

An Approach to Vectorial Total Variation based on Geometric Measure Theory

Bastian Goldluecke and Daniel Cremers
TU Munich, Germany

bastian.goldluecke,daniel.cremers@in.tum.de

Abstract

We analyze a previously unexplored generalization of the scalar total variation to vector-valued functions, which is motivated by geometric measure theory. A complete mathematical characterization is given, which proves important invariance properties as well as existence of solutions of the vectorial ROF model. As an important feature, there exists a dual formulation for the proposed vectorial total variation, which leads to a fast and stable minimization algorithm. The main difference to previous approaches with similar properties is that we penalize across a common edge direction for all channels, which is a major theoretical advantage. Experiments show that this leads to a significantly better restoration of color edges in practice.

1. Introduction

Regularity is of central importance in computer vision. Many problems, like denoising, deblurring, super-resolution and inpainting, are ill-posed, and require the choice of a good prior in order to arrive at sensible solutions. This prior often takes the form of a regularization term for an energy functional which is to be minimized. For optimization purposes, it is important that the regularizer is convex, since only then one can hope to always find a global optimum of the energy within reasonable time. Furthermore, images in the real-world can be observed to generally be piecewise smooth. For these reasons, the *total variation* (TV) of a function has emerged as a very successful regularizer for a wide range of applications. It is convex, but still discontinuity-preserving, as it assigns the same cost to sharp and smooth transitions.

While most existing work focuses on scalar valued functions, the generalization to vector valued (color or multi-channel) images remains an important challenge. In this paper, we will see that this generalization is by no means straightforward, but that a very natural one can be obtained by considering the mathematical framework of geometric measure theory.



Figure 1: For inverse problems like denoising, inpainting or superresolution the total variation is among the most powerful regularizers. We propose a novel generalization of total variation to vector-valued images which naturally arises in the context of geometric measure theory.

Vectorial TV.

For a greyscale image modeled as a differentiable function $u : \Omega \rightarrow \mathbb{R}$ on a domain $\Omega \subset \mathbb{R}^m$, the scalar total variation $\text{TV}(u)$ is defined as the integral over the Euclidean norm $|\cdot|_2$ of the gradient,

$$\text{TV}(u) = \int_{\Omega} |\nabla u|_2 \, dx. \quad (1)$$

The definition can be extended to locally integrable functions $u \in \mathcal{L}_{\text{loc}}^1(\Omega, \mathbb{R})$ using a dual formulation: Let \mathbb{E}^m denote the closed unit ball in \mathbb{R}^m , then

$$\text{TV}(u) = \sup_{\xi \in \mathcal{C}_c^1(\Omega, \mathbb{E}^m)} \left\{ \int_{\Omega} u \operatorname{div}(\xi) \, dx \right\}. \quad (2)$$

Note that the right-hand side makes sense for non-differentiable u , as derivatives are taken only of the dual vector fields ξ .

The idea of vectorial total variation is to extend the above definition to vector-valued $\mathbf{u} : \Omega \rightarrow \mathbb{R}^n$, such that in the case $n = 1$ both definitions coincide. Several variants have been proposed, which will be discussed in Section 2.

An important criterion for a good regularizer is that efficient and reliable minimization algorithms are available. For the scalar TV, such methods have been developed based on the dual formulation, pioneered by Chan *et al.* [7]. The more recent algorithm by Chambolle [6] allows to handle

the non-differentiability of $|\cdot|_2$ without need to regularize, so one can solve the exact model. In view of this, a useful generalization of scalar to vectorial TV should have a similar dual formulation available.

Contributions.

In this work, we analyze a previously unexplored variant of vectorial TV which is based on geometric measure theory. With regard to this mathematical framework, our approach is probably the most natural generalization of the scalar case. We present an in-depth mathematical characterization, which is sufficient to prove the existence and uniqueness of solutions for the vectorial ROF model. The norm has all important invariance properties. We derive a dual definition, which gives rise to an efficient optimization algorithm. Furthermore, we can show that denoising using the new regularizer leads to improved restoration of color edges. Since the regularizer can be used as a substitute for vectorial TV in any energy functional, there is a broad spectrum of further applications, see e.g. [3, 9, 13, 20, 22].

2. Related Work

Previous approaches to define total variation for vector-valued functions can roughly be divided into two classes. The first class of approaches computes the total variation channel-by-channel, and takes a suitable norm of the resulting vector. The second class of approaches integrates a pointwise function of the derivative of \mathbf{u} , and emerges when considering the Riemann geometry of the image manifold.

Channel-by-channel, l^1 -norm.

Probably the most simple and straightforward way to deal with multidimensional TV is to sum up the contributions of the separate channels [2]. This leads to the definition

$$\text{TV}_S(\mathbf{u}) := \sum_{i=1}^n \text{TV}(u_i). \quad (3)$$

The dual formulation follows immediately from (2),

$$\text{TV}_S(\mathbf{u}) = \sup_{(\xi_1, \dots, \xi_n) \in K_S} \left\{ \sum_{i=1}^n \int_{\Omega} u_i \operatorname{div}(\xi_i) \, dx \right\} \quad (4)$$

with $K_S = \mathcal{C}_c^1(\Omega, \mathbb{E}^m \times \dots \times \mathbb{E}^m)$.

The dual optimization technique of this method is a straightforward generalization of [6], fast, robust and easy to implement. However, since there is no coupling between channels, there is no preservation of color edges and significant color smearing, as we will see in the experiments. Also, there is no rotational invariance in color space. The norm has for example been used in TV- L^1 optic flow models [22], where artifacts are not immediately visible as in color image denoising.

Channel-by-channel, Euclidean norm.

Blomgren and Chan [4] define multi-dimensional TV as the Euclidean norm of the vector of channel-wise scalar TV.

$$\text{TV}_{\text{BC}}(\mathbf{u}) := \sqrt{\sum_{i=1}^n \text{TV}(u_i)^2}. \quad (5)$$

From the Euler-Lagrange equations of this norm,

$$\frac{\text{TV}(u_i)}{\text{TV}_{\text{BC}}(\mathbf{u})} \operatorname{div} \left(\frac{\nabla u_i}{|\nabla u_i|_2} \right) = 0 \text{ for all } 1 \leq i \leq n, \quad (6)$$

one can observe that there is a coupling of channels, but it is global, i.e. the same per-channel weight is used for all image pixels. Although the authors demonstrated that it has quite a few desirable properties, there is no dual formulation available. Thus, there is no efficient and exact minimization algorithm, since a regularization is required for the denominator of (6).

Riemann Geometry.

In [8], di Zenzo suggests to consider a vector-valued image as a parameterized 2-dimensional Riemann manifold in nD -space. The metric tensor of this manifold is given by

$$g_{\mu\nu} = (\partial_{\mu} \mathbf{u}, \partial_{\nu} \mathbf{u}), \quad \mu, \nu = 1, 2. \quad (7)$$

This is analogous to the structure tensor of an image, and the Eigenvector corresponding to the smaller Eigenvalue gives the direction of the vectorial edge. Several variants of anisotropic and edge-enhancing diffusion for color images have been developed using this formulation [18, 21], but in general the diffusion process does not arise as the minimizing flow of an energy.

Based on this framework, Sapiro [16, 17] suggests a family of possible definitions for the vectorial TV, which is of the form

$$\text{TV}_{\text{SR}}(\mathbf{u}) := \int_{\Sigma} f(\lambda_+, \lambda_-) \, ds, \quad (8)$$

where λ_{\pm} denote the larger and smaller Eigenvalue of $(g_{\mu\nu})$, respectively, and f is a suitable scalar-valued function. It must be noted that TV_{SR} is in general only defined for differentiable functions, only for special cases are dual formulations available to extend it to locally integrable functions.

Another approach based on Riemann geometry was pioneered by Sochen *et al.* [19]. It considers the graph of the image as a surface in $\mathbb{R}^2 \times \mathbb{R}^n$, and minimizes the surface area in order to smoothe the image. This approach leads to a diffusion equation with the direction given by the Beltrami flow. It is possible to derive a dual formulation in the case of grayscale images [5], however no such dualization is known in the vectorial case.

Pointwise Frobenius Norm.

A special case of the Sapiro-Ringach TV (8) is the choice $f(\lambda_{\pm}) = \sqrt{\lambda_+ + \lambda_-}$, i.e. the Frobenius norm of the derivative $D\mathbf{u}$,

$$\text{TV}_F(\mathbf{u}) := \int_{\Omega} \|D\mathbf{u}(x)\|_F dx. \quad (9)$$

It is remarkable because there is a convenient dual formulation, which extends the definition from differentiable to locally integrable functions,

$$\text{TV}_F(\mathbf{u}) = \sup_{(\xi_1, \dots, \xi_n) \in K_F} \left\{ \sum_{i=1}^n \int_{\Omega} u_i \text{div}(\xi_i) dx \right\} \quad (10)$$

with $K_F = \mathcal{C}_c^1(\Omega, \mathbb{E}^{n \cdot m})$.

This dual formulation is a natural generalization in view of the dual formulation for single-channel TV. Notably, definitions (10) and (4) are equivalent in the sense that they lead to the same space $\mathcal{BV}(\Omega, \mathbb{R}^n)$. However, the actual results in image processing algorithms are quite different, since there is no correlation of channels in TV_S , in contrast to TV_F , which has a desirable coupling of channels. In books, both definitions appear depending on the preference of the authors [1, 2].

Efficient minimization techniques for functionals based on (10) have been intensively studied by Bresson and Chan [5] as well as Duval *et al.* [10]. Because of its good performance, TV_F has emerged as a favourite candidate for vectorial TV, and it is often referred to as *the* vectorial total variation, although it is only a single one in the large family (8) proposed by Sapiro and Ringach.

However, it was already noted by Blomgren and Chan [4] that TV_F has some less-than-ideal properties. In particular, it actually favors grey value images over colored ones, which leads to color smearing for example in denoising applications. That the coupling of channels is not optimal can also be seen in the dual formulation. While the edge strength is correctly weighted over all channels such that common edges are not overly penalized, the edge directions can be different for the different channels.

3. Vectorial Total Variation

In the following, we will propose a novel generalization of TV to vector-valued functions which emerges naturally in the light of geometric measure theory. The resulting vectorial TV approach supports a common edge direction for all channels, comprises important invariance properties and comes with a dual formulation that allows for stable and exact minimization schemes.

Definition.

Geometric measure theory [11] studies geometric properties of the measures of sets, for example arc length and

area. One of its important concepts is the notion of a *Jacobian* J_k , a generalization of the Jacobian determinant to the case $k \leq n$. We will only require the case $k = 1$, which we are going to explain in the following. For a scalar valued differentiable function u , we have $J_1 u = |\nabla u|_2$. Thus, it is natural to define the total variation of a vector-valued function $\mathbf{u} : \Omega \rightarrow \mathbb{R}^n$ as the integral

$$\text{TV}_J(\mathbf{u}) := \int_{\Omega} J_1 \mathbf{u} dx. \quad (11)$$

The precise meaning is made clear by the following proposition.

Proposition 3.1. *For functions $\mathbf{u} \in \mathcal{C}^1(\Omega, \mathbb{R}^n)$, the vectorial total variation equals the integral over the largest singular value of the derivative matrix,*

$$\text{TV}_J(\mathbf{u}) = \int_{\Omega} \sigma_1(D\mathbf{u}) dx. \quad (12)$$

In particular, TV_J is equal to the standard total variation for real-valued functions.

Proof. See appendix. \square

Interestingly, although the motivation for our model comes from a completely different direction, Proposition 3.1 shows an intimate relationship of our model to the Sapiro-Ringach approach [18]. To see this, note that the metric tensor of the image manifold is equal to $(g_{\mu\nu}) = (D\mathbf{u})^T D\mathbf{u}$, in particular $\sigma_1(D\mathbf{u}) = \sqrt{\lambda_+}$. Thus, TV_J is a special case of (8), similar to the model TV_F . However, from the new context, we can derive a dual formulation for TV_J , which leads to a very efficient optimization method for the proposed regularizer. Furthermore, it allows to extend the definition to non-differentiable functions, which is not possible in the Sapiro-Ringach formulation.

Proposition 3.2. *On $\mathcal{C}^1(\Omega, \mathbb{R}^n)$, the vectorial total variation can be expressed as*

$$\text{TV}_J(\mathbf{u}) = \sup_{(\xi, \eta) \in K_J} \left\{ \sum_{i=1}^n \int_{\Omega} u_i \text{div}(\eta_i \xi) dx \right\} \quad (13)$$

with $K_J = \mathcal{C}_c^1(\Omega, \mathbb{E}^m \times \mathbb{E}^n)$.

The right hand side is well defined for all $\mathbf{u} \in \mathcal{L}_{loc}^1(\Omega, \mathbb{R}^n)$.

Proof. Follows from the representation (17) for $J_1 \mathbf{u}$ and the Gaussian divergence theorem. \square

Because $\sigma_1(\cdot)$ and $\|\cdot\|_F$ are equivalent norms on $\mathbb{R}^{n \times m}$, TV_J leads to the same space $\mathcal{BV}(\Omega, \mathbb{R}^n)$ of functions of bounded vectorial TV.

Convexity and lower semi-continuity.

We now investigate the mathematical properties of $\text{TV}_J(\mathbf{u})$. The first observation shows that our definition retains important basic properties of the total variation.

Setup scalar fields $u_i := f_i$, $\eta_i := \frac{f_i}{|f_i|_2}$ for $1 \leq i \leq n$, and $\xi_j = 0$ for $1 \leq j \leq m$, respectively. Then iterate until convergence:

Primal update

$$u_i \leftarrow f_i + \lambda \operatorname{div}(\eta_i \xi) \text{ for all } 1 \leq i \leq n$$

Dual gradient ascent in η and ξ

$$\eta_i \leftarrow \eta_i + \tau \xi \cdot \nabla u_i \text{ for all } 1 \leq i \leq n$$

$$\xi \leftarrow \xi + \tau \sum_{i=1}^n \eta_i \nabla u_i$$

Projection of (ξ, η) onto K_J

$$\eta \leftarrow \frac{\eta}{|\eta|_2} \text{ and } \xi \leftarrow \frac{\xi}{|\xi|_2}$$

Figure 2: Primal-dual algorithm for solving the TV_J - L^2 model.

Proposition 3.3. $TV_J : \mathcal{BV}(\Omega, \mathbb{R}^n) \rightarrow \mathbb{R}$ is a semi-norm, in particular it is convex. Furthermore, it is lower semi-continuous with respect to the weak topology.

Proof. See appendix. \square

Relations to other norms.

L^1 -regularity was shown to be superior to L^2 -regularization because deviations from the constant case are less heavily penalized. This preservation of structures lead to the celebrated results in structure-preserving denoising. Similarly the proposed vectorial TV is superior to existing vectorial TV approaches because deviations are less penalized. In particular, one can show that the proposed regularizer is systematically smaller than existing ones such that a better preservation of relevant structures is guaranteed, see Figure 3.

Proposition 3.4. For all $u \in \mathcal{BV}(\Omega, \mathbb{R}^n)$,

$$TV_J(u) \leq TV_F(u) \leq TV_S(u). \quad (14)$$

Proof. The first inequality follows directly from the primal definition because $\sqrt{\lambda_+} \leq \sqrt{\lambda_+ + \lambda_-}$. The second inequality follows from the primal-dual formulation, because the set of dual variables considered in TV_S includes and is strictly larger than the set considered in TV_F (and also than that in TV_J). \square

Invariance properties.

We show that $TV_J(\mathbf{u})$ is invariant with respect to both affine transformations in *image space* as well as orthogonal transformations in *color space*. Those important properties are not shared by all variants of the vectorial TV, see Blomgren and Chan [4] for an analysis. The first proposition shows that TV_J is invariant under affine reparametrizations of Ω .

Proposition 3.5. Let $\Omega' \subset \mathbb{R}^m$ and $A : \Omega' \rightarrow \Omega$ be affine and onto. Then for any $\mathbf{u} \in \mathcal{L}_{loc}^1(\Omega, \mathbb{R}^n)$,

$$TV_J(\mathbf{u} \circ A) = TV_J(\mathbf{u}).$$

Proof. Follows from representation (13) and the integral transformation formula. \square

Furthermore, $TV_J(\mathbf{u})$ is invariant under isometries of the color space \mathbb{R}^n . Note that this is incorrect for most variants of vectorial TV, a notable exception being TV_F .

Proposition 3.6. Let $T : \mathbb{R}^n \rightarrow \mathbb{R}^n$ be an isometry with respect to the Euclidean norm. Then

$$TV_J(T\mathbf{u}) = TV_J(\mathbf{u}).$$

Proof. Follows from the representation (17), note that $D(T\mathbf{u})(\mathbf{e}_j) = T \cdot D\mathbf{u}(\mathbf{e}_j)$ and T maps the unit sphere in \mathbb{R}^n onto itself. \square

Euler-Lagrange equations.

Of particular importance in many applications are the Euler-Lagrange equations of the functional TV_J . Formally, if one takes σ_1 as a function on $\mathbb{R}^{n \times m}$, one obtains a system of n partial differential equations, see e.g. [12]:

$$\operatorname{div} \left(\sum_{j=1}^m \partial_{ij} \sigma_1(D\mathbf{u}) \partial_j u_i \right) = 0, i = 1, \dots, n. \quad (15)$$

Unfortunately, as with $|\cdot|_2$, there are points where σ_1 is not differentiable, which means that one has to take a suitable regularization. Computationally, it is feasible only for $m \leq 2$, as then there is an explicit formula for the singular values [18]. Furthermore, the Euler-Lagrange equations are relatively expensive to evaluate, since one needs to compute the derivative of σ_1 with respect to all components of the Jacobian. It is thus fortunate that there is a fast algorithm to minimize the TV_J - L^2 denoising model based on the dual expression (13), which is analyzed in the next section.

4. The TV_J - L^2 Model

The ROF model, named after Rudin, Osher and Fatemi who introduced it in [15], is a very successful approach to image denoising, designed to restore an image which was contaminated by Gaussian noise. With the vectorial TV, it can easily be generalized to vector-valued images. This leads to the TV_J - L^2 model for denoising, which is to solve for a given noisy image $\mathbf{f} \in L^2(\Omega, \mathbb{R}^n)$ the minimization problem

$$\operatorname{argmin}_{\mathbf{u} \in \mathcal{BV}(\Omega, \mathbb{R}^n)} \left\{ TV_J(\mathbf{u}) + \frac{1}{2\lambda} \|\mathbf{u} - \mathbf{f}\|_2^2 \right\}. \quad (16)$$

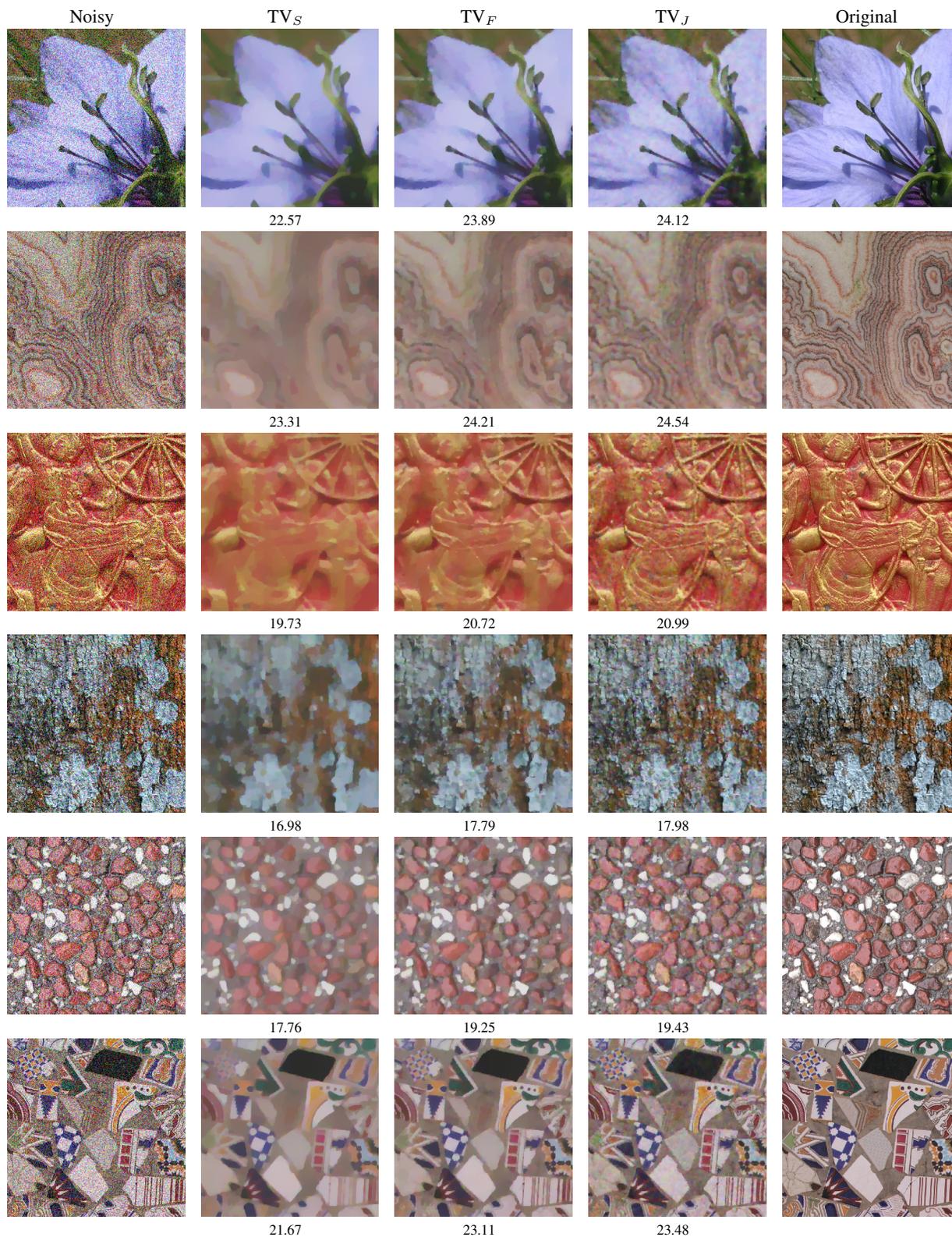


Figure 3: Denoising results for input images with additive Gaussian noise, standard deviation $\sigma = 0.2$. For each method, the value of λ which gave the best results was determined experimentally. The PSNR for each result is noted below the image. Although PSNR is only marginally better for TV_J compared to TV_F , the visual quality is significantly improved. TV_S is clearly inferior to both because of strong color smearing.



Blurred and noisy

Result

Original

Figure 4: *Deblurring using TV_J* . Colors are reproduced correctly in the result, and sharp edges are restored.

$\lambda > 0$ is a constant controlling the desired smoothness of the result - the larger, the greater the influence of the regularizer.

Existence and uniqueness of solutions.

Denote the (componentwise) mean of a function \mathbf{u} on Ω by $\mathbf{u}_\Omega := \frac{1}{|\Omega|} \int_\Omega \mathbf{u} \, dx$. We can then prove

Theorem 4.1. *Problem (16) admits a unique solution $\bar{\mathbf{u}}$ which has the same mean as \mathbf{f} .*

Proof. We give the basic ideas of the proof and refer the reader to [10] for more details. The energy is strictly convex as a sum of the convex TV_J seminorm, see Proposition (3.3), and the strictly convex norm of the L^2 -Hilbertspace. Thus, what remains to be shown is existence of a solution on the space

$$V_f := \{\mathbf{u} \in L^2(\Omega, \mathbb{R}^n) : \mathbf{u}_\Omega = \mathbf{f}_\Omega\}.$$

We already know that the energy is lower semi-continuous on V_f , Proposition (3.3). In addition, the energy is coercive on V_f : from the Poincare-inequality we have

$$\|\mathbf{u} - \mathbf{f}_\Omega\|_2 = \|\mathbf{u} - \mathbf{u}_\Omega\|_2 \leq C \cdot TV_J(\mathbf{u})$$

with a constant $C > 0$. Lower semi-continuity and coercivity imply the existence of a minimum $\bar{\mathbf{u}}$ in V_f . Since the constant function \mathbf{f}_Ω is of bounded variation and $\mathbf{f}_\Omega \in V_f$, it follows that $\bar{\mathbf{u}} \in \mathcal{BV}(\Omega, \mathbb{R}^n)$. \square

Efficient minimization.

The dual formulation of TV_J leads to a primal-dual optimization algorithm which is detailed in Figure 2. It can be derived in a similar way as the algorithm in [14]. The primal update and dual gradient ascent follow directly from the Euler-Lagrange equations for u_i , η_i and ξ , respectively, while the reprojection steps ensure that (η, ξ) remains in K_J . We found experimentally that the algorithm remains stable for $\tau = 1/8$, see also [6].

5. Experimental Results

We tested the vectorial ROF model with three different regularizers on dozens of images. The favourite alternative candidate to TV_J is TV_F for the reasons described

above, but for comparison we also included TV_S . Results are displayed in Figure 3. As expected, TV_S performs significantly worse than the other two, since there is no coupling of channels and significant smearing of colors. As previously observed by Blomgren and Chan [4], TV_F also exhibits a tendency to smear colors across edges, but it is less pronounced. The visually best results are achieved with TV_J . Since all color channels share a common edge direction, color edges are preserved and no color degradation can be observed.

To demonstrate some examples of additional applications for the regularizer, we also implemented vectorial versions of TV-deblurring [9] and TV-based superresolution [13]. Results are displayed in Figure 4 and Figure 5, respectively. One can again note that TV_J reproduces sharp color edges correctly without smearing colors.

6. Conclusion

In summary, we have explored a novel formulation for vectorial TV based on geometric measure theory. The formulation using this powerful framework yields a lot of additional theorems, which can be leveraged in future research. Interestingly, TV_J is also a special case of the family of norms proposed by Sapiro, which after TV_F makes it the second one for which a dual formulation is available. Thus, TV_J not only allows discontinuous solutions, it also gives rise to an efficient and stable minimization algorithm for the vectorial ROF model. Denoising results are superior in that color edges are preserved much better, and since TV_J can serve as a regularizer in any energy functional, there is a host of additional applications for it, which will be explored in the future.

References

- [1] L. Ambrosio, N. Fusco, and D. Pallara. *Functions of Bounded Variation and Free Discontinuity Problems*. Oxford University Press, 2000. 3
- [2] H. Attouch, G. Buttazzo, and G. Michaille. *Variational Analysis in Sobolev and BV Spaces*. MPS-SIAM Series on Optimization. Society for Industrial and Applied Mathematics, 2006. 2, 3
- [3] M. Bertalmio, L. Vese, G. Sapiro, and S. Osher. Simultaneous structure and texture image inpainting. *IEEE Transactions on Image Processing*, 12(8):882–889, 2003. 2
- [4] P. Blomgren and T. F. Chan. Color TV: Total variation methods for restoration of vector-valued images. *IEEE Trans. Image Processing*, 7:304–309, 1998. 2, 3, 4, 6
- [5] X. Bresson and T. F. Chan. Fast dual minimization of the vectorial total variation norm and applications to color image processing. *Inverse Problems and Imaging*, 2:455–484, 2008. 2, 3
- [6] A. Chambolle. An algorithm for total variation minimization and applications. *Mathematical Imaging and Vision*, 20:89–97, 2004. 1, 2, 6

(a) *Low-res and noisy*(b) *Initialization*(c) *Result*(d) *Original*

Figure 5: *Superresolution using TV_J* . (a) One of 16 low-resolution, noisy input images. (b) Initialization obtained by averaging the input images. (c) Result from running the algorithm. It can be observed that color edges are preserved very well and no color smearing occurs. (d) Original image.

- [7] T. F. Chan, G. H. Golub, and P. Mulet. A nonlinear primal-dual method for total variation-based image restoration. *SIAM Journal on Scientific Computing*, 20(6):1964–1977, 1999. 1
- [8] S. di Zenzo. A note on the gradient of a multi-image. *CVGIP*, 33:116–125, 1986. 2
- [9] D. Dobson and F. Santosa. Recovery of blocky images from noisy and blurred data. *SIAM Journal on Applied Mathematics*, 56(4):1181–1198, 1996. 2, 6
- [10] V. Duval, J.-F. Aujol, and L. Vese. Projected gradient based color image decomposition. In *Scale Space and Variational Methods in Computer Vision*, pages 295–306, May 2009. 3, 6
- [11] H. Federer. *Geometric Measure Theory*. Number 153 in Classics in Mathematics. Springer, 1996. reprint of the 1969 edition. 3
- [12] I. M. Gelfand and S. V. Fomin. *Calculus of Variations*. Dover Publications, 2003. reprint of the 1963 edition. 4
- [13] A. Marquina and S. Osher. Image super-resolution by TV-regularization and Bregman iteration. *Journal of Scientific Computing*, 37(3):367–382, 2008. 2, 6
- [14] T. Pock, T. Schoenemann, G. Graber, H. Bischof, and D. Cremers. A convex formulation of continuous multi-label problems. In *Proc. ECCV*, October 2008. 6
- [15] L. I. Rudin, S. Osher, and E. Fatemi. Nonlinear total variation based noise removal algorithms. *Physica D*, 60(1-4):259–268, 1992. 4
- [16] G. Sapiro. Color snakes. Technical report, Hewlett Packard Computer Peripherals Laboratory, 1995. 2
- [17] G. Sapiro. Vector-valued active contours. In *Proc. CVPR*, pages 680–685, 1996. 2
- [18] G. Sapiro and D. L. Ringach. Anisotropic diffusion of multi-valued images with applications to color filtering. *IEEE Trans. on Image Processing*, 5(11):1582–1586, 1996. 2, 3, 4
- [19] N. Sochen, R. Kimmel, and R. Malladi. A general framework for low level vision. *IEEE Trans. on Image Processing*, 7(3):310–318, 1998. 2
- [20] C. Vogel and M. Oman. Fast, robust total variation-based reconstruction of noisy, blurred images. *IEEE Transactions on Image Processing*, 7(6):813–824, 1998. 2
- [21] J. Weickert. Coherence-enhancing diffusion of colour images. *Image and Vision Computing*, 17(3-4):201–212, 1999. 2
- [22] C. Zach, T. Pock, and H. Bischof. A duality based approach for realtime $TV - L^1$ optical flow. In *Pattern Recognition (Proc. DAGM)*, pages 214–223, 2007. 2

Appendix

Proof of Proposition 3.1. We show how to compute $J_1 \mathbf{u}$:

$$\begin{aligned}
 J_1 \mathbf{u} &= \|\wedge_1 D\mathbf{u}\| \\
 &= \sup_{\xi \in \mathbb{E}^m} \{ |(\wedge_1 D\mathbf{u})(\xi_1 \mathbf{e}_1 + \dots + \xi_m \mathbf{e}_m)|_2 \} \\
 &= \sup_{(\xi, \boldsymbol{\eta}) \in \mathbb{E}^m \times \mathbb{E}^n} \left\{ \sum_{j=1}^m \xi_j (D\mathbf{u}(e_j), \boldsymbol{\eta}) \right\} \\
 &= \sup_{(\xi, \boldsymbol{\eta}) \in \mathbb{E}^m \times \mathbb{E}^n} \left\{ \sum_{i=1}^n \eta_i \sum_{j=1}^m \xi_j \partial_j u_i \right\} \quad (17) \\
 &= \sup_{(\xi, \boldsymbol{\eta}) \in \mathbb{E}^m \times \mathbb{E}^n} \left\{ \sum_{i=1}^n \eta_i (\boldsymbol{\xi} \cdot \nabla u_i) \right\} \\
 &= \sup_{(\xi, \boldsymbol{\eta}) \in \mathbb{E}^m \times \mathbb{E}^n} \left\{ \sum_{i=1}^n \boldsymbol{\eta}^T D\mathbf{u} \cdot \boldsymbol{\xi} \right\}
 \end{aligned}$$

The claim follows now from the singular value decomposition of $D\mathbf{u}$. \square

Proof of Proposition 3.3. Semi-norm properties follow immediately since J_1 is a norm. To show lower semi-continuity, take a sequence \mathbf{u}^n converging to \mathbf{u} in weak topology. In particular, for each $\boldsymbol{\eta}, \boldsymbol{\xi}$ and each $1 \leq i \leq n$ we have

$$\lim_{n \rightarrow \infty} \langle u_i^n, \operatorname{div}(\eta_i \boldsymbol{\xi}) \rangle = \langle u_i, \operatorname{div}(\eta_i \boldsymbol{\xi}) \rangle.$$

Thus,

$$\begin{aligned}
 &\liminf_{n \rightarrow \infty} TV_J(\mathbf{u}^n) \\
 &= \liminf_{n \rightarrow \infty} \left\{ \sup_{(\xi, \boldsymbol{\eta}) \in K_J} \sum \int_{\Omega} u_i^n \operatorname{div}(\eta_i \boldsymbol{\xi}) \, dx \right\} \\
 &\geq \sup_{(\xi, \boldsymbol{\eta}) \in K_J} \left\{ \liminf_{n \rightarrow \infty} \sum \int_{\Omega} u_i^n \operatorname{div}(\eta_i \boldsymbol{\xi}) \, dx \right\} \\
 &= \sup_{(\xi, \boldsymbol{\eta}) \in K_J} \sum \int_{\Omega} u_i \operatorname{div}(\eta_i \boldsymbol{\xi}) \, dx \\
 &= TV_J(\mathbf{u}).
 \end{aligned}$$

This completes the proof. \square