Journal of Information Hiding and Multimedia Signal Processing ©2017 ISSN 2073-4212
Ubiquitous International Volume **8**, Number **1**, January **2017**

Bayesian Compressive Sensing Based on Sparsing Important Wavelet Coefficients

Yun Wu[1], Hua Shi [1,2,*], Shaoyong Yu[1]

[1] 1College of Computer and Information Engineering
Xiamen University of Technology, Xiamen 361024, China

[2]School of Information Science and Technology
Xiamen University, Xiamen 361005, China
[*]Corresponding author:ywu@xmut.edu.cn

Received May, 2016; revised August, 2016

Abstract. *Bayesian compressive sensing introduces a new framework for compressed sensing, which has important significance. And the effect will be better when the signal being reconstructed is sparser. This paper utilizes the statistical feature of wavelet coefficients to sparse the important coefficients signals with high level because they contain much energy. Additionally, the improved method is applied to other BCS improved algorithms, and the experimental results show the significance of introducing the method.*
Keywords: Bayesian compressive censing; Wavelet coefficients; Sparsing.

1. **Introduction.** The collection process of magnetic resonance imaging is time-consuming. Using JPEG [1] or JPEG2000 [2] to compress will produce some waste because some large transformation coefficients would be threw away in the compressive process while the collection is expensive. The compressive sensing [3-7] proposed recently directly violates the Nyquist criterion, which only need a few samples captured from some random linear projections to reconstruct unknown sparse signal. The application of CS to MRI (CS-MRI) [1-3] greatly reduces the duration of scanning. Traditional Bayesian compressive sensing [8] and the derivative methods [9-12] all use probability distribution functions such as Laplace distribution to model wavelet coefficients after sparsifying transformation. Wavelet coefficients have many features [13], such as spatial frequency and direction selection, spatial clustering of high frequency, the similarity between subband coefficients, the relative between amplitude etc. The core of some Bayesian compressive algorithms such as JPEG2000 exploits the tree structure of wavelet coefficient. So far, Bayesian compressive sensing utilizing wavelet structure has been proposed [14-17]. Scale feature of wavelet coefficients is introduced in [14] and the similarity between subband coefficients is used to model sparse prior in [15-17]. In the paper, we use energy concentration of frequency domain, energy attenuation and the statistics of wavelet coefficients in different levels to improve traditional Bayesian compressive sensing and the derivative algorithms. Seen from visual appearance, the reconstructed effect is improved in details. The objective evaluate criteria data PSNR is improved by 2∼6db. In conclusion, the improved effect of the improved algorithm is obvious. The remainder of the paper is organized as follows: In section two, we introduce the energy distribution features of wavelet coefficient and the coefficient statistics in each wavelet level. In section three, we exploit the features of wavelet coefficient to obtain sparser matrix. In section four, we introduce the improved

Bayesian compressive sensing. In section five, we apply the improved idea to derivative Bayesian compressive sensing. Simulation results presented in section six testify whether the improved effect is from reconstruction image or data. Conclusions and discussions of future work are provided in section seven.

2. **The feature of wavelet coefficient.** Compressive sensing is applied in sparse signal reconstruction, but many natural signals are not sparse in time domain. We can project original signal to sparse basis and obtain sparse signal. Fourier coefficient, wavelet coefficient, total norm of bounded variation function, Gabor coefficient of oscillator signal, Curvelet coefficient of image of discontinuous edges, etc. are all sparse. In addition, we can obtain the sparse representation from redundant dictionary. Natural images include smooth region, vein region and edge information. After wavelet transformation [13] like Fig 1, energy of smooth region concentrates mostly on low frequency subband, low frequency information of smooth region and vein region concentrate mostly on low frequency subband, high frequency of vein region and edge information concentrate mostly on high frequency subband. Since natural images mostly consist of smooth region, so low frequency subbands include most of image energy.

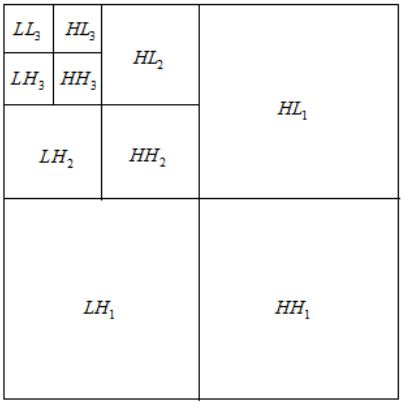

FIGURE 1. Three-level wavelet decomposition of an image

Table 1 shows energy distribution for 256×256 brain MR image in the LL, LH, HL and HH subbands. In conclusion, wavelet energy decreases level by level, energy summation of LL, HL, LH and HH subbands in the third level accounts for the most energy of image and the energy of LL subband is the highest. If subband coefficients of the third level can be reconstructed better and LL subband coefficients can be reconstructed better, then the quality of the reconstructed image will be improved obviously.

Wavelet coefficients can be modeled by two kinds of distributions [18]. θ is wavelet coefficients and ϕ is a variable of wavelet coefficient distribution. Probability density function $P(\theta)$ can be given by:

$$P(\theta)=\int_0^{\infty}P(\theta|\phi)P(\phi)d\phi=\begin{cases}\frac{\sqrt{2\lambda}}{2}\exp\left\{-\sqrt{2\lambda}\,|\theta|\right\}, & \phi \textit{ is a exponential distribution}\\ \frac{v}{2\beta\Gamma(\frac{1}{v})}\exp\left\{-\left(\frac{|\theta|}{\beta}\right)^{v}\right\}, & \phi \textit{ is not a exponential distri}-\\ & -\textit{bution}\end{cases} \tag{1}$$

Where v is peak value of coefficient distribution and β is width. In compressive sensing, if θ is sparser, then the effect of reconstruction is better. Fig 2 shows HL coefficients and

TABLE 1. Energy distribution for 256×256 brain MR image

Subband	Energy ($\times 10^5$)	Subband energy ratio(%)	Level energy ratio(%)
LL3	7852	93.34	97.54
HL3	198	2.36	
LH3	109	1.29	
HH3	46	0.54	
HL2	82	0.98	1.96
LH2	64	0.76	
HH2	19	0.22	
HL1	23	0.27	0.5
LH1	17	0.21	
HH1	2	0.02	

HH coefficients in each level after three-level wavelet decomposition for a brain MR image. From left to right, it is from the high level to the low level. We can conclude that the number of large coefficients in HL subband is bigger than the number in HH subband of the corresponding level.

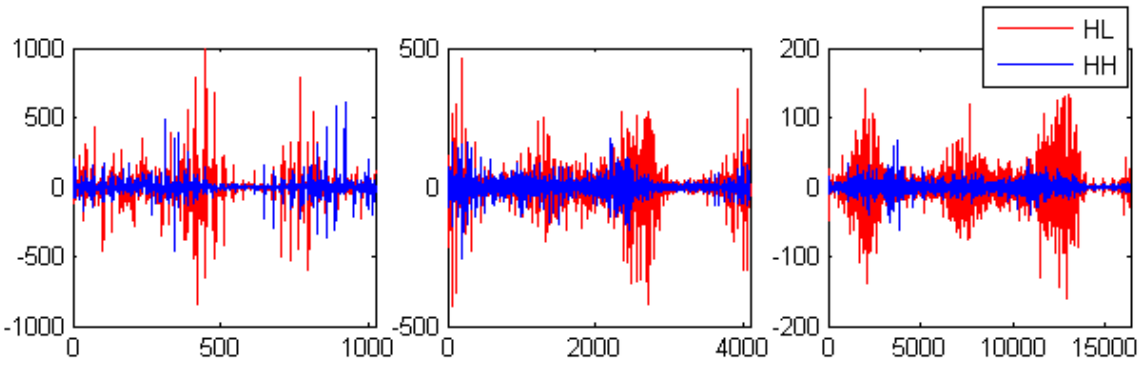

FIGURE 2. HL coefficients vs HH coefficients in every level

3. **Sparse Matrix.** The reconstruction of compressive sensing can be translated into other problem. Suppose the prior probability distribution for θ is a kind of sparse distribution like Laplace distribution and maximize posterior probability to solve l_1 norm problem. Bayesian compressive sensing mostly used sparse prior, therefore if θ is sparser, the effect of prior distribution model is better. Recent Bayesian compressive sensing [8-12,14-16]combined scale feature of wavelet coefficients [14] and tree structure [15,16] to model sparse prior distribution, but energy features of wavelet coefficients are not used. In the reconstruction process, we reconstruct row by row to decrease the number of observation matrix. Exchange HL subband and HH subband in the same level, Fig3 shows the kurtosis changing of the first sixty four row coefficients for 256×256 brain MR image before and after exchange. Kurtosis is a value to judge whether the signal is sparse or not. Kurtosis is bigger, and the signal is sparser. In Fig 3, after exchange, the kurtosis of the first thirty two row coefficients which include LL subband is small, but the kurtosis of the last thirty two row coefficients which include high level subband is bigger than before exchange. So, we can use the regularity to sparsing those important coefficients by changing the distribution of wavelet coefficients.

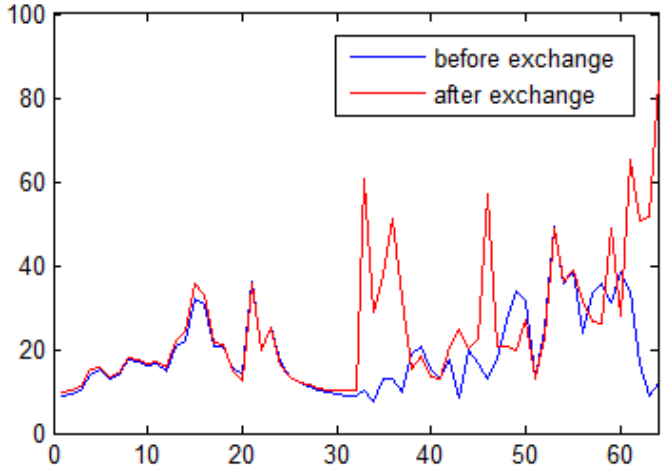

FIGURE 3. The kurtosis before and after exchanging for the first sixty four row wavelet coefficients of 256×256 brain MR image

4. **The improved Bayesian compressive sensing.** From the above sections we can see, after wavelet transformation, the wavelet coefficients of $L \times N$ image are $\theta = (\theta_1, \theta_2, ..., \theta_L)$. Exchange HH subband and HL subband of each level and add white Gaussian noise $e \tilde{} N(0, \sigma^2)$, then the reconstruction problem about compressive sensing can be translated into:

$$Y = (y_1, y_2, ..., y_L) = (\Phi w_1 + e, \Phi w_2 + e, ..., \Phi w_L + e) = \Phi w + e \tag{2}$$

where $\Phi \in \mathbb{R}^{M \times N}$ is objective matrix. This function is similar to a multi-task learning problem [11][19]. Then the Gaussian likelihood modeling is :

$$p(y_i|w_i, \sigma^2) = (2\pi\sigma^2)^{-M_i/2} \exp(-\frac{1}{2\sigma^2} \|y_i - \Phi w_i\|_2^2) \tag{3}$$

where $i = 1, 2, ..L$. Many Bayesian methods can solve function (3). We choose RVM [20, 21] and introduce hierarchy sparse prior model. The first prior can be given by:

$$p(w_i|\alpha_i) = \prod_{j=1}^{N} N(w_{i,j}|0, \alpha_{i,j}^{-1}) \tag{4}$$

where $i = 1, 2, ..L, j = 1, 2, ..N$. α_i is weight, the second prior can be given by:

$$p(\alpha_i|a, b) = \prod_{j=1}^{N} \Gamma(\alpha_{i,j}|a, b) \tag{5}$$

Therefore we can compute the sparse prior of

$$w_i p(w_i|a, b) = \prod_{j=1}^{N} \int_0^{\infty} N(w_{i,j}|0, \alpha_{i,j}^{-1}) \Gamma(\alpha_{i,j}|a, b) d\alpha_{i,j} \tag{6}$$

Gamma prior $\Gamma(\alpha_{i,0}|c, d)$ introduces noise variance $\alpha_{i,0} = 1/\sigma^2$, Gamma prior distribution is a kind of sparse prior distribution. In order to compute conveniently, we set $a = b = c = d = 0$. Suppose that α_i and $\alpha_{i,0}$ have been known, posterior density function of can approximate to a Gaussian distribution, the mean and variance are:

$$\mu_i = \alpha_{i,0} \Sigma_i \Phi^T y_i \tag{7}$$

$$\Sigma_i = \left(\alpha_{i,0}\Phi^T\Phi + A_i\right)^{-1} \tag{8}$$

where $A_i = diag(\alpha_{i,1}, \alpha_{i,2}, ..., \alpha_{i,N})$. In RVM, α_i and $\alpha_{i,0}$ and are determined by maximizing the marginal likelihood. Logarithmic function can be given by:

$$\begin{aligned} L(\alpha_i, \alpha_{i,0}) &= \log p(y_i|\alpha_i, \alpha_{i,0}) = \log \int p(y_i|w_i, \alpha_{i,0}) p(w_i|\alpha_i) dw_i \\ &= -\frac{1}{2}[M \log 2\pi + \log |C_i| + y_i^T C_i^{-1} y_i] \end{aligned} \tag{9}$$

where $C_i = \alpha_{i,0}^{-1} I + \Phi A_i^{-1} \Phi^T$. Maximize $C_i = \alpha_{i,0}^{-1} I + \Phi A_i^{-1} \Phi^T$ to estimate α_i and $\alpha_{i,0}$. To save calculating time, we update single $\alpha_{i,j}$ in each iteration process but not update the whole matrix α_i .The C_i in(7) can be:

$$C_i = \alpha_{i,0}^{-1} I + \sum_{k \neq j} \alpha_{i,k}^{-1} \Phi_k \Phi_k^T + \alpha_{i,j}^{-1} \Phi_j \Phi_j^T = C_{i,-j} + \alpha_{i,j}^{-1} \Phi_j \Phi_j^T \tag{10}$$

where $C_{i,-j}$ C_i is excluding basis function $\Phi_{i,j}$. The inverse matrix and absolute value can be given by:

$$|C_i| = |C_{i,-j}| \left|1 + \alpha_{i,j}^{-1} \Phi_j^T C_{i,-j}^{-1} \Phi_j\right| \tag{11}$$

$$C_i^{-1} = C_{i,-j}^{-1} - \frac{C_{i,-j}^{-1} \Phi_j \Phi_j^T C_{i,-j}^{-1}}{\alpha_{i,j} + \Phi_j^T C_{i,-j}^{-1} \Phi_j} \tag{12}$$

Combine with (7), we can get:

$$\begin{aligned} L(\alpha_i, \alpha_{i,0}) &= -\frac{1}{2}\left[M \log 2\pi + \log |C_{i,-j}| + y_i^T C_{i,-j}^{-1} y_i - \log\left(\frac{\alpha_{i,j}}{\alpha_{i,j} + s_{i,j}}\right) - \frac{q_{i,j}^2}{\alpha_{i,j} + s_{i,j}}\right] \\ &= L(\alpha_{i,-j}, \alpha_{i,0}) + \frac{1}{2}\left[\log\left(\frac{\alpha_{i,j}}{\alpha_{i,j} + s_{i,j}}\right) + \frac{q_{i,j}^2}{\alpha_{i,j} + s_{i,j}}\right] \\ &= L(\alpha_{i,-j}, \alpha_{i,0}) + \ell(\alpha_{i,j}, \alpha_{i,0}) \end{aligned} , \tag{13}$$

where $S_{i,j}$ and $q_{i,j}$ are $s_{i,j} = \Phi_j^T C_{i,-j}^{-1} \Phi_j$, $q_{i,j} = \Phi_j^T C_{i,-j}^{-1} y_i$.

Maximize $L(\alpha_i, \alpha_{i,0})$, differentiate $l(\alpha_i, \alpha_{i,0})$ and set the result to zero, we can get:

$$\alpha_{i,j} = \begin{cases} \dfrac{s_{i,j}^2}{q_{i,j}^2 - s_{i,j}} & if q_{i,j}^2 - s_{i,j} > 0 \\ \infty & otherwise \end{cases} \tag{14}$$

Use (7), (8) and (15) to iterate. We can obtain exchanged wavelet coefficients $w_1, w_2, ..., w_L$ after adding, deleting, updating and get reconstructed image after inverse wavelet transformation and the exchanging of HL subband and LL subband. The whole reconstruction process is shown in algorithm 1:

Algorithm 1：BCS improved algorithm

1. input: Φ， y_i
2. output: θ
3. exchange HL subband coefficients and LL subband coefficients in each level, obtain（$w_1, w_2, ..., w_L$）
4. for $i = (1, 2, ..., L)$
5. for（when convergence criterion is not satisfy）
6. choose a $\alpha_{i,j}$
7. update $\alpha_{i,j}$ using (15)
8. update Σ_i and μ_i using (7)(8)
9. update $s_{i,j}$ and $q_{i,j}$ using (14)
10. end
11. obtain w_i
12. end
13. obtain w , exchange HL subband and LL subband, obtain $(\theta_1, \theta_2, ..., \theta_L)$

5. **The extension of other algorithms.** Since Bayesian framework has been introduced into compressive sensing, many improved Bayesian compressive sensing methods based on traditional compressive sensing are proposed. One improved idea is to make prior probability distribution sparser, such as [9] and [15]. [9] introduces Laplace distribution into hierarchy model and [15] introduces spike-and-slab prior distribution which recently uses Bayesian recession and modeling [22-26]. The two distributions are all peaky distributions. Apply the improved idea above to the derivative algorithms (Laplace, BCSvb), the improved effect is well.

6. **Experiment results.** In this section, we choose three groups of MR images which include ten images in each group with the size128×128, 256×256 , 512×512 , respectively. All softwares are written in MATLAB, and run on PCs with 1.73 GHz CPU and 1 GB memory. Add white Gaussian noise to each image. Fig 4 shows the original images and reconstructed images via BCS and improved BCS. The first and the forth rows show the 256×256 and 512×512 MR images. The second and the fifth rows are reconstructed images via BCS algorithm. The third and the sixth rows are reconstructed images via improved BCS algorithm.

Fig 4 shows reduced MR images, and details are not shown clearly. Fig 5 shows the details compared original images to reconstructed images via BCS and improved BCS. Fig 5(a) is one of the ten 512×512 original MR images in Fig 4. Fig 5(b) and 5(c) is reconstructed image via BCS and improved BCS respectively. White rectangles show the details of MR images. Fig 5 shows that many details are not clear in the image reconstructed via BCS, but the corresponding details in improved BCS image are well reconstructed, which illustrates that the whole reconstruction quality is improved. Except the visual appearance, we also computed PSNR, a value used in estimating image quality.

$$PSNR = 10\log_{10}\left(\frac{255^2}{MSE}\right) \tag{15}$$

$$MSE = \frac{1}{A \times B}\sum_{i=0}^{A-1}\sum_{j=0}^{B-1}\left(x(i,j) - \widehat{x}(i,j)\right)^2 \tag{16}$$

where (i, j) is pixel of $A \times B$ image x. Fig 6 shows the PSNR values of twenty MR images reconstructed by BCS and improved BCS in Fig. 4. We conclude that the PSNR values of improved BCS algorithm increase about 2∼6db, therefore the improved effect

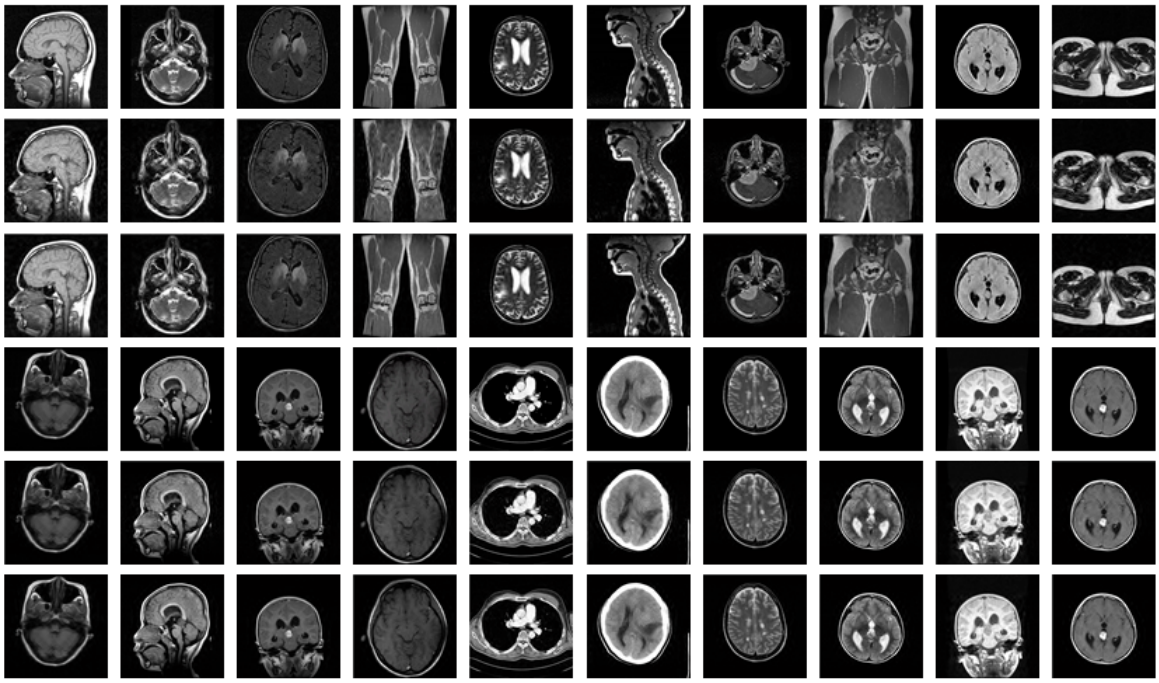

FIGURE 4. Two groups of original images and reconstructed images via BCS and improved BCS

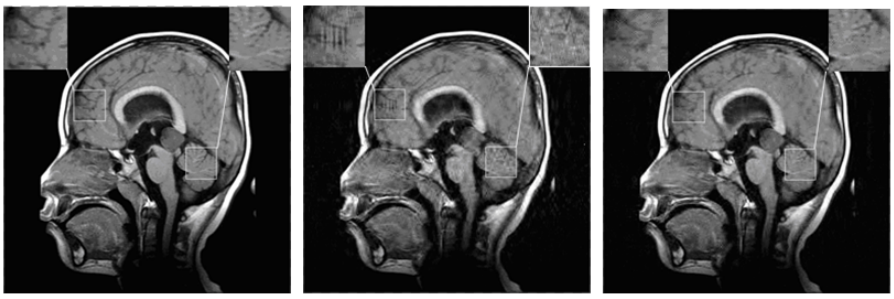

FIGURE 5. The details of reconstructed image via BCS or improved BCS

is well. According to the fifth section, the improved algorithms of extension Bayesian compressive sensing are well. Table 2 is the comparison between three types of Bayesian compressive sensing algorithms and the corresponding improved algorithms. We compute PSNR, $Rerr = \left\|\theta - \hat{\theta}\right\|_2 / \|\theta\|_2$ average values of three groups of ten images with the size of 128×128, 256×256, 512×512. It can be obviously seen that the reconstruction effect is improved via the improved method.

Fig 7 shows the Rerr and PSNR values of reconstructed images via three types of Bayesian compressive sensing algorithms and the corresponding improved algorithms in different objective values. The horizontal axis represents sample ratio, the ratio of objective values and original values. Whether the objective values are more or less, the Rerr of the improved algorithm decreases about 0.1 and PSNR increases about 3db than those of the corresponding Bayesian compressive sensing algorithms. Therefore, our improved effect is obvious.

To illustrate the effect of exploiting different subbands wavelet coefficients to our improved algorithm, we pay attention to LL subband coefficients and the third subband

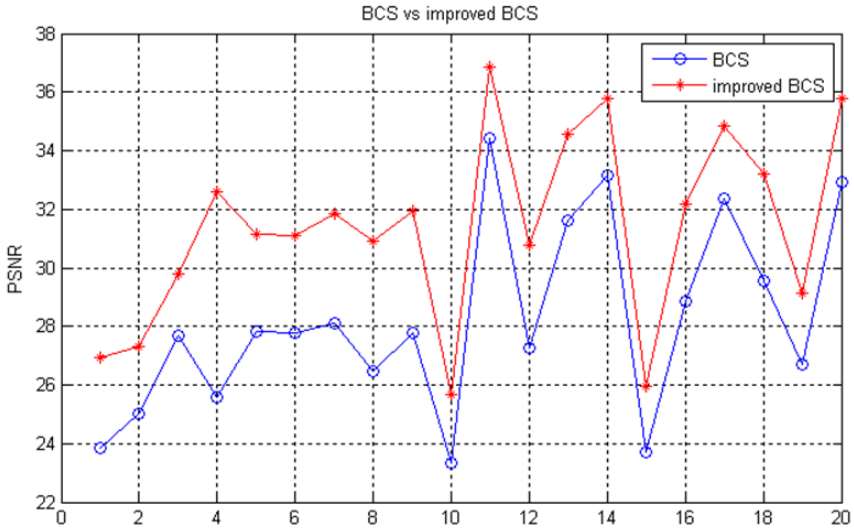

FIGURE 6. BCS vs improved BCS

TABLE 2. Comparison between three types of Bayesian compressive sensing algorithms and the corresponding improved algorithms

MR images		BCS	BCS_new	Laplace	Laplace_new	BCSvb	BCSvb_new
128×128	PSNR	14.4909	19.2956	16.9486	22.6818	17.6392	20.7513
	Rerr	0.6190	0.3921	0.4657	0.2583	0.4258	0.3110
256×256	PSNR	26.3400	29.9250	26.4841	30.4991	24.1298	28.5470
	Rerr	0.1665	0.1129	0.1642	0.1049	0.2054	0.1255
512×512	PSNR	30.0518	32.9086	30.4760	33.3587	25.4229	29.5295
	Rerr	0.1128	0.0814	0.1069	0.0770	0.1991	0.1264

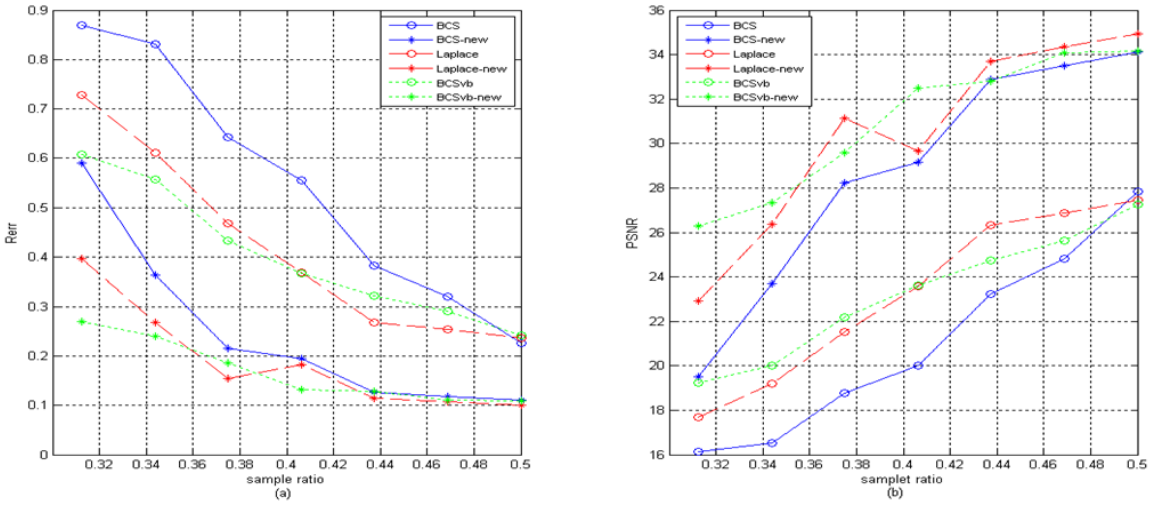

FIGURE 7. Rerr and PSNR of three Bayesian compressive sensing algorithms and corresponding improved algorithms for a random 256×256 MR image in Fig 4. (a)different Rerr value, (b)different PSNR value

coefficients which gather the most energy of an image. Fig 8 shows reconstructed wavelet coefficients of a random 128×128 MR image in Fig 4. Fig 8(a) shows the whole wavelet coefficients of three types Bayesian compressive sensing algorithms and the corresponding improved algorithms. Fig 8 shows that our improved method is not better than the original corresponding method in small coefficients, but is better in large coefficients. Fig 9 shows the first 1024 coefficients, which are the third level coefficients including LL, HL, LH and HH subbands. According to Fig 9, it can be easily seen that the effect of reconstructed wavelet coefficients in the third level via improved method is obviously better, meanwhile it testifies that our improved idea which introduces the wavelet energy feature and wavelet coefficients statistic is good for reconstruction.

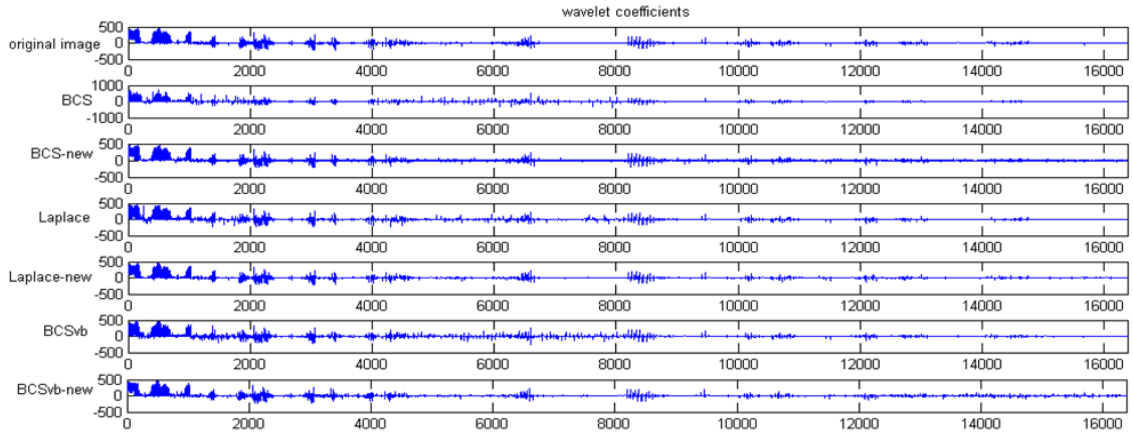

FIGURE 8. Compare to three Bayesian compressive sensing and corresponding improved algorithms:The whole wavelet coefficients, the number are 16384.

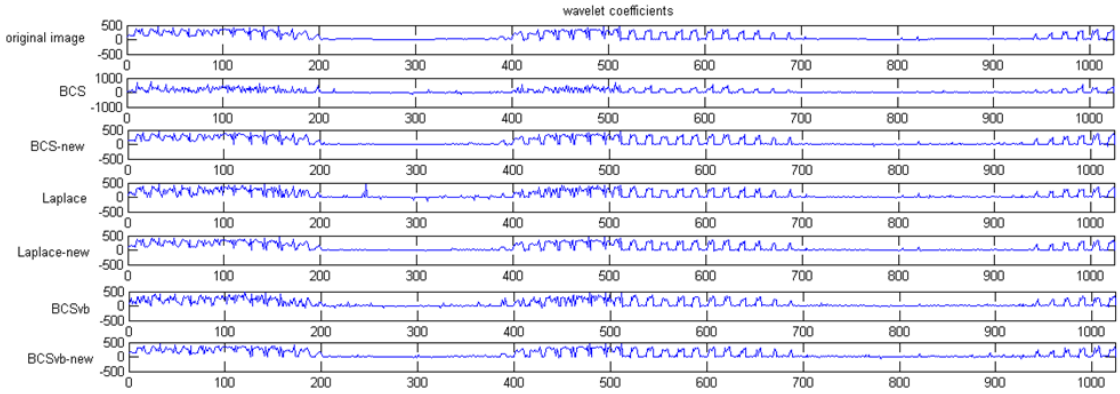

FIGURE 9. Compare to three Bayesian compressive sensing and corresponding improved algorithms: Wavelet coefficients of the third level

7. **Conclusions.** After wavelet transformation, previous Bayesian compressive did not use wavelet coefficients characteristics, but wavelet coefficients have many features and we exploit energy feature of wavelet coefficients. Energy of high level subband is the highest and its influence to image quality is the biggest. According to the wavelet coefficients statistics, we conclude that large coefficients mostly group on high subband of high energy. Exploit wavelet coefficients feature to make subband coefficients of high energy sparse.

The signal is sparser, the reconstruction effect of Bayesian compressive sensing is better. We exploited the feature and proposed improved Bayesian compressive sensing algorithm. Compared to the corresponding traditional Bayesian compressive sensing and derivative Bayesian compressive sensing algorithms, the effect of the new improved algorithm is better both from visual appearance and evaluated criterion data. In the future research, we can introduce the wavelet coefficients features more deeply into compressive sensing to improve reconstruction quality.

Acknowledgement. The work was supported by the Natural Science Foundation of Fujian Province of China (No.2013J05103).

REFERENCES

[1] G. K.Wallace, The JPEG still picture compression standard, *IEEE Transactions on Consumer Electronics,* vol. 38, pp. 18C34, 1992.
[2] C. Christopoulos, JPEG2000 tutorial, *in Proc. IEEE Int. Conf. Image Process. (ICIP), Kobe, Japan,* 1999. Online Available: http://www.dsp.toronto.edu/ dsp/JPEG2000/
[3] D. Donoho, Compressed sensing, *IEEE Trans. Inf. Theor*y, vol. 52, no. 2, pp. 1289C1306, 2006.
[4] E. J. Cands and M. B. Wakin, An introduction to compressive sampling, *IEEE Signal Process. Mag.,* vol. 25, pp. 21C30, 2008.
[5] E.bCande's, 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. 489C509, 2006.
[6] 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. 1182C1195, 2007.
[7] E. Cands and J. Romberg, Sparsity and incoherence in compressive sampling, *J. Inverse Problems*, vol. 23, pp. 969C985, 2007.
[8] Ji, S., Xue,Y., and Carin, L. Bayesian compressive sensing, *IEEE Trans. Signal Processing*, vol. 56, no. 6, pp. 2346C2356, 2008.
[9] S. D. Babacan, R. Molina, and A.K. Katsaggelos, Bayesian compressive sensing using laplace priors. *IEEE Trans. Image Processing,* vol. 19, no. 1, pp. 53-63, 2010.
[10] D. Baron, S. Sarvotham, R.G. Baraniuk, Bayesian compressive sensing via belief propagation, *IEEE Trans. Signal Processing,* vol. 58, no. 1, pp. 269C280, 2010.
[11] Ji. Shihao, D. Dunson, and L. Carin, Multi-Task Compressive Sensing, *IEEE Trans. Signal Processing,* vol. 57, no. 1, pp. 92-106, 2009.
[12] X. S. Hou, L. Zhang, A Bayesian Compressive Sensing Algorithm for Natural Images Based on Sparse Filtering in Directional Lifting Wavelet Transform Domain
[13] . *Journal of Xi'an Jiaotong Univerity,* vol. 48, no. 10, pp. 15-21, 2014.
[14] F. Rossetti, J. Katto, and M. Ohta, Improved scanning methods for wavelet coefficients of video signals, *J. Signal Processing*, vol. 8, no. 4, pp. 365-378, 1996.
[15] E. Vera, L. Mancera, S.D. Babacan, R. Molina, and A.K. Katsaggelos, Bayesian compressive sensing of wavelet coefficients using multiscale Laplacian priors. *IEEE/SP 15th Workshop. Statistical Signal Processing,* pp.229-232, 2009.
[16] L. h. He, H. Chen., and L. Carin, Tree-structured compressive sensing with variational Bayesian analysis. *IEEE. Signal Processing Letters,* vol. 17, no. 3, pp. 233-236, 2010.
[17] L. H. He and L. Carin, Exploiting structure in wavelet-based Bayesian compressive sensing, *IEEE Trans. Signal Processing,* vol. 57, no. 9, pp. 3488-3497, 2009.
[18] C. Feng., Z. H. Wei. and X. Liang, A novel image compressive sensing approach with column sparse prior. *Information Science and Engineering (ICISE), 1st*, pp: 1075-1078, 2009.
[19] E. Y. Lam, Statistical modeling of the wavelet coefficients with different bases and decomposition levels. *IEEE Proceedings. Vision, Image and Signal Processing,* vol. 151, no. 3, pp. 203-206, 2004.
[20] R. Caruana, Multitask learning, Machine Learning, vol. 28, no. 1, pp. 41C75, 1997.
[21] M.E. Tipping, Sparse Bayesian learning and the relevance vector machine, *Journal of Machine Learning Research,* vol. 1, pp. 211C244, 2001.
[22] C.M. Bishop and M. E. Tipping, Variational relevance vector machines, i*n Proc.of the 16th Conference on Uncertainty in Artificial Intelligence(UAI 16)*, pp.46C53, 2000.
[23] H. Ishwaran and J. S. Rao, Spike and slab variable selection : *Frequentist and Bayesian strategies, Ann. Statist.,* vol. 33, pp. 730C773, 2005.

[24] E. I. George and R. E. McCulloch, Variable selection via Gibbs sampling, *J. Amer. Statist. Assoc.,* vol. 88, pp. 881C889, 1993.

[25] H. Chipman, Bayesian variable selection with related predictors, Canad. *J. Statist.*, vol. 24, pp. 17C36, 1996.

[26] C. Carvalho, J. Chang, J. Lucas, Q. Wang, J. Nevins, and M. West, High-dimensional sparse factor modelling: Applications in gene expression genomics, *J. Amer. Statist. Assoc.,* 2008.

[27] M. West, Bayesian factor regression models in the large p, small n paradigm, *in Bayesian Statist. 7, J. M. Bernardo, A. P. Dawid, J. O. Berger, M. West, D. Heckerman, M. J. Bayarri, and A. F. M. Smith, Eds. Oxford* , U.K.: Oxford Univ. Press , pp. 723C732, 2003.