

Efficient Compressed Sensing Based MRI Reconstruction using Nonconvex Total Variation Penalties

D. Lazzaro, E. Loli Piccolomini, F. Zama

Department of Mathematics, University of Bologna

E-mail: fabiana.zama@unibo.it

Abstract. This work addresses the problem of Magnetic Resonance Image Reconstruction from highly sub-sampled measurements in the Fourier domain. It is modeled as a constrained minimization problem, where the objective function is a non-convex function of the gradient of the unknown image and the constraints are given by the data fidelity term. We propose an algorithm, Fast Non Convex Reweighted (FNCR), where the constrained problem is solved by a reweighting scheme, as a strategy to overcome the non-convexity of the objective function, with an adaptive adjustment of the penalization parameter. We propose a fast iterative algorithm and we can prove that it converges to a local minimum because the constrained problem satisfies the Kurdyka-Lojasiewicz property. Moreover the adaptation of non convex l0 approximation and penalization parameters, by means of a continuation technique, allows us to obtain good quality solutions, avoiding to get stuck in unwanted local minima. Some numerical experiments performed on MRI sub-sampled data show the efficiency of the algorithm and the accuracy of the solution.

1. Introduction

Compressed Sensing (CS) is based on three major components: the sparsity of representation in some known domain, the incoherent measurements, which are usually achieved by irregular sampling patterns such as random sampling, and finally the sparsity-constrained non-linear reconstruction. Thanks to these properties CS exploits the compressibility of medical images to reconstruct data sampled below the Nyquist rate (undersampled), without loss of information. The technique can be applied to Magnetic Resonance Imaging (MRI) to increase imaging speed or to Computed Tomography (CT) to reduce radiation dose. MRI is a particular good candidate for the application of compressed sensing since data acquisition is performed in a transform domain (Fourier space) instead of the image domain, which facilitates the generation of incoherent aliasing artifacts. Many MRI CS imaging applications are modeled as constrained l_1 minimization problems:

$$\min_u \|\Psi(u)\|_1 \quad \text{s.t.} \quad \Phi u = z \quad (1)$$

where Ψ is a sparsity promoting operator, such as wavelet or discrete gradient, z are the undersampled data in the Fourier domain, $\Phi \in \mathbb{R}^{M \times N}$, $M < N$ is the data acquisition matrix and u is the reconstructed image. However the number of samples required to obtain good reconstructions is much larger compared to the minimum sampling rate that can be obtained



by l_0 minimization. For this reason, some authors (see for example [1], [2]) consider non-convex approximations of the l_0 norm, in order to reduce the sampling in the Fourier space. Following this approach, we define the objective function in (1) by approximating the l_0 norm with a non-convex function of the image gradient ($F(Du)$), and we consider the Lagrangian formulation:

$$\min_u \left\{ F(Du) + \frac{1}{2\lambda} \|\Phi u - z\|_2^2 \right\}, \quad \lambda > 0 \quad (2)$$

where λ is the regularization parameter, balancing the least squares data fidelity term and the non convex penalty term. One of the most outstanding technique for the solution of the non-convex minimization problem (2) is the Iterative Reweighted (IR) ℓ_1 scheme [3] where a local linear approximation of the non-convex $F(Du)$ is considered. Hence the problem (2) is approximated by a sequence of convex minimization problems. Each convex minimization problem that we obtain is solved by means of a forward-backward splitting algorithm, where the backward step, usually the most computationally expensive, is solved by means of the Weighted Split Bregman method [4], modified for the anisotropic case and suitably accelerated. This allows us to drastically reduce the computational costs.

The paper is organized as follows. In section 2 we state the problem to be solved and we describe the proposed algorithm; in section 3 some numerical experiments on MRI undersampled synthetic data are presented and finally in section 4 we report our conclusions.

2. The method

In this section we give a precise definition of the objective function $F(Du)$ in (2) and, after analyzing its properties we formalize our algorithm. We define $F(Du)$ as follows:

$$F(Du) = \sum_{i,j} (\psi_\mu(|u_x|_{i,j}) + \psi_\mu(|u_y|_{i,j}))$$

where $Du \in \mathbb{R}^{N \times 2N}$ is the gradient of the image u , whose components (u_x, u_y) are obtained by backward finite differences approximations of the partial derivatives in the x and y spatial coordinates respectively. The function $\psi : \mathbb{R} \rightarrow \mathbb{R}$ is a continuous, symmetric, twice differentiable on $\mathbb{R} \setminus \{0\}$, increasing and non-convex function, whose goal is to improve the sparsity promoting properties of ℓ_1 norm. Among the different examples reported the literature [3, 5], in this work we choose the following function:

$$\psi_\mu(t) = \frac{1}{\log(2)} \log \left(\frac{2}{1 + e^{-\frac{|t|}{\mu}}} \right), \quad \mu > 0$$

where μ is a positive constant such that, as $\mu \rightarrow 0$, the proposed $\psi_\mu(t)$ function approaches the l_0 norm [6].

The IR scheme. Since ψ is non-convex the problem (2) can be solved by using an IR ℓ_1 scheme [3], where it is substituted by a sequence of convex minimization problems whose solutions $u^{(\ell)}$, $\ell = 1, 2, \dots$ are proved to converge. The non-convex functional F is locally approximated by its convex majorant F_ℓ obtained by linearizing F at the gradient of an approximate solution $u^{(\ell)}$:

$$F_\ell(Du) = F(Du^{(\ell)}) + f_\ell(Du) - f_\ell(Du^{(\ell)})$$

where $Du^{(\ell)} = (u_x^{(\ell)}, u_y^{(\ell)})$ is the gradient of the image $u^{(\ell)}$ and

$$f_\ell(Du) = \sum_{i,j} \left(\psi'_\mu(|u_x^{(\ell)}|_{i,j}) |u_x|_{i,j} + \psi'_\mu(|u_y^{(\ell)}|_{i,j}) |u_y|_{i,j} \right).$$

Substituting F with F_ℓ in (2) and choosing $u^0 = \Phi^T y$ we obtain a sequence of approximate solutions:

$$u^{(\ell+1)} = \arg \min_u \mathcal{P}(\lambda, f_\ell(Du), u) \quad \ell = 0, 1, \dots \quad (3)$$

where

$$\mathcal{P}(\lambda, f_\ell(Du), u) = \left\{ f_\ell(Du) + \frac{1}{2\lambda} \|\Phi u - z\|_2^2 \right\}. \quad (4)$$

The convergence of the sequence $u^{(\ell)}$ defined by (3) to a local minimum of problem (2) is proved in the following theorem.

Theorem 2.1. *The sequence $u^{(\ell)}$ defined by (3) converges to a local minimum of problem (2).*

Proof. We observe that the objective function of (2) satisfies the Kurdika-Lojasiewicz property [7]. In fact

$$\sum_{i,j} (\psi_\mu(|u_x|_{i,j}) + \psi_\mu(|u_y|_{i,j})) + \frac{1}{2\lambda} \|\Phi u - z\|_2^2$$

is an analytic function, because the function ψ is evaluated in nonnegative arguments, therefore we can restrict the domain of ψ to $\mathbb{R}^+ \cup \{0\}$. Finally the thesis immediately follows from [8], where the authors prove the convergence of the IR algorithm when the objective function satisfies the Kurdika-Lojasiewicz property. \square

The FNCR Algorithm. In order to avoid sub optimal solutions, due to the presence of many local minima, the parameters μ and λ are decreased on the basis of a continuation strategy. An outer loop, which depends on the index k , is used to decrease the parameter μ from its initial value $\mu \simeq 1$ as follows $\mu = \max(0.8\mu, 10^{-4})$, implementing in this way a graduated non-convexity procedure. Concerning the parameter λ , starting with an initial estimate $\lambda_0 = r_\lambda \|u^{(0)}\|_1$, where $0 < r_\lambda < 1$ is an assigned constant, we propose an iterative update by observing that if $\lambda_k < \lambda_{k-1}$, and $u^{(k)} = \arg \min_u \mathcal{P}(\lambda_k, f_k(Du), u)$, then it easily follows that

$$\mathcal{P}(\lambda_k, f_k(Du^{(k)}), u^{(k)}) < \mathcal{P}(\lambda_{k-1}, f_{k-1}(Du^{(k-1)}), u^{(k-1)})$$

namely, [9]

$$\lambda_{k+1} = \lambda_k \frac{\mathcal{P}(\lambda_k, f_k(Du^{(k)}), u^{(k)})}{\mathcal{P}(\lambda_{k-1}, f_{k-1}(Du^{(k-1)}), u^{(k-1)})}. \quad (5)$$

Moreover a warm starting technique is applied to the iterative solution of (3). These strategies allow us to obtain good quality solutions, avoiding to get stuck in unwanted local minima.

We include all the above observations in the Fast Non Convex Reweighting Algorithm (FNCR) and we report it in **Algorithm 2.1**. The solution of problem (3) is obtained by the iterative Forward-Backward (6)-(7) splitting approach [10] combined with a FISTA (8) acceleration strategy [11]. The forward step (6) depends on the parameter $\beta \in (0, 2/\lambda_{\max}(\Phi^t \Phi)]$, where $\lambda_{\max}(\Phi^t \Phi)$ is the maximum eigenvalue of the matrix $\Phi^t \Phi$. In MRI CS we have $\lambda_{\max} = 1$ hence $\beta \in (0, 2]$. In the present work we set $\beta = 1$ since this value satisfies the necessary condition for the convergence. The Backward step (7) consists of a weighted Total Variation Minimization and for its solution we use the Weighted Split Bregman method [4], where the fidelity term is modified by introducing suitable weights depending from the image edges. The forward-backward iterations are stopped as soon as the following **stopping condition** is true:

$$\Delta_n - \Delta_{n-1} < \gamma \lambda_k, \quad \text{where} \quad \Delta_n = \sum_{i,j=1}^N \left((w_{i,j}^x) |\hat{u}_x^{(n)}|_{i,j} + (w_{i,j}^y) |\hat{u}_y^{(n)}|_{i,j} \right). \quad (9)$$

Algorithm 2.1 (INPUT: $y, \Phi, \lambda_0, r_\lambda > 0, \gamma > 0, \mu > 0$; OUTPUT: $\lambda_{k-1}, u^{(k-1)}$).

FNCR Algorithm

$$u^{(0)} = \phi^T z, (w_{i,j}^x) = 1, (w_{i,j}^y) = 1, \lambda_0 > 0$$

$$\hat{u}^{(0)} = u^{(0)}, k = 0$$

repeat (*IR iteration*)

$$t_0 = 1, n = 0$$

repeat

(*Forward Step*)

$$v = \hat{u}^{(n)} + \beta \Phi^T (z - \Phi \hat{u}^{(n)}) \quad (6)$$

(*Backward Step*)

$$\tilde{u}^{(n+1)} = \arg \min_{u \in \mathbb{R}^{N \times N}} \left\{ \sum_{i,j=1}^N \left((w_{i,j}^x) |u_x|_{i,j} + (w_{i,j}^y) |u_y|_{i,j} \right) + \frac{1}{2\beta\lambda_k} \|u - v\|_2^2 \right\} \quad (7)$$

(*Fista Acceleration*)

$$t_{n+1} = \frac{1 + \sqrt{1 + 4t_n^2}}{2}, \alpha_n = \frac{t_n - 1}{t_{n+1}}$$

$$\hat{u}^{(n+1)} = \tilde{u}^{(n+1)} + \alpha_n (\tilde{u}^{(n+1)} - \tilde{u}^{(n)}) \quad (8)$$

$$n = n + 1$$

until stopping condition as in (9)

Update λ_{k+1} as in (5)

$$\mu = \mu \cdot 0.8 \text{ } (\mu \text{ updating})$$

$$u^{(k+1)} = \hat{u}^{(n)}$$

$$\hat{u}^{(0)} = u^{(k+1)} \text{ } (\text{Warm starting})$$

$$(w_{i,j}^x) = \psi'_\mu(|u_x^{(k+1)}|_{i,j}), (w_{i,j}^y) = \psi'_\mu(|u_y^{(k+1)}|_{i,j}), \text{ } (\text{Weights updating})$$

$$k = k + 1$$

until convergence

The parameter γ is a positive constant and its value is important to adapt the precision of the method to the difficulty of the reconstruction problem (small number of samples). In case of strong under-sampling, as reported in the numerical experiments of section 3, we use $\gamma \in [10^{-4}, 10^{-1}]$.

3. Numerical Experiments

In this section we report a few preliminary results aimed to prove the effectiveness of the FNCR algorithm in terms of reconstruction quality and computational efficiency. All the reported experiments are performed on an Intel i7-3770 CPU with 16 GB of RAM, using Matlab R2012b. We are interested in testing FNCR in very critical situations, therefore we choose a difficult test image such as the `forbild` phantom [12] represented in figure 1(a). The synthetic undersampled data z are obtained as $z = \Phi u$ where u is the phantom image and Φ is the Fourier matrix undersampled according to K-space (K) or Radial (R) sampling strategies. In the case of K-space sampling we defined the test problems K16 and K32, obtained by acquiring 16 and 32 vertical lines of the image Fourier spectrum, with a Gaussian random phase encoding. The

sampling rates are 12.5% and 25%(figure 1(c)) respectively. The Radial under-sampling is used in the test problems: R9, R10, R12 where 9, 10, 12 radial lines are acquired in the Fourier space with sampling rates 3.7%, 4.1%(figure 1(d)) and 4.8% respectively. In our experiments we focus on the best accuracy achieved by the FNCR algorithm. To this purpose we stop the outer iterations of algorithm 2.1 as soon as $psnr(U)$ reaches its maximum or when $psnr(U) \geq 100$, where U is the image reconstructed by the FNCR algorithm, u is the `forbld` phantom and

$$psnr(U) = 20 \log_{10} \left(\frac{\max_{i,j} u_{i,j}}{rmse(U)} \right), \quad rmse(U) = \frac{1}{N} \sqrt{\sum_{i,j=1}^N (u_{i,j} - U_{i,j})^2}.$$

Since in our tests we scale the values of the phantom image u in the interval $[0, 1]$, the case $psnr(U) = 100$ is obtained when $rmse(U) = 10^{-5}$, indicating a reconstruction of extremely good quality. In Table 1 we report the $psnr$ value of the initial image $u^{(0)}$ (column 3) and in columns 4 and 5, the $psnr$ values and the execution times, obtained by algorithm 2.1, with the following input parameters: $\lambda_0 = r_\lambda \|u^{(0)}\|_1$, r_λ in the range $[10^{-3}, 5 \cdot 10^{-3}]$, $\beta = 1$, γ in the range $[5 \cdot 10^{-3}, 10^{-1}]$. The $psnr$ values confirm the excellent accuracy reached by K32, R12 and R10 tests. In the cases K16 and R9, where the under-sampling is very severe, lower accuracy is obtained in both cases with $psnr \simeq 30$. Observing the reconstructed images in figures 2(b) and 3(b) we see that the critical under-sampling mainly affects the reconstruction of the small details (figures 2(c) and 3(c)), preserving good quality in the reconstruction of the coarsest features. To get a feel of the quality of our reconstructions compared to the current literature, we report the results of the Chartrand's Method [CM] in [13], obtained by replacing the anisotropic p -shrinkage (with $p = 0.5$) with the soft-thresholding, in the function `mricks.m` available at http://www.ece.rice.edu/~tag7/Tom_Goldstein/Split_Bregman_files/mricks.zip, and setting the parameters in order to obtain the most accurate results. Comparing the columns 4 and 6 in Table 1 we observe that the optimal results are never reached by CM obtaining the best accuracy in test R12 with $psnr \simeq 88$. Furthermore we see that, in the most difficult cases (K16 and R9), both algorithms have quite the same accuracy. Finally comparing the computation times, reported in Table 1 (columns 5 and 7) we see that FNCR is always faster than CM in all the test cases.

Test	%	$psnr(u^{(0)})$	FNCR		CM	
			$psnr(U)$	<i>sec.</i>	$psnr(U)$	<i>sec.</i>
K32	25%	17.38	100.05	38.5	32.75	118.0
K16	12.5%	13.50	30.70	6.4	30.65	117.0
R12	4.8%	16.36	100.12	8.3	87.80	130.0
R10	4.1%	16.32	100.1	31.9	32.02	136.4
R9	3.7%	15.69	28.41	31.0	26.87	90.5

Table 1: FCNR and CM results.

4. Conclusions

The results obtained in these preliminary tests encourage us to continue a deeper analysis of the algorithm. More synthetic and real data will be analyzed and the sensitivity of the FNCR algorithm to parameters setting and data noise will be investigated. Finally a comprehensive comparison with the methods proposed in the recent literature will be performed.



Figure 1: (a) Forbield Phantom (256×256); (b) image detail; (c) K16 test; (d) R9 test

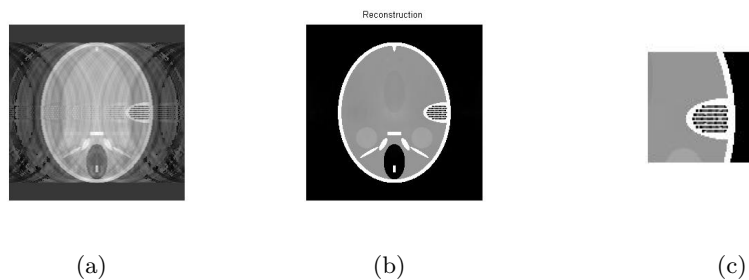


Figure 2: K16 test. (a) u^0 . (b) FNCR reconstruction. (c) Zoom

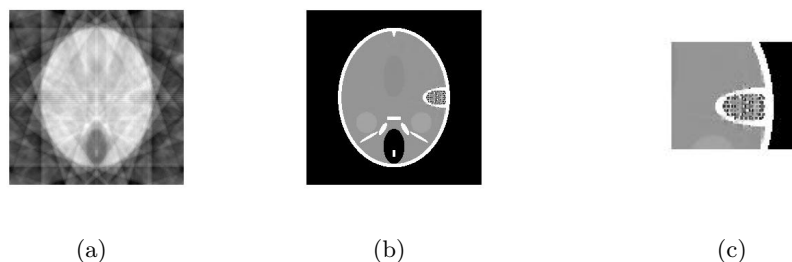


Figure 3: R9 test. (a) u^0 . (b) FNCR reconstruction. (c) Zoom

References

- [1] Chartrand R 2007 *IEEE Sig. proc. Lett.* **14** 707–710
- [2] Trzasko J and Manduca A 2007 *IEEE Sig. proc. Lett.* **14** 707–710
- [3] Candes E and Wakin M 2008 *Journal of Fourier Analysis and Applications* **14** 877–905
- [4] Zhang X, Burger M, Bresson X and Osher S 2010 *SIAM J. IMAGING SCIENCES* **3** 253–276
- [5] Mourad N and Reilly J 2010 *IEEE Transactions on Signal Processing* **58** 3485–3496
- [6] Montefusco L B, Lazzaro D and Papi S 2013 *Signal Processing* **93** 2636 – 2647 ISSN 0165-1684
- [7] Attouch H, Bolte J and Svaiter B F 2013 *Math. Program Ser. A* **137** 91–129
- [8] Ochs P, Dosovitskiy A, Brox T and Pock T 2015 *SIAM Journal on Imaging Sciences* **8** 331–372
- [9] Montefusco L B and Lazzaro D 2012 *IEEE Transactions on Image Processing* **21** 1676–1686
- [10] Combettes P L and Wajs V R 2005 *Multiscale Modeling & Simulation* **4** 1168–1200
- [11] Beck A and Tebulle M 2009 *IEEE Tran. on Imag. Proc.* **18** 2419–2434
- [12] Yu Z, Noo F, Dennerlein F, Wunderlich A, Lauritsch G and Hornegger J 2012 *Phys. Med. Biol.* **57**
- [13] Chartrand R 2009 Fast algorithms for nonconvex compressive sensing: Mri reconstruction from very few data *IEEE International Symposium on Biomedical Imaging (ISBI)*