MRI RECONSTRUCTION WITH ANALYSIS SPARSE REGULARIZATION UNDER IMPULSIVE NOISE

A. Korhan Tanc

Department of EEE
Kirklareli University
Kayali, 39100, Kirklareli, Turkey
korhan.tanc@kirklareli.edu.tr

Ender M. Eksioglu

Department of ECE
Istanbul Technical University
Maslak, 34469, Istanbul, Turkey
eksioglue@itu.edu.tr

ABSTRACT

We will be considering analysis sparsity based regularization for Magnetic Resonance Imaging reconstruction. The analysis sparsity regularization is based on the recently introduced Transform Learning framework, which has reduced complexity regarding other sparse regularization methods. We will formulate a variational reconstruction problem which utilizes the analysis sparsity regularization together with an ℓ_1 norm based data fidelity term. The use of the non-smooth data fidelity term results in robustness against outliers and impulsive noise in the observed data. The resulting algorithm with the ℓ_1 observation fidelity showcases enhanced performance under impulsive observation noise when compared to a similar algorithm utilizing the conventional quadratic error term.

Index Terms— Magnetic resonance, image reconstruction, compressed sensing, analysis sparsity, impulsive noise

1. INTRODUCTION

We consider the inverse problem of Magnetic Resonance Image (MRI) reconstruction from heavily undersampled observations under impulsive noise. Sparsity based regularization methods using a variational framework have been very popular recently. Synthesis sparsity has been on the forefront of this research torrent, with compressed sensing based data acquisition being its best known application [1]. Dictionary learning (DL) approaches together with synthesis sparsity have shown state-of-the-art performance starting from the simplest inverse imaging problem of denoising [2] and leading to more complex inverse imaging problems such as the MRI reconstruction [3]. Analysis sparsity on the other hand has been a lesser studied doppelganger for synthesis sparsity [4]. Recently, a new approach to analysis sparse modeling has been developed under the title of "Transform Learning" [TL] [5]. Transform Learning (TL) provides a framework with much reduced complexity for learning an analysis sparsity operator and for executing the analysis sparsity prior as an image model. After its initial conception in the image denoising milieu [5], the TL framework with its learned analysis sparsity model has also been successfully applied to the MRI reconstruction problem under the title of TLMRI [6]. When compared to the previous synthesis sparsity and DL based MRI reconstruction algorithms, TLMRI maintained comparable performance with greatly reduced complexity due to its avoidance of the costly sparse coding step. The TLMRI reconstruction algorithm was recently enriched by the inclusion of a more traditional global wavelet transform sparsity term, utilizing both the patch-wise analysis sparsity and the image-wide global wavelet transform sparsity [7], leading to the Globally Regularized TLMRI (G-TLMRI) algorithm. All the reconstruction algorithms mentioned thus far have used a sparsity based image model together with an ℓ_2 norm based quadratic error term for data or observation fidelity. The use of an ℓ_2 norm based data fidelity term in the overall variational framework presumes Gaussian distribution for the observation noise. In this work we want to consider the case of non-Gaussian and especially impulsive observation noise in the MRI data acquisition setting. We will modify the TLMRI reconstruction algorithm by incorporating an ℓ_1 norm data fidelity term in the overall cost function. The inclusion of the ℓ_1 norm data fidelity term results in improved reconstruction performance when compared to the original TLMRI algorithm.

2. TLMRI MEETS ℓ_1 DATA FIDELITY

The Transform Learning (TL) method as an adaptive analysis sparsity based image model was initially applied to image denoising in [5]. The application realm of TL was extended to MRI image reconstruction in [6] under the rubric of TLMRI. For the MRI reconstruction problem, the data acquisition is modeled as follows.

$$\mathbf{y} = \boldsymbol{\mathcal{F}}_u \mathbf{x}^\star + \boldsymbol{\eta} \tag{1}$$

Here, $\mathbf{y}$ is the observation or data vector in the Fourier domain. The operator $\boldsymbol{\mathcal{F}}_u$ models the downsampled Fourier transform operator from the vectorized image space to the observation space, with $\mathbf{x}^\star$ being the true vectorized image. The

vector $\boldsymbol{\eta}$ symbolizes the observation or data noise.

In TLMRI, as in most similar variational methods for image reconstruction problems, the data fidelity term is formulated as a quadratic error using the ℓ_2 norm. Under a Maximum Likelihood (ML) estimation formulation, the ℓ_2 regularization suggests a Gaussian distribution for the observation noise $\boldsymbol{\eta}$, which is the most common assumption. Here we will be considering a non-Gaussian and especially impulsive distribution for the observation noise. Such a non-Gaussian noise assumption leads to a non-smooth data fidelity term [8] from the ML perspective. Especially, a Laplacian noise distribution would result in an ℓ_1 data fidelity term, which has been studied in the literature under the least absolute deviation (LAD) regression [8]. Here, we will use the non-smooth ℓ_1 data fidelity term with an impulsive observation noise. Using the TLMRI cost function as the starting point, the cost function for the new algorithm can be given as follows.

$$\text{(P1)} \min_{\mathbf{W},\hat{\boldsymbol{\mathcal{X}}},\boldsymbol{\mathcal{A}},\mathbf{x}} \|\mathbf{W}\hat{\boldsymbol{\mathcal{X}}} - \boldsymbol{\mathcal{A}}\|_F^2 + \lambda Q(\mathbf{W}) + \beta\|\boldsymbol{\mathcal{A}}\|_1 + \tau\|\boldsymbol{\mathcal{P}}(\mathbf{x}) - \hat{\boldsymbol{\mathcal{X}}}\|_F^2 + \eta\|\boldsymbol{\mathcal{F}}_u\mathbf{x} - \mathbf{y}\|_1. \quad (2)$$

In (2), $\|\cdot\|_F$ denotes the Frobenius norm for the argument matrix. The adaptive operator $\mathbf{W} \in \mathbb{C}^{n\times n}$ is the learned sparsifying transform. The matrix $\hat{\boldsymbol{\mathcal{X}}} \in \mathbb{C}^{n\times M}$, and its column vectors $\hat{\mathbf{x}}_j \in \mathbb{C}^n$ are image patches of size $\sqrt{n} \times \sqrt{n}$ given in vectorized form. The matrix $\boldsymbol{\mathcal{A}} \in \mathbb{C}^{n\times M}$ with columns $\boldsymbol{\alpha}_j \in \mathbb{C}^n$ store the sparse codes for the patches. The function $Q(\cdot)$ implements the required penalization such that the adaptive sparsifying transform $\mathbf{W}$ avoids degenerate solutions [6]. $\boldsymbol{\mathcal{P}}$ is an operator which models the extraction of vectorized patches from the argument image matrix.

In (2), $\|\boldsymbol{\mathcal{F}}_u\mathbf{x} - \mathbf{y}\|_1$ is the data fidelity term, which enforces the connection between the observation $\mathbf{y}$ and the reconstructed image $\mathbf{x}$. All the remaining terms are about the enforcement of the TL image model onto the reconstructed image $\mathbf{x}$ using a learned sparsifying operator $\mathbf{W}$. These terms are responsible for the regularization of the reconstructed image such that the reconstructed image $\mathbf{x}$ stays sparsifiable under a learned transform $\mathbf{W}$. The cost function (2) is mostly equivalent to the formulations as utilized in [6] and [7]. The main difference in (2) is the introduction of the ℓ_1 based data fidelity term $\|\boldsymbol{\mathcal{F}}_u\mathbf{x} - \mathbf{y}\|_1$ instead of the quadratic formulation. The introduction of the ℓ_1 data fidelity leads to an algorithm which is especially robust against observation noise with heavy-tailed or impulsive distribution. In the coming section we present a modified TLMRI algorithm which solves the novel cost function (2) with the ℓ_1 data fidelity.

3. THE LAD-TLMRI ALGORITHM

We will modify the TLMRI algorithm as to handle the ℓ_1 data fidelity in the cost function (P1) (2). The problem (P1) gets separated into two main steps, one involving $\mathbf{x}$ and one without it. The first step enforces the analysis sparsity model through transform learning onto patches of the current reconstructed image.

$$\min_{\mathbf{W},\hat{\boldsymbol{\mathcal{X}}},\boldsymbol{\mathcal{A}}} \|\mathbf{W}\hat{\boldsymbol{\mathcal{X}}} - \boldsymbol{\mathcal{A}}\|_F^2 + \lambda Q(\mathbf{W}) + \beta\|\boldsymbol{\mathcal{A}}\|_1 + \tau\|\boldsymbol{\mathcal{P}}(\mathbf{x}) - \hat{\boldsymbol{\mathcal{X}}}\|_F^2. \quad (3)$$

This step is common to both the TLMRI and G-TLMRI algorithms as introduced in [6] and [7], respectively. For the purpose of completeness, we want to repeat here the solution as we introduced in [7]. We will use two distinct substeps each iterating on just two of the variables, with $\boldsymbol{\mathcal{A}}$ being common to both substeps. The two substeps are as follows [7]:

$$\min_{\boldsymbol{\mathcal{A}}} \|\mathbf{W}\hat{\boldsymbol{\mathcal{X}}} - \boldsymbol{\mathcal{A}}\|_F^2 + \beta\|\boldsymbol{\mathcal{A}}\|_1. \quad (4a)$$

$$\min_{\mathbf{W}} \|\mathbf{W}\hat{\boldsymbol{\mathcal{X}}} - \boldsymbol{\mathcal{A}}\|_F^2 + \lambda Q(\mathbf{W}). \quad (4b)$$

$$\min_{\boldsymbol{\mathcal{A}}} \|\mathbf{W}\hat{\boldsymbol{\mathcal{X}}} - \boldsymbol{\mathcal{A}}\|_F^2 + \beta\|\boldsymbol{\mathcal{A}}\|_1. \quad (5a)$$

$$\min_{\hat{\boldsymbol{\mathcal{X}}}} \|\mathbf{W}\hat{\boldsymbol{\mathcal{X}}} - \boldsymbol{\mathcal{A}}\|_F^2 + \tau\|\boldsymbol{\mathcal{P}}(\mathbf{x}) - \hat{\boldsymbol{\mathcal{X}}}\|_F^2. \quad (5b)$$

The subequations (4a) and (5a) are equivalent, and they are solved by soft thresholding. The subequations (4b) and (5b) have distinct least squares solutions [6]. Eqns. (4) and (5) have defined the substep for solving the analysis sparsity based regularization procedure given in (3).

The second substep for the solution of (P1) is a reconstruction step with an ℓ_1 data fidelity constraint.

$$\min_{\mathbf{x}} \|\boldsymbol{\mathcal{F}}_u\mathbf{x} - \mathbf{y}\|_1 + \tfrac{\tau}{\eta}\|\boldsymbol{\mathcal{P}}(\mathbf{x}) - \hat{\boldsymbol{\mathcal{X}}}\|_F^2. \quad (6)$$

Here, we would like to define an operator $\hat{\boldsymbol{\mathcal{P}}}(\cdot)$ which creates a vectorized image from a matrix of processed patches. This trick which we also utilized in [7] saves us from the burden of repeatedly switching between the patch and image domains as we iteratively search for a solution of (6). After the introduction of $\hat{\boldsymbol{\mathcal{P}}}(\cdot)$, (6) can be approximately rewritten as follows.

$$\min_{\mathbf{x}} \|\boldsymbol{\mathcal{F}}_u\mathbf{x} - \mathbf{y}\|_1 + \tfrac{\tau'}{2}\|\mathbf{x} - \hat{\boldsymbol{\mathcal{P}}}(\hat{\boldsymbol{\mathcal{X}}})\|_2^2. \quad (7)$$

In (7) we have a combination of a non-smooth data fidelity term with a smooth, quadratic regularization term. We will adopt a gradient descent algorithm for solving (7). Because of the ℓ_1 term we will be using a subgradient for the data fidelity term. We present the gradient descent step and the corresponding gradient definition in the following.

$$\mathbf{x} = \mathbf{x} - \gamma\nabla g(\mathbf{x}). \quad (8a)$$

$$\nabla g(\mathbf{x}) = \boldsymbol{\mathcal{F}}_u^H \operatorname{sgn}(\boldsymbol{\mathcal{F}}_u\mathbf{x} - \mathbf{y}) + \tau'(\mathbf{x} - \hat{\boldsymbol{\mathcal{P}}}(\hat{\boldsymbol{\mathcal{X}}})). \quad (8b)$$

The parameter γ is the step-size for the gradient descent. The operator $\boldsymbol{\mathcal{F}}_u^H$ which realizes zero-filled reconstruction is the adjoint of $\boldsymbol{\mathcal{F}}_u$. The $\operatorname{sgn}\{\cdot\}$ operator which calculates the elementwise sign function for the argument vector arises from

Algorithm 1 LAD-TLMRI Algorithm

Input: Observation, $\mathbf{y} = \boldsymbol{\mathcal{F}}_u\mathbf{x}^\star + \boldsymbol{\eta}$; parameters $\lambda, \beta, \tau, \tau', \gamma$.

Goal: $\min_{\mathbf{W},\hat{\boldsymbol{\mathcal{X}}},\boldsymbol{\mathcal{A}},\mathbf{x}} \|\mathbf{W}\hat{\boldsymbol{\mathcal{X}}} - \boldsymbol{\mathcal{A}}\|_F^2 + \lambda Q(\mathbf{W}) + \beta\|\boldsymbol{\mathcal{A}}\|_1 + \tau\|\boldsymbol{\mathcal{P}}(\mathbf{x}) - \hat{\boldsymbol{\mathcal{X}}}\|_F^2 + \eta\|\boldsymbol{\mathcal{F}}_u\mathbf{x} - \mathbf{y}\|_1$

1: Initialize $\mathbf{x} = \boldsymbol{\mathcal{F}}_u^H\mathbf{y}$.
2: **for** $i := 1, 2, \ldots$ **do** ▷ main iteration
3: Initialize $\hat{\boldsymbol{\mathcal{X}}} = \boldsymbol{\mathcal{P}}(\mathbf{x})$. ▷ denoising starts
4: Iterate (4), N_1 times.
5: Iterate (5), N_2 times.
6: Initialize $\mathbf{x} = \hat{\boldsymbol{\mathcal{P}}}(\hat{\boldsymbol{\mathcal{X}}})$. ▷ reconstruction starts
7: Iterate (8a), N_3 times.
8: **end for** ▷ end of main iteration
9: Output reconstructed MR image $\mathbf{x}$.

the subgradient of the ℓ_1 data fidelity. The reconstruction step realized by iterating (8a) completes the algorithm for solving (P1) (2). We will call the resulting overall algorithm as the Least Absolute Deviation TLMRI (LAD-TLMRI). A complete portrayal of the LAD-TLMRI is depicted in Alg.1.

4. SIMULATION RESULTS

In this section, we investigate the performance of our LAD-TLMRI algorithm through simulations. For comparison, we also simulate zero-filling reconstruction and original TLMRI algorithm [6]. In our MRI acquisition scenario we assume random sampling mask with 4-fold undersampling [7]. Furthermore we assume complex impulsive noise with mean zero and standart deviation σ, where real and imaginary parts are i.i.d. and have generalized Gaussian distribution [9, 10]:

$$f(\eta) = \frac{b}{2a\Gamma\{1/b\}} \exp\{-|\eta/b|^2\}. \tag{9}$$

Here, Γ is the complete Gamma function, b is the shape parameter, and a is the scale parameter defined by

$$a = \sigma\sqrt{\frac{\Gamma\{1/b\}}{2\Gamma\{3/b\}}}. \tag{10}$$

For the special cases of $b = 2$ and $b = 1$, the resulting noise would have Gaussian and Laplacian distributions, respectively. As b approaches zero, the distribution becomes more and more impulsive. Here, we utilize $b = 0.2$ which corresponds to a rather heavy tailed distribution. We test brain and shoulder MR images with size 256×256 and a pixel intensity range of $[0, 1]$. In the transform learning phase, we utilize 7200 maximally overlapping patches where each patch has size 6×6. Both algoritms are realized with common parameters $\lambda = 10^5$, $\beta = 0.02$, $\tau = 0.5$, and inner iteration numbers $N_1 = N_2 = 10$. For LDA-TLMRI algorithm, $\gamma = \tau' = 0.01$ and $N_3 = 5$. Our main performance metric is high-frequency error norm (HFEN) which is particularly useful for its relation with the human visual quality perception [5]. In order to calculate HFEN, the difference between the original and reconstructed images is filtered by rotationally symmetric Laplacian of Gaussian filter with size 15×15 and standart deviation 1.5. Afterwards, the Frobenius norm of the filter output is calculated. Assuming $\sigma = 50$, Figs. 1 and 2 show the original and reconstructed images for the brain and shoulder cases, respectively. Zero-filling has the worst reconstruction performance as expected. On the other hand, LAD-TLMRI provides better reconstruction when compared with original TLMRI. Fig. 3 detailing the evolution of HFEN versus iteration supports this observation. Tables 1 and 2 present HFEN values for reconstructed images with different σ values. It can be further stated that the HFEN enhancement obtained via LAD-TLMRI improves as σ increases.

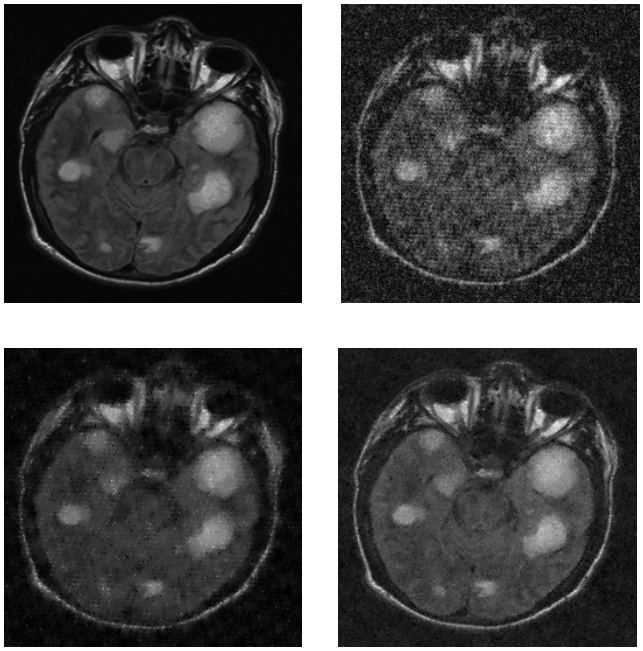

Fig. 1: Brain image results. First row: Original image (left), Zero-filling reconstruction (right). Second row: TLMRI reconstruction (left), LAD-TLMRI reconstruction (right).

Table 1: HFEN for brain image.

Reconstruction	Zero-filling	TLMRI	LAD-TLMRI
$\sigma = 20$	2.17	1.64	0.99
$\sigma = 50$	3.46	2.56	1.39
$\sigma = 100$	4.98	3.43	1.96

5. CONCLUSIONS

We have developed an MRI reconstruction algorithm with an impulsive or non-Gaussian observation noise prior. We con-

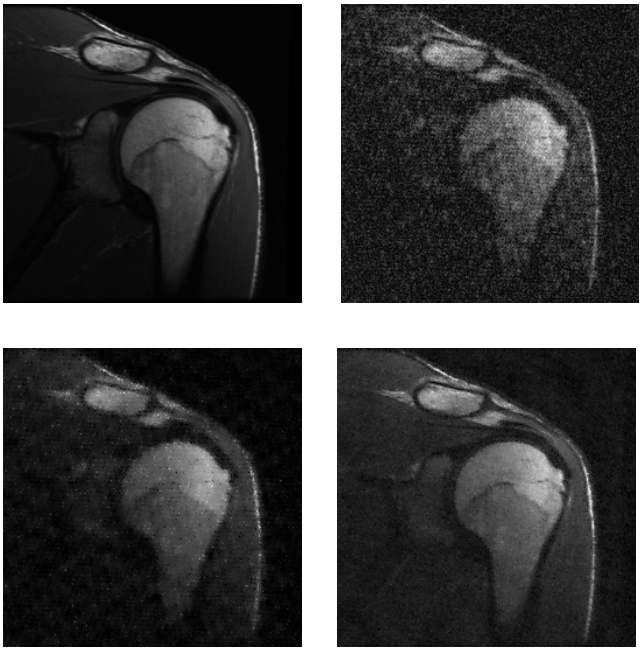

Fig. 2: Shoulder image results. First row: Original image (left), Zero-filling reconstruction (right). Second row: TLMRI reconstruction (left), LAD-TLMRI reconstruction (right).

Table 2: HFEN for shoulder image.

Reconstruction	Zero-filling	TLMRI	LAD-TLMRI
$\sigma = 20$	1.73	1.16	0.72
$\sigma = 50$	2.82	1.87	1.05
$\sigma = 100$	4.07	2.60	1.54

join an effective analysis sparsity image model of reduced complexity with a data fidelity term which uses the ℓ_1 norm. The resulting LAD-TLMRI algorithm successfully combines a very recent image model used for regularization with a non-smooth data fidelity term especially suitable for impulsive noise. Under impulsive noise, the developed LAD-TLMRI algorithm demonstrates improved reconstruction performance when compared to the original algorithm using a quadratic data fidelity assumption.

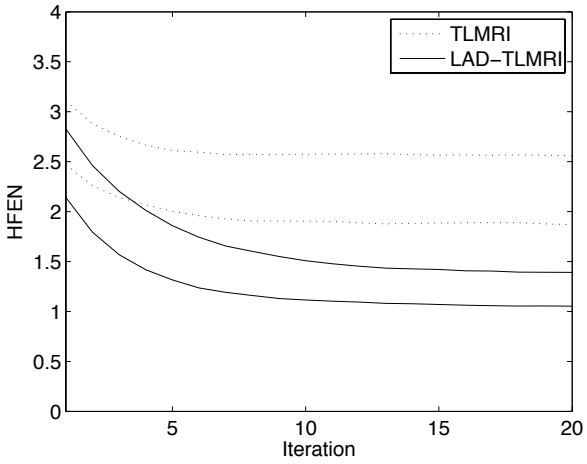

Fig. 3: Convergence of the reconstruction algorithms for brain image (upper curves) and shoulder image (lower curves).

REFERENCES

[1] M. Elad, M.A.T. Figueiredo and Y. Ma, "On the role of sparse and redundant representations in image processing," *Proc. IEEE*, vol. 98, no. 6, pp. 972–982, 2010.

[2] M. Elad and M. Aharon, "Image denoising via sparse and redundant representations over learned dictionaries," *IEEE Trans. Image Process.*, vol. 15, no. 12, pp. 3736–3745, 2006.

[3] S. Ravishankar and Y. Bresler, "MR image reconstruction from highly undersampled k-space data by dictionary learning," *IEEE Trans. Med. Imag.*, vol. 30, pp. 1028–1041, May 2011.

[4] I.W. Selesnick and M.A.T. Figueiredo, "Signal restoration with overcomplete wavelet transforms: comparison of analysis and synthesis priors," *Proc. SPIE*, vol. 7446, pp. 74460D–74460D–15, 2009.

[5] S. Ravishankar and Y. Bresler, "Learning sparsifying transforms," *IEEE Trans. Signal Process.*, vol. 61, no. 5, pp. 1072–1086, 2013.

[6] S. Ravishankar and Y. Bresler, "Sparsifying transform learning for compressed sensing MRI," in *2013 IEEE 10th International Symposium on Biomedical Imaging (ISBI)*, April 2013, pp. 17–20.

[7] A.K. Tanc and E.M. Eksioglu, "Transform learning MRI with global wavelet regularization," in *Signal Processing Conference (EUSIPCO), 2015 23rd European*, Aug 2015, pp. 1855–1859.

[8] M. Nikolova, "Minimizers of cost-functions involving nonsmooth data-fidelity terms. application to the processing of outliers," *SIAM Journal on Numerical Analysis*, vol. 40, no. 3, pp. 965–994, 2002.

[9] X. Jiang, T. Kirubarajan and W.J. Zeng "Robust sparse channel estimation and equalization in impulsive noise using linear programming," *Signal Processing*, vol. 93, no. 5, pp. 1095–1105, 2013.

[10] M. Nardon and P. Pianca "Simulating a generalized Gaussian noise with shape parameter 1/2," in *Mathematical and Statistical Methods in Insurance and Finance*, pp. 173–180, Milano: Springer, 2008.