IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received July 23, 2019, accepted August 16, 2019, date of publication August 21, 2019, date of current version September 4, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2936711

Image Restoration Models Based on Dyadic Hardy Space and Dyadic Bounded Mean Oscillation Space

TAO ZHANG AND XUTAO MO
School of Mathematics and Physics, Anhui University of Technology, Ma'anshan, China

Corresponding author: Xutao Mo (tao_ahut@163.com)

This work was supported in part by the National Natural Science Foundation of China under Grant 61701004, Grant 11504003, and Grant 11571197, in part by the Natural Science Foundation of Anhui Province under Grant 1708085QA15, and in part by the China Scholarship Council.

ABSTRACT Texture is widely existed in various images and plays an important role in many area such as medical image diagnosis, remote sensing, etc. However, the image in texture regions is tend to be deteriorated during restoration process. In this paper, we apply the dyadic Hardy space H_d^1 and dyadic Bounded Mean Oscillation (BMO) space in the texture preserving image restoration model. We propose a H_d^1 regularized minimization model to extract texture from noisy data. In this model, H_d^1 norm is taken as regularizer to enforce the prior that the local variance of the noise is below certain level depending on the regularization parameter. We also analyze the mathematical properties of this model which indicate the mechanism of H_d^1 regularizer to control the local variance. For the numerical solution of the model, we transform it into wavelet domain based on the wavelet characterization of dyadic Hardy space and dyadic BMO space, and solve it by the fixed iteration algorithm. Combing the total variation (TV) regularization method and frame based regularization method, a two-layers regularization model is proposed for edge and texture preserving, and then analyzed and solved in the frame of split Bregman method. Finally, we present various numerical results on images to demonstrate the potential of our methods.

INDEX TERMS Dyadic *BMO* space, hardy space, image restoration, mixed norm, texture.

I. INTRODUCTION

Texture is an important visual cue in interpreting images, and has been successfully used in many area such as image fusion, medical image diagnosis, biometric identification, remote sensing, etc. Natural images consist of texture, structure and smooth regions, and this makes the task of image restoration challenging when it aims at edge and texture preservation. So the modeling of texture and separating texture from non-texture parts in images play a central role in image restoration, and have been studied by several researchers [1]–[7]. In this paper, we consider the image restoration problem aiming at texture preservation.

Without loss of generality, we assume that the underlying image is grayscale and has a square domain. Let u be an original image, K be a linear operator, n be an additive Gaussian noise, and f be an observation which satisfies the relationship

$$f = Ku + n \tag{1}$$

It is well known that solving u from f is ill-posed inverse problem. It is then necessary to regularize, i.e., to introduce a prior information about the solution. A common approach is to add a regularizer to certain data fidelity, resulting the following reconstruction model:

$$\min_u \ \Phi_{reg}(u) + \lambda \Phi_{fid}(u, f) \tag{2}$$

where $\Phi_{reg}(u)$ regularizes the solution by enforcing certain prior constraints. $\Phi_{fid}(u, f)$ measures the violation of the relationship between u and f.

Traditional regularization methods include the Tikhonov regularization [8] and the total variation (TV) regularization [9], [10]. Tikhonov regularization takes the Sobolev semi norm as the regularizer. Although the minimization problems are easy to solve, Tikhonov regularization tends to make the

The associate editor coordinating the review of this article and approving it for publication was Kathiravan Srinivasan.

restored image overly smoothed and fails to preserve sharp edges. TV regularization, first introduced in [10] by Rudin, Osher, and Fatemi (ROF), takes total variation measure as the regularizer and has been shown to be suitable for preserving sharp edges.

In TV regularization models, images are often assumed to be in Bounded Variation (BV) space. This space is defined as:

$$BV(\Omega) = \{u \in L^1(\Omega) | \, J(u) < \infty\} \tag{3}$$

where

$$J(u) := \sup\{\int_\Omega u \, div \, \xi \, | \xi \in C_c^1(\Omega, R^2), \ \|\xi\|_{L^\infty(\Omega)} \leq 1\}$$

In [10], Rudin-Osher-Fatemi proposed the following model to decompose a noisy image f into a component u belonging in $BV(\Omega)$ and a component v in $L^2(\Omega)$:

$$\min_{u \in BV(\Omega)} \{J(u) + \lambda \|f - u\|_{L^2}^2\} \tag{4}$$

Traditional regularization methods make the assumption that the underlying image is smooth or piecewise constant. It is clear that this hypothesis is not necessarily satisfied for texture. So, there are several obvious drawbacks: significant small details such as textures and even large-scale fine features are often disregarded. Y. Meyer proved that TV regularization rejects the oscillatory component of f which is considered to be the texture component [1]. Meyer on one side, and Mumford-Gidas [11] on the other side advocated the use of generalized functions as distributions in dual spaces for modeling images with oscillations. Meyer suggested the generalized function spaces $G = div(L^\infty)$, $E = B^\infty_{-1,\infty}$ and $F = div(BMO)$ to model the oscillating patterns of the image, and he proposed the $BV - X$ type models, which have the general form:

$$\min_{u \in BV(\Omega)} \{J(u) + \lambda \|f - u\|_X\} \tag{5}$$

where $\| \cdot \|_X$ is the norm of space G, E or F. However, in practice it is not easy to compute the G norm or F norm. To deal with this problem, the following-up works such as Vese-Osher's model [2] and Osher-Solé-Vese's model [3], focus on the numerical computing and approximation of Meyer's model.

Inspired by Vese-Osher and Osher-Solé-Vese models, I. Daubechies et al. proposed a numerically efficient schemes by means of wavelet [12]. They replaced BV penalty term by $B^1_{1,1}$ term, and proposed to minimize the following functional:

$$\mathcal{F}^{Db} = 2\alpha \|u\|_{B^1_{1,1}} + \gamma \|v\|_{H^{-1}(\Omega)} + \|f - u - v\|_{L^2}^2 \tag{6}$$

where $\|\cdot\|_{H^{-1}(\Omega)}$ is the Sobolev semi norm. Since all the norm in functional (6) can be characterized by wavelet coefficients, the minimization of functional (6) can be solved by means of wavelet efficiently.

There are many other efficient approaches to solve Meyer's model. For example, Aujol, Albert, Blanc-Feraud and Chambolle (A^2BC) proposed to minimize the following functionals to solve $BV - G$ model [4] and $BV - E$ model [5]:

$$\mathcal{F}^{A^2BC} = J(u) + J^*(\frac{v}{\mu}) + \frac{1}{2\lambda}\|f - u - v\|_{L^2}^2 \tag{7}$$

$$\mathcal{F}^{BV-E} = J(u) + B^*(\frac{v}{\delta}) + \frac{1}{2\lambda}\|f - u - v\|_{L^2}^2 \tag{8}$$

where J^* is the dual of J, $B(v) = \|v\|_{B^1_{1,1}}$ is the Besov semi-norm, and therefore $B^*(\frac{v}{\delta}) = \chi_{\{\|v\|_{B^\infty_{-1,\infty}} \leq \delta\}}$.

Recently, low patch-rank method and sparsity regularization method are very successful to solve image restoration problems [6], [7], [13]–[15]. In [6], the authors proposed a convex prior named the block nuclear norm (BNN) for characterizing the texture components. The BNN prior is designed based on the observation that the texture enjoys a globally dissimilar but locally well-patterned nature. The BNN based method gives very impressive image decomposition result when the image contains locally well-patterned textures. In [7], the authors proposed to use two appropriate dictionaries for the representation of texture and piecewise smooth part. Both dictionaries are chosen such that they lead to sparse representation over one type of component (either texture or piecewise smooth part). However, it is very hard to choose such dictionaries for a large range of images. Furthermore, most sparsity based methods start by representing signals on a given dictionary, and process the coefficients of expansion individually. Instead of the usual independence assumption behind the l_1 norm minimization, mixed norms explicitly introduce coupling between coefficients [16]. Used as regularization terms in solving image restoration inverse problems, mixed norms can enforce some specific types of joint sparsity and diversity. In this paper, we will show that the discrete dyadic Hardy norm is a mixed norm encoding the group information, and group sparse optimization with this mixed norm can lead to better signal recovering and feature selection.

Meyer's model captures the oscillating components of the image very well. However, both texture and noise are oscillatory pattern and many examples of natural images show that it is difficult to distinguish between texture and noise. Local variance measure or local power is used to distinguish between texture and noise in several literatures [17]–[19]. Its core assumption is that the local variance of texture component is much higher than that of noise components in one image. In this paper, we consider dyadic Hardy H^1_d norm as the regularizer which can enforce the residual satisfying the local variance constraints. Assume γ be an image containing only texture v and noise w. For instance, γ is the texture + noise part of the image obtained by the ROF model $f - u$. We propose the following model (or the equivalent form in frequency domain) to extract clear textures form the noisy image γ:

$$\min_v \{F(v) = \|v\|_{H^1_d} + \frac{1}{2\lambda}\|\gamma - v\|_{L^2}^2\} \tag{9}$$

where $\|\cdot\|_{H_d^1}$ is the norm of dyadic Hardy space H_d^1. We will give detailed introduction about H_d^1 space and H_d^1 norm in Section II.

In our previous work [20], we gave the wavelet characterization of dyadic Hardy norm, and applied model (9) in image decomposition for the first time. But at that time, we did not figure out the mechanism of dyadic Hardy space for texture preserving image restoration. In [20], we solved the model in wavelet domain by Nelder-Mead algorithm, which is a simplex method for finding a local minimizer fo a function of several variables. Nelder-Mead algorithm has several drawbacks: it did not make use of derivative information and the performance of the algorithm has a remarkably dependence on the initial simplex. For these problems of [20], on the one hand, we will give the mechanism of H_d^1 space for modeling the oscillating component of the image in Section III A. On the other hand, we will give an efficient algorithm for solving model (9) based on fixed point iteration method in Section III B.

Combining the advantage of TV regularization, sparse regularization and H_d^1 regularization, we propose a two-layers image restoration model aiming at edge and texture preservation in Section IV. We give the analysis and solution of this model in the frame of Split Bregmen method.

In summary, the main contributions of this work are listed as follows:

(1) We use dyadic Hardy space H_d^1 and dyadic BMO space BMO_d for modeling the oscillating component of the image, and give an efficient algorithm based on fixed point iteration method for solving the H_d^1 norm minimization problem. This enriches the theory of Meyer's oscillating functional space modeling.

(2) Through rigorous mathematical analysis, we establish fundamental properties of the model (9) and give the regularization mechanism of H_d^1 space and BMO_d space for texture preserving image restoration problem.

(3) Combining the advantage of TV regularization, group sparse regularization and H_d^1 regularization, we propose a two- layers $TV - H_d^1$ regularization model for image restoration, and give an efficient algorithm based on split Bregmen method to solve the model.

II. WAVELET CHARACTERIZATION OF DYADIC HARDY SPACE AND DYADIC BMO SPACE

We will further assume that all images are square images. Note that this assumption is not essential, and all the theorems and properties about the proposed models can be easily extended to general cases. About the proposed algorithms for non-square images, we can use the periodically expanding operator to expand the non-square image to larger squared one, and then apply the proposed algorithm on the expanded squared image. Finally, the processed image can be sheared to the original size using shear operator.

A. SOME DEFINITIONS AND PROPERTIES ABOUT HARDY SPACE AND BMO SPACE

In this section, we will give a brief introduction of the atom Hardy space and its dual space BMO. There are many possible ways to define real Hardy spaces, the one that will be most convenient for us uses the atom decomposition.

Definition 1: Let $1 \le q \le \infty$, a measurable function $a(x)$ is called $(1, q)$ atom, if there exists a ball B in R^n, whose volume is denoted by $|B|$, such that the three following properties hold: (1) $supp\ \ a(x) \subset B$; (2) $\|a\|_q \le |B|^{1/q-1}$; (3) $\int_B a(x)dx = 0$

Definition 2: We say that a function $f \in L^1(R^n)$ belongs to atom H^1 if there exists a sequence $a_j(x)$ of $(1, q)$ atoms and a sequence λ_j of scalar coefficients such that $\sum_{j=0}^{\infty} |\lambda_j| < \infty$, and $f(x) = \sum_{j=0}^{\infty} \lambda_j a_j(x)$. i.e.,

$$H_1^{(q)} = \{f \in L^1 | f = \sum_{j=0}^{\infty} \lambda_j a_j(x)\} \tag{10}$$

with the norm defined by:

$$\|f\|_{H_1^{(q)}} = \inf\{\sum_{j=0}^{\infty} |\lambda_k|\} \tag{11}$$

where the infimum is taken over all possible atome decompositon.

Let $\mathcal{Q}$ be the collection of dyadic cubes:

$$\mathcal{Q} = \left\{ Q_{j,k} = Q_{j,k1,k2} \triangleq \left[\frac{k_1}{2^j}, \frac{k_1+1}{2^j}\right] \times \left[\frac{k_2}{2^j}, \frac{k_2+1}{2^j}\right] \right\}$$

If the atom $a_j(x)$ in Def. 2 satisfies a more restrictive condition, namely, that $a(x)$ has support in a dyadic cube $Q \in \mathcal{Q}$ such that $\|a\|_q \le |Q|^{1/q-1}$, and $\int_Q a(x)dx = 0$, the corresponding atom H^1 is called dyadic H^1 space. Dyadic H^1 is isomorphic to the usual H^1 space as a Banach space. For simplicity, we will denote the dyadic H^1 space by H_d^1.

The dual space of H_d^1 is the dyadic Bounded Mean Oscillating (BMO) space, which is defined as follows.

Definition 3: Dyadic BMO space BMO_d consists of the functions which are locally square integrable and satisfy the condition:

$$\|f\|_{BMO_d} = \sup_{Q_{j,k} \in \mathcal{Q}} \left(\frac{1}{|Q_{j,k}|} \int_{Q_{j,k}} |f(x) - m_{Q_{j,k}} f|^2 dx \right)^{\frac{1}{2}} < +\infty$$

where the upper bound is taken over the set of all dyadic cubes.

From above definition, one can see that the dyadic BMO norm is the maximum of the local standard deviation over all dyadic cubes.

For every $b(x) \in BMO_d$, it defines a linear functional l on H_d^1 by $l_b(f) = \langle f, b\rangle \triangleq \sum_{j=0}^{\infty} \lambda_j \int_\Omega b(x) a_j(x) dx$, where $a_j(x)$

are atoms, $\sum_{j=0}^{\infty}|\lambda_j|<\infty$, and $f(x)=\sum_{j=0}^{\infty}\lambda_j a_j(x)$. Conversely, every continuous linear function on H_d^1 is defined in this way.

Remark 4: For $f\in H_d^1$, $g(x)\in BMO_d$, we have

$$|l_g(f)|=|\langle f,g\rangle|\le\|f\|_{H_d^1}\|g\|_{BMO_d} \tag{12}$$

Proof: Assume $f=\sum_{k=0}^{\infty}\lambda_k a_k$ is any atom decomposition of f, and a_k is a (1,2) atom supported on dyadic cube Q. Then

$$\begin{aligned}|l_g(a_k)|&=|\int_Q a_k g dx|=|\int_Q (a_k-m_Q(a_k))g dx|\\&=|\int_Q a_k(g-m_Q(g))dx|\\&\le\|a_k\|_2\left(\int_Q|g-m_Q(g)|^2dx\right)^{\frac12}\\&\le\left(\frac{1}{|Q|}\int_Q|g-m_Q(g)|^2dx\right)^{\frac12}\\&\le\|g\|_{BMO_d}\end{aligned}$$

where $m_Q(a_k)=\frac{1}{|Q|}\int_Q a_k(x)dx$, $m_Q(g)=\frac{1}{|Q|}\int_Q g(x)dx$.

So we have

$$\begin{aligned}|l_g(f)|&=|\int_\Omega f(x)g(x)dx|\\&=|\sum_{k=0}^{\infty}\lambda_k\int_\Omega a_k(x)g(x)dx|\\&\le\sum_{k=0}^{\infty}|\lambda_k|\cdot\|g\|_{BMO_d}\end{aligned} \tag{13}$$

Take infimum for all the decompositions of f, we can get

$$|l_g(f)|=|\int_\Omega f(x)g(x)dx|\le\|f\|_{H_d^1}\|g\|_{BMO_d} \tag{14}$$

B. WAVELET CHARACTERIZATION OF H_D^1 NORM AND BMO_D NORM

Let ϕ and ψ be univariate wavelet constructed out of V_j, an r-regular multiresolution approximation of $L^2(R^2)$ with $r\ge 1$(for the exact conditions see [21]). Two dimensional wavelets can be constructed by tensor product:

$$\begin{aligned}\psi_Q^\varepsilon(x_1,x_2)&:=\psi_{j,k}^\varepsilon(x_1,x_2)\\&=2^j\psi^{\varepsilon_1}(2^jx_1-k_1)\cdot\psi^{\varepsilon_2}(2^jx_2-k_2),\end{aligned}$$

where $j\in\mathbb{Z},k=(k_1,k_2)\in\mathbb{Z}^2,\varepsilon\in\mathcal{E}:=\{\varepsilon_1,\varepsilon_2\}^2/\{0,0\}$ with $\varepsilon_j=0$ or 1, $\psi^0=\phi$ and $\psi^1=\psi$.

In this paper, our analysis is based on interpreting the image as a function f defined on the unit square $I=[0,1)^2$. Each cube $Q\in\mathcal{Q}$ is of the form $Q=2^{-k}(j+I)$. So the dyadic cube can be identified with (j,k).

One can easily construct periodic wavelets on $L^2(I)$ that can be used to decompose periodic functions f on $L^2(I)$. For the wavelet $\psi_{j,k}^\varepsilon$ we discussed above, we define its periodic version which is still denoted here by

$$\psi_Q^\varepsilon(x_1,x_2):=\tilde\psi_{j,k}^\varepsilon=\sum_{l\in\mathbb{Z}^2}\psi_{j,k}^\varepsilon(x-l) \tag{15}$$

One can show that these periodic wavelets form an orthogonal basis of $L^2(I)$. Because the translates of the scaling function form a partition of unity, we can get $\tilde\phi=1$. For the wavelet decomposition of functions in $L^2(I)$, we don't need all translates of these periodic wavelets. On the level j, we only need the translations $k\in\Gamma_j=\{0,1,\cdots,2^j-1\}^2$. The wavelet expansion for a function $f\in L^2(I)$ is

$$f=\langle f,1\rangle 1+\sum_{j=0}^{\infty}\sum_{k\in\Gamma_j}\sum_{\varepsilon\in\mathcal{E}}\langle f,\tilde\psi_{j,k}^\varepsilon\rangle\tilde\psi_{j,k}^\varepsilon \tag{16}$$

and

$$\|f\|_{L^2}^2=|\langle f,1\rangle|^2+\sum_{j=0}^{\infty}\sum_{k\in\Gamma_j}\sum_{\varepsilon\in\mathcal{E}}|\langle f,\tilde\psi_{j,k}^\varepsilon\rangle|^2 \tag{17}$$

So it follows that the constant function 1 together with the collection $\{\tilde\psi_{j,k}^\varepsilon,j\in\mathbb{Z},k\in\Gamma_j\}$ constitute an orthonormal basis of $L^2(I)$. This basis is also an unconditional basis of Hardy space $H^1(I)$. Let V_j be an r-regular multiresolution approximation of $L^2(I)$ with $r\ge 1$, then the sequence $\{1,\tilde\psi_{j,k}^\varepsilon,j\in\mathbb{N},k\in\Gamma_j\}$, of periodic wavelets constructed out of the V_j, is an unconditional basis of Hardy space $H^1(I)$.

In [21], the author gave five definitions of $H^1(R^n)$ including the atomic definition of Definition 2.1 and proved that these five definitions are equivalent. In this paper, we consider the following dyadic H^1 norm:

$$\|f\|_{H_d^1}=\left\|\left(\sum_{\varepsilon\in\mathcal{E}}\sum_{Q\in\mathcal{Q}}|Q|^{-1}|\langle f,\psi_Q^\varepsilon\rangle|^2\chi_Q(x)\right)^{\frac12}\right\|_{L^1} \tag{18}$$

We define a map T, which takes distribution f to the sequence of coefficients $\{\langle f,\psi_Q^\varepsilon\rangle\}_{Q,\varepsilon\in\mathcal{E}}$. Let h_d^1 be the collection of all sequences $\alpha=\{\alpha_Q^\varepsilon\}_{Q,\varepsilon\in\mathcal{E}}$ so that

$$\|\alpha\|_{h_d^1}=\left\|\left(\sum_{\varepsilon\in\mathcal{E}}\sum_{Q\in\mathcal{Q}}|Q|^{-1}|\alpha_Q^\varepsilon|^2\chi_Q(x)\right)^{\frac12}\right\|_{L^1} \tag{19}$$

It is a basic result shown in [22] that $f=\sum_\varepsilon\sum_Q\langle f,\psi_Q^\varepsilon\rangle\psi_Q^\varepsilon$ and $f\in H_d^1$ if and only if $Tf=\{\langle f,\psi_Q^\varepsilon\rangle\}_{Q,\varepsilon}\in h_d^1$.

Let $f\in L^2(I)\bigcap H^1(I)$. Choose N big enough such that $f\in V_{N+1}$. In [20], [23], we give the discrete wavelet representation of dyadic H^1 norm and dyadic *BMO* norm, which can be described as the following.

$$\|f\|_{H_d^1}=2^{-2N}\sum_{k\in\Gamma_N}\left(\sum_{\varepsilon\in\mathcal{E}}\sum_{j=0}^{N}|f_{j,\Lambda^{N-j}(k)}^\varepsilon|^2\cdot 2^{2j}\right)^{\frac12} \tag{20}$$

$$\|f\|_{BMO_d}=\sup_{j,k}\left(2^{2j}\sum_{\varepsilon\in\mathcal{E}}\sum_{j'\ge j}\sum_{k'\in\Lambda^{j-j'}(k)}|f_{j',k'}^\varepsilon|^2\right)^{\frac12} \tag{21}$$

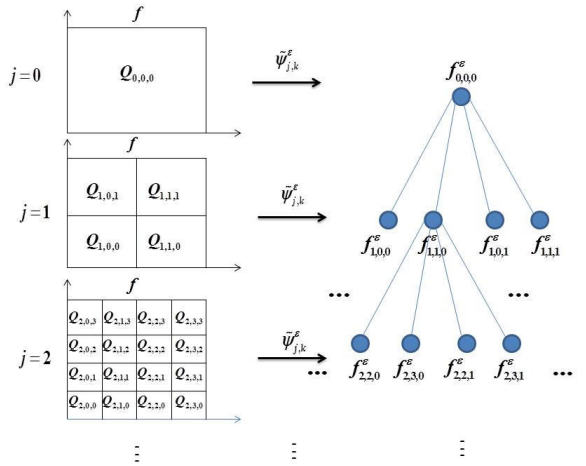


FIGURE 1. Tree-structured wavelet coefficients.

where $f^{\varepsilon}_{j,\Lambda^{N-j}(k)} = \langle f, \tilde{\psi}^{\varepsilon}_{j,\Lambda^{N-j}(k)}\rangle, f^{\varepsilon}_{j',k'} = \langle f, \tilde{\psi}^{\varepsilon}_{j',k'}\rangle$, and the supremum in Eq. (14) runs over $j \in \mathbb{Z}$, $k = (k_1, k_2) \in \Gamma_j$. $\Lambda^{j-j'}(k)$ is defined as following

$$\Lambda^{j-j'}(k) = \begin{cases} l\{k' : \lfloor \frac{k'}{2^{j'-j}} \rfloor = k\}, & j - j' \le 0 \\ \lfloor \frac{k}{2^{j-j'}} \rfloor, & j - j' \ge 0 \end{cases} \tag{22}$$

We use the following notation to rearrange the wavelet coefficients in (20). Let the vector $\alpha \in l_2(\mathbb{R}^{2^{2N}})$ be such that $\alpha = (\alpha_1, \alpha_2, \cdots, \alpha_k, \cdots)$ and for all k, $\alpha_k = \{2^j f^{\varepsilon}_{j,\Lambda^{N-j}(k)}\}_{j=0,\cdots,N.\varepsilon\in\mathcal{E}}$. Then the H^1_d norm can be rewritten as:

$$\|f\|_{H^1_d} = \|\alpha\|_{h^1_d} = \sum_{k\in\Gamma_N} \|\alpha_k\|_2 = \|\alpha\|_{l_{2,1}} \tag{23}$$

Note that the wavelet coefficients in group α_k have parent-child relationship with a tree structure[see Fig. 1]. A pair of coefficients at a certain location and adjacent scales are typically both large or small in amplitude. So this group encoding introduces some dependencies between coefficients. It is well known that texture usually has periodic pattern and the corresponding wavelet coefficients are much more strongly dependent than that of noise component. So we want to construct the group sparsity model using this tree-structure existed in wavelet domain. Based on the wavelet characterization of H^1_d norm, solving model (9) is equivalent to solving the following model in wavelet domain:

$$\min_{\alpha}\{\|\alpha\|_{h^1_d} + \frac{1}{2\lambda}\|\beta - \alpha\|_2^2\} \tag{24}$$

where $\alpha = Tv$, $\beta = T\gamma$, T is the wavelet transform. We will give the solution of above model in Section B.

III. H^1_D MINIMIZATION PROBLEM

A. SOME MATHEMATICAL PROPERTIES OF MODEL (9)

Proposition 5: Problem (9) admits a unique solution $v^* \in H^1_d$.

Proof: The existence of a solution for problem (9) is standard. It is a straightforward consequence of the fact that the functional in (9) is strictly convex and hence coercive, see e.g. [[24], Proposition 11.16].

Assume that problem (9) has two minimizer v_1 and v_2. we denote by $F(v_1) = F(v_2) = \inf_{u\in H^1_d} F(u) = M$. If $t \in (0, 1)$, then we get:

$$F(tv_1 + (1-t)v_2) = \frac{1}{2\lambda}\|t(\gamma - v_1) + (1-t)(\gamma - v_2)\|^2_{L^2} + \|tv_1 + (1-t)v_2\|_{H^1_d} \tag{25}$$

But by convexity, we have

$$\|tv_1 + (1-t)v_2\|_{H^1_d} \le t\|v_1\|_{H^1_d} + (1-t)\|v_2\|_{H^1_d} \tag{26}$$

and

$$\|t(\gamma - v_1) + (1-t)(\gamma - v_2)\|^2_{L^2} \le t\|\gamma - v_1\|^2_{L^2} + (1-t)\|\gamma - v_2\|^2_{L^2} \tag{27}$$

By substituting (26)-(27) into (25), we deduce that

$$F(tv_1 + (1-t)v_2) \le tM + (1-t)M = M \tag{28}$$

and (28) is an equality if and only if (26)-(27) are equalities. But on the other hand, we have $F(tv_1 + (1-t)v_2) \ge M$. Therefore, (28) must be an equality, as well as (26)-(27).

The functional in (27) is strictly convex. Therefore, (27) is an equality if and only if $\gamma - v_1 = \gamma - v_2$, i.e., $v_1 = v_2$. Then we get the uniqueness. □

Proposition 6: Let v^* be a solution of model (9). Then $v^* = 0$ if and only if $\|\gamma\|_{BMO_d} \le \lambda$. Thus, if $\|\gamma\|_{BMO_d} > \lambda$, then $v^* \ne 0$, and $\|\gamma - v^*\|_{BMO_d} = \lambda$, $\langle v^*, \gamma - v^*\rangle = \lambda\|v^*\|_{H^1_d}$.

Proof: $v^* = 0$ being the minimizer of (9) if and only if for any $v \in H^d_1$

$$\|v\|_{H^d_1} + \frac{1}{2\lambda}\|v - \gamma\|^2_{L^2} \ge \frac{1}{2\lambda}\|\gamma\|^2_{L^2} \tag{29}$$

Assume that $\|\gamma\|_{BMO_d} \le \lambda$. By Remark 2.1, we can get

$$|\langle v, \gamma\rangle| \le \|v\|_{H^1_d}\|\gamma\|_{BMO_d} \le \lambda\|v\|_{H^1_d} \tag{30}$$

So for any $v \in H^d_1$,

$$\begin{aligned} &\|v\|_{H^d_1} + \frac{1}{2\lambda}\|v - \gamma\|^2_{L^2} \\ &= \|v\|_{H^d_1} + \frac{1}{2\lambda}(\|\gamma\|^2_{L^2} - 2\langle v, \gamma\rangle + \|v\|^2_{L^2}) \\ &\ge \frac{1}{2\lambda}(\|\gamma\|^2_{L^2} + \|v\|^2_{L^2}) \\ &\ge \frac{1}{2\lambda}\|\gamma\|^2_{L^2} \end{aligned} \tag{31}$$

which implies that $v^* = 0$ is the minimizer.

On the other hand, if $v^* = 0$ is the minimizer, we expand the second term of the left side of (29) and have

$$\|v\|_{H^1_d} + \frac{1}{2\lambda}\|v\|^2_{L^2} \ge \frac{1}{\lambda}\langle v, \gamma\rangle \tag{32}$$

By substituting in (32) v with εv, and take $\varepsilon \to 0^+$, we have

$$\|v\|_{H^1_d} \ge \frac{1}{\lambda}\langle v, \gamma\rangle, \quad \text{i.e.,} \quad \frac{\langle v, \gamma\rangle}{\|v\|_{H^1_d}} \le \lambda$$

which implies

$$\|\gamma\|_{BMO_d} \leq \lambda \tag{33}$$

From the first assertion, we can deduce directly if $\|\gamma\|_{BMO_d} > \lambda$, then $v^* \neq 0$.

Since v^* is the solution of (9), for any $h \in H_d^1$, and $\varepsilon \in R$, we have

$$\|v^* + \varepsilon h\|_{H_d^1} + \frac{1}{2\lambda}\|\gamma - v^* - \varepsilon h\|_{L^2}^2 \geq \|v^*\|_{H_d^1} + \frac{1}{2\lambda}\|\gamma - v^*\|_{L^2}^2 \tag{34}$$

By the triangle inequality, $|\varepsilon|\|h\|_{H_d^1} + \frac{1}{2\lambda}\|\gamma - v^* - \varepsilon h\|_{L^2}^2 \geq \frac{1}{2\lambda}\|\gamma - v^*\|_{L^2}^2$.

Expanding the second term of above inequality, we obtain

$$|\varepsilon|\|h\|_{H_d^1} + \frac{1}{2\lambda}|\varepsilon|^2\|h\|_{L^2}^2 \geq \frac{\varepsilon}{\lambda}\langle h, \gamma - v^*\rangle \tag{35}$$

Dividing both side of the last equation by $\varepsilon > 0$, and taking $\varepsilon \to 0^+$, we have $\|h\|_{H_d^1} \geq \frac{1}{\lambda}\langle h, \gamma - v^*\rangle$, i.e., $\|\gamma - v^*\|_{BMO_d} \leq \lambda$

Taking $h = v^*$ in (34). Then (34) implies

$$\frac{\lambda}{\varepsilon}\langle v^*, \gamma - v^*\rangle - \frac{1}{2\lambda}\varepsilon^2\|v^*\|_{L^2}^2 \leq (|1+\varepsilon| - 1)\|v^*\|_{H_d^1} \tag{36}$$

If $\varepsilon > 0$, dividing both side of (36) by ε and then take $\varepsilon \to 0^+$, we obtain $\frac{1}{\lambda}\langle v^*, \gamma - v^*\rangle \leq \|v^*\|_{H_d^1}$.

If $\varepsilon < 0$, similarly, we obtain $\frac{1}{\lambda}\langle v^*, \gamma - v^*\rangle \geq \|v^*\|_{H_d^1}$. Altogether it gives

$$\langle v^*, \gamma - v^*\rangle = \lambda\|v^*\|_{H_d^1} \tag{37}$$

By (33), $\|\gamma - v^*\|_{BMO_d} = \sup_{v \in H_d^1} \frac{\langle v, \gamma - v^*\rangle}{\|v\|_{H_d^1}} \leq \lambda$. This implies $\|\gamma - v^*\|_{BMO_d} = \lambda$. Otherwise, equality could not reach in (37). □

From Proposition 6, we can see that the local variance of the residual $\gamma - v$ can be controlled by parameter λ. Thus, by choosing suitable parameter λ, model (9) can not only remove noise, but also can ensure that there are not too much high oscillatory texture in the residual.

B. SOLVING MODEL (9) IN WAVELET DOMAIN

By (17) and (20), model (9) can be transformed into wavelet domain. The functional in (9) is represented by

$$\begin{aligned} &F(\{v_{j,k}^{\varepsilon}\}_{0\leq j\leq N, k\in\Gamma_N}) \\ &= 2^{-2N}\sum_{k_1,k_2=0}^{2^N-1}\left(\sum_{\varepsilon\in\mathcal{E}}\sum_{j=0}^{N}|v_{j,\Lambda^{N-j}(k_1),\Lambda^{N-j}(k_2)}^{\varepsilon}|^2\cdot 2^{2j}\right)^{\frac{1}{2}} \\ &\quad + \frac{1}{2\lambda}\sum_{j=0}^{N}\sum_{k_1,k_2=0}^{2^N-1}\sum_{\varepsilon\in\mathcal{E}}|v_{j,k_1,k_2}^{\varepsilon} - \gamma_{j,k_1,k_2}^{\varepsilon}|^2 \end{aligned} \tag{38}$$

For every $0 \leq j \leq N, k = (k_1, k_2) \in \Gamma_j$, $\varepsilon \in \mathcal{E}$, we have

$$\frac{\partial F}{\partial v_{j,k}^{\varepsilon}} = \frac{1}{\lambda}(v_{j,k}^{\varepsilon} - \gamma_{j,k}^{\varepsilon}) + \sum_{K'\in\Lambda^{j-N}(k)}\frac{2^{-2(N-j)}v_{j,k}^{\varepsilon}}{\sqrt{\sum_{l=0}^{N}\sum_{\varepsilon\in\mathcal{E}}|v_{l,\Lambda^{N-l}(K')}^{\varepsilon}|^2 2^{2l}}}$$

Finding the minimizer of model (38) is reduced to solving $\frac{\partial F}{\partial v_{j,k}^{\varepsilon}} = 0$, i.e.

$$v_{j,k}^{\varepsilon} = \gamma_{j,k}^{\varepsilon} - \sum_{K'\in\Lambda^{j-N}(k)}\frac{2^{-2(N-j)}\lambda v_{j,k}^{\varepsilon}}{\sqrt{\sum_{l=0}^{N}\sum_{\varepsilon\in\mathcal{E}}|v_{l,\Lambda^{N-l}(K')}^{\varepsilon}|^2 2^{2l}}} \tag{39}$$

Particularly, when $j = N$, for every $K \in \Gamma_N$, we have

$$v_{N,K}^{\varepsilon} = \gamma_{N,K}^{\varepsilon} - \lambda v_{N,K}^{\varepsilon}\frac{1}{\sqrt{\sum_{l=0}^{N}\sum_{\varepsilon\in\mathcal{E}}|v_{l,\Lambda^{N-l}(K')}^{\varepsilon}|^2 2^{2l}}} \tag{40}$$

By (40), we know that $v_{N,K}^{\varepsilon} = 0$ if and only if $\gamma_{N,K}^{\varepsilon} = 0$. So if $\gamma_{N,K}^{\varepsilon} \neq 0$, (40) is equivalent to

$$\Leftrightarrow \quad \frac{1}{\sqrt{\sum_{l=0}^{N}\sum_{\varepsilon\in\mathcal{E}}|v_{l,\Lambda^{N-l}(K')}^{\varepsilon}|^2 2^{2l}}} = \frac{1}{\lambda}\left(\frac{\gamma_{N,K}^{\varepsilon}}{v_{N,K}^{\varepsilon}} - 1\right) \tag{41}$$

Taking sum on both side of equation (41) for all $K' \in \Lambda^{j-N}(k)$, we have

$$\sum_{K'}\frac{1}{\sqrt{\sum_{l=0}^{N}\sum_{\varepsilon\in\mathcal{E}}|v_{l,\Lambda^{N-l}(K')}^{\varepsilon}|^2 2^{2l}}} = \sum_{K'}\frac{1}{\lambda}\left(\frac{\gamma_{N,K}^{\varepsilon}}{v_{N,K}^{\varepsilon}} - 1\right) \tag{42}$$

Combining (42) with (39), we can get

$$v_{j,k}^{\varepsilon} = \gamma_{j,k}^{\varepsilon} - 2^{-2(N-j)}\lambda v_{j,k}^{\varepsilon}\sum_{K'\in\Lambda^{j-N}(k)}\frac{1}{\lambda}\left(\frac{\gamma_{N,K}^{\varepsilon}}{v_{N,K}^{\varepsilon}} - 1\right) \tag{43}$$

As $\sharp\Lambda^{j-N}(k) = 2^{2(N-j)}$, equation (43) can be further simplified as follows:

$$v_{j,k}^{\varepsilon} = \frac{2^{2(N-j)}\gamma_{j,k}^{\varepsilon}}{\sum_{K'\in\Lambda^{j-N}(k)}\frac{\gamma_{N,K'}^{\varepsilon}}{v_{N,K'}^{\varepsilon}}} \tag{44}$$

That is, wavelet coefficients $v_{j,k}^{\varepsilon}$ at all scales can be represented by the wavelet coefficients $v_{N,K}^{\varepsilon}$ at the Nth level.

By (44), we can get that for all $0 \leq l \leq N$, $K \in \Gamma_N$,

$$v_{l,\Lambda^{N-l}(K)}^{\varepsilon} = \frac{2^{N-l}\gamma_{l,\Lambda^{N-l}(K)}^{\varepsilon}}{\sum_{K'\in\triangle}\frac{\gamma_{N,K'}^{\varepsilon}}{v_{N,K'}^{\varepsilon}}} \tag{45}$$

where

$$\triangle = \Lambda^{l-N}(\Lambda^{N-l}(K)) = \{K' : \Lambda^{N-l}(K') = \Lambda^{N-l}(K)\}$$

Combining (45) with (41), we have

$$\frac{1}{\sqrt{\sum_{l=0}^{N}\sum_{\varepsilon\in\mathcal{E}}\frac{2^{4N-2l}|\gamma^{\varepsilon}_{l,\Lambda^{N-l}(K)}|^2}{(\sum_{K'\in\triangle}\frac{\gamma^{\varepsilon}_{N,K'}}{v^{\varepsilon}_{N,K'}})^2}}}=\frac{1}{\lambda}(\frac{\gamma^{\varepsilon}_{N,K}}{v^{\varepsilon}_{N,K}}-1) \tag{46}$$

Denote $x^{\varepsilon}_K=\frac{\gamma^{\varepsilon}_{N,K}}{v^{\varepsilon}_{N,K}}$, then equation (46) can be simplified as follows:

$$\frac{1}{\sqrt{\sum_{l=0}^{N}\sum_{\varepsilon\in\mathcal{E}}\frac{2^{4N-2l}|\gamma^{\varepsilon}_{l,\Lambda^{N-l}(K)}|^2}{(\sum_{K'\in\triangle}x^{\varepsilon}_{K'})^2}}}=\frac{1}{\lambda}(x^{\varepsilon}_K-1)\quad\forall K\in\Gamma_N \tag{47}$$

(47) is a system of nonlinear equations, which can be solved by fixed point iteration method. We use the following iterative scheme:

$$(x^{\varepsilon}_K)^n=1+\frac{\lambda}{\sqrt{\sum_{l=0}^{N}\sum_{\varepsilon\in\mathcal{E}}\frac{2^{4N-2l}|\gamma^{\varepsilon}_{l,\Lambda^{N-l}(K)}|^2}{(\sum_{K'\in\triangle}(x^{\varepsilon}_{K'})^{n-1})^2}}}\quad\forall K\in\Gamma_N \tag{48}$$

where $n=1,2,\cdots,(x^{\varepsilon}_K)^0$ is the iterative initial value.

Through above iterative scheme, we can get the solution of nonlinear equations (47) and further obtain the Nth level wavelet coefficients $v^{\varepsilon}_{N,K}$. Substituting it into equation (44), we can get all the wavelet coefficients, which is denoted by

$$v^{\varepsilon}_{j,k}=\frac{\gamma^{\varepsilon}_{j,k}}{(\sharp\Lambda^{j-N}(k))^{-1}\sum_{K'\in\Lambda^{j-N}(k)}x^{\varepsilon}_{K'}} \tag{49}$$

where $j=0,\cdots,N,\ k\in\Gamma_j,\ \varepsilon\in\mathcal{E}$. Through wavelet reconstruction, we can get the solution of the model (9). We define the operator W_λ by

$$W_\lambda(\gamma^{\varepsilon}_{j,k})=\frac{\gamma^{\varepsilon}_{j,k}}{(\sharp\Lambda^{j-N}(k))^{-1}\sum_{K'\in\Lambda^{j-N}(k)}x^{\varepsilon}_{K'}} \tag{50}$$

Then the solution of model (9) is given by

$$v=\sum_{j,k,\varepsilon}W_\lambda(\gamma^{\varepsilon}_{j,k})\tilde{\psi}^{\varepsilon}_{j,k} \tag{51}$$

As a summary for solving model (9) in wavelet domain, we give the following algorithm in Table 1:

C. SOLVING THE MODEL BY DUAL METHOD

Theorem 7: The minimizer of the functional F in model (9) is given by

$$v^*=\gamma-P_{\lambda E}(\gamma)=(I-P_{\lambda E})(\gamma) \tag{52}$$

where I is the identity operator, P_E is the projection operator onto the convex set

$$E=\{w\in BMO_d:\ \langle v,w\rangle\le\|v\|_{H^1_d},\quad\forall v\in H^1_d\}$$

The set E is characterized in the following Lemma:

Lemma 8 (Characterization of the Set E): $E=\{w\in BMO_d:\ \|w\|_{BMO_d}\le 1\}$.

Proof: Denote $C_1=\{w\in BMO_d:\ \langle v,w\rangle\le\|v\|_{H^1_d},\ \forall v\in H^1_d\}$, $C_2=\{w\in BMO_d:\ \|w\|_{BMO_d}\le 1\}$. We now prove $C_1=C_2$.

"$C_1\supseteq C_2$": Let $w\in C_2$. Then $\|w\|_{BMO_d}\le 1$. By Remark 2.1, we have $\langle v,w\rangle\le\|v\|_{H^1_d}\|w\|_{BMO_d}\le\|v\|_{H^1_d}$. Therefore, $C_1\supseteq C_2$.

"$C_1\subseteq C_2$": If $w\notin C_2$, we have $\|w\|_{BMO_d}>1$. we denote $\|w\|_{BMO_d}=m$. By the definition of dyadic BMO norm, we know that for arbitrary $\varepsilon>0$, there exist a dyadic cube Q, such that

$$m>\left(\frac{1}{|Q|}\int_Q(w-m_Q(w))^2dx\right)^{\frac{1}{2}}>m-\varepsilon \tag{53}$$

We choose $a:=a(x)=\frac{1}{m|Q|}(w-m_Q(w))\chi_Q(x)$. Now we check that $a(x)$ satisfies the three properties for (1, 2) atom in Definition 1:

(1) $supp\ a(x)\subset Q$.

$$\begin{aligned}(2)\ \|a(x)\|_2&=\left(\int_Q\frac{1}{m^2|Q|^2}(w-m_Q(w))^2dx\right)^{\frac{1}{2}}\\&=\frac{1}{m}|Q|^{-1/2}\left(\frac{1}{|Q|}\int_Q(w-m_Q(w))^2dx\right)^{\frac{1}{2}}\\&<|Q|^{-1/2}\end{aligned}$$

(3) $\int_Q a(x)dx=\frac{1}{m|Q|}\int_Q(w-m_Q(w))dx=0$.

So $a(x)$ is an atom of H^1_d, and $\|a(x)\|_{H^1_d}=1$.

Then

$$\begin{aligned}\langle a(x),w\rangle&=|\int_Q a\cdot wdx|=|\int_Q(a-m_Q(a))wdx|\\&=\int_Q a(w-m_Q(w))dx\\&=\frac{1}{m}\frac{1}{|Q|}\int_Q\big(w-m_Q(w)\big)^2dx\\&>\frac{1}{m}(m-\varepsilon)^2\end{aligned} \tag{54}$$

Taking $\varepsilon\to 0^+$, we can get that $\langle a,w\rangle\ge m>1=\|a\|_{H^1_d}$. So $w\notin C_1$. This implies that $C_1\subseteq C_2$. □

Before proving Theorem 7, we recall some facts from convex analysis.

Definition 9: Let X be a Banach space and $\Phi:X\to\bar{R}$ a functional. The Fenchel-Legendre transform of Φ is $\Phi^*:X^*\to\bar{R}$ and is defined by

$$\Phi^*(x^*)=\sup_{x\in X}\{\langle x,x^*\rangle_{X\times X^*}-\Phi(x)\}$$

Lemma 10: Let $\Phi:X\to\bar{R}$ be positive homogeneous of degree one. Then the Fenchel-Legendre transform of Φ has the form

$$\Phi^*(v)=\begin{cases}l0, & \text{if } v\in E\\ +\infty, & \text{if } v\notin E\end{cases} \tag{55}$$

TABLE 1. Algorithm of model (9).

Algorithm 1:
Input: $\lambda > 0$, $\gamma^{\varepsilon}_{j,k} = \langle \gamma, \psi^{\varepsilon}_{j,k} \rangle$, small tolerance $T > 0$. Initialize $(x^{\varepsilon}_K)^0 = 1$.
While $\|(x^{\varepsilon}_K)^{n+1} - (x^{\varepsilon}_K)^n\| > T$, Do
1. Fix point iteration according to (48) and get the fix point x^{ε}_K.
2. Compute all level wavelet coefficients $v^{\varepsilon}_{j,k}$ by (49).
3. Reconstruct and get the solution of model (9) by (50) and (51).
End Do

TABLE 2. Algorithm of model (60).

Algorithm 2:
Initialization: set $u_1^0 = u_2^0 = 0, d_1^0 = D_1 f, d_2^0 = D_2 0, d^0 = 0, b_1^0 = 0, b_2^0 = 0, b^0 = 0$;
Iteration: For $k = 0, 1, 2, \cdots$, until a stopping criterion is reached
$u_1^{k+1} = (\lambda I + K^T K + \lambda \nabla^T \nabla)^{-1}(K^T f - K^T K u_2^k + \lambda D_1^T (d_1^k - b_1^k) + \lambda \nabla^T (d^k - b^k))$
$u_2^{k+1} = (K^T K + I)^{-1}(K^T f - K^T K u_1^{k+1} + \lambda D_2^T (d_2^k - b_2^k))$
$d_1^{k+1} = shrink(D_1 u_1^{k+1} + b_1^k, \frac{\mu_1}{\lambda})$
$d_2^{k+1} = W_{\frac{\mu_2}{\lambda}}(D_2 u_2^{k+1} + b_2^k)$
$d^{k+1} = shrink(\nabla u_1^{k+1} + b^k, \frac{\theta}{\lambda})$
$b_1^{k+1} = b_1^k + (D_1 u_1^{k+1} - d_1^{k+1})$
$b_2^{k+1} = b_2^k + (D_2 u_2^{k+1} - d_2^{k+1})$
$b^{k+1} = b^k + (\nabla u_1^{k+1} - d^{k+1})$
Output: u_1^{k+1}, u_2^{k+1}.

and E is given by

$$E = \{v \in X^* | \quad \langle u, v \rangle_{X \times X^*} \leq \Phi(u), \forall u \in X\}$$

Now we are able to prove theorem 7.

Proof of Theorem 7: We denote $H(v) = \|v\|_{H^1_d}$. The functional F is convex, since the norms are convex. The minimizer of F is given by the solution of

$$\frac{v - \gamma}{\lambda} \in \partial H(v) \tag{56}$$

By the inverse rules for subgradients, this is equivalent to

$$v \in \partial H^*(\frac{\gamma - v}{\lambda}) \tag{57}$$

By setting $w = \gamma - v$, it comes:

$$\gamma - w \in \partial H^*(\frac{w}{\lambda}) \tag{58}$$

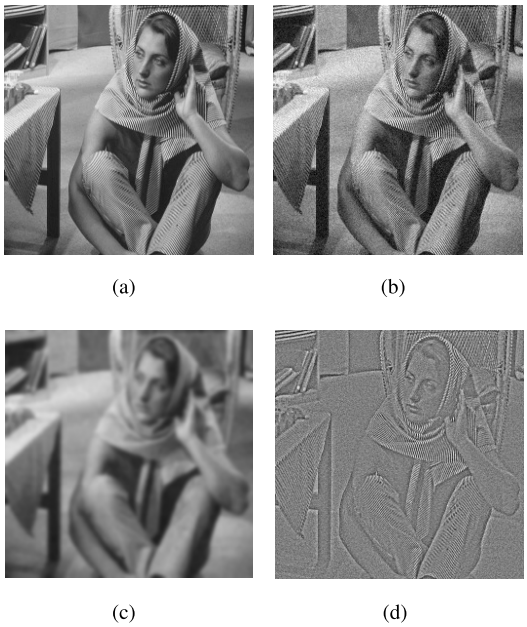
i.e., $0 \in w - \gamma + \partial H^*(\frac{w}{\lambda})$ and this means that w is a minimizer of

$$\frac{1}{2\lambda}\|w - \gamma\|^2_{L^2} + H^*(\frac{w}{\lambda}) \tag{59}$$

By Lemma 10, $H^*(\frac{w}{\lambda})$ is the indicator function of set E. So (59) indicates that $w = \gamma - v$ is the projection of γ onto the set λE. Thus, $\gamma - v = P_{\lambda E}(\gamma)$, i.e., $v = \gamma - P_{\lambda E}(\gamma)$. By Lemma 8, we can get $\lambda E = \{w \in BMO_d : \|w\|_{BMO_d} \leq \lambda\}$. □

From Theorem 7, one can see that solving model (9) is reduced to solving the projection $P_{\lambda E}(\gamma)$, which is equivalent to solving the following problem:

$$(P): \begin{cases} \min\limits_{w} & \|f - w\|^2_{L^2} \\ \text{s.t.}: & \|w\|_{BMO_d} \leq \lambda \end{cases}$$



(a) (b) (c) (d)

FIGURE 2. Image denoising result by *ROF* model: (a) original 256 × 256 Barbara image. (b) Noisy version. (c) Cartoon part. (d) Residual image.

Model (P) was first proposed in our previous work [25] to recover some clear textures from noisy data. The model is transformed into wavelet domain and solved by dual Uzawa method. The advantage of the model is that each Lagrange multiplier of the discretized model corresponds to a certain scale of dyadic region of the image. Thus, the Lagrange multipliers are space adaptive and can control the extent of denoising over dyadic image regions. For the detail of solution of model (P), we refer the readers to [25].

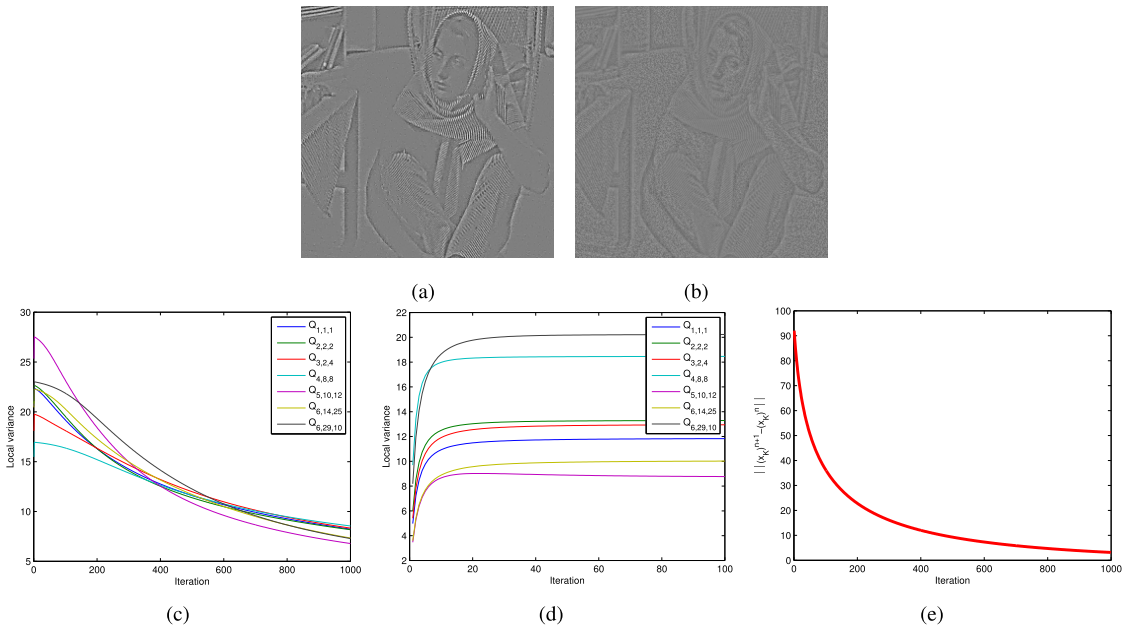


FIGURE 3. Texture extraction result on the residual image Fig. 2(d) by model (9): (a) Texture component. (b) Noise component. (c) Evolution of local variance on some dyadic regions by the proposed model. (d) Evolution of local variance on some dyadic regions by ROF model. (e) Evolution of relative change $\|(x_K^\varepsilon)^{n+1} - (x_K^\varepsilon)^n\|_2$ of algorithm 1 along with iterations.

IV. TV-HARDY REGULARIZATION MODEL

In this section, we propose the following two-layers image restoration model:

$$\min_{u_1,u_2} \{\mu_1\|D_1u_1\|_1 + \mu_2\|D_2u_2\|_{h_d^1} + \theta\|\nabla u_1\|_1 + \frac{1}{2}\|f - K(u_1+u_2)\|_2^2\} \quad (60)$$

where u_1, u_2 represent the piecewise smooth and texture part of the image respectively. D_1 is the dictionary that can represent the piecewise smooth part sparsely. D_2 is the wavelet mentioned in section *II* for characterizing dyadic Hardy space H_d^1. $\|\nabla u_1\|_1 = \|\nabla_x u_1\|_1 + \|\nabla_y u_1\|_1$ is the Total variation (TV) term, which forces the image u_1 to be close to a piecewise smooth image. The parameter μ_1 and θ control the penalty on sparsity and regularity of the cartoon component. When μ_1 is relatively larger, it tends to generate smoother images. This is because the frame coefficient D_1u_1 is quite often linked to the smoothness of the underlying image. When θ is relatively larger, the TV regularization term is more penalized. This enhances and sharpens edges, although it may introduce some artifacts. Combing TV with wavelet can damp ringing artifacts near edges.

We apply split Bregman iterations to solve this model. Bregman iteration method was first used in image processing by Osher et. al. in [26], and has been applied to TV based model, frame based model and Basis Pursuit problem in [27], [28]. Bregman iteration converges very quickly when applied to certain kind of objective functions, especially for l_1 minimization problem.

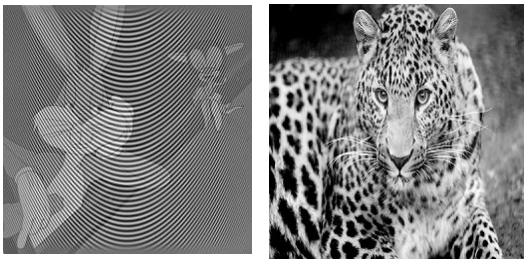


FIGURE 4. Test image for Image decomposition: (a) Original **Rabbit** image with size 256 × 256. (b) Original **Leopard** image with size 256 × 256.

We consider the following equivalent form of (60):

$$\begin{aligned}\min_{u_1,u_2,d_1,d_2,d} &\{\mu_1\|d_1\|_1 + \mu_2\|d_2\|_{h_d^1} + \theta\|d\|_1 \\ &+ \frac{1}{2}\|f - K(u_1+u_2)\|_2^2\} \\ \text{s.t.: } d_1 = D_1u_1, &\quad d_2 = D_2u_2, \\ d = (d_x, d_y) &= (\nabla_x u_1, \nabla_y u_1) \end{aligned} \quad (61)$$

Then we apply the simplified Bregman formulas and get the following split Bregman iteration:

$$\begin{aligned} u_1^{k+1} &= \arg\min_{u_1}\{\frac{1}{2}\|K(u_1+u_2^k) - f\|_2^2 \\ &\quad + \frac{\lambda}{2}\|D_1u_1 - d_1^k + b_1^k\|_2^2 + \frac{\lambda}{2}\|\nabla u_1 - d^k + b^k\|_2^2\} \\ u_2^{k+1} &= \arg\min_{u_2}\{\frac{1}{2}\|Ku_1^{k+1} + Ku_2 - f\|_2^2 \\ &\quad + \frac{\lambda}{2}\|D_2u_2 - d_2^k + b_2^k\|_2^2\} \end{aligned}$$

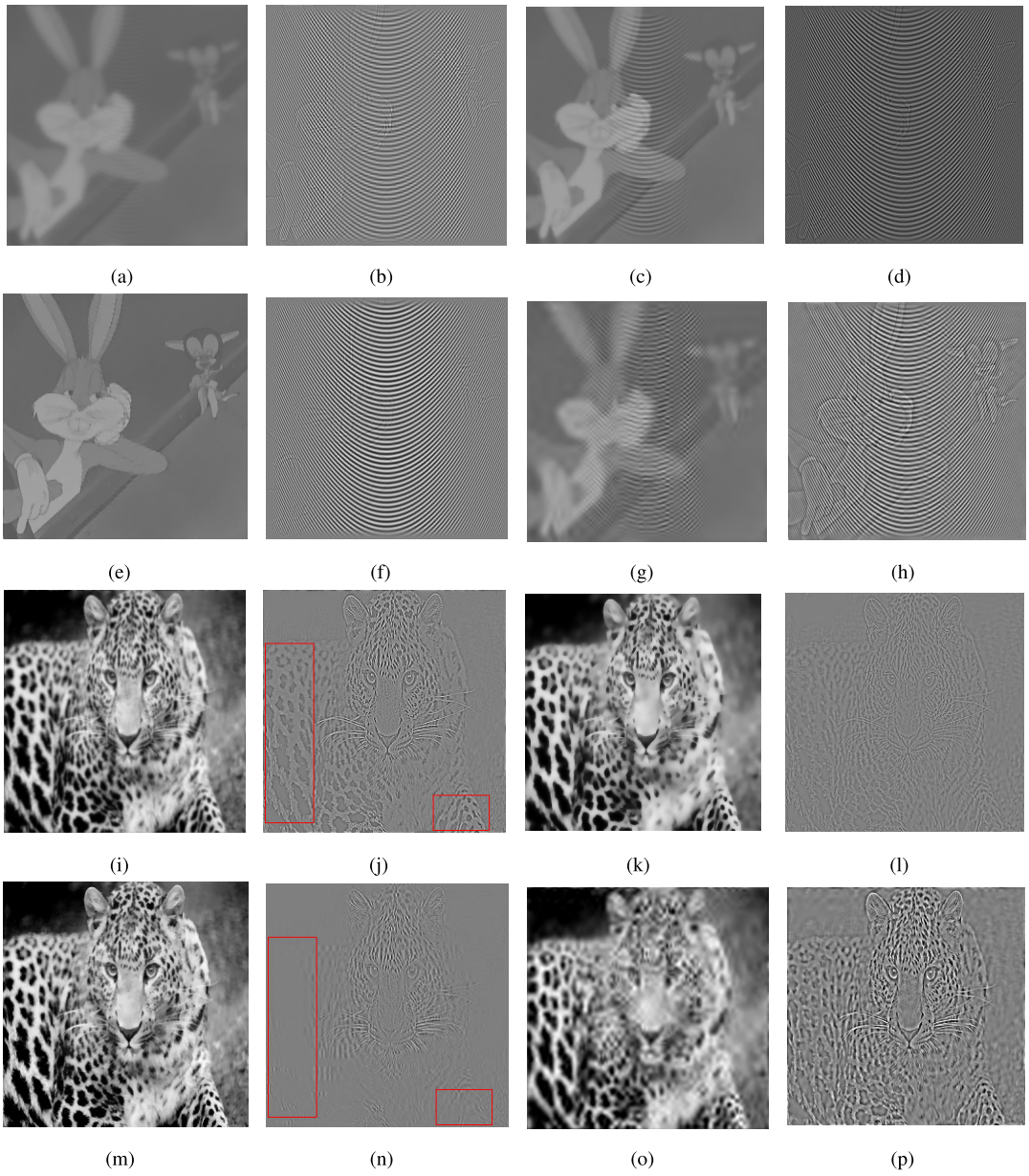


FIGURE 5. Image decomposition results on image **Rabbit** and **Leopard**: (a) Cartoon part by the proposed model (60). (b) Texture part by the proposed model (60). (c) Cartoon part by our A^2BC model. (d) Texture part by A^2BC model. (e) Cartoon part by BNN model. (f) Texture part by BNN model. (g) Cartoon part by Daubechies' model. (h) texture part by Daubechies' model;(i) cartoon part by our model (60). (j) Texture part by our model (60). (k) Cartoon part by A^2BC model. (l) Texture part by A^2BC model. (m) Cartoon part by BNN model. (n) Texture part part by BNN model. (o) Texture part by Daubechies' model. (p) Texture part by Daubechies' model.

$$
\begin{aligned}
d_1^{k+1} &= \arg\min_{d_1}\{\mu_1\|d_1\|_1 + \frac{\lambda}{2}\|D_1\, u_1^{k+1} - d_1 + b_1^k\|_2^2\}\\
d_2^{k+1} &= \arg\min_{d_2}\{\mu_2\|d_2\|_{H_1^d} + \frac{\lambda}{2}\|D_2\, u_2^{k+1} - d_2 + b_2^k\|_2^2\}\\
d^{k+1} &= \arg\min_{d}\{\theta\|d\|_1 + \frac{\lambda}{2}\|\nabla u_1^{k+1} - d + b^k\|_2^2\}\\
b_1^{k+1} &= b_1^k + \delta_b(D_1 u_1^{k+1} - d_1^{k+1})\\
b_2^{k+1} &= b_2^k + \delta_b(D_2 u_2^{k+1} - d_2^{k+1})\\
b^{k+1} &= b^k + \delta_b(\nabla u_1^{k+1} - d^{k+1})
\end{aligned}
$$

The optimality condition of the subproblem for u_1 and u_2 is:

$$
\begin{aligned}
&(\lambda I + K^TK + \lambda\nabla^T\nabla)u_1\\
&\quad = K^Tf - K^TKu_2^k + \lambda D_1^T(d_1^k - b_1^k) + \lambda\nabla^T(d^k - b^k)
\end{aligned}
\tag{62}
$$

$$
\begin{aligned}
&(K^TK + \lambda I)u_2\\
&\quad = K^Tf - K^TKu_1^{k+1} + \lambda D_2^T(d_2^k - b_2^k)
\end{aligned}
\tag{63}
$$

In the deblurring case and under the periodic boundary condition for u_1, $\nabla^T\nabla$ and K^TK are all block circulant matrix

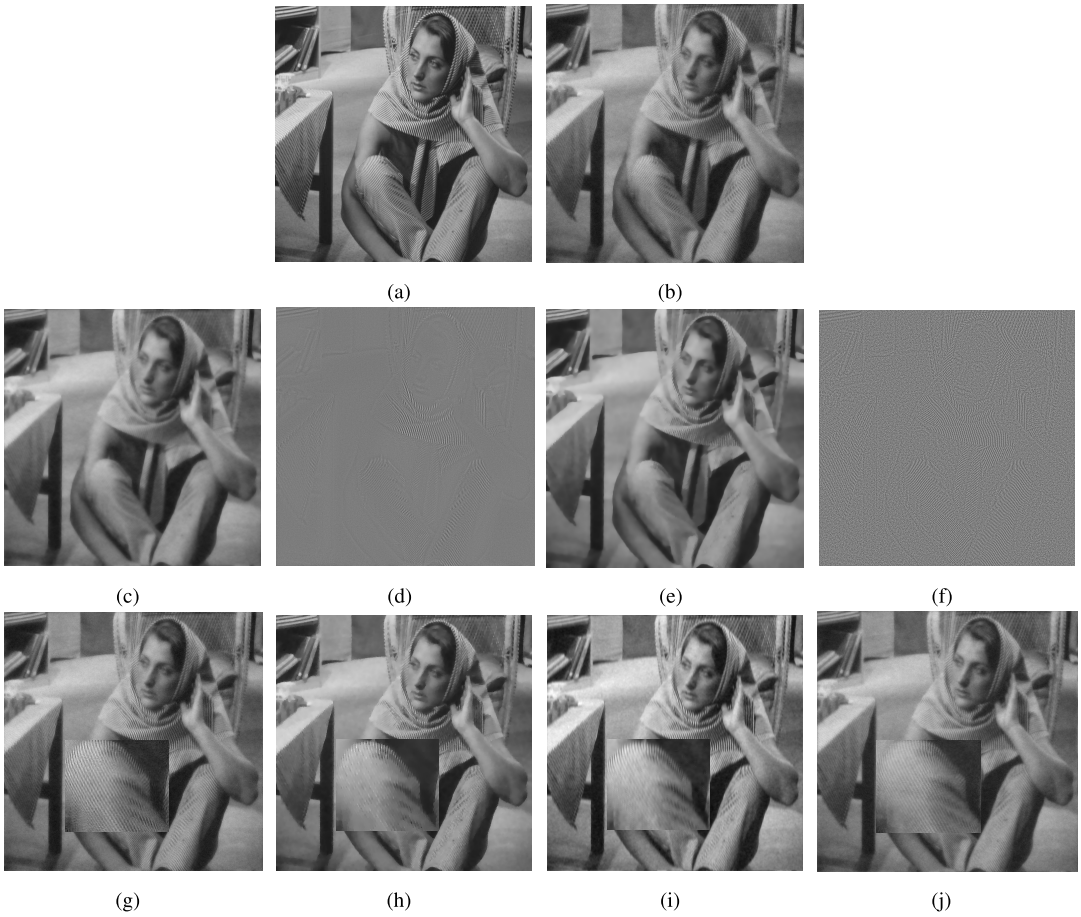


FIGURE 6. Denoising results on image **Barbara**: (a) Original Barbara image; (b) noisy version. (c) cartoon part by model (60). (d) Texture part by model (60). (e) Piecewise smooth part by A^2BC model. (f) Texture part by A^2BC model. (g) Restored image by model (60), PSNR = 27.00. (h) Restored image by TV-L^2 model, PSNR = 25.81. (i) Restored image by frame based model, PSNR = 23.75. (j) Restored image by two-layers TV-l^1 model, PSNR = 25.31.

and thus are diagonalizable by the 2D discrete Fourier transforms. So equation (62) and (63) can be solved efficiently by Fast Fourier Transform (FFT).

The subproblem for d_1, d can be solved explicitly by soft thresholding:

$$d_1^{k+1} = shrink(D_1 u_1^{k+1} + b_1^k, \frac{\mu_1}{\lambda}) \tag{64}$$

$$d_{k+1} = shrink(\nabla u_1^{k+1} + b^k, \frac{\theta}{\lambda}) \tag{65}$$

where $shrink(x, \theta) = \frac{x}{|x|} \cdot \max(|x| - \theta, 0)$.

The subproblem for d_2 can be solved by the method we proposed in section III.

As a summary, we obtain the algorithm in Table 2 for two-layers image restoration.

V. EXPERIMENTS

In this section, we present various experiments to demonstrate the performance of the proposed models for texture extraction, image decomposition, denoising and deblurring. Model (60) is compared with TV based model, BNN model and frame based regularization model. In all experiments, we choose piecewise linear B-spline framelet as the dictionary D_1 for representing piecewise smooth component, and 'Db 10' wavelet as the dictionary D_2 for representing texture component. The test images have plenty of cartoon and texture regions and are well suited for testing the restoration models which are aiming at recovering more texture while smoothing out noise.

All the experiments are implemented under Windows 7 and Matlab v2012b with Intel Pentium P6100 CPU and 2GB memory. The dynamic range of test images are normalized to [0,1].

A. TEXTURE EXTRACTION

In this subsection, we present some numerical results to demonstrate the performance of algorithm in Table 1 for extracting clear textures from noisy data. Fig. 2(a) and (b) show the Barbara image with size of 256 × 256 and its noisy version respectively. The noise added is zero-mean Gaussian noise with variance $\sigma = 10/255$. Fig. 2(c) and (d) show the cartoon component and the residual image of ROF model (also called TV-L^2 model). One can see that, the cartoon part is over-smoothing and there are a large amount of textures in the residual image. We apply model (9) to

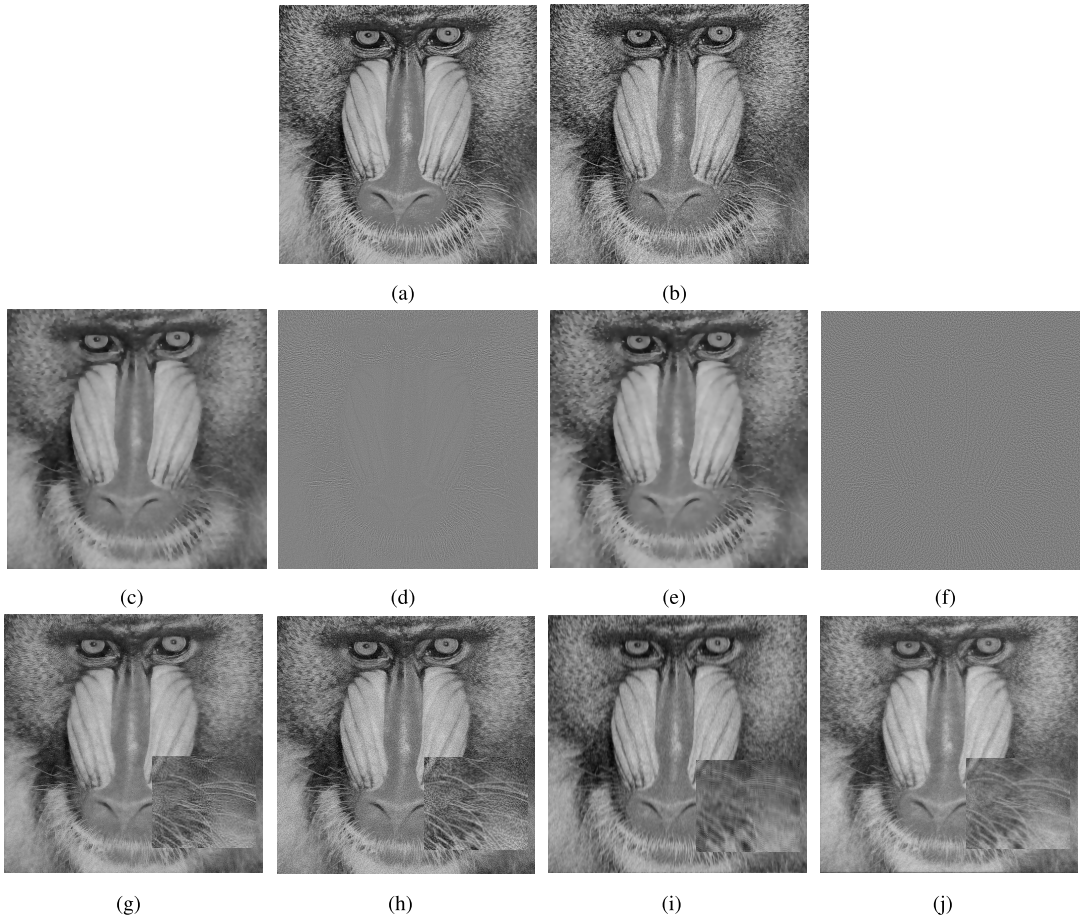


FIGURE 7. Denoising results on image **Mandrill**: (a) Original Mandrill image. (b) Noisy version. (c) Cartoon part by model (60). (d) Texture part by model (60). (e) Cartoon part by A^2BC model. (f) Texture part by A^2BC model. (g) Restored image by model (60), PSNR = 24.30. (h) Restored image by TV-L^2 model, PSNR = 23.11. (i) Restored image by frame based model, PSNR = 22.10. (j) Restored image by two-layers TV-l^1 model, PSNR = 23.72.

the residual image of ROF model and show the "texture + noise" decomposition result in Fig. 3. From Fig. 3(a), We can see that clear textures are well extracted from the residual image of ROF model. From Fig. 3(b), one can see that there are still some weak edges in the noise component. This is inevitable since some small details with low local oscillating property have the similar local statistical features with noise. But above all, the texture component extracted is mixed with very little noise. That is, the proposed model is very efficient to preserve texture with high local oscillating property.

In order to demonstrate the performance of model (9) for controlling local variances of the residual image, we show the change of local variance of residual $\gamma - v$ on different dyadic region Q_{j,k_1,k_2} along with iterations in Fig. 3(c). One can see that the local variance on dyadic regions is gradually reduced to the same level by our algorithm, and the local variance level can be controlled by parameter λ. We also give the evolution of local variance on the same dyadic regions by ROF model for comparison in Fig. 3(d). As comparison in Fig. 3(d), it can not control the local variance of the residual by turning the parameter λ of ROF model. Fig. 3(e) demonstrates the relative change $\|(x_K^\varepsilon)^{n+1} - (x_K^\varepsilon)^n\|_2$ of Algorithm 1 along with iterations, and it shows that the algorithm converges fast.

B. IMAGE DECOMPOSITION

In this subsection, we show some numerical experiments of model (60) for "cartoon+texture" image decomposition. We start with a synthetically generated image composed of a natural scene and texture. Fig. 4(a) shows the original **Rabbit** image (addition of the texture and the natural parts).

The decomposition results including cartoon part and texture part are shown in Fig. 5 (a)-(b). For comparison, in Fig. 5(c)-(h) we show the decomposition results using A^2BC model of [4], BNN model of [6] and Daubechies' model of [12]. For image **Rabbit**, the methods based on oscillating functional modeling do not perform well. The texture in the synthetic image has varying degrees of oscillation. In the middle of the image, the oscillation of texture is relatively lower, while in left and right sides of the image, the oscillation is much higher. The proposed model is based on the prior that the local oscillation of the edge and texture component is much higher than other components. As shown in Fig. 5(a), (c) and (g), the high oscillating

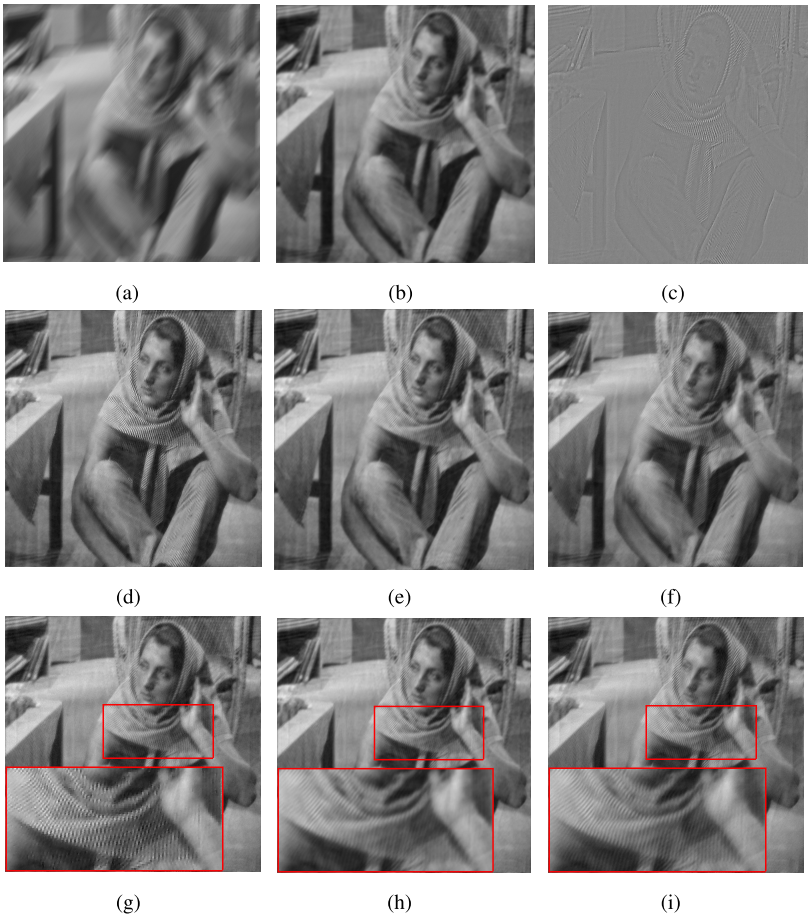


FIGURE 8. Deblurring results on image **Barbara**:(a) Blurred image, PSNR = 20.95. (b) Recovered cartoon part by model (60). (c) Recovered texture part by model (60). (d) Restored image by model (60), PSNR = 24.93. (e) Restored image by TV based model, PSNR = 24.11. (f) Restored image by frame based model, PSNR = 23.71. (g)–(i): Local enlargement of (d)–(f).

texture in the both sides of the image can be separated well from cartoon component, but the low oscillating texture in the middle of the image is failed to be separated from cartoon component. From Fig. 5 (e) and (f), one can see that the BNN method gives the best cartoon-texture decomposition result for image **Rabbit**. This is because the BNN method has a special capability of capturing globally dissimilar but locally well-patterned texture as appears in the test image **Rabbit**. However, each approach has its advantage and disadvantage. We test our method on another image **Leopard** which contains natural high oscillating textures. The original **Leopard** image is shown in Fig. 4(b), and the decomposition result by our method as well as the comparison with other methods are shown in Fig. 5(i)-(p). From Fig. 5 (j), especially from the red rectangle region, the proposed method captures more clear textures than the other methods.

C. IMAGE DENOISING

For image denoising, we add Gaussian white noise with zero mean and standard variation $\sigma = 15/255$ on the image Barbara (size 512×512) in Fig. 6(a), and get its noisy version in Fig. 6(b). The decomposition result by the proposed algorithm is compared with that by A^2BC model in Fig. 6 (c)-(f). Fig. 6(c) and (d) display the cartoon part and the texture part respectively obtained by the proposed model (60). Fig. 6(e) and (f) show the decomposition results by the A^2BC model. From comparison, we can see that the texture component is more clear and mixed with less noise. In Fig. 6(g)-(i), we compare the restored image by different methods including TV -L^2 method and frame based method (balanced method). To indicate the advantage of the h^1_d norm over the classical l^1norm, in Fig. 6(j) we show a comparison with the slightly modified version of the proposed method that the h^1_d norm in model (60) is replaced with the classical l_1 norm. We will refer to it as the two-layers $TV - l^1$ model in the following. From the comparison, especially from the local enlarged region, we can see that both TV based models and frame based model remove most of the textures and small details, while the proposed model preserves more texture and gets better subjective visual effect.

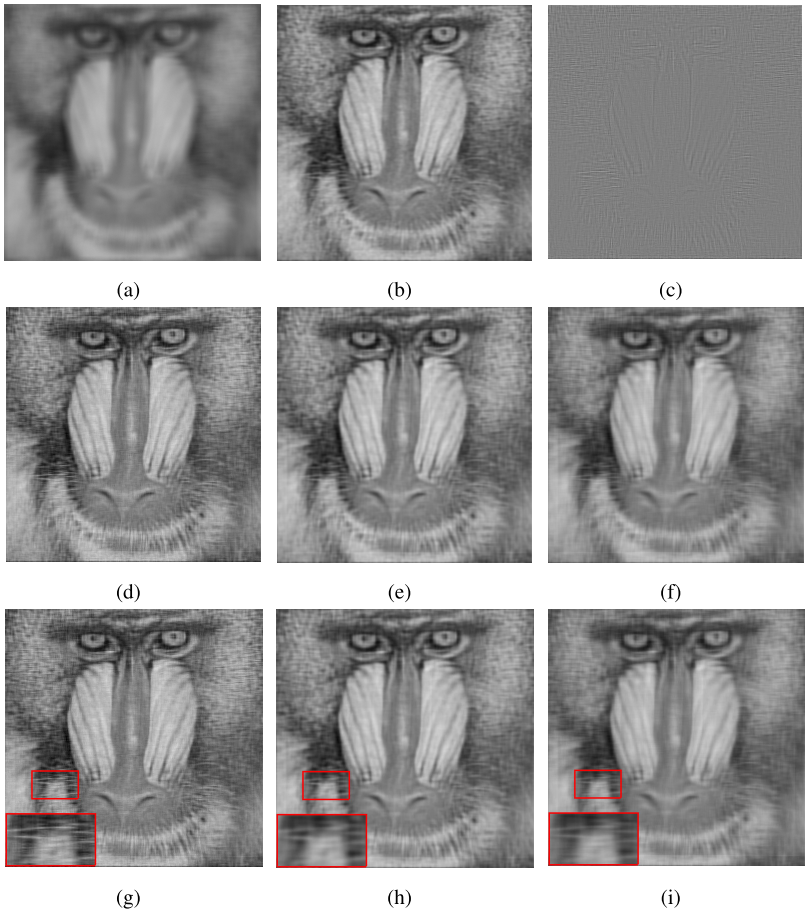


FIGURE 9. Deblurring results on image **Mandrill** by our method: (a) Blurred image, PSNR = 19.50. (b) Recovered cartoon part. (c) Recovered texture part: (d) restored image by model (60), PSNR = 21.64. (e) Restored image by TV method, PSNR = 21.00. (f) Restored image by frame based method, PSNR = 21.67. (g)-(i): Local enlargement of (d)–(f).

Fig. 7 shows the similar results on another texture image **Mandrill**, with size of 512 × 512, degraded by zero-mean Gaussian noise with standard variation $\sigma = 20/255$. Visually, textures are better recovered while denoising by the proposed two-layers denoising method. We also report the the peak signal-to-noise (PSNR) ratio result by different methods. In a whole, the proposed algorithm has higher PSNR than the other two methods.

We list all the parameters involved in the algorithm in the following:

- For Barbara image: $\mu_1 = 0.01;\ \mu_2 = 0.1;\ \lambda = 1;\ \theta = 0.1;\ \sigma = 15/255;\ \delta_b = 0.5$.
- For Mandrill image: $\mu_1 = 0.001;\ \mu_2 = 0.1;\ \lambda = 0.3;\ \theta = 0.1;\ \sigma = 20/255;\ \delta_b = 0.5$.

The peak signal-to-noise (PSNR) ratio is used to measure the quality of the recovered images which is defined as:

$$PSNR(u, f) = 10\log_{10}\left(\frac{255^2}{\frac{1}{mn}\|u - f\|_2^2}\right) \tag{66}$$

where u is the restored image, f is the true image, m and n denote the size of the image.

D. IMAGE DEBLURRING

In this subsection, we test the proposed model (60) on image deblurring problems. We test both Gaussian blur and motion blur with Gaussian white noise. We assume periodic boundary conditions since the FFT can be used to solve u^{k+1} in Algorithm 2. As comparisons, we also show the results obtained by frame based and TV based deblurring methods.

The main parameters of Algorithm 1 are list as following:

- For image Barbara: $\mu_1 = 0.001;\ \mu_2 = 0.0004;\ \lambda = 0.01;\ \theta = 0.1;\ \sigma = 0.8, \delta_b = 0.5$.
- For image Mandrill: $\mu_1 = 0.0001;\ \mu_2 = 0.00008;\ \lambda = 0.001;\ \theta = 0.5;\ \sigma = 1, \delta_b = 0.5$.

The test image in Fig. 8(a) is generated by applying a motion blur on the true image Fig. 6(a) and then adding Gaussian noise with zero mean and standard deviation $\sigma = 0.8$. The motion blur kernel is obtained by the Matlab function 'fspecial('motion',35,50)'. Fig. 8(b)-(c) show the deblurring results of the proposed two-layers image restoration model (60). We can see that both the cartoon part and texture part are well recovered. Fig. 8(d) shows the restored image obtained by the proposed model (60). As comparison,

we also include the results obtained by TV based method and frame based method (analysis based method) in Fig. 8(e)-(f). Fig. 8(g)-(i) show the local detail enlargement of Fig. 8(d)-(f). One can see that most of textures are eliminated by TV based method and frame based method, while edges and textures are better preserved by the proposed method.

The test image in Fig. 9(a) is generated by applying a Gaussian blur on the true image Fig. 7(a), and then adding Gaussian noise with zero mean and standard variation $\sigma = 1$. The Gaussian kernel is generated by the MATLAB function 'fspecial('gaussian',[20 20], 20)'. Fig. 9(b) and (c) show the restored cartoon part and texture part by the proposed algorithm. In Fig. 9((d)-(f)), we compare the proposed algorithm with TV based deblurring and frame based deburring. Fig. 9(g)-(i) show the local detail enlargement of Fig. 9(d)-(f). Again, this shows that the proposed algorithm performs better for deblurring images while preserving edges and textures.

VI. CONCLUSION

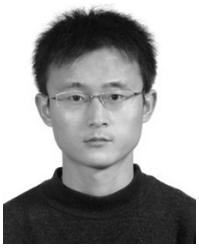
In this paper, we have introduced dyadic Hardy space and dyadic BMO space for image restoration modeling and analysis, and try to explore the regularization mechanism of these two spaces in texture preserving image restoration inverse problem. We study the wavelet characterization of the dyadic Hardy H_1^d norm, and use it as the regularizer of modeling textures. Then we propose a H_1^d minimization model which can efficiently extract clear high oscillating textures from noisy data. Combing this H_1^d regularization, TV regularization, and sparse regularization, we propose a two-layers image restoration model, in which we use two dictionaries D_1 and D_2 to represent piecewise smooth part and texture part respectively. We take D_1 as the piecewise linear framelet to represent the cartoon component and l_1 norm as the regularizer which leads to sparsity in the process of minimization. We take D_2 as wavelets to represent texture component and H_1^d norm as regularizer which leads to group sparsity in the process of minimization. Of course, there are many possible ways to choose D_1 and D_2. For example, D_1 may be curvelet, Tensor Product Complex Tight Frame (TP-CFT), and D_2 may be multi-scale Local Discrete Cosine Transform (LDCT) or Gabor Transform. These chosen of dictionaries may have better performance than that proposed in this paper. The proposed method can be seen as a specific case of a more general approach and this will be investigated in our future work.

ACKNOWLEDGMENT

The authors are very grateful to the Editor and the anonymous referees for their helpful comments and suggestions, which have led to an improvement in the quality of the paper.

REFERENCES

[1] Y. Meyer, *Oscillating Patterns in Image Processing and Nonlinear Evolution Equations: The Fifteenth Dean Jacqueline B. Lewis Memorial Lectures*. Providence, RI, USA: AMS, 2001.
[2] L. A. Vese and S. J. Osher, "Modeling textures with total variation minimization and oscillating patterns in image processing," *J. Sci. Comput.*, vol. 19, nos. 1–3, pp. 553–572, 2003.
[3] S. Osher, A. Solé, and L. Vese, "Image decomposition and restoration using total variation minimization and the H^1," *SIAM J. Multiscale Model. Simul.*, vol. 1, no. 3, pp. 349–370, 2003.
[4] J. F. Aujol, G. Auber, L. Blanc-Féraud, and A. Chambolle, "Image decomposition into a bounded variation component and an oscillating component," *J. Math. Imag. Vis.*, vol. 22, no. 1, pp. 71–88, 2005.
[5] J.-F. Aujol and A. Chambolle, "Dual norms and image decomposition models," *Int. J. Comput. Vis.*, vol. 63, no. 1, pp. 85–104, Jun. 2005.
[6] S. Ono, T. Miyata, and I. Yamada, "Cartoon-texture image decomposition using blockwise low-rank texture characterization," *IEEE Trans. Image Process.*, vol. 23, no. 3, pp. 1128–1142, Mar. 2014.
[7] J.-L. Starck, M. Elad, and D. L. Donoho, "Image decomposition via the combination of sparse representations and a variational approach," *IEEE Trans. Image Process.*, vol. 14, no. 10, pp. 1570–1582, Oct. 2005.
[8] A. N. Tikhonov, "Regularization of incorrectly posed problems," *Sov. Math. Doklady*, vol. 4, no. 6, pp. 1624–1627, 1963.
[9] D. Mumford and J. Shah, "Optimal approximations by piecewise smooth functions and associated variational problems," *Commun. Pure Appl. Math.*, vol. 42, no. 5, pp. 577–685, 1989.
[10] L. I. Rudin, S. Osher, and E. Fatemi, "Nonlinear total variation based noise removal algorithms," *Phys. D, Nonlinear Phenomena*, vol. 60, nos. 1–4, pp. 259–268, 1992.
[11] D. Mumford and B. Gidas, "Stochastic models for generic images," *Quart. Appl. Math.*, vol. 59, no. 1, pp. 85–111, 2001.
[12] I. Daubechies and G. Teschke, "Variational image restoration by means of wavelets: Simultaneous decomposition, deblurring, and denoising," *Appl. Comput. Harmon. Anal.*, vol. 19, no. 1, pp. 1–16, 2005.
[13] W. Dong, G. Shi, Y. Ma, and X. Li, "Image restoration via simultaneous sparse coding: Where structured sparsity meets Gaussian scale mixture," *Int. J. Comput. Vis.*, vol. 114, nos. 2–3, pp. 217–232, 2015.
[14] L. Zhang and W. Zuo, "Image restoration: From sparse and low-rank priors to deep priors [lecture notes]," *IEEE Signal Process. Mag.*, vol. 34, no. 5, pp. 172–179, Sep. 2017.
[15] J. Zhang, D. Zhao, and W. Gao, "Group-based sparse representation for image restoration," *IEEE Trans. Image Process.*, vol. 23, no. 8, pp. 3336–3351, Aug. 2014.
[16] M. Kowalski, "Sparse regression using mixed norms," *Appl. Comput. Harmon. Anal.*, vol. 27, no. 3, pp. 303–324, 2009.
[17] G. Gilboa, N. Sochen, and Y. Y. Zeevi, "Texture preserving variational denoising using an adaptive fidelity term," *IEEE Trans. Image Process.*, vol. 15, no. 8, pp. 2281–2289, May 2006.
[18] G. Gilles, "Noisy image decomposition: A new structure, texture and noise model based on local adaptivity," *J. Math. Imag. Vis.*, vol. 28, no. 3, pp. 285–295, 2007.
[19] N. Azzabou, N. Paragios, and F. Guichard, "Uniform and textured regions separation in natural images towards MPM adaptive denoising," in *Proc. Int. Conf. Scale Space Variational Methods Comput. Vis.*, 2007, pp. 418–429.
[20] T. Zhang, Q. Fan, and Q. Gao, "Wavelet characterization of hardy space H^1 and its application in variational image decomposition," *Int. J. Wavelets, Multiresolution Inf. Process.*, vol. 8, no. 1, pp. 71–87, 2010.
[21] Y. Meyer, *Wavelets and Operators*. Cambridge, U.K.: Cambridge Univ. Press, 1992.
[22] M. Frazier and B. Jawerth, "A discrete transform and decompositions of distribution spaces," *J. Funct. Anal.*, vol. 93, no. 1, pp. 34–170, 1990.
[23] T. Zhang and Q. B. Fan, "Wavelet characterization of dyadic BMO norm and its application in image decomposition for distinguishing between texture and noise," *Int. J. Wavelets, Multiresolution Inf. Process.*, vol. 9, no. 3, pp. 445–457, 2011.
[24] H. H. Bauschke and P. L. Combettes, *Convex Analysis and Monotone Operator Theory in Hilbert Spaces*. New York, NY, USA: Springer, 2011.
[25] T. Zhang, Q. Gao, and G. Tan, "Texture preserving image restoration with dyadic bounded mean oscillating constraints," *IEEE Signal Process. Lett.*, vol. 22, no. 3, pp. 322–326, Mar. 2015.
[26] S. Osher, M. Burger, D. Goldfarb, J. Xu, and W. Yin, "An iterative regularization method for total variation-based image restoration," *Multiscale Model. Simul.*, vol. 4, no. 2, pp. 460–489, 2005.
[27] W. Yin, S. Osher, D. Goldfarb, and J. Darbon, "Bregman iterative algorithms for ℓ_1-minimization with applications to compressed sensing," *SIAM J. Imag. Sci.*, vol. 1, pp. 142–168, Apr. 2008.
[28] J.-F. Cai, S. Osher, and Z. Shen, "Split Bregman methods and frame based image restoration," *Multiscale Model. Simul.*, vol. 8, no. 2, pp. 337–369, 2010.



TAO ZHANG received the M.S. and Ph.D. degrees from Wuhan University, China, in 2007 and 2010. Since 2010, he has been with the Department of Mathematics and Physics, Anhui University of Technology. His research interests include wavelet analysis and inverse problems in image processing.

XUTAO MO received the B.S. and Ph.D. degrees from Tianjin University, China, in 2003 and 2008, respectively. Since 2009, he has been with the Department of Mathematics and Physics, Anhui University of Technology. His research interests include the wavefront coding extended depth of field imaging and image measurement.