WAVELET INPAINTING BY NONLOCAL TOTAL VARIATION

Xiaoqun Zhang

UCLA Mathematics Department
Box 951555
Los Angeles, CA 90095-1555, USA

Tony F. Chan

The Hong Kong University of Science and Technology
Clear Water Bay
Kowloon, Hong Kong

(Communicated by Hao-Min Zhou and Xue-Cheng Tai)

Abstract. Wavelet inpainting problem consists of filling in missed data in the wavelet domain. In [17], Chan, Shen, and Zhou proposed an efficient method to recover piecewise constant or smooth images by combining total variation regularization and wavelet representations. In this paper, we extend it to nonlocal total variation regularization in order to recover textures and local geometry structures simultaneously. Moreover, we apply an efficient algorithm framework for both local and nonlocal regularizers. Extensive experimental results on a variety of loss scenarios and natural images validate the performance of this approach.

1. Introduction

The discrete wavelet transform is a popular method for transmission coding. An example is the recently developed JPEG2000 image compression standard. As conventional communication techniques can not provide an error free transmission, bit errors and data losses in received subbands (see Figure 1) may heavily affect the signal quality. This problem is often referred as *error concealment* in the transmission and communication community. In [17], Chan, Shen and Zhou addressed this problem as *wavelet inpainting*, since the problem of filling in missing or damaged wavelet coefficients is closely related to classical image domain inpainting problems. In this paper, we use the terminology *wavelet inpainting* as [17] and propose to use the nonlocal total variation for this application.

Since the pioneering work of Masnou and Morel in [39], Chan and Shen in [16] and Bertalmio, Sapiro, Caselles and Balleste in [4], geometrical and variational PDE models have been well studied for image domain inpainting based on local geometrical information [15, 16, 28, 8]. However, local edges based methods are limited to non-texture regions. Texture synthesis techniques are additionally introduced in order to recover textured natural images. In [24], Efros and Leung proposed to synthesis an unknown pixel by known pixels which have similar neighborhood as the current one. This technique is very efficient for texture synthesis, and a similar

idea is applied for image denoising by Buades, Coll and Morel [7] who proposed to use a nonlocal means filter. The texture synthesis technique has also been applied to inpainting features by Bertalmio-Vese-Sapiro -Osher in [3], where they proposed to decompose first an image into a structure part and a textured part, then to apply different techniques separately to both parts. On the other side, Demanet, Song and Chan [22] attempted to formulate the heuristic patches copy-paste technique presented in [24] in a variational framework. Recently, Aujol, Ladjal and Masnou [2] also aimed to formulate variational models for some exemplar-based algorithms and investigate the ability of reconstructing local geometric features and textures. Nevertheless, a unified and well-posed model is still an open problem. Other techniques such as morphological component analysis is also applied on simultaneous cartoon and texture image inpainting by Elad, Starck, Querre and Donoho in [26].

Motivated by a large amount of work on image domain inpainting, some authors propose to solve wavelet inpainting (error concealment) by image domain inpainting. A related scenario is JPEG standard reconstruction, which is based on the block DCT transform. Errors in the bit stream result in the loss of all the information related to the damaged blocks, and in general the damaged blocks can be well located because of the locality of the DCT transform. Once the damaged blocks are located, an image domain inpainting or interpolation algorithm is then applied. For example in [46], Rane, Sapiro, and Bertalmio first classified missing image blocks into structures and textures and then applied an image domain inpainting algorithm and texture synthesis separately. In [41], an oriented anisotropic elliptic PDE was used to repair textures as well as edges to avoid a preprocess of classification as in [46] or a decomposition in [3] for regular square loss patterns. However, the damaged regions due to wavelet loss are often not well localized and the affected regions are in general too large to fill in by block-based inpainting methods. In fact, loss in different wavelet subbands affect the image quality differently and the degradation is often spatially inhomogeneous. We show the layout of wavelet subbands in Figure 1. By the same principle as image compression, the loss of coefficients in higher frequency subbands, such as "HH_1, HH_2, LH_1, HL_1, LH_2,HL_2" in Figure 1, do not greatly affect visual quality, while the loss of other high frequencies in the coarsest subband ("HL_3,LH_3" in Figure 1) create Gibbs artifacts or other blur effects, see Figure 2 and Figure 3. The inpainting area can not be well located since the artifacts are spread along edges. Visually, the severest degradation is created by the coarsest low-low subband ("LL_3" in Figure 1). In Figure 5 and Figure 7, loss in this subband severely create regular or irregular big black squares in the image domain. In the following, we use "LL" to refer this subband. In addition, some noise might be also present in the received coefficients, in which case image domain inpainting algorithms are not applicable. In the literature of error concealment, some effort is also devoted to using explicit interpolation/extrapolation, regularization or statistic inference in the wavelet domain, such as in [34], where Hemami and Gray proposed different interpolation strategies for LL subband and higher frequencies. However, wavelet coefficients (except LL frequencies) are designed to decouple the correlation between pixels, so retained coefficients usually can not provide enough information for the missing ones. Moreover, it is well known that a small deviation of wavelet coefficients may cause displeasing artifacts along the edges in the image domain.

An alternative class of approaches, including ours proposed in this paper, is to use a hybrid method to control both image and wavelet domains. In [48], some

LL_3	HL_3	HL_2	HL_1
LH_3	HH_3		
LH_2		HH_2	
LH_1		HH_1	

FIGURE 1. DWT coefficients layout in JPEG 2000 with three levels decomposition.

constraints related to smoothness and edge continuity were imposed for the recovery of missing coefficients and a projection onto convex sets (POCS) algorithm was applied to solve the problem. Also in [1], a median filtered image constraint set was imposed without theoretical convergence proof. Based on a priori hypothesis that the original images are more or less piecewise smooth functions presenting sharp discontinuities, in [17], Chan, Shen and Zhou proposed two variational models combining the total variation (TV) minimization technique with wavelet representations. The authors have also investigated the existence and uniqueness of the proposed models, and the related numerical schemes are designed to handle the computation in wavelet domain. Another related but different work is due to Durand and Froment [23]. They considered the TV minimization under a constraint on the retained wavelet coefficients in order to remove artifacts due to the lost frequencies, created by a threshholding algorithm. The combination of the total variation and wavelet representations has been demonstrated to have effective and automatic control over geometric features, even in the presence of loss of low frequency wavelet coefficients, in which case the image is considerably affected. Other combination of total variation with harmonic representation can be also found in the application of curvelet image reconstruction [10], computerized tomography reconstruction [51, 33]. An important related topic is called *compressive sensing*(CS), which is recently emerged and quickly developed in different research area [11, 12]. The principle of compressive sensing argues that if a signal can be expressed sparsely in a proper basis, then it is possible to exactly recover the signal from a set of incomplete measurements, in a probabilistic setting. In spite of a close connection to CS, the theory developed for CS can not explain the efficiency of the total variation inpainting model. According to the CS theory, the sparse signal must be spread out in the measurement domains, such as Fourier or Gaussian measurements, while incomplete wavelet measurements mainly capture local information. Nevertheless, Chan, Shen and Zhou [17] argued that the hybrid model works because the bounded function space retains sharpness of missing edges, and the minimization of the total variation can recover the geometry information, which is not explicitly represented by classical wavelets.

Although the classical total variation is surprisingly efficient for recovering some lost wavelet coefficients as presented in [17] and confirmed in this paper, it is well-known that the total variation regularization is not suitable for images with fine structures, details and textures, just as other local gradient based methods. In this paper, we extend it to textured images and explore the feasibility of nonlocal regularization for different loss cases. The nonlocal total variation is a variational extension of the nonlocal means filter proposed by Buades, Coll and Morel in [7]. The nonlocal means filter has proven to be very efficient for Gaussian denoising, on preserving not only sharp edges, but also fine details and repetitive patterns. Gilboa and Osher [30] extended the classical total variation and H^1 norm to nonlocal functionals based on the nonlocal filters and graph Laplacian theory [18]. Zhou and Schölkopf in [52] and Elmoataz *et al.* in [27] also proposed to use a a weighted graph Laplacian in the discrete setting. The application of these functionals to different image processing tasks is an active research area, for problems such as image denoising [30], supervised segmentation [29, 5, 35], several inverse problems including image domain inpainting in [45], deconvolution and compressive sensing in [36, 50] and denoising and classification of hyperspectral imaging [13]. The nonlocal TV regularization has been shown to outperform the classical total variation by integrating global information, as long as an appropriate (initial) weight is selected.

Our main contribution in this paper is to extend the total variation based wavelet inpainting to the nonlocal total variation based model, in order to recover textures and geometry structures simultaneously. The proposed non-smooth optimization model is solved by an efficient algorithm. The efficiency of this simple model for natural images are validated by a large amount of numerical simulations. Specially, we formulate wavelet inpainting as an inverse problem and describe how some well-known algorithms for inverse problems can be easily adapted to this application. The challenges for nonlocal regularization in practice include choosing an appropriate and adapted weight, and finding a stable and efficient algorithm. As we discussed above, the received image, which is obtained by setting unknowns to zero do not in general provide the right edge and feature information, making it necessary to use weight updating strategies, as proposed in [45] and [50]. We simulate different subband loss. In particular, the loss in LL subband severely affect the image quality. On the other hand, since the lowest subband is an approximation of the image, it can sometimes be modeled as a piecewise smooth function, allowing some basic interpolation on this subband to greatly improve image quality. In this paper, we use a simple interpolation for the LL subband loss and use the result as an initial guess for further iterative regularization, such as total variation and nonlocal total variation.

The paper is organized as follows. In Section 2, we present the notations for discrete nonlocal total variation as well as the classical one. In Section 3, we introduce the unified model and general algorithm framework for both regularizers. The algorithm adopted here is based on operator splitting and Bregman iteration [42]. In Section 4, we demonstrate the performance of both classical and nonlocal total variation for different scenarios of coefficients loss for natural images, including subband loss, block loss, and random loss. In particular, we discuss two cases of random loss, one is fully random, and another one includes additional LL subband. Further numerical examples highlight the remarkable inpainting qualities of nonlocal total variation regularization for natural images.

2. Discrete Nonlocal Total variation

This section provides the basic notations of discrete formulation instead of the continuous model presented in [30]. Similar to the digital TV filter defined in [14] and graph regularization in [18, 27], we model the digital image domain by a graph (Ω, E), where Ω is a finite set of N nodes (pixels), E is the set of edges. The notation $x \sim y$ is used to denote the edge which connects the nodes x and y. An image u is then a function defined on Ω, which can be represented by a column vector, and the value at node x is denoted by $u(x)$. In the following, we consider a weight function $w(x, y)$ for all edges $x \sim y \in E$. The weight function w is required to be symmetric and it can be extended over $\Omega \times \Omega$, that is, if two nodes x and y are not connected, the weight $w(x, y)$ is set 0. In this setting, nodes may directly interact with nodes that are not neighbors, unlike classical total variation. For this reason, we use the term "nonlocal".

Now we define the weighted graph gradient and other related operators. For a given image $u(x)$ defined on Ω, the weighted graph gradient $\nabla_w u(x)$ is defined as the vector of all directional derivatives (or edge derivative) $\nabla_w u(x, \cdot)$ at x:

$$\nabla_w u(x) := (\nabla_w u(x, y))_{y \in \Omega},$$

with

$$\nabla_w u(x, y) := (u(y) - u(x))\sqrt{w(x, y)}, \quad \forall y \in \Omega.$$

The directional derivatives apply to all the nodes y since the weight $w(x, y)$ is extended to the whole domain $\Omega \times \Omega$. A graph divergence of a vector $q : \Omega \times \Omega \to R$ can be defined by the adjoint relation with the gradient operator and the standard dot product on vectors:

$$\langle \nabla_w u, q \rangle = -\langle u, \mathrm{div}_w q \rangle, \ \forall u : \Omega \to \mathcal{R}, \forall q : \Omega \times \Omega \to \mathcal{R},$$

which leads to the definition of the graph divergence $\mathrm{div}_w(q) : \Omega \to \mathcal{R}$ such that:

$$\mathrm{div}_w q(x) = \sum_{y \in \Omega}(q(x, y) - q(y, x))\sqrt{w(x, y)}.$$

The graph Laplacian is defined by:

$$\Delta_w u(x) := \frac{1}{2}\mathrm{div}_w(\nabla_w u(x)) = \sum_{y \in \Omega}(u(y) - u(x))w(x, y).$$

Note that a factor $\frac{1}{2}$ is used to be consistent with the standard Laplacian definition.

The graph-based H^1 norm and total variation semi-norm are defined by extending classical L^2 and L^1 norm of gradient operators respectively. More precisely, the nonlocal H^1 norm, and the nonlocal total variation are defined as follows:

$$(1) \qquad J^w_{NLH1}(u) \quad := \quad \sum_{x \in \Omega}|\nabla_w u(x)|^2 = \sum_{x \in \Omega}\sum_{y \in \Omega}(u(x) - u(y))^2 w(x, y),$$

$$(2) \qquad J^w_{NLTV}(u) \quad := \quad \sum_{x \in \Omega}|\nabla_w u(x)| = \sum_{x \in \Omega}\sqrt{\sum_{y \in \Omega}(u(x) - u(y))^2 w(x, y)}.$$

In this paper, we are interested in the nonlocal TV semi-norm because analogous to classical total variation, the L^1 norm is in general more efficient than the L^2 norm for sparse reconstruction. The above nonlocal H^1 norm can be viewed as a variational equivalent of the nonlocal means filter, that is, one iteration of gradient flow is equivalent to the nonlocal means filter [44]. Over the two dimensional image

domain, if we only consider the immediate neighbor for each pixel, and consider a constant weight 1 for each neighbor pair, then we have the classical total variation as proposed by Rudin-Osher-Fatemi [47]. Instead, we can consider a more general neighbors system, as the nonlocal weight used in the nonlocal means filter [7], which leads to the nonlocal TV proposed in [30]. The weight function plays an important role in the regularization functionals (1)(2). Following the notation in [7, 30, 45], for a given reference image f_0, we can compute a data-driven weight function $w_0(x, y)$ by using the difference of patches around each node. The patch $p_x(f_0)$ of size $m \times m$ (m is chosen as an odd number) centered at a pixel $x \in \Omega$ is given as

$$p_x(f_0)(t) = f_0(x + t), t = [-\frac{m-1}{2}, \cdots, \frac{m-1}{2}]^2.$$

Then the weight between two points x and y is then defined as a function of the patch distance:

$$(3) \qquad\qquad w(f_0)(x, y) = e^{-\frac{||p_x(f_0) - p_y(f_0)||^2}{2h_0^2}},$$

where h_0 is a filtering parameter. The weight defined above measures the similarities of two patches centered at x and y. The nonlocal means filter [7] is obtained by taking the average of given noisy image with a weight computed from the same noisy image. In this sense, the filter is a nonlinear data-driven filter. This choice of weight is very efficient to reduce gaussian noise while preserving textures and contrast of natural images. For a general regularization, it is crucial to choose a reference image as close as possible to the true image to introduce relevant information regarding image structures, while in practice, since a good reference image is not always available, we therefore need to choose an adapted weight by an appropriate updating scheme. We will give the detail in the next section.

3. Models and Related Algorithms

In [17], the author proposed two TV regularized wavelet inpainting models depending on whether or not noise is considered. The idea is to combine a regularization term in the image domain with a fidelity term in the transform domain. For a two-dimensional image u, let us denote the standard wavelet representation as

$$(4) \qquad\qquad u(\beta, x) = \sum_{j \in \mathcal{Z}, k \in \mathcal{Z}^2} \beta_{j,k} \phi_{j,k}(x),$$

where $x \in \Omega = [1, M] \times [1, N]$, and $\beta = (\beta_{j,k})$ denotes wavelet coefficients of u at level j and location k, and for simplicity $\phi_{j,k}$ denotes a given orthogonal or biorthogonal wavelet basis function (for two-dimensional separable wavelet basis, there are three directions of wavelet functions, but we use the same symbol ϕ). If we use an orthogonal basis, the coefficient $\beta_{j,k} = \langle u, \phi_{j,k} \rangle$, while for biorthogonal basis, the synthesis basis is different from the analysis one, that is $\beta_{j,k} = \langle u, \tilde{\phi}_{j,k} \rangle$, for $\tilde{\phi}_{j,k}$ being a dual basis of $\phi_{j,k}$, see [37]. In discrete case, let $I \subset \Omega$ be the uncorrupted known index set, $\alpha = (\alpha_{j,k})$ for $(j, k) \in I$ denotes measured coefficients, the following two models, respectively for the noiseless case and the noisy one, are considered in the paper [17]:

$$\beta^* = \arg\min_{\beta} TV(u(\beta, x)) = \arg\min_{\beta} \sum_{x \in \Omega} |\nabla_x u(\beta, x)|$$

$$(5) \qquad\qquad \text{s. t} \quad \beta_{j,k} = \alpha_{j,k}, \quad (j, k) \in I$$

and

$$(6) \quad \beta^* = \arg\min_\beta F(\beta) = \arg\min_\beta \sum_{x\in\Omega} |\nabla_x u(\beta, x)| + \sum_{(j,k)\in I} \lambda_{j,k}(\beta_{j,k} - \alpha_{j,k})^2,$$

where $\nabla_x u$ is a discreet gradient of u at x, $\lambda_{j,k}$ are regularization parameters. The solutions u^* are obtained from β^* by the reconstruction formulae (4). To be more general and simple, we rewrite the model (5) as :

$$(7) \quad (\mathbf{P_0}) \quad \beta^* = \arg\min_\beta J(W^{-1}\beta) \quad \text{s. t } P_I(\beta) = \alpha$$

where $J(u)$ is a convex regularization functional, which could be TV or nonlocal TV, W^{-1} denotes the inverse wavelet transform and P_I is the subsampling projection operator from the whole domain onto the known index set I.

Instead of the unconstrained noisy model (6), we consider the constrained model

$$(8) \quad (\mathbf{P_1}) \quad \beta^* = \arg\min_\beta J(W^{-1}\beta) \quad \text{s .t } \|P_I(\beta) - \alpha\|^2 \leq \epsilon,$$

where ϵ is the (estimated) noise level.

The formulations (7) and (8) are in terms of wavelet coefficients β. This is referred as an *synthesis prior* in [25]. We can also formulate the problem according to *analysis prior* in terms of image u as

$$(9) \quad (\mathbf{P_0'}) \quad u^* = \arg\min_u J(u) \quad \text{s. t. } P_I(Wu) = \alpha,$$

and the corresponding L^2 constrained problem for the noisy case is then written as

$$(10) \quad (\mathbf{P_1'}) \quad u^* = \arg\min_u J(u) \quad \text{s .t } \|P_I(Wu) - \alpha\|^2 \leq \epsilon,$$

If we denote $A = P_I W$, the both formulations (9) and (10) are examples of inverse problem restoration. In [50], the model (9) was considered for nonlocal TV compressive sensing, and in [45], the unconstrained Lagrangian of (10) was used for inverse problems because of its simplicity and some efficient algorithms can be applied for this formulation. Both *analysis* and *synthesis* based model for sparse reconstruction are discussed in detail by Elad, Milanfar and Rubinstein in [25]. In general the two classes of models give different solutions. Typically, when the transform is an overcomplete dictionary, it is easier to solve the synthesis model than the analysis one. In the case that the transform is non-singular square, for example, orthogonal or biorthogonal wavelet transform, the analysis models (9), (10) and the synthesis models (7), (8) are equivalent respectively. In this paper, we focus on the analysis model for the wavelet inpainting case due to its simplicity.

To solve (5), a time marching algorithm with projection was proposed in [17], which is time-consuming. In this paper, we apply the proximal Forward-Backward Splitting algorithm (PFBS) [19] and Bregman iteration [42] to solve (9) and (10), as used in [50]. Bregman iteration was originally proposed in [42] to improve TV regularized results, and it became popular for solving equality constrained problem such as (P_0) and (P_0') by solving related Lagrangian subproblems. The Bregman iteration has several advantages compared to continuation method for equality constraint because of its efficiency, especially for sparse reconstruction, and stability since the scale parameter does not need to go to 0. It is largely used for l^1 compressive sensing as Bregmanized FPC [49], linearized Bregman [21, 43], ROF TV regularization as split Bregman [31], and nonlocal TV compressive sensing with Bregmanized operator splitting [50]. The idea of PFBS is solving a sum of two convex functionals by one step forward gradient descent on one functional and one-step

backward inverting on another functional. It has received a large interest recently as a method for sparse reconstruction, especially l^1 compressive sensing such as [19, 32] because of the efficiency of soft thresholding operator for l^1 minimization. The same idea is also applied for other l^1 related regularization, such as NLTV regularization for inverse problems firstly introduced in [45] and then in [50], and TV tomographic reconstruction in [38].

We first focus on the solution of the analysis based model (P'_0). Let $A = P_I W$, by Bregman iteration, the equality constraint problem (P'_0) can be solved by the iteration scheme: $\alpha^k = \alpha$

$$(11) \qquad \begin{cases} u^{k+1} &= \arg\min_u \mu J(u) + \frac{1}{2}||Au - \alpha^k||^2, \\ \alpha^{k+1} &= \alpha^k + (\alpha - Au^{k+1}) \end{cases}$$

for a positive number $\mu > 0$. The first subproblem is solved by the PFBS algorithm, which is based on the proximal operator introduced by Moreau in [40]: let $u^{k+1,0} = u^k$, for $i \geq 1$

$$(12) \qquad u^{k+1,i} = \min_\beta \mu J(u) + \frac{1}{2\delta}\|u - (u^{k+1,i} - \delta A^T(Au^{k+1,i} - \alpha^k))\|^2,$$

where δ is a positive parameter such that $0 < \delta < \frac{2}{\|A^T A\|}$. Since $A = P_I W$, then $A^T = W^T P_I^T$. We can notice that P_I^T is the zero-padding operator, and W^T can be implemented by the inverse wavelet transform for the orthogonal case. However when the wavelet transform is not orthogonal, for example, biorthogonal, the above adjoint operator is not easy or can not be efficiently solved. In fact, we can replace A^T by the pseudo inverse A^+ for the biorthogonal case as a preconditioned algorithm [9, 50] where $A^+ = W^{-1} P_I^T$. Therefore the step (12) can be solved by two steps with forward and backward wavelet transform:

$$(13) \qquad v^{k+1,i+1} = u^{k+1,i} - \delta W^{-1} P_I^T (P_I W u^{k+1,i} - \alpha^k)$$

$$(14) \qquad u^{k+1,i+1} = \min_\beta J(u) + \frac{1}{2\delta\mu}\|u - v^{k+1,i+1}\|^2.$$

The weights computed from the initial image is not in general sufficient to give a good estimation, especially when the initial image is in a very poor quality, as in the case of LL frequency loss. Therefore it is necessary to adapt the weight used in nonlocal regularization according to (3) during the iteration. The idea of updating the weight of the nonlocal means filter is first proposed in [6] for texture pattern reconstruction. In [45, 50], an adapted weight during the minimization are proposed for non local reconstruction. In this paper, we apply the same idea, that is, the regularization functional is replaced with $J_{NLTV}^{w(u)}(u)$ instead of $J_{NLTV}^{w_0}(u)$. This functional is non-convex due to the dependence of the weight on the unknown image. In practice, we add a weight updating step in the Bregman iteration and proximal algorithm, as in [45][50]. Nevertheless, a theoretical convergence of this algorithm is not yet established.

Finally, the unified algorithm for both noise free (9) and noisy (10) cases, TV and nonlocal TV regularization, is described in Algorithm 1.

---

input: $u^0, v^0 = u^0, \alpha^0 = \alpha, h_0$: weight filter parameter, $\mu, \delta, nOuter$ (outer iterations number), $nInner$ (inner iterations number);

while *not stopping criterion* **do**

> **for** $i = 0$ *to* $nInner$ **do**
>
>> $u^{k+1,0} = u^k$;
>>
>> Update $v^{k+1,i+1} = u^{k+1,i} - \delta W^{-1} P_I^T (P_I W u^{k+1,i} - \alpha^k)$ using forward and inverse wavelet transform;
>>
>> Solving $u^{k+1,i+1}$ by ROF/NL-ROF algorithm:
>> $$u^{k+1,i+1} = \arg\min_u \mu\delta J(u) + \frac{1}{2}\|u - v^{k+1,i+1}\|^2;$$
>
> **end**
>
> New image $u^{k+1} := u^{k+1,nInner}$; If J is nonlocal TV, update the nonlocal weight according to u^{k+1}: $w^{(k)} = w(u^{k+1})$ by (3);
>
> Update $\alpha^{k+1} = \alpha^k + \alpha - P_I W u^{k+1}$;

end

Algorithm 1: Nonlocal TV and TV Wavelet Inpainting Algorithm

4. NUMERICAL RESULTS

Now we present some numerical results. We use the standard Peak Signal to Noise (PSNR) to quantify the performance of wavelet coefficient filling:

$$(15) \qquad \text{PSNR}(u, \overline{u}) = 10 \log_{10}(\frac{1}{\|u - \overline{u}\|^2})$$

where 1 is the maximum intensity value of gray scale images since we normalize the images to $[0, 1]$), $\overline{u}$ is the noise free reference image, and $\| \cdot \|$ is the standard l^2 norm. A higher PSNR will imply a better performance. In all examples presented in this section, we apply the Daubechies 7-9 biorthogonal wavelets with symmetric extension, which is used in standard JPEG2000 for lossy compression. We use the WaveLab[1] to implement the forward and the backward biorthogonal wavelet transforms. The size of the coarsest subband is 32×32 for 256×256 images and 64×64 for $512 * 512$ images.

The computation of the nonlocal weights (3) over the whole domain is expensive and difficult to compute. As [7], we choose a searching window of size $d \times d$ and the size of patch to be $m \times m$ for two integers d, m. Also we adopt two strategies to accelerate the computation without degrading the image quality. First, the difference between every two patches are implemented by a fast translation algorithm proposed by Darbon *et al.* [20]. Secondly, in practice, a number of best neighbors are enough to recover a pixel with a good quality, as is done in most nonlocal applications [7, 30, 36, 45, 50]. On the other hand, decreasing the number of neighbors can dramatically speed up the efficiency of minimization of the nonlocal TV norm. Therefore we only use the k-best neighbors for each pixel to compute the nonlocal weight. More precisely, for each pixel $x = (x_1, x_2)$, we consider a neighborhood $N(x)$ containing k most similar neighbors in the $d \times d$ searching window around x (sort the distance between $p_x(f)$ and $p_z(f)$ for each z belongs to the $d \times d$ searching

---

[1] http://www-stat.stanford.edu/~wavelab/

window and take the k smallest distances neighbors) and the four nearest neighbors. Note that it may happen that $x \in N(y)$ but $y \notin N(x)$. To make the weight matrix symmetric, we set $x \in N(y)$ if $y \in N(x)$, and vice versa. By this construction, we have $w(x, y) = w(y, x)$, and the maximum number of neighbors for each pixel is up to $2k + 4$.

For the total variation denoising algorithm, we adopt PDHG [53] based ROF code [2] to solve the second subproblem, although other ROF solvers can be also applied. The patch size and the searching window for the nonlocal weight are fixed as $m = 5$ and $d = 15$ and we choose $k = 10$ best neighbors and four nearest neighbors for each pixel. The split bregman algorithm [31] adapted for NLTV regularization [50] is used for the denoising step. For all the experiments presented in the following, we use $\mu = 0.05$ for TV and $\mu = 0.01$ for nonlocal TV since empirically these choices give good results. By default, the parameter δ used in proximal iteration is set to 1. Since the image is essentially not noisy, we only use 5 steps of NLTV denoising, and use the default setting of PDHG for TV denoising. The inner iteration step is set as 10. The stopping criterion is $\|P_I W u - \alpha\| < 10^{-5}$ for the noiseless case. In the following, we simulate several cases with different loss scenarios as well as with the presence of noise. In all experiments, the received image is compared to the reconstruction by setting unknowns to be zeros. The code for the experiments presented here is available online [3].

4.1. LH AND HL SUBBANDS LOSS. It is well known that any loss in the coarsest subband (LL suband) will heavily degrade visual image quality since the coarsest band is a smoothed estimation of the image, where the big structure information is contained. In this subsection, we consider the lowest LH and HL subband loss (see Figure 1) simulation and deal with the LL subband loss in the subsection 4.3. For both cases, the maximum outer iterations are set as 15. In Figure 2 and Figure 3, the received images are of fair quality with small oscillations, and they are used as the initial guess for both TV and NLTV reconstruction. The images and their zoomed in views show that TV and NLTV regularization can remove small oscillations created by lost coefficients, while NLTV gives more than 2db higher PSNR and better edge reconstruction.

4.2. BLOCK LOSS. In the JPEG-2000 compression standard, the image is decomposed into wavelet subbands and then each subband is divided into codeblocks. Hence, loss or corruption of bits could affect the whole codeblock it belongs to. We crop a subimage of size $128 * 128$ from the Barbara image and we simulate a block loss mask, shown in Figure 4. The outer iterations number is chosen as 100 for both TV and NLTV. The results of reconstructed images and two zoomed in regions are shown in Figure 5. Both TV and NLTV improve the PSNR dramatically compared to the received image and NLTV performs even 7.42db higher(it could be even higher if we stop earlier for NLTV) than TV. Due to the loss of some low frequencies, there is a black square region (red zone) in the received image. Both TV and nonlocal TV can recover the corrupted low frequencies, but there is still a dark line presented in TV reconstructed image. For the second zoomed zone, the cloth pattern is corrupted. The TV regularization is unable to recover the pattern, while NLTV reconstructs perfectly the direction of cloth patterns.
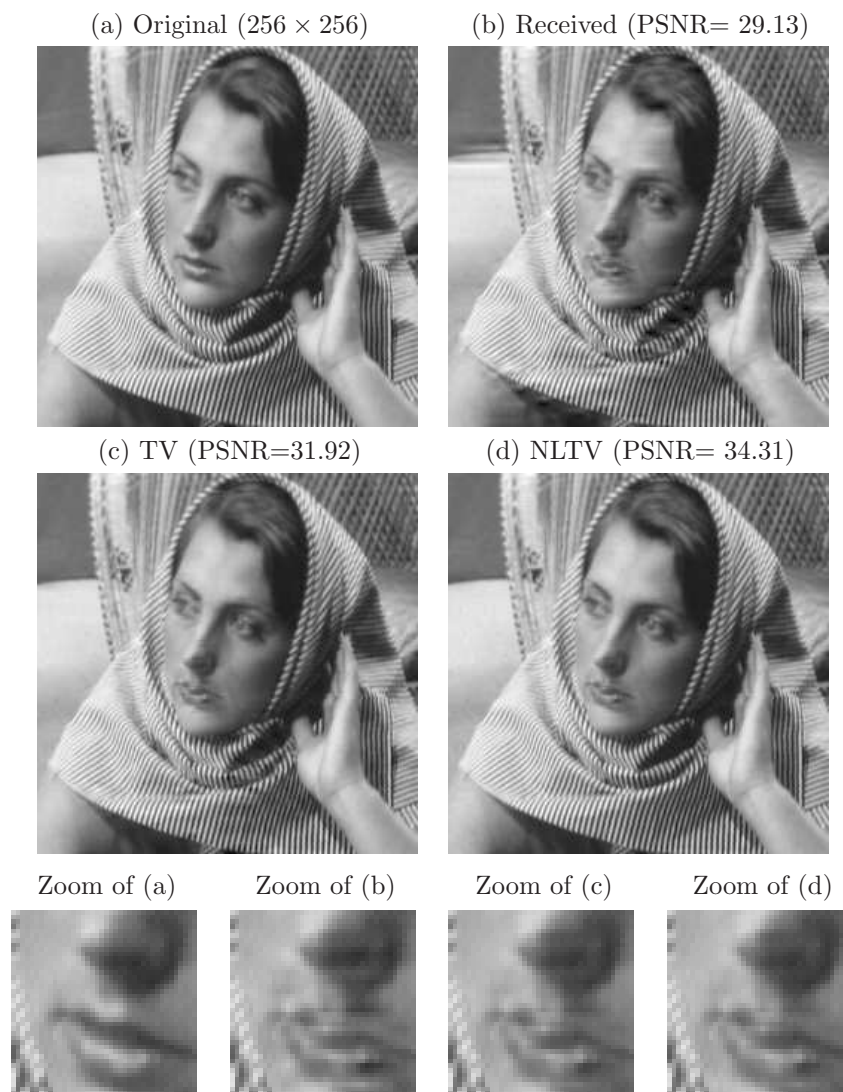
---

[2]http://pages.cs.wisc.edu/~swright/TVdenoising/index.html
[3]http://www.math.ucla.edu/~xqzhang/html/code.html

(a) Original (256 × 256) (b) Received (PSNR= 29.13)

(c) TV (PSNR=31.92) (d) NLTV (PSNR= 34.31)

Zoom of (a)   Zoom of (b)   Zoom of (c)   Zoom of (d)

FIGURE 2. Whole HL subband (32 ∗ 32) loss. We can observe the artifacts on the forehead and along the mouth in (b). Both TV(c) and NLTV(d) remove almost all the artifacts, and NLTV achieves 2db higher PSNR than TV does.

4.3. RANDOM LOSS. Analogous to compressive sensing application, we consider random loss of wavelet coefficients. We first consider the case of random loss but keeping all LL frequencies. Figure 6 shows the results with 50% randomly chosen coefficients and all LL frequencies. The random loss in the higher frequencies causes edge artifacts and incomplete structure, but the initial image is of fair quality. Thus we can use it to estimate an initial weight. As above, NLTV outperforms TV regularization in terms of PSNR and visual image quality, especially for the repeating cloth pattern.
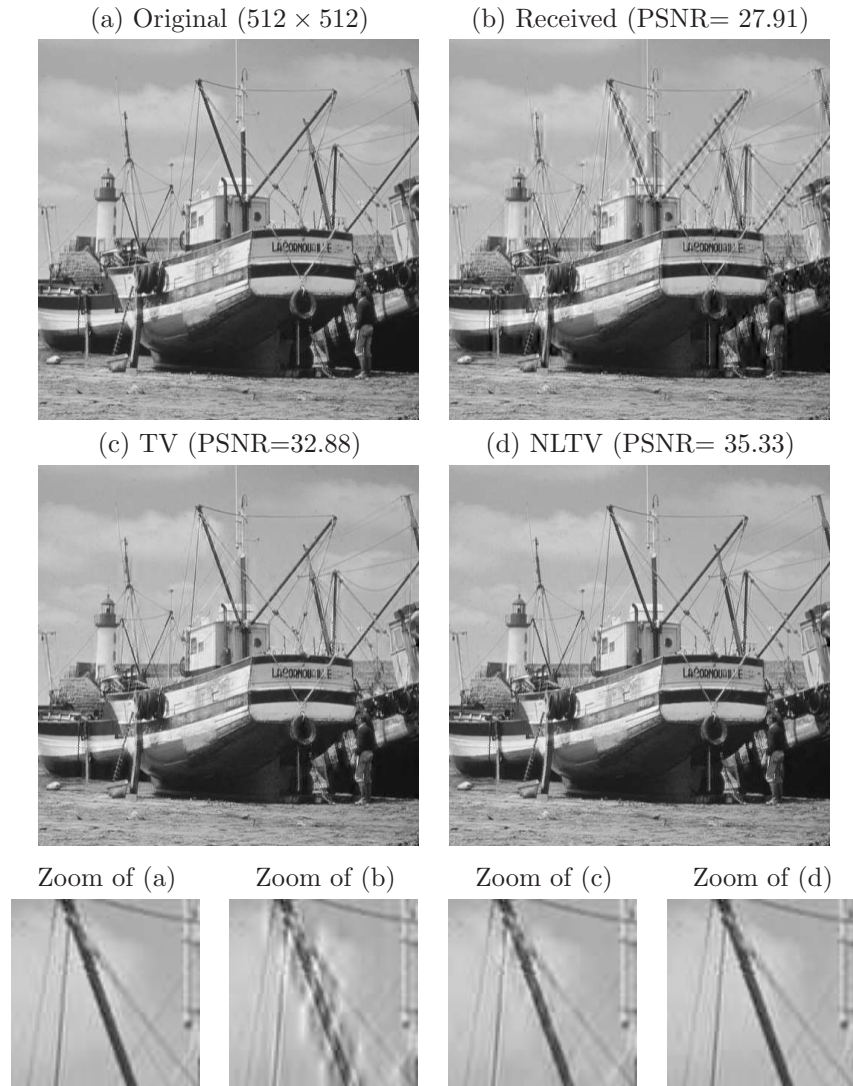
(a) Original (512 × 512)   (b) Received (PSNR= 27.91)

(c) TV (PSNR=32.88)   (d) NLTV (PSNR= 35.33)

Zoom of (a)   Zoom of (b)   Zoom of (c)   Zoom of (d)

FIGURE 3. Whole LH subband (64 ∗ 64) loss. Similar to Figure 2, TV and NLTV regularization recover cleaner edges and NLTV has even higher PSNR than the one of TV.

We also simulate a completely random loss in the whole wavelet domain. Since the received image is in very poor quality because of LL frequencies loss, we consider an interpolated image as the initial guess for TV and nonlocal TV reconstruction instead of the received one. Even though the loss in LL is very crucial, a simple interpolation (nearest neighbor) only on LL subband allows us to remove those ill-posed black squares. One important reason to choose this initialization is that the algorithm for NLTV regularization approximately solve a non-convex functional, thus we need to choose a good weight to avoid a bad local minimum. The initialization is very cheap to obtain and the weight computed from this image, which is
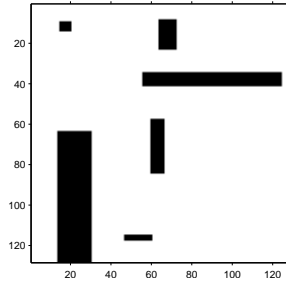
FIGURE 4. Inpainting mask: white (kept coefficients, 88%).

used for the NLTV, is better conditioned. Figure 7 compares these results. Both TV and NLTV can recover much better results than the direct interpolation and NLTV outperforms TV.

Moreover, in Figure 8, we plot PSNR versus sampling rate for the cases of random loss with LL, and random loss on the Barbara image. The nonlocal TV regularization gives a higher PSNR than TV since NLTV can recover more global information than TV. In the random case, we also compare the different initializations for NLTV reconstruction. The initialization with TV is slightly better than interpolated one, but it is more expensive to obtain and it depends on TV minimization performance.

4.4. Noisy data. In Figure 9, we test the case when there is noise in the random loss data. As with the noise free case, the (c) of Figure 9 is obtained by a nearest neighbor interpolation on the LL subband, and it is dramatically better than the received one (b). The figure (d) is obtained by TV regularization with (c) as the initial guess. With the presence of the noise, the cloth pattern in the image can barely be reconstructed by TV regularization since it penalizes noises as well as fine details. For NLTV reconstructed images (e) and (f), we test two initial guesses. One is with (c) and one is with (d). Both have similar PSNR, while (e) has a better cloth pattern inherited from (c) and (f) has a better edge inherited from (e). This result is due to the fact the weight-updating model by itself is nonconvex, and the initial weight more or less dominates the final result.

5. Discussion

We studied the performance of TV and nonlocal TV regularization for wavelet coefficient inpainting of natural images. We discussed different loss scenarios in the wavelet domain, with and without noise. The numerical examples have shown that the nonlocal TV regularization is very effective not only for restoring geometric features but also for filling in some image patterns with global similarity.

Acknowledgements

(a) Original (128 × 128)   (b) Received (PSNR= 17.51)

(c) TV (PSNR=28.64)   (d) NLTV (PSNR= 36.06)

Zoom 1 of (a)   Zoom 1 of (b)   Zoom 1 of (c)   Zoom 1 of (d)

Zoom 2 of (a)   Zoom 2 of (b)   Zoom 2 of (c)   Zoom 2 of (d)
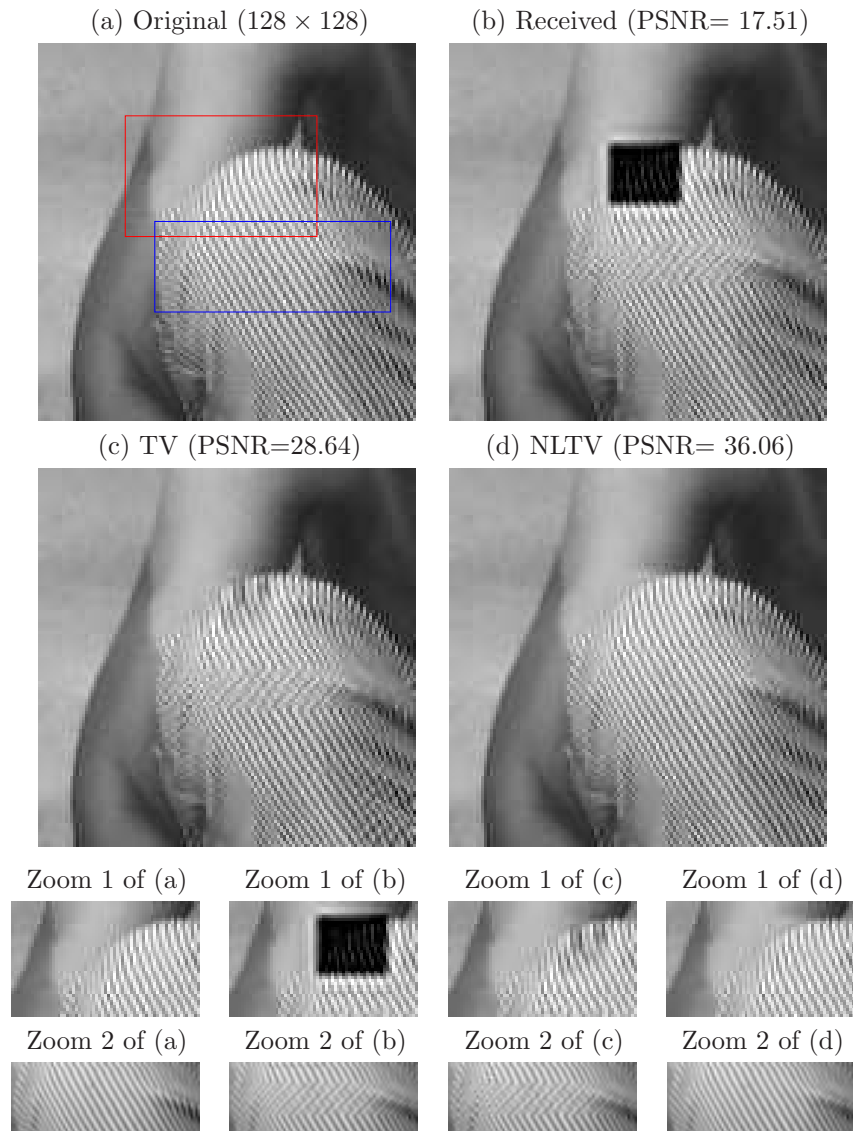
FIGURE 5. Block loss. For both TV and NLTV, the initial guess is
the received image (b). TV can almost recover the dark region due
to some LL frequencies loss (Zoom 1 of (c)), but it can not recover
correctly the cloth pattern(Zoom 2 of (c)). Comparatively, NLTV
recovers perfectly the dark region and the directions of cloth pat-
tern, using non-local geometry information. And PSNR of NLTV
is dramatically higher (7.42db) than TV's.

REFERENCES

[1] (MR2128290) L. Atzori, G. Ginesu, and A. Raccis. *JPEG2000-coded image error concealment
    exploiting convex sets projections*, IEEE Trans. Image Process, **14** (2005), 487–498.

(a) Original (256 × 256)   (b) Received (PSNR=22.84)

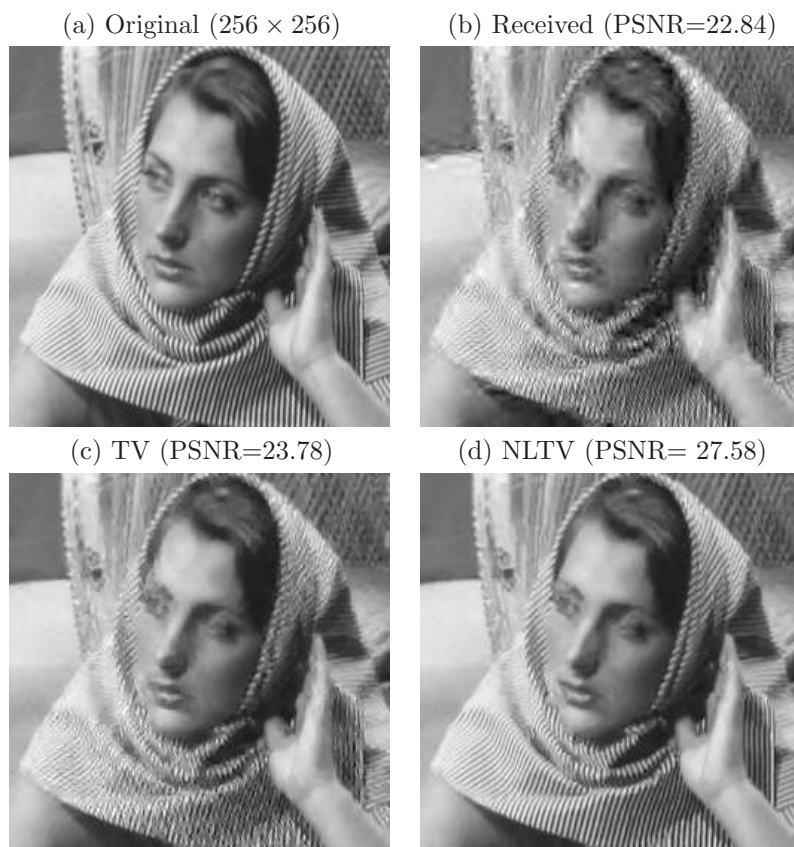(c) TV (PSNR=23.78)   (d) NLTV (PSNR= 27.58)

FIGURE 6. 50% random loss keeping all LL subband frequencies. The edges are globally affected in the observed image (b). NLTV performs better than TV for removing those global artifacts and the difference of PSNR is about 4db.

[2] J.-F. Aujol, S. Ladjal, and S. Masnou. *Exemplar-based inpainting from a variational point of view*, UCLA CAM Report, (2009), 09-04.

[3] M. Bertalmio, L. Vese, G. Sapiro, and S. Osher. *Simultaneous structure and texture image inpainting*, IEEE Trans. Image Process, **12** (2003), 882–889.

[4] M. Bertalmo, G. Sapiro, V. Caselles, and C. Ballester. *Image inpainting*. in "Proceedings of SIGGRAPH 2000", (2000), 417–424.

[5] X. Bresson and T. F. Chan. *Non-local unsupervised variational image segmentation models*, UCLA CAM Report, (2008), 08-67.

[6] T. Brox and D. Cremers. *Iterated nonlocal means for texture restoration*, in "Scale Space and Variational Methods in Computer Vision", **4485** (2008), Springer Berlin / Heidelberg, 13–24.

[7] A. Buades, B. Coll, and J-M. Morel. *A review of image denoising algorithms, with a new one*, Multiscale Model. Simul., **4** (2005), 490–530.

[8] M. Burger, L. He, and C. Schoenlieb. *Cahn-hilliard inpainting and a generalization for gray-value images*, SIAM J. Imaging Sci. **2** (2009), 1129-1167.

[9] (MR2486529) J.-F. Cai, S. Osher, and Z. Shen. *Linearized bregman iterations for frame-based image deblurring*, SIAM J. Imaging Sci. **2** (2009), 226–252.

[10] E. J. Candès. and F. Guo. *New multiscale transforms, minimum total variation synthesis: Applications to edge-preserving image reconstruction*. Signal Processing, **82** (2004), 1519–1543.

(a) Received   (b) Interpolated (PSNR=18.02)

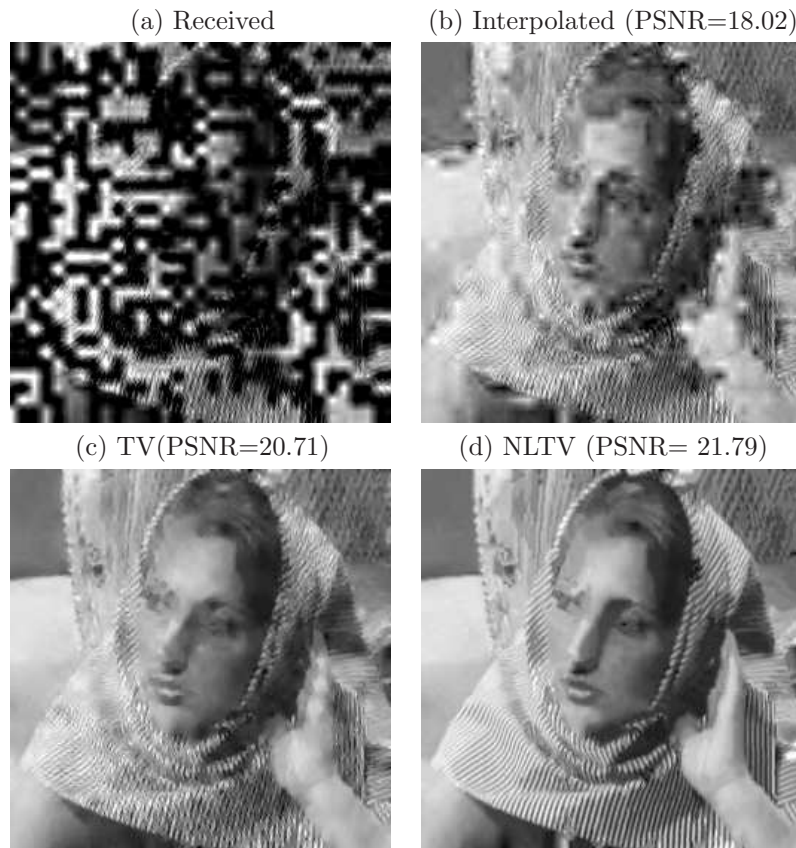(c) TV(PSNR=20.71)   (d) NLTV (PSNR= 21.79)

Figure 7. Random loss: 50% random chosen frequencies. The received image (a) is heavily degraded. The initial guess (b) for both TV and NLTV is obtained by applying the nearest neighbor interpolation on the LL subband. NLTV (d) recover better edges and features than TV (c).

[11] E. J. Candès and J. Romberg. *Practical signal recovery from random projections.* in " Wavelet Applications in Signal and Image Processing XI" (Proc. SPIE Conf.), **5914** (2004).

[12] (MR2236170) E. J. Candès, J. Romberg, and T. Tao. *Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information*, IEEE Trans. Inform. Theory, **52** (2006), 489–509.

[13] A. Castrodad. *Graph-based denoising and classification of hyperspectral imagery using non-local operators*, in "Proc. SPIE", **7334** (2009), 73340E.

[14] T. F. Chan, S. Osher, and J. Shen. *The digital TV filter and nonlinear denoising*, IEEE Trans. Image Process., **10** (2001), 231–241.

[15] T. F. Chan and J. Shen. *Mathematical models for local non-texture inpainting*, SIAM J. Appl. Math, **62** (2001), 1019–1043.

[16] T. F. Chan and J. Shen. *Morphologically invariant PDE inpaintings*, UCLA CAM Report, (2001), 01-15.

[17] (MR2254441) T. F. Chan, J. Shen, and H-M Zhou. *Total variation wavelet inpainting.* J. Math. Imaging Vision, **25** (2006), 107–125.

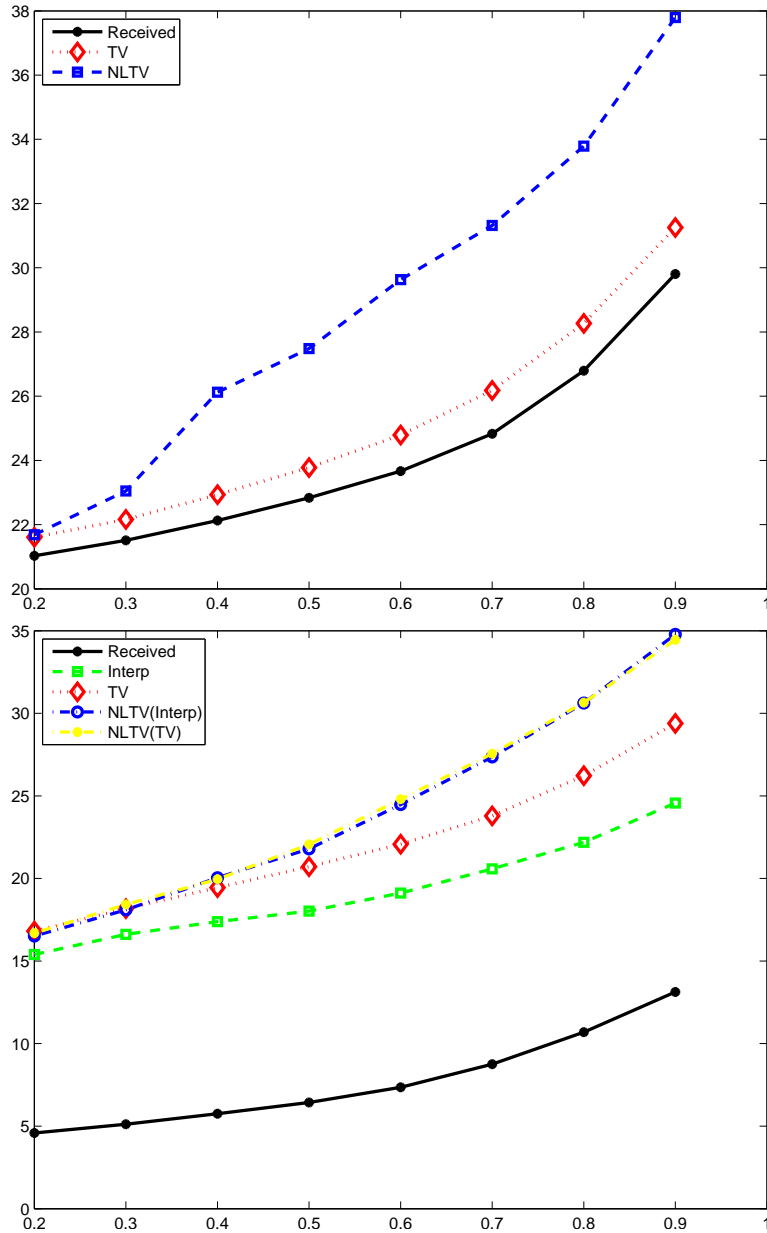[18] (MR1421568) F. R. K. Chung. "Spectral graph theory", American Mathematical Society, Providence, RI, 1997.

FIGURE 8. PSNR *vs* sampling rates (on Barbara 256×256, Figure 6, *nOuter* = 25). Top: random loss keeping LL frequencies. Both TV and NLTV improve the image quality by removing artifacts and recovering clean edges, and NLTV has much higher PSNR than the other two. Bottom: random loss. The interpolation, TV and NLTV can dramatically improve the PSNR by removing dark regions created by LL loss. NLTV regularization with TV and the interpolated one as an initial guess gives similar results and they both perform more effectively than TV.
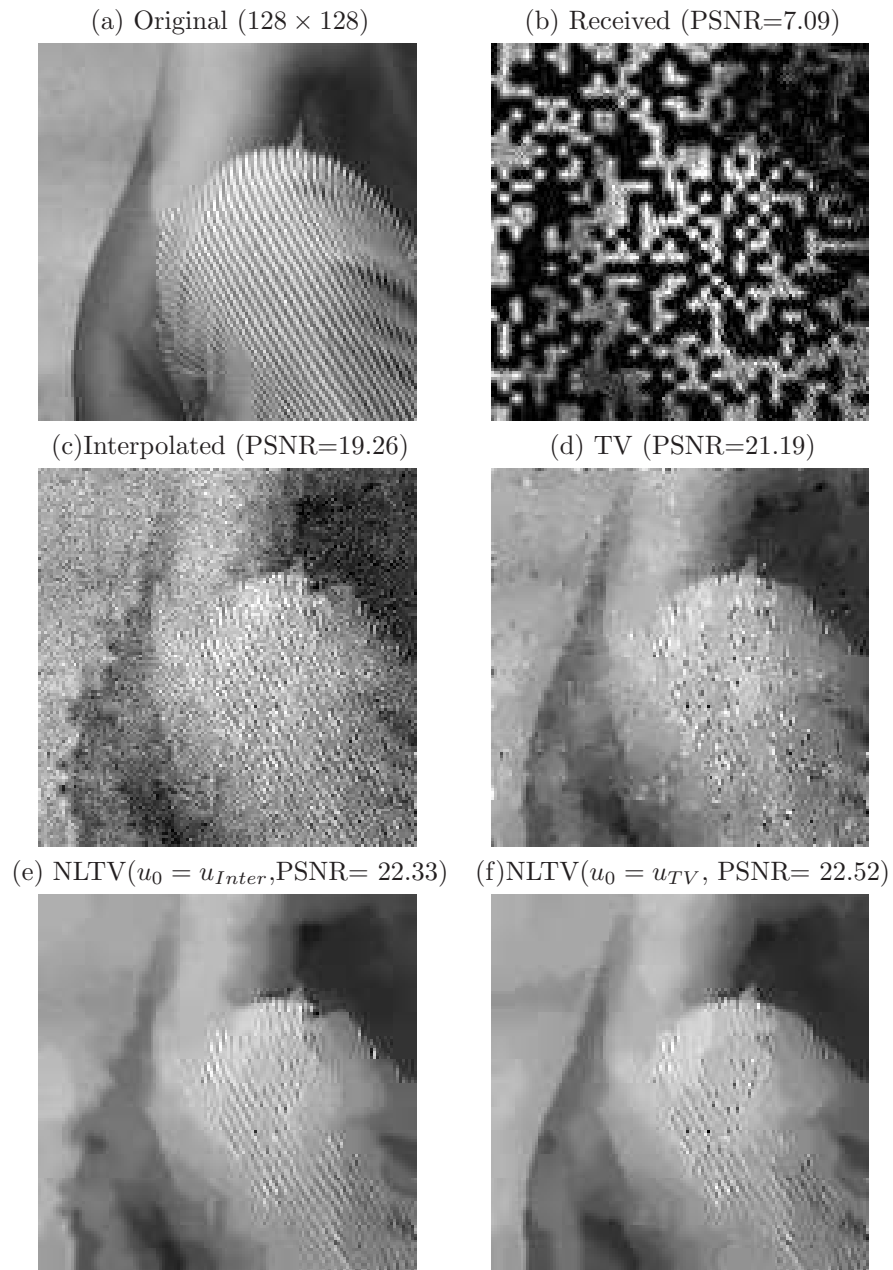
(a) Original (128 × 128)

(b) Received (PSNR=7.09)

(c)Interpolated (PSNR=19.26)

(d) TV (PSNR=21.19)

(e) NLTV($u_0 = u_{Inter}$,PSNR= 22.33)

(f)NLTV($u_0 = u_{TV}$, PSNR= 22.52)

FIGURE 9. Random loss with noise: 50% random kept frequencies, noise level $\sigma = 0.1$. NLTV with different initial guesses have similar PSNR, while (e) has better cloth patterns inherited from (c) and (f) has better edges inherited from (d).

[19] (MR2203849) P. L. Combettes and V. R. Wajs. *Signal recovery by proximal forward-backward splitting*, Multiscale Model. Simul., **4** (2005), 1168–1200 .

[20] J. Darbon, A. Cunha, T.F. Chan, S. Osher, and G.J. Jensen. *Fast nonlocal filtering applied to electron cryomicroscopy*, in "Proceedings of ISBI", (2008), 1331-1334.

[21] J. Darbon and S. Osher. *Fast discrete optimization for sparse approximations and deconvolutions*, UCLA preprint, (2007).

[22] L. Demanet, B. Song, and T. Chan. *Image inpainting by correspondence maps: a deterministic approach*, in "Proc. VLSM conf.", Nice, (2003).

[23] (MR1978159) S. Durand and J. Froment. *Reconstruction of wavelet coefficients using total variation minimization*, SIAM J. Sci. Comput., **24** (2003), 1754–1767.

[24] A. A. Efros and T. K. Leung. *Texture synthesis by non-parametric sampling*, in " Proc. Int. Conf. Computer Vision", **2** (1999), 1022–1038.

[25] M. Elad, P. Milanfar, and R. Rubinstein. *Analysis versus synthesis in signal priors*, Inverse Problems, **23** (2007), 947–968.

[26] (MR2186449) M. Elad, J. Starck, P. Querre, and D. Donoho. *Simultaneous cartoon and texture image inpainting using morphological component analysis (MCA)*. Appl. Comput. Harmon. Anal., **19** (2005), 340–358.

[27] A. Elmoataz, O. Lezoray, and S. Bougleux. *Nonlocal discrete regularization on weighted graphs: a framework for image and manifold processing*, IEEE Trans. Image Process, **17** (2008), 1047–1060.

[28] (MR1925256) S. Esedoglu and J. Shen. *Digital inpainting based on the mumford-shah-euler image model*, European J. Appl. Math., **13** (2002), 353–370.

[29] (MR2338496) G. Gilboa and S. Osher. *Nonlocal linear image regularization and supervised segmentation*, Multiscale Model. Simul., **6** (2007), 595–630.

[30] (MR2480109) G. Gilboa and S. Osher. *Nonlocal operators with applications to image processing*, Multiscale Model. Simul., **7** (2008), 1005–1028.

[31] (MR2496060) T. Goldstein and S. Osher. *The split bregman method for l^1 regularized problems*, SIAM J. Imaging Sci., **2** (2009), 323–343.

[32] E. T. Hale, W. Yin, and Y. Zhang. *A fixed-point continuation method for l^1-regularized minimization with applications to compressed sensing*, SIAM Journal on Optimization, **19** (2008), 1107-1130.

[33] L. He, T.-C. Chung, S. Osher, T. Fang, and P. Speier. *MR image reconstruction by using the iterative refinement method and nonlinear inverse scale space methods*, UCLA CAM report, (2006), 06-35.

[34] S. S. Hemami and R.M. Gray. *Subband coded image reconstruction for lossy packed networks*, IEEE Trans. Image Process., 6 (1997), 1466–1477.

[35] N. Houhou, X. Bresson, A. Szlam, T. F. Chan, and J-P Thiran. *Semi-supervised segmentation based on non-local continuous min-cut*, in "Proceedings of the 2nd International Conference on Scale Space and Variational Methods in Computer Vision", **5567** (2009), 112-123

[36] Y. Lou, X. Zhang, S. Osher, and A. Bertozzi. *Image recovery via nonlocal operators*, Journal of Scientific Computing, 10.1007/s10915-009-9320-2, 2009

[37] (MR1614527) S. Mallat. " A Wavelet Tour of Signal Processing", Academic Press, Inc., San Diego, CA, 1998.

[38] Y. Mao, B. Fahimian, S. Osher, and J. Miao. *Development and optimization of regularized tomographic reconstruction algorithms utilizing equally-sloped tomography*, IEEE Trans. Image Process, to appear.

[39] S. Masnou and J.-M. Morel. *Level lines based disocclusion*, in " International Conference on Image Processing", **3** (1998), 259–263.

[40] (MR0144188) J.-J. Moreau. *Fonctions convexes duales et points proximaux dans un espace hilbertien*, C. R. Acad. Sci. Paris, **255** (1962), 2897–2899.

[41] Y. Niu and T. Poston. *Using an oriented PDE to repair image textures*, in "Variational, Geometric, and Level Set Methods in Computer Vision" (Eds,N. Paragios et al.), Springer Berlin / Heidelberg, **3752** (2005), 61–72.

[42] (MR2162864) S. Osher, M. Burger, D. Goldfarb, J. Xu, and W. Yin. *An iterative regularization method for total variation-based image restoration*, Multiscale Model. and Simu., **4** (2005), 460–489.

[43] S. Osher, Y. Mao, B. Dong, and W. Yin. *Fast linearized bregman iteration for compressive sensing and sparse denoising*, Communications in Mathematical Sciences, **8** (2010), 93-111.

[44] (MR2443009) G. Peyré. *Image processing with nonlocal spectral bases*, Multiscale Model. Simul., **7** (2008), 703–730.

[45] G. Peyré, S. Bougleux, and L. Cohen. *Non-local regularization of inverse problems*, in "Proceedings of ECCV'08", (2008), 57-68.

[46] (MR1982918) S.D. Rane, G. Sapiro, and M. Bertalmio. *Structure and texture filling-in of missing image blocks in wireless transmission and compression*, IEEE Trans. Image Process., **12** (2003), 296–303.

[47] L. I. Rudin, S. Osher, and E. Fatemi. *Nonlinear total variation based noise removal algorithms*, Physica D., 60 (1992), 259–268.

[48] H. Sun and W. Kwok. *Concealment of damaged block transform coded images using projections onto convex sets*, IEEE Trans. Image Process., **4** (1995), 470–477.

[49] (MR2475828) W. Yin, S. Osher, D. Goldfarb, and J. Darbon. *Bregman iterative algorithms for ℓ^1 minimization with applications to compressed sensing*, SIAM J. Imaging Sci., **1** (2008), 143–168.

[50] X. Zhang, M. Burger, X. Bresson, and S. Osher. *Bregmanized nonlocal regularization for deconvolution and sparse reconstruction*, UCLA CAM Report, (2009), 09-03.

[51] X. Zhang and J. Froment. *Constrained total variation minimization and application in computerized tomography*, in "EMMCVPR05", Springer Berlin / Heidelberg, **3757** (2005), 456–472.

[52] D. Zhou and B. Schölkopf. *Regularization on discrete spaces*, Pattern Recognition, **361** (2005), 361–368.

[53] M. Zhu and T. Chan. *An efficient primal-dual hybrid gradient algorithm for total variation image restoration.* UCLA CAM Report, (2008), 08-34.

E-mail address: xqzhang@math.ucla.edu
E-mail address: tonyfchan@ust.hk