Classification of Cardiac Ultrasound Image Sequences Based on Sparse Representation

Xiaofang Hou
Department of Computer Science and Engineering, North China Institute of Aerospace Engineering, Langfang, China
Email: houxf1208@163.com

Penghua Zhu
Department of Computer Science and Engineering, North China Institute of Aerospace Engineering, Langfang, China
Email: lf_zhupenghua@163.com

Yanxin Ma
School of Computer Science and Technology, Harbin Institute of Technology, Harbin China
Email:ma_yan_xin@qq.com

***Abstract*—To classify thrombosis and pectinate muscle in cardiac ultrasound image sequences, a classification method based on sparse representation is proposed. This method extracts GLCM based texture features to form the sample set and compute the sparse solution with coefficients how a test sample be represented by the training set. After that, two kinds of constraints and classification strategy are added to achieve the classification. Experiment results shows that the proposed approach can achieve a classification accuracy of 91.92%, significantly higher than other popular classifiers.**

***Index Terms*—Sparse representation, Image sequence, Texture feature, Thrombosis, Pectinate muscle**

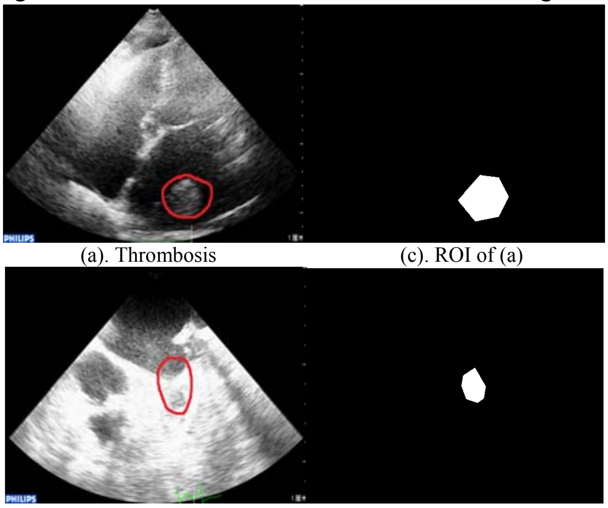
I. Introduction

Cardiac thrombus has been a serious threat to human health and can also lead to myocardial infarction. Cardiac thrombus is more likely to be in left atrial appendage. We usually diagnose thrombus by ultrasound, such as transthoracic echocardiography and transesophageal echocardiography. Transesophageal echocardiography has a high sensitivity of thrombosis detection up to 99%.

It's difficult to differ thrombus with pectinate muscles, sometimes leads to misdiagnosis. With the rapid development of computer-aided diagnostic technology, combining CAD and cardiac ultrasound image sequences to detect blood clots become an inevitable trend. Cardiac thrombus mostly appears in the left atrium and left atrial appendage. Pectinate muscle is normal muscle tissue, and sometimes it's difficult to distinguish it and thrombus in some clinical situations.

Thrombosis and pectinate muscles has some different characteristics. Fresh thrombus has low ultrasound echo and is active, as old thrombus has strong ultrasound echo and is not active. Repeatedly formed thrombus has uneven echo intensity and different stratification.

Pectinate muscles has even echo intensity and is active,and it is consistent with the surrounding tissue. Based on the above identification features, distinguishing thrombosis and pectinate muscles in the image sequence will be easier. We invited experienced physicians to calibrate the region of interest (ROI) area, shown in Figure 1. We will do the feature extraction in the region.



(a). Thrombosis (c). ROI of (a)

(b). Pectinate muscles (d). ROI of (b)

Figure 1. Thrombosis, pectinate muscle and their corresponding ROI

(a) and (b) are respectively original thrombosis image and pectinate muscle image taken by ultrasound machine. We invited experienced physicians to calibrate the region of interest (ROI) area, respectively shown in (c) and (d).

Classic computer-aided diagnostic techniques can be divided into two major steps: feature extraction and classification. In medical image classification and recognition, we often provide new feature extraction methods and classification methods to improve it. And we choose to improve the classification method.

LangFang Science and Technology Bureau Fund Project (Project No. 2013011003, No. 2011011006).

Address: Aimin Road No.133, 130(15) box, Langfang, Hebei, China. Telephone number: 15833163645.

doi:10.4304/jsw.9.9.2309-2314

Sarkar and Chaudhuri improved the texture feature extraction based on box-counting method of fractal dimension, and proposed differential box counting algorithm [1], with fast and precise characteristics. Markov random field (Markov Random Field) theory is also widely used in texture feature extraction [2]. Gabor filter based texture feature extraction [3] and wavelet-based texture feature extraction [4] also provides us a new direction to extract texture feature.

Over the last decade, sparsity has become one of the most popular concept in signal processing and application. Recently, researchers have been studied over-complete signal representation. The charm of redundancy signal representation is that it can economically represent a large class of signal. The interests in sparsity come from the new sampling theory called compressed sensing. Compressed sensing is a succedaneum of Shannon sampling theorem, and it gives strong theoretical support to the approach of seeking coefficient solution.

This paper studies the classification method based on sparse representation, which is a hot spot of compressed sensing and signal processing. The main idea of sparse representation is to linearly represent the test sample with the subspace of sample space with the same class as test samples, so the representation coefficient is sparse when using the entire sample space to represent the test sample. This article uses two popular classifiers SVM (Support Vector Machine) [5] and artificial neural networks (Artificial Neural Networks) [6] to do comparative experiments.

II. Texture Feature Extraction

A. Texture Feature Extraction Based on GLCM

Texture features are widely used in medical image processing and recognition. We extract some GLCM-based texture features as image features.

Texture describes some regular changes of the image pixel, color and grayscale and reflects the depth of groove of the image, grayscale differences, complexity of the shape and texture. In 1970s, Haralick proposed the concept of GLCM (Gray Level Co-Occurrence Matrix) based method [7], which is used to estimate the second-order conditional combination probability density of grayscale images. Let $\mathbf{F}=f(x, y)$ represents the grayscale image with size of a two-dimensional space of W×H, and it's gray level is L. Then the gray level co-occurrence matrix **G** of the image **F** is a square matrix with a dimension of L×L, which the element $p(i, j)$ in the position (i, j) in **G** represents the probability that grayscale of image **F** changes from i to j through a displacement. And the displacement can be expressed as (d, θ), including the two variables direction and distance. We take d = 1, 2, 3， θ = 0°, 45°, 90°, 135° to construct the GLCM.

Values of GLCM of coarse texture image are mainly concentrated in the main diagonal nearby, while values of GLCM of fine texture image scattered throughout the matrix. After establishing the GLCM, we can then extract meaningful statistics as texture features. Commonly used features are as follows:

(1) Contrast is the difference of adjacent points.

$$\text{Contrast} = \sum_{i,j} |i-j|^2 \; p(i,j)$$

(2) Correlation is the similarity of row or column.

$$\text{Correlation} = \sum_{i,j} \frac{(i-\mu i)(j-\mu j)p(i,j)}{\sigma_i \sigma_j}$$

(3) Energy is squared sum of elements.

$$\text{Energy} = \sum_{i,j} p(i,j)^2$$

(4) Homogeneity reflects the proximity of elements and the diagonal.

$$\text{Homogeneity} = \sum_{i,j} \frac{p(i,j)}{1+|i\text{-}j|}$$

(5) Entropy reflects the complexity or uniformity of the texture.

$$\text{Entropy} = -\sum_{i,j} p(i,j) \log p(i,j)$$

In addition, we also extracted the mean and variance of GLCM as texture features.

B. Texture Features in Image Sequences

With the above features, we can form the feature space in image sequences through a special and reasonable way.

Looking for features that can be used for classification of thrombosis and pectinate muscles in the image sequence is not easy. People often use two successive images or frames to measure the continuity and the difference. This idea has been widely used in target tracking. In this paper, we use a more simple and effective way to measure the difference between the two continuous images. Using GLCM-based texture features and take d = 1, 2, 3， θ = 0°, 45°, 90°, 135°, we calculate contrast, correlation, energy, homogeneity, entropy, variance and mean to composite a 51-dimensional texture feature vector. Then we stack feature vector of the reference image I and the floating image I_k to composite a 102-dimensional feature vector. During the experiment, the floating images are taken as I_1, I_2 and I_3. Thus, each image sequence has 3 set of texture features.

III. Sparse Representation Based Classification

A. Principle of Sparse Representation

In this section, we are going to show what sparse representation is and how it functions.

The main idea of sparse representation is to linearly represent the test sample with the subspace of sample space with the same class as test samples, so the representation coefficient is sparse when using the entire sample space to represent the test sample. Assuming the number of categories of over-complete dictionary is k, and the number of samples is n, and the number of training sample of i-th class is n_i (n is the sum of all n_i). Matrix A_i represents all samples of the i-th class as follows:

$$A_i = [v_{i,1}, v_{i,2}, \ldots, v_{i,n_i}] \in \mathbb{R}^{m \times n_i}$$

where $v \in \mathbb{R}^m$ is a m-dimensional vector that represents a sample. And $m = w \times h$ is a sub-image whose size is much smaller than the size of the original image.

If the number of samples in class i is enough, then we can approximately represent any new sample y belonging to class i with A_i:

$$y = \alpha_{i,1} v_{i,1} + \alpha_{i,2} v_{i,2} + \ldots + \alpha_{i,n_i} v_{i,n_i}$$, or

$$y = A_i \alpha_i \in \mathbb{R}^m$$ 。

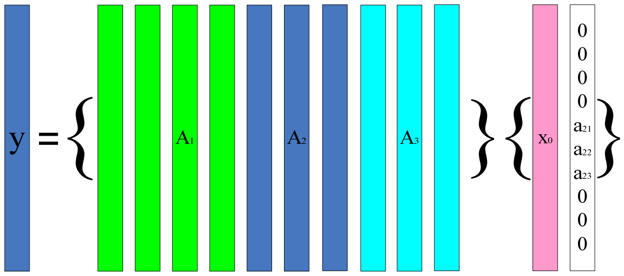


Figure 2. Encode y with x0

But in fact, class label of new sample y is unknown. Then matrix A is defined to represent all samples: $A = [A_1, A_2, ..., A_k]$. Then $y = A_i \alpha_i \in \mathbb{R}^m$ can be represented with A as formula (1):

$$y = Ax_0 \in \mathbb{R}^m \tag{1}$$

where $x_0 = [0, ..., 0, \alpha_{i,1}, \alpha_{i,2}, ..., \alpha_{i,n_i}, 0, ..., 0]^T \in \mathbb{R}^n$. In the coefficient vector x_0, only coefficients corresponding to class i are not zero. Therefore we can say, the coefficient vector x_0 encodes the sample y, shown in Figure 2. In Figure 2, there are 3 classes of training samples, each is represented as A_1, A_2, A_3. Knowing that new sample y belongs to the second class, so you can use the 3 training samples of the second class to linearly represent y. Non-zero Coefficients of x_0 are corresponding to A_2, that are a_{21}, a_{22} and a_{23}.

So we can see that x_0 contains information that can correctly classify the sample y. Then the above problem can be transformed into solving linear equations:

$$y = Ax \tag{2}$$

where $A \in \mathbb{R}^{m \times n}$. In this problem, condition $m < n$ is generally satisfied, then equation (2) is underdetermined, so that the solution of the equations is not unique. ℓ^2 -norm minimization can be used to solve underdetermined equations, as formula (3). But ℓ^2 -norm minimization problem can be solved through the pseudo-inverse matrix. But the solution of ℓ^2 -norm minimization is usually more dense, and its non-zero value are associated with multiple classes, which make it difficult to correctly classify sample y, as shown in the left figure of Figure 3.

$$(\ell^2): \hat{x}_2 = \arg\min \|x\|_2 \quad s.t. \quad Ax = y \tag{3}$$

However, we found that if the number of class k is properly large, then the representation of the sample y is naturally sparse. The more sparse this representation is, the easier to classify y. In this case, our problem can be solved by ℓ^0 -norm minimizing, as the formula (4), where $\|\bullet\|_0$ represents ℓ^0 -norm, that is the number of non-zero elements. It has been proven that when the number of non-zero elements in the solution x of ℓ^0 -norm minimum is less than $m/2$, x is also the only sparse solution [8], that is $\hat{x}_0 = x$.

$$(\ell^0): \hat{x}_0 = \arg\min \|x\|_0 \quad s.t. \quad Ax = y \tag{4}$$

But ℓ^0 -norm minimizing is NP-hard, so it's difficult to solve our problem even use the approximation algorithms are difficult to solve [9]. Recent research of compressed sensing and sparse representation show that [10-12], when the solution x_0 is sparse enough, formula (4) and formula (5) have equal solution. And formula (5) can be solved via linear programming optimization in polynomial time [13].

$$(\ell^1): \hat{x}_1 = \arg\min \|x\|_1 \quad s.t. \quad Ax = y \tag{5}$$

The result of ℓ^1 -norm minimizing is shown in right figure of Figure 3.

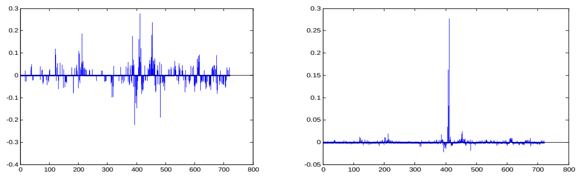


Figure 3. Solution of ℓ^2 and ℓ^1 norm minimizing. We can see that the solution of ℓ^1 norm minimizing is more sparse.

B. Dealing With Noise and Feature Extraction

In practice, image contains noise. Therefore, formula (2) is rewritten as:

$$y = Ax_0 + z \tag{6}$$

where $z \in \mathbb{R}^m$ is a noise item, whose energy satisfies $\|z\|_2 < \varepsilon$. Then we can also use ℓ^1 -norm minimizing to calculate x_0:

$$(\ell^1_S): \hat{x}_1 = \arg\min \|x\|_1 \quad s.t. \quad \|Ax - y\|_2 \le \varepsilon \tag{7}$$

The optimization problem above can also be solved via SOCP [13].

Advantage of sparse representation is that feature extraction is not the key, so you do not have to study the complex feature extraction methods. Assuming the map from the image space to the feature space is $R \in \mathbb{R}^{d \times m}$, where d is the dimension of the feature space, m is the dimension of the image space, and $d \ll m$. Then multiplied both ends of equation (2) with **R** and we have underdetermined equation (8).

$$\tilde{y} = Ry = RAx_0 \in \mathbb{R}^d \tag{8}$$

We can still use ℓ^1 -norm minimization to solve equations (8), as formula (9) below. Here $\varepsilon > 0$ is the error tolerance. We only need to replace matrix A and sample y in equation (2) with $RA \in \mathbb{R}^{d \times n}$ and Ry to solve the equations.

$$(\ell^1_r): \hat{x}_1 = \arg\min \|x\|_1 \quad s.t. \quad \|RAx - \tilde{y}\|_2 \le \varepsilon \tag{9}$$

In this paper, we use GLCM based texture features to construct sample space, and use FPC (Fixed-point Continuation) algorithm [14] to solve the ℓ^1 norm minimization problem.

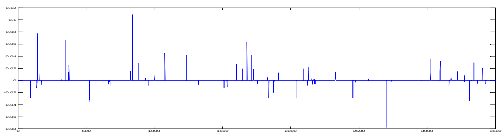
C. Sparse Representation Based Classification

The basic idea of our method is adding two non-negative constraint P/A to sparse solution $\hat{x}_1$, and then use two classification strategy S/M to achieve the classification. The purpose of non-negative constraint is to make all coefficients in sparse solution vector greater than or equal to zero, while classification strategy decide which class a sample belongs.

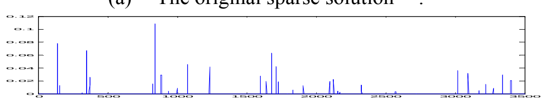
First, the non-negative constraint P/A makes negative coefficients zeroes or their absolute values, shown in Figure 4. In figure 4, (a) shows the original sparse solution $\hat{x}_1$, in the following figure are the results of P and A constraint.

In this paper, we translate the classification problem of thrombosis and pectinate muscles to classification problem of thousands of classes. Every sequence of image is a class, and the 3 corresponding feature vectors can be seen as 3 samples, and then add the non-negative constraint.

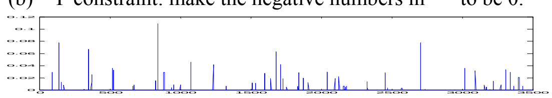
Finally we use S/M strategy to calculate the sum or maximum value coefficient. For strategy S, respectively calculate the sum of the coefficients corresponding to pectinate muscles and thrombotic, and the new sample is classified as the class corresponding to the larger sum. And for strategy M, the new sample is classified as the class corresponding to the largest coefficients.



(a) The original sparse solution $\hat{x}_1$.



(b) P constraint: make the negative numbers in $\hat{x}_1$ to be 0.



(c) A constraint: calculate the absolute value of $\hat{x}_1$.

Figure 4. Adding non-negative constraint to $\hat{x}_1$

IV. EXPERIMENTAL AND ANALYSIS

A. Unbalanced Data Set

We use MATLAB R2010b for Windows and developed a software platform to achieve the experiments. We use FPC algorithm in [14] to solve the ℓ^1 norm minimization problem, so that we can achieve the classification fast and effectively.

Our software platform runs on a well-equipped DELL workstation with Windows Server 2008. First of all, we invited experienced physicians to calibrate all the ROIs, and then we calculate texture features of all image sequences in database over their corresponding ROIs using the method introduced in section II-(B).

In this paper, our database includes thrombus (T) and pectinate muscles (M), and there are 1300 available sequences, including 740 pectinate muscle image sequences and 560 thrombus image sequences. Each sequence is a class, there are 3 feature vectors of each sequence, so we have 1300 × 3 = 3900 available samples.

In this paper, a randomized cross-validation method is used to achieve experiment. First, we randomly and respectively select 50 classes from Class M and T, that is 300 samples. And the remaining samples are used to construct the over-complete dictionary. Then calculate the sparse solution, and finally use the proposed classification method based on sparse representation to classify and calculate classification accuracy. The above procedure will be repeated for 15 times to calculate the average accuracy. We repeat the randomized cross-validation for 15 times because there are totally 3900 samples, and each time we randomly select 300 as the test samples, and 3900/300 = 13. When the randomized cross-validation is achieved for 15 times, we believe that every sample have the opportunity to participate in the experiment as a test sample.

Combining the proposed non-negative constraint P/A and the classification strategy S/M, we can form different classification methods, such as “P+S”, “P+M”, “A+S” and “A+M”. We compared our experiments proposed method with Ma Yi’s method that minimizes the residual [15], SVM and ANN. Table 1 only shows the experimental results comparing the proposed method with Ma Yi’s method in [15].

TABLE I.
RESULT OF SPARSE REPRESENTATION BASED CLASSIFICATION

%	[15]	The proposed method				average
		P + S	A + S	P + M	A + M	
average	85.65	89.78	89.76	86.26	85.07	87.31
M	94.76	96.91	97.84	89.89	88.13	93.51
T	76.53	82.64	81.68	82.63	82.01	81.10

In Table 1, the first row is average accuracy. The second row and the third row respectively represents the accuracy of Class M and T. The average accuracy A_{avg}, accuracy of class M, denoted as A_M and accuracy of class T, denoted as A_T satisfy: $A_{avg}=(A_M+A_T)/2$.

As can be seen from Table 1, the proposed classification method has the highest accuracy of 89.78%. And MA Yi’s classification methods can only achieve a accuracy of 85.65%. In addition, we use the ANN and

SVM classification and respectively achieve a accuracy of 87.86% and 85.69%.

From Table 1, we also found a strange phenomenon. For all the classifiers, accuracy for M class is much higher than T class. We speculate that the reason may be that the number of samples of Class M is much more than Class T. So we achieve another experiments over balanced data set.

B. Balanced Data Set

In this experiment, we randomly and respectively select 50 classes from class M and T as test set. But we randomly and respectively select 500 classes from the remaining samples of class M and T as training set. So that class M and T have the same size of data set. We still achieve 15 randomized cross-validation, and the results of this experimental is shown in Table 2.

In Table 2, we add the results of unbalanced set (N) to compare with balance for the set (E). We can see that the accuracy of class M decreases, but the accuracy of class T increases a lot. So that the average accuracy of classification improves greatly, and the highest accuracy even reached 91.92%.

TABLE II.
RESULT OF SPARSE REPRESENTATION BASED CLASSIFICATION
(BALANCED DATA SET)

%		[15]	P + S	A + S	P + M	A + M	avg
avg	E	85.65	89.78	89.76	86.26	85.07	87.31
	N	87.21	91.92	91.05	87.35	86.95	88.94
M	E	94.76	96.91	97.84	89.89	88.13	93.51
	N	92.83	94.80	93.46	88.40	87.86	91.47
T	E	76.53	82.64	81.68	82.63	82.01	81.10
	N	81.58	89.05	88.63	86.30	86.03	86.41

V. CONCLUSION

In this paper, we propose two types of non-negative constraint strategy and two classification strategy based on sparse representation to classify thrombosis and pectinate muscles. Experimental results show that the proposed classification strategy can achieve classification accuracy up to 91.92%, higher than the other classifiers. In conclusion, the proposed classification method based on sparse representation has obvious advantages, and can be directly applied to computer-aided diagnosis.

ACKNOWLEDGMENT

In this paper, the research was sponsored by the LangFang Science and Technology Bureau Fund Project (Project No. 2013011003), the LangFang Science and Technology Bureau Fund Project (Project No. 2013011005) and LangFang Science and Technology Bureau Fund Project (Project No. 2011011006).

REFERENCES

[1] Sarkar N, Chaudhuri B B. An efficient differential box-counting approach to compute fractal dimension of image[J]. Systems, Man and Cybernetics, IEEE Transactions on, 1994, 24(1): 115-120.
[2] Cross G R, Jain A K. Markov random field texture models[J]. Pattern Analysis and Machine Intelligence, IEEE Transactions on, 1983 (1): 25-39.
[3] Manjunath B S, Ma W Y. Texture features for browsing and retrieval of image data[J]. Pattern Analysis and Machine Intelligence, IEEE Transactions on, 1996, 18(8): 837-842.
[4] Chang T, Kuo C C J. Texture analysis and classification with tree-structured wavelet transform[J]. Image Processing, IEEE Transactions on, 1993, 2(4): 429-441.
[5] Cortes C, Vapnik V. Support-vector networks[J]. Machine learning, 1995, 20(3): 273-297.
[6] Rumelhart D E, Hintont G E, Williams R J. Learning representations by back-propagating errors[J]. Nature, 1986, 323(6088): 533-536.
[7] Haralick R M, Shanmugam K, Dinstein I H. Textural features for image classification[J]. Systems, Man and Cybernetics, IEEE Transactions on, 1973 (6): 610-621.
[8] D. Donoho and M. Elad, "Optimal sparse representation in general (nonorthogonal) dictionaries via `1 minimization," Proceedings of the National Academy of Sciences of the United States of America, pp. 2197–2202, March 2003.
[9] E. Amaldi and V. Kann, "On the approximability of minimizing nonzero variables or unsatisfied relations in linear systems," Theoretical Computer Science, vol. 209, pp. 237–260, 1998.
[10] D. Donoho, "For most large underdetermined systems of linear equations the minimal l1-norm solution is also the sparsest solution," Comm. on Pure and Applied Math, vol. 59, no. 6, pp. 797–829, 2006.
[11] E. Cand`es, J. Romberg, and T. Tao, "Stable signal recovery from incomplete and inaccurate measurements," Comm. on Pure and Applied Math, vol. 59, no. 8, pp. 1207–1223, 2006.
[12] E. Cand`es and T. Tao, "Near-optimal signal recovery from random projections: Universal encoding strategies?" IEEE Trans. Information Theory, vol. 52, no. 12, pp. 5406–5425, 2006.
[13] S. Chen, D. Donoho, and M. Saunders, "Atomic decomposition by basis pursuit," SIAM Review, vol. 43, no. 1, pp. 129–159, 2001.
[14] E. T. Hale, W. Yin, and Y. Zhang, Fixed-point continuation for L1-minimization: Methodology and convergence, SIAM J. Optimiz., 19 (2008), 1107-1130.
[15] Wright J, Yang A Y, Ganesh A, et al. Robust face recognition via sparse representation[J]. Pattern Analysis and Machine Intelligence, IEEE Transactions on, 2009, 31(2): 210-227.
[16] LiYing Lang and XueKe Jing, "Application of Non-negative sparse matrix factorization in occluded face recognition," Journal of Computers, vol. 6, no. 12, pp. 2675-2679, 2011.
[17] Shuangcheng Deng, Yunhua Li, Lipei Jiang, Yingyu Cao, and Junwen Zhang, "Surface Reconstruction from Sparse & Arbitrarily Oriented Contours in Freehand 3D Ultrasound," Journal of Computers, vol. 8, no. 5, pp. 1279-1285, 2013.
[18] Zhe Liu, Xiaoxian Zhen, and Cong Ma, "A Novel Image Reconstruction Algorithm Based on Concatenated Dictionary," Journal of Computers, vol. 7, no. 2, pp. 444-449, 2012.

Xiaofang Hou from Xiaoyi City Shanxi Province, was born in December, 1980. From 2000 to 2004, she majored in computer science and technology in Shanxi Normal University. And from 2004 to 2007, she pursued her master's degree in Taiyuan University of Technology, with the major of computer application technology. Her research field covers image processing and information security.

From June 2007, she has been a teacher in Department of Computer Science and Engineering in North China Institute of Aerospace Engineering. In 2010, she attained the lecturer title. And now, she has been engaged in the research of image processing and digital watermarking.

Penghua Zhu from Handan City Hebei Province, was born in September, 1980. From 1999 to 2003, he majored in computer science and technology in Hebei University of Technology. And from 2007 to 2011, he pursued his master's degree in engineering in Graduate University of Chinese Academy of Sciences, with the major of computer technology. His research field covers computer network and information security.

Yanxin Ma from 2006 to 2010, he majored in computer science and technology in Harbin Institute of Technology. And from 2010 to 2013, he pursued his master's degree in School of Computer Science and Technology, with the major of computer technology. His research field covers pattern identification.