PARAMETER ESTIMATION FOR HYBRID WAVELET-TOTAL VARIATION REGULARIZATION

Lotfi Chaari[1], *Jean-Christophe Pesquet*[2], *Jean-Yves Tourneret*[3] *and Philippe Ciuciu*[4]

[1] INRIA Rhône-Alpes
655 avenue de l'Europe, Montbonnot
38334 Saint Ismier Cedex, France
lotfi.chaari@inria.fr

[2] Université Paris-Est, IGM and
UMR-CNRS 8049, 77454
Marne-la-Vallée cedex, France
pesquet@univ-mlv.fr

[3] University of Toulouse,
IRIT/ENSEEIHT/TSA
31071 Toulouse, France
jean-yves.tourneret@enseeiht.fr

[4] CEA/DSV/I^{2}BM/Neurospin, CEA
Saclay, Bbt. 145, Point Courrier 156,
91191 Gif-sur-Yvette cedex, France
philippe.ciuciu@cea.fr

ABSTRACT

In many image restoration/reconstruction problems, using redundant linear decompositions also named as *frames* may be fruitful. Moreover, Total Variation (TV) is also widely used in the edge-preserving regularization literature. Associating these two tools in a joint regularization framework may be of great interest since they are somehow complementary. However, estimating the regularization parameters in this case becomes a tricky issue which cannot be performed by using standard estimators. In this work, a hierarchical model is introduced to solve this problem within a fully Bayesian framework. A hybrid MCMC algorithm is subsequently proposed to sample from the derived posterior distribution. We show that this algorithm allows the regularization parameters to be determined accurately. We finally investigate its application to parallel MRI reconstruction, where the use of a joint wavelet-TV regularization is also novel.

Index Terms— Bayesian estimation, regularization, MCMC, parameter estimation, sparsity, wavelets, frame, Total Variation.

1. INTRODUCTION

In image restoration or reconstruction, data representation is crucial. In this respect, many decompositions have been proposed in order to obtain suitable representations in other domains than the original one. Since the 90's, Wavelet Transforms (WT) [1] have been widely used in the regularization literature since they allow one to detect local image features and details through the sparse representation they provide. WTs allow one to decompose signals into insightful scale-space elements which are easier to interpret and process. Moreover, it has been observed in the image processing literature [2, 3] that overcomplete wavelet representations (frames) are generally more advantageous than wavelet bases. On the other hand, Total Variation (TV) [4] has also been widely used in edge-preserving regularization [5]. A major difficulty when using frames is to estimate the regularization parameters. As outlined in [6], since frame synthesis operators are generally not injective, the determination of the frame coefficients is an underdetermined problem even in the case of a perfectly known signal. Moreover, because of the TV properties, this parameter estimation problem becomes more difficult when coupling TV to frame representations (FRs) in a joint regularization framework. In this paper, we propose a hierarchical approach in order to address this problem from a Bayesian viewpoint. This approach can be applied to noisy data when only inaccurate measurements of the signal are available. Our work takes advantage of the developments in Markov Chain Monte Carlo (MCMC) algorithms [7]. In particular, we will consider hybrid MCMC algorithms combining Metropolis-Hastings (MH) and Gibbs moves to sample according to the posterior distribution of interest.

This rest of paper is organized as follows. We first describe the general context of our study in Section 2. The statistical problem and the proposed hierarchical Bayesian method are presented in Section 3. Section 4 is devoted to the application of the proposed algorithm to pMRI reconstruction. Finally, conclusions are drawn in Section 5.

2. GENERAL CONTEXT

2.1. Problem statement

Let $\overline{\boldsymbol{y}} \in \mathbb{R}^{M\times N}$ be the image to be recovered from its observation $\boldsymbol{z} \in \mathbb{R}^{M'\times N'}$ degraded by the additive noise $\boldsymbol{n} \in \mathbb{R}^{M'\times N'}$ and a linear operator $H : \mathbb{R}^{M\times N} \longrightarrow \mathbb{R}^{M'\times N'}$ describing the physical laws linking $\overline{\boldsymbol{y}}$ to the measurement $\boldsymbol{z}$:

$$\boldsymbol{z} = H\overline{\boldsymbol{y}} + \boldsymbol{n}. \tag{1}$$

Our inverse problem will consist of retrieving the image $\overline{\boldsymbol{y}}$ based on the observation $\boldsymbol{z}$ and the knowledge of H. However, due to the noise level or the linear operator H, the problem may be ill-posed. In this case, one generally resorts to regularization techniques [8], which consist of introducing some prior knowledge about the target solution. However, this knowledge may be easier to formulate using some image features either in the original space such as the image TV, or in a transformed space penalizing the frame coefficients. It is well known [2, 3] that WTs provide us with a good performance by preserving edges without introducing over-smoothing effects. However, it is also known that wavelet regularization may introduce some ringing artifacts along the image contours. On the other hand, TV regularization is well adapted to piecewise smooth regions, but may over-smooth image details and introduce staircase effects. These two edge-preserving regularizations are generally used separately. However, in order to benefit from their advantages, they have been combined in some recent works [9, 10].

2.2. Combined wavelet-TV regularization

Let us denote by $F\colon \mathbb{R}^{M\times N} \to \mathbb{R}^{K\times L}$ and $F^*\colon \mathbb{R}^{K\times L} \to \mathbb{R}^{M\times N}$ the wavelet linear operator and its adjoint, respectively. Note that we concentrate here on wavelet representations, but the proposed method can be easily extended to any kind of FR (e.g., Modulated Complex Lapped Transform). Introducing an FR in the observation model in Eq. (1) gives:

$$\boldsymbol{z} = HF^*\boldsymbol{x} + \boldsymbol{n}, \tag{2}$$

where $\overline{\boldsymbol{y}} = F^*\boldsymbol{x}$ and $\boldsymbol{x} \in \mathbb{R}^{K\times L}$ is the frame coefficient vector associated to $\overline{\boldsymbol{y}}$. A general joint Wavelet-Total Variation (W-TV) regularization amounts to recovering the image $\widehat{\boldsymbol{y}} = F^*\widehat{\boldsymbol{x}}$ from the estimated wavelet coefficients $\widehat{\boldsymbol{x}}$ obtained by minimizing the following penalized criterion:

$$\mathcal{J}(\boldsymbol{x}) = D(HF^*\boldsymbol{x}, \boldsymbol{z}) + g(\boldsymbol{x}) + h(F^*\boldsymbol{x}), \tag{3}$$

where $D(HF^*\cdot, \boldsymbol{z})$ is a discrepancy measure between the observed data $\boldsymbol{z}$ and the solution $\boldsymbol{x}$, $g(\cdot)$ and $h(F^*\cdot)$ are two penalty functions. In the case of Gaussian noise, we have for instance $D(HF^*\boldsymbol{x}, \boldsymbol{z}) = \|HF^*\boldsymbol{x} - \boldsymbol{z}\|^2$. After reindexing the wavelet coefficients so that $\boldsymbol{x} = (x_k)_{k\in\mathbb{K}}$ with $\mathbb{K} = \{1, \ldots, KL\}$, we will focus on the case where $g(\boldsymbol{x}) = \sum_{k\in\mathbb{K}} |x_k|^{p_k}/\gamma_k$ (i.e., generalized Gaussian prior [6]) and $h(F^*\boldsymbol{x}) = \lambda\|F^*\boldsymbol{x}\|_{\mathrm{TV}}$ with $\lambda > 0$, $\gamma_k > 0$ and $p_k > 1$ (to ensure the convexity of $\mathcal{J}$), $\|\cdot\|_{\mathrm{TV}}$ being the TV seminorm [4]. Moreover, as explained in [6], the wavelet coefficients can be split into G groups (typically subbands) having similar statistical properties. Consequently, the optimality criterion in Eq. (3) reads:

$$\mathcal{J}(\boldsymbol{x}) = D(HF^*\boldsymbol{x}, \boldsymbol{z}) + \sum_{g\in\mathbb{G}}\sum_{k\in S_g} \frac{|x_k|^{p_g}}{\gamma_g} + \lambda\|F^*\boldsymbol{x}\|_{\mathrm{TV}}, \tag{4}$$

where the summation covers the index set S_g of the gth group containing n_g elements, and $\mathbb{G} = \{1, \ldots, G\}$. The major difficulty here stems from the estimation of the regularization parameters λ, $\boldsymbol{\gamma} = (\gamma_1, ..., \gamma_G)^\intercal$ and $\boldsymbol{p} = (p_1, ..., p_G)^\intercal$. In fact, using *redundant* wavelet representations makes the reference wavelet coefficients not directly observable even if a reference image is available since the wavelet operator is not bijective. Moreover, the TV penalty makes the use of standard estimators such as the Maximum Likelihood (ML) estimator very complicated regarding to the inherent maximization problem. This difficulty has been outlined in [11] where a hybrid regularization is applied combining an ℓ_1 sparsity term and a TV penalization. This problem is addressed within a Bayesian viewpoint by proposing an algorithm based on a hierarchical model which allows us to estimate these parameters when a redundant FR is used in combination with a TV penalization. Note that when only a TV regularization term is considered, some automatic methods may be used (e.g., SURE, L-curve) to estimate the TV regularization parameter [12]. However, the large number of parameters in hybrid regularization makes the extension of these methods complicated.

3. REGULARIZATION PARAMETER ESTIMATION

3.1. Hierarchical Bayesian Model

In a probabilistic setting, $\boldsymbol{z}$ and $\boldsymbol{x}$ are assumed to be realizations of random vectors $\boldsymbol{Z}$ and $\boldsymbol{X}$. Our goal will be to characterize the posterior probability distribution of $\boldsymbol{X}$ given $\boldsymbol{Z}$, by considering some parametric probabilistic model and by estimating the associated hyperparameters. In this context, we will denote by $\boldsymbol{\Theta}$ the random variable associated with the hyperparameter vector $\boldsymbol{\theta}$ whose prior probability density function (pdf) is denoted by $f(\boldsymbol{\theta})$. We also assume that a reference image is available. In this context, the linear operator H in Eq. (2) reduces to the identity, and the error $\boldsymbol{n}$ is modeled by imposing that $\boldsymbol{x}$ belongs to the closed convex set $C_\delta = \{\boldsymbol{x} \in \mathbb{R}^{K\times L} \mid \|\boldsymbol{z} - F^*\boldsymbol{x}\| \le \delta\}$ where $\delta > 0$ generally takes small values since it reflects the noise level (see [6] for motivations). The conditional pdf $f(\boldsymbol{z}|\boldsymbol{x})$ is thus assumed to be a *uniform* distribution on the closed convex set $D_\delta = \{\boldsymbol{z} \in \mathbb{R}^{M\times N} \mid \|\boldsymbol{z} - F^*\boldsymbol{x}\| \le \delta\}$. Using the hierarchical structure between $\boldsymbol{Z}, \boldsymbol{X}$ and $\boldsymbol{\Theta}$, the conditional distribution of $(\boldsymbol{X}, \boldsymbol{\Theta})$ given $\boldsymbol{Z}$ can be written as

$$f(\boldsymbol{x}, \boldsymbol{\theta}|\boldsymbol{z}) \propto f(\boldsymbol{z}|\boldsymbol{x})f(\boldsymbol{x}|\boldsymbol{\theta})f(\boldsymbol{\theta}), \tag{5}$$

where $\propto$ means "proportional to". The following prior distribution will be assigned to the frame coefficients of the image assumed to be:

$$\begin{aligned} f(\boldsymbol{x}|\boldsymbol{\theta}) &= \frac{e^{-(\lambda\|F^*\boldsymbol{x}\|_{\mathrm{TV}} + \sum_{k\in\mathbb{K}} \frac{|x_k|^{p_k}}{\gamma_k})}}{C(\boldsymbol{\theta})} \\ &= \frac{e^{-\lambda\|F^*\boldsymbol{x}\|_{\mathrm{TV}}}}{C'(\boldsymbol{\theta})} \prod_{g\in\mathbb{G}} \frac{e^{\frac{-\sum_{k\in S_g}|x_k|^{p_g}}{\gamma_g}}}{\gamma_g^{n_g/p_g}}, \end{aligned} \tag{6}$$

where $\boldsymbol{\theta} = (\lambda, \boldsymbol{\gamma}, \boldsymbol{p})$ and $C(\boldsymbol{\theta}) = C'(\boldsymbol{\theta})\prod_{k\in\mathbb{K}} \gamma_k^{1/p_k} = C'(\boldsymbol{\theta})\prod_{g\in\mathbb{G}} \gamma_g^{n_g/p_g}$ is a normalization constant which ensures that $f(\boldsymbol{x}|\boldsymbol{\theta})$ is a density and leads to a tractable expression of the posterior distribution. The hierarchical Bayesian model for the FR is completed by the following hyperprior

$$f(\boldsymbol{\theta}) \propto C'(\boldsymbol{\theta})1_{[0,10]}(\lambda)\prod_{g\in\mathbb{G}}\left[\frac{1}{\gamma_g}1_{\mathbb{R}^+}(\gamma_g)1_{[0,3]}(p_g)\right], \tag{7}$$

which is improper due to the hyperprior on γ_g, and where for a set A, $1_A(\xi) = 1$ if $\xi \in A$ and 0 otherwise. This hyperprior is of practical interest since Jeffrey's prior on γ_g is noninformative, and the intervals $[0, 3]$ and $[0, 10]$ cover all possible values of p_g and λ usually encountered in practice. The introduction of the normalization constant $C'(\boldsymbol{\theta})$ in the hyperprior simplifies the calculation of the posterior in Eq. (5). Bayesian estimators such as the maximum a posteriori (MAP) or the minimum mean square error (MMSE) associated with the posterior distribution in Eq. (5) have no simple closed-form expressions. The solution developed in this paper for estimating the unknown model parameters is to use MCMC methods to draw samples of $\boldsymbol{x}$ and $\boldsymbol{\theta}$ from the posterior pdf in Eq. (5), and compute estimates from these generated samples.

3.2. Sampling strategy

Sampling according to the posterior distribution in Eq. (5) will be performed by using a hybrid Gibbs sampler to iteratively sample according to $f(\boldsymbol{x}|\boldsymbol{\theta}, \boldsymbol{z})$ and $f(\boldsymbol{\theta}|\boldsymbol{x}, \boldsymbol{z})$.

3.2.1. Frame coefficient sampling

Straightforward calculations yield the conditional distribution

$$f(\boldsymbol{x}|\boldsymbol{\theta}, \boldsymbol{z}) \propto 1_{C_\delta}(\boldsymbol{x}) e^{-\lambda \|F^* \boldsymbol{x}\|_{\mathrm{TV}}} \prod_{g \in \mathbb{G}} e^{\frac{-\sum_{k \in S_g} |x_k|^{p_g}}{\gamma_g}}. \quad (8)$$

Sampling directly according to this truncated distribution is not easy to perform because of the TV term and since the adjoint frame operator F^* is usually of large dimension. Akin to [6], we propose to use an MH move exploiting the algebraic properties of the frame operator F. More precisely, we can write $\boldsymbol{x} = \boldsymbol{x}_H + \boldsymbol{x}_{H^\perp}$, where $\boldsymbol{x}_H$ and $\boldsymbol{x}_{H^\perp}$ are realizations of random vectors taking their values in $H = \mathrm{Ran}(F)$ and $H^\perp = [\mathrm{Ran}(F)]^\perp = \mathrm{Null}(F^*)$, respectively. The used proposal distribution allows us to generate samples $\boldsymbol{x}_H \in H$ and $\boldsymbol{x}_{H^\perp} \in H^\perp$ separately (see [6] for more details).

3.2.2. Hyperparameter sampling

Instead of sampling $\boldsymbol{\theta}$ according to $f(\boldsymbol{\theta}|\boldsymbol{x}, \boldsymbol{z})$, we propose to iteratively sample according to $f(p_g|\gamma_g, \lambda, \boldsymbol{x}, \boldsymbol{z})$, $f(\lambda|(\gamma_g, p_g)_{g\in\mathbb{G}}, \boldsymbol{x}, \boldsymbol{z})$ and $f(\gamma_g|p_g, \lambda, \boldsymbol{x}, \boldsymbol{z})$ which, after straightforward calculations, are given by:

- $f(p_g|\gamma_g, \lambda, \boldsymbol{x}, \boldsymbol{z}) \propto e^{-\sum_{k\in S_g} \frac{|x_k|^{p_g}}{\gamma_g}} 1_{[0,3]}(p_g)$,
- $f(\lambda|(\gamma_g, p_g)_{g\in\mathbb{G}}, \boldsymbol{x}, \boldsymbol{z}) \propto e^{-\lambda\|F^*\boldsymbol{x}\|_{\mathrm{TV}}} 1_{[0,10]}(\lambda)$,
- $f(\gamma_g|p_g, \lambda, \boldsymbol{x}, \boldsymbol{z}) \propto \mathcal{IG}\left(\frac{n_g}{p_g}, \sum_{k\in S_g} |x_k|^{p_g}\right)$,

where $\mathcal{IG}(a, b)$ is the inverse gamma distribution with parameters a and b. Sampling according to the pdfs $f(p_g|\gamma_g, \lambda, \boldsymbol{x}, \boldsymbol{z})$ and $f(\lambda|(\gamma_g, p_g)_{g\in\mathbb{G}}, \boldsymbol{x}, \boldsymbol{z})$ is achieved by using two MH moves whose proposals $q(p_g \mid p_g^{(i-1)})$ and $q(\lambda \mid \lambda^{(i-1)})$ are Gaussian distributions truncated on the intervals $[0, 3]$ and $[0, 10]$ with standard deviations $\sigma_{p_g} = 0.05$ and $\sigma_\lambda = 0.01$, respectively (values fixed based on our empirical observations). The resulting sampler is summarized in Algorithm 1.

The validation of the proposed estimation strategy on synthetic data (akin to [6, Section V.A]) is complicated due to the TV term. Instead, the next section considers an application to parallel Magnetic Resonance Imaging reconstruction (pMRI).

Algorithm 1 Proposed MCMC algorithm

Initialize $\boldsymbol{\theta}^{(0)} = \left(\lambda^{(0)}, (\gamma_g^{(0)}, p_g^{(0)})_{g\in\mathbb{G}}\right)$, $\boldsymbol{x}^{(0)} \in C_\delta$ and $i = 1$.

repeat

1) ***Sampling of $\boldsymbol{x}$***:

- Given $\boldsymbol{x}^{(i-1)}$ generate $\boldsymbol{x}_H^{(i)}$ and $\boldsymbol{x}_{H^\perp}^{(i)}$ and set $\widetilde{\boldsymbol{x}}^{(i)} = \boldsymbol{x}_H^{(i)} + \boldsymbol{x}_{H^\perp}^{(i)}$ (see [6] for details).
- Compute the acceptance ratio $r(\widetilde{\boldsymbol{x}}^{(i)}, \boldsymbol{x}^{(i-1)}) = \dfrac{f(\widetilde{\boldsymbol{x}}^{(i)}|\boldsymbol{\theta}^{(i-1)}, \boldsymbol{z})\ q(\boldsymbol{x}^{(i-1)}|\widetilde{\boldsymbol{x}}^{(i)})}{f(\boldsymbol{x}^{(i-1)}|\boldsymbol{\theta}^{(i-1)}, \boldsymbol{z})\ q_\eta(\widetilde{\boldsymbol{x}}^{(i)}|\boldsymbol{x}^{(i-1)})}$ and accept the proposed candidate $\widetilde{\boldsymbol{x}}^{(i)}$ with probability $\min\{1, r(\widetilde{\boldsymbol{x}}^{(i)}, \boldsymbol{x}^{(i-1)})\}$.

2) ***Sampling of $\boldsymbol{\theta}$***:

for $g = 1$ to G **do**

- Generate $\gamma_g^{(i)} \sim \mathcal{IG}\left(\frac{n_g}{p_g^{(i-1)}}, \sum_{k\in S_g} |x_k^{(i)}|^{p_g^{(i-1)}}\right)$.

- Simulate $p_g^{(i)}$ as follows:

- Generate $\widetilde{p}_g^{(i)} \sim q(\cdot \mid p_g^{(i-1)})$
- Compute the ratio $r(\widetilde{p}_g^{(i)}, p_g^{(i-1)}) = \frac{f(\widetilde{p}_g^{(i)}|\gamma_g^{(i)}, \boldsymbol{x}^{(i)}, \boldsymbol{z}) q(p_g^{(i-1)}|\widetilde{p}_g^{(i)})}{f(p_g^{(i-1)}|\gamma_g^{(i)}, \boldsymbol{x}^{(i)}, \boldsymbol{z}) q(\widetilde{p}_g^{(i)}|p_g^{(i-1)})}$ and accept the proposed candidate with the probability $\min\{1, r(\widetilde{p}_g^{(i)}, p_g^{(i-1)})\}$.

end for

Simulate $\lambda^{(i)}$ as follows:

- Generate $\widetilde{\lambda}^{(i)} \sim q(\cdot \mid \lambda^{(i-1)})$.
- Compute the ratio

$r(\widetilde{\lambda}^{(i)}, \lambda^{(i-1)}) = \frac{f(\widetilde{\lambda}^{(i)}|(\gamma_g, p_g)_{g\in\mathbb{G}}, \boldsymbol{x}^{(i)}, \boldsymbol{z}) q(\lambda^{(i-1)}|\widetilde{\lambda}^{(i)})}{f(\lambda^{(i-1)}|(\gamma_g, p_g)_{g\in\mathbb{G}}, \boldsymbol{x}^{(i)}, \boldsymbol{z}) q(\widetilde{\lambda}^{(i)}|\lambda^{(i-1)})}$
and accept the proposed candidate with the probability $\min\{1, r(\widetilde{\lambda}^{(i)}, \lambda^{(i-1)})\}$.

until Convergence

4. APPLICATION TO PMRI RECONSTRUCTION

PMRI [13] is a fast acquisition technique which is particularly useful in functional MRI (fMRI) to improve the spatio-temporal resolution. To this end, N_c receiver coils with complementary spatial sensitivities are employed to acquire N_c MRI signals at the same time. The received signal by a given coil ℓ corresponds to the Fourier transform of the desired 2D field $\overline{\boldsymbol{y}}$ weighted by the corresponding coil sensitivity profile. In pMRI, the frequency domain (i.e., k-space) is sampled along the phase encoding direction at a rate that is R times lower than the Nyquist one. Because of this low sampling rate, registered data suffer from aliasing artifacts in the image domain, which increase with the reduction factor. The challenge here is to unfold the received images by exploiting the complementarity between the sensitivity profiles of the coils, and to reconstruct a non-aliased full Field of View (FoV) image. SENSitivity Encoding (SENSE) [13] was one of the early reconstruction methods operating in the spatial domain. It relies on the observation model in Eq. (1) where the linear operator H corresponds to the sensitivity operator. Since the considered inverse problem is ill-posed due to the observation noise and the ill-conditioning of the sensitivity operator, a regularization is thus necessary to achieve better reconstruction results under severe experimental conditions (high reduction factor, low magnetic field,...). We propose to apply a W-TV regularization in order to combine their advantages, in contrast with the current pMRI regularization literature where they are always used separately. Taking into account that the observed signal is complex-valued, and since the noise is Gaussian, the reconstructed image is obtained by:

$$\widehat{\boldsymbol{y}} = F^{*}[\underset{\boldsymbol{x}\in\mathbb{C}^{K\times L}}{\arg\min} \|HF^{*}\boldsymbol{x}-\boldsymbol{z}\|^2 + \sum_{g\in\mathbb{G}}\sum_{k\in S_g}\Big[\frac{|\mathrm{Re}(x_k)|^{p_g^{Re}}}{\gamma_g^{Re}} + \frac{|\mathrm{Im}(x_k)|^{p_g^{Im}}}{\gamma_g^{Im}}\Big] + \lambda^{Re}\|\mathrm{Re}(F^{*}\boldsymbol{x})\|_{\mathrm{TV}} + \lambda^{Im}\|\mathrm{Im}(F^{*}\boldsymbol{x})\|_{\mathrm{TV}}], \tag{9}$$

where p_g^{Re}, γ_g^{Re} and λ^{Re} (resp., p_g^{Im}, γ_g^{Im} and λ^{Im}) are the regularization parameters to be estimated, and $\mathrm{Re}(x_k)$ (resp. $\mathrm{Im}(x_k)$) denote the real (resp. imaginary) part of the scalar $x_k \in \mathbb{C}$. The regularization parameters for the real and imaginary parts are first estimated using Algorithm 1 based on a SENSE reconstructed image (as a reference assuming that H is the identity operator). Note that the frame coefficients estimated by Algorithm 1 are not used here. These parameters are then injected in the optimality criterion in Eq. (9), which is minimized using the Parallel ProXimal Algorithm (PPXA) [10] in order to obtain the estimated frame coefficients $\widehat{\boldsymbol{x}}$, and thus the image $\widehat{\boldsymbol{y}}$. In the experiments given below, the used FR was the union of two orthonormal bases with Daubechies and shifted Daubechies filter of lengths 4 and 8, respectively. Three resolution levels have been used, which means that $G = 20$ groups of wavelet coefficients are considered. In these experiments, reconstruction results are given for TV, wavelet and the proposed hybrid W-TV regularization using hyperparameters estimated by Algorithm 1. As shown in Fig. 1, our results obtained on anatomical gradient echo images acquired at a 1.5 Tesla magnetic field and a reduction factor $R = 4$ suggest that the combined W-TV approach outperforms the other regularized reconstructions in terms of reconstruction artifacts (see the middle part of the slice). The zoom presented in Fig. 1 (bottom row) clearly shows that the hybrid regularization allows to avoid over-smoothing and irregularities caused by the TV and wavelet regularizations, respectively. Also, the gain in Signal to Noise Ratio (SNR) has to be noticed. In the displayed results, the remaining artifacts have been attenuated by using the variational approach proposed in [14]. Note that the hybrid W-TV regularization allows us to achieve an SNR improvement of 2.59 and 0.33 dB with respect to TV or wavelet regularization, respectively.

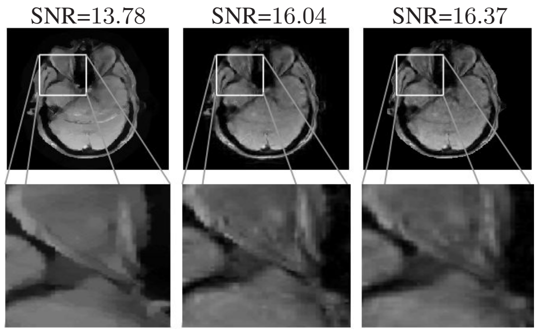

Fig. 1. Reconstructed images using the TV (left), wavelet (middle) and the proposed W-TV regularization method (right): whole images (top row); zoom (bottom row).

5. CONCLUSION

We proposed an MCMC algorithm to estimate the parameters for a hybrid wavelet-TV regularization if a noisy observation of a reference image is available. The problem was addressed from a Bayesian viewpoint and relied on a hierarchical model. Our experiments showed that the proposed algorithm provided accurate parameter estimation, which leads to improved regularization performance for pMRI. Future work will extend the proposed hierarchical Bayesian model to involve the observation linear operator.

6. REFERENCES

[1] S. Mallat, *A wavelet tour of signal processing*, Academic Press, San Diego, CA, USA, 2nd edition, 1998.

[2] R. Coifman and D. Donoho, “Translation-invariant de-noising,” in *Wavelets and Statistics*, Springer Lecture Notes in Statistics 103, pp. 125–150. Springer, New York, 1995.

[3] S. Mallat, “Geometrical grouplets,” *Appl. Comput. Harm. Anal.*, vol. 26, no. 2, pp. 161–180, 2009.

[4] A. Chambolle and P. L. Lions, “Image recovery via total variation minimization and related problems,” *Numerische Mathematik*, vol. 76, no. 2, pp. 167–188, Apr. 1997.

[5] C. Louchet and L. Moisan, “Total variation denoising using posterior expectation,” in *EUSIPCO*, Lausanne, Aug. 25-29, 2008.

[6] L. Chaari, J.-C. Pesquet, J.-Y. Tourneret, Ph. Ciuciu, and A. Benazza-Benyahia, “A hierarchical Bayesian model for frame representation,” *IEEE Trans. on Signal Process.*, vol. 18, no. 11, pp. 5560–5571, 2010.

[7] C. Robert and G. Castella, *Monte Carlo statistical methods*, Springer, New York, 2004.

[8] A. Tikhonov, “Tikhonov regularization of incorrectly posed problems,” *Soviet Mathematics Doklady*, vol. 4, pp. 1624–1627, 1963.

[9] F. C. Tony, S. Jianhong, and Z. Hao-Min, “Total variation wavelet inpainting,” *Journal of Mathematical Imaging and Vision*, vol. 25, no. 1, pp. 107–125, Jul. 2006.

[10] P. L. Combettes and J.-C. Pesquet, “A proximal decomposition method for solving convex variational inverse problems,” *Inverse Problems*, vol. 24, no. 6, Dec. 2008, 27 p.

[11] N. Pustelnik, C. Chaux, and J.-C Pesquet, “Parallel proximal algorithm for image restoration using hybrid regularization,” *arXiv:0911.1536v1*, 2010, Submitted.

[12] S. Ramani, T. Blu, and M. Unser, “Monte-carlo sure: A black-box optimization of regularization parameters for general denoising algorithms,” *IEEE Trans. on Image Process.*, vol. 17, no. 9, pp. 1540 – 1554, Sep. 2008.

[13] K. P. Pruessmann, M. Weiger, M. B. Scheidegger, and P. Boesiger, “SENSE: sensitivity encoding for fast MRI,” *Magnetic Resonance in Medicine*, vol. 42, no. 5, pp. 952–962, Jul. 1999.

[14] L. Chaari, J.-C. Pesquet, A. Benazza-Benyahia, and Ph. Ciuciu, “A wavelet-based regularized reconstruction algorithm for SENSE parallel MRI with applications to neuroimaging,” *Medical Image Analysis*, vol. 15, no. 2, pp. 185–201, Nov. 2011.