

Multivariate Compressive Sensing for Image Reconstruction in the Wavelet Domain: Using Scale Mixture Models

Jiao Wu, *Student Member, IEEE*, Fang Liu, *Senior Member, IEEE*, L. C. Jiao, *Senior Member, IEEE*, Xiaodong Wang, and Biao Hou, *Member, IEEE*

Abstract—Most wavelet-based reconstruction methods of compressive sensing (CS) are developed under the independence assumption of the wavelet coefficients. However, the wavelet coefficients of images have significant statistical dependencies. Lots of multivariate prior models for the wavelet coefficients of images have been proposed and successfully applied to the image estimation problems. In this paper, the statistical structures of the wavelet coefficients are considered for CS reconstruction of images that are sparse or compressive in wavelet domain. A multivariate pursuit algorithm (MPA) based on the multivariate models is developed. Several multivariate scale mixture models are used as the prior distributions of MPA. Our method reconstructs the images by means of modeling the statistical dependencies of the wavelet coefficients in a neighborhood. The proposed algorithm based on these scale mixture models provides superior performance compared with many state-of-the-art compressive sensing reconstruction algorithms.

Index Terms—Compressive sensing, multivariate model, scale mixture model, wavelet transform.

I. INTRODUCTION

ONE of the stable statistical structures inherent in real-world images is sparsity after some orthogonal transforms, with the wavelet transform [1] being an important example. This important consequence of sparsity has led to compressive sensing (CS) [2], [3], which generates a tremendous amount of excitement in the signal processing community. The theory of CS demonstrates that, under the sparse transform, images can be sampled significantly below

Manuscript received March 14, 2010; revised August 29, 2010 and December 23, 2010; accepted April 19, 2011. Date of publication May 05, 2011; date of current version November 18, 2011. This work was supported in part by the National Natural Science Foundation of China (Grant No. 60971112, 60971128, 60970067, 61072106, and 61072108), in part by the Fund for Foreign Scholars in University Research and Teaching Programs (the 111 Project) (Grant No. B07048), and in part by the Fundamental Research Funds for the Central Universities (Grant No. JY10000902001, JY10000902045, and K50510020001). The associate editor coordinating the review of this manuscript and approving it for publication was Margaret Cheney.

J. Wu, F. Liu, and X. Wang are with the School of Computer Science and Technology, and also with the Key Laboratory of Intelligent Perception and Image Understanding of Ministry of Education of China, Xidian University, Xi'an 710071, China (e-mail: wu_jiao@yahoo.cn; lf204310@163.com; xdwang@mail.xidian.edu.cn).

L. C. Jiao and B. Hou are with the Key Laboratory of Intelligent Perception and Image Understanding of Ministry of Education of China, Xidian University, Xi'an 710071, China (e-mail: jlc1023@163.com; avodec@163.com).

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/TIP.2011.2150231

the Nyquist/Shannon limit and still be reconstructed to satisfaction by solving the following l_1 -regularized optimization problem

$$\min \|\mathbf{x}\|_1 \quad \text{s.t. } \Phi \mathbf{x} = \mathbf{y} \quad (1)$$

where \mathbf{x} is an M -dimension transform coefficient vector of an image \mathbf{f} under a certain $M \times M$ orthogonal basis Ψ , that is $\mathbf{x} = \Psi \mathbf{f}$. Φ is a $K \times M$ sensing matrix, \mathbf{y} is a K -dimension measurement vector with $K \ll M$. Several recent algorithms [4]–[8] based on a multiscale CS method [9] in the wavelet domain are employed to the image reconstruction from CS by solving the aforementioned l_1 inversion.

This l_1 inversion may be viewed as a maximum a posteriori (MAP) estimate for \mathbf{x} under the assumption that each component of \mathbf{x} is drawn i.i.d. from a Laplace prior [8]. Many CS reconstruction algorithms are performed under this independency assumption of the wavelet coefficients. Although the orthonormal wavelet was found to be fairly well-decorrelated, a residual dependency structure always remains among the wavelet coefficients of images [10]. Recently, some new multivariate models and methods have been developed to characterize the statistical dependency among the wavelet coefficients. Examples of such techniques include the bivariate model [11], the wavelet domain hidden Markov tree (HMT) model [12], the multivariate scale mixture of Gaussians (MGSM) models [13], [14], as well as the multivariate model based on the elliptically contoured distribution family (ECDF) [15]. These models have been successfully applied to image compression and denoising problems. More recently, the tree structure and HMT model of the wavelet coefficients have been applied to CS reconstruction problem [16]–[18], and have provided superior performance compared to traditional CS algorithms that do not employ the structure of the wavelet coefficients.

Inspired by these applications in image processing, the multivariate prior models are considered for CS reconstruction of images in this paper. A multivariate CS sampling scheme is proposed first, and then a multivariate pursuit algorithm (MPA) based on the multivariate prior models is developed for CS reconstruction. The proposed algorithm can be viewed as a multivariate extension of Bayesian pursuit algorithm (BPA) based on the univariate model, which has been discussed in our preliminary work [19]. MPA, which estimates the MAP of the wavelet coefficients of images, is of a simple iterative form and only uses

the derivative information of a prior model. When a new multivariate prior model is available, it can be easily adopted for CS reconstruction of images with the proposed algorithm. Several multivariate scale mixture models are utilized as cases to test the influence on the CS reconstruction performance of the proposed algorithm and the multivariate priors. These multivariate scale mixture models include a multivariate Laplacian (ML) distribution [20], a multivariate K (MK) distribution [21], a multivariate normal inverse Gaussian (MNIG) distribution [22], and a multivariate t -type (MTT) distribution [23]. The ML distribution is a special case of the MK distribution. ML, MK, and MNIG distributions are the MGSM using, respectively, the exponential, the Gamma, and the inverse Gaussian distributions as priors for the scale factor. MTT distribution is a scale mixture of the multivariate Kotz-type (MKT) distribution using an inverse generalized gamma distribution as prior for the scale factor.

These multivariate scale mixture models have many very attractive and useful properties, for example, they can be utilized to model the statistical dependency of the wavelet coefficients of images, because they belong to ECDF [24] and include the longer tailed distributions. These models also have flexibility at the same time. By adjusting the parameters of them, we can capture the behavior of a large number of important multivariate non-Gaussian distributions, e.g., the gamma distribution, the Rayleigh distribution, the exponential distribution, the Student's t distribution, the Student's power t distribution [25], the generalized t distributions family [26], and the multivariate generalized t distributions family [27] and the MKT distribution [28], to name but a few. Some of them have been applied to image processing [29], [30].

The reconstruction performance of the proposed multivariate algorithm is tested by CS reconstruction experiments. Our method reconstructs the images by means of modeling the multivariate probability distribution of the wavelet coefficients in a neighborhood. We find that the proposed algorithms based on these multivariate models are computationally tractable and provide superior performance compared with many state-of-the-art compressive sensing reconstruction algorithms. In addition, we discuss the problem of the reconstructed image quality assessment, and utilize a state-of-the-art image quality assessment method [31] to assess the reconstructed images. Finally, we give an example of noise image reconstruction from CS. The results demonstrate that better reconstruction can be obtained by the proposed algorithm and noise can be removed to some degree.

The outline of the rest of the paper is as follows. Section II gives the multivariate CS sampling scheme and the derivation of MPA. MPA based on several multivariate scale mixture models are given in Section III. Section IV reports the experimental results, with comparisons to many of the state-of-the-art CS algorithms, and provides a sufficient discussion. Conclusions are discussed in Section V.

II. COMPRESSIVE SENSING BASED ON MULTIVARIATE PRIOR MODEL

MPA to be derived is a multivariate extension of BPA [19], which use the univariate model in CS recovery. It is different from the univariate CS; a multivariate sampling scheme will be

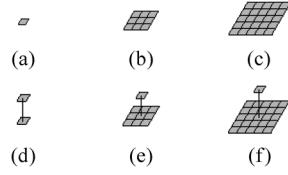


Fig. 1. Neighborhood structures of wavelet coefficients. The top row shows the intra-scale neighborhoods: (a) 1×1 ; (b) 3×3 ; (c) 5×5 . The bottom row shows the inter-scale neighborhoods: (d) $1 + 1 \times 1$; (e) $1 + 3 \times 3$; (f) $1 + 5 \times 5$.

developed to facilitate taking advantage of the statistical dependency among the wavelet coefficients of images in CS recovery. The main procedure of MPA is the same as that of BPA; the key modifications are that we replace estimating one wavelet coefficient in sub-problems of BPA with estimating a small neighborhood of the wavelet coefficients, and utilize the multivariate prior models to depict the statistical dependency of the coefficients in neighborhood. After integrating the dependency structure in CS recovery, more accurate recovery can be achieved using the same number of measurements. The multivariate compressive sampling scheme is discussed in the next subsection. Section II-B focuses on the derivation of MPA.

A. Multivariate Compressive Sampling

The traditional CS samples the measurements in the way that we rearrange the coefficients of images to a vector first, then projects it with sensing matrix. Each component of this vector only contains one coefficient, and is independent of others. Here we rearrange the coefficients according to the partitioned neighborhoods.

Two type wavelet neighborhood structures, the intra- and inter-scale neighborhoods, can be utilized. The intra-scale neighborhood consists of the coefficients in a small window with the size $d \times d$ at the same scale [e.g., Fig. 1(b) and (c)]. The inter-scale neighborhood consists of the coefficients in a small window together with a parent coefficient that is at the same location and orientation as the central one in the small window but at the next scale [e.g., Fig. 1(d)–(f)]. Some multivariate probability models, that depict the statistical dependency of the coefficients in these neighborhoods, have been applied to Bayesian image estimation problems [14], [15], [32]. The intra-scale neighborhood structures are adopted in this work.

Assume that the wavelet coefficients have been partitioned into M' small neighborhoods with the size of $d \times d$. Let $q = d \times d$, $\mathbf{x}_i = (x_{i1}, \dots, x_{iq})^T$ ($i = 1, \dots, M'$) denotes the i th neighborhood. Then the wavelet coefficients can be represented as an $M' \times q$ coefficients matrix $\mathbf{X} = (\mathbf{x}_1, \dots, \mathbf{x}_{M'})^T$, where $M' = \lceil M/q \rceil$, M is the total number of the wavelet coefficients, and $\lceil \cdot \rceil$ is the ceiling function. Note that the case $q = 1$ corresponds to the coefficients vector of traditional compressive sampling, which is depicted by the univariate models. Given the $K \times M'$ sensing matrix Φ ($K \ll M'$), the multivariate measurements are taken as $\mathbf{Y} = \Phi \mathbf{X}$, where \mathbf{Y} is a $K \times q$ measurements matrix. The number of measurements is $K \times q$.

B. Multivariate Pursuit Algorithm

1) *Procedure of MPA:* As the multivariate extension of BPA [19], MPA has the same framework with BPA. The wavelet co-

TABLE I
PROCEDURE OF MPA

Input: Sensing matrix Φ , measurements matrix \mathbf{Y} , prior models of \mathbf{X}^n ($n=0,1,\dots,N$).
Output: Approximation $\tilde{\mathbf{X}}$ to true coefficients matrix \mathbf{X} .
Step 1: $n=0$, initialize $\tilde{\mathbf{X}}=0$, $\mathbf{R}\mathbf{Y}^0=\mathbf{Y}$.
Step 2: Let $\tilde{\mathbf{X}}^n=0$, $\mathbf{R}^0=\mathbf{R}\mathbf{Y}^n$. Compute the residual correlations $\mathbf{U}=\Phi^T\mathbf{R}^0$.
Step 3: Determine the index set $S=\{s_1,\dots,s_{M'}\}$, such that $\left \sum_{j=1}^q u_{s_j j}\right > \dots > \left \sum_{j=1}^q u_{s_{M'} j}\right $, $s_i \in \{1,\dots,M'\}$.
Step 4: According to S , estimate the components of \mathbf{X}^n , successively.
Step 4.1: $i=1$. Let $\mathbf{R}^i=\mathbf{R}^0$.
Step 4.2: Estimate $\tilde{\mathbf{x}}_{s_i}^n$ by solving the sub-problem $\mathbf{R}^i = \phi_{s_i} \mathbf{x}_{s_i}^{nT} + \epsilon^i$.
Step 4.3: Update the residual $\mathbf{R}^{i+1} = \mathbf{R}^i - \phi_{s_i} \tilde{\mathbf{x}}_{s_i}^{nT}$.
Step 4.4: $i=i+1$. If $i \leq M'$, go to Step 4.2.
Step 5: Update $\tilde{\mathbf{X}} = \tilde{\mathbf{X}} + \tilde{\mathbf{X}}^n$, $\mathbf{R}\mathbf{Y}^{n+1} = \mathbf{R}^{M'}$.
Step 6: $n=n+1$. If $\frac{1}{K \cdot q} \sum_{k=1}^K \sum_{j=1}^q \mathbf{R}\mathbf{Y}_{k,j}^{n+1} ^2 < \varepsilon_0$, or $n=N+1$, stop; else, go to Step 2.

coefficients \mathbf{X} (with the size of $M' \times q$) is viewed as the sum of a main part \mathbf{X}^0 and some residual parts $\mathbf{X}^1, \mathbf{X}^2, \dots$, that is $\mathbf{X} = \sum_{n=0}^{\infty} \mathbf{X}^n$. The main energy of wavelet coefficients is approximated by estimating the main part in the first iteration, then the residual energy is approximated by iteratively estimating the residual parts. After $N+1$ iterations, the CS reconstruction of \mathbf{X} is approximated by $\sum_{n=0}^N \tilde{\mathbf{X}}^n$, where $\tilde{\mathbf{X}}^n$ is the approximation of \mathbf{X}^n ($n = 0, 1, \dots, N$).

The procedure of MPA is given in Table I. Starting with the initial residual $\mathbf{R}\mathbf{Y}^0 = \mathbf{Y}$, MPA performs the reconstructions for the main part and residual parts $N+1$ times. In the $(n+1)$ th iteration ($n = 0, 1, \dots, N$), let $\mathbf{R}^0 = \mathbf{R}\mathbf{Y}^n$, according to a descending sort order of the residual correlations, we successively estimate the MAP of the components of \mathbf{X}^n by solving a sequence of sub-problems. The i th sub-problem is $\mathbf{R}^i = \phi_{s_i} \mathbf{x}_{s_i}^{nT} + \epsilon^i$, where \mathbf{R}^i is the current residual, ϕ_{s_i} is the s_i th column of Φ , $\mathbf{x}_{s_i}^n$ is the s_i th component of \mathbf{X}^n , ϵ^i is the noise vector. After updating the residual to $\mathbf{R}\mathbf{Y}^{n+1}$, we check the halting criterion [$n = N+1$, or $(1/(K \cdot q)) \sum_{k=1}^K \sum_{j=1}^q |\mathbf{R}\mathbf{Y}_{k,j}^{n+1}|^2 < \varepsilon_0$, ε_0 is the positive constant]; if it is not yet time to stop, we set $n = n+1$ and go to the next reconstruction of residual part. If it is time to stop, $\tilde{\mathbf{X}}$ is the final approximation of the reconstruction of \mathbf{X} .

2) *Prior Assumptions and Solving the Sub-Problems of MPA:* Solving the sub-problems in Step 4.2 is based on the following assumptions. Assume that \mathbf{R}^i has a matrix normal distribution $MN_{K,q}(\phi_{s_i} \mathbf{x}_{s_i}^{nT}, \sigma_0^2 \mathbf{I}_q)$ [33], its probability density function (pdf) has the form

$$P(\mathbf{R}^i | \mathbf{x}_{s_i}^n, \sigma_0^{-2} \mathbf{I}_q) = (2\pi\sigma_0^2)^{-qK/2} \times \exp \left\{ -\frac{1}{2}\sigma_0^{-2} \text{tr} \left[(\mathbf{R}^i - \phi_{s_i} \mathbf{x}_{s_i}^{nT}) (\mathbf{R}^i - \phi_{s_i} \mathbf{x}_{s_i}^{nT})^T \right] \right\}. \quad (2)$$

Assume that the main part has a zero-mean multivariate distribution with covariance matrix equal to $\sigma_m^{-2} \mathbf{I}_q$. This means that the elements of $\mathbf{x}_{s_i}^0$ are uncorrelated but not independent. Its pdf is denoted as

$$P(\mathbf{x}_{s_i}^0) = C_q \cdot h(t) \quad (3)$$

where C_q is the normalizing constant, $t = \sigma_m^{-2} \mathbf{x}_{s_i}^{0T} \mathbf{x}_{s_i}^0$. In Section III, we will discuss several multivariate scale mixture distributions, which are used as the prior of the main part. In addition, the prior of the residual part $\mathbf{x}_{s_i}^n$ ($n = 1, \dots, N$) is assumed to be a zero-mean multivariate Gaussian distribution with covariance matrix equal to $\sigma_r^{-2} \mathbf{I}_q$; its pdf has the form

$$P(\mathbf{x}_{s_i}^n) = (2\pi)^{-q/2} \sigma_r^q \exp \left\{ -\frac{1}{2}\sigma_r^2 \mathbf{x}_{s_i}^{nT} \mathbf{x}_{s_i}^n \right\}. \quad (4)$$

In order to concisely describe the solving process of the sub-problems, the superscripts and subscripts corresponding to the main part and residual parts are neglected in the following context. We denote \mathbf{x} as a component of the main part \mathbf{X}^0 or the residual part \mathbf{X}^n at certain iteration. \mathbf{R} is the corresponding residual matrix. ϕ is a column vector of Φ corresponding to \mathbf{x} .

Given the aforementioned assumptions, the classical MAP of the main part and residual parts can be estimated by maximizing the log-posterior of \mathbf{x}

$$\begin{aligned} \tilde{\mathbf{x}}_{\text{MAP}} &= \arg \max_{\mathbf{x}} \{ \log P(\mathbf{x} | \mathbf{R}) \} \\ &= \arg \max_{\mathbf{x}} \{ \log P(\mathbf{R} | \mathbf{x}) + \log P(\mathbf{x}) \}. \end{aligned} \quad (5)$$

This is equivalent to solving the equation

$$\frac{\partial}{\partial \mathbf{x}} [\log P(\mathbf{R} | \mathbf{x}) + \log P(\mathbf{x})] = 0. \quad (6)$$

From (2), it is easy to get

$$\frac{\partial}{\partial \mathbf{x}} \log P(\mathbf{R} | \mathbf{x}) = \sigma_0^{-2} (\mathbf{R}^T \phi - (\phi^T \phi) \mathbf{x}). \quad (7)$$

For the pdf of the main part given in (3), the MAP estimation of \mathbf{x} can be obtained with a simple iterative process. From (3), we have

$$\frac{\partial}{\partial \mathbf{x}} \log P(\mathbf{x}) = 2\sigma_m^{-2} \mathbf{x} \frac{\partial}{\partial t} \log h(t). \quad (8)$$

Substituting (7) and (8) into (6) gives

$$\mathbf{x} = \left(\phi^T \phi - 2\frac{\sigma_0^2}{\sigma_m^2} \frac{\partial}{\partial t} \log h(t) \right)^{-1} \mathbf{R}^T \phi. \quad (9)$$

The MAP estimator $\tilde{\mathbf{x}}_{\text{MAP}}$ can be obtained via the following successive substitution method:

- 1) Set $l = 0$ and $\mathbf{x}^{[l]} = (\phi^T \phi)^{-1} \mathbf{R}^T \phi$.
- 2) Compute $t^{[l]}$ by substituting $\mathbf{x}^{[l]}$ into $t = \sigma_m^{-2} \mathbf{x}^T \mathbf{x}$.
- 3) Set $l = l + 1$ and compute $\mathbf{x}^{[l+1]}$ by substituting $t^{[l]}$ into (9).
- 4) Find the difference $e = \|\mathbf{x}^{[l+1]} - \mathbf{x}^{[l]}\|/q$.
- 5) If e is small enough or $l \leq L$ (L is the maximum number of iterations. In our experiments, we set $L = 10$, $e = 1e-3$), then terminate the iteration. Otherwise, go to (2).

For the Gaussian prior of the residual parts, a closed form solution of MAP estimation can be obtained. From (4), we have

$$\frac{\partial}{\partial \mathbf{x}} \log P(\mathbf{x}) = -\sigma_r^{-2} \mathbf{x}. \quad (10)$$

Substituting (7) and (10) into (6) gives

$$\tilde{\mathbf{x}} = \left(\phi^T \phi + 2\sigma_0^2 \sigma_r^{-2} \right)^{-1} \mathbf{R}^T \phi. \quad (11)$$

III. MPA USING SCALE MIXTURE MODELS: EXAMPLES

As mentioned above, most existing multivariate scale mixture models are versatile tools for modeling heavy-tailed random processes. In this section, MAP estimation of the main part of MPA will be derived using several different multivariate scale mixture priors.

A. Multivariate K Distribution

The MK distribution is an MGSM model with a Gamma distributed scale factor [21]. $\text{MK}_q(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \alpha, \lambda)$ denotes a q -dimension MK distribution with parameters $\boldsymbol{\mu}$, $\boldsymbol{\Sigma}$, α , and λ . Note that $\alpha = 0$ gives a multivariate Laplace (ML) distribution [20], which is an MGSM using an exponential prior for the scale factor. The MK and ML prior have advantages in the sense that the model becomes easy to analyze, and a closed form solution can be obtained.

$\text{MK}_q(\mathbf{0}, \sigma_m^2 \mathbf{I}_q, \alpha, 1)$ used here has the following pdf:

$$\begin{aligned} P(\mathbf{x}) &= \frac{2\sigma_m^{-q}}{(2\pi)^{q/2}\Gamma(\alpha+1)} \left(\sqrt{\frac{1}{2}t} \right)^{-q/2+\alpha+1} \\ &\quad \times K_{-q/2+\alpha+1}(\sqrt{2t}) \\ &= C_q \cdot h_{\text{MK}}(t) \end{aligned} \quad (12)$$

where $t = \sigma_m^{-2}\mathbf{x}^T\mathbf{x}$, $K_v(\cdot)$ is the v -order modified Bessel function of the second kind, $C_q = 2\sigma_m^{-q}/(2\pi)^{q/2}\Gamma(\alpha+1)$ is the normalizing constant, and $h_{\text{MK}}(t) = (\sqrt{(1/2)t})^{-q/2+\alpha+1}K_{-q/2+\alpha+1}(\sqrt{2t})$.

Let $r = \sqrt{2t}$, we have

$$h_{\text{MK}}(t) = 2^{q/2-\alpha-1}h_{\text{MK}}(r) \quad (13)$$

where $h_{\text{MK}}(r) = r^{-q/2+\alpha+1}K_{-q/2+\alpha+1}(r)$. By means of the property of Bessel function [34]

$$K_v(x) \sim \sqrt{\frac{\pi}{2x}} \exp(-x) \quad \text{when } |x| \rightarrow \infty \quad (14)$$

we find that

$$h_{\text{MK}}(r) \sim r^{-q/2+\alpha+1} \cdot \sqrt{\frac{\pi}{2}} r^{-1/2} \exp(-r) \quad \text{for large } r \quad (15)$$

thus

$$\begin{aligned} \frac{\partial}{\partial t} \log h_{\text{MK}}(t) &= \frac{1}{r} \frac{\partial}{\partial r} \log h_{\text{MK}}(r) \\ &= \frac{K'_{-q/2+\alpha+1}(r)}{r K_{-q/2+\alpha+1}(r)} - \frac{-q/2+\alpha+1}{r^2} \\ &\approx \frac{-q+2\alpha+1}{2r^2} - \frac{1}{r}. \end{aligned} \quad (16)$$

Substituting (16) into (9) gives

$$\mathbf{x} = \left(\boldsymbol{\phi}^T \boldsymbol{\phi} - \frac{\sigma_0^2}{\sigma_m^2} \left(\frac{-q+2\alpha+1}{r^2} - \frac{2}{r} \right) \right)^{-1} \mathbf{R}^T \boldsymbol{\phi}. \quad (17)$$

Furthermore, a solution of analytical form can be obtained as follows. Let $w = \sqrt{\mathbf{x}^T \mathbf{x}}$, then $r = \sqrt{2\sigma_m^{-2}}w$. Substituting r into (17), it follows that

$$\mathbf{x} = \left(\boldsymbol{\phi}^T \boldsymbol{\phi} - \sigma_0^2 \left(\frac{-q+2\alpha+1}{2w^2} - \frac{\sqrt{2}}{\sigma_m w} \right) \right)^{-1} \mathbf{R}^T \boldsymbol{\phi}. \quad (18)$$

Thus, w can be expressed as

$$\begin{aligned} w &= \sqrt{\mathbf{x}^T \mathbf{x}} \\ &= \left(\boldsymbol{\phi}^T \boldsymbol{\phi} - \sigma_0^2 \left(\frac{-q+2\alpha+1}{2w^2} - \frac{\sqrt{2}}{\sigma_m w} \right) \right)^{-1} \\ &\quad \times \sqrt{\boldsymbol{\phi}^T \mathbf{R} \mathbf{R}^T \boldsymbol{\phi}}. \end{aligned} \quad (19)$$

Substituting (18) into (19) gives the estimation of \mathbf{x}

$$\tilde{\mathbf{x}} = (w \mathbf{R}^T \boldsymbol{\phi}) / \sqrt{\boldsymbol{\phi}^T \mathbf{R} \mathbf{R}^T \boldsymbol{\phi}}. \quad (20)$$

To compute $\tilde{\mathbf{x}}$, we need to compute w first. Equation (19) is a quadratic equation with respect to w . All that remains is the following solution of (19) in experiments (since this solution can lead to much better performance)

$$w = \frac{1}{2}C + \frac{1}{2} \left(C^2 - 2(\boldsymbol{\phi}^T \boldsymbol{\phi})^{-1} \sigma_0^2 (-q+2\alpha+1) \right)^{1/2} \quad (21)$$

$$\text{where } C = (\boldsymbol{\phi}^T \boldsymbol{\phi})^{-1} (\sqrt{\boldsymbol{\phi}^T \mathbf{R} \mathbf{R}^T \boldsymbol{\phi}} - \sqrt{2}\sigma_0^2 \sigma_m^{-1}).$$

B. Multivariate Normal Inverse Gaussian Distribution

The MNIG distribution [22] is a recent MGSM using an inverse Gaussian distribution as prior for the scale factor. It belongs to a more general class of distributions known as the multivariate generalized hyperbolic (MGH) distributions [35]. The symmetric MNIG distribution belongs to the class of ECDF. There has been an increasing interest in such models for financial [36] and signal processing applications [37].

A symmetric MNIG distribution $\text{MNIG}_q(\mathbf{0}, \sigma_m^2 \mathbf{I}_q, \delta, \gamma)$ used here is of the form

$$\begin{aligned} P(\mathbf{x}) &= \frac{\delta^{1/2} \exp(\sqrt{\delta\gamma})}{2^{(q-1)/2} \pi^{(q+1)/2}} \cdot \left(\sqrt{\frac{1}{\gamma}(\delta+t)} \right)^{-(q+1)/2} \\ &\quad \cdot K_{-(q+1)/2}(\sqrt{\gamma(\delta+t)}) \\ &= C_q \cdot h_{\text{MNIG}}(t) \end{aligned} \quad (22)$$

where $t = \sigma_m^{-2}\mathbf{x}^T\mathbf{x}$, $K_v(\cdot)$ is the v -order modified Bessel function of the second kind, $C_q = \delta^{1/2} \exp(\sqrt{\delta\gamma}) / 2^{(q-1)/2} \pi^{(q+1)/2}$ is the normalizing constant, and $h_{\text{MNIG}}(t) = (\sqrt{(1/\gamma)(\delta+t)})^{-(q+1)/2} K_{-(q+1)/2}(\sqrt{\gamma(\delta+t)})$.

Let $r = \sqrt{\gamma(\delta+t)}$, we have

$$h_{\text{MNIG}}(t) = \gamma^{(q+1)/2} h_{\text{MNIG}}(r) \quad (23)$$

where $h_{\text{MNIG}}(r) = r^{-(q+1)/2} K_{-(q+1)/2}(r)$. Using again the property of the Bessel function

$$h_{\text{MNIG}}(r) \sim r^{-(q+1)/2} \cdot \sqrt{\frac{\pi}{2}} r^{-1/2} \exp(-r) \text{ for large } r \quad (24)$$

we have

$$\begin{aligned} \frac{\partial}{\partial t} \log h_{\text{MNIG}}(t) &= \frac{\gamma}{2r} \frac{\partial}{\partial r} \log h_{\text{MNIG}}(r) \\ &= \frac{K'_{-(q+1)/2}(r)}{K_{-(q+1)/2}(r)} \frac{\gamma}{2r} - \frac{(q+1)\gamma}{4r^2} \frac{\gamma}{2r} \\ &\approx -\frac{(q+2)\gamma}{4r^2} - \frac{\gamma}{2r}. \end{aligned} \quad (25)$$

Substituting (25) into (9) gives

$$\mathbf{x} = \left(\boldsymbol{\phi}^T \boldsymbol{\phi} + \frac{\gamma \sigma_0^2}{\sigma_m^2} \left(\frac{q+2}{2r^2} + \frac{1}{r} \right) \right)^{-1} \mathbf{R}^T \boldsymbol{\phi}. \quad (26)$$

Substituting $r = \sqrt{\gamma(\delta+t)}$ into (26), we can obtain

$$\mathbf{x} = \left(\boldsymbol{\phi}^T \boldsymbol{\phi} + \frac{\sigma_0^2}{\sigma_m^2} \left(\frac{q+2}{2(\delta+t)} + \sqrt{\frac{\gamma}{\delta+t}} \right) \right)^{-1} \mathbf{R}^T \boldsymbol{\phi}. \quad (27)$$

C. Multivariate T-Type Distribution

The MTT distribution is a scale mixture of the Kotz-type distribution with an inverse generalized gamma distributed scale factor. As the case of the MGSM distribution, Arslan [23] uses the mixing approach to the symmetric MKT distribution to produce a multivariate scale mixture of the MKT distribution family. Since this family includes the Student's power t , the generalized t , and the generalization of the multivariate t distributions, it is called multivariate t -type (MTT) distribution. $\text{MTT}_q(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \alpha, \beta, \lambda, p, \gamma)$ denotes that a q -dimension random vector \mathbf{x} possesses the MTT distribution with the parameters $\boldsymbol{\mu}$, $\boldsymbol{\Sigma}$, α , β , λ , p , and γ .

The pdf of $\text{MTT}_q(\mathbf{0}, \sigma_m^2 \mathbf{I}_q, 1, 1, 1, p, \gamma)$ used here has the form

$$\begin{aligned} P(\mathbf{x}) &= \frac{p \cdot \Gamma(q/2) \cdot \gamma^\gamma}{\pi^{q/2} \cdot B(\gamma, q/2p) \cdot \sigma_m^q} (\gamma + t^p)^{-(\gamma+q/2p)} \\ &= C_q \cdot h_{\text{MTT}}(t) \end{aligned} \quad (28)$$

where $t = \sigma_m^{-2} \mathbf{x}^T \mathbf{x}$, $C_q = (p \cdot \Gamma(q/2) \cdot \gamma^\gamma) / (\pi^{q/2} \cdot B(\gamma, q/2p) \cdot \sigma_m^q)$, $B(\cdot, \cdot)$ is the beta function, and $h_{\text{MTT}}(t) = (\gamma + t^p)^{-(\gamma+q/2p)}$. Thus, we have

$$\frac{\partial}{\partial t} \log h_{\text{MTT}}(t) = -\frac{(2p\gamma + q)t^{p-1}}{2(\gamma + t^p)}. \quad (29)$$

Substituting (29) to (9) gives

$$\mathbf{x} = \left(\boldsymbol{\phi}^T \boldsymbol{\phi} + \frac{\sigma_0^2}{\sigma_m^2} \frac{(2p\gamma + q)t^{p-1}}{\gamma + t^p} \right)^{-1} \mathbf{R}^T \boldsymbol{\phi}. \quad (30)$$

IV. EXPERIMENTS AND DISCUSSION

The performance of MPA with varying multivariate scale mixture priors will be tested in this section. All the reconstruction results are obtained based on the average of five trials. For each trial, a $K \times M'$ sensing matrix Φ ($K \ll M'$) is generated by randomly selecting K rows from an $M' \times M'$ orthogonal



Fig. 2. Natural images of seven groups: peoples, animals, buildings, landscape, boat, indoor, and plane/truck.

random matrix with i.i.d. draws of a Gaussian distribution $N(0,1)$. We use the 9/7 orthonormal wavelet in the following experiments. The reconstruction performance is evaluated in terms of reconstructed PSNR, CPU time and visually. All tests are run under Windows XP and MATLAB v7.8 (R2009a) on PCs with Pentium Dual Core CPU at 2.60 GHz and 2 GB of memory.

A. Experiments

Experiment 1: This experiment compares the reconstruction of MPA-ML to seven CS reconstruction algorithms based on univariate models to show the performance gain of multivariate algorithm. MPA-ML, a special case of MPA-MK, is used here, because it can obtain a closed form solution and requires designing fewer parameters. The univariate algorithms include: basis pursuit (BP) [4], matching pursuit (MP) [5], stagewise orthogonal matching pursuit (StOMP) [7], Bayesian compressive sensing (BCS) [8], Lasso-modified least angle regression (LARS/LASSO) [38] (from the *Sparselab* package [39]), total variation (TV) [40] (from the l_1 -Magic toolbox [41]), as well as BPA [19].

An image database, which includes 50 natural images (available at <http://decsai.ugr.es/cvg/dbimagenes/index.php>) with size 256×256 , is used for testing. To provide the detail reconstruction results, these images are divided into seven groups: people, animal, building, landscape, boat, indoor, and plane/truck. Each of the first three groups contains ten images. Each of the last four groups contains five images. The images are depicted in Fig. 2.

We here assume the CS measurements consist of two parts. The first part is the scaling coefficients, which are measured directly, and the second is the compressed samples of the wavelet coefficients, which are made by projecting the wavelet coefficients with the sensing matrix Φ . Taking 40% CS measurements rate for example, MPA-ML is implemented with the neighborhood structures 3×3 and 5×5 [Fig. 1(b) and (c)]. The multivariate wavelet coefficients are decomposed into three layers, that is, $\mathbf{X} \approx \sum_{n=0}^2 \mathbf{X}^n$. From the experiments, we find that for the fixed σ_0^2 and σ_r^2 , the small fluctuation in σ_r^2 has almost no effect on the reconstruction accuracy to all images used here. Therefore, for each image, given $\sigma_r^2 = 3$, we try the following values of the pair (σ_0^2, σ_m^2) : $\sigma_0^2, \sigma_m^2 = 1e-2, 5e-3, 1e-3, \dots, 5e-6, 1e-6$. The first and second rows of Fig. 3 show two examples of the average reconstructed PSNR as the functions of σ_0^2 and σ_m^2 for seven image groups in the case of 3×3

TABLE II
COMPARISON AVERAGE PERFORMANCE FOR CS RECONSTRUCTION OF EACH IMAGE GROUP USING VARYING ALGORITHMS IN TERMS OF PSNR AND TIME

Reconstruction Algorithm	People		Animal		Building		Boat		Indoor		Plane/Truck		Landscape	
	PSNR (dB)	Time (s)												
BP	32.85	12.2	29.91	12.7	28.85	13.1	27.76	13.6	31.21	12.4	30.64	11.7	26.30	11.5
MP	32.82	14.7	29.92	14.6	28.90	14.8	27.82	15.0	31.22	14.6	30.67	15.2	26.39	14.6
StOMP(CFAR)	32.15	0.5	29.37	0.5	28.39	0.4	27.19	0.5	30.20	0.6	30.15	0.5	26.10	0.4
TV	32.75	217.8	29.33	225.1	28.62	223.4	27.58	232.9	30.69	219.4	30.08	228.7	25.99	230.9
BCS	32.24	29.3	29.31	32.0	28.24	33.3	27.20	32.5	30.75	26.5	30.06	32.1	25.79	34.3
LARS/LASSO	32.29	10.6	29.43	11.2	28.40	11.0	27.33	11.4	30.73	11.3	30.14	11.6	25.89	11.4
BPA	32.90	0.5	30.04	0.5	28.99	0.5	27.86	0.5	31.28	0.5	30.82	0.5	26.48	0.5
MPA-ML (3×3)	33.81	0.8	30.44	0.9	29.36	0.9	28.46	0.8	31.78	0.8	31.07	0.9	26.95	0.9
MPA-ML (5×5)	33.57	0.3	30.57	0.3	29.38	0.3	28.44	0.3	31.61	0.3	31.21	0.3	26.68	0.3

*The numbers of CS measurements for the univariate and multivariate models are as follows: Univariate algorithm: 26 224; Multivariate algorithm (3 × 3): 26 212; Multivariate algorithm (5 × 5): 26 209.

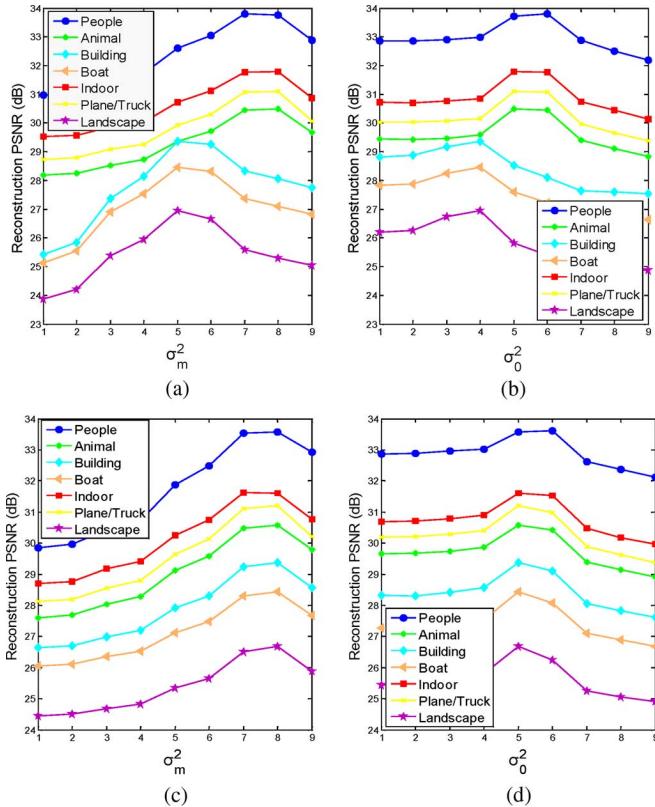


Fig. 3. Changes of reconstructed PSNR as the functions of σ_0^2 and σ_m^2 for seven image groups with 3×3 and 5×5 neighborhood structures. The x-axis from left to right corresponds to the nine value of σ_0^2 and σ_m^2 that are $1e-2, 5e-3, 1e-3, \dots, 5e-6, 1e-6$. (a) Case of 3×3 : (a) shows the changes of PSNR with varying values of σ_m^2 and fixed σ_0^2 ($\sigma_0^2 = 5e-5$ for “people”, “animal”, “indoor”, and “plane/truck”; $\sigma_0^2 = 5e-4$ for “building”, “boat”, and “landscape”). (b) shows the changes of PSNR with varying values of σ_0^2 and fixed σ_m^2 ($\sigma_m^2 = 1e-5$ for “people”, “animal”, “indoor”, and “plane/truck”; $\sigma_m^2 = 1e-4$ for “building”, “boat”, and “landscape”). Case of 5×5 : (c) shows the changes of PSNR with $\sigma_0^2 = 1e-4$ and varying value of σ_m^2 . (d) shows the changes of PSNR with $\sigma_m^2 = 5e-6$ and varying value of σ_0^2 .

and 5×5 neighborhoods, respectively. The left plots are the results with the fixed σ_0^2 and varying values of σ_m^2 . The right plots are the results with the fixed σ_m^2 and varying values of σ_0^2 . Note that in the case of 3×3 , the fixed σ_0^2 and σ_m^2 are taken as $5e-5$

and $1e-5$ for “people”, “animal”, “indoor”, and “plane/truck”, $5e-4$ and $1e-4$ for “building”, “boat”, and “landscape”, respectively; in the case of 5×5 , for each image group, $\sigma_0^2 = 1e-4$ and $\sigma_m^2 = 5e-6$. Fig. 3 demonstrates that MPA-ML produces the best reconstructed performance in the average sense with the fixed value of σ_0^2 and σ_m^2 as mentioned above. In addition, from the reconstructed result of each image, we find that the values of σ_0^2 and σ_m^2 are relatively large for the images with the weak sparsity, and vice versa.

The parameters of the other CS algorithms are designed as follows: For BP, TV algorithms, default parameters (if required as input arguments) are used. For StOMP algorithm, we use CFAR thresholding with the false-alarm rate specified as 0.1. The initial noise variance of BCS is set to 1. For MP algorithm, the maximum number of iterations to perform is set to $6 * K_{CS}$, and for LASSO algorithm, the maximum number of Lars iterations to perform is set to K_{CS} , where K_{CS} is the number of the compressed samples of the wavelet coefficients.

The comparisons of reconstruction performance for the univariate algorithms and MPA-ML are shown in Table II. The reconstructed PSNR and CPU times of all algorithms are the average values based on all images of each group. Comparing the results of the algorithms based on univariate model, we can see that BP, MP, and BPA have better performance than other univariate algorithms. However, BP and MP are far slower than BPA. StOMP has fast reconstructed speed at the cost of reconstruction quality. TV, BCS, and LARS/LASSO all run slowly. As a multivariate extension of BPA, MPA-ML significantly improves CS reconstruction quality of images and remains the high computationally effective. MPA-ML performs about 0.2–0.5 dB better than the univariate algorithm with the best reconstruction performance, and only runs about 0.3–0.9 s. The performance gain means that multivariate models successfully improve the CS reconstruction quality of natural images by means of captured joint statistical information among wavelet coefficients.

Experiment 2: To be fair, we compare the performance of MPA based on the varying multivariate scale mixture priors (MPA-ML, -MK, -MNIG, and -MTT) to two recently developed CS algorithms: the model-based CoSaMP (Model-

TABLE III
COMPARISON PERFORMANCE FOR CS RECONSTRUCTION OF SIX TEST IMAGES USING VARYING ALGORITHMS IN TERMS OF PSNR AND CPU TIME

Reconstruction Algorithm		Lena		House		Peppers		Boat		Barbara		Baboon	
		PSNR (dB)	Time (s)										
Model-CoSaMP		32.16	724.7	33.96	299.2	32.61	424.0	28.84	572.6	29.21	549.3	20.81	693.3
TSWCS		32.58	1285.6	35.37	1251.1	34.26	1217.9	29.57	1269.2	30.45	1288.1	22.44	1382.4
MPA-ML	3×3	34.94	1.0	37.04	1.0	34.06	0.9	31.34	0.9	33.29	1.0	23.54	1.2
	5×5	34.60	0.4	36.86	0.5	33.05	0.4	31.20	0.4	33.26	0.3	24.16	0.4
MPA-MK	3×3	34.96	1.1	37.04	1.0	34.06	1.1	31.32	1.1	33.26	1.1	23.57	1.1
	5×5	34.60	0.4	36.86	0.4	33.01	0.4	31.14	0.4	33.28	0.4	24.12	0.4
MPA-MNIG	3×3	34.91	1.5	36.92	1.4	34.07	1.5	31.35	1.6	33.30	1.5	23.80	1.5
	5×5	34.25	0.6	36.15	0.6	32.39	0.5	30.47	0.7	32.68	0.7	24.03	0.6
MPA-MTT	3×3	34.97	1.6	36.87	1.5	34.10	1.4	31.32	1.6	33.27	1.6	23.54	1.7
	5×5	34.41	0.6	36.68	0.7	32.99	0.5	30.93	0.7	32.91	0.6	24.00	0.5

*The numbers of CS measurements for all algorithms are as follows: Model-CoSaMP: 32 752; TSWCS: 32 764; MPAs(3 × 3): 32 746; MPAs(5 × 5): 32 734.

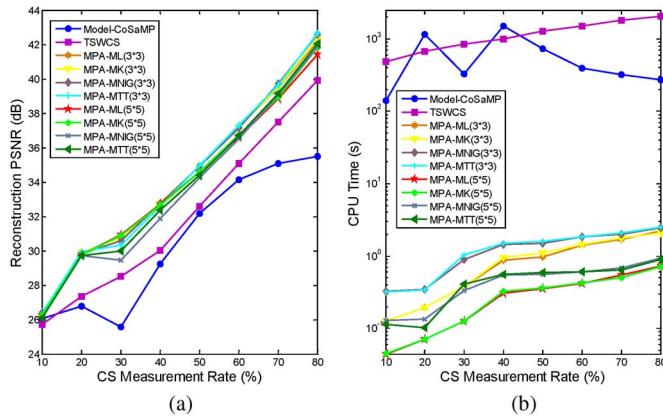


Fig. 4. Performance comparisons for Lena image with varying CS measurement rates. (a) Reconstructed PSNR as a function of CS measurement rate. (b) Associated reconstructed CPU time.

CoSaMP) [18] and the tree-structured wavelet compressive sensing (TSWCS) [17]. These two algorithms use the tree structure in the wavelet coefficients, and provide superior performance compared to traditional CS algorithms. Six standard test images (namely, Lena, Peppers, House, Boat, Barbara, and Baboon) with the size 256 × 256 are used in the following experiment. In the case of 3 × 3 neighborhood, the parameters of MPAs are defined as $\sigma_0^2 = 5e-5$, $\sigma_m^2 = 1e-5$, and $\sigma_r^2 = 3$ for the first four images, and $\sigma_0^2 = 5e-4$, $\sigma_m^2 = 1e-4$, and $\sigma_r^2 = 3$ for Baboon image. In the case of 5 × 5 neighborhood, we take $\sigma_0^2 = 1e-4$, $\sigma_m^2 = 5e-6$, and $\sigma_r^2 = 3$ for each image. For Model-CoSaMP and TSWCS, the default parameters are used.

Taking Lena image as an example, a comparison is performed for the aforementioned algorithms with varying CS measurement rates in Fig. 4. We can see that several MPAs based on different priors produce similar performances in terms of reconstructed PSNR and CPU time. The reconstructed PSNRs of MPAs are significantly better than those of Model-CoSaMP and TSWCS, and MPAs have absolute advantage in speed.

Table III presents the comparison performance of all algorithms for six test images with 50% CS measurement rate. It is clear that all MPAs substantially outperform Model-CoSaMP

and TSWCS. For example, for Lena image, MPA-MTT(3 × 3) is 2.81 dB better than Model-CoSaMP, and 2.39 dB better than TSWCS. For House image, MPA-ML(3 × 3) is 3.08 and 1.67 dB better than Model-CoSaMP and TSWCS. For Peppers image, TSWCS produces the best reconstructed PSNR. The reconstructed PSNR of MPA-MTT(3 × 3) is slightly worse than that of TSWCS. For Boat image, MPA-MNIG(3 × 3) is 2.51 and 1.78 dB better than Model-CoSaMP and TSWCS. For Barbara image, MPA-MNIG(3 × 3) is 4.09 and 2.85 dB better than Model-CoSaMP and TSWCS. For Baboon image, MPA-ML(5 × 5) is 3.35 and 1.72 dB better than Model-CoSaMP and TSWCS.

Comparing the reconstruction results of MPA based on the varying scale mixture priors, we can see that several MPAs are about the same reconstructed accuracies. Relative to MPA-ML, the additional adjustable parameters of MPA-MK, -MNIG, and -MTT make the models more flexible so that they are suitable for different images; however, they increase the complexity of the models at the same time. Considering the convenience of use, MPA-ML is the best choice due to its advantage of simplicity and fewer parameters.

Fig. 5 shows the visual effect comparisons of Table III for Model-CoSaMP, TSWCS, and MPA. For MPAs based on different priors and neighborhoods, we only give the reconstructed images with the best PSNR. From these local enlarged reconstructed images of three algorithms, we can see that MPA produces the reconstruction results with a best visual effect. In addition, the losses of the edge and texture in the reconstructed images with Model-CoSaMP and TSWCS are more serious than that of the MPA. MPA can better preserve the definition and continuity of the edge and texture. It can be seen from the edge of Lena's hat, Boat's mast and Baboon's beard, and the texture of House's surface and Barbara's scarf. TSWCS preserves Peppers' surface better than MPA, so as to produce better PSNR than MPA.

Experiment 3: The following experiment is the CS reconstruction of medical images in the wavelet domain. Medical images have higher sparsity than natural images. An MRI image and a CT image are taken as the examples to test. These two medical images are the size of 256 × 256,



Fig. 5. Comparisons of the reconstructed natural images with 50% CS measurements rate. The images from left to right are original images and reconstructed images with Model-CoSaMP, TSWCS, and MPA, respectively. The PSNR and CPU time are the best result of five trials for each algorithm. (a) Lena. (b) House. (c) Peppers. (d) Boat. (e) Barbara. (f) Baboon.

and shown in the first column of Fig. 6. The methods of CS sampling and reconstruction are the same as Experiment 2. The

total CS measurements rate is taken about 40%. The comparison performance of these two images are shown in Table IV,

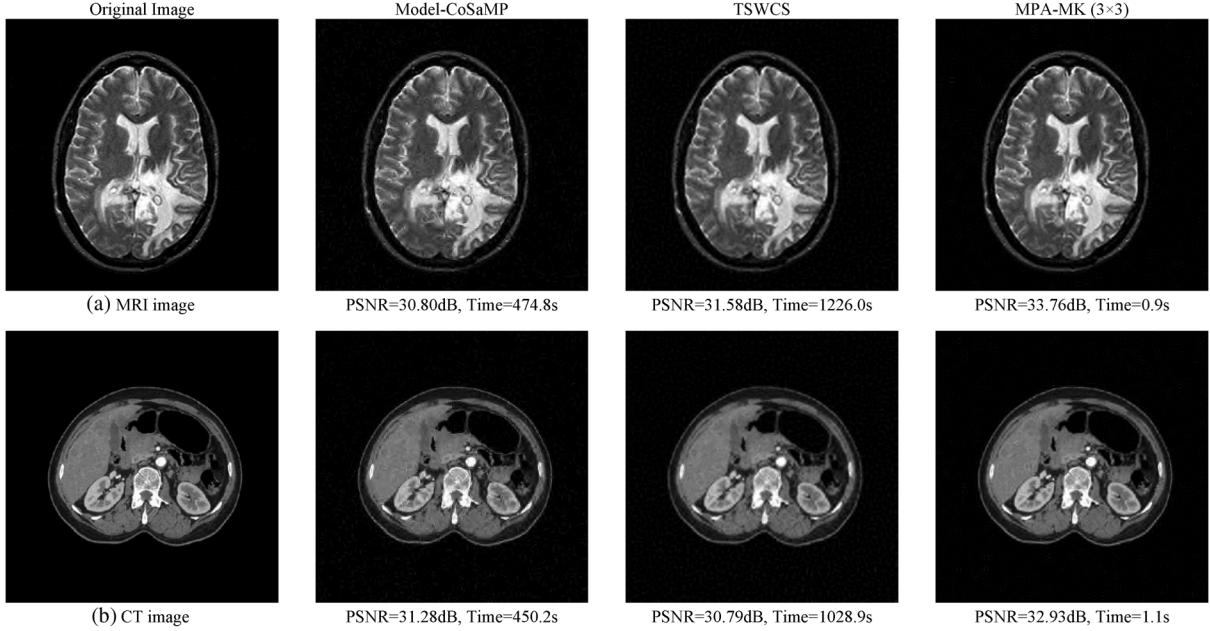


Fig. 6. Comparisons of the reconstructed MRI and CT images with 40% CS measurements rate. The images from left to right are original images and reconstructed images with Model-CoSaMP, TSWCS, and MPA-MK(3×3), respectively. The PSNR and CPU time are the best result of five trials for each algorithm.

TABLE IV
COMPARISON PERFORMANCE FOR CS RECONSTRUCTION OF TWO MEDICAL IMAGES USING VARYING ALGORITHMS IN TERMS OF PSNR AND CPU TIME

Reconstruction Algorithm		MRI image		CT image	
		PSNR (dB)	Time (s)	PSNR (dB)	Time (s)
Model-CoSaMP		30.80	474.8	31.28	450.2
TSWCS		31.53	1132.8	30.66	1039.3
MPA-ML	3×3	33.53	0.9	32.86	0.9
	5×5	33.45	0.4	32.57	0.3
MPA-MK	3×3	33.56	1.0	32.90	1.0
	5×5	33.45	0.3	32.60	0.3
MPA-MNIG	3×3	33.46	1.4	32.70	1.5
	5×5	32.57	1.0	31.85	0.5
MPA-MTT	3×3	33.40	1.3	32.77	1.4
	5×5	32.95	0.6	32.20	0.6

*The numbers of CS measurements for all algorithms are as follows:
Model-CoSaMP: 26 224; TSWCS: 26 212; MPA(3×3): 26 212; MPA(5×5): 26 209.

and associated reconstructed images are shown in Fig. 6. Obviously, MPAs produce better reconstruction performance than Model-CoSaMP and TSWCS.

B. Discussion

As shown in the experiments outlined above, these multivariate algorithms proposed in this paper present excellent performance for CS reconstruction problem of images. Exploring the intra-scale statistical dependencies among wavelet coefficients contributes much to the improvement of the reconstruction performance. Specifically, the edge and texture of the images can be well preserved by the multivariate algorithms. In the above experiments, the size of the test images is 256×256 . The images with a larger size (e.g., 512×512) have also been tested. We find that for the images with the size of 256×256 , the 3×3 neighborhood structure is effective enough to those with higher compressibility in the wavelet domain (e.g. Lena,

House, Peppers, Boat, and medical images), the 5×5 neighborhood structure is suitable for those images with lower compressibility (e.g., Baboon) and larger size.

A worthwhile noticeable problem in CS reconstruction is the reconstructed image quality assessment (IQA). Several conventional IQA methods, e.g., PSNR and MSE, are often utilized to assess the quality of the reconstructed images, but they have not been in agreement with perceived quality measurement widely [42], [43]. For example, from CS reconstruction experiments, we find that a certain algorithm recovers fewer coefficients and produces the bad visual effect, but the reconstructed image has a high PSNR. Therefore, an IQA method proposed in [31] is used as an example to assess the quality of the reconstructed image here. The framework for this IQA method mimics the human visual system (HVS) by incorporating the merits from the multiscale geometric analysis (MGA), contrast sensitivity function (CSF), and the Weber's law of just noticeable difference (JND). Q is defined as the evaluation index of IQA based on MGA. The larger value of Q means the better quality of the reconstructed image. Taking Lena and Peppers images for examples, PSNR and Q of the reconstructed images with 50% CS measurement rate are shown in Table V. For Lena image, MPA-MTT(3×3) performs best both in PSNR and Q. For Peppers image, TSWCS performs 0.16 dB better than MPA-MTT(3×3), but its Q is worse than that of all MPAs.

As mentioned above, many multivariate models have been successfully applied to image denoising problems [13]–[15], [31]. Therefore, we also try an experiment on noise image. By means of the characteristic of multiscale CS method [9], the recovery and denoising are combined in the process of image reconstruction. We implement denoising to the wavelet coefficients, which are kept directly in CS measurements, and recover the wavelet coefficients, which are measured randomly. From these experiments, we find that the reconstructed results based on recovery and denoising are much better than that without

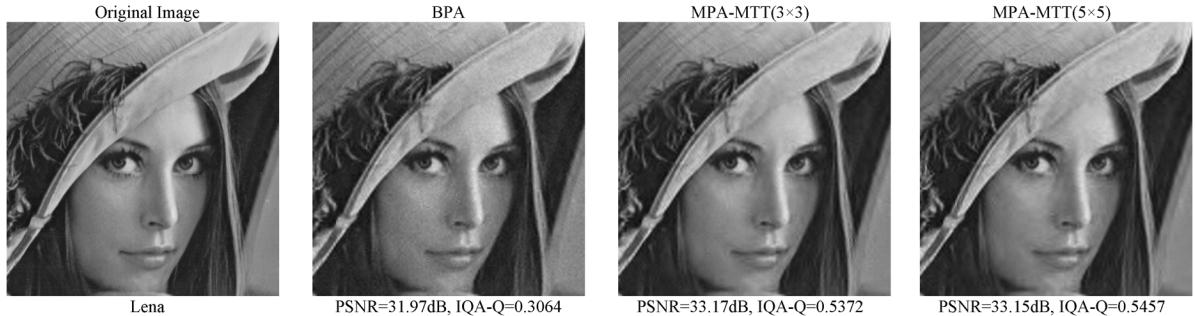


Fig. 7. Local enlarged CS reconstruction images of Lena (with additive Gaussian white noise of standard deviation 10, $\text{PSNR} = 28.14 \text{ dB}$, $\text{IQA} - Q(\alpha = 3) = 0.2930$). The images from left to right are the original Lena image and the reconstructed images with BPA, MPA-MTT(3×3), and MPA-MTT(5×5) from $K = 105\,088$, $104\,956$, and $104\,836$ CS measurements, respectively.

TABLE V
COMPARISON PERFORMANCE FOR CS RECONSTRUCTION OF LENA AND PEPPERS IMAGES USING VARYING ALGORITHMS IN TERMS OF PSNR AND Q

Reconstruction Algorithm		Lena		Peppers	
		PSNR (dB)	IQA-Q ($\alpha=3$)	PSNR (dB)	IQA-Q ($\alpha=3$)
Model-CoSaMP		32.16	0.5213	32.61	0.6460
TSWCS		32.58	0.6189	34.26	0.7111
MPA-ML	3×3	34.94	0.8884	34.06	0.8759
	5×5	34.60	0.8742	33.05	0.8460
MPA-MK	3×3	34.96	0.8916	34.06	0.8731
	5×5	34.60	0.8655	33.01	0.8531
MPA-MNIG	3×3	34.91	0.8754	34.07	0.8589
	5×5	34.25	0.8721	32.39	0.8498
MPA-MTT	3×3	34.97	0.9175	34.10	0.8764
	5×5	34.41	0.8761	32.99	0.8257

* is a tuning parameter of Q corresponding to different types of distortion.
For more details, see [31].

** We only list the best reconstruction results of the multivariate algorithms.

using denoising. This means that making further research on how to combine the other image process method with CS recovery is very significant. Here we only take Lena image as an example to show the performance of MPAs for noise image. The image is the size of 512×512 , and contaminated by additive Gaussian white noise with standard deviation 10. The total CS measurement rate is about 40%. The comparison results of MPA-MTT and BPA are shown in Fig. 7. It is clear that all MPA-MTTs outperform BPA both in PSNR and Q. MPA-MTTs provide better visual effect than BPA at the same time. The reconstruction results for noise image demonstrate that multivariate algorithms can give better reconstruction and remove the noise to some degree.

V. CONCLUSION AND FUTURE WORK

In this paper, CS reconstruction problem of images are discussed from a multivariate point of view. Most conventional wavelet-based CS reconstruction methods assume that the wavelet coefficients are mutual independent. However, significantly statistical dependency exists among the wavelet coefficients of images. A new CS reconstruction algorithm based on the multivariate models is developed. Several scale mixture models are utilized to exploit the intra-scale dependencies of the wavelet coefficients. Multivariate algorithms

proposed here present superior performance compared with some state-of-the-art CS algorithms. They consistently produce superior CS reconstruction quality and have higher computational efficiency. The multivariate algorithms based on inter-scale neighborhood structures are also considered. However, we found that the reconstructed results are a little worse than those due to intra-scale neighborhood structures. This means that the multivariate distributions used in this paper are suitable to model the intra-scale dependency of the wavelet coefficients. More suitable multivariate models to exploit the inter-scale dependency of wavelet coefficients are to be found and applied to CS reconstruction problem in future research. In addition, exploiting the statistical structures of the images in other popular transforms (e.g., curvelet and contourlet) for CS reconstruction of images is also part our research interest.

ACKNOWLEDGMENT

The authors would like to thank the anonymous reviewers for their valuable suggestions, which have helped to greatly improve this paper.

REFERENCES

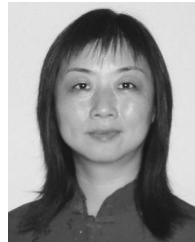
- [1] S. Mallat, *A Wavelet Tour of Signal Processing*, 2nd ed. New York: Academic, 1998.
- [2] D. L. Donoho, "Compressed sensing," *IEEE Trans. Inf. Theory*, vol. 52, no. 4, pp. 1289–1306, Apr. 2006.
- [3] E. Candès, J. Romberg, and T. Tao, "Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information," *IEEE Trans. Inf. Theory*, vol. 52, no. 2, pp. 489–509, Feb. 2006.
- [4] S. Chen, D. L. Donoho, and M. A. Saunders, "Atomic decomposition by basis pursuit," *SIAM J. Sci. Comput.*, vol. 20, no. 1, pp. 33–61, 1999.
- [5] S. Mallat and Z. Zhang, "Matching pursuits with time-frequency dictionaries," *IEEE Trans. Signal Process.*, vol. 41, no. 12, pp. 3397–3415, Dec. 1993.
- [6] J. A. Tropp and A. C. Gilbert, "Signal recovery from random measurements via orthogonal matching pursuit," *IEEE Trans. Inf. Theory*, vol. 53, no. 12, pp. 4655–4666, Dec. 2005.
- [7] D. L. Donoho, Y. Tsaig, I. Drori, and J. C. Starck, Sparse Solution of Underdetermined Linear Equations by Stagewise Orthogonal Matching Pursuit, Stanford Univ., Dept. Statist., Mar. Tech. Rep. 2006-02.
- [8] S. Ji and L. Carin, "Bayesian compressive sensing and projection optimization," in *Proc. 24th Int. Conf. Machine Learning*, 2007, vol. 277, pp. 337–384.
- [9] Y. Tsaig and D. L. Donoho, "Extensions of compressed sensing," *Signal Process.*, vol. 86, no. 3, pp. 549–571, 2006.
- [10] E. P. Simoncelli, "Statistical models for images: Compression, restoration and synthesis," in *Proc. 31st Asilomar Conf. Signals, Systems, Computers*, Pacific Grove, CA, Nov. 1997, vol. 1, pp. 673–678.

- [11] L. Sendur and I. W. Selesnick, "Bivariate shrinkage functions for wavelet-based denoising exploiting interscale dependency," *IEEE Trans. Signal Process.*, vol. 50, no. 11, pp. 2744–2756, Nov. 2002.
- [12] M. S. Crouse, R. D. Nowak, and R. G. Baraniuk, "Wavelet-based statistical signal processing using hidden Markov models," *IEEE Trans. Signal Process.*, vol. 46, no. 4, pp. 886–902, Apr. 1998.
- [13] J. Portilla, V. Strela, M. J. Wainwright, and E. P. Simoncelli, "Image denoising using scale mixtures of Gaussians in the wavelet domain," *IEEE Trans. Image Process.*, vol. 12, no. 11, pp. 1338–1351, Nov. 2003.
- [14] S. Tan and L. C. Jiao, "Multishrinkage: Analytical form for a Bayesian wavelet estimator based on the multivariate Laplacian model," *Opt. Lett.*, vol. 32, no. 17, pp. 2583–2585, 2007.
- [15] S. Tan and L. C. Jiao, "Multivariate statistical models for image denoising in the wavelet domain," *Int. J. Comput. Vision*, vol. 75, no. 2, pp. 209–230, 2007.
- [16] M. F. Duarte, M. B. Wakin, and R. G. Baraniuk, "Wavelet-domain compressive signal reconstruction using a hidden Markov tree model," in *Proc. IEEE Int. Conf. Acoustics, Speech and Signal Processing (ICASSP)*, 2008, pp. 5137–5140, IEEE Press.
- [17] 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.
- [18] R. G. Baraniuk, V. Cevher, M. F. Duarte, and C. Hegde, "Model-based compressive sensing," *IEEE Trans. Inf. Theory*, vol. 56, no. 4, pp. 1982–2001, Apr. 2010.
- [19] J. Wu, F. Liu, L. C. Jiao, and X. Wang, "SAR image reconstruction from compressive sensing based on Bayesian framework and evolutionary computation," *IEEE Trans. Image Process.*, to be published.
- [20] T. Eltoft, T. Kim, and T. Lee, "On the multivariate Laplace distribution," *IEEE Signal Process Lett.*, vol. 13, no. 5, pp. 300–303, May 2006.
- [21] T. Eltoft, T. Kim, and T. Lee, "Multivariate scale mixture of Gaussians modeling," *Independent Component Analysis and Blind Signal Separation*, vol. 3889, Lecture Notes in Computer Science, pp. 799–806, 2006.
- [22] T. A. Øigård, A. Hanssen, R. E. Hansen, and F. Godtliebsen, "EM-estimation and modeling of heavy-tailed processes with the multivariate Normal Inverse Gaussian distribution," *Signal Process.*, vol. 85, no. 8, pp. 1655–1673, 2005.
- [23] O. Arslan, "A new class of multivariate distributions: Scale mixture of Kotz-type distributions," *Statist. Probabil. Lett.*, vol. 75, no. 1, pp. 18–28, 2005.
- [24] K. Fang and Y. Zhang, *Generalized Multivariate Analysis*. Beijing, China: Springer, 1990.
- [25] R. Haro-López and A. F. M. Smith, "On robust Bayesian analysis for location and scale parameters," *J. Multivariate Anal.*, vol. 70, no. 1, pp. 30–56, 1999.
- [26] R. B. Arellano-Valle and H. Bolfarine, "On some characterizations of the t distribution," *Statist. Probabil. Lett.*, vol. 25, no. 1, pp. 79–85, 1995.
- [27] O. Arslan, "Family of multivariate generalized t distributions," *J. Multivariate Anal.*, vol. 89, no. 2, pp. 329–337, 2004.
- [28] S. Kotz, "Multivariate distributions at a cross-road," in *Statistical Distributions in Scientific Work*, G. P. Patil, S. Kotz, and J. K. Ord, Eds. Dordrecht, The Netherlands: D. Reidel, 1975, pp. 247–270.
- [29] J. Huang, "Statistics of natural images and models," Ph.D. dissertation, Div. Appl. Math., Brown Univ., Providence, RI, 2000.
- [30] G. Verdoolaege, S. De Backer, and P. Scheunders, "Multiscale colour texture retrieval using the geodesic distance between multivariate generalized Gaussian models," in *Proc. 15th IEEE Int. Conf. Image Processing, ICIP'08*, San Diego, CA, 2008, pp. 169–172.
- [31] X. Gao, W. Lu, D. Tao, and X. Li, "Image quality assessment based on multiscale geometric analysis," *IEEE Trans. Image Process.*, vol. 18, no. 7, pp. 1409–1423, Jul. 2009.
- [32] S. Tan, L. Jiao, and I. A. Kakadiaris, "Wavelet-based Bayesian image estimation: From marginal and bivariate prior models to multivariate prior models," *IEEE Trans. Image Process.*, vol. 17, no. 4, pp. 469–481, Apr. 2008.
- [33] K. Fang and R. Li, "Bayesian statistical inference on elliptical matrix distributions," *J. Multivariate Anal.*, vol. 70, pp. 66–85, 1990.
- [34] M. Abramowitz and C. A. Stegun, *Handbook of Mathematical Functions With Formulas, Graphs, and Mathematical Tables*. New York: Dover, 1972.
- [35] O. E. Barndorff-Nielsen, "Exponentially decreasing distributions for the logarithm of particle size," *Proc. Roy. Soc. London. Series A, Math. Phys. Sci.*, vol. 353, no. 1674, pp. 401–419, 1977.
- [36] O. E. Barndorff-Nielsen, *Normal Inverse Gaussian Processes and the Modeling of Stock Returns*, Dept. Theoret. Statist., Inst. Math., Univ. Aarhus, Aarhus, Denmark, Research Rep. No. 300, 1995.
- [37] A. Hanssen and T. A. Øigård, "The normal inverse Gaussian distribution as a flexible model for heavy-tailed processes," in *Proc. IEEE-EURASIP Workshop Nonlinear Signal and Image Processing*, Baltimore, MD, Jun. 2001, 5 pp.
- [38] B. Efron, T. Hastie, I. Johnstone, and R. Tibshirani, "Least angle regression," *Ann. Statist.*, vol. 32, no. 2, pp. 407–499, 2004.
- [39] D. L. Donoho, I. Drori, V. C. Stodden, and Y. Tsaig, Sparselab, 2007. [Online]. Available: <http://sparselab.stanford.edu/>.
- [40] M. Lustig, D. Donoho, and J. M. Pauly, "Sparse MRI: The application of compressed sensing for rapid MR imaging," *Magn. Reson. Med.*, vol. 58, no. 6, pp. 1182–1195, 2007.
- [41] E. Candès, J. Romberg, and S. Becker, l1-Magic, 2007. [Online]. Available: <http://www.acm.caltech.edu/l1magic/>.
- [42] I. Avcibas, B. Sankur, and K. Sayood, "Statistical evaluation of image quality measures," *J. Electron. Imag.*, vol. 11, no. 2, pp. 206–213, 2002.
- [43] A. M. Eskicioglu and P. S. Fisher, "Image quality measures and their performance," *IEEE Trans. Commun.*, vol. 43, no. 12, pp. 2959–2965, Dec. 1995.



Jiao Wu (S'09) received the B.S. degree and the M.S. degree in applied mathematics from Shaanxi Normal University, Xi'an, China, in 1999 and 2002, respectively. She is currently working towards the Ph.D. degree in computer application technology at the School of Computer Science and Technology, Xidian University and the Key Laboratory of Intelligent Perception and Image Understanding of Ministry of Education of China, Xi'an, China.

Her research interests include image processing, machine learning, statistics learning theory, and algorithms.



Fang Liu (M'07–SM'07) received the B.S. degree in computer science and technology from Xi'an Jiaotong University, Xi'an, China, in 1984, and the M.S. degree in computer science and technology from Xidian University, Xi'an, in 1995.

Currently, she is a Professor with the School of Computer Science, Xidian University, Xi'an, China. She is the author or coauthor of five books and more than 80 papers in journals and conferences. Her research interests include signal and image processing, synthetic aperture radar image processing, multiscale geometry analysis, learning theory and algorithms, optimization problems, and data mining.



L. C. Jiao (SM'89) received the B.S. degree from Shanghai Jiaotong University, Shanghai, China, in 1982, and the M.S. and Ph.D. degrees from Xi'an Jiaotong University, Xi'an, China, in 1984 and 1990, respectively.

He is currently a Distinguished Professor with the School of Electronic Engineering, Xidian University, Xi'an, China. His research interests include signal and image processing, natural computation, and intelligent information processing. He has led approximately 40 important scientific research projects and published more than ten monographs and 100 papers in international journals and conferences. He is the author of three books: *Theory of Neural Network Systems* (Xi'an, China: Xidian Univ. Press, 1990), *Theory and Application on Nonlinear Transformation Functions* (Xi'an, China: Xidian Univ. Press, 1992), and *Applications and Implementations of Neural Networks* (Xi'an, China: Xidian Univ. Press, 1996). He is the author or coauthor of more than 150 scientific papers.

Prof. Jiao is a member of the IEEE Xi'an Section Executive Committee, and the Chairman of Awards and Recognition Committee and an executive committee member of the Chinese Association of Artificial Intelligence.



Xiaodong Wang received the B.S. degree from Harbin Institute of technology, Harbin, China, in 1998, and the M.S. degree from Inner Mongolia University of Technology, Hohhot, China, in 2007. He is currently working toward the Ph.D. degree in computer application technology at the School of Computer Science and Technology, Xidian University and the Key Laboratory of Intelligent Perception and Image Understanding of Ministry of Education of China, Xi'an, China.

His current research interests include convex optimization and its applications in compressive sensing and wireless networks.



Biao Hou (M'07) received the B.S. and M.S. degrees in mathematics from Northwest University, Xi'an, China, in 1996 and 1999, respectively, and the Ph.D. in circuits and systems from Xidian University, Xi'an, China, in 2003.

Since 2003, he has been with the Key Laboratory of Intelligent Perception and Image Understanding of Ministry of Education of China in Xidian University, where he is currently a Professor.

His research interests include multiscale geometric analysis and SAR image processing.