A Novel Image Reconstruction Algorithm Based on Concatenated Dictionary

Zhe Liu
School of Science, Northwestern Polytechnical University, Xi'an, 710072, China
liuzhe@nwpu.edu.cn

Xiaoxian Zhen, Cong Ma
School of Science, Northwestern Polytechnical University, Xi'an, 710072, China
zhenxiaoxian@126.com, Macong19861010@163.com

***Abstract*—Image sparse representation plays a vital role in the process of image reconstruction. In recent years, several pioneering works suggested that signals/images could be represented sparsely with a redundant dictionary. The selection of components from the dictionary directly influences the precision of the reconstructed image, while the scale of dictionary influences the computational efficiency. This paper presents a novel method for image reconstruction, which decomposites the image by concatenating a redundant dictionary of several bases and then reconstruct the image efficiently by means of Matching Pursuit algorithm. The proposed method constructs the concatenated dictionary with cosine bases, wavelet bases and contourlet bases, which will lead to a better approximation of the original image. The experimental results show that the proposed algorithm can greatly reduce the computational complexity and generate a better reconstruction effect compared with previous methods.**

***Index Terms*—component; concatenated dictionary; sparse decomposition; matching pursuit; orthogonal bases**

I. INTRODUCTION

The task of decomposing signals into their building atoms is of great interest in many applications such as compression, enhancement, restoration, and so on with its benefits of fewer decomposition coefficients and high flexibility. Image is a kind of complex 2-dimensional signal which usually contains many types of features such as edges, contours, textures and so on. Traditional methods for image sparse decomposition are based on orthogonal linear transforms which are not suitable for the complex features presented in natural images. A typical example is that it is hopeless to analyze a mixture of impulses and sinusoids with a single base such as DCT or Fourier transform, because each phenomenon needs its own appropriate basis. This would be similar to fitting a round peg into a square hole if we decompose the signal into either basis alone. Therefore, it is natural to consider that constructing a representation with a combination of several bases would be more effective than only with single one.

Sparse representations over redundant dictionaries came forth in 1990s. In 1993, Mallat and Zhang[1] originally proposed the idea of signal sparse decomposition with redundant dictionaries and first introduced the Matching Pursuit (MP) algorithm. In 1994, Mallat put forward the matching pursuit algorithm based on image sparse decomposition which motivated the research effort on sparse decomposition applications in the image processing field. Many achievements have been obtained so far on the image sparse decomposition[2]. Decomposing a signal/image based upon redundant dictionaries is a new method which approximates a signal with an overcomplete function system instead of an orthogonal basis to provide a sufficient choice for adaptive sparse decompositions. The overcomplete function system of redundant functions is referred to as redundant dictionary and the functions are called atoms. The new signal decomposition method results in a higher matching degree between signal and atoms, and also a greater flexibility in capturing the inherent structure of the natural images with the redundant dictionaries.

Finding the best approximation of the original image is equivalent to solving the l_0 norm minimization problem, thus the image reconstruction based on redundant dictionaries can be formulated as

$$\min\|\theta\|_0 \quad s.t. \quad f = D\theta = \sum_{k=1}^{K} c_k g_k \tag{1}$$

where f is the original signal or image, D is the redundant dictionary, θ represent the coefficient vector which is the projection of f onto the selected atoms. $\|\theta\|_0$ denotes the number of nonzero entries in the coefficient vector θ, namely the l_0 norm. Since l_0 norm formula is non-convex, finding the sparsest representation of a given signal over a redundant dictionary is NP-hard. Therefore, researchers had developed many suboptimal approximation methods instead. Currently, the most widely used methods are Matching Pursuit[1] which will be accounted in the section II. In section III we discuss the image reconstruction algorithm based on the proposed concatenated dictionary and section VI addresses the numerical experiments to validate our algorithm's performance. Conclusions are drawn in the last section.

Corresponding author email: liuzhe@nwpu.edu.cn

doi:10.4304/jcp.7.2.444-449

II. MATCHING PURSUIT

As is well known, Matching Pursuit[1], Basis Pursuit[3] and their variants[4-5] have been widely applied for image reconstruction. The MP algorithm is much faster and much easier to implement, which makes it an attractive alternative for image recovery problems. However, its computational complexity is still very high, because it must project the image or residual image onto the redundant dictionary for each iteration.

Matching Pursuit (MP) is an adaptive greedy algorithm that optimizes the signal approximation by iteratively selecting atoms which best match the image structures at each step from the redundant dictionary. Thus the linear combination of the selected atoms is an approximate of the given image. After *k*-th iteration, the image f can be decomposed as

$$f = \sum_{k=1}^{K} \left\langle R^k f, g_{\gamma_k} \right\rangle g_{\gamma_k} + R^{K+1} f \tag{2}$$

where $R^{k+1} f$ defines the residue at order $k+1$, and $R^0 f = f$. Thus, we obtain an energy conservation

$$\|f\|^2 = \sum_{k=1}^{K} \left| \left\langle R^k f, g_{\gamma_k} \right\rangle \right|^2 + \left\| R^{K+1} f \right\|^2 \tag{3}$$

As $K \to \infty$, the energy $\left\| R^{K+1} f \right\|$ decreased exponentially to 0. According to [6], there exists $\alpha, \beta \in (0,1]$ to ensure

$$\left\| R^{k+1} f \right\| \le \left(1 - \alpha^2 \beta^2\right)^{1/2} \left\| R^k f \right\| \tag{4}$$

for $k>0$, where α is an optimal factor which is relative to the strategy adopted to select the optimal atom from the dictionary; β depends on the selection of the dictionary which represents the capability of capturing the structure features of signal *f*. Experientially, the convergence performance of the MP algorithm is relative to both searching strategies and the selection of dictionary. Therefore, it is very important to construct appropriate redundant dictionary.

III. MATCHING PURSUIT ALGORITHM BASED ON CONCATENATE DICTIONARY

A. Redundant concatenate dictionary with several bases

In this paper we consider structure characteristics of natural images and construct the redundant dictionary by concatenating several bases. Let $D = \{D_1, D_2, \cdots D_L\}$ be a concatenate dictionary and $D_i = \{g_k^i, k = 1, 2, \cdots, N\}$ is the *i*th basis function system, $i = 1, 2, \cdots, L$. $N = I_r \times I_c$ is the size of the given image, thus *D* is a redundant dictionary concatenated by *L* bases. We hope to find $K(K<<N)$ coefficients to approximate f by

$$f = \sum_{k=1}^{K} c_k g_k \ . \tag{5}$$

We choose the cosine bases, contourlet bases, orthogonal wavelet bases 'sym1-sym10' (symlet series wavelet bases), to construct the redundant dictionary. Since all the chose bases have fast algorithms, which will greatly decrease the computational complexity.

The Discrete Cosine Transform (DCT) is an orthogonal transform, which is known as the best transform for first order Markov stationary signals. Its coefficients essentially represent the frequency contents, similar to the Fourier coefficients. Thus, it's suitable to provide the sparse representation for the smooth or periodic part of the images. Discrete orthogonal wavelet transform has a good ability in analyzing the time-frequency localization, thus it could capture the singularity features of image. Contourlet is a new image multi-scale geometric analysis tool, which could represent image sparsely and excellently approximate the geometry structure of image. Simultaneously, Contourlet possess the property of multi-scale, localization, critical sampling, multi-directionality and anisotropy. The Contourlet basis functions are multi-directional and anisotropic, with flexible aspect ratio, thus enable describing smooth contours near optimally. Moreover, Contourlet could approximate geometric structures of images efficiently, with much less coefficients than wavelet transform. The ultimate purpose of Contourlet transform is to approximate the initial image with basis functions similar to contour segments, which makes it more suitable to capture the contours and edge details[7]. Therefore, it could represent the natural structure much better. Especially when we retain much fewer important coefficients, it could obtain better vision effect and signal to noise ratio (SNR) compared to wavelets transform, for it obtains more detail information of the image[8].

B. Image reconstruction based on concatenate dictionary

Image reconstruction on concatenated dictionary is a special case of image sparse decomposition based on redundant dictionary or overcomplete dictionary. In this case, the concatenated dictionary is constructed by concatenating several transform bases. Because of the redundancy of concatenated dictionary, image recovery based on concatenated dictionary could approximate the image more flexibly, and possibly obtain a better reconstruction, whereas it is a NP-hard problem. The traditional method for image reconstruction based on redundant dictionary, such as Matching pursuit (MP) and Basis pursuit (BP), could also be used for image reconstruction based on redundant concatenated dictionary.

The main idea of our algorithm can be described as follows:

Firstly, the redundant dictionary $D = \{D_1, D_2, \cdots D_L\}$ is constructed by concatenating cosine bases, symlet series of wavelet bases (sym1-sym10) and Contourlet bases. Then matching the original image with each set of bases to select an optimal set of bases matching the original image in terms of the inherent structure of image;

Secondly, projecting the image over the optimal basis D_{i1} and obtain the first set of important coefficients $c_k^{i_1}, k = 1, \cdots, K_{i_1}$, and the optimal approximation of f

according to (5) as well as residual image $R^1 f = R^0 f - f_1$, where $R^0 f = f$.

Finally, performing the projection on the residual image $R^1 f$ with the same process and obtain the next set of important coefficients.

At each iteration, we select a set of bases optimally matching the residual image and thus approximate the image more exactly. The procedure is implemented iteratively until achieving the precision requirement of residual image $R^k f$ or the number of decomposition coefficients.

The detailed implementation of our algorithm consists of three steps:

(1) Constructing a redundant dictionary $D=\{D_1, D_2, \cdots D_L\}$;

(2) Projecting the initial image f on the redundant dictionary by selecting a set of bases optimally matching the image and thus obtain a set of important coefficients $c_k^{i_s}, k = 1,\cdots,K_{i_s}$ and the residual image $R^1 f = R^0 f - f_1$, where the optimal approximate image $f_1 = \sum_{k=1}^{K_{i_1}} c_k^{i_1} g_k^{i_1}$;

(3) Decomposing the residual $R^n f$ (n=1,2,…,N) into

$$\mathrm{R}^n f = \sum_{k=1}^{K_{i_{n+1}}} c_k^{i_{n+1}} g_k^{i_{n+1}} + \mathrm{R}^{n+1} f, \quad (6)$$

which also defines the residual image $R^{n+1} f = R^n f - f_{n+1}$, where $f_{n+1} = \sum_{k=1}^{K_{i_{n+1}}} c_k^{i_{n+1}} g_k^{i_{n+1}}$ is a optimal approximation of $R^n f$;

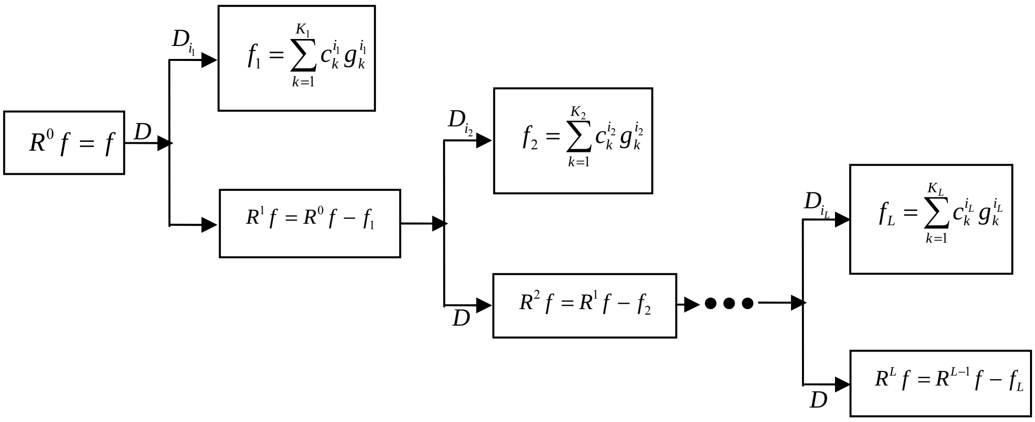
(4) Performing the decomposition in step (3) until the stop condition is satisfied. Finally, K coefficients $c_1^{i_1}, c_2^{i_1}, \cdots c_{K_{i_1}}^{i_1}; c_1^{i_2}, c_2^{i_2}, \cdots, c_{K_{i_2}}^{i_2}; \cdots; c_1^{i_L}, c_2^{i_L}, \cdots, c_{K_{i_L}}^{i_L}$ are selected, classify them as L groups according to subscript:

$c_1 : \{c_1^{i_1}, c_1^{i_2}, \cdots c_1^{i_L}\}, \cdots, c_{K_{i_L}} : \{c_{K_{i_L}}^{i_1}, c_{K_{i_L}}^{i_2}, \cdots c_{K_{i_L}}^{i_L}\}$,

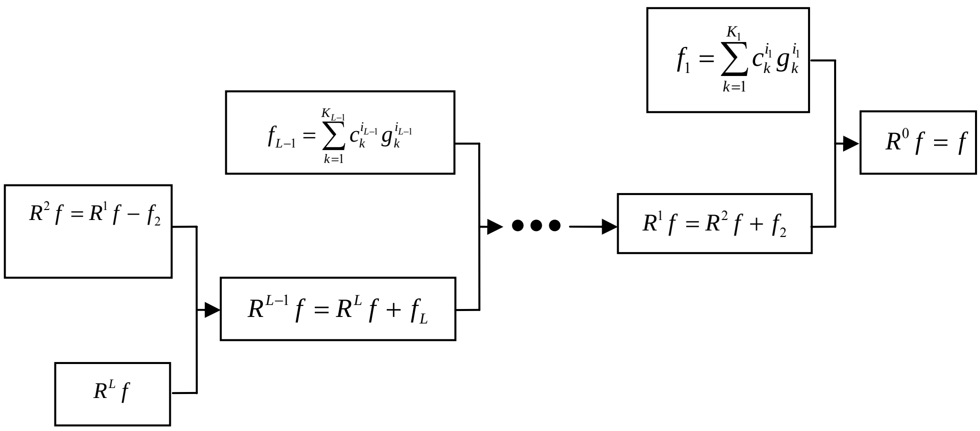
and optimal approximation of image f is obtained:

$$f \approx f_K = \sum_{k=1}^{K} c_k g_k \quad (7)$$

To observe the processing intuitively, Fig.1 shows the flow chart of image sparse decomposition and reconstruction based on concatenated dictionary.



(a) Image decomposition process



(b) Image reconstruction process

Figure1. Flow chart of image sparse decomposition and reconstruction based on concatenated dictionary

IV. Simulation Experiments and Analysis

To validate the proposed algorithm, we perform a series of experiments on a 4GHz Pentium IV (4G RAM) computer with Windows XP and Matlab7.0.

A. Image reconstruction performance

The purpose of this first experiment is to demonstrate that the proposed algorithm achieves a better reconstruction effect. Taking the 512×512 Mondrian image for example and construct redundant dictionary by concatenating cosine bases, symlet series of wavelet bases (Sym1-Sym10) and contourlet bases. We select $K = [\mathbf{0.125} \times I_r \times I_c]$. Tab.1 shows a detailed comparison of the mean square error (MSE) and the peak signal-to-noise ratio (PSNR) based on single wavelet base on the Mondrian image, where the MSE is formulated by

$$MSE = \| f - f_K \| = \sum_{i=1}^{I_r} \sum_{j=1}^{I_c} [f_{i,j} - (f_K)_{i,j}] \Big/ (I_r \times I_c) \quad (8)$$

here $f_{i,j}$ denote the grayscale value of image at position (i, j) and $N = I_r \times I_c$ is the size of image. Sym2 is shown to be superior compared with other bases in approximating the original image, with MSE=6.6258 and PSNR=39.918 dB.

Fig.2 illustrates the comparison of image reconstruction performance between the proposed algorithm and the Wavelet, Contourlet transform. The experiments are performed on the 512×512 Peppers image which has a sparse geometrical structure and the Fingerprint image which has abundant texture. The Wavelet in the experiments uses single optimal wavelet basis. We can clearly see that the proposed algorithm is superior on both Peppers image and Fingerprint image by comparing the PSNR quantitatively. This attributes to the redundant concatenated dictionary which could match the image structural characteristics more flexibly.

We have also applied our method to remote sensing images. Fig.3 shows the reconstructions of a remote sensing image 041_1(512×512). As can be seen, the reconstruction image based on the new algorithm is superior than both others from the visual sense, with more clear visual effect. Also, from the perspective of PSNR, our algorithm outperforms the Wavelet and Contourlet transform.

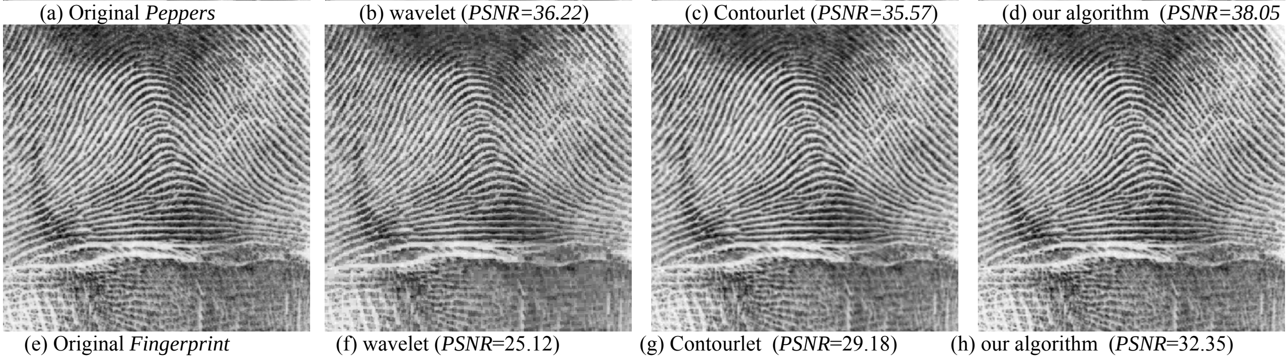
To contrast the three algorithms comprehensively, we applied them on four 256×256 grayscale images (bridge, Lena, house, angiogram) reserving 1024, 2048, 4096 and 8192 coefficients respectively. The experimental results are showed in table 1 from which we can derive a detailed comparison.

As can be seen, the PSNR of the new algorithm is more than that of Wavelet transform by 0.91-4.10 dB, that of Contourlet transform by 0.66-2.51 dB, when the retained coefficient $K = \mathbf{1024}$; For $K = \mathbf{2048}$, the new algorithm is more than that of Wavelet transform by 0.17-0.82 dB, that of Contourlet transform by 0.55-2.38 dB; For $K = \mathbf{4096}$, the proposed algorithm is more than that of Wavelet transform by 0.19-0.82 dB, that of Contourlet transform by 1.02-4.62 dB and for $K = \mathbf{8192}$, the proposed algorithm is more than that of Wavelet transform by 0.16-0.59 dB, that of Contourlet transform by 1.10-5.90 dB. Since the proposed algorithm retains fewer coefficients, it would have a better performance in the signal encoding.

TABLE I. *PSNR* AND *MSE* OF MONDRIAN'S RECONSTRUCTION RESULTS BASED ON SINGLE WAVELET BASE

wavelet	Sym1	Sym2	Sym3	Sym4	Sym5	Sym6	Sym7	Sym8	Sym9	Sym10
MSE	6.6321	6.6258	7.053	7.3115	7.505	7.8875	8.331	8.5845	9.0362	9.261
PSNR/dB	39.914	39.918	39.647	39.491	39.377	39.161	38.924	38.794	38.571	38.464

(a) Original *Peppers* (b) wavelet (*PSNR=36.22*) (c) Contourlet (*PSNR=35.57*) (d) our algorithm (*PSNR=38.05*



(e) Original *Fingerprint* (f) wavelet (*PSNR*=25.12) (g) Contourlet (*PSNR*=29.18) (h) our algorithm (*PSNR*=32.35)

Figure2. Comparison of reconstruction performance on natural images "*Mondrian* and *Babara*"

(a) Original remote sensing image (b) wavelet (*PSNR=34.06*) (c) Contourlet (*PSNR=33.12*) (d) our algorithm (*PSNR=35.27*)

Figure3. Comparison of reconstruction performance on remote sensing images “041-1”

TABLE II. THE COMPARISON OF PSNR PERFORMANCE ON IMAGES

PSNR(db)	Size	wavelet	Contourlet	Our algorithm
bridge	*K*=8192	27.46	26.75	27.85
	K=4096	24.90	24.60	25.62
	K=2048	23.09	22.72	23.87
	K=1024	21.55	19.95	22.46
Lena	*K*=8192	34.48	31.74	34.64
	K=4096	29.87	28.42	30.06
	K=2048	26.15	25.89	26.44
	K=1024	21.11	23.82	24.48
house	*K*=8192	38.36	32.77	38.67
	K=4096	34.46	30.08	34.70
	K=2048	29.95	27.74	30.12
	K=1024	23.00	25.50	27.10
angiogram	*K*=8192	35.64	34.31	36.23
	K=4096	31.66	31.29	32.48
	K=2048	28.64	28.61	29.46
	K=1024	26.17	25.26	27.30

B. Performance with the change of coefficients retained

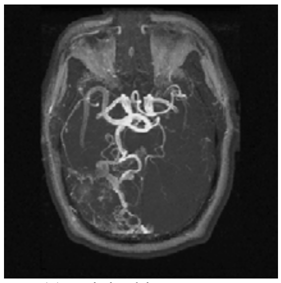
To contrast the three algorithms further, another set of experiments is implemented.

Fig.4 shows the comparison of image reconstructions with the change of coefficients retained，from which we can observe that the reconstruction is greatly affected by the quantity of retained coefficients. If more coefficients are retained, a clearer image with higher PSNR could be obtained.

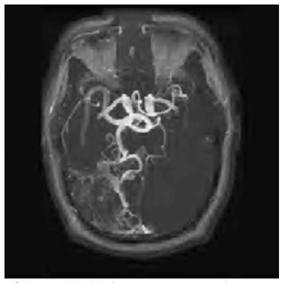
Fig.5 compares the performance of PSNR and MSE with the change of coefficients retained. As can be seen the value of PSNR grows and that of MSE decreases with the count of coefficients increasing for all algorithms. But the new algorithm is superior than the others obviously. Put differently, the proposed algorithm requires fewer coefficients for the same reconstruction quality.

C. Time comparison

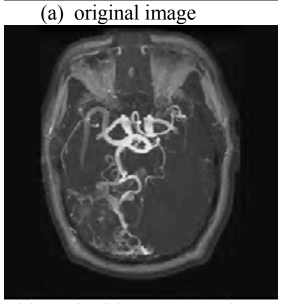
This last experiment is to verify that the proposed algorithm is preferable on running time, too. Tab.3 shows the running time of reconstructing five 256×256 grayscale images (Peppers, Lena, Barbara, Boats, house and Angiogram) based on concatenated dictionary and Gabor dictionary respectively. In all our image reconstruction experiments based on Gabor dictionary, we represent the images by 112 modulated Gabor atoms. The numerical results show that the image reconstruction based on concatenated dictionary are about 35 times faster than the algorithm based on Gabor dictionary with the same PSNR.



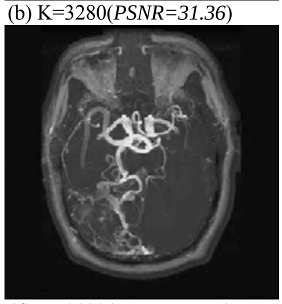
(a) original image



(b) K=3280(*PSNR=31.36*)

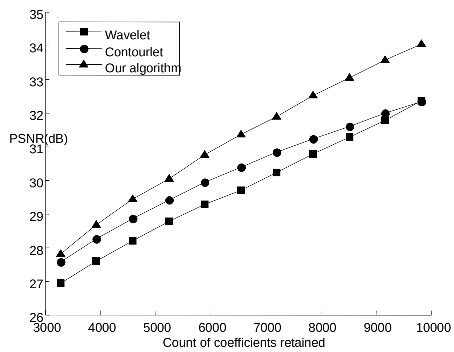


(c) K=6553(*PSNR=34.85*)

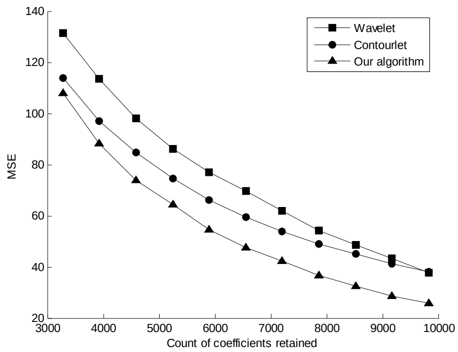


(d) K=9833(*PSNR=37.39*)

Figure4. Reconstructions with the change of sampling ratio



(a)



(b)

Figure5. Performance with the change of coefficients retained

TABLE III. THE COMPARISON OF COST TIME ON IMAGES

Time(s)	Gabor dictionary	concatenate dictionary
Peppers	397.52	11.33
Lena	386.55	11.25
Barbara	382.61	11.34
Boats	381.22	11.20
House	380.09	11.21
Angiogram	390.89	11.28

V. DISCUSSION AND CONCLUSION

So far, overcomplete representations have a wide range of applications. Many heuristic and theoretical arguments have been advanced to support the benefits of dictionary redundancy. The research on image reconstruction based on redundant dictionaries is mainly focused on the following two aspect: (1) designing more efficient algorithms to reduce the computational complexity; (2) constructing more appropriate redundant dictionaries or concatenate dictionaries in terms of the structure characteristics of given images. Considering the above two aspect, this paper concatenated the cosine bases, contourlet bases and Symlet series wavelet bases (sym1-sym10) to construct the redundant dictionary. The simulation results showed that the performance of the reconstructed images bases on concatenated dictionary is better than those based on general dictionaries, such as Gabor dictionary, at the same compression ratio. Since the concatenated dictionary may be constructed on possible bases, it seems reasonable to further investigate the construction of concatenated dictionaries using cosine bases, ridgelet bases, bandlet bases, directionlet bases and all those has fast algorithms. On the other hand, it is imperative to design more efficient optimality approximation algorithms in terms of the structure characteristics of redundant dictionary, such as Orthogonal Matching Pursuit (OMP) [7], Tree Matching Pursuit (TMP) [8] and their variants.

ACKNOWLEDGMENT

This work was supported by the National Natural Science Foundation of China under Grant No. 61071170, and Program for New Century Excellent Talents in University.

REFERENCES

[1] Mallat S, Zhang Z, " Matching pursuits with time-frequency dictionaries," *IEEE Trans Signal Process*.1993, 41(12):3397~3415.

[2] Bergeau F, Mallat S, " Matching Pursuit of images," *In Proceedings of IEEE Signal Processing*. Piladelphia, USA, 1994:330-333.

[3] Chen S, Donoho D, Saunders M, " Atomic Decomposition by Basis Pursuit ," *SIAM J Sci Comput* .1999, 20 (1): 33-61.

[4] Pati Y C, Rezaiifar R, Krishnaprasad P S, " Orthogonal Matching Pursuit: Recursive Function Approximation with Applications to Wavelet Decomposition," *Proceedings of the 27th Annual Asilomar Conference in Signals, Systems and Computers. Los Alamitos: IEEE*, 1993, 1(11): 40-44.

[5] C.La and M. N. Do, "Signal reconstruction using sparse tree representation," *Proceedings of SPIE. San Diego, CA, United States: International Society for Optical Engineering*. 2005,5914:1-11.

[6] Vandergheynst P, Frossard P, " Efficient image representation by anisotropic refinement in matching pursuit," *Proceedings of IEEE on ICASSP. USA:Salt Lake City*, 2001: 1757-1760.

[7] M. N. Do, M. Vetterli, "The contourlet transform: an efficient directional multiresolution image representation," *IEEE Transactions Image on Processing*, Dec. 2005, 14(12): 2091-2106.

[8] Shenqiu Zhang, Moloney C, "The nonredundant contourlet transform (NRCT): A multiresolution and Multidirection image representation," *Electrical and Computer Engineering, 2008. CCECE 2008*. Canadian Conference on 4-7 May 2008 Page(s):001323 – 001326.

Zhe Liu was born in Shaanxi Province, China, in 1970. She received her Ph.D degree from Northwestern Polytechnical University in 2002. She is now a professor at Department of Applied Mathematics, School of Science, Northwestern Polytechnical University. Her research interests include compressed sensing, image processing and intelligent information processing.

XiaoXian Zhen was born in Shanxi Province, China, in 1988. She is a graduate student at Department of Applied Mathematics, School of Science, Northwestern Polytechnical University. Her research interests include compressed sensing and image processing.

Cong Ma was born in Henan Province, China, in 1986. He is a graduate student at Department of Applied Mathematics, School of Science, Northwestern Polytechnical University. His research interests include compressed sensing and signal processing.