SUPER-RESOLUTION FROM A LOW- AND PARTIAL HIGH- RESOLUTION IMAGE PAIR

M. Hidane[1,3,4]*, J.-F. Aujol*[1,3]*, Y. Berthoumieu*[2,4]*, C. Deledalle*[1,3]

[1]Univ. Bordeaux, Talence, France
[2]IPB/ENSEIRB-Matmeca, Talence, France
[3]CNRS, IMB, UMR 5251, Talence, France
[4]CNRS, IMS, UMR 5218, Talence, France

ABSTRACT

The classical super-resolution (SR) setting starts with a set of low-resolution (LR) images related by subpixel shifts and tries to reconstruct a single high-resolution (HR) image. In some cases, partial observations about the HR image are also available. Trying to complete the missing HR data without any reference to LR ones is an inpainting (or completion) problem. In this paper, we consider the problem of recovering a single HR image from a pair consisting of a complete LR and incomplete HR image pair. This setting arises in particular when one wants to fuse image data captured at two different resolutions. We propose an efficient algorithm that allows to take advantage of both image data by first learning nonlocal interactions from an interpolated version of the LR image using patches. Those interactions are then used by a convex energy function whose minimization yields a super-resolved complete image.

***Index Terms*—** super-resolution; inpainting; nonlocal patch-based methods; graph-regularization; total variation; Douglas-Rachford algorithm.

1. INTRODUCTION

We consider in this paper the problem of recovering a high-resolution (HR) image from a pair consisting of a complete low-resolution (LR) image and an incomplete HR one. Before entering technical details, we first describe the context of our work as well as some related works.

1.1. Motivations and Context

The application that motivates our work concerns the estimation of petrophysical parameters of petroleum reservoirs through image processing techniques. In this context, a cylindrical sample of a rock is extracted from a reservoir and then imaged through computed tomography (CT) scanning in order to produce a three-dimensional digital representation. This 3D acquisition is further analyzed by an image analysis software in order to automatically segment and classify different regions according to visual cues related to some underlying petrophysical parameters. A typical example concerns the automatic classification of different textured regions, which in turn relates to different *porosity levels* of the underlying material.

In a practical setting, the limited resolution provided by the non invasive 3D reconstruction from CT acquisition does not allow for automatic inspection. Thus a need for higher resolution image data emerges. In the context we are interested in, this need is alleviated by performing a second acquisition, this time focusing on a specific part of the volume. This second acquisition thus provides incomplete but higher resolution slices of the extracted cylindrical object. Now, the goal is to combine the image data coming from both acquisitions in order to obtain a *complete* high-resolution volume.

1.2. Modeling

In this work, we put aside the 3D aspect and focus on the case of 2D images. Furthermore, we adopt a simplified framework where we assume that the forward model for the HR/LR pair is completely known. In particular, we do not discuss possible registration issues. Letting $y_1 \in \mathbb{R}^p$ and $y_2 \in \mathbb{R}^n$ respectively denote the lexicographical ordering of the complete LR and partial HR images, we adopt the following discrete forward model:

$$\begin{cases} y_1 &= SHf_0 + \eta_1 \in \mathbb{R}^p, \\ y_2 &= M(f_0 + \eta_2) \in \mathbb{R}^n, \end{cases} \tag{1}$$

where $0 < p < n$; $f_0 \in \mathbb{R}^n$ denotes the unknown HR image; $S \in \mathbb{R}^{p \times n}$ stands for spatial downsampling by a dyadic factor $r = 2^k$ in each direction; $H \in \mathbb{R}^{n \times n}$ accounts for spatial blurring of the image, modelled in our case by a circular convolution with a known point spread function h: $Hx = h \circledast x$; $M = \text{diag}(m_1, \ldots, m_n) \in \mathbb{R}^{n \times n}$ accounts for some sort of occlusion in the image in the form of a binary mask: $m_i = 1$, if pixel i is observed, and $m_i = 0$ otherwise; the vectors $\eta_1 \in \mathbb{R}^p$ and $\eta_2 \in \mathbb{R}^n$ are white Gaussian noise.

This study has been carried out with financial support from the French State, managed by the French National Research Agency (ANR) in the frame of the "Investments for the future" Programme IdEx Bordeaux - CPU (ANR-10-IDEX-03-02). J-F. Aujol is a member of Institut Universitaire de France.

1.3. Related Work

Considering the reconstruction of f_0 from y_1 alone is a *zooming* problem, often also called magnification, interpolation, upsampling or single-image super-resolution. This problem has been the subject of many investigations and it would be difficult within this note to sketch the different trends developed. We instead refer the interested reader to the recent works [1, 2, 3] and to references therein.

The reconstruction of f_0 from y_2 alone is an *inpainting* problem. As for the zooming problem, inpainting has been extensively studied; the interested reader can consult [5] as a starting point for this large research theme. Broadly speaking, inpainting algorithms can be categorized as either geometry-oriented [6, 7] or texture-oriented [8, 9]; the latter class being also termed examplar-based (see [10, 11] and references therein).

Considering the reconstruction of f_0 from both y_1 and y_2 is a super-resolution problem that can be seen either as an inpainting problem with additional LR consistency constraints or as a zooming problem with HR (soft) constraints. While the traditional super-resolution setting consists in the reconstruction of f_0 from *multiple* LR images related by *subpixel shifts*, we note that this setting becomes appropriate when we consider the full 3D problem sketched earlier. We refer the interested reader to [12] for a review article and to [13] for a recent monograph.

1.4. Plan of the paper

Our contributions are presented in Section 2 where we introduce a strategy for exploiting both the available HR and LR data for the super-resolution task. Therein, we instantiate a nonlocal variational problem whose solution gives an estimation of the HR image. Links to total variation regularization [14] are discussed and a minimization using Douglas-Rachford algorithm is presented. We perform a set of numerical experiments in Section 3, and show that it allows to recover fine detail structures. We conclude in Section 4.

1.5. Notations

In all this paper, a standard LR image is denoted $y_1 \in \mathbb{R}^p$ while $y_2 \in \mathbb{R}^n$ denotes a standard HR image, $p < n$. The matrices M, S, H are the ones defined above. Let $\Omega = \{1, \dots, n\}$ denote the set of HR pixels, $D \subset \Omega$ denote the set of pixels where the HR information is missing, so that D^c denotes the set of pixels where the HR information is available (see Figure1). For an image f and a pixel i, $p_i(f)$ denotes an image patch extracted from f and centered at i.

For a matrix A and a vector y, we write $\{A\,.\, = y\}$ for the set $\{x, Ax = y\}$. For a set C, ι_C denotes the indicator function: $\iota_C(x) = 1$ if $x \in C$, and $\iota_C(x) = +\infty$ otherwise. We denote by $\Gamma_0(\mathbb{R}^n)$ the set of all convex lower semi-continuous and proper functions on $\mathbb{R}^n$. Further notations will be introduced as needed.



(a) (b)

Fig. 1: Example of a partial HR (a) and complete LR (b) image pair. Our goal is to complete (a) using information from (b), or equivalently, to super-resolve (b) using information from (a).

2. NONLOCAL VARIATIONAL FRAMEWORK

The recovery of an estimate $\hat{f}$ of f_0 from the measurements y_1 and y_2 according to the forward model (1) is an ill-posed inverse problem. As is usual for such inverse problems in imaging, we formulate the estimation task in a variational setting leading to the minimization an energy function of the form $F(f) = F_1(f) + F_2(f)$. The function F_1 is a regularization term forcing the solutions to have pre-specified properties, while the term $F_2(f)$ penalizes the discrepancy between f and y_1, y_2, according to the forward model (1). In the presence of Gaussian noise, the latter term is usually taken, in its penalized form, as the squared ℓ_2 distance and can be written in our case as $F_2(f) = \lambda_1 \|SHf - y_1\|^2 + \lambda_2 \|Mf - y_1\|^2$. If the noise level in the HR image can be neglected, we will instead consider the following form: $F_2(f) = \lambda \|SHf - y_1\|^2 + \iota_{\{M\,.\, = y_2\}}(f)$.

2.1. LR/HR-Driven Graph Regularizer

To obtain an estimate $\hat{f}$ of f_0 we need to define a sensible prior F_1. The regularizer we propose in this paper tries to enforce the coherence of patches in a way similar to the NL-Means algorithm [15]. The first step is to interpolate the LR image $y_1 \in \mathbb{R}^p$ to match the domain of the incomplete HR image $y_2 \in \mathbb{R}^n$. To this end, we use bicubic interpolation and get $\tilde{y}_1 \in \mathbb{R}^n$. The image $\tilde{y}_1$ is used to build similarities between patches whose centers lie in D and patches whose centers lie outside D. In practice, we form a *directed* graph G whose vertices are the (locations of the) pixels of y_2 and whose weighted adjacency matrix $W \in \mathbb{R}^{n \times n}$ satisfies $w_{i,j} \neq 0$ if and only if $(i, j) \in D \times D^c$ and the patch $p_j(\tilde{y}_1)$ belongs the k-nearest neighbors of $p_i(\tilde{y}_1)$ according to the Euclidean distance. In the experiments we performed, we opted for binary weights $w_{i,j} \in \{0, 1\}$.

Equipped with the graph encoded by the weighted adjacency matrix W, we propose the following regularizer, in-

spired by [16, 17]: $F_1(f) = \sum_{i=1}^{n}(\sum_{j=1}^{n} w_{i,j}(f_i - f_j)^2)^{1/2}$. Note that, due to the discussed properties of W, the outer sum actually runs over D, while for $i \in D$, the inner sum runs only over a subset of D^c consisting of the neighbors of i. This is in contrast with the works in [16, 17], where the matrix W is assumed symmetric.

To summarize, we use an interpolated version $\tilde{y}_1$ of the LR image y_2 to infer a non-symmetric binary relation between the occluded patches and the visible HR patches and use this relation to infer the values of missing pixels by adopting F_1 as a prior and F_2 as a data fidelity measure.

Let us note that if we allow to define edges in $\Omega \times \Omega$ instead of restricting to $D \times D^c$, then one can recover the isotropic total variation (TV) regularizer [14] by connecting each pixel to its right and bottom neighbors. Works using the TV prior for image interpolation and super-resolution can be found in [18, 19, 20].

2.2. Douglas-Rachford Algorithm in a Product-Space

From the previous section we are led to consider the following convex optimization problem:

$$\underset{f\in\mathbb{R}^n}{\text{minimize}}\; F_1(f) + \frac{\lambda}{2}\|SHf - y_1\|^2 + \iota_{\{M\,.\,=y_2\}}(f). \quad (2)$$

To proceed with the derivation of the algorithm, we rewrite the penalizer F_1 in the form $F_1(f) = \|\nabla_w(f)\|_{1,2}$, where $\nabla_w : \mathbb{R}^n \mapsto \mathbb{R}^{n\times n}$ is given, for $i, j \in \{1, \ldots, n\}$, by $(\nabla_w f)_{i,j} = \sqrt{w_{i,j}}(f_j - f_i)$ and $\|.\|_{1,2}$ is the mixed-norm given, for $p \in \mathbb{R}^{n\times n}$, by $\|p\|_{1,2} = \sum_{i=1}^{n}\sqrt{\sum_{j=1}^{n} p_{i,j}^2}$.

As the energy in (2) is not differentiable, we use proximal splitting algorithms [21] in order to tackle the optimization. The base operation in this class of algorithms is the computation of the proximity operator of a function $F \in \Gamma_0(\mathbb{R}^n)$: $\text{prox}_F z = \underset{x\in\mathbb{R}^n}{\text{argmin}} \frac{1}{2}\|x - z\|^2 + F(x)$.

The energy in (2) is composite (mixing the operators M, S, H) and contains two nonsmooth terms. The direct application of the forward-backward or the primal-dual algorithm of Chambolle and Pock [22] would lead to iterative algorithms with difficult optimization sub-problems at each iteration. In order to take advantage of the properties of the involved operators, and to allow partial splitting, we adopt the strategy suggested in [23] by transcripting the Douglas-Rachford algorithm in a product space. More precisely, letting $X = \mathbb{R}^{n\times n} \times \mathbb{R}^{n\times n} \times \mathbb{R}^{n\times n}$ and $C = \{(\nabla_w f', Hf', f'), f' \in \mathbb{R}^n\}$, we rewrite (2) in the form

$$\underset{(p,g,f)\in X}{\text{minimize}}\; \|p\|_{1,2} + \frac{\lambda}{2}\|Sg - y_1\|^2 + \iota_{\{M\,.\,=y_2\}}(f) + \iota_C(p, g, f). \quad (3)$$

We solve problem (3) by applying the Douglas-Rachford algorithm [21]. Letting $\mathcal{F}_1(p, g, f) = \|p\|_{1,2} + \frac{\lambda}{2}\|Sg - y_1\|^2 + \iota_{\{M\,.\,=y_2\}}(f)$ and $\mathcal{F}_2(p, g, f) = \iota_C(p, g, f)$, we recall that in each iteration of the Douglas-Rachford algorithm we need to compute $\text{prox}_{\mathcal{F}_1}$ and $\text{prox}_{\mathcal{F}_2}$. Let us start with the evaluation of $\text{prox}_{\mathcal{F}_1}$. Since $\mathcal{F}_1$ is separable, the evaluation of $\text{prox}_{\mathcal{F}_1}(p, g, f)$ amounts to evaluating each part separately. The first component, $\text{prox}_{\|.\|_{1,2}}$, corresponds to the matrix form of soft-thresholding [24]. The matrix S being diagonal, the evaluation of $\text{prox}_{\|S.-y1\|^2}$ amounts to solving a diagonal linear system of equations. Finally, the evaluation of the third proximity operator is immediate since its corresponds to projecting on the set $\{M\,.\,= y_2\}$. The evaluation of $\text{prox}_{\mathcal{F}_2}$ leads to a linear system whose matrix is $A = \nabla_w^t\nabla_w + H^tH + I$, where ∇_w^t denotes the adjoint of ∇_w, with respect to the standard Euclidean inner product. We solve this inner system with a conjugate gradient method, using a 'warm start' strategy, meaning that we start the conjugate gradient solver at iteration $k+1$ with the result it yielded at iteration k. The algorithm we use can now be summarized as follows:

1: $\gamma > 0, \beta \in]0, 2[$
2: $p_{1,0} \in \mathbb{R}^{n\times n}, g_{1,0} \in \mathbb{R}^n, f_{1,0} \in \mathbb{R}^n$
3: $p_{2,n} = \text{ST}_\alpha(p_{1,n})$ (matrix soft-thresholding)
4: $g_{2,n} = (\gamma\lambda S^tS + I)^{-1}(g_{1,n} + \gamma\lambda S^t y_1)$
5: $f_{2,n} = \text{proj}_{M\,.\,=y_2}(f_{1,n})$ (Euclidean projection)
6: $a_n = A^{-1}(\nabla^t(2p_{2,n} - p_{1,n}) + H^t(2g_{2,n} - g_{1,n}) + 2f_{2,n} - f_{1,n})$ (using conjugate gradient)
7: $p_{1,n+1} = p_{1,n} + \beta(\nabla a_n - p_{2,n})$
8: $g_{1,n+1} = g_{1,n} + \beta(Ha_n - g_{2,n})$
9: $f_{1,n+1} = f_{1,n} + \beta(a_n - f_{2,n})$

As remarked at the end of the previous section, if we allow to define edges in $\{1, \ldots, N\} \times \{1, \ldots, N\}$ instead of $D \times D^c$ only, then one can recover the isotropic total variation regularizer. Then, by adopting circular boundary conditions, the operator $\nabla_w^t\nabla_w$ corresponds to the 4-stencil discretization of the Euclidean Laplacian operator. Hence, the matrix A in step 6 becomes diagonal in the discrete Fourier domain and thus the system can be solved by applying the discrete Fourier transform, modulating, and applying the inverse discrete Fourier transform. This is the version we use when we compare in the next section our regularization scheme with TV-based regularization.

3. EXPERIMENTS

We illustrate in this section the behaviour of our method on a set of 3 images. The first and second images are part of two standard test images; the last image is an image of a slice of a rock captured by an electronic microscope [1].

In each case we start with the ground truth full HR image f_0 and simulate the incomplete image y_2 by masking part of f_0. Similarly, we simulate the complete LR image y_1 by applying a Gaussian convolution kernel of bandwidth σ followed by spatial downsampling of factor r in each direction.

[1]The authors would like to thank the team "Sismage" from the Group TOTAL for providing CT data.

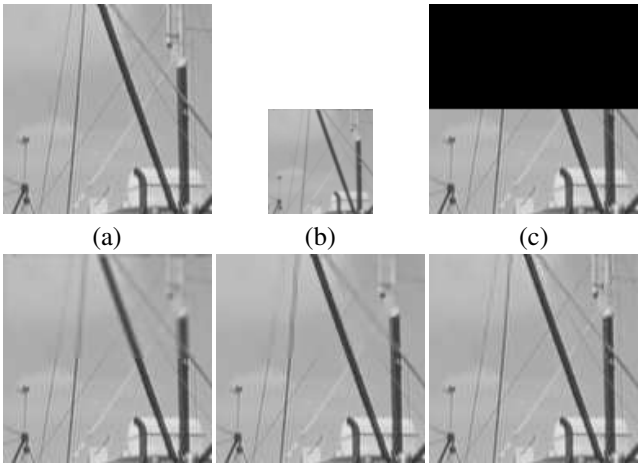


Fig. 2: (a) ground-truth image 128×128; (b) LR image ($\sigma = 1$ and $r = 2$); (c) masked HR (50%percent missing); (d) bicubic interpolation of missing part (psnr=31.3);(e) TV (psnr=34.69); (f) our result (psnr=37.66).

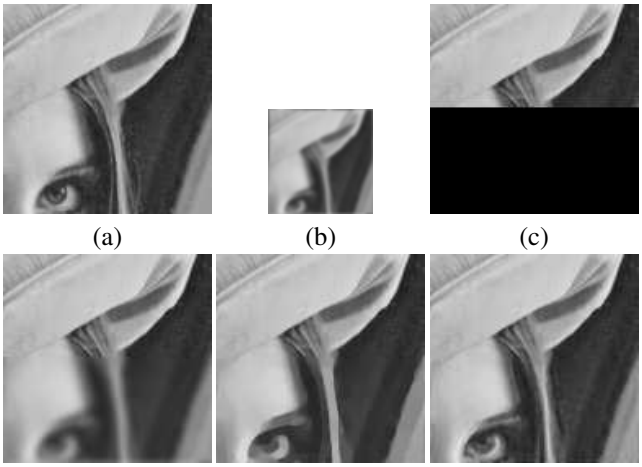


Fig. 3: (a) ground-truth image 128×128; (b) LR image ($\sigma = 2$ and $r = 2$); (c) masked HR (50%percent missing); (d) bicubic interpolation of missing part (psnr=29.27);(e) TV (psnr=33.39); (f) our result (psnr=34.91).

The parameters used are given in the caption of corresponding figure. We then interpolate y_1 to the size of y_2 and get $\tilde{y}_1$. Then, we construct a weighted graph as explained in Section 2, based on 5×5 patches extracted from $\tilde{y}_1$. We choose 15 nearest neighbors in each case. Once the corresponding graph has been estimated, we minimize the energy (2) using the algorithm of Section 2. We finally compare the solution we get with the one obtained by adopting the total variation as a prior. The corresponding energy is minimized again using the algorithm of Section 2, the difference being that the inner system of linear equation can be solved by Fourier transform as explained above. We also compare those two solutions with a simple recopy of $\tilde{y}_1$ (see, e.g., image (d) in Figure 2). Regarding the regularization parameter λ, it has been chosen by trial and error, each time seeking to optimize the peak signal to noise ration between the solution and the ground truth.

From the Figures 2, 3 and 4 we see that our approach invariably performs better than the bicubic interpolation and the TV-based super-resolution both in term of PSNR and visual inspection. In particular, for the textured area of the slice rock image, our method allows to recreate the textured part in a faithful manner.

4. CONCLUSION

We have studied in this paper an inverse problem which combines the single-image super-resolution and the inpainting/completion problems. We have motivated the importance of this problem and one of its possible applications. Then we developed a method capable of taking advantage of both the HR and LR image data in order to perform super-resolution. The proposed method consists in minimizing a convex but not differentiable energy defined over a directed graph. We have shown how the Douglas-Rachford algorithm can be used to solve the problem. Finally we performed a series of numerical simulations showing the benefit of our method.

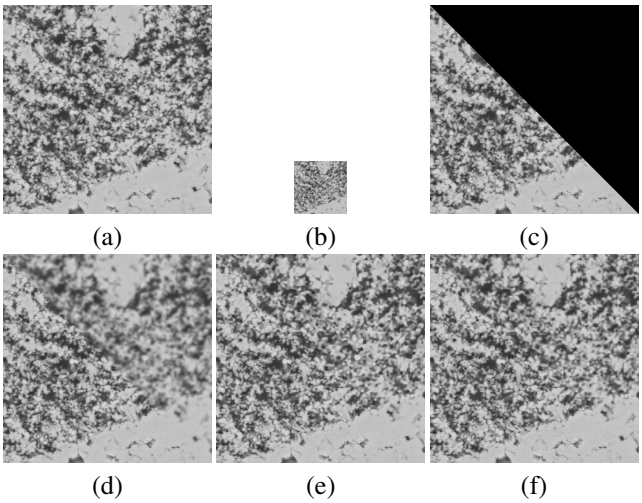


Fig. 4: (a) ground-truth image 256×256; (b) LR image ($\sigma = 2$ and $r = 4$); (c) masked HR (50%percent missing); (d) bicubic interpolation of missing part (psnr=24.43);(e) TV (psnr=29.52); (f) our result (psnr=30.16).

5. REFERENCES

[1] A. Bourquard and M. Unser, "Anisotropic interpolation of sparse generalized image samples," *Image Processing, IEEE Transactions on*, vol. 22, no. 2, pp. 459–472, Feb 2013.

[2] Stéphane Mallat and Guoshen Yu, "Super-resolution with sparse mixing estimators," *Image Processing, IEEE Transactions on*, vol. 19, no. 11, pp. 2889–2900, 2010.

[3] Daniel Glasner, Shai Bagon, and Michal Irani, "Super-resolution from a single image," in *Computer Vision, 2009 IEEE 12th International Conference on*. IEEE, 2009, pp. 349–356.

[4] Christine Guillemot and Olivier Le Meur, "Image inpainting: Overview and recent advances," *Signal Processing Magazine, IEEE*, vol. 31, no. 1, pp. 127–144, 2014.

[5] Jianhong Shen and Tony F Chan, "Mathematical models for local nontexture inpaintings," *SIAM Journal on Applied Mathematics*, vol. 62, no. 3, pp. 1019–1043, 2002.

[6] Simon Masnou, "Disocclusion: a variational approach using level lines," *Image Processing, IEEE Transactions on*, vol. 11, no. 2, pp. 68–76, 2002.

[7] Homan Igehy and Lucas Pereira, "Image replacement through texture synthesis," in *Image Processing, 1997. Proceedings., International Conference on*. IEEE, 1997, vol. 3, pp. 186–189.

[8] Alexei A Efros and Thomas K Leung, "Texture synthesis by non-parametric sampling," in *Computer Vision, 1999. The Proceedings of the Seventh IEEE International Conference on*. IEEE, 1999, vol. 2, pp. 1033–1038.

[9] Pablo Arias, Gabriele Facciolo, Vicent Caselles, and Guillermo Sapiro, "A variational framework for exemplar-based image inpainting," *International journal of computer vision*, vol. 93, no. 3, pp. 319–347, 2011.

[10] Jean-François Aujol, Saïd Ladjal, and Simon Masnou, "Exemplar-based inpainting from a variational point of view," *SIAM Journal on Mathematical Analysis*, vol. 42, no. 3, pp. 1246–1285, 2010.

[11] 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.

[12] Peyman Milanfar, *Super-resolution imaging*, CRC Press, 2010.

[13] Leonid I Rudin, Stanley Osher, and Emad Fatemi, "Nonlinear total variation based noise removal algorithms," *Physica D: Nonlinear Phenomena*, vol. 60, no. 1, pp. 259–268, 1992.

[14] Antoni Buades, Bartomeu Coll, and Jean-Michel Morel, "Image denoising methods. a new nonlocal principle," *SIAM review*, vol. 52, no. 1, pp. 113–147, 2010.

[15] Abderrahim Elmoataz, Olivier Lézoray, and Sébastien Bougleux, "Nonlocal discrete regularization on weighted graphs: A framework for image and manifold processing," *IEEE Transactions on Image Processing*, vol. 17, no. 7, pp. 1047–1060, jul 2008.

[16] G. Gilboa and S. Osher, "Nonlocal operators with applications to image processing," *Multiscale Modeling and Simulation*, vol. 7, no. 3, pp. 1005–1028, 2008.

[17] François Malgouyres and Frederic Guichard, "Edge direction preserving image zooming: a mathematical and numerical analysis," *SIAM Journal on Numerical Analysis*, vol. 39, no. 1, pp. 1–37, 2001.

[18] Tony F Chan, Michael K Ng, Andy C Yau, and Andy M Yip, "Superresolution image reconstruction using fast inpainting algorithms," *Applied and Computational Harmonic Analysis*, vol. 23, no. 1, pp. 3–24, 2007.

[19] Antonio Marquina and Stanley J Osher, "Image super-resolution by tv-regularization and bregman iteration," *Journal of Scientific Computing*, vol. 37, no. 3, pp. 367–382, 2008.

[20] Patrick L. Combettes and Jean-Christophe Pesquet, "Proximal splitting methods in signal processing," in *Fixed-Point Algorithms for Inverse Problems in Science and Engineering*, vol. 49, pp. 185–212. Springer, 2011.

[21] A. Chambolle and T. Pock, "A first-order primal-dual algorithm for convex problems with applications to imaging," *Journal of Mathematical Imaging and Vision*, vol. 40, no. 1, pp. 120–145, 2011.

[22] Patrick L Combettes and Jean-Christophe Pesquet, "A proximal decomposition method for solving convex variational inverse problems," *Inverse problems*, vol. 24, no. 6, pp. 065014, 2008.

[23] M. Hidane, O. Lézoray, and A. Elmoataz, "Nonlinear multilayered representation of graph-signals," *Journal of Mathematical Imaging and Vision*, pp. 1–24, 2012.