# MEDICAL IMAGE SUPER-RESOLUTION WITH NON-LOCAL EMBEDDING SPARSE REPRESENTATION AND IMPROVED IBP

*Mingli Zhang*[1], *Christian Desrosiers*[1], *Qiang Qu*[2], *Fenghua Guo*[3], *Caiming Zhang*[3,4]

[1] École de technologie supérieure, Montreal, Canada, H3C 1K3
[2]Data Science and Information Systems (Dainfos) Lab, Innopolis University
[3]School of Computer Science and Technology, Shandong University, Jinan, China
[4]Lab of Digital Media Technology, Shandong University of Finance and Economics, Jinan, China

## ABSTRACT

This paper proposes a novel super-resolution method that exploits the sparse representation and non-local similarity of patches for the effective reconstruction of images. High-resolution images are reconstructed from low resolution observations with an efficient technique based on the alternating direction method of multipliers (ADMM). A robust iterative back-projection approach is used in a post-processing step to remove residual noise and artifacts in the reconstructed image. Experiments on benchmark medical images illustrate the advantage of our method, in terms of PSNR and SSIM, compared to state of the art approaches.

***Index Terms*—** Non-local sparse representation, non-local embedding, improved IBP, high-resolution image

## 1. INTRODUCTION

Medical imaging provides a non-invasive way to obtain rich information about various anatomical structures in the body [1–3]. The acquisition of high-resolution and noise free medical images is essential to many post-processing steps, such as the segmentation and registration of structures in these images. Having high quality images is also critical to the automated detection and diagnosis of various diseases. However, due to limitations of the imaging hardware, quality-limiting factors (e.g., signal to noise ratio) and time requirements, obtaining such images can be challenging.

An alternative solution is to use image processing techniques to enhance the spatial resolution and general quality of images. To reach this goal, conventional interpolation methods are widely used. However, these methods may blur sharp edges representing the contour of anatomical structures, introduce ringing artifacts, and fail to recover fine details in the image [1]. A data-driven solution to this problem is super-resolution (SR), which aims at reconstructing a high-resolution (HR) image from a low-resolution (LR) one. SR methods can be divided in two broad categories: single image algorithms [2, 4, 5], which recover the HR image from a single LR image, and multi-frame algorithms [6, 7] which use different LR observations of an HR image to reconstruct this image.

Formally, the process of acquiring LR images $y \in \mathbb{R}^{N_l}$ from an HR image $x \in \mathbb{R}^{N_h}$ can be formulated as

$$y \;=\; DHx + v \tag{1}$$

where, $H \in \mathbb{R}^{N_h \times N_h}$ is a blurring operator, $D \in \mathbb{R}^{N_l \times N_h}$ is a down-sampling operator, and $v$ is additive noise. In this paper, we suppose the noise to be zero-mean Gaussian, and the blur kernel to be the Dirac delta function (i.e., $H$ is the identity matrix). In this case, the single image SR problem can be defined as finding the HR image $\widehat{x}$ such that

$$\widehat{x} \;=\; \operatorname*{argmin}_{x} \; \frac{1}{2}\|y - Dx\|_2^2 \;+\; \lambda\mathcal{R}(x). \tag{2}$$

In this formulation, $\mathcal{R}(x)$ is a regularization term encoding prior knowledge on the HR image, and $\lambda$ is used to balance the fidelity term and regularization term.

In recent years, sparse representation models have shown promising results for improving the robustness of super-resolution methods [4, 8, 9]. Such models are based on the hypothesis that small patches of $\sqrt{M} \times \sqrt{M}$ pixels in an image $x$ can be approximated as a sparse linear combination of atoms within a dictionary $\Phi \in \mathbb{R}^{M \times K}$, where $K$ in the number of atoms [4]. Denote as $p_i$ the patch corresponding to pixel $i$ in $x$ (patches corresponding to different pixels may overlap), and let $R_i \in \mathbb{R}^{M \times N_h}$ be the patch extraction matrix such that $p_i = R_i x$. Using the $l_1$-norm to model sparsity, the sparse representation formulation of SR corresponds to the following problem:

$$\begin{aligned} &\operatorname*{argmin}_{\alpha_{\bullet}, x} \; \frac{1}{2}\|y - Dx\|_2^2 \;+\; \lambda \sum_{i=1}^{N_h} \|\alpha_i\|_1 \\ &\text{s.t. } R_i x \;=\; \Phi\alpha_i, \quad i = 1, \ldots, N_h. \end{aligned} \tag{3}$$

Reconstruction methods can often be improved by adding non-local self-similarity constraints [4, 10–14]. The idea behind this approach is that patches in an image are similar to other patches in the same image, due to the regularity and recurrence of structures (e.g., see Figure 1). Formally, a patch $p_i$ can be approximated as a convex combination of $S$

non-local similar patches $p_i^s$:

$$p_i \approx \sum_{s=1}^{S} \omega_i^s \, p_i^s, \;\; \text{s.t} \sum_{s=1}^{S} \omega_i^s = 1, \;\; \omega_i^s \geq 0, \; \forall s. \qquad (4)$$

The relation between similar patches, encoded in weights $\omega_i^s$ can be used to impose constraints on the reconstructed HR image [10–12] or the sparse codes [4, 13, 14].

**Fig. 1**. Example of similar patches in an image.

In this paper, we propose a novel single-image super-resolution method using non-local sparse representation and iterative back-projection. The detailed contributions of this work are the following:

- An efficient algorithm, based on the alternating direction method of multipliers (ADMM) [15], is presented for the reconstruction of HR images. As other sparse representation methods, the proposed approach uses a compact dictionary to effectively model the subspace of image patches and regularizes the reconstruction by imposing sparsity constraints. Our approach further regularizes the reconstruction process by adding non-local similarity constraints on the patches and their sparse codes.
- Super-resolution methods can lead to residual noise and reconstruction artifacts in the form of ringing patterns. To remove such artifacts, we enhance our model by incorporating a post-processing step based on robust iterative back-projection (IBP) [16].

In the next section, we present the proposed SR model, the ADMM approach used to recover the HR image, and the IBP post-processing step to remove residual noise and artifacts.

## 2. THE PROPOSED METHOD

The workflow of our SR method is composed of three steps: 1) computation of non-local similar patches, 2) HR image recovery using sparse representation and non-local embedding, and 3) post-processing removal of artifacts using IBP. The following subsections describe each of these steps.

### 2.1. Computation of non-local similar patches

To compute the non-local embedding of patches, we first approximate the HR image from the LR one using bicubic interpolation. We then find for each HR patch $p_i$ the $S$ most similar patches in the image based on the weighted Euclidean distance. To speed-up this process, the search is limited to a small window around target pixel $i$, for instance, of size $100 \times 100$ pixels. The process can be further accelerated by finding a shortlist of candidates in the LR image [17].

Based on Eq. (4), a patch $p_i$ is then encoded as a convex combination of its similar patches $p_i^s$. Let $P_i = [p_i^1, \ldots, p_i^S]$, the non-local similarity weights of $p_i$ can be computed by solving the following constrained least-square problem:

$$\underset{\omega_i}{\operatorname{argmin}} \; \frac{1}{2}\|p_i - P_i\omega_i\|_2^2, \; \text{s.t.} \; \|\omega_i\|_1 = 1, \; \omega_i \geq 0. \qquad (5)$$

Note that, since $S$ is typically small, the solution can be obtained rapidly using a standard quadratic program solver. The non-local embedding of all patches can be represented using a matrix $\Omega$ such that

$$\Omega_{i,j} = \begin{cases} \omega_i^s, \; \text{if } p_j \text{ is the } s\text{-th non-local similar patch of } p_i \\ 0, \; \text{otherwise.} \end{cases} \qquad (6)$$

Matrix $\Omega$ is used to constrain the recovered HR image, as described in the following subsection.

### 2.2. Image recovery via non-local sparse representation

Our super-resolution approach is related to the sparse representation model proposed by Dong et al. in [4]. As their model, we use similarities between non-local patches, as defined in matrix $\Omega$, to constrain their sparse code representation. Let $P = [p_1, \ldots, p_{N_h}]$ be the matrix of all patches, where $p_i = R_i x$. From Eq. (4), we know that $P \approx P\Omega^\top$. Moreover, since the sparse representation of each patch $p_i$ is $\Phi\alpha_i$, we also have that $P = \Phi A$, where $A = [\alpha_1, \ldots, \alpha_{N_h}]$ is the matrix of all sparse codes. Combining these two relations gives a non-local embedding of sparse codes, $\Phi A \approx \Phi A \Omega^\top$, which is used to enhance the standard sparse representation SR model of Eq. (3) as follows:

$$\underset{A,\,x}{\operatorname{argmin}} \; \frac{1}{2}\|y - Dx\|_2^2 \, + \, \lambda\|A\|_1 \, + \, \frac{\gamma}{2}\|\Phi A - \Phi A\Omega^\top\|_F^2$$
$$\text{s.t. } R_i x \, = \, \Phi\alpha_i, \quad i = 1, \ldots, N_h. \qquad (7)$$

Parameters $\lambda \geq 0$ and $\gamma \geq 0$ control the trade-off between the sparsity of patches in the dictionary space and their non-local embedding.

An optimization strategy based on the alternating direction method of multipliers (ADMM) [15] is used to recover the sparse codes $A$ and HR image $x$. The general principle of this method is to split a hard to solve problem into easier to solve sub-problems, the solutions of which are connected using auxiliary variables. In our case, we add auxiliary variables $Y$, $Z$ and reformulate the problem as

$$\underset{A,\,Y,\,Z,\,x}{\operatorname{argmin}} \; \frac{1}{2}\|y - Dx\|_2^2 \, + \, \lambda\|Y\|_1 \, + \, \frac{\gamma}{2}\|Z - Z\Omega^\top\|_F^2$$
$$\text{s.t. } R_i x \, = \, \Phi\alpha_i, \; i = 1, \ldots, N_h, \; Y = A, \; Z = \Phi A. \qquad (8)$$

The constraints are then moved to the cost function by added three augmented Lagrangian terms with multipliers $U, V, W$, and parameters $\mu_1, \mu_2, \mu_3 \geq 0$:

$$\underset{\substack{A,Y,Z,x,\\U,V,W}}{\operatorname{argmin}} \ \frac{1}{2}\|y - Dx\|_2^2 \ + \ \lambda\|Y\|_1 \ + \ \frac{\gamma}{2}\|Z - Z\Omega^\top\|_F^2 \\ + \ \frac{\mu_1}{2}\sum_{i=1}^{N_h}\|R_i x - \Phi\alpha_i + u_i\|_2^2 \ + \ \frac{\mu_2}{2}\|A - Y + V\|_F^2 \\ + \ \frac{\mu_3}{2}\|\Phi A - Z + W\|_F^2. \quad (9)$$

In practice, $\mu_1$, $\mu_2$ and $\mu_3$ affect the convergence of the algorithm, both in terms of speed and optimality. Too large values will lead to a fast convergence but a sub-optimal solution, while too small ones give a slow convergence. However, ADMM methods are not overly sensitive to those parameters, and little tuning is required. As in standard ADMM methods, we update variables $A$, $Y$, $Z$, $x$, $U$, $V$ and $W$, alternatively, until convergence is reached.

### 2.2.1. *Updating $x$*

Image $x$ is updating by solving the following problem:

$$\underset{x}{\operatorname{argmin}} \ \frac{1}{2}\|y - Dx\|_2^2 + \frac{\mu_1}{2}\sum_{i=1}^{N_h}\|R_i x - (\Phi\alpha_i - u_i)\|_2^2. \quad (10)$$

This corresponds to an unconstrained quadratic program, having a closed form solution

$$x \ = \ \Big(D^\top D + \mu_1\sum_{i=1}^{N_h}R_i^\top R_i\Big)^{-1}\Big(D^\top y + \mu_1\sum_{i=1}^{N_h}R_i^\top(\Phi\alpha_i - u_i)\Big) \quad (11)$$

Since matrix $M = D^\top D + \mu_1\sum_{i=1}^{N_h}R_i^\top R_i$ is fixed, we can use Cholesky factorization in a pre-processing step to decompose it as $M = CC^\top$, where $C$ is lower triangular. The HR image can then be updated as

$$x \ = \ C^\top \setminus \Big(C \setminus \big(D^\top y \ + \ \mu_1\sum_{i=1}^{N_h}R_i^\top(\Phi\alpha_i - u_i)\big)\Big), \quad (12)$$

where operator $\setminus$ corresponds to a simple backward/forward substitution.

### 2.2.2. *Updating $A$*

To update $A$, we first reconstruct patch matrix $P$ such that $p_i = R_i x$, where $x$ is the updated HR image. The task of updating $A$ can then be expressed as follows:

$$\underset{A}{\operatorname{argmin}} \ \frac{\mu_1}{2}\|\Phi A - (P + U)\|_F^2 \ + \ \frac{\mu_2}{2}\|A - (Y - V)\|_F^2 \\ + \ \frac{\mu_3}{2}\|\Phi A - (Z - W)\|_F^2. \quad (13)$$

The closed form solution to this problem is given by

$$A \ = \Big((\mu_1 + \mu_3)\Phi^\top\Phi \ + \ \mu_2 I\Big)^{-1}\Big(\mu_1\Phi^\top(P + U) \\ + \ \mu_2(Y - V) \ + \ \mu_3\Phi^\top(Z - W)\Big). \quad (14)$$

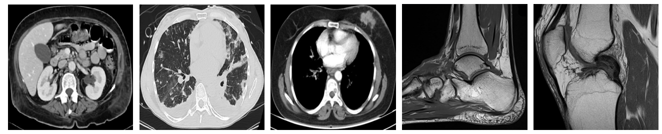


**Fig. 2**. The five test images: 1) abdomen CT, 2) thorax CT, 3) chest CT, 4) ankle MRI, 5) knee MRI.

Since the linear system to solve has only $K \leq 100$ equations, updating $A$ is fast.

### 2.2.3. *Updating $Y$*

Updating $Y$ corresponds to solving an $l_1$-norm proximal problem

$$\underset{Y}{\operatorname{argmin}} \ \lambda\|Y\|_1 \ + \ \frac{\mu_2}{2}\|Y - (A + V)\|_F^2, \quad (15)$$

the solution of which is obtained via the soft-thresholding operator $\mathcal{S}_\eta(X) = \text{sign}(X) \cdot \max(|X| - \eta, 0)$:

$$Y \ = \ \mathcal{S}_{\lambda/\mu_2}(A + V). \quad (16)$$

### 2.2.4. *Updating $Z$*

To update $Z$, we consider the following problem:

$$\underset{Z}{\operatorname{argmin}} \ \frac{\gamma}{2}\|Z - Z\Omega^\top\|_F^2 \ + \ \frac{\mu_3}{2}\|Z - (\Phi A + W)\|_F^2, \quad (17)$$

Let $Q = I - \Omega^\top$, the closed form solution this problem is

$$Z \ = \ \mu_3\big(\Phi A + W\big)\big(\gamma QQ^\top + \mu_3 I\big)^{-1}. \quad (18)$$

Once again, Cholesky decomposition can be used to accelerate the update of $Z$ (see Section 2.2.1).

### 2.2.5. *Updating $U$, $V$ and $W$*

Lastly, the Lagrangian multiplier are updated following the usual ADMM method:

$$\begin{aligned} U' &= U + (P - \Phi A) \\ V' &= V + (A - Y) \\ W' &= W + (\Phi A - Z) \end{aligned} \quad (19)$$

## 2.3. Post-processing artifact removal

In a low noise setting, the generative model of Eq. (1) imposes that $y \approx DHx$. Due to the sparsity and non-local embedding constraints, the reconstructed HR image $x$ may not satisfy this model exactly. This may lead to residual noise or artifacts like ringing patterns in the image. Following [18], this can be addressed with a post-processing step based on iterative back projection (IBP). Let $x_0$ be the image reconstructed from the

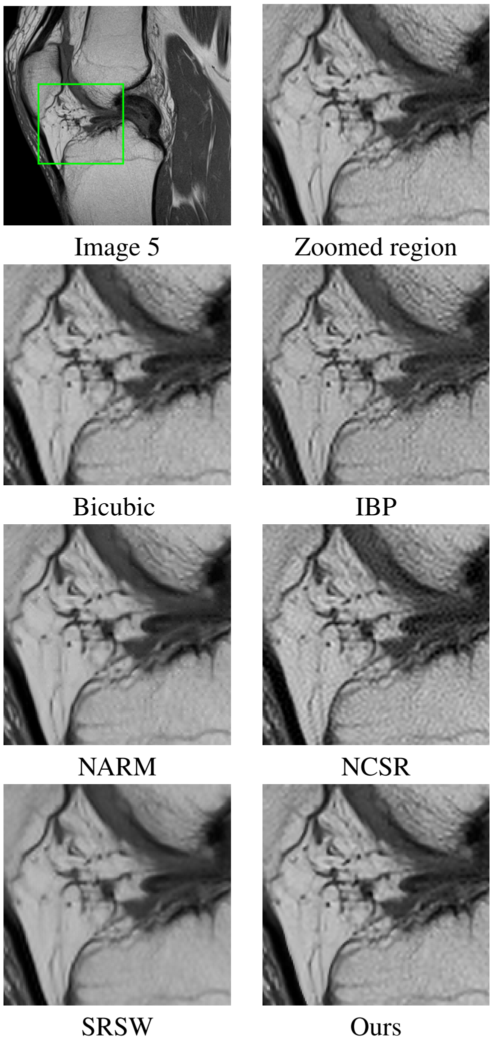


**Fig. 3**. Examples of reconstructions obtained for a zoomed region of Image 5 (knee MRI).

**Table 1**. PSNR (dB) and SSIM obtained by the SR methods on the test images of Figure 2 (scale = 2).

| | Bicubic | IBP | NARM | NCSR | SRSW | Ours |
|---|---|---|---|---|---|---|
| Image 1 | 36.45 | 31.62 | 38.96 | 32.80 | 35.19 | **39.84** |
| | 0.989 | 0.967 | 0.987 | 0.855 | 0.972 | **0.991** |
| Image 2 | 33.69 | 27.74 | 36.09 | 31.85 | 30.07 | **37.17** |
| | 0.965 | 0.929 | 0.965 | 0.830 | 0.908 | **0.980** |
| Image 3 | 35.15 | 29.53 | 39.46 | 32.97 | 32.99 | **39.71** |
| | 0.981 | 0.953 | 0.977 | 0.869 | 0.953 | **0.987** |
| Image 4 | 31.76 | 28.91 | 31.88 | 30.48 | 30.86 | **33.57** |
| | 0.926 | 0.894 | 0.917 | 0.835 | 0.909 | **0.951** |
| Image 5 | 31.76 | 28.91 | 34.64 | 31.26 | 27.21 | **36.15** |
| | 0.926 | 0.894 | 0.915 | 0.806 | 0.828 | **0.951** |
| **Average** | 33.76 | 29.34 | 36.21 | 31.87 | 31.26 | **37.29** |
| | 0.957 | 0.927 | 0.952 | 0.839 | 0.914 | **0.972** |

SR method, the post-processing step consists in finding a new image $x_{\mathrm{new}}$ such that

$$x_{\mathrm{new}} \;=\; \operatorname*{argmin}_{x} \; \frac{1}{2}\|x - x_0\|, \;\; \text{s.t. } y \;=\; DHx. \qquad (20)$$

The solution to this problem can be obtained efficiently using IBP. Starting with $x_0$, the HR image is updated iteratively using the following scheme:

$$x_{t+1} \;=\; x_t \;+\; B(y \,-\, DHx_t), \qquad (21)$$

where $B$ is a linear operator corresponding to upsampling and back-projection.

Standard IBP methods can produce reconstruction artifacts by amplifying noise from one iteration to another. To overcome this limitation, we use the improved IBP method proposed in [16]. The general principle of this method is that reconstruction artifacts are often characterized by high variance regions in the residual $e_t = y - DHx_t$. To avoid such artifacts, high variance terms of $e_t$ are attenuated before applying the back-projection operator $B$.

## 3. EXPERIMENTAL RESULTS

We evaluated our proposed method on the five benchmark medical images of Figure 2, previously used in [14]. We compared our method to five well-known SR approaches: simple bicubic interpolation, IBP [19], NARM [4], NCSR [13], and SRSW [14]. For all experiments, we have used patches of $25 \times 25$ pixels (i.e., $M = 625$), and $S = 17$ for the number of non-local similar patches. Following [4], dictionary $\Phi$ was generated by clustering patches with the k-means algorithm, using $K = 60$ as the number of clusters. Finally, regularization parameters $\lambda$ and $\gamma$ were selected using a validation set.

Table 1 gives the performance obtained by the methods on the benchmark images, in terms of peak signal-to-noise ratio (PSNR) and structural similarity (SSIM) [20]. Based a one-sided paired t-test, our method outperforms all other approaches, both in terms of PSNR (max. $p$-value of $0.0065$) and SSIM (max. $p$-value of $0.0185$). Figure 3 shows examples of reconstructions obtained by the tested SR methods on Image 5 of Figure 2. We can see that our method leads to fewer ringing artifacts than other approaches, in particular bicubic interpolation, IBP and NCSR. In comparison to NARM and SRSW, our method is better at recovering fine details in the image, which could be useful for assessing the microstructure of tissues.

## 4. CONCLUSION

We presented a novel method[1] for the image super-resolution problem. Our method combines sparse representation and non-local patch embedding in a single model, and uses an efficient optimization algorithm based on ADMM to recover the high-resolution image. A post-processing step, using a robust iterative back-projection technique, is proposed to remove residual artifacts in the reconstructed image. Experiments on benchmark medical images show the advantage of our method compared to several state of the art approaches.

[1]This work is partly supported by the National Nature Science Foundation of China (61373078,61472220).

## 5. REFERENCES


[1] JD Van Ouwerkerk, "Image super-resolution survey," *Image and Vision Computing*, vol. 24, no. 10, pp. 1039–1052, 2006.

[2] Sung Cheol Park, Min Kyu Park, and Moon Gi Kang, "Super-resolution image reconstruction: a technical overview," *Signal Processing Magazine, IEEE*, vol. 20, no. 3, pp. 21–36, 2003.

[3] Kunio Doi, "Computer-aided diagnosis in medical imaging: historical review, current status and future potential," *Computerized medical imaging and graphics*, vol. 31, no. 4, pp. 198–211, 2007.

[4] Weisheng Dong, Lei Zhang, Rastislav Lukac, and Guangming Shi, "Sparse representation based image interpolation with nonlocal autoregressive modeling," *Image Processing, IEEE Transactions on*, vol. 22, no. 4, pp. 1382–1394, 2013.

[5] Mading Li, Jiaying Liu, Jie Ren, and Zongming Guo, "Adaptive general scale interpolation based on weighted autoregressive models," *Circuits and Systems for Video Technology, IEEE Transactions on*, vol. 25, no. 2, pp. 200–211, 2015.

[6] Sina Farsiu, Dirk Robinson, Michael Elad, and Peyman Milanfar, "Advances and challenges in super-resolution," *International Journal of Imaging Systems and Technology*, vol. 14, no. 2, pp. 47–57, 2004.

[7] Sean Borman and Robert L Stevenson, "Super-resolution from image sequences – a review," in *MWCAS*. IEEE, 1998, p. 374.

[8] Scott Shaobing Chen, David L Donoho, and Michael A Saunders, "Atomic decomposition by basis pursuit," *SIAM journal on scientific computing*, vol. 20, no. 1, pp. 33–61, 1998.

[9] Julien Mairal, Francis Bach, Jean Ponce, Guillermo Sapiro, and Andrew Zisserman, "Non-local sparse models for image restoration," in *Computer Vision, 2009 IEEE 12th International Conference on*. IEEE, 2009, pp. 2272–2279.

[10] José V Manjón, Pierrick Coupé, Antonio Buades, Vladimir Fonov, D Louis Collins, and Montserrat Robles, "Non-local MRI upsampling," *Medical image analysis*, vol. 14, no. 6, pp. 784–792, 2010.

[11] François Rousseau, Alzheimers Disease Neuroimaging Initiative, et al., "A non-local approach for image super-resolution using intermodality priors," *Medical image analysis*, vol. 14, no. 4, pp. 594–605, 2010.

[12] Hong Chang, Dit-Yan Yeung, and Yimin Xiong, "Super-resolution through neighbor embedding," in *Computer Vision and Pattern Recognition, 2004. CVPR 2004. IEEE Conference on*. IEEE, 2004, vol. 1, pp. I–I.

[13] Weisheng Dong, Lei Zhang, Guangming Shi, and Xin Li, "Nonlocally centralized sparse representation for image restoration," *Image Processing, IEEE Transactions on*, vol. 22, no. 4, pp. 1620–1630, 2013.

[14] Dinh-Hoan Trinh, Marie Luong, Franccoise Dibos, Jean-Marie Rocchisani, Canh-Duong Pham, and Truong Q Nguyen, "Novel example-based method for super-resolution and denoising of medical images," *Image Processing, IEEE Transactions on*, vol. 23, no. 4, pp. 1882–1895, 2014.

[15] Stephen Boyd, Neal Parikh, Eric Chu, Borja Peleato, and Jonathan Eckstein, "Distributed optimization and statistical learning via the alternating direction method of multipliers," *Foundations and Trends® in Machine Learning*, vol. 3, no. 1, pp. 1–122, 2011.

[16] Xin Yang, Yan Zhang, Dake Zhou, and Ruigang Yang, "An improved iterative back projection algorithm based on ringing artifacts suppression," *Neurocomputing*, vol. 162, pp. 171–179, 2015.

[17] Pierrick Coupé, Pierre Yger, Sylvain Prima, Pierre Hellier, Charles Kervrann, and Christian Barillot, "An optimized blockwise nonlocal means denoising filter for 3-D magnetic resonance images," *Medical Imaging, IEEE Transactions on*, vol. 27, no. 4, pp. 425–441, 2008.

[18] Jianchao Yang, John Wright, Thomas S Huang, and Yi Ma, "Image super-resolution via sparse representation," *Image Processing, IEEE Transactions on*, vol. 19, no. 11, pp. 2861–2873, 2010.

[19] Michal Irani and Shmuel Peleg, "Improving resolution by image registration," *CVGIP: Graphical models and image processing*, vol. 53, no. 3, pp. 231–239, 1991.

[20] Zhou Wang, Alan Conrad Bovik, Hamid Rahim Sheikh, and Eero P Simoncelli, "Image quality assessment: from error visibility to structural similarity," *Image Processing, IEEE Transactions on*, vol. 13, no. 4, pp. 600–612, 2004.