Published in final edited form as:

Phys Med Biol. ; 62(18): N428–N435. doi:10.1088/1361-6560/aa837d.

A sequential solution for anisotropic total variation image denoising with interval constraints

Jingyan Xu¹ and Frédéric Noo²

¹Department of Radiology, Johns Hopkins University

²Department of Radiology and Imaging Sciences, University of Utah

Abstract

We show that two problems involving the anisotropic total variation (TV) and interval constraints on the unknown variables admit, under some conditions, a simple sequential solution. Problem 1 is a constrained TV penalized image denoising problem; Problem 2 is a constrained fused lasso signal approximator. The sequential solution entails finding first the solution to the unconstrained problem, and then applying a thresholding to satisfy the constraints. If the interval constraints are uniform, this sequential solution solves problem 1. If the interval constraints furthermore contain zero, the sequential solution solves problem 2. Here uniform interval constraints refer to all unknowns being constrained to the same interval. A typical example of application is image denoising in x-ray CT, where the image intensities are non-negative as they physically represent linear attenuation coefficient in the patient body. Our results are simple yet seem unknown; we establish them using the Karush-Kuhn-Tucker conditions for constrained convex optimization.

Keywords

denoising; smoothing; total variation; lasso; fused lasso; interval constraints; box constraints; constrained optimization

1. Introduction

Total variation (TV) is widely used in signal processing and in image denoising, restoration, and reconstruction, as an edge-preserving penalty to cope with ill-posedness in the problem. The TV penalty, both the anisotropic model and the isotropic model, is convex but nonsmooth, thus causes difficulty for numerical optimization. Some numerical algorithms replace the TV by a smoothed version of TV so that conventional gradient-descent type algorithms can be applied. These algorithms introduce a nuisance parameter $\varepsilon > 0$ to the TV penalty; the value of ε can be used to trade off computational accuracy and algorithm efficiency [1]. Other methods, such as the primal-dual type [2], the operator splitting type [3], and numerous others [4–6], have been developed to deal with the nonsmooth TV directly without introducing any nuisance parameters.

In many imaging applications, the image intensity values are known to be non-negative. For example, in x-ray CT, the pixel values can only be non-negative as they represent the x-ray linear attenuation coefficient of human tissues. In nuclear medicine, the pixel values represent radioactivity uptakes, which are also non-negative. Incorporating this prior knowledge not only produces more interpretable results [7–9], it also improves image quality; see *e.g.*, [10–14]. In SPECT and CT, it is also often possible to identify an upper bound on the image intensity values. Such interval constraints have been shown to reduce reconstruction time and generate more accurate attenuation maps in the case of truncated SPECT transmission data [15, 16].

TV problems with interval constraints have also received attention from a numerical optimization perspective, see, *e.g.*, [10, 17–19]. The existing algorithms for the constrained problems often consider a generic image restoration problem with either the anisotropic or the isotropic TV model. In many practical implementations, the interval constraints are simply handled in the following ad-hoc manner. First, the solution to the unconstrained TV problem is obtained. Second, the solution from the first step is thresholded to satisfy the constraints [13,20]. In this work, we show that this seemingly naive solution is indeed accurate for a subclass of constrained TV problems, namely, when the problem is a denoising problem using the anisotropic TV and when the interval constraints are uniform, *i.e.*, all unknowns have the same interval constraint $[L, u]$, $L < u$. Note that this subclass includes TV denoising with the non-negativity constraint as a special case, *i.e.*, $L = 0$ and $u = \infty$.

The rest of the paper is organized as follows. In Sec. 2 we define Problem 1 and establish that the two-step solution mentioned above is exact. Note that this exactness is true for nonsmooth TV minimization, *i.e.*, without introducing any nuisance parameter ϵ . Counter examples with smooth TV or with nonuniform interval constraints are given in Appendix. The analysis in Sec 2 is more detailed and serves as a model for Sec. 3, which discusses a generalization of Problem 1, a constrained fused lasso signal approximator. The constrained fused lasso also admits a sequential solution when the interval constraint is uniform and contains zero. We discuss some related development in the literature in Sec 4 and conclude in Sec. 5.

2. TV-penalized denoising with uniform interval constraints

Consider the following problem

$$\min_{\{x_i\}} \left\{ \Theta_1(\gamma) = \sum_{i=1}^N \frac{w_i}{2} (x_i - y_i)^2 + \gamma \sum_i \iota_{[L,u]}(x_i) + \sum_i \sum_{j \in N_i} \beta_{ij} |x_i - x_j| \right\} \quad (1)$$

where $x_i, y_i, i = 1, \dots, N$ are the unknowns[‡] and the noisy input data, respectively, and where N_i is a fixed set of neighbors for each i . The weight $w_i > 0$ is often chosen to be inversely proportional to the variance of y_i and $\beta_{ij} > 0$ is the penalty weight.

Note that we use ι_C to denote the indicator function of the convex set C , i.e., $\iota_C(x) = 0$ if $x \in C$, and $\iota_C(x) = \infty$ when $x \notin C$. In (1), the parameter γ of the indicator function is either 0 or 1, so that the objective function is unconstrained $\Theta_1(0)$ or constrained $\Theta_1(1)$, respectively. When $\gamma = 1$, the unknowns in (1) are all uniformly constrained to the interval $[l, u]$. We assume $l < u$ and we allow $l = -\infty$ or $u = \infty$, which means the constraint can be one-sided. Problem (1) is a general smoothing or denoising problem.

Since the objective function (1) is strongly convex, there is a unique global minimum to $\Theta_1(\gamma)$ for $\gamma = \{0, 1\}$. To simplify notation, we write $z \equiv \{z_j\}$, where z is a generic vector variable. The solution x^* to the constrained problem $\Theta_1(1)$ can be obtained from the Karush-Kuhn-Tucker (KKT) conditions for convex nonsmooth problems [21, Theorem 3.34]: if the triple $x^*, \lambda^{1*}, \lambda^{2*}$ satisfies the following conditions

$$w_i(x_i^* - y_i) + \sum_{j \in N_i} \beta_{ij} t_{ij}(x^*) - \sum_{j: i \in N_j} \beta_{ji} t_{ji}(x^*) - \lambda_i^{1*} + \lambda_i^{2*} = 0, \forall i \quad (2a)$$

$$x_i^* \leq u, \quad \lambda_i^{2*} \geq 0, \quad \lambda_i^{2*}(x_i^* - u) = 0, \forall i \quad (2b)$$

$$x_i^* \geq l, \quad \lambda_i^{1*} \geq 0, \quad \lambda_i^{1*}(x_i^* - l) = 0, \forall i \quad (2c)$$

where $t_{mn}(x)$ is the subdifferential of $|x_m - x_n|$ at x , then x^* is the minimizer of $\Theta_1(1)$. By definition, $t_{mn}(x) = \text{sign}(x_m - x_n)$ if $x_m \neq x_n$, and $t_{mn}(x) = [-1, 1]$ if $x_m = x_n$. The auxiliary variables $\lambda_i^{1*}, \lambda_i^{2*}, i = 1, \dots, N$, are the (optimal) Lagrange multipliers introduced for the constraints.

Denote by $\tilde{x}^*$ the solution to minimizing $\Theta_1(0)$, i.e., the difference between $\tilde{x}^*$ and x^* is that x^* satisfies the constraints $l \leq x_i^* \leq u, \forall i$. We establish that $x_i^* = \text{median}\{\tilde{x}_i^*, l, u\}$, which expands to

$$x_i^* = \begin{cases} l & \tilde{x}_i^* \leq l \\ \tilde{x}_i^* & l < \tilde{x}_i^* < u \\ u & \tilde{x}_i^* \geq u \end{cases} \quad (3)$$

In other words, x^* can be obtained in a two-step manner. First we solve (1) without considering the constraints; second, we perform pixelwise thresholding.

The solution $\tilde{x}^*$ from the first step satisfies the first order optimality condition, i.e.,

[‡]Note that x_j is not necessarily a 1-D signal; it can be a multi-dimensional signal, represented using lexicographical ordering of the entries.

$$w_i(\tilde{x}_i^* - y_i) + \sum_{j \in N_i} \beta_{ij} t_{ij}(\tilde{x}^*) - \sum_{j: i \in N_j} \beta_{ji} t_{ji}(\tilde{x}^*) \ni 0, \forall i \quad (4)$$

and the second (thresholding) step in (3) is the solution to the following constrained minimization problem

$$\min_{\hat{x}_i \in [l, u]} \frac{1}{2} \sum_i (\tilde{x}_i^* - \hat{x}_i)^2. \quad (5)$$

By applying the KKT conditions to (5), we find that the solution $\hat{x}$ satisfies

$$(\hat{x}_i - \tilde{x}_i^*) - \hat{\lambda}_i^1 + \hat{\lambda}_i^2 = 0, \forall i \quad (6a)$$

$$\hat{x}_i \leq u, \quad \hat{\lambda}_i^2 \geq 0, \quad (\hat{x}_i - u) \hat{\lambda}_i^2 = 0, \forall i \quad (6b)$$

$$\hat{x}_i \geq l, \quad \hat{\lambda}_i^1 \geq 0, \quad (\hat{x}_i - l) \hat{\lambda}_i^1 = 0, \forall i \quad (6c)$$

for some $\{\hat{\lambda}_i^1\}, \{\hat{\lambda}_i^2\}$. We show that $\{\hat{x}_i\}, \{w_i \hat{\lambda}_i^1\}, \{w_i \hat{\lambda}_i^2\}$ is a triple that satisfies the KKT conditions (2). Hence, $\hat{x} = x^*$, the unique minimizer of $\Theta_1(1)$.

Given the relation in (6) between $\tilde{x}^*$ and $\hat{x}$, we must have

$$t_{ij}(\tilde{x}^*) \subseteq t_{ij}(\hat{x}) \quad (7)$$

This is because the thresholding from $\tilde{x}^*$ to $\hat{x}$ is order-preserving when all pixels are subject to the same constraints $[l, u]$. In other words, if $\tilde{x}_i^* \geq \tilde{x}_j^*$, then $\hat{x}_i \geq \hat{x}_j$. More specifically,

- If $\hat{x}_i < \hat{x}_j$, then $t_{ij}(\hat{x}) = \text{sign}(\hat{x}_i - \hat{x}_j) = t_{ij}(\tilde{x}^*)$;
- If $\hat{x}_i = \hat{x}_j$, then $t_{ij}(\tilde{x}^*) \subseteq t_{ij}(\hat{x}) = [-1, 1]$.

If we multiply (6a) by w_i and add it to (4), then

$$w_i(\hat{x}_i - y_i) + \sum_{j \in N_i} \beta_{ij} t_{ij}(\tilde{x}^*) - \sum_{j: i \in N_j} \beta_{ji} t_{ji}(\tilde{x}^*) - (w_i \hat{\lambda}_i^1) + (w_i \hat{\lambda}_i^2) \ni 0 \quad (8)$$

which is almost (2a), except that in (8) t_{ij} are dependent on $\hat{x}^*$. The subset relation in (7) ensures that we can replace $t_{ij}(\hat{x}^*)$ by $t_{ij}(\hat{x})$ in (8) and the relation still holds. Then our claim is established as the triple $\{\hat{x}_I\}, \{w_i \hat{\lambda}_i^1\}, \{w_i \hat{\lambda}_i^2\}$ obviously satisfies (2b) and (2c).

3. Denoising with the fused lasso penalty subject to uniform intervals

We define the following minimization problem

$$\min_{\{x_i\}} \left\{ \Theta_2(\gamma, \eta) = \sum_{i=1}^N \frac{w_i}{2} (x_i - y_i)^2 + \gamma \sum_i \iota_{[l, u]}(x_i) + \eta \sum_i |x_i| + \sum_i \sum_{j \in N_i} \beta_{ij} |x_i - x_j| \right\}. \quad (9)$$

The I_1 penalty encourages sparse coefficients in the unknown x , with hyperparameter $\eta = 0$. When $w_i = w$ for all i , $\Theta_2(0, \eta)$ is the objective function of the fused lasso signal approximator (FLSA) [22]. When $\eta = 0$, we have $\Theta_2(\gamma, 0) \equiv \Theta_1(\gamma)$, thus (9) includes (1) as a special case. Again, since the objective function is strongly convex, there is a unique global minimizer to $\Theta_2(\gamma, \eta)$ for $\gamma = \{0, 1\}$ and all $\eta = 0$.

Applying the KKT conditions for convex nonsmooth problems [21, Theorem 3.34], the solution to (9) satisfies

$$w_i(x_i^* - y_i) + \eta s(x_i^*) + \sum_{j \in N_i} \beta_{ij} t_{ij}(x_i^*) - \sum_{j: i \in N_j} \beta_{ji} t_{ji}(x_i^*) - \lambda_i^{1*} + \lambda_i^{2*} \ni 0, \forall i \quad (10a)$$

$$x_i^* \leq u, \quad \lambda_i^{2*} \geq 0, \quad \lambda_i^{2*}(x_i^* - u) = 0, \forall i \quad (10b)$$

$$x_i^* \geq l, \quad \lambda_i^{1*} \geq 0, \quad \lambda_i^{1*}(x_i^* - l) = 0, \forall i \quad (10c)$$

Comparing (10) with (2), the only addition is the term related to η , where $s(x_i)$ is the subdifferential of $|x_i|$, namely $s(x_i) = \text{sign}(x_i)$ if $x_i \neq 0$, and $s(x_i) = [-1, 1]$ otherwise.

On the other hand, the solution to the unconstrained counterpart satisfies the first order optimality condition:

$$w_i(\tilde{x}_i^* - y_i) + \eta s(\tilde{x}_i^*) + \sum_{j \in N_i} \beta_{ij} t_{ij}(\tilde{x}_i^*) - \sum_{j: i \in N_j} \beta_{ji} t_{ji}(\tilde{x}_i^*) \ni 0, \forall i \quad (11)$$

The second (thresholding) step, which produces the output $\hat{x}_I$ from $\tilde{x}_i^*$ in (11), is the same as in (6) and will not be repeated. We claim that $\{\hat{x}_I\}, \{w_i \hat{\lambda}_i^1\}, \{w_i \hat{\lambda}_i^2\}$ is a triple that satisfies

the KKT conditions (10). Hence, we must have $\hat{x} = x^*$, the unique solution of $\Theta_2(1, \eta)$. To prove this, we multiply both sides of (6) by w_i and add it to (11) to obtain

$$w_i(\hat{x}_i - y_i) + \eta s(\tilde{x}_i^*) + \sum_{j \in N_i} \beta_{ij} t_{ij}(\tilde{x}^*) - \sum_{j: i \in N_j} \beta_{ji} t_{ji}(\tilde{x}^*) - (w_i \hat{\lambda}_i^1) + (w_i \hat{\lambda}_i^2) \ni 0 \quad (12)$$

which is almost (10a) except that the subdifferentials are evaluated at $\tilde{x}_i^*$ and not $\hat{x}_i$. As discussed in Sec 2, since all pixels are subject to the same constraint $[L, u]$ and since the thresholding step is order-preserving, we can replace $t_{ij}(\tilde{x}^*)$ by $t_{ij}(\hat{x})$ without affecting the validity of (12). If in addition, the interval $[L, u]$ contains zero, *i.e.*, $0 \in [L, u]$, then the thresholding step (3) is sign-preserving as well. Therefore we have $s(\hat{x}_i) \supseteq s(\tilde{x}_i^*)$ for all i . Indeed, if $\hat{x}_i > 0$ then $\tilde{x}_i^* > 0$ and thus $s(\hat{x}_i) = s(\tilde{x}_i^*)$; if $\hat{x}_i = 0$, then $\tilde{x}_i^*$ can be nonzero if $L = 0$ or $u = 0$, but we always have $s(\tilde{x}_i^*) \subseteq s(\hat{x}_i) = [-1, 1]$; if $\hat{x}_i < 0$, then $\tilde{x}_i^* < 0$ and thus $s(\hat{x}_i) = s(\tilde{x}_i^*)$. In summary, replacing $\tilde{x}^*$ by $\hat{x}$ maintains the correctness of (12). Hence our claim is

established as the triple $\{\hat{x}_i\}, \{w_i \hat{\lambda}_i^1\}, \{w_i \hat{\lambda}_i^2\}$ obviously satisfies (10b) and (10c).

It is interesting to point out that the FLSA itself, *i.e.*, the problem of minimizing $\Theta_2(0, \eta)$ for $\eta > 0$, admits a two-step solution when all weights $w_i = w$ are equal. First, the solution to the problem of minimizing $\Theta_2(0, 0)$ is obtained, then pixelwise soft-thresholding can be applied to the first step solution to obtain the solution to $\Theta_2(0, \eta)$ [23]. Combining this observation with our result, the solution to minimizing $\Theta_2(1, \eta)$ can be obtained in a sequential manner. Starting from the solution to $\Theta_2(0, 0)$, we apply pixelwise soft-thresholding (to minimize $\Theta_2(0, \eta)$) and then another thresholding to satisfy the interval constraints, *i.e.*, to solve $\Theta_2(1, \eta)$.

4. Discussion

Throughout the text, we have assumed that a solver for the unconstrained TV penalized problem is available. Our contribution has shown that applying thresholding to the result from this solver yields the solution to the problem with uniform interval constraints. Any solver for the nonsmooth TV problem can be used. When the problem has additional structures, *e.g.*, if the unknowns x_i form a chain [14, 24], a tree or an acyclic graph [25], a direct (noniterative) solution to the unconstrained problem can be obtained. Applying our result then produces a direct solution to the constrained counterpart.

Some problem instances in sparse reconstruction and machine learning are known to admit a sequential solution, see *e.g.*, [26–28]. These solutions can be addressed in a general setting using proximal decomposition [29]. We did not take the approach of proximal decomposition as that would limit us to uniform data weighting $w_i = w > 0$. Our sequential characterization of the solution to (1) and (9) works under general nonuniform data weighting.

5. Conclusions

We have discussed constrained image denoising problems that involve the anisotropic total variation (TV) penalty. Two specific problems were considered: (i) the constrained TV denoising problem, and (ii) the constrained fused lasso problem. When all unknowns are subject to the same interval constraint, and when the interval contains 0 for problem (ii), we have shown that the solution to the constrained problem is the thresholded version of the solution to the unconstrained problem. Our work thus provides justification for this approach, which is often adopted in an ad-hoc manner. However, note that the sequential approach is in general not correct for non-uniform interval constraints, nor is it correct in general for smooth TV using the Huber potential (see the counter examples in Appendix). Whether or not the sequential approach works with isotropic TV remains an open question. We also do not know if the sequential solution is valid for a generic image reconstruction problem. However, an image reconstruction algorithm often contains image denoising in an inner loop. Within such an inner loop, our sequential solution is applicable to solve the constrained image denoising problem.

Acknowledgments

The work of F. Noo was supported by Siemens Medical Solutions, USA; the concepts presented in this paper are based on research and are not commercially available.

Appendix

Examples

We provide two 1-D counter-examples showing that the two-step approach is incorrect for (i) non-uniform interval constraints and (ii) a smooth TV using the Huber function. We consider the following objective function,

$$\sum_{i=1}^N \frac{1}{2} w_i (x_i - y_i)^2 + \beta \sum_{i=1}^{N-1} \phi(x_{i+1} - x_i), \quad (13)$$

where $\phi(\cdot)$ is a generic penalty function penalizing neighboring differences. The followings are common parameters for both examples.

- Number of data points, $N=3$;
- Input $y = \{y_1, y_2, y_3\} = \{-1, 0, 1\}$;
- Data weight $w_1 = w_2 = w_3 = 1$;
- Penalty weight $\beta = 0.5$;

Example (i)

Here $\phi(s) = |s|$, then (13) is a special case of (1) in Sec 2. The unknowns x_1, x_2, x_3 are subject to the nonuniform constraints: $x_1 \leq -0.5$, $x_2 \leq 0$, and $x_3 \leq 0.5$. That is, we let the upper bound $u = \infty$ in this example. The unconstrained solution is $\bar{x}^* = \{-0.5, 0, 0.5\}$, as can be

verified by checking the first order optimality condition. To satisfy the nonuniform constraints, we need to threshold $\tilde{x}_1^*$ and obtain $\hat{x}^* = \{0.5, 0, 0.5\}$ which has an objective function value of 1.75. However, this is not optimal as another candidate $\{0.5, 0.5, 0.5\}$ achieves a lower objective value of 1.375. Thus, the 2-step solution is incorrect when the interval constraints are not the same for all unknowns.

Example (ii)

We replace the TV penalty by the Huber function, given by $\phi(s) = \inf_t \{ |t| + (s - t)^2/(2\varepsilon) \}$, where $\varepsilon > 0$ is a nuisance parameter. As $\varepsilon \downarrow 0$, $\phi(s) \uparrow |s|$. Alternatively, we can write $\phi(s) = s^2/(2\varepsilon)$ if $|s| \leq \varepsilon$, and $\phi(s) = |s| - \varepsilon/2$ otherwise. We set $\varepsilon = 0.1$ in this example. The unknown $x = \{x_1, x_2, x_3\}$ is subject to the non-negativity constraint, i.e., $x_i \geq 0$ for $i = 1, 2, 3$. It can be verified that the unconstrained solution $\tilde{x}^* = \{-0.5, 0, 0.5\}$ which is then thresholded to give $\hat{x} = \{0, 0, 0.5\}$ and the objective function at $\hat{x}$ is 0.85. This solution is not optimal as another candidate $\{0, 1/12, 0.5\}$ achieves a lower objective value of ≈ 0.83 . Thus the 2-step solution is not correct for smoothed TV in general, even with uniform interval constraints.

References

1. Nesterov Y. Smooth minimization of non-smooth functions. *Mathematical Programming*. Dec.2004 103:127–152.
2. Chambolle A, Pock T. A first-order primal-dual algorithm for convex problems with applications to imaging. *Journal of Mathematical Imaging and Vision*. 2011; 40(1):120–145.
3. Goldstein T, Osher S. The split bregman method for l1-regularized problems. *SIAM journal on imaging sciences*. 2009; 2(2):323–343.
4. Chambolle A. An algorithm for total variation minimization and applications. *Journal of Mathematical Imaging and Vision*. 2004; 20(1):89–97.
5. Tai, X-C., Wu, C. Augmented Lagrangian Method, Dual Methods and Split Bregman Iteration for ROF Model. Berlin, Heidelberg: Springer Berlin Heidelberg; 2009. p. 502-513.
6. Kunisch K, Hintermuller M. Total bounded variation regularization as a bilaterally constrained optimization problem. *SIAM Journal on Applied Mathematics*. 2004; 64(4):1311–1333.
7. Schafer RW, Mersereau RM, Richards MA. Constrained iterative restoration algorithms. *Proceedings of the IEEE*. Apr.1981 69:432–450.
8. Persson M, Bone D, Elmqvist H. Total variation norm for three-dimensional iterative reconstruction in limited view angle tomography. *Physics in Medicine and Biology*. 2001; 46(3):853. [PubMed: 11277230]
9. Dey N, Blanc-Feraud L, Zimmer C, Kam Z, Olivo-Marin JC, Zerubia J. A deconvolution method for confocal microscopy with total variation regularization. 2004 2nd IEEE International Symposium on Biomedical Imaging: Nano to Macro (IEEE Cat No. 04EX821). Apr.2004 2:1223–1226.
10. Krishnan D, Lin P, Yip AM. A primal-dual active-set method for non-negativity constrained total variation deblurring problems. *IEEE Transactions on Image Processing*. Nov.2007 16:2766–2777. [PubMed: 17990753]
11. Vogel, C. Computational Methods for Inverse Problems. Society for Industrial and Applied Mathematics; 2002.
12. Chen X, Ng MK, Zhang C. Non-Lipschitz ℓ_p -regularization and box constrained model for image restoration. *IEEE Transactions on Image Processing*. Dec.2012 21:4709–4721. [PubMed: 23008251]
13. Hansis E, Schafer D, Dossel O, Grass M. Evaluation of iterative sparse object reconstruction from few projections for 3-d rotational coronary angiography. *IEEE Transactions on Medical Imaging*. Nov.2008 27:1548–1555. [PubMed: 18955171]

14. Storath M, Brandt C, Hofmann M, Knopp T, Salamon J, Weber A, Weinmann A. Edge preserving and noise reducing reconstruction for magnetic particle imaging. *IEEE Transactions on Medical Imaging*. Jan.2017 36:74–85. [PubMed: 27455521]
15. Byrne C. Iterative algorithms for deconvolution and deblurring with constraints. *Inverse Problems*. 1998; 14:1455–1467.
16. Narayanan MV, Byrne CL, King MA. An interior point iterative maximum-likelihood reconstruction algorithm incorporating upper and lower bounds with application to SPECT transmission imaging. *IEEE transactions on medical imaging*. 2001; 20(4):342–353. [PubMed: 11370901]
17. Chartrand, R., Wohlberg, B. Total-variation regularization with bound constraints; 2010 IEEE International Conference on Acoustics, Speech and Signal Processing; Mar. 2010 p. 766-769.
18. Vogel C. Solution of linear systems arising in nonlinear image deblurring. *Scientific Computing: Proceedings of the Workshop*. 1998:148–158.
19. Fu H, Ng MK, Nikolova M, Barlow JL. Efficient minimization methods of mixed l_2 – l_1 and l_1 – l_1 norms for image restoration. *SIAM Journal on Scientific Computing*. 2006; 27(6):1881–1902.
20. Beck A, Teboulle M. Fast gradient-based algorithms for constrained total variation image denoising and deblurring problems. *IEEE Transactions on Image Processing*. Nov.2009 18:2419–2434. [PubMed: 19635705]
21. Ruszczy ski, AP. Nonlinear optimization. Vol. 13. Princeton University Press; 2006.
22. Tibshirani R, Saunders M, Rosset S, Zhu J, Knight K. Sparsity and smoothness via the fused lasso. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*. 2005; 67(1):91–108.
23. Friedman J, Hastie T, Hfling H, Tibshirani R. Pathwise coordinate optimization. *Ann. Appl. Stat.* 2007; 1:302–332. 12.
24. Condat L. A Direct Algorithm for 1-D Total Variation Denoising. *IEEE Signal Processing Letters*. Nov.2013 20:1054–1057.
25. Loeliger HA. An introduction to factor graphs. *IEEE Signal Processing Magazine*. Jan.2004 21:28–41.
26. Jenatton R, Mairal J, Obozinski G, Bach F. Proximal Methods for Hierarchical Sparse Coding. *Journal of Machine Learning Research*. Jul.2011 12:2297–2334.
27. Chartrand, R., Wohlberg, B. A nonconvex admm algorithm for group sparsity with sparse groups; 2013 IEEE International Conference on Acoustics, Speech and Signal Processing; May. 2013 p. 6009-6013.
28. Beygi S, Mitra U, Ström EG. Nested sparse approximation: Structured estimation of v2v channels using geometry-based stochastic channel model. *IEEE Transactions on Signal Processing*. Sep. 2015 63:4940–4955.
29. Yu, Y-L. On decomposing the proximal map. In: Burges, CJC. Bottou, L. Welling, M. Ghahramani, Z., Weinberger, KQ., editors. *Advances in Neural Information Processing Systems 26*. Curran Associates, Inc.; 2013. p. 91-99.