TGV regularization for variational approaches to quantitative susceptibility mapping

Kristian Bredies, Florian Knoll and Christian Langkammer

***Abstract*—Algorithms for solving variational problems with total generalized variation (TGV) regularization are presented and discussed. The TGV functional shares convenient properties of the well-known total-variation (TV) functional, but is also able to deal with higher-order smoothness and therefore leads to more natural solutions. We present a minimization algorithm for the solution of general linear inverse problems with TGV-regularization with special attention on deconvolution. Finally, the application to quantitative susceptibility mapping is discussed. In particular, a variational model for recovering the susceptibility distribution from wrapped phase images is proposed.**

***Keywords*—Total generalized variation, variational modelling, minimization algorithms, deconvolution, quantitative susceptibility mapping.**

I. Introduction

VARIATIONAL methods have shown to be very successful for solving mathematical imaging problems. The major reason might be the great flexibility variational modelling has to offer. Last but not least, the availability of efficient minimization algorithms for the solution of the underlying optimization problems also plays an important role.

The basic idea of variational methods is to model the imaging problem which has to be solved as the minimization of the sum

$$\min_u \; F(u) + \Phi(u)$$

where F and Φ are data fidelity and regularization functionals, respectively. For instance, if it is known that some given data f can be inferred from the unknown sought quantity u^* by a linear map K, i.e., the aim is to solve $Ku = f$ a typical approach would to choose the data fidelity as the discrepancy

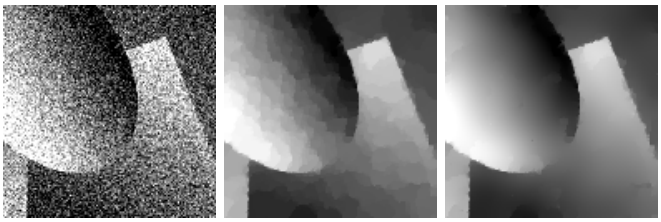
$$F(u) = \frac{\|Ku - f\|^2}{2}$$

for general u and let the minimization procedure determine the correct u^*. However, the inversion of K might be ill-posed: On the one hand, the data f might be incomplete and therefore insufficient to uniquely recover u^*. On the other hand, even in case of complete data, a solution may not depend stably on the data.

K. Bredies is with the Institute of Mathematics and Scientific Computing, University of Graz, Heinrichstrasse 36, A-8010 Graz, Austria. E-mail: kristian.bredies@uni-graz.at.

F. Knoll is with the Department of Radiology, New York University Medical Center, 660 First Avenue, New York, NY 10016. E-mail: Florian.Knoll@nyumc.org.

C. Langkammer is with the Department of Neurology, Medical University of Graz, Auenbruggerplatz 22, A-8036 Graz, Austria. E-mail christian.langkammer@medunigraz.at.



noisy image TV denoising TGV denoising

Fig. 1. The effect of TGV compared to TV on variational denoising.

In both cases, regularization by adding a functional Φ is necessary. A good choice for Φ should reflect appropriate model assumptions on u^*. Here, the total variation is enjoying much attention [1], i.e., the choice

$$\Phi(u) = \alpha \,\mathrm{TV}(u) = \alpha \int_\Omega |\nabla u|$$

for some $\alpha > 0$. The total variation is an appropriate model for piecewise constant functions, a property which is often assumed for images. It can in particular be used for "compressed sensing" applications in which the data is highly incomplete [2]. There, TV-regularization has the effect that the missing data is often almost perfectly recovered.

However, TV has some drawbacks: As it only takes the first derivative ∇u into account, it is unaware of higher-order smoothness. For images which are not piecewise constant, this leads to blocky solutions exhibiting so-called "staircase artifacts". One way to avoid these artifacts is to regularize with the total generalized variation which does take higher-order derivative information of u into account [3]:

$$\mathrm{TGV}^2_\alpha(u) = \min_w \; \alpha_1 \int_\Omega |\nabla u - w| + \alpha_0 \int_\Omega |\mathcal{E} w|$$

where $\alpha_0, \alpha_1 > 0$, w is a vector field on Ω for which the symmetrized derivative $\mathcal{E}w$ is a symmetric matrix with entries $(\mathcal{E}w)_{ij} = \frac{1}{2}(\frac{\partial w_i}{\partial x_j} + \frac{\partial w_j}{\partial x_i})$. It generalizes TV in the sense that it is a good model for piecewise smooth images, i.e., there is no tendency towards "staircase" images. In particular, TGV can be employed for "compressed sensing" applications [4].

In the following, we focus on the algorithmic aspects of TGV-regularization for linear inverse problems.

II. Linear inverse problems with TGV-regularization

Assume that K is a linear operator for which we like to solve the linear inverse problem

$$Ku = f \tag{1}$$

for given data f by means of Tikhonov regularization with second-order total generalized variation, i.e.,

$$\min_u \ \frac{\|Ku - f\|^2}{2} + \mathrm{TGV}^2_\alpha(u). \tag{2}$$

Generally, this minimization problem can be solved by any suitable optimization method. We will present one particular in the following. It bases on the saddle-point formulation

$$\min_{u,w} \max_{v,p,q} \ \langle Ku - f,\, v\rangle + \langle \nabla u - w,\, p\rangle + \langle \mathcal{E}w,\, q\rangle - \frac{\|v\|^2}{2} - I_{\{\|\cdot\|_\infty \le \alpha_1\}}(p) - I_{\{\|\cdot\|_\infty \le \alpha_0\}}(q) \tag{3}$$

where I_M denotes the indicator functional with respect to M, i.e. $I_M(z) = 0$ if $z \in M$ and ∞ otherwise. This problem is equivalent to (2) in the sense that primal-dual solution pairs (u^*, w^*) and (v^*, p^*, q^*) give a u^* which is solving (2).

With $x = (u, w)$ and $y = (v, p, q)$ the problem (3) fits into the framework

$$\min_x \max_y \ \langle \mathcal{K}x,\, y\rangle + \mathcal{G}(x) - \mathcal{F}(y)$$

where $\mathcal{F}$ and $\mathcal{G}$ are convex functionals. A solution can approximately be found by iterating as follows [5]: Choose $x^0, \bar{x}^0 = 0$, $y^0 = 0$, $\sigma, \tau > 0$ such that $\sigma\tau < 1/\|\mathcal{K}\|^2$ and compute, for $k \ge 0$,

$$\begin{cases} y^{k+1} = (I + \sigma\partial\mathcal{F})^{-1}(y^k + \sigma\mathcal{K}\bar{x}^k) \\ x^{k+1} = (I + \tau\partial\mathcal{G})^{-1}(x^k - \tau\mathcal{K}^* y^{k+1}) \\ \bar{x}^{k+1} = 2x^{k+1} - x^k. \end{cases} \tag{4}$$

This requires that the resolvents

$$(I + \sigma\partial\mathcal{F})^{-1}(y') = \arg\min_y \ \frac{\|y - y'\|^2}{2} + \sigma\mathcal{F}(y), \tag{5}$$

$$(I + \tau\partial\mathcal{G})^{-1}(x') = \arg\min_x \ \frac{\|x - x'\|^2}{2} + \tau\mathcal{G}(x) \tag{6}$$

as well as the application of $\mathcal{K}$ and $\mathcal{K}^*$ can be computed efficiently. Moreover, a good estimate on

$$\|\mathcal{K}\|^2 = \sup_{\|x\|\le 1} \|\mathcal{K}x\|^2$$

is required.

In case of (3), the algorithm (5) is realized by easy computations [6]. Assuming that the application of K and its adjoint K^* can be computed and there is an estimate $\|K\|_{\mathrm{est}} \ge \|K\|$ available, it can be written as a convergent update scheme according to Algorithm 1.

The method can in particular be used for images on a uniform grid in an arbitrary dimension d. The variable u is then a scalar-valued image, while w, p are $\mathbf{R}^d$-valued images and q takes values in $S^{d\times d}$, the space of symmetric $d \times d$ matrices. A usual realization for the gradient operator ∇ is given by forward differences with homogeneous Neumann boundary conditions, for instance, in two dimensions as

$$(\partial_1^+ u)_{i,j} = \frac{u_{i+1,j} - u_{i,j}}{h_1}, \quad (\partial_2^+ u)_{i,j} = \frac{u_{i,j+1} - u_{i,j}}{h_2} \tag{7}$$

Algorithm 1 (TGV-Tikhonov minimization):

function TGVTIKHONOV

Input: f data
(α_0, α_1) regularization parameters
eval$K$ forward operator
evalK^* adjoint operator
$\|K\|_{\mathrm{est}}$ norm estimate

Output: u approximate solution

$u, \bar{u} \leftarrow 0$, $w, \bar{w} \leftarrow 0$
$v \leftarrow 0$, $p \leftarrow 0$, $q \leftarrow 0$
Choose $\sigma, \tau > 0$ such that

$$\sigma\tau \frac{\|K\|^2_{\mathrm{est}} + 2\|\nabla\|^2 + \sqrt{(\|K\|^2_{\mathrm{est}} - 1)^2 + 4\|\nabla\|^2} + 1}{2} \le 1$$

repeat
$K\bar{u} \leftarrow \mathrm{eval}K(\bar{u})$
$v \leftarrow \dfrac{v + \sigma(K\bar{u} - f)}{1 + \sigma}$
$p \leftarrow P_{\alpha_1}\big(p + \sigma(\nabla\bar{u} - \bar{w})\big)$
$q \leftarrow P_{\alpha_0}(q + \sigma\mathcal{E}\bar{w})$
$u_{\mathrm{old}} \leftarrow u$
$w_{\mathrm{old}} \leftarrow w$
$K^*v \leftarrow \mathrm{eval}K^*(v)$
$u \leftarrow u + \tau(\mathrm{div}_1\, p - K^*v)$
$w \leftarrow w + \tau(p + \mathrm{div}_2\, q)$
$\bar{u} \leftarrow 2u - u_{\mathrm{old}}$
$\bar{w} \leftarrow 2w - w_{\mathrm{old}}$
until convergence of u
return u

TABLE I. ALGORITHM FOR THE SOLUTION OF TGV-REGULARIZED LINEAR INVERSE PROBLEMS.

for $1 \le i \le N_1$, $1 \le j \le N_2$ if the image has size $N_1 \times N_2$ and grid sizes $h_1, h_2 > 0$. Here, $u_{N_1+1,j} = u_{N_1,j}$ as well as $u_{i,N_2+1} = u_{i,N_2}$. This gives

$$\nabla u = \begin{pmatrix} \partial_1^+ u \\ \partial_2^+ u \end{pmatrix}, \qquad \|\nabla\|^2 < 4\Big(\frac{1}{h_1^2} + \frac{1}{h_2^2}\Big). \tag{8}$$

An adaptation to higher dimensions is straightforward. Once a gradient operator has been defined, $\mathcal{E}$ is immediately given by

$$(\mathcal{E}w) = \begin{pmatrix} \partial_1^+ w^1 & \frac{1}{2}(\partial_2^+ w^1 + \partial_1^+ w^2) \\ \frac{1}{2}(\partial_2^+ w^1 + \partial_1^+ w^2) & \partial_2^+ w^2 \end{pmatrix}. \tag{9}$$

The divergence operators div_1 and div_2 are then the negative adjoint of ∇ and $\mathcal{E}$, respectively. The choices (7)–(9) lead to

$$(\partial_1^- u)_{i,j} = \frac{u_{i,j} - u_{i-1,j}}{h_1}, \quad (\partial_2^- u)_{i,j} = \frac{u_{i,j} - u_{i,j-1}}{h_2} \tag{10}$$

for $1 \le i \le N_1$, $1 \le j \le N_2$ where $u_{0,j} = u_{N_1,j} = u_{i,0} = u_{i,N_2} = 0$. Consequently,

$$\mathrm{div}_1\, p = \partial_1^- p^1 + \partial_2^- p^2, \quad \mathrm{div}_2\, q = \begin{pmatrix} \partial_1^- q^{11} + \partial_2^- q^{12} \\ \partial_1^- q^{12} + \partial_2^- q^{22} \end{pmatrix}. \tag{11}$$

Finally, the projection operators P_{α_1} and P_{α_0} read in this setting as the pointwise operations

$$P_{\alpha_1}(p) = \frac{p}{\max\bigl(1, \sqrt{|p^1|^2 + |p^2|^2}/\alpha_1\bigr)}, \tag{12}$$

$$P_{\alpha_0}(q) = \frac{q}{\max\bigl(1, \sqrt{|q^{11}|^2 + |q^{22}|^2 + 2|q^{12}|^2}/\alpha_0\bigr)}. \tag{13}$$

This also generalizes to arbitrary dimensions in a straightforward manner. Observe, however, the coincidence of entries for symmetric matrices.

Finally, note that from the operator K, all what is needed are evaluation functions eval$K$ and evalK^* which take some data as argument and return K or K^* applied to this data, respectively. This allows to use Fast Fourier Transform operators, for instance.

III. Applications in Magnetic Resonance Imaging

A. Undersampling reconstruction

Reconstructing an MR image from an incomplete subset Σ of arbitrarily sampled k-space data from n coils amounts to solving the underdetermined equation

$$\mathcal{F}(\sigma_i u)|_\Sigma = f_i, \qquad i = 1, \dots, n \tag{14}$$

where f_i is the k-space data on Σ and σ_i the sensitivity from the ith coil, respectively. Choosing

$$Ku = \bigl(\mathcal{F}(\sigma_1 u)|_\Sigma, \dots, \mathcal{F}(\sigma_n u)|_\Sigma\bigr) \tag{15}$$

and regularizing with TGV, i.e., solving (2) then leads to a "compressed sensing" approach which is aware of the piecewise smooth structure of the data [4].

In order to employ Algorithm 1 in an efficient way, a Non-Uniform Fast Fourier Transform NUFFT$_\Sigma$ and its adjoint NUFFT$_\Sigma^*$ has to be available, for instance from [7]. Based on this, the evaluation operators eval$K$ and evalK^* may be implemented according to Algorithm 2. A estimate for $\|K\|$ is then given by

$$\|K\|_{\text{est}} = \bigl(\max_{i=1,\dots,n} \|\sigma_i\|_\infty\bigr)\|\text{NUFFT}_\Sigma\|.$$

The results of this method to sample data and a comparison of TGV to TV and unregularized solutions is depicted in Figure 2.

B. Deconvolution of phase images

Deconvolution is the problem of solving

$$u * k = f \tag{16}$$

with given k and f. Choosing $Ku = u * k$, it has the form (1) and hence, Algorithm 1 can be utilized for TGV-regularized deconvolution according to (2) [6]. In particular, if the support of k is large, it might be beneficial to employ the Fast Fourier Transform for fast evaluation. The corresponding functions eval$K$ and evalK^* can then by implemented according to Algorithm 3. With the exact estimate

$$\|K\|_{\text{est}} = \|K\| = \|\mathcal{F}k\|_\infty,$$

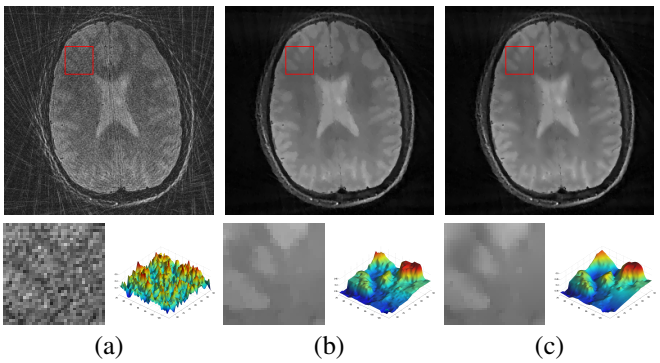
the ingredients for performing Algorithm 1 are complete.

Algorithm 2 (Fast MR undersampling operators):

function EVALK
Input: u data, σ_i sensitivities, Σ sampling
Output: Ku according to (15)
for $i = 1 \to n$ **do**
$(Ku)_i \leftarrow \text{NUFFT}_\Sigma(\sigma_i u)$
return $\bigl((Ku)_1, \dots, (Ku)_n\bigr)$

function EVALK^*
Input: $(v_1, \dots, v_n)$ data, σ_i sensitivities, Σ sampling
Output: K^*v with K^* the adjoint of K in (15)
$u \leftarrow 0$
for $i = 1 \to n$ **do**
$u \leftarrow u + \overline{\sigma_i}\text{NUFFT}_\Sigma^*(v_i)$
return u

TABLE II. Pseudocode for the fast evaluation of the undersampling operator and its adjoint.



(a) (b) (c)

Fig. 2. MR radial undersampling reconstruction example. (a) Direct inversion via NUFFT, (b) TV "compressed sensing" regularization, (c) TGV regularization. The images are reconstructed from radial k-space data (24 spokes, 12 coils), see [4] for more details.

The method has been tested with the convolution kernel δ appearing in quantitative susceptibility mapping [8] which relates, up to a multiplicative factor, the susceptibility distribution and the measured phase field. Its Fourier transform reads as

$$(\mathcal{F}\delta)(k_x, k_y, k_z) = \frac{\frac{1}{3}(k_x^2 + k_y^2) - \frac{2}{3}k_z^2}{k_x^2 + k_y^2 + k_z^2}. \tag{17}$$

The results for a synthetically generated phantom data set disturbed by noise can be found in Figure 3.

IV. An integrative approach for quantitative susceptibility mapping

We like to reconstruct the susceptibility distribution χ directly from single-echo wrapped phase image ϕ_0^{wrap} by solving a suitable optimization problem. For that purpose, observe [9] that for the unwrapped phase image ϕ_0, we have $\mathrm{e}^{\mathrm{i}\phi_0} = \mathrm{e}^{\mathrm{i}\phi_0^{\text{wrap}}}$, hence

$$\Delta\phi_0 = \text{Imag}\bigl((\Delta \mathrm{e}^{\mathrm{i}\phi_0^{\text{wrap}}})\mathrm{e}^{-\mathrm{i}\phi_0^{\text{wrap}}}\bigr).$$

Algorithm 3 (FFT-based convolution):

function EVALK
Input: u data, $\mathcal{F}k$ Fourier transformed kernel
Output: $u * k$
return IFFT$\big($FFT$(u) \cdot \mathcal{F}k\big)$

function EVALK^*
Input: u data, $\mathcal{F}k$ Fourier transformed kernel
Output: $u * k(- \cdot)$
return IFFT$\big($FFT$(u) \cdot \overline{\mathcal{F}k}\big)$

TABLE III. PSEUDOCODE FOR THE FAST EVALUATION OF CONVOLUTION AND ADJOINT CONVOLUTION OPERATORS.

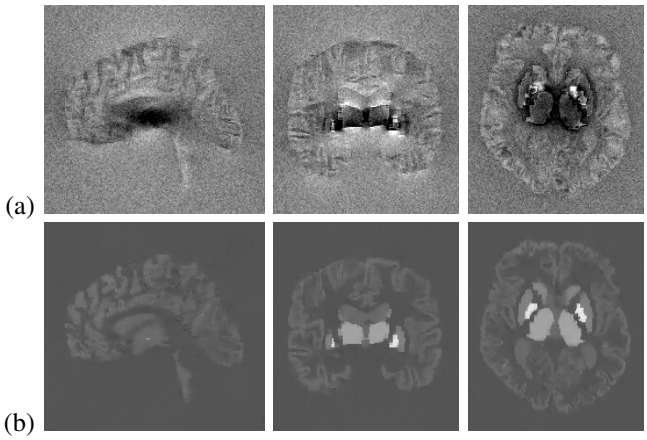


Fig. 3. TGV-regularized deconvolution of noisy data. (a) Smoothly modulated and convolved phantom image with heavy noise (kernel according to (17), scale from -0.02 to 0.02ppm). (b) Result of Algorithm 1 (scale from -0.07 to 0.14ppm).

Up to a multiplicative constant, $\Delta\phi_0$ is linked to the susceptibility distribution χ by

$$\Box\chi = \frac{1}{3}\Big(\frac{\partial^2\chi}{\partial x^2} + \frac{\partial^2\chi}{\partial y^2}\Big) - \frac{2}{3}\frac{\partial^2\chi}{\partial z^2} = \Delta\phi_0,$$

see, for instance, [8]. Since χ is supported on the mask Ω', one can uniquely recover it from the knowledge of $\Box\chi$.

We like to solve $\Box\chi = \Delta\phi_0$ in Ω' by a variational method. As the second derivatives of χ and ϕ_0 are distributions, we measure, however, the discrepancy with respect to a negative norm. This can be realized by taking $\|\psi\|^2/2$ on Ω' where ψ satisfies $\Delta\psi = \Box\chi - \Delta\phi_0$ in Ω'. Regularized with TGV, the functional to minimize then reads as

$$\begin{cases} \min\limits_{\chi,\psi} \ \dfrac{1}{2}\displaystyle\int_{\Omega'} |\psi|^2 \ \mathrm{d}x + \mathrm{TGV}_\alpha^2(\chi) \\ \text{subject to} \quad \Delta\psi = \Box\chi - \Delta\phi_0 \quad \text{in} \quad \Omega'. \end{cases} \tag{18}$$

In order to solve this constrained problem, Algorithm 1 has to be modified to incorporate the equality constraints. This can be done by introducing a Lagrange multiplier η in (3) deriving an update rule for η from (4).

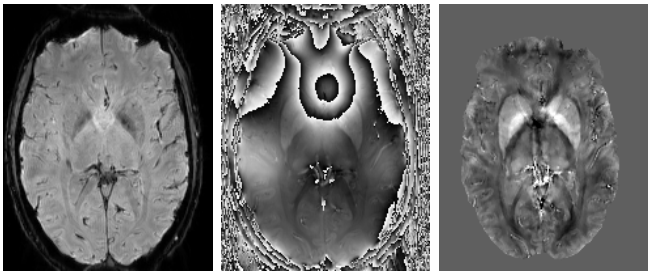


Fig. 4. TGV-regularized susceptibility reconstruction from wrapped phase data. (a) Magnitude image (for brain mask extraction). (b) Input phase image ϕ_0^{wrap} (single gradient echo, 27ms, 3T). (c) Result of the integrative approach (18) (scale from -0.15 to 0.25ppm).

The variational approach (18) and the associated minimization method has been tested on raw phase data, see Figure 4.

V. CONCLUSION

Variational methods offer a great flexibility for the solution of mathematical image processing problems. The total generalized variation as regularization functional leads to results of high visual quality and can also be used for compressed sensing applications. Simple and efficient numerical algorithms can be found for the solution of the associated minimization problems. These methods are suitable for problems in MRI, for instance, for undersampling reconstruction and QSM.

REFERENCES

[1] L. I. 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, 1992.

[2] M. Lustig, D. Donoho, and J. M. Pauly, "Sparse MRI: The application of compressed sensing for rapid MR imaging," *Magnetic Resonance in Medicine*, vol. 58, no. 6, pp. 1182–1195, 2007.

[3] K. Bredies, K. Kunisch, and T. Pock, "Total generalized variation," *SIAM Journal on Imaging Sciences*, vol. 3, no. 3, pp. 492–526, 2010.

[4] F. Knoll, K. Bredies, T. Pock, and R. Stollberger, "Second order total generalized variation (TGV) for MRI," *Magnetic Resonance in Medicine*, vol. 65, no. 2, pp. 480–491, 2011.

[5] A. Chambolle and T. Pock, "A first-order primal-dual algorithm for convex problems with applications to imaging," *Journal of Mathematical Imaging and Vision*, vol. 40, no. 1, pp. 120–145, 2011.

[6] K. Bredies, "Recovering piecewise smooth multichannel images by minimization of convex functionals with total generalized variation penalty," University of Graz, Tech. Rep., 2012.

[7] J. Fessler, "Image reconstruction toolbox," [Online] http://web.eecs.umich.edu/~fessler/code/.

[8] L. de Rochefort, T. Liu, B. Kressler, J. Liu, P. Spincemaille, V. Lebon, J. Wu, and Y. Wang, "Quantitative susceptibility map reconstruction from MR phase data using Bayesian regularization: Validation and application to brain imaging," *Magnetic Resonance in Medicine*, vol. 63, no. 1, pp. 194–206, 2010.

[9] M. A. Schofield and Y. Zhu, "Fast phase unwrapping algorithm for interferometric applications," *Opt. Lett.*, vol. 28, no. 14, pp. 1194–1196, 2003.