Model-Based Nonuniform Compressive Sampling and Recovery of Natural Images Utilizing a Wavelet-Domain Universal Hidden Markov Model

Behzad Shahrasbi, *Member, IEEE*, and Nazanin Rahnavard, *Member, IEEE*

***Abstract*—In this paper, a novel model-based compressive sampling (CS) technique for natural images is proposed. Our algorithm integrates a *universal hidden Markov tree (uHMT)* model, which captures the relation among the sparse wavelet coefficients of images, into both *sampling* and *recovery* steps of CS. At the *sampling* step, we employ the uHMT model to devise a nonuniformly sparse measurement matrix Φ_{uHMT}. In contrast to the conventional CS sampling matrices, such as dense Gaussian, Bernoulli or uniformly sparse matrices that are oblivious to the signal model and the correlation among the signal coefficients, the proposed Φ_{uHMT} is designed based on the signal model and samples the coarser wavelet coefficients with higher probabilities and more sparse wavelet coefficients with lower probabilities. At the *recovery* step, we integrate the uHMT model into two state-of-the-art Bayesian CS recovery schemes. Our simulation results confirm the superiority of our proposed *HMT model-based nonuniform compressive sampling and recovery*, referred to as *uHMT-NCS*, over other model-based CS techniques that solely consider the signal model at the recovery step. This paper is distinguished from other model-based CS schemes in that we take a novel approach to simultaneously integrating the signal model into both CS *sampling* and *recovery* steps. We show that such integration greatly increases the performance of the CS recovery, which is equivalent to reducing the required number of samples for a given reconstruction quality.**

***Index Terms*—Compressed sensing, wavelet coefficients, image sampling.**

I. Introduction

TODAY'S multimedia-rich applications have dramatically increased the traffic flow in the communication networks. To deal with this overwhelmingly large amount of data and reduce the computational complexity, new compression techniques are in demand. In this regard, the emerging field of compressive sampling (CS) [1], [2] that has revolutionized the traditional concept of sensing and sampling established by the Nyquist sampling theorem has attracted a lot of attention. According to the CS theory, the signals that have a sparse representation over a proper basis can be recovered from a small set of linear measurements.

Manuscript received August 15, 2014; revised March 24, 2015, January 10, 2016, and September 6, 2016; accepted September 7, 2016. Date of publication September 29, 2016; date of current version October 28, 2016. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Sergios Theodoridis. This work was supported by the National Science Foundation under Grants ECCS-1418710 and CCF-1439182.

The authors are with the School of Electrical Engineering and Computer Science, University of Central Florida, Orlando, FL 32816 USA (e-mail: behzad.shahrasbi@knights.ucf.edu; nazanin@eecs.ucf.edu).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TSP.2016.2614654

The compression efficiency of CS cannot compete with conventional codec such as JPEG2000 or MPEG4 when dealing with already acquired image or video signals with high resolution and quality [3]. However, CS is still desirable in applications in which sensing is expensive (such as MRI or infra-red imaging). The conventional CS algorithms merely exploit signal *sparsity* in their designs. Nevertheless, it has recently been shown that in addition to the sparsity, we can utilize the extra *knowledge* about the signal structure as a priori information in the CS *recovery* step to enhance the overall CS recovery performance compared to the conventional CS recovery algorithms [4]–[7]. In a very recent work [8], Indyk and Razenshteyn proved that for the signals with tree-structured sparsity (such as natural images), the recovery is achievable with fewer measurements compared to the recovery of general sparse signals. In [4], the tree structure of wavelet coefficients is used to design a CS recovery algorithm based on weighted ℓ_1 minimization, called HMT-based IRWL1, for one dimensional piecewise smooth signals. In [6], the tree structure of wavelet coefficients is exploited to create a statistical model for the sparse coefficients that results in more accurate recovery. In [7], the authors have modified the novel approximate message passing algorithm [9] such that the tree structure of the wavelet coefficients is utilized in the recovery process.

In this paper, we take one step further and show that by exploiting the signal model in the *sampling* step of the CS in addition to the recovery step, we can achieve greater gains. Our paper mainly focuses on natural images. It is well known that the signal coefficients of natural images in the wavelet domain are not only sparse but also are correlated by a tree structure as described in [10], [11]. Romberg *et al.* developed a *universal hidden Markov model (uHMT)* for normalized natural images in [11]. The uHMT model provides relatively accurate predictions about the wavelet coefficients. We will demonstrate how uHMT model can be integrated into the design of novel sampling and recovery schemes.

There have been some recent studies on unconventional measurement matrices, such as structured measurement matrices [12] or Toeplitz matrices [13]) to mimic the real-world applications in which the acquisition of samples by a random Gaussian or Bernoulli sampling matrices are infeasible [12] (e.g., multipath channel estimation [13]). However, only a few contributions deliberately modify the sampling matrix structure with the goal of improving the overall CS performance [14]–[17]. In [14], the authors have shown that separately measuring each scale of signals' wavelet coefficients improves the

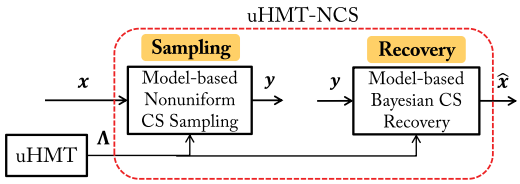


Fig. 1. The block diagram of the proposed uHMT-NCS scheme. The vectors $\boldsymbol{x}$, $\boldsymbol{y}$, and $\hat{\boldsymbol{x}}$ correspond to the image signal, the CS measurements, and the recovered image, respectively. The uHMT model parameters ($\boldsymbol{\Lambda}$) are utilized at both CS Sampling and CS recovery steps.

recovery performance. In [15], the authors have exploited the visual importance of different areas of an image in the sampling step and have modified the block-CS algorithm (proposed in [18]) to generate CS measurements with different compression levels. In [16], the authors have proposed a structurally random structure for the measurement matrix which provides fast computation and low complexity compared to random measurement matrices. In [19], the authors proposed a model-based dictionary learning considering the tree structure of the wavelet coefficients. In [17], the measurement matrix is designed such that it is matched to the dictionary that is learned through training images. Such a match reduces the coherence between the measurement matrix and the dictionary and enhances the CS recovery performance.

To the best of our knowledge [14], [15] are the only studies that modify the sampling step based on the sparse signal model (in the wavelet domain) to improve the performance. In our work, we take a novel and more effective approach to generate the *CS measurement matrix* to be applied to natural images utilizing the properties of multi-scale wavelet transform as side information.

A. Contribution of this Paper

In this paper, we exploit the uHMT model of wavelet coefficients of natural images to modify *both CS sampling and recovery steps* and significantly enhance the performance. The contribution of this paper is two-fold. First, using the uHMT model, we propose a *nonuniform CS measurement matrix* that generates CS measurements such that they include the significant coefficients *with a higher probability* compared to the non-significant coefficients. Second, we propose two model-based recovery algorithms that employ the uHMT model to improve the CS recovery performance even further. The novelty of this work is in proposing a uHMT-model-based *nonuniform* CS sampling, and in *simultaneously* employing the tree structure of wavelet coefficients at both CS sampling and CS recovery steps. Fig. 1 depicts the block diagram of the sampling and the recovery steps in our proposed scheme, referred to as *uHMT-NCS*. Although this work mainly focuses on natural images, the ideas from this work can be extended for the design of new CS matrices when dealing with different signal models.

The structure of this paper is as follows: Section II provides a brief background on CS and the uHMT model of images' 2D wavelet coefficients. In Section III, the details on forming a novel CS measurement matrix $\boldsymbol{\Phi}_{uHMT}$ based on the uHMT model is discussed. In Section IV, we modify two message-passing-based CS recovery algorithms to adapt to our proposed model-based CS measurements. In Section V, we compare the performance of our proposed algorithm with state-of-the-art and other existing model-based CS schemes. Finally, Section VI concludes the paper.

II. Background

In this section, we provide a brief introduction to the compressive sensing and uHMT model for wavelet coefficients of natural images that we will use later to develop our proposed schemes.

A. Overview of Compressive Sensing

Let us define a discrete-time signal of length n as $\boldsymbol{x} = [x_1, x_2, \ldots, x_n]$. The signal $\boldsymbol{x}$ is said to be k-*sparse* in some orthonormal basis $\boldsymbol{\Psi} = [\boldsymbol{\psi}_1, \boldsymbol{\psi}_2, \ldots, \boldsymbol{\psi}_n]$ if $\boldsymbol{\theta} = \boldsymbol{\Psi}^T\boldsymbol{x}$ has at most $k \ll n$ non-zero coefficients. The *sparsity rate* of a k-sparse signal is defined as $s = k/n$. In this paper, $\boldsymbol{x}$ corresponds to the vectorized pixel values of an image, $\boldsymbol{\Psi}$ is the wavelet basis, and $\boldsymbol{\theta} = \boldsymbol{\Psi}^T\boldsymbol{x}$ corresponds to the sparse wavelet coefficients of the image.[1] The CS paradigm suggests that instead of sampling all the n coefficients of $\boldsymbol{x}$, we can recover $\boldsymbol{x}$ from only $m = \mathcal{O}(k\log(n/k)) \ll n$ random measurements [1]. The random measurements are generated by $\boldsymbol{y} = \boldsymbol{\Xi}\boldsymbol{x} = \boldsymbol{\Phi}\boldsymbol{\Psi}^T\boldsymbol{x}$, where $\boldsymbol{\Phi} = \boldsymbol{\Xi}\boldsymbol{\Psi} = [\varphi_{i,j}]$. $\boldsymbol{\Phi}$ is an $m \times n$ matrix with $m \ll n$ and is called the measurement matrix. *Signal recovery* solves $\hat{\boldsymbol{x}} = \arg\min \|\boldsymbol{\Psi}^T\boldsymbol{x}\|_1$ such that $\boldsymbol{y} = \boldsymbol{\Xi}\boldsymbol{x}$.

Numerous sparse recovery algorithms are proposed in the literature. Among them, Bayesian-based methods [20], [21] gained a considerable interest in recent years due to development of message passing algorithms that can efficiently approximate the Bayesian solution [9], [20], [22]–[24]. Bayesian compressive sensing via belief propagation (CSBP) [20] and approximate message passing (AMP) [9] are two of the most efficient CS recovery algorithms that employ message passing. We modify these two algorithms and use them to solve our recovery problem. CSBP requires a *sparse* measurement matrix to perform efficiently, and AMP can work with the sparse measurement matrices without imposing any penalty on the number of measurements [7]. This is an important property because as we will see in Section III, in our proposed algorithm the measurement matrix itself needs to be *sparse*.

B. The Universal Hidden Markov Tree (uHMT) Model

The multi-resolution wavelet decomposition of images has many applications in image processing (e.g., JPEG2000 standard). The wavelet-domain coefficients of an image exhibit both sparsity and a *tree structure* [10], [11]. Fig. 2 shows a

[1]Given $\boldsymbol{x}$ and $\boldsymbol{\theta}$ are vectorized versions of image $\boldsymbol{X}$ and its 2D wavelet coefficients $\boldsymbol{\Theta}$ (i.e., $\boldsymbol{x} = \text{vec}(\boldsymbol{X})$ and $\boldsymbol{\theta} = \text{vec}(\boldsymbol{\Theta})$), respectively, we have $\boldsymbol{\theta} = \text{vec}(\boldsymbol{\Theta}) = \text{vec}(\boldsymbol{\Psi}_2\boldsymbol{X}\boldsymbol{\Psi}_2^T) = \boldsymbol{\Psi}_2 \otimes \boldsymbol{\Psi}_2\text{vec}(\boldsymbol{X}) = \boldsymbol{\Psi}^T\boldsymbol{x}$, where $\boldsymbol{\Theta} = \boldsymbol{\Psi}_2\boldsymbol{X}\boldsymbol{\Psi}_2^T$ represents the 2D wavelet decomposition of image $\boldsymbol{X}$, $\boldsymbol{\Psi}^T = \boldsymbol{\Psi}_2 \otimes \boldsymbol{\Psi}_2$, and $\otimes$ represents Kronecker product.

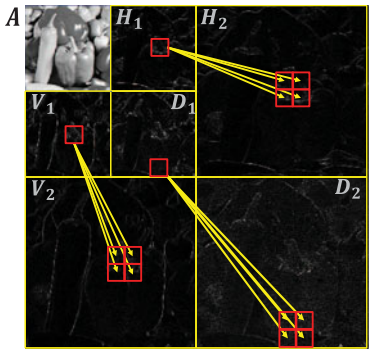


Fig. 2. The two-dimensional wavelet transform representing an image in terms of approximate coefficients (A), and wavelet coefficients in horizontal (H), vertical (V), and diagonal (D) directions. The wavelet coefficients form quad trees with each parent coefficient having four children in the finer scale.

two-dimensional wavelet transform of an image with two scales. The top-left block represents the approximate coefficients (A). The next two top blocks (H_1 and H_2), the two blocks along the left side of the image (V_1 and V_2), and the two diagonal blocks (D_1 and D_2) represent the tree structures along the horizontal, the vertical, and the diagonal directions, respectively. As it can be seen in Fig. 2, quad trees form in horizontal, vertical, and diagonal directions.

Let $\boldsymbol{\theta}|_J = [\theta^b_{j,k}]$ represent the coefficients of the two-dimensional discrete wavelet transform up to scale J of an image represented by $\boldsymbol{x}$. For a coefficient $\theta^b_{j,k}$, $b \in \{A, H, V, D\}$, where A, H, V, D stand for the approximate coefficients, horizontal, vertical, and diagonal subband coefficients, respectively, $j = 0, \ldots, J$ represents the scale of the coefficient with $j = 0$ indicating that the coefficient is an approximate coefficient, and $k = 1, \ldots, n4^{-J-1+\max(1,j)}$ represents the index of the coefficient at direction b and scale j. Generally, the first $\alpha_0 n$ wavelet coefficients in $\boldsymbol{\theta}|_J$ correspond to the approximate coefficients. The remaining $\alpha_1 n, \alpha_2 n, \ldots, \alpha_J n$ coefficients, where $\sum_{j=0}^{J} \alpha_j = 1$, correspond to the coefficients in wavelet scales $1, 2, \ldots, J$, respectively. For example, given the image in Fig. 2, $\alpha_0 n$ is equal to the number of coefficients in A, $\alpha_1 n$ is equal to the number of coefficients in H_1, V_1, and D_1 combined, and finally, $\alpha_2 n$ is equal to the number of coefficients in H_2, V_2, and D_2 combined. For a two-dimensional wavelet transform with J scales, we have $\alpha_0 = 4^{-J}$ and $\alpha_j = 3\alpha_0 4^{j-1}$ for $j = 1, \ldots, J$. The coefficients at scale 1 are called the *coarsest coefficients* because they only represent a rough estimate of the image.

In [10], [11] a complete analysis of the properties of wavelet coefficients is provided, and for them a hidden Markov tree (HMT) model has been developed. According to this model, every wavelet coefficient $\theta^b_{j,k}$ corresponds to a hidden state variable, which can be in either state small ($\mathcal{S}$) or large ($\mathcal{L}$). The HMT model suggests that a tree structure exists among state variables and each coefficient $\theta^b_{j,k}$ for $j = 1, \ldots, J$ has a two-state mixture Gaussian probability density function (pdf) given by

$$f(\theta^b_{j,k}) = \pi^b_{j,k}\mathcal{N}\left(0, {\sigma^b_{\mathcal{S},\{j,k\}}}^2\right) + (1-\pi^b_{j,k})\mathcal{N}\left(0, {\sigma^b_{\mathcal{L},\{j,k\}}}^2\right), \quad (1)$$

where $\mathcal{N}(0, \sigma^2)$ represents a zero-mean Gaussian distribution with variance σ^2, and $\pi^b_{j,k}$ is the probability that the state of $\theta^b_{j,k}$ is small (in the statistical sense) and we have ${\sigma^b_{\mathcal{S},\{j,k\}}}^2 \ll {\sigma^b_{\mathcal{L},\{j,k\}}}^2$.

The state dependency between $\theta^b_{j,k}$ and its parent $\theta^b_{j-1,\lceil k/4 \rceil}$ is modeled by a state transition probability matrix $\boldsymbol{A}^b_{j,k}$ given by

$$\boldsymbol{A}^b_{j,k} = \begin{bmatrix} p^{\mathcal{S}\to\mathcal{S}}_{j,k,b} & p^{\mathcal{S}\to\mathcal{L}}_{j,k,b} \\ p^{\mathcal{L}\to\mathcal{S}}_{j,k,b} & p^{\mathcal{L}\to\mathcal{L}}_{j,k,b} \end{bmatrix}, \quad (2)$$

with $p^{\mathcal{S}\to\mathcal{L}}_{j,k,b} = 1 - p^{\mathcal{S}\to\mathcal{S}}_{j,k,b}$ and $p^{\mathcal{L}\to\mathcal{S}}_{j,k,b} = 1 - p^{\mathcal{L}\to\mathcal{L}}_{j,k,b}$, where $p^{\mathcal{S}\to\mathcal{S}}_{j,k,b}$ is the probability that $\theta^b_{j,k}$ is in state $\mathcal{S}$ given its parent is in state $\mathcal{S}$ and $p^{\mathcal{L}\to\mathcal{L}}_{j,k,b}$ is the probability that $\theta^b_{j,k}$ is in state $\mathcal{L}$ given its parent is in state $\mathcal{L}$. Using $\boldsymbol{A}^b_{j,k}$, we can formulate $\pi^b_{j,k}$ based on the state of its parent as follows,

$$\pi^b_{j,k} = \pi^b_{j-1,\lceil k/4 \rceil} p^{\mathcal{S}\to\mathcal{S}}_{j,k,b} + \left(1 - \pi^b_{j-1,\lceil k/4 \rceil}\right) p^{\mathcal{L}\to\mathcal{S}}_{j,k,b}, \quad (3)$$

given the probability of being small at root coefficients ($\pi^b_{1,k}$) is known for all values of b, k. Later in this section, we will determine $\pi^b_{1,k}$ according to the image size and the number of wavelet scales.

Although the HMT model is very powerful in capturing the properties of wavelet coefficients, it requires at least $4n$ parameters to be specified. However, in [10], the authors proposed to reduce the number of parameters to $4J$ by assuming all coefficients within one wavelet scale have similar statistical parameters. Therefore, the parameters reduce to

$$\pi^b_{j,k} = \pi_j,\ {\sigma^b_{\mathcal{S},\{j,k\}}}^2 = \sigma^2_{\mathcal{S},j},\ {\sigma^b_{\mathcal{L},\{j,k\}}}^2 = \sigma^2_{\mathcal{L},j},\ \text{and}\ \boldsymbol{A}^b_{j,k} = \boldsymbol{A}_j, \quad (4)$$

for all b, j, k. These parameters can be estimated using a set of training images and exploiting the Expectation-Maximization (EM) algorithm as done in [10]. However, it has been shown in [11] that leveraging additional wavelet-domain image structure (such as exponential decay across scale), a reduced-parameter HMT model can be developed that is represented with only 9 meta parameters independent of the size of the image and the number of wavelet scales. Further, it has been shown in [11] that these 9 parameters take similar values for real-world images, allowing to fix a set of universal set of parameters, resulting in a *universal HMT (uHMT)*. Employing uHMT, the image-specific training is avoided.[2]

Let the uHMT model be represented by a hyper-parameter $\boldsymbol{\Lambda} = [\alpha_{\mathcal{S}}, \alpha_L, C_{\sigma_{\mathcal{S}}}, C_{\sigma_{\mathcal{L}}}, \gamma_{\mathcal{S}}, \gamma_{\mathcal{L}}, C_{\mathcal{SS}}, C_{\mathcal{LL}}, \pi'_1]$ [11]. These 9 parameters are used to determine *a priori* pdfs (as given by (14)) for all the wavelet coefficients. The variances of large and small coefficients vary over the scales [11]:

$$\sigma^2_{\mathcal{L},j} = C_{\sigma_{\mathcal{L}}} 2^{-(J_d+j)\alpha_L}, \quad (5a)$$

$$\sigma^2_{\mathcal{S},j} = C_{\sigma_{\mathcal{S}}} 2^{-(J_d+j)\alpha_S}, \quad (5b)$$

where $C_{\sigma_{\mathcal{L}}} \gg C_{\sigma_{\mathcal{S}}}$ and $J_d = \log_4 n - J$. Four parameters $C_{\sigma_{\mathcal{L}}}, \alpha_L, C_{\sigma_{\mathcal{S}}}$, and α_S characterize the variances in the marginal

[2]The uHMT parameters are accurate when $J \leq \log_4 n - 3$ [11], which complies with our proposed sensing scheme.

densities of the wavelet coefficients [11]. Using the uHMT model, $\boldsymbol{A}_j$ is given by [11],

$$\boldsymbol{A}_j = \begin{cases} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} & j \leq 3 - J_d, \\ \begin{bmatrix} 1 - C_{\mathcal{SS}} 2^{-(J_d+j)\gamma_{\mathcal{S}}} & C_{\mathcal{SS}} 2^{-(J_d+j)\gamma_{\mathcal{S}}} \\ \frac{1}{2} - C_{\mathcal{LL}} 2^{-(J_d+j)\gamma_{\mathcal{L}}} & \frac{1}{2} + C_{\mathcal{LL}} 2^{-(J_d+j)\gamma_{\mathcal{L}}} \end{bmatrix} & j > 3 - J_d. \end{cases} \tag{6}$$

Unlike [11], we perform the wavelet decomposition up to an arbitrary scale J and not necessarily up to scale $\log_4 n$ (which is the full wavelet decomposition). Therefore, we add the deficit term $J_d = \log_4 n - J$ in Equations (5)–(6) to compensate for this incomplete decomposition.

The parameter π'_1 in $\boldsymbol{\Lambda}$ is the probability that the root coefficients are small when the wavelet decomposition is carried out up to the *coarsest* scale (i.e., $J = \log_4 n$). In this case, we find the probability that a coefficient at scale j is in the small state, π'_j, using the following recursive formula

$$\pi'_j = \pi'_{j-1} p_j^{'\mathcal{S}\to\mathcal{S}} + (1 - \pi'_{j-1}) p_j^{'\mathcal{L}\to\mathcal{S}}, \tag{7}$$

where $p_j^{'\mathcal{S}\to\mathcal{S}}$ and $p_j^{'\mathcal{L}\to\mathcal{S}}$ are entries of transition matrix $\boldsymbol{A}_j$ given $J = \log_4 n$. In our problem, since we perform partial wavelet decomposition (i.e., $J < \log_4 n$), we set $\pi_1 = \pi'_{\log_4 n - J + 1}$ to compensate for the scale deficit. From π_1, we can determine π_j's using (3) and (6) for $j = 2, \ldots, J$.

From (3) and (6), we note that π_j is an increasing function of j. Therefore, the sparsity rate of wavelet coefficients at scale j, given by $s_j = 1 - \pi_j$, is a decreasing function of j. In the next section, we will exploit this *nonuniform sparsity* property of wavelet coefficients to design a novel CS measurement matrix.

As an example, for a 128×128 image with $J = 4$ and using the parameters of uHMT model from [11] ($\gamma_{\mathcal{S}} = \gamma_{\mathcal{L}} = 1, C_{\mathcal{SS}} = 0.2$, $C_{\mathcal{LL}} = 0.4$, $\alpha_{\mathcal{S}} = \alpha_{\mathcal{L}} = 2.5$ and $\pi'_1 = 0.5$), we find $\pi_1 = 0.731$, $\pi_2 = 0.858$, $\pi_3 = 0.925$ and $\pi_4 = 0.961$. Therefore, the sparsity rates of different scales are found as $s_1 = 0.269$, $s_2 = 0.142$, $s_3 = 0.075$, $s_4 = 0.039$. The overall sparsity rate of the wavelet coefficients excluding the approximate coefficients is found as $\sum_{j=1}^{J} \alpha_j s_j = 0.055$ and we see the sparsity rate of lower scales are much higher than the overall sparsity rate of the image.

III. Model-based Nonuniform Compressive Sampling

In this section, we introduce our proposed *nonuniform* measurement matrix $\boldsymbol{\Phi}_{uHMT}$ for natural images. Our approach is based on the integration of the uHMT model of natural images' wavelet coefficients into the design of the measurement matrix.

A. Nonuniform Sampling

We propose a nonuniform CS sampling matrix that can be implemented in the design of CS-based image sensors. Among many proposed techniques in the literature, on-chip realizations of CS-based image sensors with single shot image capture is highly desirable [25]. Employing CS-based image sensors with these realizations can reduce the energy consumption of the image acquisition compared to current designs that are not based on compressive sensing [25].

The compressive sampling process in the *pixel-domain* is performed using the sampling matrix $\boldsymbol{\Xi}$. It is worth noting that the wavelet coefficients of the image are not available for sampling and the goal is to obtain compressive samples from the analog readings of the image sensor in pixel-domain. The relationship between $\boldsymbol{\Xi}$ and $\boldsymbol{\Phi}_{uHMT}$ is given in the following.

$$\boldsymbol{y} = \boldsymbol{\Xi}\boldsymbol{x} = \boldsymbol{\Phi}_{uHMT}\boldsymbol{\Psi}^T\boldsymbol{x}, \tag{8}$$

where $\boldsymbol{\Psi}$ is the 2D wavelet transform matrix. In other words, the compressive sensing measurement matrix that is applied to image $\boldsymbol{x}$ is $\boldsymbol{\Xi} = \boldsymbol{\Phi}_{uHMT}\boldsymbol{\Psi}^T$. Therefore, applying the measurement matrix $\boldsymbol{\Phi}_{uHMT}$ to wavelet-domain coefficients is equivalent to applying the compressive sampling matrix $\boldsymbol{\Xi}$ in pixel-domain given the *incoherency condition* holds between $\boldsymbol{\Phi}_{uHMT}$ and wavelet sparsifying basis (this condition will be addressed Section III-B).

It is known that the wavelet coefficients of an image show an *exponential decay* along the scales of the wavelet tree [10], [11]. This means that most of images' energy is carried by the approximate coefficients and the wavelet coefficients of the coarser scales. In addition, the sparsity rate of wavelet coefficients at scale j, given by $s_j = 1 - \pi_j$, is a decreasing function of j. Therefore, it seems rational to sample the coefficients at coarser scales, which are the initial coefficients in $\boldsymbol{\theta}|_J$, with a higher probability and decrease the probability that a coefficient is sampled by a CS measurement as we increase j. For these purposes, we propose a novel measurement matrix $\boldsymbol{\Phi}_{uHMT}$ that integrates the uHMT model of wavelet coefficients into its design.

The proposed measurement matrix *directly* samples the $\alpha_0 n = 4^{-J} n$ approximate coefficients and nonuniformly samples the remaining wavelet coefficients using *a nonuniformly sparse matrix* $\boldsymbol{\Phi}_{NU} = [\boldsymbol{\Phi}_1, \boldsymbol{\Phi}_2, \ldots, \boldsymbol{\Phi}_J]$, where sub-matrix $\boldsymbol{\Phi}_j$ corresponds to the $\alpha_j n = 3 \times 4^{-(J-j+1)} n$ wavelet coefficients at scale j, for $j = 1, 2, \ldots, J$. Here $J \leq \log_4 n$ represents the depth of the wavelet trees (i.e., the level up to which the wavelet decomposition has been done). It is worth noting that there is no need to perform wavelet transform and obtain approximate coefficients in the sampling process.

In addition, we set every row of $\boldsymbol{\Phi}_{NU}$ to have L non-zero coefficient, chosen from a Gaussian distribution with zero mean and variance $\frac{1}{L}$. The structure of the proposed measurement matrix $\boldsymbol{\Phi}_{uHMT}$ is shown in Fig. 3, in which $\boldsymbol{I}_{\alpha_0 n}$ is an identity matrix of size $\alpha_0 n \times \alpha_0 n$.

If we take m CS measurements, $m - \alpha_0 n$ of them will be generated nonuniformly from the wavelet coefficients other than the approximate coefficients. To impose the nonuniform sampling through $\boldsymbol{\Phi}_{NU}$, we set each row of $\boldsymbol{\Phi}_j$ to have exactly L_j nonzero entries, whose locations are selected uniformly at random out of $\alpha_j n$ columns (similar to [20]) such that $\sum_{j=1}^{J} L_j = L$.

Let $\boldsymbol{y} = \boldsymbol{\Phi}_{uHMT}\boldsymbol{\theta}|_J$ denote the compressive samples of wavelet coefficients of an image. For a measurement y_i ($\alpha_0 n + 1 \leq i \leq m$), the number of contributing coefficients from scale j is L_j. Since the significance of wavelet coefficients decreases

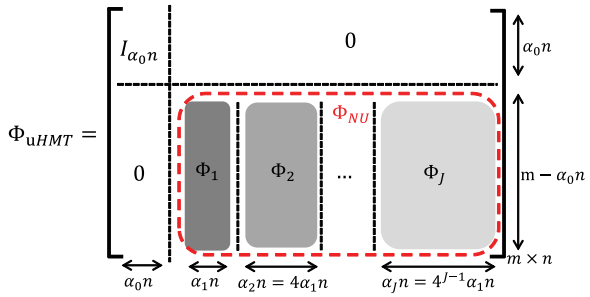


Fig. 3. The proposed nonuniform measurement matrix for compressive sensing of wavelet coefficients of an image. A darker color corresponds to a denser matrix.

as j increases, we set

$$\frac{L_1}{\alpha_1} > \frac{L_2}{\alpha_2} > \ldots > \frac{L_J}{\alpha_J}. \tag{9}$$

This will ensure that the coefficients at lower scales contribute to more measurements and correspondingly they have more contribution in each measurement.

In our proposed scheme, we relate L_j to the sparsity of wavelet coefficients as follows:

$$L_j = \begin{cases} \left\lceil \dfrac{\alpha_j s_j}{\sum_{l=1}^{J} \alpha_l s_l} L \right\rceil & j = 1, \ldots, J-1 \\ L - \sum_{l=1}^{J-1} L_j & j = J. \end{cases}, \tag{10}$$

where $s_j = 1 - \pi_j$ is the sparsity rate of the wavelet coefficients at scale j and π_j is found using the uHMT model discussed in Section II-B. In (10), L_J will be greater than zero if we have $L \geq \frac{\sum_{j=1}^{J} \alpha_j s_j}{\alpha_J s_J}(J-1)$. Since $J = \mathcal{O}(\log n)$ represents the levels of the wavelet transform, choosing L to be at least $\mathcal{O}(\log n)$ suffices to have $L_J \geq 0$. As we mentioned in the proposed nonuniform sampling, we set $L = \frac{n}{c}$. Therefore for sufficiently large n, L_j will be greater than or equal to zero.

It should be noted that proposed sensing matrix is universal given the size of the image size (determined by the image acquisition hardware) and the number of wavelet scales. This is rooted in the properties of uHMT model, which is independent of the image size and number of wavelet scales [11].

B. Mutual Coherence of $\mathbf{\Phi}_{uHMT}$ and $\mathbf{\Psi}$

In order to have an efficient sampling and recovery process, the measurement matrix and the sparsifying basis must be mutually incoherent. The *mutual coherence* μ of the measurement matrix and the sparsifying matrix is defined as $\mu(\mathbf{\Phi}, \mathbf{\Psi}) = \sqrt{n} \max_{k,j} |\langle \boldsymbol{\phi}_k, \boldsymbol{\psi}_j \rangle|$ [26], where $\boldsymbol{\phi}_k$ is the k^{th} row $\mathbf{\Phi}$ and $\boldsymbol{\psi}_j$ is the j^{th} column of $\mathbf{\Psi}$ and $\|\boldsymbol{\phi}_k\|_2 = 1$ for all $k = 1, \ldots, m$ and $\|\boldsymbol{\psi}_j\|_2 = 1$ for all $j = 1, 2, \ldots, n$.

As shown in Fig. 3, our proposed measurement matrix can be written as $\mathbf{\Phi}_{uHMT} = \begin{bmatrix} [I_{\alpha_0 n} & | & \mathbf{0} \;]_{\alpha_0 n \times n} \\ [\;\mathbf{0} & | & \mathbf{\Phi}_{NU}]_{(m-\alpha_0 n)\times n} \end{bmatrix}$. Therefore, the mutual coherence of $\mathbf{\Phi}_{uHMT}$ and $\mathbf{\Psi}$ is

$$\mu(\mathbf{\Phi}_{uHMT}, \mathbf{\Psi}) = \max\{\mu([\boldsymbol{I}_{\alpha_0 n}|\mathbf{0}], \mathbf{\Psi}), \mu([\mathbf{0}|\mathbf{\Phi}_{NU}], \mathbf{\Psi})\}. \tag{11}$$

Given $\mathbf{\Psi}$ is an orthonormal basis with $|\psi_{ij}| = \mathcal{O}(\frac{\log n}{\sqrt{n}})$ for all $i, j = 1, \ldots, n$, we have $\mu([\boldsymbol{I}_{\alpha_0 n}|\mathbf{0}], \mathbf{\Psi}) = \mathcal{O}(\log n)$. In addition using the following lemma, we find the mutual coherence of $[\mathbf{0}|\mathbf{\Phi}_{NU}]$ and any orthonormal bases.

Lemma 1: Let $\mathbf{\Phi}$ be an m-by-n sparse random matrix. Assume every row of $\mathbf{\Phi}$ has $L = \frac{n}{c}$ non-zero entries (not necessarily uniformly distributed) where each non-zero entry is chosen from $\mathcal{N}(0, \frac{1}{L})$ and c is a large constant. For any arbitrary orthonormal basis $\mathbf{\Psi}_{n\times n}$, we have $\mathbf{\Xi} = \mathbf{\Phi}\mathbf{\Psi}$ is an iid zero-mean Gaussian matrix with variance less than or equal to $\frac{1}{L}$.

Proof: Assume $\boldsymbol{\phi}_k$ is the k^{th} row $\mathbf{\Phi}$ and $\boldsymbol{\psi}_j$ is the j^{th} column of $\mathbf{\Psi}$ and the L indices of non-zero elements in $\boldsymbol{\phi}_k$ are denoted by set R_k ($R_k \subset \{1, \ldots, n\}$). The matrix $\mathbf{\Xi} = \mathbf{\Phi}\mathbf{\Psi}$ with entries $\Xi_{kj} = \sum_{i=1}^{n} \phi_{ki}\psi_{ij} = \sum_{i\in R_k} \phi_{ki}\psi_{ij}$ is defined. Therefore, it can be easily seen that for all $j = 1, \ldots, n$ and $k = 1, \ldots, m$, Ξ_{kj} is a linear combination of L Gaussian random variables. Accordingly, Ξ_{kj} itself is a Gaussian random variable with mean $E\{\Xi_{kj}\} = \sum_{i\in R_k} E\{\phi_{ki}\}\psi_{ij}$ and variance $\mathrm{Var}(\Xi_{kj}) = \sum_{i\in R_k} \psi_{ij}^2 \mathrm{Var}(\phi_{ki})$. Hence, we have the following,

$$E\{\Xi_{kj}\} = 0,$$
$$\mathrm{Var}(\Xi_{kj}) = \frac{1}{L} \sum_{i\in R_k} \psi_{ij}^2 \leq \frac{1}{L}$$

■

Lemma 1 implies that $[\mathbf{0}|\mathbf{\Phi}_{NU}] \times \mathbf{\Psi}$ is a Gaussian random matrix with iid zero-mean entries and variance less than or equal to $1/L$. Using the union bound for the maximum absolute magnitude of a Gaussian matrix, $|\Xi_{kj}|$ for all $j = 1, \ldots, n$ and $k = 1, \ldots, m$ can be bounded as follows [16],

$$P\left(\max_{1\leq k\leq m, 1\leq j\leq n} |\Xi_{kj}| \geq t\right) \preceq 2\,nm \exp\left(-\frac{t^2}{2\sigma^2}\right) \leq 2n^2 \exp\left(-\frac{t^2}{2\sigma^2}\right), \tag{12}$$

where $\sigma^2 \leq \frac{1}{L} = \frac{c}{n}$ and $\preceq$ represents *asymptotically smaller than or equal*. Choosing $t = \sqrt{\frac{2c\log\left(\frac{2n^2}{\delta}\right)}{n}}$, the inequality in (12) becomes,

$$P\left(\max_{1\leq k\leq m, 1\leq j\leq n} |\Xi_{kj}| \leq \sqrt{\frac{2c\log\left(\frac{2n^2}{\delta}\right)}{n}}\right) \succeq 1 - \delta. \tag{13}$$

Inequality (13) shows that $\mu([\mathbf{0}|\mathbf{\Phi}_{NU}], \mathbf{\Psi}) = \mathcal{O}(\sqrt{\log(\frac{n}{\sqrt{\delta}})})$, with probability at least $1-\delta$. Consequently, $\mu(\mathbf{\Phi}_{uHMT}, \mathbf{\Psi}) = \max\{\mu([\boldsymbol{I}_{\alpha_0 n}|\mathbf{0}], \mathbf{\Psi}), \mu([\mathbf{0}|\mathbf{\Phi}_{NU}], \mathbf{\Psi})\} = \mathcal{O}(\sqrt{\log(\frac{n}{\sqrt{\delta}})})$, which is close to optimal bound except for the $\log n$ factor. It is worth mentioning that the Haar wavelet satisfies the condition $|\psi_{ij}| = \mathcal{O}(\frac{\log n}{\sqrt{n}})$ for all $i, j = 1, \ldots, n$ [27]. Therefore, the Haar wavelet basis, $\mathbf{\Psi}_{Haar}$, satisfies $\mu(\mathbf{\Phi}_{uHMT}, \mathbf{\Psi}_{Haar}) = \mathcal{O}(\sqrt{\log(\frac{n}{\sqrt{\delta}})})$.

Algorithm 1 summarizes our proposed algorithm for nonuniform compressive sampling of 2D wavelet coefficients of an image.

Algorithm 1: Nonuniform compressive sampling of wavelet coefficients of an image.

1: Initialize $\alpha_0 = 4^{-J}$, $\alpha_j = 3 \times 4^{-(J-j+1)}$, $s_j = 1 - \pi_j$ for $j = 1, \ldots, J$. The values of π_j's are found from the uHMT model discussed in Section II-B.
2: Generate $\mathbf{\Phi}_{NU} = [\mathbf{\Phi}_1 \ \mathbf{\Phi}_2 \ \ldots \ \mathbf{\Phi}_J]$ based on the structure in Fig. 3. Each row of $\mathbf{\Phi}_j$ has L_j non-zero entries, which are chosen as iid zero-mean and unit variance Gaussian random variables.
3: Directly sample $\alpha_0 n$ initial coefficients of $\boldsymbol{\theta}$.
4: Sample the remaining $(1 - \alpha_0)n$ coefficients of $\boldsymbol{\theta}$ using the updated $\mathbf{\Phi}_{NU}$.

IV. Model-based Nonuniform CS Recovery

In Section III, we integrated the uHMT signal model into the design of a novel nonuniform measurement matrix. In this section, we utilize the same model at the recovery step. We consider two Bayesian recovery algorithms (i.e., the CSBP algorithm [20] and the AMP algorithm [9]) and modify them to exploit the uHMT model.

A. CSBP-uHMT: Integrating the uHMT Model into CSBP

One of the advantages of the CSBP recovery algorithm [20] is its ability to accommodate *a priori* knowledge about the signal model in the CS recovery process. In the conventional CSBP, all the variable nodes are assigned *the same a priori* pdf that considers the *sparsity rate* $s = k/n$ as the probability that each coefficient is at the large state.

In contrast, in our proposed CSBP-uHMT, we assign a *different* prior to each variable node based on the uHMT model. Specifically, each wavelet coefficient $\theta^b_{j,k}$ receives an *a priori* mixture Gaussian pdf

$$f(\theta^b_{j,k}) = \pi_j \mathcal{N}(0, \sigma^2_{\mathcal{S},j}) + (1 - \pi_j)\mathcal{N}(0, \sigma^2_{\mathcal{L},j}). \tag{14}$$

The hyper-parameter $\mathbf{\Lambda} = [\alpha_{\mathcal{S}}, \alpha_L, C_{\sigma_{\mathcal{S}}}, C_{\sigma_{\mathcal{L}}}, \gamma_{\mathcal{S}}, \gamma_{\mathcal{L}}, C_{\mathcal{SS}}, C_{\mathcal{LL}}, \pi'_1]$ is used to determine π_j, $\sigma^2_{\mathcal{S},j}$, and $\sigma^2_{\mathcal{L},j}$ as discussed in Section II-B.

B. AMP-uHMT: Integrating the uHMT into AMP

The AMP algorithm [9] and its model-based version (Turbo AMP) [7] are proven to have very competitive recovery performances and very low computational complexity. Unlike Turbo AMP that employs a learning-based approach to obtain the statistical parameters of HMT model, we directly apply the uHMT parameters to the AMP algorithm. In Turbo AMP, the variances $\sigma^b_{\mathcal{S},\{j,k\}}{}^2$ and $\sigma^b_{\mathcal{L},\{j,k\}}{}^2$ and the probabilities $\pi^b_{j,k}$ are assumed to be random variables with known distributions [7, Eq. (3)–(8)]. In our proposed work we treat them as fixed and known parameters found by uHMT model using Equations (3) and (5).

Clearly the uHMT model may not be as accurate as the training-based approaches. However, it has been shown in [11] that using uHMT has a negligible degrading effect on the accuracy of images.

TABLE I
Properties of Different CS Sampling and Recovery Schemes

Scheme	Model-based Recovery	Model-Based Sampling	Measurement Matrix
Model-based CS [5]	✓	x	Full Gaussian
TSWCS [6]	✓	x	Direct sampling+ full Gaussian
Turbo AMP [7]	✓	x	Full Gaussian
AMP [9] w/ sparse random $\mathbf{\Phi}$	x	x	Sparse random $\mathbf{\Phi}$ ($L = 40$)
uHMT-NCS w/ AMP	x	✓	$\mathbf{\Phi}_{uHMT}$
Sparse Random Sampling w/ AMP-uHMT	✓	x	Sparse random $\mathbf{\Phi}$ ($L = 40$)
uHMT-NCS w/ AMP-uHMT	✓	✓	$\mathbf{\Phi}_{uHMT}$

V. Simulation Results and Discussion

In this section, we compare the recovery performance of our proposed model-based uHMT-NCS algorithms (which is based on the integration of the HMT model into both sampling and recovery steps) with other state-of-the-art model-based CS recovery schemes that only integrate the model at the recovery step and use conventional CS at the sampling step. The algorithms from the literature that we have chosen are Turbo AMP [7], model-based CS [5], and TSWCS-MCMC [6]. To implement these algorithms, we use the full Gaussian measurement matrices with iid entries with these algorithms (See Table I). We consider a 128×128 image ($n = 16,384$) for our simulations (given in Figs. 7(a) and 8(a)). The sparsifying basis, $\mathbf{\Psi}$, is considered to be the two-dimensional Haar wavelet basis. The parameters of uHMT model are set as described in [11] (we verified these parameters for a set of 128×128 test natural images and Haar wavelet): $\alpha_{\mathcal{S}} = \alpha_L = 2.5, C_{\sigma_{\mathcal{S}}} = 2^7, C_{\sigma_{\mathcal{L}}} = 2^{13}$, $\gamma_{\mathcal{S}} = \gamma_{\mathcal{L}} = 1$, $C_{\mathcal{SS}} = 0.2$, $C_{\mathcal{LL}} = 0.4$, and $\pi'_1 = 0.5$. We set $L = 40$ and $J = 3$. Therefore, we find $\pi_1 = 0.858, \pi_2 = 0.925$, and $\pi_3 = 0.961$. For uHMT-NCS simulations, the matrix $\mathbf{\Phi}$ is generated using Algorithm 1.

In Fig. 4, we have shown the normalized recovery error (NRE $= \frac{\|\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}\|_2}{\|\boldsymbol{\theta}\|_2}$) versus the number of measurements. We depicted the performance of our proposed uHMT-NCS (with CSBP-uHMT and AMP-uHMT) and Turbo AMP, model-based CS, and TSWCS-MCMC schemes. For these algorithms, the tunable parameters such as wavelet levels are set for the best performance. As we see in Fig. 4, Bayesian recovery-based algorithms such as TSWCS-MCMC [6] and our proposed uHMT-NCS have very good performances even with small number of measurements. Among all schemes uHMT-NCS with AMP-uHMT recovery performs the best.

A similar simulation is performed over sample images the Microsoft object class recognition database,[3] which includes thousands of weakly labeled images in 18 categories. From every category either one or two images are randomly chosen,

[3]All images are cropped to be rectangular, and resized to 128×128. The database is available for download at "http://research.microsoft.com/en-us/projects/objectclassrecognition/"

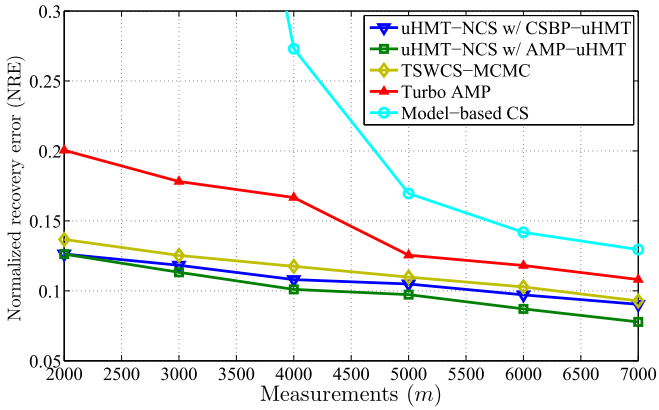


Fig. 4. Comparison between the recovery performances of our proposed schemes (uHMT-NCS w/ CSBP-uHMT and uHMT-NCS w/ AMP-uHMT) and different CS recovery algorithms (Turbo AMP [7], Model-based CS [5], and TSWCS-MCMC [6]).

Fig. 5. A set of 27 sample images randomly chosen from different categories of Microsoft object class recognition database. The numbers in each image are solely for referencing and are not part of the image.

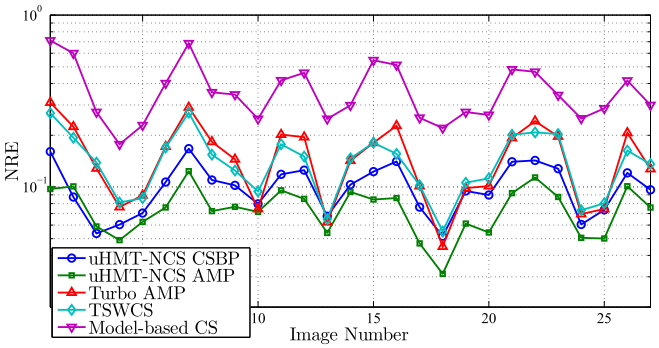
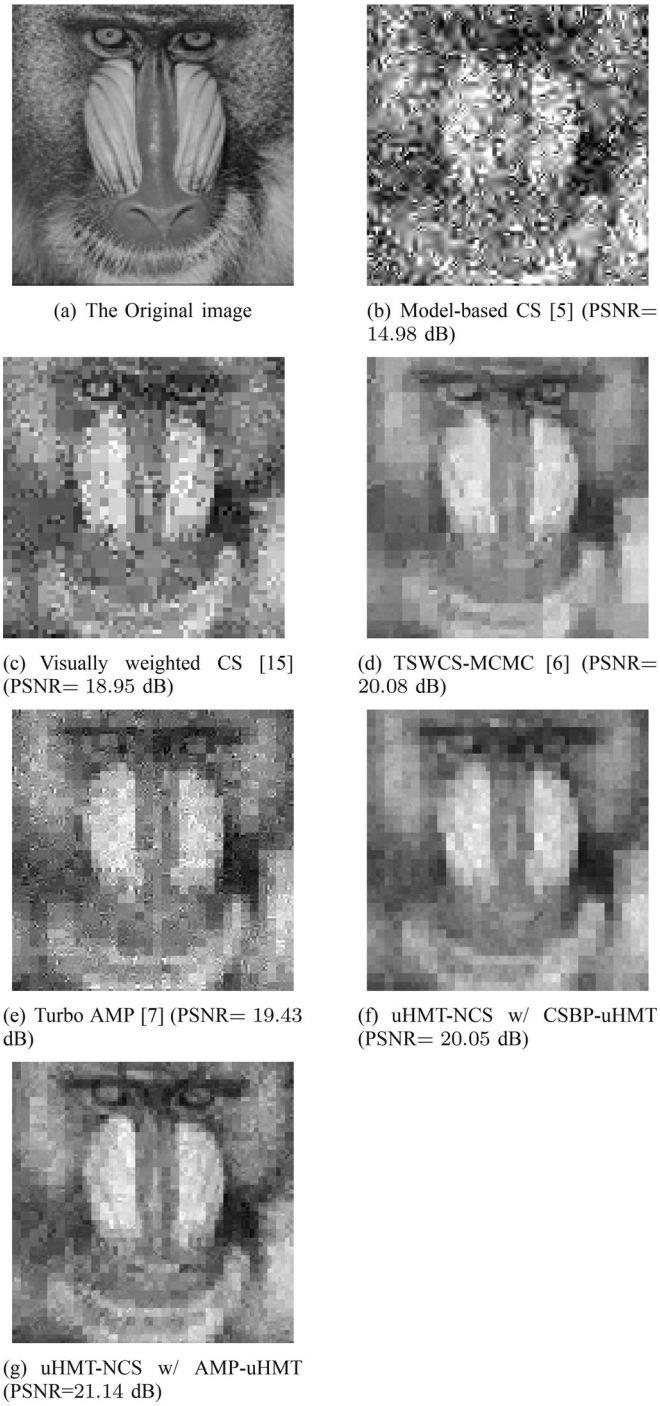


Fig. 6. NRE performance for the sample images in Fig. 5.

resulting in a total of 27 images (see Fig. 5). We applied our proposed uHMT-NCS (with CSBP-uHMT and AMP-uHMT) as well as Turbo AMP, model-based CS, and TSWCS-MCMC to these images. The NRE performance of each image using each algorithm is shown in Fig. 6. For this simulation we set $m = 4000$ and $J = 3$ and 2D Haar wavelet in our own algorithms. However, the model-based CS algorithm is tuned for its best performance which is achieved with $J = 6$.

As we see in Fig. 6, in 25 out of 27 images our proposed uHMT-NCS performs better than other algorithms in comparison.



(a) The Original image

(b) Model-based CS [5] (PSNR= 14.98 dB)

(c) Visually weighted CS [15] (PSNR= 18.95 dB)

(d) TSWCS-MCMC [6] (PSNR= 20.08 dB)

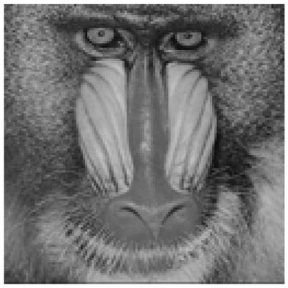
(e) Turbo AMP [7] (PSNR= 19.43 dB)

(f) uHMT-NCS w/ CSBP-uHMT (PSNR= 20.05 dB)

(g) uHMT-NCS w/ AMP-uHMT (PSNR=21.14 dB)

Fig. 7. Comparing the visual performance of different CS schemes at $m = 4000$ measurements.

Figs. 7 and 8 provide the visual comparison of the recovery performance of uHMT-NCS with state-of-the-art algorithms. The number of measurements in Figs. 7 and 8 are $m = 4000$ and $m = 6000$, respectively. Figs. 7(a) and 8(a) show the original image, Figs. 7(b)–8(b), 7(c)–8(c), Figs. 7(d)–8(d), and Figs. 7(e)–8(e) show the visual performance of model-based CS [5], visually weighted CS [15], TSWCS-MCMC Bayesian algorithm [6], and Turbo AMP [7], respectively. Finally, Figs. 7(f)–8(f), and 7(g)–8(g) represent our proposed

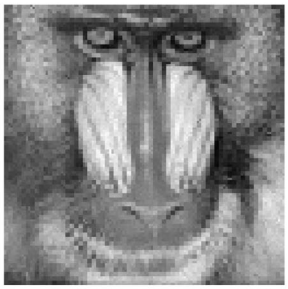


(a) The Original image

(b) Model-based CS (PSNR= 20.59 dB)

(c) Visually weighted CS (PSNR= 21.63 dB)

(d) TSWCS-MCMC (PSNR= 22.98 dB)



(e) Turbo AMP (PSNR= 22.83 dB)

(f) uHMT-NCS w/ CSBP-uHMT (PSNR= 23.55 dB)

(g) uHMT-NCS w/ AMP-uHMT (PSNR= 24.15 dB)

Fig. 8. Comparing the visual performance of different CS schemes at $m = 6000$ measurements.

uHMT-NCS algorithm with CSBP-uHMT and AMP-uHMT recovery algorithms, respectively. As we see, uHMT-NCS with both CSBP-uHMT and AMP-uHMT recovery algorithms has a significantly smaller visual distortion compared to the others.

In the next simulation, we investigate the effect of employing our proposed nonuniform $\mathbf{\Phi}_{uHMT}$ in the overall performance of the proposed uHMT-NCS algorithm. In other words, we would like to see how much of the improved performance is due to the integration of the model at the sampling step through our proposed nonuniform $\mathbf{\Phi}_{uHMT}$. In Fig. 9, we compare the NRE performance of AMP [9] with the cases when the uHMT model is added to only sampling step (uHMT-NCS w/ AMP), only recovery step (sparse random sampling w/ AMP-uHMT), and both steps (uHMT-NCS w/ AMP-uHMT) (See Table I). When uHMT is only applied to the recovery step the measurement matrix is sparse random matrix with $L = 40$ non-zero entries per row. As we see in Fig. 9, including the uHMT model only at the sampling step (uHMT-NCS w/ AMP) improves the performance compared to the case when it is solely added to the recovery step. Therefore, *integrating the model at the sampling phase* is even *more effective* than such integration at the *recovery phase*. Clearly, when uHMT model is included in both sampling and recovery steps (denoted by uHMT-NCS w/ AMP-uHMT) the most performance improvements is achieved.

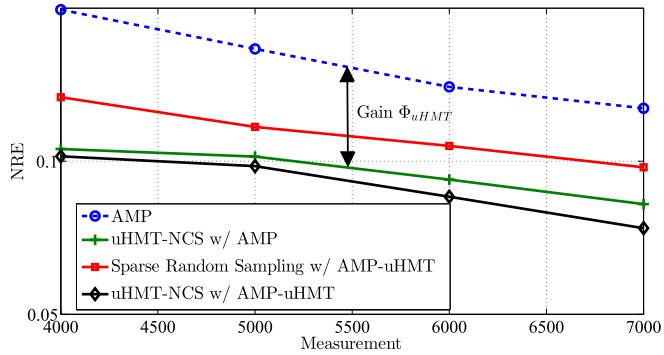


Fig. 9. Improvement over the conventional AMP technique by exploiting the uHMT signal model at the sampling step (uHMT-NCS w/ AMP), the recovery step (sparse random sampling w/ AMP-uHMT), and both (uHMT-NCS w/ AMP-uHMT).

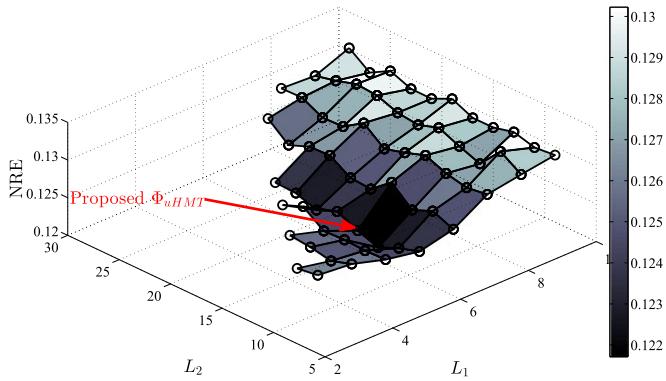


Fig. 10. NRE versus L_1 and L_2. Choosing L_1 and L_2 using (10) results in almost 6% performance improvement.

In the next simulation, we consider the effect of changing the row weights of $\mathbf{\Phi}_j$ s on the performance of the uHMT-NCS. We set $m = 4000$, $L = 40$ and $J = 3$ and change the row weights of $\mathbf{\Phi}_j$ s. Accordingly, we find the NRE of uHMT-NCS with CSBP-uHMT changing L_1 and L_2 parameters. As shown in Fig. 10, employing the values of L_1 and L_2 given in (10) (i.e. $L_1 = 6$ and $L_2 = 12$) results in a smaller NRE which is about 6% better than the uniform sampling case (i.e. $L_1 = 2, L_2 = 8$). However,

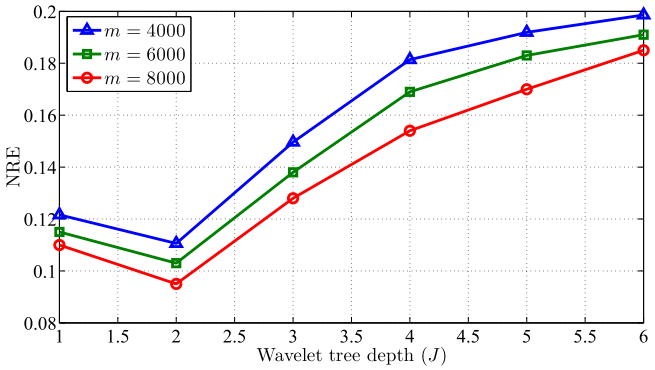


Fig. 11. Recovery performance of uHMT-NCS with CSBP-uHMT versus different depths of the wavelet tree J.

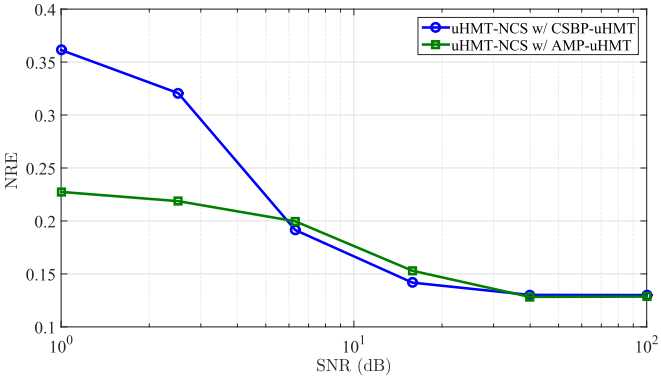


Fig. 12. Recovery performance of uHMT-NCS with CSBP-uHMT and AMP-uHMT versus signal to measurement noise ratio.

for this particular example for $4 \leq L_1 \leq 6$ and $10 \leq L_2 \leq 18$ the results are within the 2% of its minimum value.

In the next simulation, we study the effect of the parameter J (the depth of wavelet trees) in the performance. In Fig. 11, we have depicted the NRE of our proposed uHMT-NCS algorithm with the CSBP-uHMT recovery scheme versus J for the sample image of Fig. 7(a) for different values of m.

As shown, by increasing tree depth, NRE initially decreases and then increases. The optimal performance is obtained at $J = 2$. This observation can be explained using the *exponential decay* property of wavelet coefficients. This property suggests that most of images energy are located in the initial signal coefficients. When $J = 1$, direct sampling dominates the measurements. Therefore, many coefficients will not be included in any measurement and that results in poor performance. As J increases, fewer coefficients are measured directly and more nonuniform CS samples are generated. However, by increasing J the sparsity rate of higher scales decreases exponentially and it is inefficient to allocate too many measurements to sample these coefficients. We used optimal value $J = 2$ for our previous simulations.

Finally, we consider the performance of the proposed algorithms in case of noisy measurements. Similar to the previous simulations, we set $m = 4000$, $L = 40$ and $J = 3$. In this simulation we assumed the measurements are corrupted by an additive white Gaussian noise process. As shown in Fig. 12, the normalized recovery error of uHMT-NCS with CSBP recovery and AMP recovery are depicted versus different values of the SNR. SNR is calculated as the ratio of the average signal value to the standard deviation of the Gaussian noise. The results of this simulation show that for SNR levels higher than 6 dB the performance of both proposed uHMT-NCS CSBP and AMP algorithms are very similar to each other. However, in low SNR scenarios the performance of the uHMT-NCS AMP is considerably better than uHMT-NCS CSBP. This observation roots in the more accurate noise model of AMP algorithm in the case of AWGN compared to the CSBP [9]. Therefore for low SNR scenarios, the uHMT-NCS with AMP-uHMT recovery should be employed due to more robustness to noise.

VI. Conclusion

In this paper, we have developed a *model-based CS nonuniform sampling and recovery* scheme (uHMT-NCS) for natural images that exploits the universal hidden Markov tree (uHMT) model of wavelet coefficients in *both* CS sampling and CS recovery steps. As we have shown, not only the signal model can be utilized to optimize the initial priors for Bayesian CS recovery algorithms, but it can also be employed in the design of new CS measurement matrices. The results of our numerical experiments suggest a significant performance gain compared to the state-of-the-art model-based CS algorithms. To the best of our knowledge, our work is one of the first Bayesian-based algorithms to consider the signal model in the design of the CS measurement matrix as well as the recovery scheme. Although we considered the natural images as our underlying signal, our approach is not limited to the image processing applications and the wavelet sparsifying bases. After the model extraction, similar procedure can be adopted for compressive sampling and recovery of other real-world signals.

References

[1] D. Donoho, "Compressed sensing," *IEEE Trans. Inf. Theory*, vol. 52, no. 4, pp. 1289–1306, Apr. 2006.

[2] E. Candès, "Compressive sampling," in *Proc. Int. Congr. Mathematicians*, 2006, vol. 1, no. 3, pp. 1433–1452.

[3] A. Schulz, L. Velho, and E. A. B. Da Silva, "On the empirical rate-distortion performance of compressive sensing," in *Proc. 16th IEEE Int. Conf. Image Process.*, Nov. 2009, pp. 3049–3052.

[4] M. Duarte, M. Wakin, and R. Baraniuk, "Wavelet-domain compressive signal reconstruction using a hidden Markov tree model," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, Mar. 2008, pp. 5137–5140.

[5] R. Baraniuk, V. Cevher, M. Duarte, and C. Hegde, "Model-based compressive sensing," *IEEE Trans. Inf. Theory*, vol. 56, no. 4, pp. 1982–2001, Apr. 2010.

[6] L. He, H. Chen, and L. Carin, "Tree-structured compressive sensing with variational Bayesian analysis," *IEEE Signal Process. Lett.*, vol. 17, no. 3, pp. 233–236, Mar. 2010.

[7] S. Som and P. Schniter, "Compressive imaging using approximate message passing and a Markov-tree prior," *IEEE Trans. Signal Process.*, vol. 60, no. 7, pp. 3439–3448, Jul. 2012.

[8] P. Indyk and I. Razenshteyn, "On model-based RIP-1 matrices," in *Automata, Languages, and Programming* (ser. Lecture Notes in Computer Science), vol. 7965, F. Fomin, R. Freivalds, M. Kwiatkowska, and D. Peleg, Eds. Berlin, Germany: Springer-Verlag, 2013, pp. 564–575. [Online]. Available: http://dx.doi.org/10.1007/978-3-642-39206-1_48

[9] D. Donoho, A. Maleki, and A. Montanari, "Message-passing algorithms for compressed sensing," *Proc. Nat. Acad. Sci.*, vol. 106, no. 45, pp. 18914–18919, 2009.

[10] M. Crouse, R. Nowak, and R. Baraniuk, "Wavelet-based statistical signal processing using hidden Markov models," *IEEE Trans. Signal Process.*, vol. 46, no. 4, pp. 886–902, Apr. 1998.
[11] J. Romberg, H. Choi, and R. Baraniuk, "Bayesian wavelet-domain image modeling using hidden Markov trees," in *Proc. Int. Conf. Image Process.*, Oct. 1999, vol. 1, pp. 158–162.
[12] M. Duarte and Y. Eldar, "Structured compressed sensing: From theory to applications," *IEEE Trans. Signal Process.*, vol. 59, no. 9, pp. 4053–4085, Sep. 2011.
[13] J. Haupt, W. Bajwa, G. Raz, and R. Nowak, "Toeplitz compressed sensing matrices with applications to sparse channel estimation," *IEEE Trans. Inf. Theory*, vol. 56, no. 11, pp. 5862–5875, Nov. 2010.
[14] Y. Tsaig and D. Donoho, "Extensions of compressed sensing," *Signal Process.*, vol. 86, no. 3, pp. 549–571, 2006. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0165168405002215
[15] H. Lee, H. Oh, S. Lee, and A. Bovik, "Visually weighted compressive sensing: Measurement and reconstruction," *IEEE Trans. Image Process.*, vol. 22, no. 4, pp. 1444–1455, Apr. 2013.
[16] T. Do, L. Gan, N. Nguyen, and T. Tran, "Fast and efficient compressive sensing using structurally random matrices," *IEEE Trans. Signal Process.*, vol. 60, no. 1, pp. 139–154, Jan. 2012.
[17] M. Zhou *et al.*, "Nonparametric Bayesian dictionary learning for analysis of noisy and incomplete images," *IEEE Trans. Image Process.*, vol. 21, no. 1, pp. 130–144, Jan. 2012.
[18] L. Gan, "Block compressed sensing of natural images," in *Proc. 15th Int. Conf. Digit. Signal Process.*, Jul. 2007, pp. 403–406.
[19] R. Jenatton, J. Mairal, G. Obozinski, and F. Bach, "Proximal methods for hierarchical sparse coding," *J. Mach. Learn. Res.*, vol. 12, pp. 2297–2334, Jul. 2011. [Online]. Available: http://dl.acm.org/citation.cfm?id=1953048.2021074
[20] D. Baron, S. Sarvotham, and R. Baraniuk, "Bayesian compressive sensing via belief propagation," *IEEE Trans. Signal Process.*, vol. 58, no. 1, pp. 269–280, Jan. 2010.
[21] D. Baron, M. Duarte, S. Sarvotham, M. Wakin, and R. Baraniuk, "An information-theoretic approach to distributed compressed sensing," in *Proc. 43rd Allerton Conf. Commun., Control, Comput.*, 2005, pp. 814–825.
[22] L. He and L. Carin, "Exploiting structure in wavelet-based Bayesian compressive sensing," *IEEE Trans. Signal Process.*, vol. 57, no. 9, pp. 3488–3497, Sep. 2009.
[23] M. Bayati and A. Montanari, "The dynamics of message passing on dense graphs, with applications to compressed sensing," *IEEE Trans. Inf. Theory*, vol. 57, no. 2, pp. 764–785, Feb. 2011.
[24] S. Rangan, P. Schniter, E. Riegler, A. Fletcher, and V. Cevher, "Fixed points of generalized approximate message passing with arbitrary matrices," in *Proc. IEEE Int. Symp. Inf. Theory*, Jul. 2013, pp. 664–668.
[25] Y. Oike and A. El Gamal, "CMOS image sensor with per-column $\Sigma\Delta$ ADC and programmable compressed sensing," *IEEE J. Solid-State Circuits*, vol. 48, no. 1, pp. 318–328, Jan. 2013.
[26] E. Candès and J. Romberg, "Sparsity and incoherence in compressive sampling," *Inverse Problems*, vol. 23, no. 3, pp. 969–985, 2007.
[27] G.-S. Cheon and B. L. Shader, "Sparse orthogonal matrices and the haar wavelet," *Discrete Appl. Math.*, vol. 101, nos. 1–3, pp. 63–76, 2000. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0166218X99001821

Behzad Shahrasbi (S'07–M'16) received the Bachelor's degree from the Amirkabir University of Technology, Tehran, Iran, in 2006, the Master's degree from Oklahoma State University, Stillwater, OK, USA, in 2011, and the Ph.D. degree from the University of Central Florida, Orlando, FL, USA, in 2015. His research interests include sparse signal representations, compressed sensing and recovery algorithms, low-rank matrix recovery, and approximation.

Nazanin Rahnavard (S'97–M'10) received the Ph.D. degree from the School of Electrical and Computer Engineering, Georgia Institute of Technology, Atlanta, GA, USA, in 2007. She is currently an Associate Professor in the Department of Electrical and Computer Engineering at the University of Central Florida, Orlando, FL, USA. Her interest and expertise include a variety of research topics in the communications, networking, and signal processing areas. She received the NSF CAREER Award in 2011. She serves on the editorial board of the Elsevier *Journal on Computer Networks* and on the Technical Program Committee of several prestigious international conferences.