IMPROVED HIGHER DEGREE TOTAL VARIATION (HDTV) REGULARIZATION

Yue Hu and Mathews Jacob†*

*Department of Electrical and Computer Engineering, University of Rochester, NY, USA
†Department of Electrical and Computer Engineering, University of Iowa, IA, USA

ABSTRACT

The main focus of this paper is to further improve the performance of the recently introduced higher degree total variation (HDTV) penalties, which are L_1-$L_p; p \geq 1$ norms of directional image derivatives. We generalize this class as the L_1-L_p norms of image responses to rotated versions of an arbitrary derivative operator. We show that several penalties proposed by other researchers are special cases of the generalized isotropic penalties ($p = 2$), when the derivative operator is chosen appropriately; our experiments show that the anisotropic ($p = 1$) versions of these penalties provide improved reconstructions. In addition, we optimize the derivative operator for improved orientation selectivity, thus further improving the ability of the resulting penalties to provide high quality image reconstructions. We also focus on the efficient discretization of HDTV penalties, which are specified in the continuous domain. Specifically, we approximate the derivative operators as the sum of partial derivatives of an almost isotropic B-spline window. Our numerical experiments confirm the benefit of the improved discretization and the optimization of the operator.

Index Terms— Higher degree total variation, B-spline, compressed sensing

1. INTRODUCTION

The total variation (TV) regularization penalty is widely used in several biomedical imaging applications, including image denoising, deblurring, and reconstruction [1]. Since TV regularization favors images with sparse first order derivatives, the reconstructed images often suffer from staircase and patchy artifacts. To overcome this problem, we introduced a class of image regularization penalties termed as HDTV functionals [2]. The HDTV penalties are essentially the L_1-$L_p; p \geq 1$ norm of the n^{th} degree directional image derivatives. We showed that this class of penalties inherit the desirable properties of standard TV regularizers, including preservation of discontinuities, invariance to rotations and translations, and simplicity. We also introduced efficient image reconstruction algorithms, which exploit the rotation steerability of directional derivatives. The validation of the HDTV scheme on denoising and MR image recovery shows that the completely separable anisotropic (when $p = 1$) HDTV (termed as A-HDTV) provides consistently improved reconstructions compared to standard TV, current second degree regularization penalties, HDTV penalty with $p > 1$, and wavelet regularization. The main focus of this work is to further improve the performance of A-HDTV penalty in practical biomedical problems by **(a)** considering generalized derivative operators and **(b)** by efficiently discretizing the penalty.

We generalize the HDTV penalties by considering rotations of general steerable differential operators, which are linear combinations of n^{th} degree partial image derivatives. Since the directional derivative is the rotation of a high degree differential operator, the previous HDTV penalties are special cases of the generalization. We show that the generalized family of HDTV penalties includes many of the current regularization functionals introduced by other researchers [3, 4], when $p = 2$. We observe that the performance of these schemes can be further improved by considering the corresponding anisotropic penalties ($p = 1$). In addition, we introduce a design procedure to improve the HDTV penalty. Specifically, we optimize the shape of the steerable derivative operator, which is used in the penalty, to improve its orientation selectivity. Operators with improved orientation selectivity can provide better preservation of line-like, elongated features (edges and ridges) in the image.

In our earlier work, we specified the HDTV penalties in the continuous domain [2], similar to the classical TV penalty; the image was assumed to be a function $f(\mathbf{x})$ of the real variable $\mathbf{x}$, while the penalty is the continuous integral of the norms. We now focus on the discretization of this penalty and study the effect of discretization on the quality of the reconstructions. We observe that the standard finite difference approximation of partial derivatives, which is widely used in standard TV methods, is not rotation steerable. Specifically, the directional derivatives of the image along any orientation cannot be expressed as the weighted linear combination of the partial derivatives obtained using finite differences. Hence, the discretization of the TV and HDTV penalties using such operators is a poor approximation of the continuous domain formulation; this will in turn translate to poor reconstruc-

This work is supported by NSF awards CCF-0844812 and CCF-1116067.

tions. We propose to use partial derivatives of tensor product of B-spline windows to approximate the partial derivatives, in an effort to improve the steerability of the operators. Since B-spline windows are approximately isotropic, its partial derivatives are approximately steerable. At the same time, the discrete operators closely approximate the derivatives.

We compare the utility of the proposed penalties in the context of MR image recovery from sparse Fourier samples. We observe that the higher degree anisotropic ($p = 1$) penalties considerably reduced the patchy image artifacts in the reconstructions. We observe that using partial derivatives of B-spline windows provides improved reconstructions compared to classical finite difference approximations, thanks to the improved rotation steerability of such operators. We also observe that the use of optimized derivative operators provide better preservation of line-like features in the image.

2. BACKGROUND

We consider the recovery of a continuously differentiable complex image f from its noisy and undersampled measurements $\mathbf{b}$, specified by $\mathbf{b} = \mathcal{A}(f) + \mathbf{n}$. Here, $\mathbf{n}$ is Gaussian distributed white noise with the standard deviation of σ. One way to solve this ill-conditioned problem is to pose it as an optimization problem:

$$\hat{f} = \arg\min_f \underbrace{\|\mathcal{A}(f) - b\|^2 + \lambda\,\mathcal{J}(f)}_{\mathcal{C}(f)}. \tag{1}$$

The parameter λ is chosen such that $\|\mathcal{A}(\hat{f}) - b\|^2 \approx \sigma^2$. The standard TV regularization is essentially the L_1 norm of the image gradient, which is specified as $\mathcal{J}_1 = \int_\Omega |\nabla f(\mathbf{r})| d\mathbf{r}$. We derived the higher degree total variation (HDTV) regularization penalties in [5, 2]:

$$\mathcal{J}_{n,p}(f) = \int_\Omega \left(\frac{1}{2\pi}\int_0^{2\pi} |f_{\theta,n}(\mathbf{r})|^p \, d\theta\right)^{\frac{1}{p}} d\mathbf{r} \tag{2}$$

where $f_{\theta,n}(\mathbf{r})$ is the rotation steerable n^{th} order directional derivative along the unit vector $\mathbf{u}_\theta = (\cos\theta, \sin\theta)$. Note that (2) is a fully separable L_1-L_1 penalty when $p = 1$, which is thus called anisotropic HDTV (A-HDTV). Our experiments show that the A-HDTV penalty provides better image reconstructions than other HDTV penalties ($p > 1$), mainly because it is completely separable and hence is capable of better smoothing along line-like image features . We term the case with $p = 2$ as the isotropic version (I-HDTV).

3. THEORY

3.1. Generalized HDTV using derivative operators

The higher degree TV penalty in (2) is the sum of the norms of the directional image derivatives; the second degree directional derivative operator along a specific orientation is essentially the rotated version of the second derivative along the y axis. We now generalize this class by considering the rotated versions of an arbitrary differential operator. We focus on second order case in this paper. The general operator in the space of second degree derivatives is given by $\mathcal{D}_{0,2}(\mathbf{r}) = \alpha_1\,\partial_{xx}(\mathbf{r}) + \alpha_2\,\partial_{yy}(\mathbf{r}) + \alpha_3\,\partial_{xy}(\mathbf{r})$. The generalized second degree TV penalty for $p \geq 1$ is specified by

$$\mathcal{G}_{\mathcal{D},p}(f) = \int_\Omega \left(\frac{1}{2\pi}\int_0^{2\pi} |\mathcal{D}_{\theta,2}(\mathbf{r}) * f(\mathbf{r})|^p \, d\theta\right)^{\frac{1}{p}} d\mathbf{r}, \tag{3}$$

where $\mathcal{D}_{\theta,2}(\mathbf{r})$ is the rotated version of this operator. We now show that this definition includes many of the current extensions of standard TV norm, when $p = 2$.

3.1.1. Laplacian penalty

The operator $\mathcal{D}_{0,2}(\mathbf{r}) = \partial_{xx}(\mathbf{r}) + \partial_{yy}(\mathbf{r})$ is the Laplacian when $\alpha_1 = \alpha_2 = 1; \alpha_3 = 0$. The corresponding penalty

$$\mathcal{G}_{\mathcal{D},2}(f) = \int_\Omega |\triangle f(\mathbf{r})| d\mathbf{r}, \tag{4}$$

was introduced for image denoising in [4]. This penalty term has two major disadvantages. First of all, it has a large null space. Specifically, any function that satisfies the Laplace equation ($\triangle f(\mathbf{r}) = 0$) will result in $\mathcal{G}_{\mathcal{D},2}(f) = 0$. As a result, the use of this regularizer to constrain general ill-posed inverse problems is not desirable. Another problem is that the detector being isotropic, its use results in the enhancement of point-like features rather than line-like features.

3.1.2. Lysaker's second degree penalty

Another interesting case corresponds to $\alpha_1 = 1$, $\alpha_2 = 2\sqrt{2} - 3$, and $\alpha_3 = 0$. The isotropic ($p = 2$) penalty is thus given by

$$\mathcal{G}_{\mathcal{D},2}(f) = \int_\Omega \sqrt{\frac{1}{2\pi}\int_0^{2\pi} |f_{\theta,2}(\mathbf{r}) - \alpha_2 f_{\theta^\perp,2}(\mathbf{r})|^2 \, d\theta}, d\mathbf{r}, \tag{5}$$

where $f_{\theta^\perp,2}(\mathbf{r})$ is the second derivative of f along $\theta^\perp = \theta + \frac{\pi}{2}$. Using the rotation steerability of second degree directional derivatives $f_{\theta,2}(\mathbf{r}) = f_{xx}(\mathbf{r})\cos^2\theta + f_{yy}(\mathbf{r})\sin^2\theta + 2f_{xy}(\mathbf{r})\cos\theta\sin\theta$, the expression for $\mathcal{G}_{\mathcal{D},2}(f)$ simplifies to

$$\mathcal{G}_{\mathcal{D},2}(f) = c\int_\Omega \sqrt{|f_{xx}(\mathbf{r})|^2 + |f_{yy}(\mathbf{r})|^2 + 2\,|f_{xy}(\mathbf{r})|^2}\, d\mathbf{r}. \tag{6}$$

where $c = \sqrt{(6 - 4\sqrt{2})}$. This functional can be expressed as $\mathcal{G}_{\mathcal{D},2}(f) = \int_\Omega \|\nabla^2 f\|_F \, d\mathbf{r}$, where $\nabla^2 f$ is the Hessian matrix of $f(\mathbf{r})$ and $\|\cdot\|_F$ is the Frobenius norm. This second order penalty was proposed by [3]. (6) can also be thought of as the straightforward extension of the classical second-degree Duchon's seminorm [6]. Our experiments demonstrate that the corresponding anisotropic penalty $\mathcal{G}_{\mathcal{D},1}(f)$ provides superior reconstructions than $\mathcal{G}_{\mathcal{D},2}(f)$. We will consider the derivation of the optimal penalty in Section 3.2.1.

3.2. Steerable approximation of derivatives

The TV and HDTV penalties are essentially defined for continuous images, using continuous domain differentials and integrals. For practical implementations, the standard practice is to approximate the derivatives using finite differences. For example, the derivative of the 2-D signal along the x dimension is approximated as $q[k,l] = f[k+1,l] - f[k,l] = \triangle_1 * f$. This approximation can be viewed as the convolution of f by $\triangle_1[k] = \varphi\left(k+\frac{1}{2}\right)$, where $\varphi(x) = \partial\beta^1(x)/\partial x$ [7]. Here β^1 is a linear B-spline function. Note that this approximation does not possess rotation steerabilty; the first degree discrete directional derivatives cannot be obtained as the linear combination of the finite differences along x & y directions.

To obtain discrete operators that are approximately rotation steerable, we approximate the partial derivatives of the signal as the convolution of the signal with the partial derivatives of a window, which is the tensor product of Bspline functions:

$$q_{n_1,n_2}[k_1,k_2] = \underbrace{\left[\beta^d_{n_1}(k_1+\delta)\otimes\beta^d_{n_2}(k_2+\delta)\right]}_{\beta^d_{n_1,n_2}(k_1,k_2)} * f[k_1,k_2]. \tag{7}$$

Here $\beta^d_n(x)$ denotes the n^{th} derivative of a d^{th} degree B-spline. We choose $d = n_1 + n_2$ and

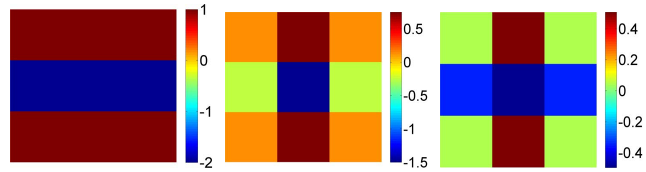
$$\delta = \begin{cases} \frac{1}{2} & \text{if} \quad d \text{ is odd} \\ 0 & \text{else} \end{cases} \tag{8}$$

Higher order tensor product B-spline functions approximate Gaussians and hence are approximately isotropic. Thus, the derivatives of B-spline functions are approximately rotation steerable. The shift δ is chosen to obtain short filters; the shift implies that we are evaluating the image derivatives at the intersection of the voxels and not at the voxels midpoints. While the proposed discrete operators are not steerable in the strict sense, this approximation is better than the current finite difference schemes.

3.2.1. Optimization of the derivative operator

We now optimize the orientation selectivity of the second degree operator $\mathcal{D}_{0,2}(\mathbf{r}) = \alpha_1\,\beta^2_{2,0}(\mathbf{r}) + \alpha_2\,\beta^2_{0,2}(\mathbf{r})$ to improve the penalty. Ideally, the steerable operator should behave as a derivative operator along the y direction, while it should be maximally elongated along x direction. The improved orientation selectivity of the detector will encourage the preservation of line-like features (edges/ridges) in the image. The ideal 3×3 discrete operator $\mathcal{D}_{\text{ideal}}(\mathbf{r})$ is shown in Fig.1.(a). Since this detector is not rotation steerable, we will determine the parameters of the operator $\mathcal{D}_{0,2}(\mathbf{r}) = \alpha_1\,\beta^2_{2,0}(\mathbf{r}) + \alpha_2\,\beta^2_{0,2}(\mathbf{r})$ to approximate $\mathcal{D}_{\text{ideal}}(\mathbf{r})$.

We determine the parameters of $\mathcal{D}_{0,2}$ such that $\|\mathcal{D}_{\text{ideal}}(\mathbf{r}) - \mathcal{D}_{\text{opt}}(\mathbf{r})\|^2$ is minimized. We normalize the coefficients α_1 and α_2 such that $\|\mathcal{D}_{\text{opt}}(\mathbf{r})\|^2 = 1$. As a result, the optimal operator



(a) $\mathcal{D}_{\text{ideal}}$ (b) B-spline (c) optimized B-spline

Fig. 1. Different discrete second order derivative operators (along the y direction.) (a) is the ideal finite difference approximation $\mathcal{D}_{\text{ideal}}$, which is maximally elongated along x. (b) is the second derivative of the quadratic tensor-product B-spline $\beta^2_{2,0}$. (c) is the optimized elongated B-spline operator $\mathcal{D}_{\text{opt}}$. Compared with ideal operator (a), the B-spline derivative operator (b) is more steerable. The optimized B-spline operator (c) inherits the property of approximate steerability of (b), while being more elongated compared to (b).

$\mathcal{D}_{\text{opt}}(\mathbf{r})$ can be determined as:

$$\{\alpha_1,\alpha_2\} = \arg\min_{\alpha_1,\alpha_2} \|\mathcal{D}_{\text{ideal}} - \alpha_1\beta^2_{2,0} - \alpha_2\beta^2_{0,2}\|^2 \quad \text{such that } \|\mathcal{D}_{\text{opt}}\|^2 = 1 \tag{9}$$

Using Lagrange multipliers, we reformulate this problem as:

$$\arg\min_{\mathbf{a}} \mathbf{V}^T\mathbf{a} + \lambda\mathbf{a}^T\mathbf{Q}\mathbf{a}, s.t. \|\mathbf{a}^T\mathbf{Q}\mathbf{a}\|^2 = 1 \tag{10}$$

where $\mathbf{a} = [\alpha_1,\alpha_2]^T$, $\mathbf{V} = [\langle\mathcal{D}_{\text{ideal}},\,\beta^2_{2,0}\rangle, \langle\mathcal{D}_{\text{ideal}},\,\beta^2_{0,2}\rangle]^T$, and

$$\mathbf{Q} = \begin{bmatrix} \|\beta^2_{2,0}\|^2 & \langle\beta^2_{2,0},\,\beta^2_{0,2}\rangle \\ \langle\beta^2_{0,2},\,\beta^2_{2,0}\rangle & \|\beta^2_{0,2}\|^2 \end{bmatrix}$$

We thus obtain the optimized second degree operator as $\mathcal{D}_{\text{opt}} = 0.5895\beta^2_{2,0} - 0.2586\beta^2_{0,2}$. Thus, the corresponding directional derivative is specified by

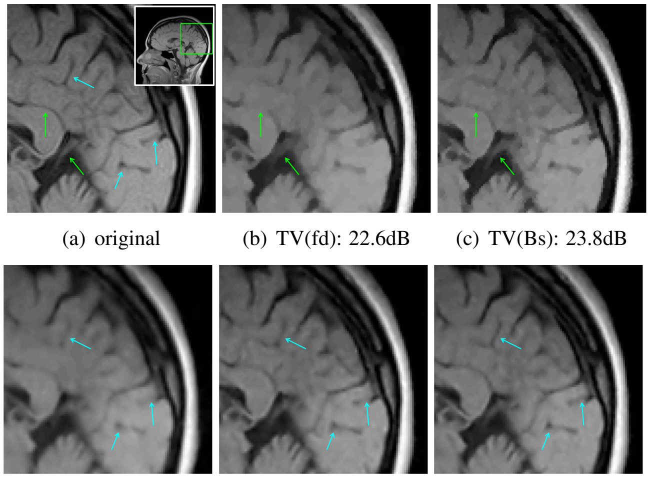
$$\begin{aligned} f_{\theta,n} &= f_{xx}(\alpha_1\cos^2\theta + \alpha_2\sin^2\theta) + f_{yy}(\alpha_1\sin^2\theta + \alpha_2\cos^2\theta) \\ &+ 2(\alpha_1-\alpha_2)f_{xy}\cos\theta\sin\theta \end{aligned} \tag{11}$$

Substituting (11) into (2), we obtain the corresponding optimized HDTV penalties. The ideal operator $\mathcal{D}_{\text{ideal}}$, standard B-spline derivative $\beta^2_{2,0}$, and optimized detectors $\mathcal{D}_{\text{opt}}$ are shown in Fig. 1. Note that the optimized operator is more elongated than the standard B-spline discrete differential operator.

4. RESULTS

In order to determine the utility of the improved TV and HDTV penalties, we apply the proposed schemes in the context of compressed sensing recovery of MRI data. We retrospectively downsample fully sampled MRI datasets using a variable density random sampling pattern such that the acceleration is approximately $A = 4.35$. We choose the regularization parameter λ such that $\|\mathcal{A}(\hat{f}) - b\|^2 \approx \sigma^2$. We compute the signal to error ratio (SER) of the reconstructions as

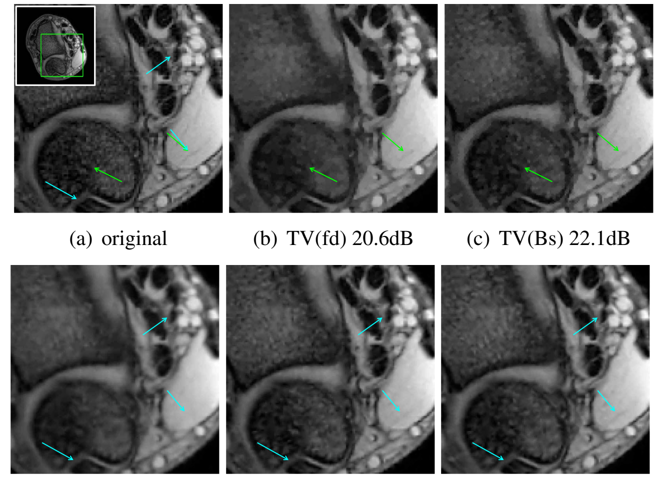
$$\text{SER} = -10\log_{10}\left(\frac{\|f_{\text{orig}} - \hat{f}\|_F^2}{\|f_{\text{orig}}\|_F^2}\right), \tag{12}$$



(a) original (b) TV(fd): 22.6dB (c) TV(Bs): 23.8dB

(d) HDTV(fd): 19.9dB (e) HDTV(Bs) 24.3dB (f) HDTV(op) 24.6dB

Fig. 2. Comparison of different derivative approximation operators. (a) is the actual brain MR image. (b) and (c) show the TV reconstructions using finite difference operator (fd) and B-spline operator (Bs), respectively. (d), (e), and (f) show the HDTV reconstructions using finite difference, B-spline, and the optimized B-spline operator (op), respectively.



(a) original (b) TV(fd) 20.6dB (c) TV(Bs) 22.1dB

(d) HDTV(fd) 19.0dB (e) HDTV(Bs) 22.3dB (f) HDTV(op) 22.5dB

Fig. 3. Comparison of different derivative approximation operators. The original wrist MR image is shown in (a). (b) and (c) illustrate the TV reconstructions using finite differences (fd) operator and B-spline (Bs) operator, respectively. (d) to (f) show the HDTV reconstructions using finite differences operator, B-spline operator and the optimized B-spline (op) derivative operator, respectively.

where $\hat{f}$ is the reconstructed image; f_{orig} is the original image; $\|\cdot\|_F$ is the Frobenius norm.

We compare the TV & HDTV based schemes in Fig. 2 using a brain MR image. We show the reconstructions of the image using standard TV penalty with finite difference approximation and B-spline operator, while we compare the finite difference operator, the B-spline operator, and the proposed optimized B-spline operator, in the context of HDTV penalty. The original image is shown in (a). We observe that for TV penalty, the B-spline operator approximation (c) provides better ridge/edge preservation, compared to the finite difference approximation (b), indicated in green arrows. Similarly, for HDTV penalty, the finite difference operator based reconstruction (d) is more blurred than the B-spline operator reconstruction (e). In contrast, the optimized B-spline operator preserves the ridges better than standard B-spline operator (see blue arrows). Fig. 3 compares the reconstructions of a MR wrist image using TV & HDTV regularization with different derivatives approximation operators. We observe that using B-spline operators, the SER is improved by around 2-3dB than finite difference operator. The optimized B-spline operator captures the subtle details more effectively and improves the SER by 0.2dB over the standard B-spline operator.

5. CONCLUSION

We generalized the HDTV penalties by considering rotated versions of an arbitrary derivative operator instead of directional derivative operators. We show that many of the current higher degree penalties are special cases of the generalized isotropic HDTV penalties. Our experiments show that the anisotropic counterparts of these schemes provide improved reconstructions. We also optimize the derivative operator to improve its orientation selectivity, thus further improving the performance of the resulting HDTV penalty. We also consider efficient discretization of the penalties, which are specified in the continuous domain. Our numerical experiments show a significant improvement in performance, offered by the improved discretization and optimization of the penalty.

6. REFERENCES

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

[2] Y. Hu and M. Jacob, "Higher degree total variation (HDTV) regularization for image recovery," *IEEE Transactions on Image Processing, accepted*.

[3] M. Lysaker, A. Lundervold, and X.-C. Tai, "Noise removal using fourth-order partial differential equation with applications to medical magnetic resonance images in space and time," *IEEE Transactions on Image Processing*, vol. 12, no. 12, pp. 1579–1590, Dec 2003.

[4] T. Chan, A. Marqina, and P. Mulet, "Higher-order total variation-based image restoration," *SIAM J. Sci. Computing*, vol. 22, no. 2, pp. 503–516, Jul 2000.

[5] Y. Hu and M. Jacob, "Image recovery using improved total variation regularization," *IEEE International Symposium on Biomedical Imaging*, pp. 1154–1157, 2011.

[6] J. Kybic, T.Blu, and M. Unser, "Generalized sampling: a variational approach, part i," *IEEE Trans. Signal Processing*, vol. 50, pp. 1965–1976, 2002.

[7] M. Unser and T. Blu, "Wavelet theory demystified," *IEEE Transactions on Signal Processing*, vol. 51, no. 2, pp. 470–483, Feb 2003.