Image Zoom Completion

Moncef Hidane, Mireille El Gheche, Jean-François Aujol, Yannick Berthoumieu, Charles Deledalle

Image Zoom Completion

Moncef Hidane, Mireille El Gheche, Jean-François Aujol, Yannick Berthoumieu, Charles Deledalle

***Abstract*—We consider the problem of recovering a high-resolution image from a pair consisting of a complete low-resolution image and a high-resolution but incomplete one. We refer to this task as the *image zoom completion* (IZC) problem. After discussing possible contexts in which this setting may arise, we study three regularization strategies, giving full details concerning the numerical optimization of the corresponding energies, and discussing the benefits and shortcomings of each method. As the application that leads us to consider the IZC setting concerns images with strong textural content, we evaluate the performance of the proposed methods on a set of Brodatz textures, comparing the results we get with those obtained with two recent state-of-the-art single-image super-resolution algorithms.**

I. Introduction

Image restoration problems are ubiquitous in the field of image processing [1]. Therein, the goal is to estimate an underlying image from a set of related, degraded, and possibly incomplete measurements. The first step towards this goal is the identification of the imaging device parameters as well as those pertaining to the scene being imaged. Considering a linear forward model, either known *a priori* or properly estimated, with additive noise, the restoration problem can be cast as an inverse problem where the goal is to recover an image $\mathbf{f} \in \mathbb{R}^n$, $n \geqslant 1$, related to measurements $\mathbf{y} \in \mathbb{R}^m$, $m \leqslant n$, through the following forward model

$$\mathbf{y} = \mathbf{A}\mathbf{f} + \boldsymbol{\eta}, \tag{1}$$

where $\mathbf{A} \in \mathbb{R}^{m \times n}$ is a known linear operator and $\boldsymbol{\eta}$ is unknown and accounts for both modeling and sensing errors.

Depending on the nature of the matrix $\mathbf{A}$ in (1), different restoration problems are encountered. In all cases, the very nature of imaging devices leads to problems of the form (1) which are either undetermined ($m < n$) or ill-conditioned (the condition number of $\mathbf{A}$ is very high). A common strategy to deal with such ill-posed inverse problems is through the variational approach. In this setting, specific prior information about the sought-after image is selected by choosing a specific *penalty function* J. An estimated image $\hat{\mathbf{f}}$ is then obtained by minimizing an overall cost function imposing a trade-off between the prior information (through J) and a data fidelity term. The role of the data fidelity term is to control the discrepancy between the estimated image $\hat{\mathbf{f}}$ and the measurements $\mathbf{y}$ according to the forward model (1).

Moncef Hidane is with INSA Centre Val de Loire and with Laboratoire d'Informatique de l'Université de Tours.

Mireille El Gheche is with the IMS and IMB laboratories of Université de Bordeaux.

Jean-François Aujol and Charles Deledalle are with IMB CNRS-Université de Bordeaux.

Yannick Berthoumieu is with Institut Polytechnique de Bordeaux and with the IMS Laboratory.

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 acknowledges the support of the Institut Universitaire de France.

A. Image Zoom Completion

In this paper, we study a specific restoration problem that we term *image zoom completion* (IZC). Here, the problem is to recover a high-resolution (HR) image from a pair consisting of a complete low-resolution (LR) image and a HR but incomplete one. To the best of our knowledge, this problem has been introduced for the first time by the authors of the present paper in [2].

The application that leads us to investigate a solution to the IZC problem concerns the estimation of petro-physical parameters through image processing techniques. In this context, a 3D non-invasive acquisition of a cylindrical sample of a rock is performed. This acquisition is further analyzed in order to automatically segment and classify different regions according to visual cues related to underlying petro-physical parameters. A typical example concerns the automatic classification of different textured regions, which in turn relate to different *porosity levels* of the underlying material. In a practical setting, the resolution provided by this acquisition is not sufficient 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 provides incomplete but higher resolution slices of the object of interest. Now, the next step is to combine the image data coming from both acquisitions in order to obtain a *complete* HR volume.

The setting we have just described may also arise in the context of digital photography when one uses a digital camera to capture an image of a given scene and then uses the *optical zoom* of the same camera to capture a subset of the same scene. The second acquisition is incomplete but it provides higher resolution than the first one since moving the lens further from the sensors array leads a smaller portion of the scene to be seen by the sensors, thus achieving optical zoom. The IZC problem is now to *compute* a *full* HR image with the aid of the complete LR and incomplete HR data. Figure 1 illustrates this setting. The IZC problem we have just described is related to two classical image restoration problems: single-image super-resolution (SR) and image inpainting. We briefly review in the next subsection some approaches related to these two problems. Due to the large body of existing literature we concentrate on specific approaches related to our present proposal.

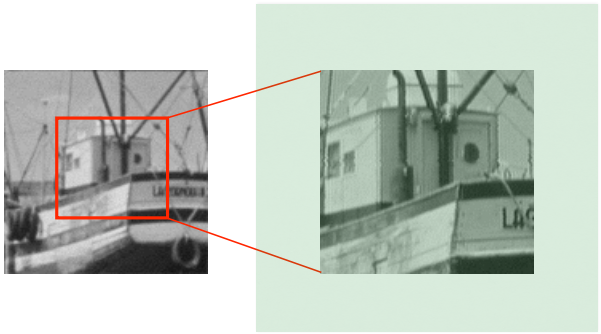

Fig. 1. Setting for the IZC problem. The left image represents a complete LR acquisition. The right image represents a zoom of a subset of the left image and is thus referred to as an *incomplete* HR image. The goal of IZC is to combine both acquisitions in order to complete the missing HR information.

B. Related Work

In the single-image super-resolution setting, the goal is to recover an HR image from a *single* input LR frame. In this context, there are two degradation sources relating the sought-after image to the observed one: blurring and coarse sampling. A possible forward model is given by (1) with $\mathbf{A} = \mathbf{S} \circ \mathbf{H}$ where $\mathbf{H} \in \mathbb{R}^{n \times n}$ is a convolution (or more generally a linear filtering) operator and $\mathbf{S} \in \mathbb{R}^{m \times n}$ is a spatial sub-sampling operator. This problem is under-constrained, thus SR algorithms generally make use of various kinds of prior knowledge.

The baseline method for SR is through linear interpolation methods, especially bicubic interpolation. In order to outperform bicubic interpolation, many variational approaches exploiting smoothness and geometric regularity of images have been proposed (see [3] and references therein). For example, in [4], the authors study the use of the total variation (TV) semi-norm [5] for single-image SR. The use of self-similarity of images patches to regularize the single-image SR problem has been in introduced in [6]. This idea has been pushed further in [7] by using patch redundancy across different scales. Approaches based on image sparsity, either in analysis or synthesis forms [8], [9] or through Gaussian mixtures [10] have also been recently considered. The application of ideas combining sparsity and self-similarity [11], [12] currently leads to the best performing methods.

A different line of work involves sparse representation of LR/HR pairs of patches. This kind of approach has been initiated in [13], [14] where the authors propose to infer the sparse code of each sought-after HR patch from the sparse code of its corresponding observed LR version. To enforce the invariance of LR/HR sparse codes, two coupled LR/HR dictionaries are learned from a large training data set.

While the IZC problem can be seen as a single-image SR problem where part of the HR image is known, it can also be seen as an image completion problem where the complete image is known at a lower resolution. Trying to complete a missing HR image without any further reference to an LR version is an inpainting problem. More precisely, the inpainting problem consists in modifying the image values of the pixels of a given region Ω so that this region does not stand out with respect to its surroundings. The purpose of inpainting might be to restore damaged portions of an image or to remove disturbing unwanted elements present in the image, *e.g.* occlusions.

Most inpainting methods can be classified as being either geometry- or texture-oriented. In the first category [15], [16], [17], interpolation is performed by imposing and continuing a smoothness assumption to the missing part. Approaches in the second category rely directly on a sample of the desired texture to perform the synthesis. The value of a target pixel is copied from the center of a patch in the sample image. The interested reader may consult [18], [19], [20] and references therein.

C. Organization of the Paper

The paper is organized as follows. We recall in Section II the tools we need to carry the numerical optimization of the proposed energies.

We formulate in Section III the IZC problem as an inverse one, and introduce the notations we use throughout the rest of the paper. We also briefly examine the ill-posed nature of the IZC problem.

We adopt in Sections IV, V, VI a variational approach in which we seek a super-resolved and complete image through the minimization of appropriate energy functions combining prior information and data fidelity terms related to the observation model. More specifically, we study *three priors* for the regularization of the IZC problem, leading, in turn, to three solutions. As we proceed, we give full details concerning the numerical optimization of the corresponding energy functions.

More specifically, Section IV is devoted to a first regularization strategy employing the total variation as a penalty function. This solution is targeted towards the class of *piecewise smooth* images with low textural content. Full details about the numerical optimization of the corresponding energy are given. In particular, we show that by using variable splitting and deploying the Douglas-Rachford (DR) algorithm in a product-space, we can devise a fast and inner-loop-free algorithm.

The second method we develop, detailed in Section V, targets classes of images with rich but *homogeneous* textural content. Rather than imposing a sophisticated prior on the sought-after image, we obtain separately the structure and texture components of the super-resolved image, imposing different priors on each component. We adopt the recent model introduced by Schaeffer and Osher in [21], making use of TV penalization on the structure part and a low nuclear norm on the patch-map of the texture part (see Section V). The corresponding energy function is then minimized with the DR algorithm by using variable splitting and rewriting the functional in a product space as was done in Section IV.

In Section VI we develop a regularization strategy targeting images with rich, non-necessarily homogeneous, textural content. To this end, we adopt the self-similarity prior, namely the observation that natural image patches tend to recur several times inside the same image. This third method, termed *nonlocal* hereafter, tries to enforce this prior knowledge by exploiting patch similarities and imposing nonlocal smoothness for the super-resolved image. In practice, only one part

of the HR patches is visible and this makes inferring patch similarities difficult. In order to circumvent this difficulty, we propose to compute patch similarities from an interpolated version of the available LR image. Once this is done, we formulate a functional which makes use of the above nonlocal similarities. We carry out the minimization using a recent primal dual proximal algorithm [22].

We perform a set of numerical simulations in Section VII, discussing both the benefits and shortcomings of each of the proposed methods and comparing our results to two recent state-of-the-art single-image SR algorithms. We finally conclude this paper in Section VIII, pointing out to possible future work.

A preliminary version of this paper has been published in [2]. Therein, the IZC problem is termed "super-resolution from low- and partial high-resolution image pair" and only the nonlocal regularization strategy was developed. Furthermore, the nonlocal regularization presented therein consists in building a pointwise estimate of the image where each missing pixel is connected to a set of pixels that lies in the HR region. In this work, instead, we regularize each pixel by using the most similar pixels in a local neighborhood and the closest ones lying in the HR part.

II. Tools from Convex Analysis and Optimization

We recall in this section some notations and algorithms that we will use for the numerical optimization of the energy functions appearing in the rest of the paper, namely, the concept of proximity operator and the Douglas-Rachford algorithm. We refer the reader to [23] and references therein for further details and proofs.

A. Definitions and Notations

Let $(\mathcal{H}, \langle .,. \rangle)$ be a finite-dimensional inner-product space. The associated norm is denoted by $\|.\| := \sqrt{\langle .,. \rangle}$. A function $f : \mathcal{H} \to]-\infty, +\infty]$ is said to be proper if its domain, $\mathrm{dom}(f) := \{x \in \mathcal{H} : f(x) < +\infty\}$, is nonempty. The function f is said to be convex if its epigraph, $\mathrm{epi}(f) := \{(x, a) \in \mathcal{H} \times \mathbb{R} : f(x) \leqslant a\}$, is convex; it is said to be lower semicontinuous if $\mathrm{epi}(f)$ is closed. The set of all proper convex and lower semicontinuous functions from $\mathcal{H}$ to $]-\infty, +\infty]$ is denoted by $\Gamma_0(\mathcal{H})$.

Let C denote a subset of $\mathcal{H}$. We denote by $\mathrm{ri}(C)$ the relative interior of C, that is, its interior relative to its affine hull. The indicator function of C, denoted ι_C, is defined for all $x \in \mathcal{H}$ by $\iota_C(x) = 0$ if $x \in C$ and $\iota_C(x) = +\infty$ otherwise. Note that $\iota_C \in \Gamma_0(\mathcal{H})$ if and only if C is nonempty, closed and convex.

For $f \in \Gamma_0(\mathcal{H})$ and $z \in \mathcal{H}$, the function $x \in \mathcal{H} \mapsto \frac{1}{2}\|x - z\|^2 + f(x)$ achieves its infimum at a unique point called proximity operator of f at point z and denoted by $\mathrm{prox}_f z$:

$$\mathrm{prox}_f z := \underset{x \in \mathcal{H}}{\mathrm{argmin}} f(x) + \frac{1}{2}\|x - z\|^2. \tag{2}$$

If $f = \iota_C$, then one recovers the definition of the Euclidean convex projection operator on C, denoted proj_C: $\mathrm{proj}_C z := \mathrm{prox}_{\iota_C} z$.

An important property concerning proximal operators is their decomposability in orthonormal bases. In particular, letting $(\mathcal{H}_i)_{1 \leqslant i \leqslant n}$ be a family of Hilbert spaces, $\mathcal{H} = \mathcal{H}_1 \times \ldots \mathcal{H}_n$, $f \in \Gamma_0(\mathcal{H})$, $f_i \in \Gamma_0(\mathcal{H}_i)$ for $i \in \{1, \ldots, n\}$, such that $f(x) = \sum_{i=1}^n f_i(x_i)$, we have $\big(\mathrm{prox}_f(x)\big)_i = \mathrm{prox}_{f_i}(x_i)$ for $i \in \{1, \ldots, n\}$. Such a property makes it possible to devise parallel proximal splitting algorithms as in [24], [25], [26], [27], [28], [29], [22], [30]. We will come back to this point in the sequel.

B. Douglas-Rachford

The strategy we adopt in this paper towards the minimization of ℓ_1-based energy functions (Sections IV and V) relies on variable splitting and employing the DR algorithm in a product space. Consider the following variational problem:

$$\underset{x \in \mathcal{H}}{\mathrm{minimize}} \quad F_1(x) + F_2(x), \tag{3}$$

- $F_1, F_2 \in \Gamma_0(\mathcal{H})$;
- $\mathrm{ri}(\mathrm{dom} F_1) \cap \mathrm{ri}(\mathrm{dom} F_2) \neq \varnothing$;
- $\lim_{\|x\| \to +\infty} F_1(x) + F_2(x) = +\infty$.

Then, the set of minimizers of $F_1 + F_2$ over $\mathcal{H}$ is nonempty. Furthermore, letting $(x_n)_n, (y_n)_n \in \mathcal{H}^{\mathbb{N}}$ constructed as follows

$$\begin{cases} y_0 \in \mathcal{H} \\ x_n = \mathrm{prox}_{\nu F_2}(y_n) \\ y_{n+1} = y_n + \xi \left(\mathrm{prox}_{\nu F1}(2x_n - y_n) - x_n\right) \end{cases} \tag{4}$$

with $0 < \xi < 2$ and $\nu > 0$, the sequence (x_n) converges to a minimizer x^* of $F_1 + F_2$ over $\mathcal{H}$.

For the minimization of the linearly-constrained quadratic nonlocal energy appearing in Section VI we use the accelerated projected gradient introduced in [31].

We are now ready to tackle the IZC problem. The next section is dedicated to the formalization of the problem.

III. The Image Zoom Completion Problem

The setting for IZC is a complete LR image and a corresponding incomplete HR image. Let $\mathbf{y}^{(1)} \in \mathbb{R}^p$ and $\mathbf{y}^{(2)} \in \mathbb{R}^n$, $p < n$, respectively denote the lexicographical ordering of the complete LR and partial HR images. We adopt the following discrete forward model:

$$\begin{cases} \mathbf{y}^{(1)} &= \mathbf{S}\mathbf{H}\mathbf{f}_0 + \boldsymbol{\eta}_1 \in \mathbb{R}^p, \\ \mathbf{y}^{(2)} &= \mathbf{M}(\mathbf{f}_0 + \boldsymbol{\eta}_2) \in \mathbb{R}^n, \end{cases} \tag{5}$$

where

- $\mathbf{f}_0 \in \mathbb{R}^n$ denotes the unknown HR image;
- $\mathbf{S} \in \mathbb{R}^{p \times n}$ stands for spatial downsampling by a factor r in each direction ($n = p\,r^2$); Note that $\mathbf{S}^\top \in \mathbb{R}^{n \times p}$ corresponds to upsampling by the same factor and that the matrix $\mathbf{S}^\top\mathbf{S} \in \mathbb{R}^{n \times n}$ is diagonal with binary diagonal elements;
- $\mathbf{H} \in \mathbb{R}^{n \times n}$ accounts for spatial blurring of the image, modeled in our case by a discrete circular convolution with a *known* point spread function $\mathbf{h}$: $\mathbf{H}\mathbf{x} = \mathbf{h} \circledast \mathbf{x}$;
- $\mathbf{M} = \mathrm{diag}(m_1, \ldots, m_n) \in \mathbb{R}^{n \times n}$ is a binary HR mask indicating which HR pixels are observed: $m_i = 1$, if

pixel i belongs to the observed area of the HR image, and $m_i = 0$ otherwise;
- the vectors $\boldsymbol{\eta}_1 \in \mathbb{R}^p$ and $\boldsymbol{\eta}_2 \in \mathbb{R}^n$ are samples of a white Gaussian noise accounting for acquisition and model noise.

Throughout the paper, we assume that the forward model (5) relating the HR/LR pair to the sought-after image $\mathbf{f}_0$ is completely known. In particular, we discuss neither the identification of the convolution kernel nor possible registration issues when the partial zoom is performed.

Under the setting described above, the IZC problem corresponds to the recovery of an estimate $\hat{\mathbf{f}}$ of $\mathbf{f}_0$ from the measurements $\mathbf{y}^{(1)}$ and $\mathbf{y}^{(2)}$, according to the forward model (5). Depending on the downsampling factor r, the support of the blurring kernel and the area of the visible HR part, this problem can be ill-posed, with more unknowns than equations. In all cases, the presence of the convolution operator makes it ill-conditioned.

As usual for such inverse problems in imaging [1], we formulate the estimation task in a variational setting leading to the minimization of an energy function of the form $E(\mathbf{f}) = R(\mathbf{f}) + D(\mathbf{f})$. The function R is a regularization term forcing the solutions to have pre-specified properties, while the term $D(\mathbf{f})$ penalizes the discrepancy between $\mathbf{f}$ and the observations $\mathbf{y}^{(1)}, \mathbf{y}^{(2)}$, according to the forward model (5). In the presence of Gaussian noise, the latter term is usually taken, in its penalized form, as the squared ℓ_2 distance, leading to an optimization problem [1] of the form

$$\underset{\mathbf{f}\in\mathbb{R}^n}{\text{minimize}} \quad R(\mathbf{f}) + \lambda_1\|\mathbf{SHf} - \mathbf{y}^{(1)}\|^2 + \lambda_2\|\mathbf{Mf} - \mathbf{y}^{(2)}\|^2. \quad (6)$$

When the noise level in the observed part of the HR image can be neglected, or if one does not want to modify the observed HR part, one can instead consider the constrained problem

$$\underset{\mathbf{f}\in\mathbb{R}^n}{\text{minimize}} \quad R(\mathbf{f}) + \lambda\|\mathbf{SHf} - \mathbf{y}^{(1)}\|^2 + \iota_{\{\mathbf{M}\cdot=\mathbf{y}^{(2)}\}}(\mathbf{f}), \quad (7)$$

where we write $\{\mathbf{A}\cdot = \mathbf{y}\}$ for the set $\{\mathbf{x} \in \mathbb{R}^n : \mathbf{Ax} = \mathbf{y}\}$ and $\lambda > 0$. This latter setting is the one we adopt in the rest of the paper.

To summarize, we tackle the IZC problem by minimizing an energy function of the form given in (7). The problem now amounts to devising suitable regularizers R as well as the practical optimization of the corresponding energies. We begin by investigating the use of the TV prior.

IV. IZC via TV Regularization

Image restoration techniques should be able to preserve image edges. When adopting a variational setting, this can be achieved by penalizing the total variation of the sought-after image. First let us introduce relevant notations. Let

$$\sigma : \{1, \dots, \sqrt{n}\} \times \{1, \dots, \sqrt{n}\} \to \{1, \dots, n\}$$

denote the lexicographical pixel enumeration[2]. We consider the discrete gradient operator $\nabla : \mathbb{R}^n \to \mathbb{R}^{n\times 2}$ given, for $1 \leqslant k, l \leqslant \sqrt{n} - 1$ by

$$\begin{cases} (\nabla\mathbf{f})_{\sigma(l,k),1} &= \mathbf{f}_{\sigma(l,k+1)} - \mathbf{f}_{\sigma(l,k)}, \\ (\nabla\mathbf{f})_{\sigma(l,k),2} &= \mathbf{f}_{\sigma(l+1,k)} - \mathbf{f}_{\sigma(l,k)}. \end{cases} \quad (8)$$

We adopt *circular boundary conditions* for $k = \sqrt{n}$ and/or $l = \sqrt{n}$ so that $-\nabla^\top\nabla$ corresponds to the 4-stencil discretization of the Laplacian operator with circular boundary conditions, and is thus diagonal in the discrete Fourier domain. Here $\nabla^\top$ denotes the adjoint of ∇, relatively to the standard inner products on $\mathbb{R}^n$ and $\mathbb{R}^{n\times 2}$.

We write $\|\cdot\|_{1,2}$ for the following norm

$$(\forall \mathbf{p} \in \mathbb{R}^{n\times m}) \quad \|\mathbf{p}\|_{1,2} = \sum_{i=1}^{n} \|\mathbf{p}_{i,\cdot}\|_2 = \sum_{i=1}^{n} \sqrt{\sum_{j=1}^{m} p_{i,j}^2}. \quad (9)$$

We adopt the following definition for the discrete total variation of an image $\mathbf{f} \in \mathbb{R}^n$

$$J(\mathbf{f}) = \|\nabla\mathbf{f}\|_{1,2}. \quad (10)$$

Adopting the total variation as a prior for our IZC problem leads to the following convex optimization problem

$$\underset{\mathbf{f}\in\mathbb{R}^n}{\text{minimize}} \quad \|\nabla\mathbf{f}\|_{1,2} + \frac{\lambda}{2}\|\mathbf{SHf} - \mathbf{y}^{(1)}\|^2 + \iota_{\{\mathbf{M}\cdot=\mathbf{y}^{(2)}\}}(\mathbf{f}). \quad (11)$$

A. A DR Algorithm in a Product Space

The energy in (11) is composite: it mixes the operators $\mathbf{M}, \mathbf{S}, \mathbf{H}$, and it contains two nonsmooth terms. In order to take advantage of the properties of the involved operators, we adopt the strategy suggested in [32] by using variable splitting and working in a product space. For the case of TV regularization, we show that this approach makes it possible to devise a fast iterative algorithm with *no inner loops*.

Letting $\mathcal{H} = \mathbb{R}^{n\times 2} \times \mathbb{R}^n \times \mathbb{R}^n$ and $K = \{(\nabla\mathbf{f}, \mathbf{Hf}, \mathbf{f}), \mathbf{f} \in \mathbb{R}^n\}$, we rewrite (11) in the form

$$\underset{\mathbf{x}=(\mathbf{p},\mathbf{u},\mathbf{f})\in\mathcal{H}}{\text{minimize}} \quad F_1(\mathbf{x}) + F_2(\mathbf{x}), \quad (12)$$

where

$$F_1(\mathbf{x}) = \|\mathbf{p}\|_{1,2} + \frac{\lambda}{2}\|\mathbf{Su} - \mathbf{y}^{(1)}\|^2 + \iota_{\{\mathbf{M}\cdot=\mathbf{y}^{(2)}\}}(\mathbf{f}), \quad (13)$$

and

$$F_2(\mathbf{x}) = \iota_K(\mathbf{p}, \mathbf{u}, \mathbf{f}). \quad (14)$$

Note that $F_1, F_2 \in \Gamma_0(\mathcal{H})$. In order to apply DR, we need to evaluate $\text{prox}_{\nu F_1}$ and $\text{prox}_{\nu F_2}$, $\nu > 0$, at each iteration. Let us start with the evaluation of $\text{prox}_{\nu F_1}$.

Since F_1 is separable, the evaluation of $\text{prox}_{\nu F_1}(\mathbf{p}, \mathbf{u}, \mathbf{f})$ amounts to evaluating each part separately. The first component, $\text{prox}_{\nu\|\cdot\|_{1,2}}$, corresponds to the vector-field *soft-thresholding* function given, for all $\mathbf{p} \in \mathbb{R}^{n\times 2}$, by

$$(\text{ST}_\nu\mathbf{p})_{i,j} = \begin{cases} 0 & \text{if } \|\mathbf{p}_{i,\cdot}\|_2 \leqslant \nu, \\ \left(1 - \frac{\nu}{\|\mathbf{p}_{i,\cdot}\|}\right) p_{i,j} & \text{else.} \end{cases} \quad (15)$$

[1] We suppose that $\boldsymbol{\eta}_1$ and $\boldsymbol{\eta}_2$ are the realizations of two *independent* Gaussian vectors.

[2] For simplicity of notations, we work with square images.

The evaluation of $\text{prox}_{\nu\|\mathbf{S}\cdot-\mathbf{y}_1\|^2}$ amounts to solving a system of linear equations whose matrix is given by $\mathbf{A}_1 = \frac{\lambda_1}{\nu}\mathbf{S}^\top\mathbf{S} + \mathbf{I}$. This computation can be done in linear time since $\mathbf{A}_1$ is diagonal.

Finally, the evaluation of the third component of the proximity operator of F_1 is immediate since its corresponds to projecting on the set $\{\mathbf{M}\cdot = \mathbf{y}^{(2)}\}$:

$$\left(\text{proj}_{\{\mathbf{M}\cdot=\mathbf{y}^{(2)}\}}(\mathbf{f})\right)_i = \begin{cases} y_i^{(2)} & \text{if } m_i = 1, \\ f_i & \text{if } m_i = 0. \end{cases} \tag{16}$$

The evaluation of $\text{prox}_{\nu F_2}$ amounts to projecting on the constraint set K. It is straightforward to see that it leads to a system of linear equations whose matrix is $\mathbf{A}_2 = \nabla^\top\nabla + \mathbf{H}^\top\mathbf{H} + \mathbf{I}$. Due to the circular boundary choice for both ∇ and $\mathbf{H}$, the matrix $\mathbf{A}_2$ is 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.

B. Discussion

The results obtained by the TV approach we have just introduced are further discussed in Section VII. We discuss in this subsection the qualitative behavior of the TV-regularized solutions to the IZC problem along with the iteration complexity of the minimization we have just derived.

From a computational point of view, the complexity of a single iteration of the algorithm we have just derived is dominated by the projection on the constraints set K. As discussed above, due to the properties of the matrix $\mathbf{A}_2$, this is done by applying Fourier methods, leading to a direct and an inverse Fourier transform at each iteration.

From a qualitative point of view, the TV-regularized solution suffers from the usual straircasing effect typical of TV regularization. This is the case even for the simulations we carried with the exact knowledge of the blurring matrix $\mathbf{H}$ and without addition of noise. This model is thus relevant only for the class of *piecewise smooth images*. As we target reconstructing larger classes of images, typically with mixed patterns, we now describe a first step towards dealing with more textured images.

V. IZC via TV + Low-Patch-Rank Decomposition

Rather than trying to impose a single prior for the sought-after HR image, as was done in the previous section, we adopt in this section a decomposition model for the IZC problem. The idea here is to compute separately two different components of the HR image by imposing different priors on each. The resulting image we obtain is then given by the sum of these two components. More specifically, we adopt the *structure-texture decomposition framework* [33], [34], [35], [36], [37]. In this context, an image $\mathbf{f} \in \mathbb{R}^n$ is decomposed in the form $\mathbf{f} = \mathbf{u} + \mathbf{v}$, where the structure component $\mathbf{u}$ contains only salient image structures while the texture component $\mathbf{v}$ contains the "oscillating patterns" present in $\mathbf{f}$.

Structure-texture decompositions can be obtained by solving variational problems of the form

$$\begin{aligned} \underset{(\mathbf{u},\mathbf{v})\in\mathbb{R}^n\times\mathbb{R}^n}{\text{minimize}} \quad & F_1(\mathbf{u}) + F_2(\mathbf{v}) \\ \text{subject to} \quad & \mathbf{f} = \mathbf{u} + \mathbf{v}, \end{aligned} \tag{17}$$

for appropriate choices of F_1 and F_2, namely:

- $F_1(\mathbf{u}) \ll F_2(\mathbf{u})$ if $\mathbf{u}$ is piecewise constant;
- $F_2(\mathbf{v}) \ll F_1(\mathbf{v})$ if $\mathbf{v}$ is composed mainly of oscillating textures.

While the choice for F_1 has now settled to the TV functional, several works have investigated different choices for F_2. Of particular interest to us in the present paper is the TV+low-patch-rank decomposition model of [21] which we now briefly describe.

A. TV+Low-Patch-Rank Decomposition Model

This model introduces a new formulation for the texture penalization term F_2 in (17). More precisely, due to the repetitive aspect of natural textures, it is reasonable to assume that some patches of the texture part $\mathbf{v}$ of an image $\mathbf{f}$ will repeat themselves at different locations. In order to integrate this observation in a variational setting, the authors of [21] propose to extract all $\sqrt{q}\times\sqrt{q}$ non-overlapping patches of an image $\mathbf{f}\in\mathbb{R}^n$, placing each of them as the column of a matrix $\mathcal{P}\mathbf{f}\in\mathbb{R}^{q\times m}$. Here, the *patch-map* $\mathcal{P}:\mathbb{R}^n\to\mathbb{R}^{q\times m}$, $n = qm$, allows to pass from an image in $\mathbf{f}\in\mathbb{R}^n$ to a matrix $\mathcal{P}\mathbf{f}\in\mathbb{R}^{q\times m}$ where q is the number of pixels inside a patch and m is the number of $\sqrt{q}\times\sqrt{q}$ non-overlapping patches contained in $\mathbf{f}$. Let us note at this point that, due to the fact that $\mathcal{P}$ is essentially a permutation of image pixels, it is an isometry with respect to all point-wise norms on $\mathbb{R}^{q\times m}$. This fact will be important for deriving our minimization algorithm.

Equipped with the patch-map $\mathcal{P}$, the authors of [21] propose the following decomposition model

$$\begin{aligned} \underset{(\mathbf{u},\mathbf{v})\in\mathbb{R}^n\times\mathbb{R}^n}{\text{minimize}} \quad & \lambda TV(\mathbf{u}) + \|\mathcal{P}\mathbf{v}\|_* \\ \text{subject to} \quad & \mathbf{f} = \mathbf{u} + \mathbf{v}, \end{aligned} \tag{18}$$

where $\|\mathbf{A}\|_*$ denotes the nuclear (or trace) norm of a matrix $\mathbf{A}$, that is, the sum of its singular values. In this context, the nuclear norm acts as a convex relaxation for the combinatorial rank function. Thus, adopting the term $\|\mathcal{P}\mathbf{v}\|_*$ favors an overall repetition of texture patches. Further details can be found in [21].

In [21], the decomposition model (18) is applied to denoising, deblurring and sparse reconstruction. The authors use the split Bregman algorithm [38] which is closely related to the algorithm we use in this paper. Let us finally note that a generalized version of the decomposition model of [21], where patches are allowed to overlap, has been recently proposed in [39].

B. TV+Low-Patch-Rank for IZC

We adopt in this subsection the following model for constructing a solution to our IZC problem:

$$\underset{(\mathbf{u},\mathbf{v})\in\mathbb{R}^n\times\mathbb{R}^n}{\text{minimize}} \quad \alpha TV(\mathbf{u}) + \beta\|\mathcal{P}(\mathbf{v})\|_* + \frac{\lambda}{2}\|\mathbf{S}\mathbf{H}(\mathbf{u}+\mathbf{v}) - \mathbf{y}^{(1)}\|^2 + \iota_{\{\mathbf{M}\cdot=\mathbf{y}^{(2)}\}}(\mathbf{u}+\mathbf{v}), \tag{19}$$

with $\alpha, \beta, \lambda > 0$.

The advantage of using (19) instead of the TV-regularized problem (11) is that the additional term $\|\mathcal{P}(\mathbf{v})\|_*$ allows us to have better control on the texture part of the sought-after HR image, thus partly avoiding the staircasing effect mentioned above. Let us note that due to the constraint imposed by the term $\iota_{\{\mathbf{M}\cdot=\mathbf{y}^{(2)}\}}(\mathbf{u}+\mathbf{v})$, we expect the visible part of HR image to be decomposed into a structure-texture image pair. Forcing the texture component of the sought-after HR image to have an overall low patch-rank will favor invisible HR patches to be linear combinations of visible HR ones. This can be seen as an implicit strategy for making use of the available HR image in order to complete the unobserved part. We now detail our approach to the minimization of (19).

C. A DR Algorithm

We work in the Hilbert space $\mathcal{H} = \mathbb{R}^{n\times 2}\times\mathbb{R}^n\times\mathbb{R}^n\times\mathbb{R}^n\times\mathbb{R}^n$. Let

$$K = \{(\nabla\mathbf{u}, \mathbf{u}, \mathbf{v}, \mathbf{H}(\mathbf{u}+\mathbf{v}), \mathbf{u}+\mathbf{v}) \in \mathcal{H} \,:\, \mathbf{u}\in\mathcal{X}, \mathbf{v}\in\mathcal{X}\},$$

and for $\mathbf{x} = (\mathbf{p}, \mathbf{u}, \mathbf{v}, \mathbf{g}, \mathbf{w}) \in \mathcal{H}$

$$F_1(\mathbf{x}) = \alpha\|\mathbf{p}\|_{1,2} + \beta\|\mathcal{P}(\mathbf{v})\|_* + \frac{\lambda}{2}\|\mathbf{S}\mathbf{g} - \mathbf{y}^{(1)}\|^2 + \iota_{\{\mathbf{M}\cdot=\mathbf{y}^{(2)}\}}(\mathbf{w}), \tag{20}$$

$$F_2(\mathbf{x}) = \iota_K(\mathbf{x}). \tag{21}$$

The problem is now to minimize $F_1 + F_2$ over $\mathcal{H}$. In order to apply DR, we need to evaluate $\text{prox}_{\nu F_1}$ and $\text{prox}_{\nu F_2}$, $\nu > 0$, at each iteration. Let us start with the evaluation of $\text{prox}_{\nu F_1}$.

As in Section IV-A, F_1 is separable, thus the evaluation of $\text{prox}_{\nu F_1}(\mathbf{p}, \mathbf{u}, \mathbf{v}, \mathbf{g}, \mathbf{w})$ amounts to evaluating each part separately. The evaluation of the first component corresponds to the soft-thresholding function defined in (15). The second component corresponds to the identity since F_1 does not depend on $\mathbf{u}$. The evaluation of the fourth component involves solving a linear system with a diagonal matrix. The evaluation of the last component is given in (16).

The evaluation of the third proximity operator is more intricate since it involves the pre-composition with $\mathcal{P}$ and the nuclear norm. The generic problem is to compute $\text{prox}_{\mu\|\mathcal{P}\cdot\|_*}$ for $\mu > 0$. Due the fact that $\mathcal{P}$ is an isometry with respect to all point-wise norms, we have (Proposition 11 of [40]) that

$$\text{prox}_{\mu\|\mathcal{P}(\cdot)\|_*}(\mathbf{v}) = \mathcal{P}^\top \text{prox}_{\mu\|\cdot\|_*}\mathcal{P}(\mathbf{v}) \tag{22}$$

It has been shown in [41] that the latter prox computation can be done by singular value thresholding:

$$\text{prox}_{\mu\|\cdot\|_*}(\mathbf{v}) = SVT_\mu(\mathbf{v}) := \mathbf{U}\max(\mathbf{\Sigma} - \mu\mathbf{I}, 0)\mathbf{V}^*, \tag{23}$$

where $\mathbf{v} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^*$ is a singular value decomposition (SVD) of $\mathbf{v}$.

Regarding the evaluation of $\text{prox}_{\nu F_2}$, it is easy to see that it leads to a system of linear equations with matrix

$$A_3 := \left[\begin{array}{c|c} \nabla^\top\nabla + \mathbf{H}^\top\mathbf{H} + 2\mathbf{I} & \mathbf{H}^\top\mathbf{H} + \mathbf{I} \\ \hline \mathbf{H}^\top\mathbf{H} + \mathbf{I} & \mathbf{H}^\top\mathbf{H} + 2\mathbf{I} \end{array}\right]. \tag{24}$$

While each block of $\mathbf{A}_3$ is diagonal in the Fourier domain, the overall matrix is not. We instead solve this linear system with a conjugate gradient method.

D. Discussion

From a computational point of view, the iteration complexity of the algorithm we have derived is dominated, on the one hand, by the computation of the singular value decomposition of the patch matrix, and on the other hand, by the conjugate gradient iterations used to solve a linear system whose matrix is $\mathbf{A}_3$ given in (24).

Concerning the SVD computation, it is important to note that it applies to a matrix with much fewer rows (numbers of pixels in a patch) than columns (number of non-overlapping patches).

Concerning the conjugate gradient solver, in practice we perform only 4 iterations and use a 'warm start' strategy [42], meaning that we start the conjugate gradient solver at iteration $k+1$ with the result it yielded from iteration k. Each conjugate gradient iteration involves a matrix-vector product $\mathbf{A}_3\mathbf{x}$, so it involves the evaluation of one Laplacian filter ($\nabla^\top\nabla$) and one low-pass filter ($\mathbf{H}^\top\mathbf{H}$).

As it will be shown in Section VII, when the sought-after image contains different types of texture, the global constraint on the rank of the patch map is not sufficient to recover these textures.

Next section is devoted to a nonlocal regularization approach targeting images with rich, possibly non-homogeneous textures.

VI. IZC via Nonlocal Regularization

A. The Nonlocal Framework

Classical variational techniques employed in the field of image processing rely on the regularity of the underlying image in terms of *local* relationships between neighboring pixels. Nonlocal regularization techniques [43], [44], [45], [46], [47], [48], [49], [50], [51], [52], [53], [54], [55], [56] replace the local regularity assumption by the self-similarity prior for natural images. This recent trend of ideas is inspired by the nonlocal-means algorithm for image denoising [49].

The exploitation of the self-similarity hypothesis of natural images in the context of inverse problems relies on its incorporation in a variational setting. In the context of denoising, the work of Kindermann, Osher, and Jones [43] interpreted the class of neighborhood filters as regularization based on nonlocal functionals. Equipped with those functionals, general inverse problems can be tackled; for instance, the work of [43] already considered its applications to image deblurring.

A systematic framework for nonlocal regularization has been proposed in [44], [45] where nonlocal versions of Laplacian and TV regularizations are formulated. While the functionals of [44], [45] are continuous, their discretization

Fig. 2. Notations for the nonlocal regularization of the IZC problem.

leads to the discrete graph regularization framework of [46], [47], [48] which allows to tackle the problem of denoising and smoothing signals defined on irregular domains, as long as a sensible neighborhood relation can be found. This latter line of works also considers the larger class of p-Laplacians on graphs and the diffusion processes associated with them. Further references about the nonlocal approach, targeted specifically towards the regularization of inverse problems in image processing, can be found in [50], [51], [52], [53], [54], [55], [57], [58], [59], [60], [61].

Non-Local Total Variation (NLTV) is known to preserve textures, details and fine structures better than the standard TV. This arises from the fact that TV is based on the discrete gradient, while NLTV is based on a discrete difference operator whose orientations are driven by the image itself. Such directions are chosen for each pixel independently, based on the similarity between their neighbors.

In the following, we will consider the classical $\ell_{1,2}$-NLTV, expressed as

$$\text{NLTV}(\mathbf{f}) = \sum_{i\in\Omega} \sqrt{\sum_{j\in\mathcal{N}_i} w_{i,j}\|f_i - f_j\|^2}, \tag{25}$$

where $\Omega = \{1,\dots,n\}$ denotes the set of the image pixels and $\mathcal{N}_i \subset \mathcal{W}_i$ is a subset of positions located into a $Q \times Q$ window $\mathcal{W}_i \subset \Omega\backslash\{i\}$ centered at i. For every $i \in \Omega$, we design the support $\mathcal{N}_i$ by selecting the pixels $j \in \mathcal{W}_i$ that are most similar to i according to a Euclidean distance between patches surrounding the pixels i and j: $\|p_i(\mathbf{f}) - p_j(\mathbf{f})\|$. The corresponding edges are weighted according to

$$w_{i,j} = e^{-\|p_i(\mathbf{f})-p_j(\mathbf{f})\|^2/2\sigma^2}, \tag{26}$$

where $\sigma > 0$ and $p_i(\mathbf{f}) \in \mathbb{R}^q$ denotes an image patch extracted from $\mathbf{f}$ and centered at i.

B. Nonlocal regularization for IZC

In order to enforce the spatial coherence of pixels, we first need to define a similarity graph of the HR image. As only a part of the HR image is sensed, the information about all HR patches is incomplete, and thus it cannot be used to build the graph. In order to circumvent this difficulty, we take advantage of the available LR image $\mathbf{y}^{(1)} \in \mathbb{R}^p$. The first step in this strategy is to interpolate $\mathbf{y}^{(1)} \in \mathbb{R}^p$ to match the definition of $\mathbf{y}^{(2)} \in \mathbb{R}^n$. To this end, we use bicubic interpolation and get $\tilde{\mathbf{y}}^{(1)} \in \mathbb{R}^n$. Once we have $\tilde{\mathbf{y}}^{(1)}$, the similarities between patches can be estimated by the following two different approaches.

The first approach is based on the hypothesis that the interpolated patches generally exhibit similar spectral structure and maintain the same coherence. So, for each missing pixel we compute the weights on patches lying in a local neighborhood. For the regularization, we use the function introduced in Eq. (25). The NLTV regularization can be expressed as the $\ell_{1,2}$-norm composed with a discrete difference operator, yielding

$$\text{NLTV}(\mathbf{f}) = \|\mathbf{L}_1\mathbf{f}\|_{1,2}, \tag{27}$$

where

$$\mathbf{L}_1\mathbf{f} = \begin{bmatrix} \big[w_{1,j}(f_1 - f_j)\big]_{j\in\mathcal{N}_1} \\ \vdots \\ \big[w_{n,j}(f_n - f_j)\big]_{j\in\mathcal{N}_n} \end{bmatrix} \begin{matrix} \} \in \mathbb{R}^k \\ \vdots \\ \} \in \mathbb{R}^k. \end{matrix} \tag{28}$$

The second approach consists of adding to the previous one a new set of connections. To do so, we denote the set of pixels where the HR information is missing by $D \subset \Omega$ and the set of sensed HR pixels by D^c (see Figure 2). We connect each unobserved HR pixel $i \in D$ to its k-nearest observed HR neighbors in D^c. The new nonlocal gradient operator is expressed as

$$\mathbf{L}_2\mathbf{f} = \begin{bmatrix} \big[w^c_{1,j}(f_1 - y^{(2)}_j)\big]_{j\in\mathcal{N}^c_1} \\ \vdots \\ \big[w^c_{n,j}(f_n - y^{(2)}_j)\big]_{j\in\mathcal{N}^c_n} \end{bmatrix} \begin{matrix} \} \in \mathbb{R}^k \\ \vdots \\ \} \in \mathbb{R}^k, \end{matrix} \tag{29}$$

where $\forall i \in \Omega$, $\mathcal{N}^c_i$ is a subset of positions in D_c, $w^c_{i,j}$ are the weights computed between the patches $p_i(\tilde{y}^{(1)})$ and $p_j(y^{(2)})$. Therefore, the second regularizer is defined in function of $\mathbf{W} = [\mathbf{L}_1^\top \ \mathbf{L}_2^\top]^\top$, leading to

$$\text{NLTV}^+(\mathbf{f}) = \|\mathbf{W}\mathbf{f}\|_{1,2}. \tag{30}$$

Incorporating data consistence terms in the form of a penalization for the LR image and a constraint for the HR one leads to the problem

$$\underset{\mathbf{f}\in\mathbb{R}^n}{\text{minimize}} \quad \|\mathbf{W}\mathbf{f}\|_{1,2} + \frac{\lambda}{2}\|\mathbf{S}\mathbf{H}\mathbf{f} - \mathbf{y}^{(1)}\|^2 + \iota_{\{\mathbf{M}\cdot = \mathbf{y}^{(2)}\}}(\mathbf{f}). \tag{31}$$

Therefore, we use an interpolated version $\tilde{\mathbf{y}}^{(1)}$ of the LR image $\mathbf{y}^{(1)}$ to infer a weighted adjacency relations and use it to infer the values of the missing pixels by solving (31). Note that each node in D has *at least* k neighbors in a local region of Ω and *at least* k neighbors in D^c.

C. Minimization Using FBPD

The problem of Eq. (31) can be solved using proximal algorithms, such as the Forward-backward Primal Dual method reported in Algorithm 1. Such a method requires: the proximity operator of the NLTV discussed in Eq. (15); the projection onto the constraint set reviewed in Eq. (16); the gradient of $g(\mathbf{f}) = \frac{\lambda}{2}\|\mathbf{S}\mathbf{H}\mathbf{f} - \mathbf{y}^{(1)}\|^2$ given by

$$\nabla g(\mathbf{f}) = \lambda\mathbf{H}^\top\mathbf{S}^\top\mathbf{S}\mathbf{H}\mathbf{f} - \lambda\mathbf{H}^\top\mathbf{S}^\top\mathbf{y}^{(1)}, \tag{32}$$

whose Lipschitz constant is equal to λ.[3]

[3]Since $\mathbf{S}^\top\mathbf{S}$ is diagonal with binary entries, we have that $|||\mathbf{S}^\top\mathbf{S}||| = 1$. Using a normalized kernel $\mathbf{H}$ leads to $|||\mathbf{H}^\top||| = |||\mathbf{H}||| = 1$. Where $|||\cdot|||$ denotes the spectral norm.

Algorithm 1 FBPD method [22]

INITIALIZATION

$$\left\lfloor \begin{array}{l} \text{choose } \left(x^{[0]}, y^{[0]}\right) \in \mathbb{R}^n \times \mathbb{R}^{2kn} \\ \text{set } \tau > 0 \text{ and } \omega > 0 \text{ such that} \\ \qquad \tau\left(\lambda/2 + \omega\|\mathbf{W}\|^2\right) < 1 \end{array} \right.$$

FOR $l = 0, 1, \ldots$

$$\left\lfloor \begin{array}{l} \widehat{x}^{[l]} = \nabla g(x^{[l]}) + \mathbf{W}^\top y^{[l]} \\ x^{[l+1]} = \text{proj}_{\{\mathbf{M}\cdot = \mathbf{y}^{(2)}\}}\left(x^{[l]} - \tau\, \widehat{x}^{[l]}\right) \\ \widehat{y}^{[l]} = \mathbf{W}\left(2x^{[l+1]} - x^{[l]}\right) \\ y^{[l+1]} = \text{prox}_{\omega\|\cdot\|_{1,2}}\left(y^{[l]} + \omega\, \widehat{y}^{[l]}\right) \end{array} \right.$$

If the set of the solutions to Problem (31) is nonempty, then any sequence $(x^{[l]})_{l\in\mathbb{N}}$ generated by Algorithm 1 converges to an element of this set (under a suitable weak qualification condition).

D. Discussion

The results obtained by the nonlocal approach we have just introduced are further discussed and compared to the TV-based and decomposition-based approaches of Sections VII and V, respectively. In this subsection, we discuss the iteration complexity of the minimization we have just derived. We also discuss two foreseeable shortcomings of the proposed nonlocal approach.

From a computational point of view, the iteration complexity of the above algorithm depends on the evaluation of $\nabla g(\mathbf{f})$ and $\mathbf{W}\mathbf{f}$. The former is dominated by the application of two low-pass filters ($\mathbf{H}$ and $\mathbf{H}^\top$). The latter depends on the sparsity of $\mathbf{W}$, which is related to the number of unobserved HR pixels and to the number k of nearest-neighbors. Furthermore, there is an overhead for computing the patch similarities between the unobserved and observed HR parts.

As described earlier, patch similarities that drive the nonlocal regularization are based on (an interpolated version of) the LR image. When the super-resolution factor r is important and/or when the blurring matrix attenuates too much the high frequencies of the HR scene, these weights can be erroneous and their incorporation into the regularization functional can lead to bad reconstructions. While a solution to the first problem (large super-resolution factor) can consist in performing zoom completion recursively using a dyadic factor, the only solution to the second problem (strong attenuation of high frequencies) is to recompute the weights after a fixed number of iterations.

Another case where the nonlocal approach we proposed may fail is when the typical patterns present in the visible HR part are different from the ones that make the invisible part. In this case, the nearest-neighbors graphs provides very limited information and any use of the corresponding weights may lead to bad reconstructions.

VII. Experimental Results

We evaluate in this section the effectiveness of the three methods we have introduced on a set of 256×256 images. As the application that lead us to consider the IZC setting concerns images with strong textural content, we decided to test our methods on a set of 8 images taken from the Brodatz image data set (see Fig. 3).

In order to assess the relative performances of the three approaches detailed in this paper, we start with a ground truth full HR image f_0 and simulate the incomplete image $y^{(2)}$ by masking part of f_0. The binary mask used to hide a part of the HR image is shown in Figures 4, 5 and 6. The observed HR pixels amount to 25% of the total pixels. Similarly, we simulate the complete LR image $y^{(1)}$ by blurring f_0, downsampling the result by a factor r in each direction, and finally adding white Gaussian noise. In all the experiments we carried, the blurring kernel corresponds to a normalized Gaussian $e^{-d^2/2s^2}$ with $s = 1.2$. The kernel is truncated to a 5×5 window. The super-resolution factor r is fixed to 2. Computations on images are done in the range $[0, 255]$. The standard deviation of the additive white Gaussian noise is fixed to 2.5.

The results we report are obtained by optimizing the parameter λ in (11), (19), (31), as well as α, β, γ in (19) in black box manner, each time maximizing the peak signal to noise ration (PSNR) to the ground truth image. For the nonlocal approach detailed in Section VI, three additional parameters are involved: the size of compared patches, the number k of nearest neighbors, and the parameter σ in (26). In our experiments, similarities between pixels were computed based on 5×5 surrounding patches. The number k has been fixed to 14 and Q to 25, while σ has been optimized in black box manner. Optimal values of σ lie in the interval $[10, 50]$. Furthermore, our nearest neighbors computation is performed exactly but we note that fast approximate computations can be performed *e.g.* using the algorithm of [62]. For TV+low-patch rank approach, a single additional parameter is involved, namely the size of the extracted patches. In our experiments, we tried patch sizes of 4×4, 8×8 and 16×16. With the images we used, the 8×8 size gave the best results. Regarding the parameters of the DR algorithms, we use $\nu = \zeta = 1$. For TV regularization we used 500 iterations, which in all cases provided a stable energy profile. For the TV+low-patch rank and the nonlocal approaches we used 1000 iterations.

Since, to the best of our knowledge, the IZC setting has not been considered in the literature (except in [2]), we cannot perform fair comparisons with other methods. Indeed, in the framework of single-image SR, there is no available HR data and single-image SR algorithms rely only the LR image. However, we selected two state-of-the-art single-image SR methods whose implementations are freely available and decided to compare the performance yielded by our three algorithms with the ones yielded by these two methods. The methods we selected are those of Yang et al. [13] and of Dong et al. [12].

The method of [13] uses sparse code invariance of LR/HR patches over learned dictionaries. As the degradation model assumed in [13] is different from the one we adopted in this paper, we re-learned a dictionary of size 1024 form LR/HR examples synthesized from the same data set proposed in [13] but this time using our degradation model. Furthermore,

we optimized the Lagrange multiplier appearing in the back-projection step in order to yield the best results.

The method of Dong et al. [12] combines ideas from clustering and sparse coding with the nonlocal approach. In the first iteration of the algorithm, patches of an interpolated version of the LR image serve as a training set. They are clustered using K-means. A global dictionary is obtained, along the corresponding sparse codes, by concatenating PCA sub-dictionaries from each cluster. The computed sparse codes are further refined by averaging with similar patches in a manner similar to the nonlocal-means approach [49], leading to sparse codes for each input interpolated patch. These sparse codes are further refined through a back-projection step aiming to enforce consistency with regard to the degradation model. A first estimation of the HR image is obtained by averaging overlapping patches. The overall procedure is iterated a fixed number of times, each time taking the output of the previous iteration as input for the next one.

In Table I we report PSNR ans SSIM values between the 8 images of Figure 3 and the results obtained with the 6 methods outlined in the previous paragraph. These values are computed only on the reconstructed HR part. We also report the results obtained with bicubic interpolation. For each of our proposed methods, we also include reconstruction obtained using the algorithms of Sections IV, V, VI but without using the available HR data. This corresponds to using the same algorithms with the null matrix as a binary mask. The corresponding columns are labeled "no HR" in Table I.

the quality of the observed data is improved by using a nonlocal

As can be seen in Table I, the NLTV$^+$ achieves the best performance in terms of PSNR. We also note that the nonlocal approaches achieve the best results (in terms of SSIM) when the HR part is taken into account. Regarding the methods we proposed, we note that the TV+$\|\mathcal{P}\cdot\|_*$ decomposition method consistently outperforms the TV-based approach. In some cases, it also achieves better performance than the dictionary-based method of Yang et al. This numerical evidence can further be confirmed by inspecting the IZC reconstructions of Figures 4, 5, 6. Notice in particular how the zoom completion performed by the nonlocal approach on Figures 4 and 6 is hardly distinguishable from the sensed HR part. This is also confirmed in Fig. 7 where the quality of the observed data is improved by using nonlocal technique on a slice of a CT scan of a rock [4]. The experimental setting is the same as for the Brodatz images: we start with a ground truth full HR image f_0 and simulate the incomplete image $y^{(2)}$ by masking 25% of f_0. A complete LR image $y^{(1)}$ is obtained by blurring f_0, downsampling the result by a factor 2 in each direction, and finally adding white Gaussian noise.

VIII. Conclusion

An inverse problem which combines single-image SR and inpainting has been studied in this paper. We have motivated

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

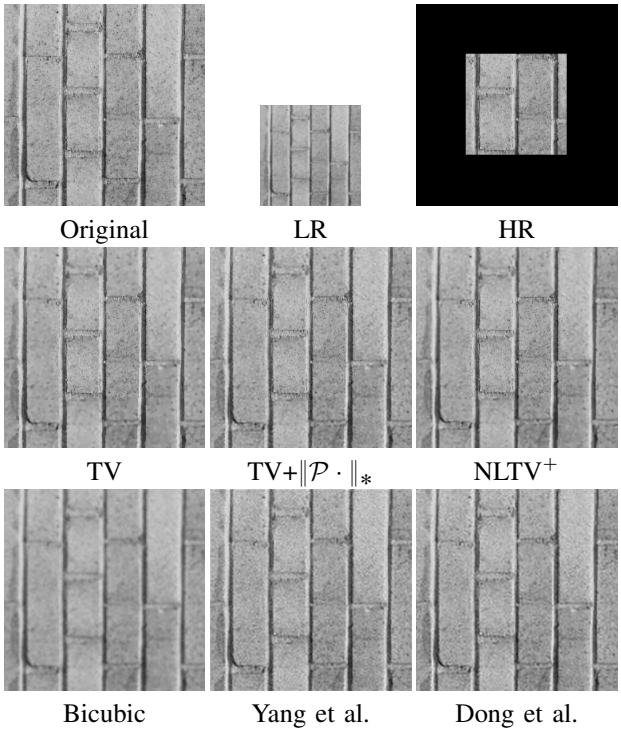

Fig. 4. IZC for the "brick" image. From left to right and form top to bottom: ground-truth, LR input, incomplete HR input, TV (psnr = 27.27), TV+$\|\mathcal{P}\cdot\|_*$ (psnr = 27.33), NLTV$^+$ (psnr = **27.79**), bicubic (psnr = 25.38), Yang et al. [13] (psnr = 27.12), Dong et al. [12] (psnr = 27.41).

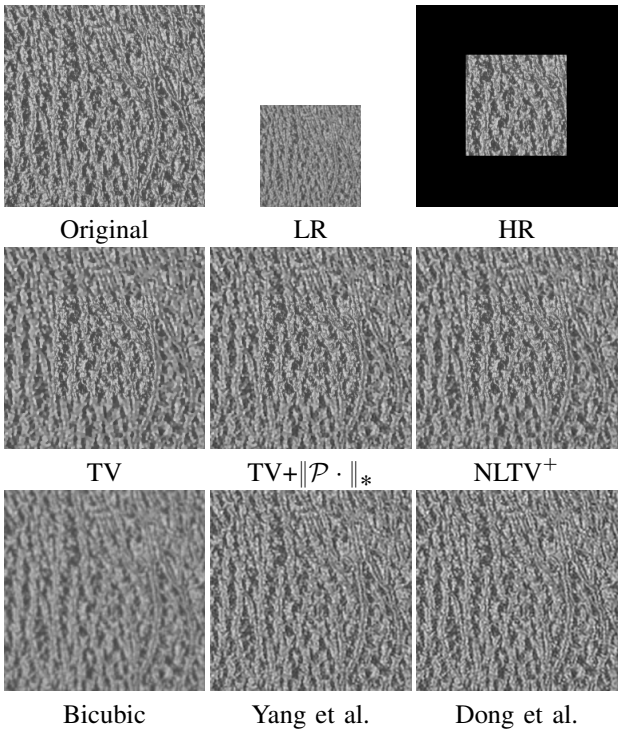

Fig. 5. IZC for the "leather" image. From left to right and form top to bottom: ground-truth, LR input, incomplete HR input, TV (psnr = 22.62), TV+$\|\mathcal{P}\cdot\|_*$ (psnr = 22.98), NLTV$^+$ (psnr = **23.24**), bicubic (psnr = 20.25), Yang et al. [13] (psnr = 23.01), Dong et al. [12] (psnr = 23.14).

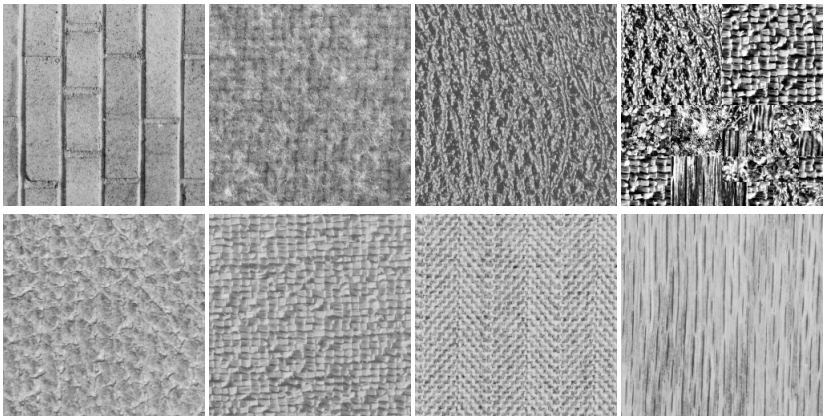

Fig. 3. A subset from the Brodatz image data set we use as input to SR algorithms. From left to right and from top to bottom : "brick", "coth", "leather", "mosaic", "pigskin", "raffia", "weave" and "wood".

TABLE I
PSNR (DB) AND SSIM ON THE RECONSTRUCTED HR AREA FOR DIFFERENT SR METHODS

	TV		TV+$\|\mathcal{P}\cdot\|_*$		**NLTV**		**NLTV**$^+$		Bicubic	Yang et al.	Dong et al.
	no HR	with HR	no HR	with HR	no HR	with HR	no HR	with HR			
Brick	27.26	27.27	27.29	27.33	27.44	27.43	27.71	**27.79**	25.38	27.12	27.41
	0.639	0.640	0.639	0.641	0.687	0.840	0.697	**0.847**	0.608	0.636	0.650
Coth	26.21	26.22	26.30	26.30	26.50	26.49	26.56	**26.58**	25.21	26.34	26.34
	0.640	0.640	0.657	0.657	0.699	0.850	0.706	**0.856**	0.577	0.664	0.664
Leather	22.60	22.62	22.93	22.98	22.90	23.06	23.10	**23.24**	20.25	23.01	23.14
	0.706	0.707	0.728	0.730	0.760	0.886	0.772	**0.889**	0.502	0.721	0.739
Mosaic	18.90	18.93	19.19	19.25	19.10	19.40	19.38	**19.51**	15.62	18.99	19.27
	0.644	0.645	0.671	0.672	0.805	**0.910**	0.811	0.909	0.490	0.675	0.685
Pigskin	28.09	28.09	28.16	28.18	28.38	28.51	28.63	**28.65**	26.27	28.10	28.35
	0.734	0.735	0.748	0.749	0.788	0.901	0.799	**0.902**	0.644	0.748	0.758
Raffia	27.88	27.91	28.08	28.13	28.22	28.50	28.63	**28.79**	24.86	27.95	28.59
	0.820	0.821	0.830	0.832	0.865	0.934	0.879	**0.939**	0.698	0.828	0.848
Weave	26.05	26.06	26.33	26.38	26.43	26.63	26.66	**26.74**	23.05	26.35	26.72
	0.792	0.792	0.805	0.807	0.848	0.927	0.859	**0.929**	0.643	0.807	0.823
Wood	26.40	26.49	26.47	26.64	27.00	27.51	27.72	**28.35**	24.74	26.82	27.98
	0.794	0.794	0.751	0.759	0.773	0.893	0.801	**0.907**	0.808	0.729	0.787
Average	25.43	25.44	25.59	25.64	25.74	25.94	26.04	**26.20**	23.17	25.58	25.97
	0.820	0.821	0.822	0.825	0.778	0.893	0.791	**0.897**	0.722	0.817	0.842

the importance of this problem and one of its possible applications. Then we developed three regularization methods and examined the practical optimization of each corresponding energy functions.

While the TV-based solution is appropriate for the class of piecewise smooth images, the nonlocal and the TV+low-patch rank approaches allow targeting larger image classes, in particular images with strong repetitive textures. Both approaches organize the patches of the sought-after image following the evidence provided by the complete LR and incomplete HR image pair.

The TV+$\|\mathcal{P}\cdot\|_*$ method allows to go one step further than the TV-based approach. However, we have seen that in order to truly leverage the available data, it is necessary to resort to explicit patch comparisons and to incorporate them in a nonlocal cost function. In this paper, patch similarities that drive the nonlocal regularization were based on an interpolated version of the LR image. When the super-resolution factor r is important and/or when the blurring matrix attenuates too much the high frequencies of the HR image, these weights can be erroneous and their incorporation into the regularization functional can lead to bad reconstructions.

The approaches introduced in Sections IV and VI for regularizing the IZC problem relied on the smoothness (local or nonlocal) of the underlying sought-after image. In section V, we augmented this smoothness assumption by adopting a decomposition model and by trying to enforce a low-rank hypothesis on the patches of the texture component. We note that all the proposed methods do not rely on sparse representations over learned dictionaries and as such they can be easily adapted when the degradation model changes, *e.g.* when the blurring filter or the super-resolution factor change.

Possible future work concerning the IZC problem, in particular when the blurring filter and the super-resolution factor are fixed, can concentrate on developing IZC strategies based on sparse and redundant representations [63]. As the setting for IZC is a LR/HR image pair, one can consider training both LR and visible HR patches in order to learn an adapted dictionary. Another possible direction for future work is to consider, for the nonlocal approach, an iterative minimization process alternating between weight computation and zoom completion steps as was done in [19] for image inpainting.

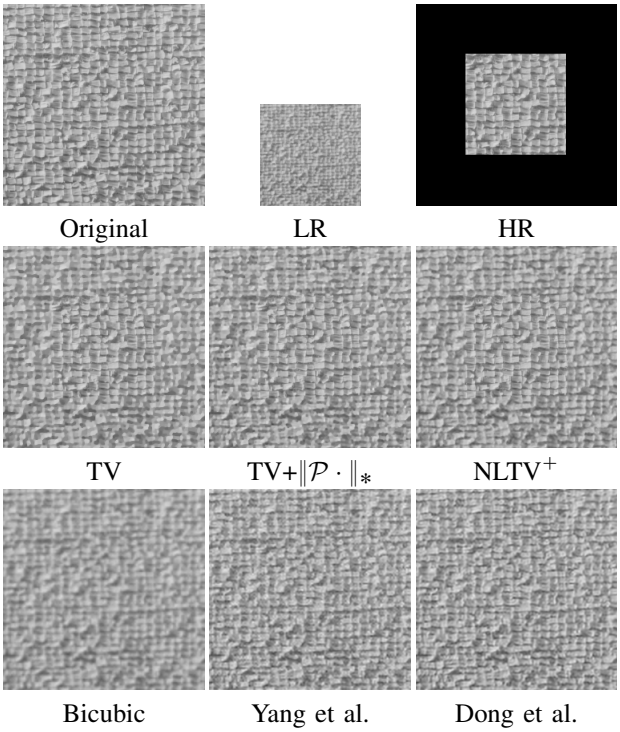

Fig. 6. IZC for the "raffia" image. From left to right and form top to bottom: ground-truth, LR input, incomplete HR input, TV (psnr = 27.91), TV+$\|\mathcal{P}\cdot\|_*$ (psnr = 28.13), NLTV$^+$ (psnr = **28.79**), bicubic (psnr = 24.86), Yang et al. [13] (psnr = 27.95), Dong et al. [12] (psnr = 28.59).

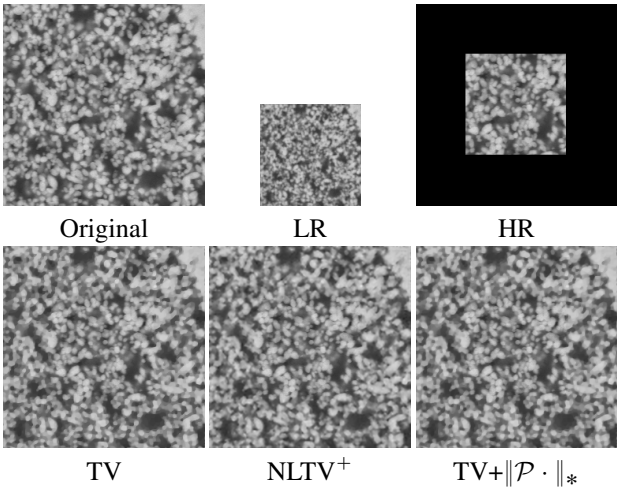

Fig. 7. IZC for a CT image scan. From left to right and from top to bottom : ground truth, LR image, partial HR image, TV-regularization (psnr = 31.05), nonlocal regularization (psnr = **32.17**), TV+low-patch rank regularization (psnr = 31.70).

REFERENCES

[1] M. Bertero and P. Boccacci, *Introduction to inverse problems in imaging*. CRC press, 2010.

[2] M. Hidane, J.-F. Aujol, Y. Berthoumieu, and C.-A. Deledalle, "Super-resolution from a low-and partial high-resolution image pair," in *International Conference on Image Processing*, 2014.

[3] J. Yang and T. Huang, "Image super-resolution: Historical overview and future challenges," *Super-resolution imaging*, pp. 20–34, 2010.

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

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

[6] M. Ebrahimi and E. R. Vrscay, "Solving the inverse problem of image zooming using self-examples?" in *Image analysis and Recognition*. Springer, 2007, pp. 117–130.

[7] D. Glasner, S. Bagon, and M. Irani, "Super-resolution from a single image," in *Computer Vision, 2009 IEEE 12th International Conference on*. IEEE, 2009, pp. 349–356.

[8] A. Adler, Y. Hel-Or, and M. Elad, "A shrinkage learning approach for single image super-resolution with overcomplete representations," in *Computer Vision–ECCV 2010*. Springer, 2010, pp. 622–635.

[9] S. Hawe, M. Kleinsteuber, and K. Diepold, "Analysis operator learning and its application to image reconstruction," *Image Processing, IEEE Transactions on*, vol. 22, no. 6, pp. 2138–2150, 2013.

[10] G. Yu, G. Sapiro, and S. Mallat, "Solving inverse problems with piecewise linear estimators: From gaussian mixture models to structured sparsity," *Image Processing, IEEE Transactions on*, vol. 21, no. 5, pp. 2481–2499, 2012.

[11] K. Zhang, X. Gao, D. Tao, and X. Li, "Multi-scale dictionary for single image super-resolution," in *Computer Vision and Pattern Recognition (CVPR), 2012 IEEE Conference on*. IEEE, 2012, pp. 1114–1121.

[12] W. Dong, L. Zhang, G. Shi, and X. Li, "Nonlocally centralized sparse representation for image restoration." *IEEE Transactions on Image Processing*, vol. 22, no. 4, pp. 1620–1630, 2013.

[13] J. Yang, J. Wright, T. S. Huang, and Y. Ma, "Image super-resolution via sparse representation," *Image Processing, IEEE Transactions on*, vol. 19, no. 11, pp. 2861–2873, 2010.

[14] J. Yang, Z. Wang, Z. Lin, S. Cohen, and T. Huang, "Coupled dictionary training for image super-resolution," *Image Processing, IEEE Transactions on*, vol. 21, no. 8, pp. 3467–3478, 2012.

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

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

[17] C. Ballester, M. Bertalmio, V. Caselles, G. Sapiro, and J. Verdera, "Filling-in by joint interpolation of vector fields and gray levels," *Image Processing, IEEE Transactions on*, vol. 10, no. 8, pp. 1200–1211, 2001.

[18] J.-F. Aujol, S. Ladjal, and S. Masnou, "Exemplar-based inpainting from a variational point of view," *SIAM Journal on Mathematical Analysis*, vol. 42, no. 3, pp. 1246–1285, 2010.

[19] P. Arias, G. Facciolo, V. Caselles, and G. Sapiro, "A variational framework for exemplar-based image inpainting," *International journal of computer vision*, vol. 93, no. 3, pp. 319–347, 2011.

[20] A. Bugeau, M. Bertalmío, V. Caselles, and G. Sapiro, "A comprehensive framework for image inpainting," *Image Processing, IEEE Transactions on*, vol. 19, no. 10, pp. 2634–2645, 2010.

[21] H. Schaeffer and S. Osher, "A low patch-rank interpretation of texture," *SIAM Journal on Imaging Sciences*, vol. 6, no. 1, pp. 226–262, 2013.

[22] L. Condat, "A primal-dual splitting method for convex optimization involving lipschitzian, proximable and linear composite terms," *Journal of Optimization Theory and Applications*, vol. 158, no. 2, pp. 460–479, 2013.

[23] P. L. Combettes and J.-C. Pesquet, "Proximal splitting methods in signal processing," in *Fixed-point algorithms for inverse problems in science and engineering*. Springer, 2011, pp. 185–212.

[24] J.-C. Pesquet and N. Pustelnik, "A parallel inertial proximal optimization methods," *Pac. J. Optim*, vol. 8, no. 2, pp. 273–305, Apr. 2012.

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

[26] L. M. Briceño-Arias and P. L. Combettes, "A monotone + skew splitting model for composite monotone inclusions in duality," *SIAM J. Optim.*, vol. 21, no. 4, pp. 1230–1250, Oct. 2011.

[27] P. L. Combettes and J.-C. Pesquet, "Primal-dual splitting algorithm for solving inclusions with mixtures of composite, Lipschitzian, and parallel-sum type monotone operators," *Set-Valued Var. Anal.*, vol. 20, no. 2, pp. 307–330, Jun. 2012.

[28] H. Raguet, J. Fadili, and G. Peyré, "Generalized Forward-Backward splitting," Preprint Hal-00613637, Tech. Rep., 2011. [Online]. Available: http://hal.archives-ouvertes.fr/hal-00613637/

[29] B. C. Vũ, "A splitting algorithm for dual monotone inclusions involving cocoercive operators," vol. 38, no. 3, pp. 667–681, Apr. 2013.

[30] N. Komodakis and J.-C. Pesquet, "Playing with duality: An overview of recent primal-dual approaches for solving large-scale optimization problems," 2015, to appear.

[31] A. Beck and M. Teboulle, "A fast iterative shrinkage-thresholding algorithm for linear inverse problems," *SIAM Journal on Imaging Sciences*, vol. 2, no. 1, pp. 183–202, 2009.

[32] P. L. Combettes and J.-C. Pesquet, "A proximal decomposition method for solving convex variational inverse problems," *Inverse problems*, vol. 24, no. 6, p. 065014, 2008.

[33] J.-F. Aujol, G. Aubert, L. Blanc-Féraud, and A. Chambolle, "Image decomposition into a bounded variation component and an oscillating component," *J. Math. Imaging Vis.*, vol. 22, no. 1, pp. 71–88, Jan. 2005.

[34] J.-F. Aujol and A. Chambolle, "Dual norms and image decomposition models," *Int. J. Comput. Vision*, vol. 63, no. 1, pp. 85–104, Jun. 2005.

[35] T. F. Chan and S. Esedoglu, "Aspects of total variation regularized l 1 function approximation," *SIAM J. Appl. Math*, 2005.

[36] J.-F. Aujol, G. Gilboa, T. Chan, and S. Osher, "Structure-texture image decomposition–modeling, algorithms, and parameter selection," *Int. J. Comput. Vision*, vol. 67, no. 1, pp. 111–136, Apr. 2006.

[37] A. Buades, T. Le, J. Morel, and L. Vese, "Fast cartoon+ texture image filters," *Image Processing, IEEE Transactions on*, vol. 19, no. 8, pp. 1978–1986, 2010.

[38] T. Goldstein and S. Osher, "The split bregman method for l1-regularized problems," *SIAM Journal on Imaging Sciences*, vol. 2, no. 2, pp. 323–343, 2009.

[39] S. Ono, T. Miyata, and I. Yamada, "Cartoon-texture image decomposition using blockwise low-rank texture characterization," *Image Processing, IEEE Transactions on*, vol. 23, no. 3, pp. 1128–1142, March 2014.

[40] P. L. Combettes and J.-C. Pesquet, "A Douglas–Rachford splitting approach to nonsmooth convex variational signal recovery," *Selected Topics in Signal Processing, IEEE Journal of*, vol. 1, no. 4, pp. 564–574, 2007.

[41] J.-F. Cai, E. J. Candès, and Z. Shen, "A singular value thresholding algorithm for matrix completion," *SIAM Journal on Optimization*, vol. 20, no. 4, pp. 1956–1982, 2010.

[42] M. V. Afonso, J. M. Bioucas-Dias, and M. A. Figueiredo, "Fast image recovery using variable splitting and constrained optimization," *Image Processing, IEEE Transactions on*, vol. 19, no. 9, pp. 2345–2356, 2010.

[43] S. Kindermann, S. Osher, and P. W. Jones, "Deblurring and denoising of images by nonlocal functionals," *Multiscale Modeling & Simulation*, vol. 4, no. 4, pp. 1091–1115, 2005.

[44] G. Gilboa and S. Osher, "Nonlocal linear image regularization and supervised segmentation," *Multiscale Modeling & Simulation*, vol. 6, no. 2, pp. 595–630, 2007.

[45] ——, "Nonlocal operators with applications to image processing," *Multiscale Modeling & Simulation*, vol. 7, no. 3, pp. 1005–1028, 2008.

[46] A. Elmoataz, O. Lézoray, and S. 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.

[47] S. Bougleux, A. Elmoataz, and M. Melkemi, "Local and nonlocal discrete regularization on weighted graphs for image and mesh processing," *International journal of computer vision*, vol. 84, no. 2, pp. 220–236, 2009.

[48] O. Lézoray, V.-T. Ta, and A. Elmoataz, "Partial differences as tools for filtering data on graphs," *Pattern Recognition Letters*, vol. 31, no. 14, pp. 2201 – 2213, 2010.

[49] A. Buades, B. Coll, and J.-M. Morel, "Image denoising methods. A new non-local principle," *SIAM Review*, vol. 52, no. 1, pp. 113–147, 2010.

[50] G. Peyré, S. Bougleux, and L. Cohen, "Non-local regularization of inverse problems," in *Computer Vision ECCV 2008*, ser. LNCS. Springer Berlin Heidelberg, 2008, vol. 5304, pp. 57–68.

[51] G. Peyré, S. Bougleux, L. D. Cohen *et al.*, "Non-local regularization of inverse problems," *Inverse Problems and Imaging*, vol. 5, no. 2, pp. 511–530, 2011.

[52] Y. Lou, X. Zhang, S. Osher, and A. Bertozzi, "Image recovery via nonlocal operators," *Journal of Scientific Computing*, vol. 42, no. 2, pp. 185–197, 2010.

[53] X. Zhang, M. Burger, X. Bresson, and S. Osher, "Bregmanized nonlocal regularization for deconvolution and sparse reconstruction," *SIAM Journal on Imaging Sciences*, vol. 3, no. 3, pp. 253–276, 2010.

[54] X. Zhang and T. F. Chan, "Wavelet inpainting by nonlocal total variation," *Inverse problems and Imaging*, vol. 4, no. 1, pp. 191–210, 2010.

[55] O. Lézoray and L. Grady, *Image processing and analysis with graphs: theory and practice*. CRC Press, 2012.

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

[57] M. Werlberger, T. Pock, and H. Bischof, "Motion estimation with nonlocal total variation regularization." IEEE, 2010, pp. 2464–2471.

[58] G. Peyré, "A review of adaptive image representations." *J. Sel. Topics Signal Processing*, vol. 5, no. 5, pp. 896–911, 2011.

[59] C. Couprie, L. Grady, L. Najman, J.-C. Pesquet, and H. Talbot, "Dual constrained TV-based regularization on graphs," *SIAM J. on Imaging Sciences*, vol. 6, no. 3, pp. 1246–1273, Oct. 2013.

[60] G. Chierchia, N. Pustelnik, J.-C. Pesquet, and B. Pesquet-Popescu, "Epigraphical splitting for solving constrained convex optimization problems with proximal tools," *J. on Image and Video Proc.*, 2014.

[61] J. Duran, A. Buades, B. Coll, and C. Sbert, "A nonlocal variational model for pansharpening image fusion," *SIAM Journal on Imaging Sciences*, vol. 7, no. 2, pp. 761–796, 2014.

[62] C. Barnes, E. Shechtman, D. B. Goldman, and A. Finkelstein, "The generalized PatchMatch correspondence algorithm," in *European Conference on Computer Vision*, Sep. 2010.

[63] M. Elad, *Sparse and redundant representations: from theory to applications in signal and image processing*. Springer, 2010.