|_{L^2} \leqslant \frac{\|w\|_{L^2}}{\lambda} \leqslant \alpha_0.$$

As a result, the properties of $v_{\lambda,w}$ are governed by those of $v_{\lambda,0}$, and it turns out that the properties of $v_{\lambda,0}$ are governed, in the low noise regime, by those of the minimal L^2 norm solution to $(\mathcal{D}_0(f))$, as the next result hints.

Proposition 1. *[2] If the source condition holds, then* $\lim_{\lambda\to 0}\|v_{\lambda,0}-v_0\|_{L^2}=0$ *where* v_0*, the minimal norm certificate is defined as in* (2)*.*

The main point in studying the dual problems is that their solutions $v_{\lambda,w}$ are related to the primal solutions $u_{\lambda,w}$ by the extremality relations

$$v_{\lambda,w}\in\partial J(u_{\lambda,w}),\qquad v_{\lambda,w}=\frac{1}{\lambda}(f-u_{\lambda,w}),$$

which makes possible the analysis of the support of $Du_{\lambda,w}$, as described by the following proposition. The intuition behind our study of the minimal norm certificate is that when $v_{\lambda,w}$ is close to v_0, the support of $Du_{\lambda,w}$ will be governed by v_0.

Proposition 2 ([2])**.** *Let* $f,v\in L^2(\mathbb{R}^2)$*. For* $t>0$*, let* $F^{(t)}\overset{\text{def.}}{=}\{f\geqslant t\}$ *and for* $t<0$*, let* $F^{(t)}\overset{\text{def.}}{=}\{f\leqslant t\}$*. The following hold.*

(i) $v\in\partial J(f)$ *if and only if* $v\in\partial J(0)$ *and the level sets of* f *satisfy*

$$\forall t>0,\quad P(F^{(t)})=\int_{F^{(t)}}v,\quad \forall t<0,\quad P(F^{(t)})=-\int_{F^{(t)}}v.$$

(ii) $\mathrm{Supp}(Df)=\overline{\bigcup\left\{\partial^*F^{(t)}\ ;\ t\in\mathbb{R}\setminus\{0\}\right\}}\subseteq\overline{\bigcup\left\{\partial^*F\ ;\ |F|<+\infty\quad\textit{and}\quad P(F)=\pm\int_F v\right\}}$.

Examples of minimal norm certificates

- If C is a calibrable set, then $v_0=h_C\mathbb{1}_C$.
- If C is a convex set such that ∂C is of class $C^{1,1}$ and $\operatorname{ess\,sup}_p\kappa_{\partial C}(p)\leqslant c$ for some $c>0$, then there exists some R^* such that C_{R^*} (as defined in Section 1.1) is the maximal calibrable set inside C. In this case, $v_0(x)=1/r$ for $x\in\partial C_r$ and $r\in[0,R^*]$ and $v_0(x)=0$ otherwise.

3. Support stability

In this section, we first define the extended support and state our main result.

3.1. The extended support

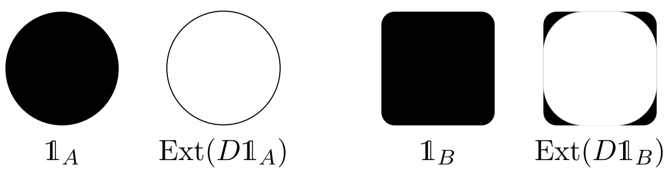
Motivated by Proposition 2, the extended support of $f\in L^2(\mathbb{R}^2)$ with respect to its minimal norm certificate v_0 is defined as follows.

Definition 1.

$$\mathrm{Ext}(Df)\overset{\text{def.}}{=}\overline{\bigcup\{\mathrm{supp}(Dg)\ ;\ v_0\in\partial J(g)\}}=\overline{\bigcup\left\{\partial^*E\ ;\ |E|<\infty,\ \pm\int_E v_0=P(E)\right\}}.$$

Examples of the extended support

- If C is a calibrable set, then $\mathrm{Ext}(D\mathbb{1}_C)=\partial C$.
- If C is a smooth convex set, then $\mathrm{Ext}(D\mathbb{1}_C)=C\setminus\mathrm{int}(C_{R^*})$, where C_{R^*} is the maximal calibrable set inside C.



Journal of Physics: Conference Series **756** (2016) 012007 doi:10.1088/1742-6596/756/1/012007

3.2. Main result

We are now ready to present our main result which relates support stability to the extended support.

Theorem 1. *[2] As $\lambda \to 0$ and $\|w\|_{L^2}/\lambda \to 0$,*

$$\sup_{x \in \mathrm{supp}(Df)} \mathrm{dist}(x, \mathrm{supp}(Du_{\lambda,w})) \to 0, \qquad \sup_{x \in \mathrm{supp}(Du_{\lambda,w})} \mathrm{dist}(x, \mathrm{Ext}\, Df) \to 0,$$

Furthermore, by denoting by $U^{(t)}_{\lambda,w}$ the level sets of $u_{\lambda,w}$, for almost every $t \in \mathbb{R}$, as $\lambda \to 0$ and $\|w\|_{L^2}/\lambda \to 0$,

$$\left| U^{(t)}_{\lambda,w} \Delta F^{(t)} \right| \to 0, \quad \text{and} \quad \lim_{n \to +\infty} \partial U^{(t)}_{\lambda,w} \to \partial F^{(t)},$$

and the last equality holds in the sense of Hausdorff convergence.

Remark 1. The first part of the theorem above implies that for all $r > 0$, there exists λ_r and α_r such that for all $(\lambda, r) \in D_{\lambda_r,\alpha_r}$, $\mathrm{Supp}(Du_{\lambda,w}) \subseteq T_r \overset{\text{def.}}{=} \left\{ x \in \mathbb{R}^2 \,;\, \mathrm{dist}(x, \mathrm{Ext}(Df)) \leqslant r \right\}$, and $\mathrm{Supp}(Df) \subseteq \left\{ x \in \mathbb{R}^2 \,;\, \mathrm{dist}(x, \mathrm{Supp}(Du_{\lambda,w})) \leqslant r \right\}$. Thus, one can always define an arbitrarily small tube T_r around the extended support such that $u_{\lambda,w}$ is constant outside T_r for all (λ, w) in some low noise regime.

3.3. Rate of convergence

By definition of v_0, there exists some vector field $z_0 \in L^\infty(\mathbb{R}^2; \mathbb{R}^2)$ such that $v_0 = -\operatorname{div} z_0$ and $\|z_0\|_{L^\infty} \leqslant 1$. This vector field is not unique and our main result is not dependent on its behaviour. However, our next result shows that when the non-degeneracy of z_0 is known (as described in Section 1.1), we can make explicit the relationship between the width of the tube T_r, the decay of $\|v_{\lambda,w} - v_{0,0}\|_{L^2}$ and the nondegeneracy of z_0. Assume now that for all $r > 0$, there exists $\delta_r > 0$ such that

$$1 - \delta_r = \operatorname*{ess\,sup}_{x \notin T_r} |z_0(x)| \,. \tag{4}$$

Theorem 2. *[2] If $\lambda > 0$ and $w \in L^2(\mathbb{R}^2)$ is such that $\|v_{\lambda,w} - v_0\|_{L^2} \leqslant \delta_{r/2} \min\{r/(2C),\ 2\sqrt{\pi}\}$, then* $\mathrm{Supp}(Du_{\lambda,w}) \subset T_r$.

Remark 2. Let us remark on the condition (4). As a result of the constructions described in [8], if C is a calibrable set, then the vector field z_0 of minimal norm certificate associated with $\mathbb{1}_C$ is such that every $K \subset C$ compact, $\sup_K |z_0| < 1$. Furthermore, we can show [2] that for a convex set C with $\mathscr{C}^2$ boundary, $|z_0(s,r)| \lesssim \min\{1 - \kappa_{\partial C}(s) r,\ 1 - r^2\}$ outside C.

Example: If $f = \mathbb{1}_{B(0,R)}$ and $\|w\|_{L^2} \leqslant \lambda r^2/(8R^2)$, then $\mathrm{supp}(Du_{\lambda,w}) \subset T_r$.

4. Numerical examples of the extended support

Figure 1 presents some numerical examples of denoised solutions, the extended support and the vector field z_0 associated with several indicator functions of sets. These examples are computed by solving the dual problem $(\mathcal{D}_\lambda(f))$ using the projected gradient descent algorithm of [9].

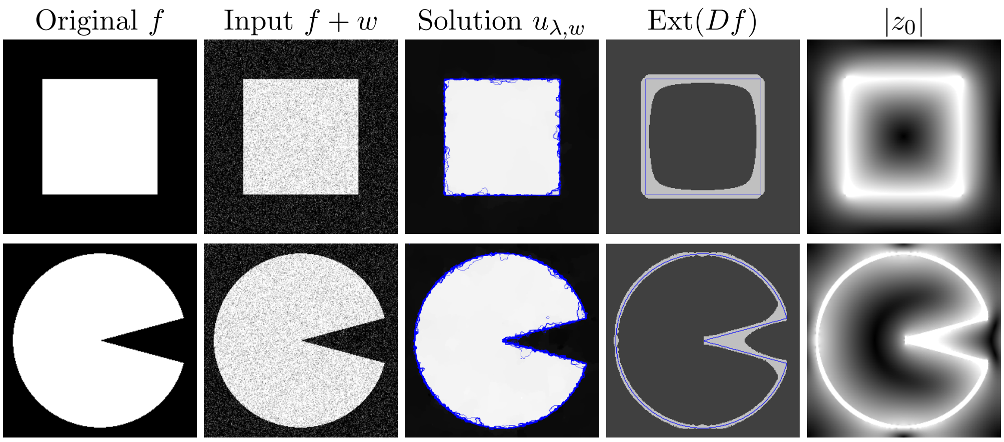


Figure 1: The level lines of each denoised solution are shown in blue. The extended support for each indicator function is shown in gray, and the absolute value of the associated vector field is shown such that white corresponds to one and black corresponds to zero. Observe that the decay of the vector field is linked to the curvature of the boundary as suggested by our remark in the previous section.

5. Conclusion

Our main result shows that although the support of TV regularized solutions are in general not stable and staircasing effects can be introduced, in the low noise regime, the instabilities are confined to a small neighbourhood of the extended support. Furthermore, for the indicator set of a calibrable set C, the support of TV regularized solutions cluster around ∂C. Finally, although the situation where the source condition is not satisfied was not discussed in the present paper, our approach can be extended to characterize the support stability for various functions without the source condition, such as $\mathbb{1}_{[0,1]^2}$ (see [2] for details).

Acknowledgements

V.D and C.P. acknowledge support from the CNRS project CAVALIERI. G.P. acknowledges support from the European Research Council (project SIGMA-Vision). C.P. acknowledges support from the FSMP.

References

[1] Rudin L I, Osher S and Fatemi E 1992 *Physica D: Nonlinear Phenomena* **60** 259–268
[2] Chambolle A, Duval V, Peyré G and Poon C 2016 *Preprint* URL `http://arxiv.org/abs/1602.00087`
[3] Caselles V, Chambolle A and Novaga M 2007 *Multiscale Modeling and Simulation* **6** 879–894
[4] Valkonen T 2015 *SIAM J. Math. Analysis* **47** 2587–2629
[5] Alter F, Caselles V and Chambolle A 2005 *Math. Ann.* **7** 29–53
[6] Scherzer O, Grasmair M, Grossauer H, Haltmeier M and Lenzen F 2009 *Variational Methods in Imaging* 1st ed Applied Mathematical Sciences (Springer) ISBN 0387309314
[7] Burger M and Osher S 2004 *Inverse problems* **20** 1411
[8] Giusti E 1978 *Invent. Math.* **46** 111–137
[9] Chambolle A 2004 *Journal of Mathematical imaging and vision* **20** 89–97