A sparse and prior based method for 3D image denoising

Juan F P J Abascal, S. Si-Mohamed, Philippe Douek, C. Chappard, Françoise Peyrin

A sparse and prior based method for 3D image denoising

1st J. F. P. J. Abascal
Univ Lyon, INSA-Lyon, Universit Claude Bernard Lyon 1, UJM-Saint Etienne, CNRS
Lyon, France
juan.abascal@creatis.insa-lyon.fr

2nd S. Si-Mohamed
Univ Lyon, INSA-Lyon, Universit Claude Bernard Lyon 1, UJM-Saint Etienne, CNRS
Lyon, France
Salim.Si-Mohamed@creatis.insa-lyon.fr

3rd P. Douek
Univ Lyon, INSA-Lyon, Universit Claude Bernard Lyon 1, UJM-Saint Etienne, CNRS
Lyon, France
philippe.douek@creatis.insa-lyon.fr

4th C. Chappard
B2OA, UMR CNRS 7052
University Paris Diderot, Lyon, France
christine.chappard@inserm.fr

5th F. Peyrin
Univ Lyon, INSA-Lyon, Universit Claude Bernard Lyon 1, UJM-Saint Etienne, CNRS, Lyon, France
peyrin@esrf.fr

***Abstract*—Denoising algorithms via sparse representation are among the state-of-the art for 2D image restoration. In this work, we propose a novel sparse and prior-based algorithm for 3D image denoising (SPADE). SPADE is a modification of total variation (TV) problem with an additional functional that promotes sparsity with respect to a prior image. The prior is obtained from the noisy image by combining information from neighbor slices. The functional is minimized using the split Bregman method, which leads to an efficient method for large scale 3D denoising, with computational cost given by three FFT per iteration. SPADE is compared to TV and dictionary learning on the Shepp-Logan phantom and on human knee data acquired on a spectral computerized tomography scanner. SPADE converges in approximately ten iterations and provides comparable or better results than the other methods. In addition, the exploitation of the prior image avoids the patchy, cartoon-like images provided by TV and provides a more natural texture.**

***Index Terms*—image denoising, total variation, split Bregman, spectral CT**

I. Introduction

Sparse methods are among the state-of-the art denoising methods. Total variation (TV) preserves edges and was termed as a "proper norm" for images [1]. Other L1-norm methods exploited sparsity using wavelet transform or frames and also provided competitive results [2]. However, TV leads to the so called "staircase artefacts" when images are highly corrupted by noise. To mitigate these effects generalizations of TV, such as total generalized variation [3] and nonlocal TV [4], were proposed. Similarly to nonlocal TV, patch-based [5] and dictionary learning (DL) [6] methods exploit redundancy across patches at different locations on the image. This adds an extra dimension along which images have a sparser representation.

Exploiting sparsity in a high dimensional space is common in compressed sensing applications where data redundancy is present along temporal or other higher dimensions [7], [8], [9]. In these applications, prior-based methods exploit sparsity by subtracting the unknown image to a prior image that is built from data in the high dimensional space. When restricted to 2D or 3D, the best is to use patch-based or similar strategies.

The goal of this work is to present a new paradigm to exploit sparsity in 3D and, in specific, to propose a novel sparse and prior-based algorithm for 3D image denoising (SPADE), with the additional benefit of being computationally efficient for application to experimental clinical images. SPADE minimizes both a TV functional and an additional functional that promotes sparsity with respect to a prior image. The prior is obtained from the noisy image by combining information from neighbor slices. In this study, we used a running average to provide a specific prior image for each slice. The functional is minimized using the split Bregman method, which is an efficient method for L1-norm problems [10]. SPADE is compared to TV and to DL method. The comparison with TV shows the improvement of adding the proposed prior-based regularization functional to a given algorithm. Although the DL method is a 2D-denoising algorithm, it allows to compare the recovery of image texture with a more complex algorithm than TV. Methods are assessed on Shepp-Logan phantom and on human knee experimental data acquired on a spectral computerized tomography (SCT) scanner. SCT is a new imaging modality with the potential of improving conventional CT but reconstructed images are affected by noise [11], [12], [13]. In particular, we are interested in using SCT for knee osteoarthritis, as SCT may provide sufficient soft-tissue contrast to visualize and quantify cartilage tissue. Quantification requires image segmentation and improvement in image quality is thus essential.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Sklodowska-Curie grant agreement N° 701915. It was also performed within the framework of the LabEx PRIMES (ANR-11-LABX-0063) of University de Lyon and under the support of the ANR project SALTO (ANR-17-CE19-0011-01). The project has also received funding from the European Union's Horizon 2020 research and innovation programme under the grant agreement N 668142.

II. METHODS

A. TV denoising

Let $\tilde{u}$ be the 3D noisy image and u the 3D denoised image, the constrained total variation denoising problem is posed as follows:

$$\min_u \|\nabla u\|_1 \qquad \text{such that} \qquad \|u-\tilde{u}\|_2^2 \leq \sigma^2, \tag{1}$$

where the first term is 3D TV functional and the second term constrains the denoised image to lie close to the noisy image. We assume that noise follows a Gaussian distribution.

The TV functional is nonlinear and nondifferentiable, so it requires stable and efficient iterative algorithms. It is common to approximate TV functional but this can lead to slow convergence or non optimal results [14]. The split Bregman method provides a framework for efficient minimization of convex nondifferentiable functional [10] (see section II-C).

B. SPADE denoising

SPADE imposes both spatial regularization and sparsity with respect to a prior image by solving the following problem:

$$\begin{aligned}\min_u (1-\alpha)\|\Psi u\|_1 + \alpha\|\Phi(u-u_p)\|_1 \\ \text{such that} \qquad \|u-\tilde{u}\|_2^2 \leq \sigma^2,\end{aligned} \tag{2}$$

where u_p is a prior image that is assumed to be known. Operators Ψ and Φ are chosen to be equal to ∇ in order to compare to TV denoising and for computational purposes, as described in section II-C. With this choice, SPADE minimizes a combination of $TV(u)$ and TV$(u-u_p)$.

The prior image should be a good estimate of the unknown image in most but not in all pixels and should have low noise. In this study, the prior is obtained by combining information from neighbor slices by using a running average to provide a specific prior image for each slice, that is, $u_p = (u_p^1, \ldots, u_p^N)$, where u_p^i represents the ith slice and N is the number of slices along the z-dimension in the image. Each prior slice u_p^i is built from the 3D noisy image $\tilde{u}$ as follows

$$u_p^i = \sum_j \beta_{ij}\tilde{u}_j, \tag{3}$$

where $\beta_{ij} = \exp(-\gamma|j-i|)$, $j = 1, \ldots, N$ is the index of the slice, and $\gamma = 0.04$, which was set empirically in order to provide slow decaying coefficients. The prior image, which is a low-pass filtered version of the noisy image, has low noise but it is obviously blurred.

This methodology provides an easy way to exploit structure similarity across slices. The further away from the current slice the lower the similarity. Thus, we propose a stack based denoising approach where a given number of slices are denoised at each time. We used stack of ten slices ($N = 10$).

C. Split Bregman formulation to solve SPADE and TV problems

We implemented SPADE and TV methods using the Split Bregman method, which is an efficient method for L1-regularization problems [10], [15], [7]. The nondifferentiability of TV is handled by introducing 3D auxiliary variables d_i and e_i, $i = x, y, z$, such that problem (2) becomes

$$\begin{aligned}\arg\min_{u,d_i,e_i} (1-\alpha)\|(e_x,e_y,e_z)\|_1 + \alpha\|(d_x,d_y,d_z)\|_1 \\ \text{such that } \|u-\tilde{u}\|_2^2 \leq \sigma^2, d_i = \nabla_i(u-u_p), \\ e_i = \nabla_i u, i = x,y,z,\end{aligned} \tag{4}$$

adopting isotropic TV [10]. Constrained problem (4) is easily managed using an iterative regularization scheme where constraints are imposed by adding Bregman iterations b_i, c_i:

$$\begin{aligned}(u^{k+1}, d_i^{k+1}, e_i^{k+1}) = \arg\min_{u,d_i,e_i} \alpha\|(d_x,d_y,d_z)\|_1 + \\ (1-\alpha)\|(e_x,e_y,e_z)\|_1 + \frac{\mu}{2}\|u-\tilde{u}^k\|_2^2 + \sum_i \frac{\lambda}{2}\|d_i - \\ \nabla_i(u-u_p) - b_i\|_2^2 + \sum_i \frac{\lambda}{2}\|e_i - \nabla_i u - c_i\|_2^2,\end{aligned} \tag{5a}$$

$$\tilde{u}^{k+1} = \tilde{u}^k + \tilde{u} - u^{k+1}, \tag{5b}$$

$$b_i^{k+1} = b_i^k + \nabla_i(u^{k+1} - u_p) - d_i^{k+1} \tag{5c}$$

$$c_i^{k+1} = c_i^k + \nabla_i u^{k+1} - e_i^{k+1} \tag{5d}$$

Variables u, d_i, and e_i are decoupled so they can be solved in alternating steps. Solution of u that leads to a linear system is efficiently solved in the Fourier domain and auxiliary variables are solved using shrinkage formulas. This leads to an efficient method for large scale 3D denoising, with computational cost given by three FFT per iteration. In addition, it can be parallelized across stacks, which provides additional computational efficiency. We remark that the linear system can be solved in the Fourier domain as $\Phi'\Phi$ and $\Psi'\Psi$ have Fourier representation. Other choices for Φ or Ψ that meet this criterion are unitary operators such as the wavelet transform. A previous work proposed a Gauss-Seidel method for efficient resolution of the linear system for a banded Hessian matrix that arises in a Bregmanized nonlocal TV formulation [4].

For methods hyperparameters, we used $\mu = \lambda = 1$ for both TV and SPADE. This parameter selection does not affect convergence, only convergence speed (see [10], [15], [16] for further details). The effect of α is analyzed in section III-A.

D. Dictionary learning denoising

For the DL method [6] we used the KSVD toolbox downloaded from http://www.cs.technion.ac.il/ elad/software/ with the following parameters: block size $b = 8$, redundancy factor $R = 4$, number of atoms in the dictionary Rb^2, estimate of noise SD σ and ten KSVD iterations. Number of iterations could be made smaller to accelerate convergence but did not affect results. We checked varying parameter σ in the range from 5 to 20 and present best results. We remark that while TV and SPADE are 3D-denoising methods, the DL code used is a 2D-denoising method.

E. Methods analysis and image evaluation

The proposed method SPADE (2) is a modification of TV, so convergence and solution error at different iteration numbers

are shown for both methods. Then, SPADE, TV and DL methods are compared in terms of mean-squared error (MSE) with respect to the ground truth image, streak artefact measure (SAM), and noise. SAM is the total variation of the denoised image subtracted to the ground truth image, so it provides an estimate of how well edges are recovered [7]. SAM accounts for noise, artefacts and texture recovery. Noise is computed as the standard deviation (SD) on a small homogeneous region. As DL is a 2D method, we also provide metrics for TV method in 2D (TV-2D), for comparison.

F. Data sets

Methods were tested on a numerical phantom and experimental data (figure 1). The phantom was built from the Shepp-Logan phantom where a slice was replicated along z-dimension and 5% additive Gaussian noise was added. Experimental data consisted of a excised human knee image acquired at Philips prototype SCT scanner, at CERMEP, Lyon [17]. Knee sample was taken from the Institut d'Anatomie Paris Descartes and provided by B2OA (Bioingénierie et Bioimagerie ostéo-articulaire) CNRS 7052, Paris Diderot University. The study was approved by the ethics committee of Descartes University, Paris. Two levels of noise were considered: the original image and a second noisier data set with 5% additive Gaussian noise. For the latter the image provided by the scanner was adopted as ground truth. Images were 640×640 pixels and we worked with one stack of 10 slices.

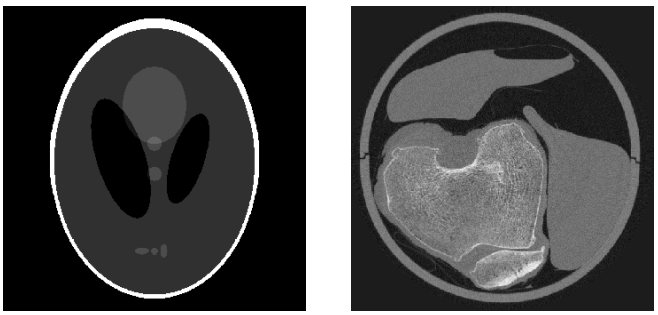


Fig. 1. Left: Slice of Shepp-Logan phantom. Right: Slice of experimental SCT image of human knee.

III. Results

A. Shepp-Logan phantom

SPADE is first analyzed in comparison to TV method on the Shepp-Logan phantom. Figure 2 shows solution error norm versus the number of iterations for TV and for SPADE for different values of α. Both methods converged in less than 10 iterations. SPADE presented lowest errors and more stable convergence for intermediate values of α ($0.2 \leq \alpha \leq 0.4$); solution error was 34% lower than TV error. Very low or large values of α resulted more unstable. We emphasize that TV becomes unstable after few iterations because of high noise. On the other hand, SPADE had smoother convergence for a wide range of values of α. From now onwards we use $\alpha = 0.4$.

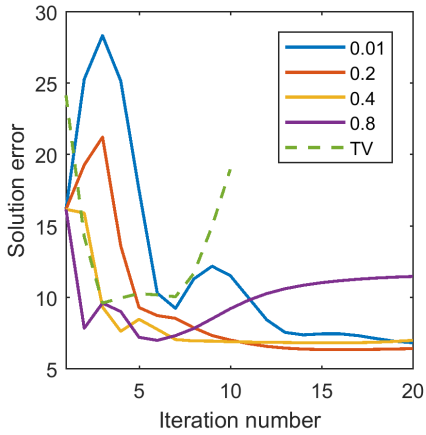


Fig. 2. Solution error norm versus the iteration number for TV and SPADE for different values of α.

TABLE I
METRICS FOR DENOISING OF SHEPP-LOGAN PHANTOM (LEFT) AND EXPERIMENTAL KNEE DATA (RIGHT).

	MSE %	SAM	SD	MSE %	SAM	SD
TV	6.5	4.1	0.010	6.7	2.7	45.2
SPADE	4.3	3.8	0.011	6.4	2.7	57.4
Prior	4.4	5.8	0.016	8.3	3.3	68.7
DL	4.8	3.2	0.008	6.8	2.7	48.0
TV-2D	8.9	6.3	0.017	7.6	3.0	57.0

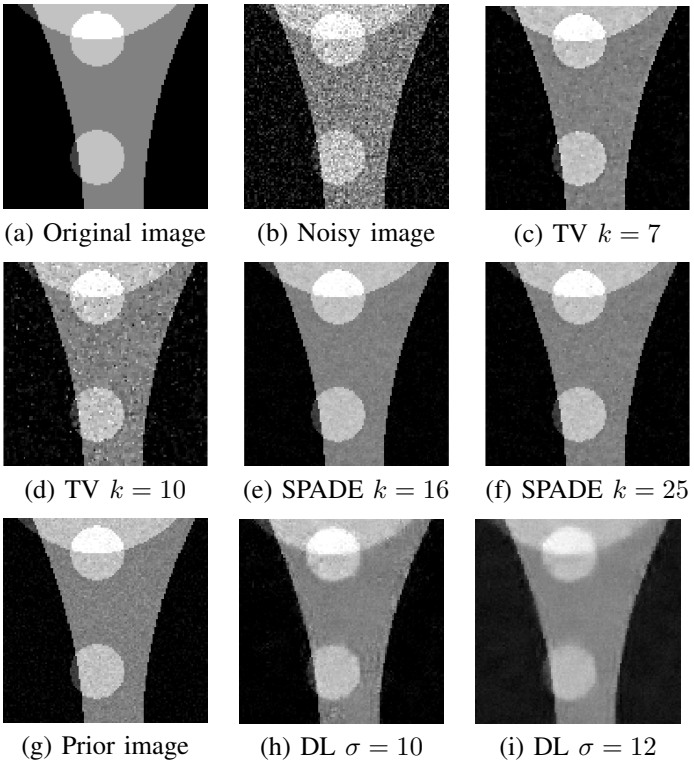
Figure 3 shows denoised images by TV and SPADE for different number of Bregman iterations k and the prior image used by SPADE. TV presented the lowest MSE for $k = 3$ but edges were not well recovered. Higher number of iterations allowed to recover edges at the expense of noise and "staircase artefacts". SPADE had the lowest MSE for $k = 16$, for which most edges were recovered. Increasing the iteration number recovered edges without detrimental noise.

DL reduced noise considerably but it did not recover edges as efficiently as SPADE and presented less natural texture. Best results for DL were obtained for $10 \leq \sigma \leq 12$, where 10 provided sharper objects but also more artefacts, and 12 led to larger noise reduction but blurred edges. Using the simulated noise SD, ($\sigma = 12.75$ for gray-level scale images), would have led to a excessively blurred image.

Table I (left) presents MSE, SAM and noise for TV ($k = 3$), SPADE ($k = 16$), prior image used by SPADE, TV-2D ($k = 3$) and DL (10 iterations, $\sigma = 12$). SPADE led to the lowest MSE among all methods and DL led to the lowest noise reduction and SAM. SPADE improved all metrics with respect to TV except to noise, which was similar.

B. Experimental data

Figure 4 shows denoised images for the knee experimental data set with higher noise (5% additive Gaussian noise). TV and DL were quite efficient in removing noise but they presented less natural image texture than SPADE, which



(a) Original image (b) Noisy image (c) TV $k = 7$

(d) TV $k = 10$ (e) SPADE $k = 16$ (f) SPADE $k = 25$

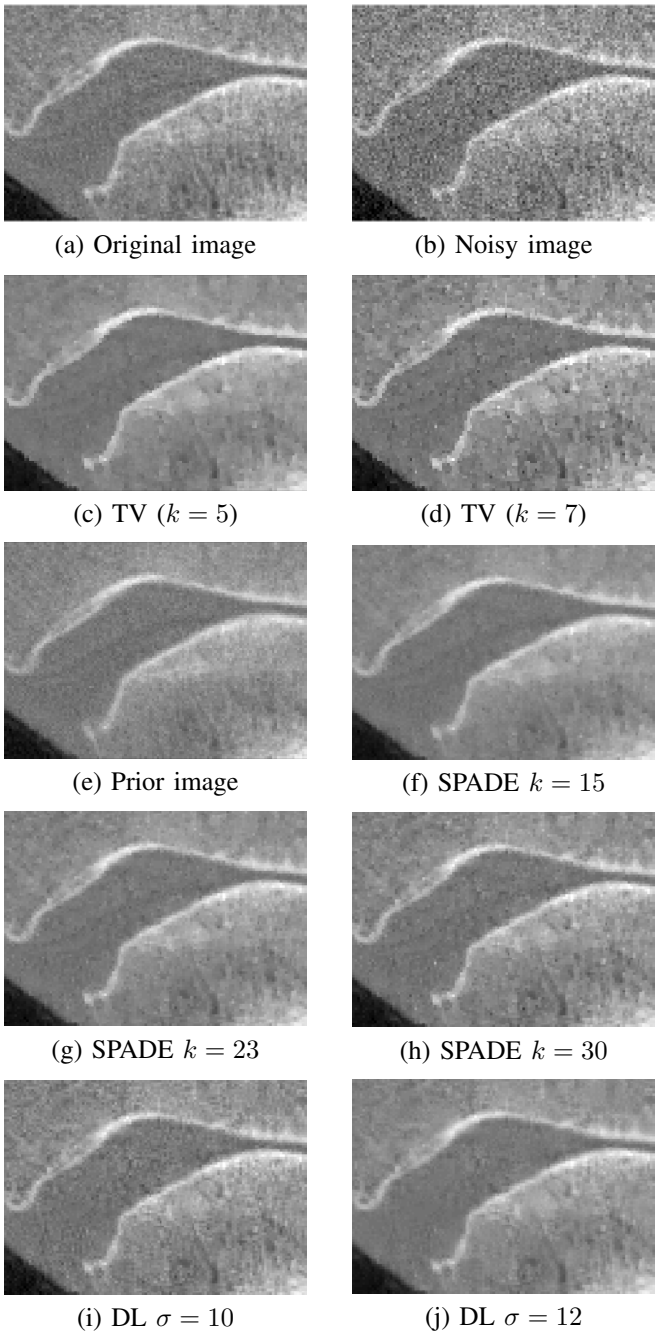
(g) Prior image (h) DL $\sigma = 10$ (i) DL $\sigma = 12$

Fig. 3. a) Detail of Shepp-Logan image, b) noisy image with 5% Gaussian noise, and denoised images by c-d) TV, e-f) SPADE for $\alpha = 0.4$, and h-i) DL. g) Prior image used by SPADE.

recovery is guided by the prior image. Denoised images at different iteration numbers for TV and SPADE show that SPADE presented more robust convergence than TV, where solutions with lowest MSE for TV and SPADE were for $k = 5$ and $k = 23$, respectively. For $k = 15$, SPADE has suppressed noise and recovered most features, and for $k = 23$ it recovered all bone details, which can be appreciated on the bottom part of the image. A large iteration number ($k = 30$) led to moderate increase of noise in comparison with TV and without TV-like artefacts. DL presented lowest error for $\sigma = 12$; it provided better image quality than TV but it contained a large number of artefacts. Overall, SPADE presented best trade-off, being quite effective in removing noise while maintaining good image quality and a natural texture.

Table I (right) shows MSE, SAM and noise metrics for knee data. SPADE led to lower reduction in MSE than other methods. TV and DL provided the largest noise reduction but all methods presented similar SAM, as it accounts for both noise and artefacts.

Codes were run on a Windows computer with a 64-bit operating system, Intel Xeon(R) E5-1650 v4 3.60 GHz CPU and 128 GB RAM. We provide in table II algorithm computation times with codes ran in MATLAB. We also give the average time per iteration as most codes can be tuned to achieve faster convergence. Time for denoising the whole volume was computed only for SPADE; times for other methods were



(a) Original image (b) Noisy image

(c) TV ($k = 5$) (d) TV ($k = 7$)

(e) Prior image (f) SPADE $k = 15$

(g) SPADE $k = 23$ (h) SPADE $k = 30$

(i) DL $\sigma = 10$ (j) DL $\sigma = 12$

Fig. 4. Detail of a) original image, b) noisy image with 5% Gaussian noise, and denoised images by c-d) TV, f-h) SPADE and i-j) DL. e) Prior image used by SPADE.

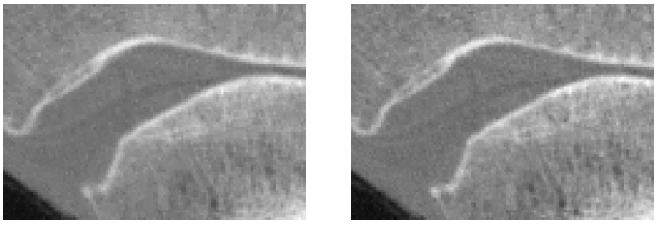
extrapolated. Computation times for SPADE can be further reduced by parallelization and exploiting GPU, as heaviest computation time is the FFT.

Figure 5 shows denoising of the original image provided by the spectral CT scanner (figure 4-a), which has lower noise than previous data set (figure 4-b). SPADE improved image quality both in terms of reduced noise and increased tissue contrast. Cartilage tissue is clearly seen and better

TABLE II
COMPUTATION TIME FOR THE KNEE DATA.

	Iteration (s)	Slice (s)	Volume (min)
TV	0.05	0.4	2
SPADE	0.08	1.1	5.2
DL	2.4	29	192

delineated than in the original image, which is the goal for the medical application. Images are shown at two different iteration numbers. We can see that SPADE already provides high quality image after 10 iterations and larger number of iterations are needed to recover all fine details.



(a) $k = 10$ (b) $k = 50$

Fig. 5. Denoising on experimental knee image (figure 4-a) by SPADE for different iteration number.

IV. DISCUSSION AND CONCLUSION

We have presented a new methodology for 3D image denoising, in specific, we proposed SPADE method and compared it to TV and DL methods. Results show that adding a prior-based functional avoids artefacts that appeared in TV and provides a more natural image texture than TV and DL methods. We note that while the prior image is not actually set a priori, as it is learned from the data, it guides the denoising algorithm to reduce noise and maintain texture. In addition, the prior makes SPADE to be more robust to the iteration number than TV method, where both methods were solved using the split Bregman formulation [7], [10], [18].

SPADE also provided competitive results in terms of computation time, allowing to denoise the whole volume in few minutes. We found that a higher iteration number led to improved results for the less noisy data. We remark that SPADE can further benefit from GPU implementation and that the possibility of processing the 3D volume in stacks allows for further parallelization.

This work is subject to few limitations. The comparison of SPADE, which is intrinsically a 3D-denoising method, to DL is not fair, as DL is a 2D-denoising method. However, we showed the improvement of adding a prior-based regularization functional to TV method, so we could expect similar improvement by applying this approach to DL or other methods such as total generalized variation. In addition, the use of DL for 3D images should be time consuming which could be a severe restriction for clinical applications. With regard to the Shepp-Logan phantom, it was built from a 2D slice, which makes it an ideal situation for the prior image as it is less blurred than in practice. On the contrary, results on knee data show the feasibility of the proposed methodology on clinical data.

Several generalizations of the proposed method can be designed. For instance, one may think of more general transforms that combine neighbor slices.

In conclusion, we have presented a new paradigm for 3D image denoising that allows to exploit sparsity in a easy and efficient manner. From the results shown on experimental data, we expect SPADE to be widely used for 3D image processing.

REFERENCES

[1] L. I. Rudin, S. Osher *et al.*, "Nonlinear total variation based noise removal algorithms," *Physica D*, vol. 60, no. 1-4, pp. 259–268, 1992.

[2] J.-F. Cai, S. Osher *et al.*, "Split Bregman Methods and Frame Based Image Restoration," *Multiscale Model. Simul.*, no. 2, pp. 337–369.

[3] K. Bredies, K. Kunisch, and T. Pock, "Total Generalized Variation," *SIAM J. Imaging. Sci.*, vol. 3, no. 3, pp. 492–526, 2010.

[4] X. Zhang, M. Burger *et al.*, "Bregmanized Nonlocal Regularization for Deconvolution and Sparse Reconstruction," *SIAM J. Imaging. Sci.*, vol. 3, no. 3, pp. 253–276, 2010.

[5] H. Schaeffer and S. Osher, "A Low Patch-Rank Interpretation of Texture," *SIAM J. Imaging. Sci.*, vol. 6, no. 1, pp. 226–262, 2013.

[6] M. Aharon, M. Elad, and A. Bruckstein, "K-SVD: An algorithm for designing overcomplete dictionaries for sparse representation," *IEEE Trans. Signal Process.*, vol. 54, no. 11, pp. 4311–4322, 2006.

[7] J. Abascal, M. Abella *et al.*, "A novel prior- and motion-based compressed sensing method for small-animal respiratory gated CT," *PLoS One*, vol. 11, no. 3, 2016.

[8] G.-H. Chen, J. Tang, and S. Leng, "Prior image constrained compressed sensing (PICCS): A method to accurately reconstruct dynamic CT images from highly undersampled projection data sets," *Med. Phys.*, vol. 35, no. 2, pp. 660–663, 2008.

[9] J. F. P. J. Abascal, M. Desco, and J. Parra-Robles, "Incorporation of prior knowledge of the signal behavior into the reconstruction to accelerate the acquisition of MR diffusion data," *IEEE Trans. Med. Imaging*, pp. 547–556, 2017.

[10] T. Goldstein and S. Osher, "The Split Bregman Method for L1-Regularized Problems," *SIAM J. Imaging. Sci.*, vol. 2, no. 2, pp. 323–343, 2009.

[11] B. J. Heismann, B. T. Schmidt, and T. Flohr, *Spectral computed tomography*. SPIE, 2012.

[12] N. Ducros, J. F. P.-J. Abascal *et al.*, "Regularization of nonlinear decomposition of spectral x-ray projection images," *Med. Phys.*, vol. 44, no. 9, pp. e174–e187, 2017.

[13] J. F. P. J. Abascal, N. Ducros, and F. Peyrin, "Nonlinear material decomposition using a regularized iterative scheme based on the Bregman distance," *Inverse Probl.*, vol. 34, no. 12, p. 124003, 2018.

[14] M. Burger and S. Osher, "A Guide to the TV Zoo," *Lecture Notes in Mathematics*, vol. 2090, pp. 1–70, 2013.

[15] J. F. P. J. Abascal, J. Chamorro-Servent *et al.*, "Fluorescence diffuse optical tomography using the split Bregman method," *Med. Phys.*, vol. 38, no. 11, pp. 6275–6284, 2011.

[16] J. Abascal, M. Abella, A. Sisniega *et al.*, "Investigation of different sparsity transforms for the PICCS algorithm in small-animal respiratory gated CT," *PLoS One*, vol. 10, no. 4, 2015.

[17] S. Si-Mohamed, D. Bar-Ness *et al.*, "Review of an initial experience with an experimental spectral photon-counting computed tomography system," *Nucl. Instr. Meth. Phys. Res. B*, vol. 873, pp. 27–35, 2017.

[18] J.-J. Abascal, J. Chamorro-Servent, J. Aguirre *et al.*, "Fluorescence diffuse optical tomography using the split Bregman method," *Med. Phys.*, vol. 38, no. 11, 2011.