proceedings

Proceeding

The Analysis of Compressed Sensing for Total Variation Minimization and Bregman †

Xinyu Liu [1,*], Gongliu Yang [2]

[1] Beihang University, Beijing 100191, China; liuxy0306@buaa.edu.cn
[2] Tsinghua University, Beijing 100084, China; yanggongliu@mails.tsinghua.edu.cn
* Correspondence: liuxy0306@buaa.edu.cn; Tel.: +86-131-2668-0606

† Presented at the 4th International Electronic Conference on Sensors and Applications (ECSA 2017), 15–30 November 2017; Available online: https://sciforum.net/conference/ecsa-4.

Published: 14 November 2017

Abstract: 1) CS introduces a framework for simultaneous sensing and compression of big size vectors that applies in a range of applications including Optical Imaging and Synthetic Aperture Radar. 2) Total variation minimization; split Bregman; linearized Bregman and sparse reconstruction propose extremely efficient methods for solving optimization problems; which transform l_1-norm constrained problems into unconstrained problems by adding penalty term.. In the paper; the main principles of several algorithms are firstly introduced; then optimization iteration steps for algorithms are presented in detail. 3) Next; to research the performances of the algorithms in terms of the convergence and reconstruction precision; a series of numerical experiments for the above algorithms clearly show visual qualities of reconstructed images.4) we analyze the influence of the parameters μ and $\boldsymbol{\sigma}$ on iterative performances as well as the difficulties of controlling parameters; making clear the advantage of The Minimum total variation compared to other algorithms; and the low-complexity of Bregman .

Keywords: compressed sensing; total variation minimization; Bregman; sparse reconstruction

1. Introduction

In the application of image processing, when the acquisition process of measuring information is linear, reconstructing the target image from the measured data may be simplified as a linear system of equations. Using mathematical concepts, The relationship measured data $\mathbf{y} \in \boldsymbol{R}^m$ with signal vector $\boldsymbol{x} \in \boldsymbol{R}^{\mathrm{N}}$ can be described as:

$$\boldsymbol{y} = \boldsymbol{A}\boldsymbol{x} \quad (1)$$

Matrix $\boldsymbol{A} \in \boldsymbol{R}^{m\times \mathrm{N}}$（m≪N）establishes linear measurement process (referred to as measurement matrix), by solving the above linear equation restores the original signal x. If m < N, the classical linear algebra pointed out that problem (1) is underdetermined [1], and there are a number of solutions (assuming equation solution, at least one). It is impossible to reconstruct the signal x from y in case of $\mathrm{m} < \mathrm{N}$, which must meet that the sampling frequency of a continuous time signal must be twice more than the highest frequency to ensure the reconstruction.

In the following discussion, we often use norm. For a signal vector $\boldsymbol{x} =$ （$x_{1,\ldots,}x_n$）,its l_p norm denotes [2]:

$$\|\boldsymbol{x}\|_p = (\sum_{i=1}^{n}|x_i|^p)^{\frac{1}{p}}\ p \in [1,\infty] \quad (2)$$

Equation (1) are generally converted to a minimum l_0 norm optimization problem [2–4]:

$$\hat{x} = \arg\min \|x\|_0 \quad \text{st} \quad y = Ax \tag{3}$$

$$\min_{x \in R^N} \mu J_R(x) + \|Ax - y\|_2^2 \tag{4}$$

Equation (4) is a second-order cone programming problem. Using interior-point method can realize cone programming problem。

2. Experiments

2.1. Total Variation Minimization

Briefly, a function of the total variation is the Euclidean norm integral of functional gradient:

$$\|f\|_{BV} = \int |\nabla f(x)| dx \tag{5}$$

If f is discrete, it can be written as

$$\|f\|_{BV} = \sum_{i,j} \sqrt{|(\delta_1 f)_{i,j}|^2 + |(\delta_2 f)_{i,j}|^2} \tag{6}$$

where $(\delta_1 f)_{i,j} = f_{i,j} - f_{i-1,j}$, δ_2 is similar.

Total variation minimization solves the optimization problem $\min_{x} J(x)$ as follow [5]:

$$J(x) = \mu J_R(x) + \frac{1}{2}\|Ax - y\|_2^2 \tag{7}$$

where $\mu J_R(x)$ is called penalty term and $J_R(x) = \|x\|_{BV}$.

In following, we introduce mathematical methods called a half second neat to solve complex penalty term.

Conference [6] comes to using the new price of half a second structured functional to solve the optimization problem:

$$J^*(x, b^x, b^y) = \|Ax - y\|^2 + \omega^2 \sum_k \left[b_k^x (D^x x)_k^2 + \Phi(b_k^x) \right] + \omega^2 \sum_k \left[b_k^y (D^y x)_k^2 + \Phi(b_k^y) \right] \tag{8}$$

where the new half-quadratic regularization transform $J(x)$ with total variation term into where $(D^x x)_{i,j} = (x_{i,j+1} - x_{i,j}) / \delta, (D^y x)_{i,j} = (x_{i+1j} - x_{i,j}) / \delta.$ Iterations of the algorithm start with $x^0 = \mathbf{0}$. The iteration of b^x and b^y is

$$\left(b_{n+1}^x\right)_k = \frac{\varphi'\left((D^x x^n)_k\right)}{2(D^x x^n)_k}, \left(b_{n+1}^y\right)_k = \frac{\varphi'\left((D^x x^n)_k\right)}{2(D^x x^n)_k} \tag{9}$$

The iteration of x is solve the following equation

$$A^T A - \omega^2 C^{n+1} x^{n+1} = A^T y \tag{10}$$

where $C^{n+1} = -D_x^T B_x^{n+1} D_x - D_y^T B_y^{n+1} D_y$.

where D_x and D_y are convolution matrix generated from $\frac{1}{\delta}[1-1]$和$\frac{1}{\delta}[1-1]^T$.

In order to effectively solve (10) the result of the iterative process, we adopt the conjugate gradient algorithm [7].

2.2. Linearized Bregman

Because the design of measurement matrix A is mostly linear, so the above iterative process is simplified to

$$x^{n+1} = \arg\min J(x) + \frac{1}{\mu}\|Ax - y\|_2^2 \tag{11}$$

$$p^{n+1} = p^n - \frac{1}{\mu}\nabla\|Ax - y\|_2^2 \tag{12}$$

When the penalty term $J(\boldsymbol{x})=\|\boldsymbol{x}\|_1$, using linearized Bregman iteration [8] convert the above process into

$$x^{n+1} = \|x\|_1 - \left\langle \frac{1}{\mu}\nabla\|Ax - y\|_2^2 - p^n, x\right\rangle + \frac{1}{2\delta}\left\|x - x^n\right\|_2^2 \tag{13}$$

$$p^{n+1} = p^n - \frac{1}{\mu}\nabla\|Ax - y\|_2^2 \tag{14}$$

2.3. Split Bregman

Split Bregman iteration solve minimum l_1 norm, like

$$x = \arg\min\|\varphi(x)\|_1 + H(x) \tag{15}$$

where H is convex, φ is convex and differentiable. The basic idea is to put the problem down into the following questions [9,10]:

$$(x, d) = \arg\min\|d\|_1 + H(x) \quad \textit{subject to} \quad \varphi(x) = d \tag{16}$$

We add l_2 norm, then get unconstrainted problem

$$\arg\min_x \|d\|_1 + H(x) + \frac{1}{2}\|d - \varphi(x)\|^2 \tag{17}$$

The decomposition is introduced by Wang and Dr.Yin Zhang (FTVd)[].

We need a way to modify (17) the unconstrained problem and get accurate solution. For the problem (17), a simplified iteration method is given [11]:

$$(x^{k+1}, d^{k+1}) = \arg\min_x \|d\|_1 + H(x) + \frac{1}{2}\left\|d - \varphi(x) - b^k\right\|^2 \tag{18}$$

$$b^{k+1} = b^k + (\varphi(x) - d^{k+1}) \tag{19}$$

3. Results

1) Convergence [11]: the evolutions of x^k slowly keep convergence with the algorithm iteration, The minimization of two iterative results is determined by

$$\delta_x(k) \triangleq \sqrt{\frac{1}{M\times N}}\left\|x^k - x^{k-1}\right\|_F \tag{20}$$

2) Reconstruction precision [11]: reconstruction quality is measured by the mean variance error

$$\sigma \triangleq \frac{1}{N\times Iter}\sum_{k=1}^{Iter}\left\|\hat{x}_k - x_k^*\right\|_2^2 \tag{21}$$

where x_k^* is original image, $\hat{x}_k$ is the result of final iteration.

Assuming matrix $\boldsymbol{D}$ represent sparse manipulation[12–14], $\boldsymbol{D}^T$ transform sparse image into original image. Considering $\boldsymbol{\alpha} = \boldsymbol{Dx}$,so $\boldsymbol{\alpha}$ is called the sparse representation of image $\boldsymbol{x}$.

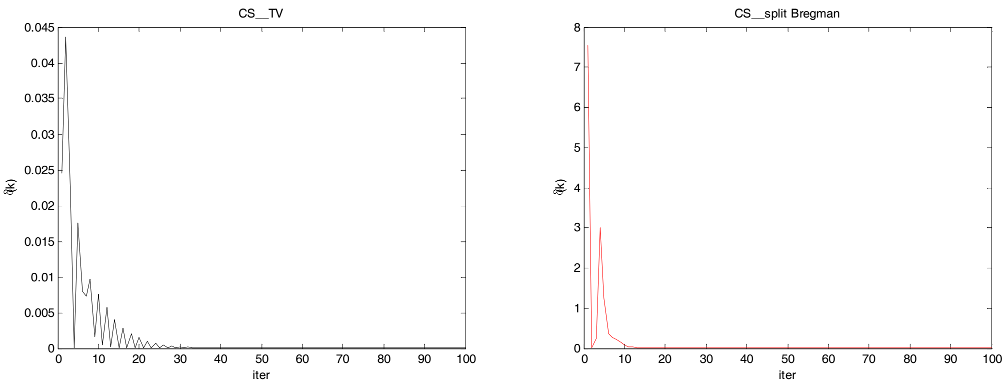
Total variation minimization is denoted CS_{TV} , Split Bregman has two schemes, which $J_R(\boldsymbol{x}) = \boldsymbol{x}$ denotes $CS_{splitBregman}$, $J_R(\boldsymbol{x}) = \boldsymbol{Dx}$ denotes CS_{sparse_model1} . Linearized Bregman chooses $J_R(\boldsymbol{x}) = \boldsymbol{\alpha}$ that (7) is rewritten as $\boldsymbol{\alpha} = argmin\|\boldsymbol{\alpha}\|_1 + \frac{\mu}{2}\|\boldsymbol{AD}^T\boldsymbol{\alpha} - \boldsymbol{y}\|_2^2$, which is denoted CS_{sparse_model2} .

3.1. Experiment 1 Simulation

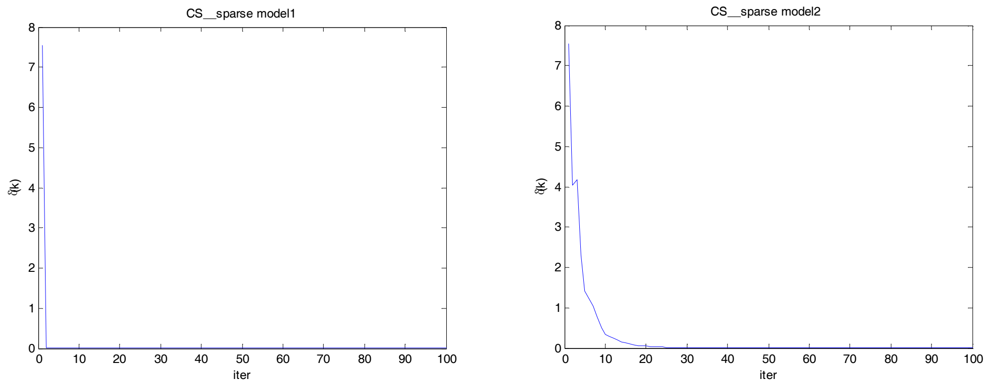
$M \times N$ measurement matrix is extracted from sample matrix with single pixel camera experiment in Rice University. （http://dsp.rice.edu/cscamera）,Haar wavelet transform produces $N \times L$ sparse matrix, and test data by $\boldsymbol{y}_k = \boldsymbol{Ax}_k$.

(**a**)Original image (**b**) CS_{TV} (**c**) $CS_{splitBregman}$ (**d**) CS_{sparse_model1} (**e**) CS_{sparsr_model2}

Figure 1. Simulation for letter.



(**a**) CS_{TV} (**b**) $CS_{splitBregman}$



(**c**) CS_{sparse_model1} (**d**) CS_{sparse_model2}

Figure 2. The relationship of $\delta_x(k)$ and $Iter$.

3.2. Experiment 2 Test Data Reconstruction

We recover images using measurement data obtained by experiment Rice single pixel camera makes. Measurement data are generally less than 60% of the reconstructed data, so next, measurement data with $M = 2048$ get reconstruction images with 64×64 ($N = 4096$) pixels, where experiments based on the algorithm of CS_{TV} and CS_{sparse_model2} are given.

4. Discussion

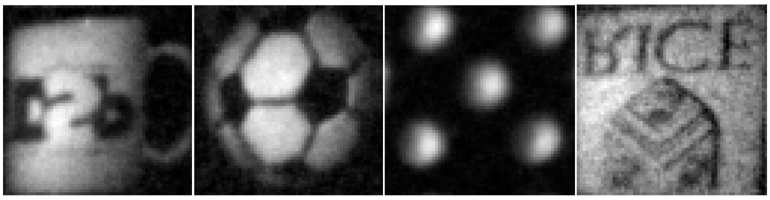


Figure 3. The reconstruction for mug, ball, dice and logo.

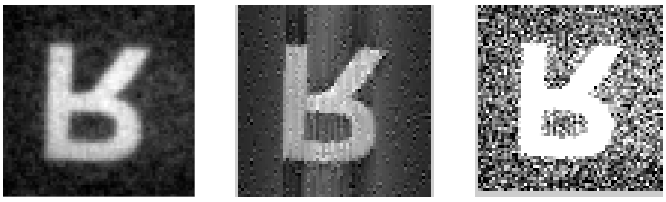


Figure 4. Synthetic data reconstruction.

Table 1. Algorithm Convergence M = 2048, N = 4096, K = 10,F = 2.

$\delta_x(k)Iter$	CS_{TV}	$CS_{splitbregman}$	CS_{sparse_model1}	CS_{sparse_model2}
10	5.62×10^{-7}	1.37×10^{-04}	3.43×10^{-05}	2.89×10^{-04}
50	2.31×10^{-9}	1.41×10^{-06}	3.1210^{-05}	2.52×10^{-05}
100	6.01×10^{-16}	1.40×10^{-09}	2.97×10^{-05}	2.01×10^{-05}

Table 2. Algorithm Accuracy M = 2048, N = 4096, K = 10,F = 2.

σ $Iter$	CS_{TV}	$CS_{splitbregman}$	CS_{sparse_model1}	CS_{sparse_model2}
10	0.05707	1.30×10^{-04}	0.01236	3.70×10^{-03}
50	0.05097	1.05×10^{-08}	0.01695	2.86×10^{-03}
100	0.05098	6.69×10^{-09}	0.03139	2.48×10^{-03}

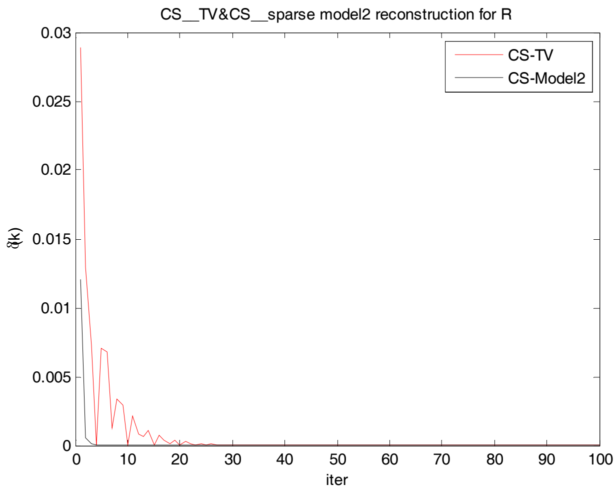


Figure 5. Convergence for CS_{TV} and CS_{SSB2} .

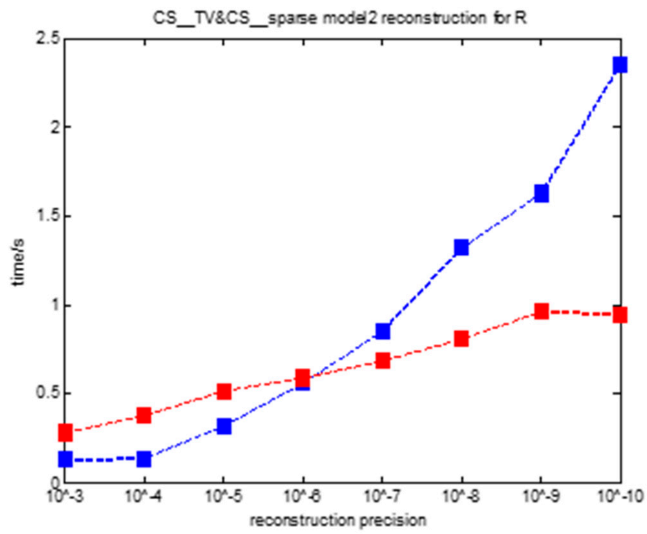


Figure 6. Relationship between accuracy and time.

Total variation minimization always keeps excellent performances no matter how synthetic data change, and its stability of convergence is relatively high compared to three other algorithms, which are susceptible to the effects of measurement noise. Split Bregman and linearized Bregman need control the parameters ,such as μ and σ , which determine convergence precision and the quality of reconstruction. But, Bregman has itself outstanding advantages that accelerates the speed of convergence and greatly simplifies the complexities for reconstruction process. Fig 5 shows the speeds of convergence for CS_{TV} and CS_{sparsr_model2} . The relationship between convex accuracy and time is given in Fig 6. The only drawback is the choices for appropriate parameters take a lot of work. So, in the next tasks, parameters for Bregman are important direction on research.

5. Conclusions

Compared to other algorithm, total variation minimization always maintains good convergence for different images and measurement data, and can obtain fine quality of images, which are susceptible to the effects of measurement noise. For Bregman, the parameters of μ and σ determine the qualities of images recovery precision and the speed of convergence.

Acknowledgments: This research was financially supported by the National Science Foundation.

Author Contributions: Paper's main work was accomplished by Liu Xinyu independently, teacher Yang Gongliu did careful guidance for the principles of the algorithms.

Conflicts of Interest: The authors declare no conflict of interest. The founding sponsors had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, and in the decision to publish the results.

References

1. Foucart, S.; Rauhut, H. *A Mathematical Introduction to Compressive Sensing*; Birkhäuser: Basel, The Switzerland, 2013.
2. Yonina, C.; Kutyniok, G.E.. *Compressed Sensing Theory and Applications*; Cambridge University Press: Cambridge, England, 2012.
3. Herzet, C.; Drémeau, A. Bayesian pursuit algorithms. In Proceedings of the 18th European Signal Processing Conference, Aalborg, Denmark, 23–27 August 2010; pp. 1474–1478.
4. Schniter, P.; Potter, L.C.; Ziniel, J. Fast Bayesian matching pursuit. In Proceedings of the Information Theory and Applications Workshop, 27 January–1 February 2008; pp. 326–333.
5. Candès, E.J.; Guo, F. New multiscale transforms, minimum total variation synthesis: Applications to edge-preserving image reconstruction. *Sig. Proc.* **2002**, *82*, 1519–1543.
6. Mouyan, Z. *Deconvolution and Aignal Recovery*; National defense Industry Press: Beijing, China, 2001.

7. Li, G.; Tang, C.; Li, L. High-efciency improved symmetric successive over-relaxation preconditioned conjugate gradient method for solving large-scale finite element linear equations. *Appl. Math. Mech*. **2013**, *34*, 1225–1236.
8. Yin, W.; Osher, S.; Goldfarb, D.; Darbon, J. Bregman iterative algorithms for l_1 minimization with applications to Compressed Sensing. *SIAM J.Imag. Sci*. **2008**, *1*, 143–168.
9. Goldstein, T.; Osher, S. The split Bregman method for L1-regularized problems. *SIAM J.Imag. Sci*. **2009**, *2*, 323–34.
10. Xu, Y.; Huang, T.Z.; Liu, J.; Lv, X.G. Split Bregman iteration algorithm for image deblurring using fourth-order total bounded variation regularization model. *J. Appl. Math*. **2013**, *3*, 417–433.
11. Daubechies, I.; DeVore, R.; Fornasier, M.; Güntürk, C.S. Iteratively reweighted least squares minimization for sparse recovery. *Commun. Pure Appl. Math.* **2010**, *63*, 1–38.
12. Rubinstein, R.; Zibulevsky, M.; Elad, M. Efficient implementation of the K-SVD algorithm using batch orthogonal matching pursuit. *CS Tech.* **2008**, *40*, 1–15.
13. Indyk, P. Explicit constructions for compressed sensing of sparse signals. In Proceedings of the Nineteenth annual ACM-SIAM Symposium on Discrete Algorithms, Society for Industrial and Applied Mathematics, San Francisco, CA, USA, 20–22 January 2008; pp. 30–33.
14. Rauhut, H. Compressive sensing and structured random matrices. In *Theoretical Foundations and Numerical Methods for Sparse Recovery*; Fornasier M, Ed.; Walter de Gruyter: Berlin, Germany, 2010; pp. 1–92.