Medical Images Modality Classification using Multi-scale Dictionary Learning

M. Srinivas
Computer Science and Engineering
Indian Institute of Technology Hyderabad
Hyderabad, India
Email: cs10p002@iith.ac.in

C. Krishna Mohan
Computer Science and Engineering
Indian Institute of Technology Hyderabad
Hyderabad, India
Email: ckm@iith.ac.in

***Abstract*—In this paper, we proposed a method for classification of medical images captured by different sensors (modalities) based on multi-scale wavelet representation using dictionary learning. Wavelet features extracted from an image provide discrimination useful for classification of medical images, namely, diffusion tensor imaging (DTI), magnetic resonance imaging (MRI), magnetic resonance angiography (MRA) and functional magnetic resonance imaging (FRMI). The ability of On-line dictionary learning (ODL) to achieve sparse representation of an image is exploited to develop dictionaries for each class using multi-scale representation (wavelets) feature. An experimental analysis performed on a set of images from the ICBM medical database demonstrates efficacy of the proposed method.**

Keywords*—*Multi-scale Dictionary Learning, Medical X-ray image, MRI, MRA, FMRA, DTI, Multi-scale representation, Sparse representation, ODL, Wavelet.

I. Introduction

Modern medical diagnostic techniques like radiology, histopathology and computerized tomography generate a lot of medical images that need to be indexed, archived and stored for future use. The medical image classification systems available today classify medical images based on modality, body part, disease or orientation. The enormous amount of medical images with a wide variety of image modalities such as diffusion tensor imaging (DTI), magnetic resonance imaging (MRI), magnetic resonance angiography (MRA) and functional magnetic resonance imaging (FRMI) are available on medical databases. Effectively and efficiently searching and retrieving of medical image data in these different modality image collections poses significant technical challenges as the characteristics of the medical images differ from other general purpose images. Solving this problem with traditional text based image retrieval (TBIR) approach has many practical limitations [1] like the images in the collection have to be annotated manually which becomes very difficult as the size of the image collection increases and time consuming. Another important limitation of TBIR is inadequacy in representing the image content [2]. Content based image retrieval approaches were proposed by [3] to overcome the limitations of text based image retrieval. Content Based Image Retrieval (CBIR) gives a way of searching similar images in a large image repository on the basis of their visual content. When applied for medical images, CBIR can be used to retrieve images of similar nature (like same modality and disease) and characteristics and this process is known as Content Based Medical Image Retrieval (CBMIR).

Medical image classification is an important task in CBMIR. Automatic medical image classification is a technique for assigning a medical image to an appropriate class among a number of medical image classes. In medical image classification, several methods and algorithms have been presented in the literature [4]-[6]. One approach to content based medical image retrieval is proposed in [4], in which medical images are classified based on body orientation, biological system, anatomical region and image modality. The performance of the classification is evaluated on IRMA database and the best classification result is achieved by using distorted tangent distance in a kernel density classifier. The CBMIR system can achieve better performance by filtering out the images of irrelevant classes from the medical database because it reduces the search space and time for retrieving similar type of images. This establishes the importance of image classification in a CBMIR system. In literature, it has been suggested that modality is one of the most important filters that can limit the search and retrieval time [7].

Content based medical image classification (CBMIC) overcomes the need for manual annotation and human perception. Also, finding similar images in large volumes of medical image databases is a difficult task. Modality based classification of medical images enables the efficient retrieval of relevant images from the large database and reduces the search space and time. Multimodality during capturing images suffers from significant contrast variation between the images of the same scene. Due to this large variation, existing image classification and retrieval algorithms do not perform well for different type of modality images.

Selection of features for adequately representing the class specific information is an important process in medical image classification. The classification performance mostly depends on the extracted features. Commonly, there exists a semantic gap between the content of an image and its visual features. Thus, decreasing the semantic gap through extracting more effective features has still remained as a challenging topic in content based image classification and retrieval task. With the help of high level features to overcome the semantic gap between low level and high level features [8]. Features extracted from sub-bands in a multi-resolution space are useful for extracting some high level features. And capturing images of various modalities suffers from significant contrast variation between the images of the same organ or body part. Due to this large variation, existing image classification and retrieval algorithms do not perform well for different modality images. In this paper, we propose a new classification technique, namely, sparse representation based multi-scale dictionary learning to classify the different type of modality images. Multi scale image representation can handle the semantic gap and intensity variations of the different modality images.

An X-ray image categorization and retrieval method using patch-based visual word representations is proposed in [9]. The feature extraction process is based on local patch representation of the image content and a bag-of-features approach for defining image categories, with a kernel based SVM classifier. The method is especially effective in discriminating orientation and body regions in X-ray images, and in medical visual retrieval. Modality classification and its use in text based image retrieval in medical databases is proposed in [10]. Visual descriptors and text features are used for classifying the medical images. Medical image classification is then done with the help of support vector machines classifier. In [11], explore different type of medical image modality and retrieval strategies. Bags of visual words and fisher vectors representations are integrated to perform medical modality classification. Wavelet optimization techniques for content based image retrieval in medical database are described in [12]. In [13], multiple features are used for medical image indexing and retrieval. In this approach, combines the edge and patch based feature extraction methods. And based on similarity measure retrieve similar type of images. The following subsection describes the feature extraction using wavelet packet decomposition.

A. Feature Extraction

The performance of a content based image classification system depends on the representation of an image as a feature vector. Generally, content based image classification techniques use fundamental visual features like images color, shape and texture yielding vectors with thousands of features. But using these features directly, one cannot classify images easily. In the proposed method, multi-scale wavelet packet decomposition based feature extraction method is used. Wavelet packet decomposition can be implemented by progressively applying two channel filter banks. At every stage each filter bank comprises of a low-pass (L) and a high-pass (H) filter and whose sampling frequency is half of that of the previous stage. As a consequence, the original image can be decomposed into four sub-images, namely, both horizontal and vertical directions have low-frequencies (LL), the horizontal direction has low frequencies and the vertical one has high-frequencies (LH), the horizontal direction has high frequencies and the vertical one has low frequencies (HL) and both horizontal and vertical directions have high-frequencies (HH) sub-images. Next, construct a gradient vector for each sub-image. Similar approach applied for the entire training and testing database images to form a feature vector.

The procedure for feature extraction is as follows:
1. Applying a wavelet packet decomposition on an original image to obtain the LL, LH, HL and HH sub-images.
2. Construct a gradient vector for each sub-image.
3. repeat (1) and (2) steps for all original training and testing images to form a gradient feature vector
4. Combine the similar sub-bands(e.g. LL)from all the images of each class to form a feature vector matrix. This will generate four feature vector matrices for the four sub-bands for each class. The following subsection describes introduction about sparse representation.

B. Sparse representation

Sparse representation has received a lot of attention from the research in signal and image processing. Sparse coding involves the representation of an image as a linear combination of some atoms in a dictionary [14]. It is a powerful tool for efficiently representing data. This is mainly due to the fact that signals and images of interest tend to enjoy the property of being sparse in some dictionary. These dictionaries are often learned directly from the wavelet coefficients of training data. Several algorithms like On-Line Dictionary Learning (ODL) [15], K-SVD [16] and Method of Optimal Directions (MOD) [17] have been developed to process training data. Sparse representation is used to match the input query image with the appropriate class. Linear discriminant analysis (LDA) based selection and feature extraction algorithm for classification using wavelet packet has been proposed by Etemand and Chellappa [18]. Recently, similar algorithms for simultaneous sparse signal representation and discrimination have also been proposed [19],[20]. In [21], a method for simultaneously learning a set of dictionaries that optimally represent each cluster is proposed. This approach was later extended by adding a block incoherence term in their optimization problem to improve the accuracy of sparse coding. Multi-scale dictionary learning is proposed in [22]. It combines the advantages of generic multi-scale representations with the K-SVD dictionary learning method.

In this paper, we propose a modality based classification method for International Consortium for Brain Mapping (ICBM) database [23] using wavelet based on-line dictionary learning approach. Learned dictionaries are used to represent datasets in sparse model of ICBM medical images. Dictionaries are designed to represent each class. For a given N number of classes, we design N dictionaries to represent the classes. Each image associated with a dictionary provides the best sparsest representation. For every image in the given set of images $\{\mathbf{y}_i\}_{i=1}^{n}$, ODL is used to seek the dictionary D that has the sparsest representation for the image. We define $l(\hat{\mathbf{D}}, \hat{\Phi})$ as the optimal value of the l_1-lasso sparse coding problem [24]. This is accomplished by solving the following optimization problem:

$$l(\hat{\mathbf{D}}, \hat{\Phi}) = \arg\min_{\mathbf{D},\Phi} \frac{1}{N} \sum_{i=1}^{N} \frac{1}{2} \|\mathbf{Y}_i - \mathbf{D}\Phi_i\|_2^2$$

$$\text{subject to } \|\Phi_i\|_1 \leq \lambda \tag{1}$$

where Y is the matrix whose columns are y_i and λ is the sparsity parameter. D denotes the learned dictionary, Φ represents the sparse representation vectors, N denotes the number of classes and Y represents the training database. The ODL algorithm alternates between sparse coding and dictionary update steps. Several efficient pursuit algorithms have been proposed in the literature for sparse coding [17],[25]. The simplest one is the l_1-lasso algorithm [24]. Main advantage with ODL algorithm is its computational speed as it uses l_1-lasso algorithm for sparse representation. In sparse coding step, dictionary D is fixed and representation vectors $\mathbf{\Phi_i}$ are identified for each example y_i. Then, the dictionary is updated atom by atom in an efficient way.

The rest of the paper is organized as follows. Section 2 presents the proposed method. Experiments of modality based medical image classification application using sparse representation are discussed in detail in section 3. Finally, we draw the conclusions in section 4.

II. Medical Image Classification Using Sparse Representation and ODL Algorithm

The present work provides a method for medical image classification using the framework of multi-scale dictionary learning. There are many advantages to this approach. Firstly, the feature extracted from sub-bands in a multi-resolution space are useful for extracting some high level features. With the help of high level features to overcome the semantic gap. Secondly, the entire dataset is represented with the help of fixed small size of dictionary which greatly reduces computational time. Moreover, performance improves because of the uniform dictionary size irrespective of number of training images. The following subsection describes sparsest representation based classification method.

A. Sparsity based medical image classification

In this proposed method, we introduce a sparsity based medical image classification by representing the test data as a sparse linear combination of training data from a dictionary. In this paper, each class $C_i = [c_{ib1}, \ldots, c_{ib4}]$ (each class contains 4 sub-bands feature vector matrices i.e. for class $C_1 = [c_{1b1}, c_{1b2}, c_{1b3}, c_{1b4}]$) consists of all classes training samples collected directly from the wavelet coefficient of same sub-bands. In the proposed sparsity model, images belonging to the same class are assumed to lie approximately in a low dimensional subspace. Given N training classes, the p^{th} class has K_p training images $\{\mathbf{y}_i^N\}$ i=1,..., K_p. Let r be an image belonging to the p^{th} class, then it is represented as a linear combination of these training samples:

$$r = \mathbf{D}^p \Phi^p, \tag{2}$$

where D^p is $m \times K^p$ a dictionary whose columns are the training samples in the p^{th} class and Φ^p is a sparse vector.

Proposed method consists of two steps:

1) *Dictionary Construction:* In the wavelet packet decomposition domain contains a collection of coefficient images or sub-bands. The different wavelet coefficient images contain data are different scales and orientations. As such it makes sense that separate dictionaries be used to represent these images. Construct the dictionary for each sub-band of class (D_{ib}) (where i is the number of classes i.e. i=1,..,4 and b is number of sub-bands in each class i.e. b=1,...,4) using on-line dictionary learning algorithm [15].

Then, the dictionaries for all training class on same sub-band is $D_b = [D_{1b}, \ldots, D_{4b}]$ (if b=1, then D_{4b} means fourth class and first sub-band dictionary) and computed using the equation:

$$(\hat{\mathbf{D}}_i, \hat{\Phi}_i) = \arg \min_{\mathbf{D}_i, \Phi_i} \frac{1}{N} \sum_{b=1}^{4} \sum_{i=1}^{N} \frac{1}{2} \|\mathbf{C}_{ib} - \mathbf{D}_{ib}\Phi_{ib}\|_2^2 + \lambda \|\Phi_{ib}\|_1, \tag{3}$$

satisfying $\mathbf{C}_i = \hat{\mathbf{D}}_i \hat{\Phi}_i, \quad i = 1, 2, \ldots, N.$

2) *Classification:* In this classification process, the sparse vector Φ for given test image is found in the test dataset $Z = [z_1, \ldots, z_l]$. Using the dictionaries of training samples of each class on same sub-band is $D_b = [D_{1b}, \ldots, D_{4b}]$, the sparse representation Φ satisfying $D_b\Phi$=Z is obtained by solving the following optimization problem:

$$\Phi^l = \arg\min_{\Phi} \textstyle\sum_{b=1}^{4} \frac{1}{2} \|\mathbf{z}_{lb} - \mathbf{D}_b \Phi_{bl}\|_2^2 \quad \text{subject to} \|\Phi_l\|_1 \leq T_1,$$

$$\text{and}$$

$$\hat{i} = \arg\min_i \|\mathbf{z}_l - \mathbf{D}\delta_i(\mathbf{\Phi^l})\|_\mathbf{2}^\mathbf{2} \;\; l = \mathbf{1}, \cdots, t, \tag{4}$$

where δ_i is a characteristic function that selects the coefficients. Then z_l is assigned to C_i associated with the i^{th} dictionary. It means, finding the sparsest dictionary for a given test data using l_1 -lasso algorithm. Then, test data is assigned to the class associated with this sparsest dictionary.

In the classification phase, each sub-image acquired from the test image is matched with the trained dictionaries of only that sub-image. The class which yields maximum sparsity is chosen as the class for that sub-band. Once all the sub-images are evaluated, the class which agrees with the majority of the sub-bands is chosen as the category for the test image.

III. Experimental Results

In this section, we show the effectiveness of the proposed modality based medical image classification method using multi-scale dictionary learning and sparse representation. Data used in the preparation of this work were obtained from the International Consortium for Brain Mapping (ICBM) database (www.loni.usc.edu/ICBM). The ICBM project (Principal Investigator John Mazziotta, M.D., University of California, Los Angeles) is supported by the National Institute of Biomedical Imaging and BioEngineering. ICBM is the result of efforts of co-investigators from UCLA, Montreal Neurologic Institute, University of Texas at San Antonio, and the Institute of Medicine, Juelich/Heinrich Heine University - Germany.

Experiments are carried out on ICBM medical database, in which each image is of size 200×200 pixels. Majority of medical images are generally gray scale images such as X-ray, FMRI, MRI etc. The main problem in classifying medical radiological images is high inter class overlap and intra class variability in some of the classes [2]. For tackling this problem, wavelet packet decomposition based feature extraction method is used to overcome semantic gap between low level features and high level features. Moreover, the proposed method works for images with various sensors. ICBM database consisting of a four different type of image modalities such as Diffusion Tensor Imaging (DTI), magnetic resonance imaging (MRI), Magnetic resonance angiography (MRA) and functional magnetic resonance imaging (FRMI) are available. Entire database images are divided into 70% training and 30% testing images for each class and experiments are run through 5-fold cross validation. Each class consists of 5587 training and 1482 testing images. Proposed method tested with various wavelet families, namely, Harr, Daubechies, Coiflets, Symlets, Discrete Meyer, and Biorthogonal. The best results obtained from these experiments are presented in Table I. The proposed method was tested with dictionaries of size 60, 80 and 100. Generally, accuracy improves for larger sized dictionaries. However, after a certain point, increase in dictionary size does not yield better classification accuracy. The dictionary size at this point of time gives the best possible sparse representation of the given feature descriptor. In our case, recognition rate of 91.6% was obtained for dictionary size of 80.

Table. I Classification accuracy (%) of proposed method using wavelet decomposition based features and different dictionary sizes.

Wavelet Families	**60 Dict**	**80 Dict**	**100 Dict**
Daubechies(db4)	86.3	86.9	85.6
Daubechies(db10)	85.6	85.9	84.4
Harr(db2)	90.3	**91.6**	90.7
Discrete Meyer	86.7	86.2	86.6
Coiflets	87.2	87	86.8
Symlets2	90.2	89.5	87
Biorthogonal	87	87.8	87

The confusion matrices for proposed, SVM, KNN and Bayesian classification method with highest accuracy results using Haar wavelet are shown in Figure 1-4 respectively.

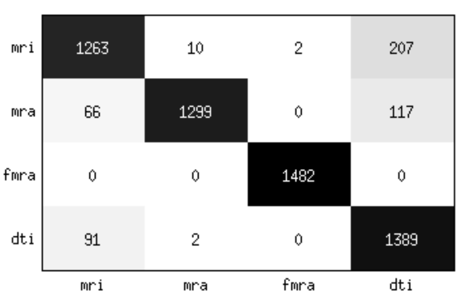

Fig. 1. Confusion matrix for proposed method with haar wavelet feature.

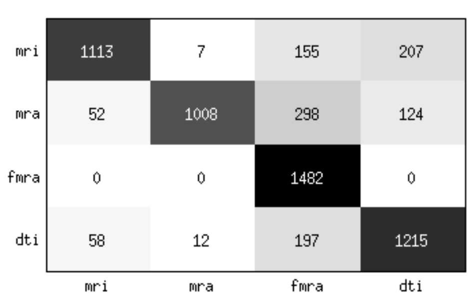

Fig. 2. Confusion matrix using SVM classification method with haar wavelet feature.

The proposed method gives best classification results of 91.6% as compared to other image classification techniques

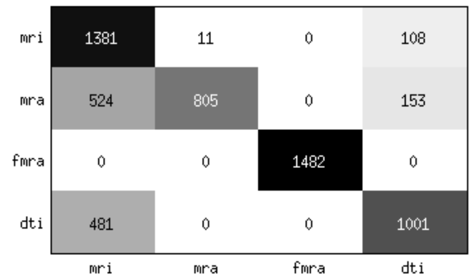

Fig. 3. Confusion matrix using Neural Network (BP) classification method with haar wavelet feature.

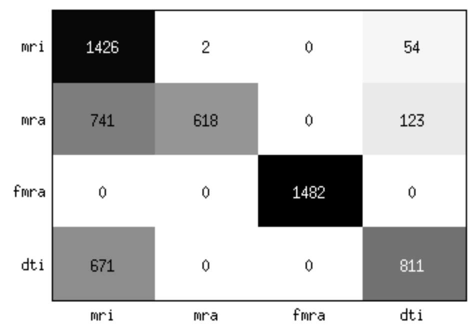

Fig. 4. Confusion matrix using Bayesian classification method with haar wavelet feature.

such as SVM, Neural Network (BP), and Bayes Classifier (BC). The classification performance of different classifiers are shown in Table II.

Table. II Performance of the proposed method with different classifiers on same dataset.

Classifier	**Accuracy (%)**
SVM	81.2
Neural Network(BP)	78.3
Bayesian	73.1
Proposed	**91.6**

Wavelet packet decomposition generates gradient vectors individually for each of the four sub-bands. Although distinct, these gradient vectors by themselves do not have enough discriminative capabilities. Using different combinations of the gradient vectors may yield different discriminating characteristics [26].

Classification accuracy of different possible combinations of the gradient vectors extracted from the four sub-bands are presented in Table III. We may notice that LL sub band contains more information among the four sub-bands. The classification accuracy based on the gradient vectors extracted from the LL sub-band is 84.3% The classification accuracy based on the gradient vectors extracted from the LH, HL, and HH sub-bands were 73.4, 70.2, and 73.8 %, respectively. To increase the classification accuracy, we can combine all sub-bands sparsity results. Various combination sequences were tried and best classification accuracy of 91.6% was achieved after combining the dictionaries from all the sub-bands. And based on majority of the sparsity results classify the testing images. With the combination based sparsity results best classification accuracy of 91.6% can be achieved.

Table. III Classification accuracy of proposed method based on individual and all combination of the sub-bands obtained from wavelet decomposition.

Subband	**Accuracy (%)**
LL	84.3
LH	73.4
HL	70.2
HH	73.8
LL+LH+HL+HH	**91.6**

IV. Conclusion

we proposed a method for classification of medical images captured by different sensors (modalities) based on multi-scale wavelet representation using dictionary learning. We have exploited the ability of ODL to achieve sparse representation of an image, to develop dictionaries for each class using wavelet features. Other classifiers, namely, SVM, NN and Bayes were also examined. The medical images database containing four different type of modality(sensors) images, namely, diffusion tensor imaging (DTI), magnetic resonance imaging (MRI), magnetic resonance angiography (MRA) and functional magnetic resonance imaging (FRMI) was used for training and testing the models. Experimental results indicate that the wavelet packet decomposition based feature can provide useful information for discriminating the classes. Preliminary computational results are promising and have the potential for practical image classification. The proposed method has achieved best performance of 91.6%. The experimental results suggest that the proposed method is better than other well known classification algorithms.

Acknowledgement

Data collection and sharing for this project was provided by the International Consortium for Brain Mapping (ICBM; Principal Investigator: John Mazziotta, MD, PhD). ICBM funding was provided by the National Institute of Biomedical Imaging and BioEngineering. ICBM data are disseminated by the Laboratory of Neuro Imaging at the University of Southern California.

References

[1] G. Csurka, S. Clinchant, G. Jacquet. Medical Image Modality Classification and Retrieval. In Proc. IEEE Int. Conf CBMI 2011; 193-198.

[2] H. Pourghassem, H. Ghassemian. Content based medical image classification using a new hierarchical merging scheme. Computerized Medical Imaging and Graphics 2008; 32(8), 651-661.

[3] A. W. M. Smeulder, M. Worring, S. Santini, A. Gupta, R. Jain. Content based image retrieval at the end of the early years. IEEE Trans. Pattern Analysis Machine Intell 2000; 22(12):1349-1380.

[4] D. Keysers, J. Dahmen, H. Ney, B. B. Wein, T. M. Lehmann. Statistical frame work for model based image retrieval in medical applications. J Electronic Image 2003; 12(1): 59-68.

[5] M. M. Rahman, P. Bhattacharya, B. C. Desai. A frame work for medical image retrieval using machine learning and statistical similarity matching techniques with relevance feedback. IEEE Trans. Inf Technology Biomed 2007; 11(1): 58-69.

[6] H. Greenspan, A. T. Pinhas. Medical image categorization and retrieval for PACS using the GMM-KL frame work. IEEE Trans. Inf Technology Biomed 2007; 11(2): 190-202.

[7] Viktor. G. l. Etienne Kerre. Mike Nachtegael. Multiple kernel learning based modality classification for medical images. Int Conf. 2012.

[8] Esmat Rashedi, Hossein Nezamabadi-pou. Saeid Saryazdi. A simultaneous feature adaptation and feature selection method for content-based image retrieval systems. J Knowledge-Based Systems Vol. 39, pp. 8594, February 2013,

[9] U. Avni et al., X-ray image categorization and retrieval using patch-based visual words representation. In Proc. IEEE Int. Symp. Biomed. Image 2009; 1(2): 350-353.

[10] Pierre. Tirilly, Kun. Lu, Xiangming. Mu, Tian. Zhao, Yu. Cao, On modality classification and its use in text-based image retrieval in medical databases, In Proc. IEEE Int. workshop on Content-Based Multimedia Indexing (CBMI), 13-15 June 2011.

[11] Gabriela .Csurka, Stephane. Clinchant, Guillaume. Jacquet, Medical Image Modality Classification and Retrieval. In Proc. IEEE Int. workshop on Content-Based Multimedia Indexing (CBMI), 13-15 June 2011.

[12] G. Quellec, M. Lamard, G. Cazuguel, B. Cochener, C. Roux. Wavelet optimization for content-based image retrieval in medical databases. Journal. Medical Image Analysis 2010; (14): 227-241.

[13] M. Srinivas, C. Krishna Mohan. Medical Image Indexing and Retrieval Using Multiple Features. In Proceedings of CIIT conference 2013; Elsevier.

[14] I. Ramirez, P. Sprechmann, Guillermo. Spiro. Classification and Clustering via Dictionary Learning with Structured Incoherence and Shared Features. In IEEE Int. Conf. Computer Vision and Pattern Recognition (CVPR) 2010; 3501-3508.

[15] J. Mairal, F. Bach, J. Ponce, G. Sapiro. Online dictionary learning for sparse coding. In Proceedings of the 26th Annual International Conference on Machine Learning ICML 2009; 689-696.

[16] M. Aharon, M. Elad, A. M. Bruckstein. The K-SVD: An Algorithm for Designing of Over complete Dictionaries for Sparse Representation. IEEE Trans. Signal Process 2006; 54(11): 4311-4322.

[17] K. Engan, S. Aase, J. Hakon-Husoy. Method of optimal directions for frame design. In Proc. IEEE Int. Conf. Acoust., Speech, Signal Process 1999; (5): 2443-2446.

[18] K. Etemand, R. Chellapa. Separability Based Multiscale Basis Selection and Feature Extraction for Signal and Image Classification. IEEE Trans. Image Proc. 1998; 7(10): 1453-1465.

[19] F. Rodriguez, G. Sapiro, “Sparse Representation for Image Classification: Learning Discriminative and Reconstructive non Parametric Dictionaries. Tech. Report, University of Minnesota, Dec. 2007.

[20] K. Huang, S. Aviyente. Sparse representation for signal classification. NIPS 2007; (19): 609-616.

[21] P. Sprechmann, G. Spiro. Dictionary Learning and Sparse Coding for Unsupervised Clustering. In Proc. IEEE Int. Conf.Acoust., Speech, Signal Process. 2010; 2042-2045.

[22] Boaz. Ophi. Michael. Lustig. Michael. Elad. Multi-Scale Dictionary Learning Using Wavelets. IEEE Trans. Signal Process, Vol. 5, No. 5,pp.1014-1024, Sept 2011.

[23] https://ida.loni.usc.edu/collaboration/access/appLicense.jsp.

[24] H. Lee, A. Battle, R. Raina, A. Y. Ng. Efficient sparse coding algorithms. In Proc. Advances in Neural Information Processing Systems 2007; (19): 801-808.

[25] S. Mallat, Z. Zhang. Matching Pursuits with Time Frequency Dictionaries. IEEE Trans. Signal Processing 1993; 41(12): 3397-3415.

[26] P.W. Huanga. S.K. Daia. P.L. Linb. Texture image retrieval and image segmentation using composite sub-band gradient vectors. J Visual Communication and Image Representation Vol. 17, No. 5, pp. 94795, October 2006,