Pulmonary Emphysema Classification based on an Improved Texton Learning Model by Sparse Representation

Min Zhang*[a], Xiangrong Zhou[a], Satoshi Goshima[b], Huayue Chen[c], Chisako Muramatsu[a], Takeshi Hara[a], Ryujiro Yokoyama[d], Masayuki Kanematsu[b,d] and Hiroshi Fujita[a]

[a]Department of Intelligent Image Information, Division of Regeneration and Advanced Medical Sciences, Graduate School of Medicine, Gifu University, Gifu-shi, 501-1194 Japan
[b]Department of Radiology, Gifu University Hospital, Gifu-shi, 501-1194 Japan
[c]Department of Anatomy, Division of Disease Control, Graduate School of Medicine, Gifu University, Gifu-shi,501-1194 Japan
[d]Department of Radiology services, Gifu University Hospital, Gifu-shi, 501-1194 Japan

ABSTRACT

In this paper, we present a texture classification method based on texton learned via sparse representation (SR) with new feature histogram maps in the classification of emphysema. First, an overcomplete dictionary of textons is learned via K-SVD learning on every class image patches in the training dataset. In this stage, high-pass filter is introduced to exclude patches in smooth area to speed up the dictionary learning process. Second, 3D joint-SR coefficients and intensity histograms of the test images are used for characterizing regions of interest (ROIs) instead of conventional feature histograms constructed from SR coefficients of the test images over the dictionary. Classification is then performed using a classifier with distance as a histogram dissimilarity measure. Four hundreds and seventy annotated ROIs extracted from 14 test subjects, including 6 paraseptal emphysema (PSE) subjects, 5 centrilobular emphysema (CLE) subjects and 3 panlobular emphysema (PLE) subjects, are used to evaluate the effectiveness and robustness of the proposed method. The proposed method is tested on 167 PSE, 240 CLE and 63 PLE ROIs consisting of mild, moderate and severe pulmonary emphysema. The accuracy of the proposed system is around 74%, 88% and 89% for PSE, CLE and PLE, respectively.

Keywords: emphysema, computed tomography (CT), texture classification, texton, sparse representation, dictionary learning

1. INTRODUCTION

Chronic obstructive pulmonary disease (COPD) is a disease of pulmonary system and is a growing health problem worldwide. Coughing up mucus is an early sign of COPD. The term chronic obstructive pulmonary disease or COPD is often used to describe patients who have chronic and largely irreversible airways obstruction, most commonly associated with some combination of emphysema and chronic bronchitis. Emphysema is one of the important kinds of COPD. It is defined histologically as permanent enlargement of the airspace distal to the terminal bronchioles and destruction of the alveolar wall (See Fig.1 for illustration). Areas of the lung that are affected by emphysema have reduced CT attenuation coefficients. Emphysema may be further characterized as centrilobular emphysema (CLE), panlobular emphysema (PLE) and paraseptal emphysema (PSE). CLE affects the lobules around the central respiratory bronchioles and is the most common type of smoking-related emphysemas. CLE is typically found in the upper lung zone. It is depicted on CT images as a low-attenuation area surrounded by normal lung attenuation. CLE predominates in older subjects, whereas PSE predominated in younger subjects. PLE uniformly affects the entire secondary lobules. It appears as a generalized decrease in CT attenuation, predominantly in the lower lobe. The vessels in affected regions of the lung are reduced in number and caliber. Although PLE is typically associated with α1-antitrypsin deficiency, it also may be seen in severe smoking-related emphysemas.
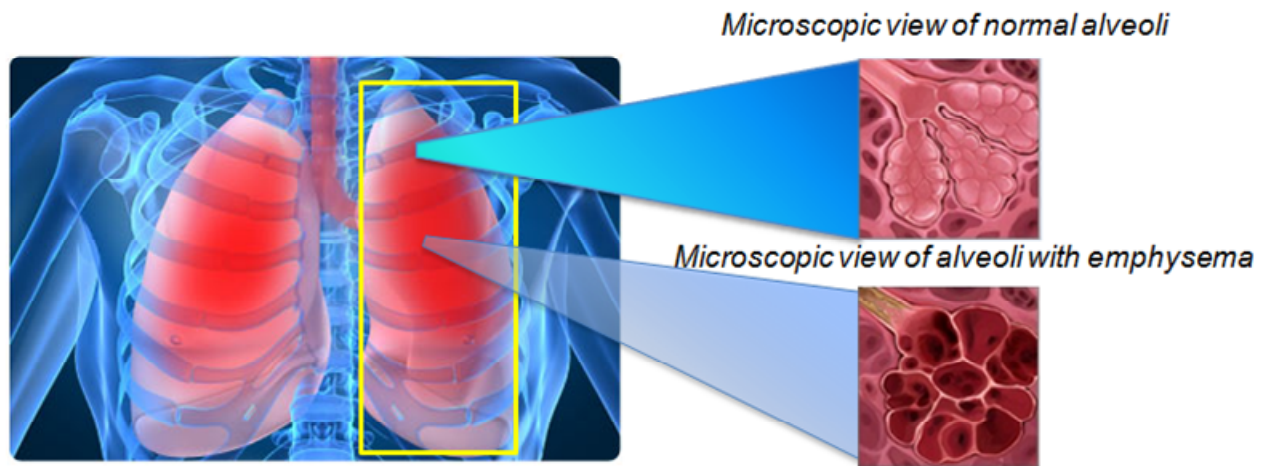
*min@fjt.info.gifu-u.ac.jp; zhangmin@mailst.xjtu.edu.cn

Figure 1. Ilustration of the microscopic view of normal alveoli and alveoli with emphysema [1,2].

Pulmonary function tests (PFTs) are used for the evaluation of patients with suspected COPD. In addition, PFTs are used to determine the severity of the airflow limitation, to assess the response to medications, and to follow up disease progression. Another common measure is the relative area of emphysema (RA), which measures the relative amount of lung parenchyma pixels that have attenuation values below a certain threshold. RA method relies on a single intensity threshold on individual pixels, thus ignoring any interrelations between pixels. RA of emphysema cannot be capable of detecting early stages of emphysema and relies on a hand-picked parameter sometimes [3].

Computerized diagnosis and quantification of emphysema are very important. Clinically, high resolution computed tomography (HRCT) can provide an accurate assessment of lung tissue patterns. Three-dimensional (3D) image of the pulmonary volumes produced by HRCT can avoid the superposition of anatomic structures, and is suitable for the assessment of lung tissue texture. HRCT is a reliable tool for demonstrating the pathology of emphysema, even in subtle changes within secondary pulmonary lobules. Another traditional method is based on the histogram of CT attenuation values. The quantitative measures of the degree of emphysema are derived from this histogram [4]. Another way to objectively characterize the emphysema morphology is to describe the local image structure using texture analysis techniques. An adaptive multiple feature method (AMFM) for examining the lung parenchyma from HRCT scans is proposed [5]. The AMFM is a texture-based method that combines statistical texture measures with a fractal measure. Seventeen measures of texture were used, including the grey level distribution measures, run-length measures, co-occurrence matrix measures, and a geometric fractal dimension (GFD). The optimal feature selection is performed by using the divergence measure along with correlation analysis and the classification is performed by a bayesian approach. Then this method is extended from a 2D texture based tissue classification method (AMFM) to 3D texture based method [6]. In this way, 3D AMFM provides a means of discriminating subtle differences between smokers and nonsmokers both with normal PFTs. Recently, small-sized local operators, such as local binary patterns (LBPs) and patch representation in texton-based approaches are also introduced for classification and quantification of COPD [7,8]. Small-sized local operators are especially desirable in situations where the region of interest (ROI) is rather small, which is often the case in texture analysis in medical imaging, where pathology can be localized in small areas.

In our previous work [9], we used a texture classification method based on textons learned via sparse representation in the classification of emphysema. This method is inspired by the great success of l_1-norm minimization based sparse representation (SR). The dictionary of textons is learned by applying SR to image patches in the training dataset. The SR coefficients of the test images over the dictionary are used to construct the histograms for texture classification.

However textons learned in dictionary learning stage are not sensitive to intensity difference. When dealing with CT images, intensity is of a physical property of the tissue which may be related to pathological pattern. Hence, intensity information should be included in the feature histogram. In this paper, 3D joint SR-coefficient-and-intensity histogram maps are proposed as the features used in the classification of emphysema. The joint SR coefficient and intensity histogram map capture information about at which intensity levels the different textons reside to improve discrimination

results. With the proposed method, a texton training dataset is first constructed from descriptors in the training images, and then an over-complete dictionary of textons is computed under the SR framework. A histogram map of joint SR-coefficient-and-intensity can be extracted for texture classification by coding the texture image with the texton dictionary and the image intensity. The test subjects, including mild, moderate and severe emphysema with three different subtypes, are used to evaluate the effectiveness and robustness of the proposed method. Annotated ROIs extracted from 14 test subjects, including 6 PSE subjects, 5 CLE subjects and 3 PLE subjects of different stages, are used to evaluate the effectiveness and robustness of the proposed method.

2.   METHOD

2.1 Related work

Recently, the theory and algorithms of sparse coding or sparse representation (SR) [10, 11] have been successfully used in image processing and pattern recognition [12-14]. The principle of SR reveals that a given natural signal can be often sparsely represented as the linear combination of an over-complete dictionary. The success of SR largely owes to the fact that natural signals are intrinsically sparse in some domain.

Sparseness is one of the reasons for the extensive use of popular transforms such as the discrete fourier transform (DFT), the wavelet transform and the singular value decomposition (SVD). The aim of these transforms is often to reveal certain structures of a signal and to represent these structures in a compact and sparse representation. However, wavelet transform is limited in characterizing the various local structures in natural images. In order to overcome the shortcomings of wavelet transform, more advanced multi-scale transforms such as Curvelet [15], Contourlet [16] and Bandlet [17] are developed. Although these transforms can offer multi-scale and multi-direction representations for natural images, these representations are not adaptive to the image contents. They can only handle some specific classes of images optimally such as piece-wise smooth images, while natural images often have sharp edge structures. If a redundant dictionary of bases can be learned from example images, a good representation of the input signal can be expected by using this redundant dictionary. SR and the associated dictionary learning techniques can be employed to this end. Various dictionary learning methods have been proposed such as the K-SVD [11] and dual Lagrange methods [18].

With K-SVD, for a given signal $x \in R^m$, we say that x has a sparse approximation over a dictionary $D = [d_1, d_2, ..., d_l] \in R^{m \times l}$, if we can find a linear combination of only "a few" atoms from D that is "close" enough to the signal x. Under this assumption, the sparsest representation of x over D is the solution of

$$\min_\alpha \|\alpha\|_0 \, s.t. \|x - D\alpha\|_2^2 \leq \varepsilon ,\tag{1}$$

where $\|\alpha\|_0$ represents the number of non-zeros in the coding vector α. The process of solving the above optimization problem is commonly referred to as "sparse coding." However, the l_0-norm sparse coding is a non-convex and NP-hard problem, and approximate solutions are often found by algorithms such as matching pursuit (MP) [19] and orthogonal matching pursuit (OMP) [20].

The selection of dictionary D plays an important role in sparse coding. The wavelet, curvelet and contourlet bases can be used as the dictionary to represent signals. However, these analytically designed universal dictionaries are too generic to be effective enough for a specific task, such as texture representation. In many applications of computer vision and pattern classification, we may have training samples, and a more effective dictionary D can be learned from a training dataset, denoted by $X = [x_1, x_2, ..., x_n] \in R^{m \times n}$. It is expected that each training sample (i.e., each column of X) can be sparsely and faithfully represented over the dictionary D, i.e., $x_i \approx D\alpha_i$ and only a few elements in α_i are significant. Sparse representations account for most or all information of a signal with a linear combination of a small number of elementary signals called atoms. Often, the atoms are chosen from a so called over-complete dictionary. Formally, an over-complete dictionary is a collection of atoms such that the number of atoms exceeds the dimension of the signal space, so that any signal can be represented by more than one combination of different atoms.

2.2 The proposed method

The motivation of the proposed work is from the following aspects. First, from above analysis, HRCT can provide an accurate assessment of lung tissue patterns. And based upon the previous work, texture features are quite useful in the classification and quantification of emphysema in CT images. Second, recently, there is a growing interest in the use of sparse representations for signals. Texture classification inspired by dictionary learning and sparse representation has achieved great success. It shows impressive performance in the natural texture image classification. Sparsity in an overcomplete dictionary is the basis for all kinds of highly effective signal and image presentations. It is based on the suggestion that natural signals can be efficiently represented as linear combinations of prespecified atom signals with linear sparse coefficients.

The proposed method mainly includes four stages: 1) ROI image pre-processing; 2) construction of a dictionary of textons via sparse representation; 3) texton histogram learning with the training set; and 4) ROI image classification. The details of the proposed method are as follows.

1. *ROI preprocessing*

   The test images were obtained from 14 subjects, including 6 PSE subjects, 5 CLE subjects and 3 PLE subjects of different stages. Totally 470 ROIs with 40×40 pixels were extracted from these subjects.

2. *Texton learning by K-SVD*

   a) Before texton learning, all training ROI images are normalized to have zero mean and unit standard deviation. The normalization offers certain amount of invariance to the illumination changes. A square neighborhood around each pixel in the image is cropped and is stretched to a vector. For each type of ROIs, a training dataset $X = [x_1, x_2, ..., x_n] \in R^{m \times n}$ is constructed, where $x_i, i = 1, 2, \cdots, n$, is the patch vector at a position in a training sample image of this type. The dictionary of textons, denoted by $D = [d_1, d_2, ..., d_l] \in R^{m \times l}$, can be learned from the constructed training dataset X, where $d_j, j = 1, 2, \cdots, k$ is one of the k textons. In this stage, a high-pass filter is applied to exclude the patches reside in smooth regions.

   b) Then an overcomplete dictionary of textons is learned by optimizing D and $d_j, j = 1, 2, \cdots, k$ of the function below by using K-SVD [11]

   $$\min_{\alpha} \|\alpha\|_0 \ s.t. \ \|x - D\alpha\|_2^2 \leq \varepsilon,$$ (1)

   where the coding vector α is the SR objective function. In this stage, orthogonal matching pursuit (OMP) is used to update the dictionary. Figure 2 (a) shows the flowchart for the training stage.

3. *Feature extraction*

   Dictionary learning is not sensitive to intensity difference due to the normalization process in the dictionary learning stage. When dealing with CT images, intensity is of a physical property of the tissue may be related to pathological pattern. Hence, intensity information should be included in the feature histogram by forming the joint histogram between the SR coefficient and the intensity in the center pixels. So in this work, new 3D joint SR coefficient and intensity histogram maps are used as the extracted feature for the classification of emphysema. The joint SR coefficient and intensity histogram captures information about at which intensity levels the different textons reside to improve discrimination results. Figure 2 (b) illustrates the extraction of the proposed histogram map.

4. *ROI classification*

   Finally, the ROI image can be classified into the corresponding class by the selected classifier using the proposed feature histogram maps.
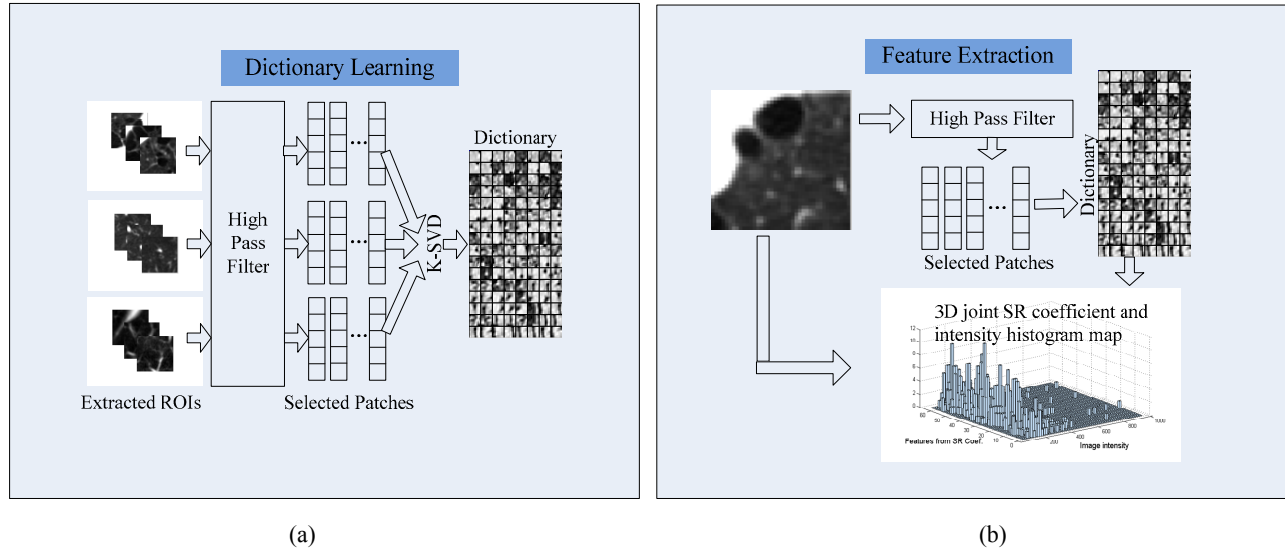
Figure 2. (a) The illustration of the dictionary learning using K-SVD and (b) the generation of 3D histogram map.

3. EXPERIMENTS AND RESULTS

3.1 Data preparation

The dictionary of texton is learned via applying SR to image patches in the training dataset. The SR coefficients of the test images over the dictionary are used to construct the histograms for texture classification. The proposed scheme was applied to 14 patient cases of non-contrast CT images. Each CT image covers the whole torso region with an isotopic spatial resolution of 0.63 [mm] and a 12 [bits] density resolution. The test images were obtained from 14 subjects with three subtypes of pulmonary emphysema. Totally 470 40×40 region of interests (ROIs) are extracted. Table 1 shows details of the test dataset with different emphysema.

Table 1. Details of test dataset.

	ROI Numbers	ROI Size	Texton Size
PSE	167 (6 subjects)		
CLE	240 (5 subjects)	40×40	8×8
PLE	63 (3 subjects)		

3.2 Experimental results and comparison

In this section, we present the results of the proposed method. In the experiments, training set was constructed only by 47 ROIs, which account for 10% of total ROIs. This training set was used for the texton learning by sparse representation to build 128 textons separately. Patch size of 8×8 are used in the experiments.

Preliminary results are shown in Table 2 and Table 3 along with the result obtained based on the proposed classification framework. The confusion matrix in Table 3 shows that the result of proposed method generally agrees with the class labels in most cases. The classification accuracy for PSE is relatively low. The discrimination between PSE and CLE needs further improvement.

Table 2. The average performance for different types of emphysema classification.

	PSE	CLE	PLE
Average accuracy (%)	73.6	87.5	88.9

Table 3. Confusion matrix showing the true label vs. label assigned by the classifier for the proposed method.

Estimated labels	True labels		
	PSE	CLE	PLE
PSE	123	25	2
CLE	38	210	5
PLE	6	5	56

4. CONCLUSION

This paper presents to use 3D joint SR coefficient and intensity histogram maps in pulmonary emphysema classification based on texton learning via sparse representation. The dictionary of texton is learned via applying K-SVD to image patches in the training dataset. The SR coefficients of the test images over the dictionary joint intensities are used to construct the 3D histogram maps for texture classification. The dataset, including 470 annotated ROIs consisting of PSE, CLE and PLE emphysema of mild, moderate and severe, is used. The performance of the proposed system obtained an accuracy of around 74%, 88% and 89% for PSE, CLE and PLE, respectively.

ACKNOWLEGEMENT

REFERENCES

[1] http://depositphotos.com/2627823/stock-photo-Lung-infection.html.
[2] http://www.modernmedicalguide.com/emphysema/.
[3] N. L. Müller, C. A. Staples, R. R. Miller, and R. T. Abboud, "Density mask. An objective method to quantitate emphysema using computed tomography, " Chest, vol.94, no.4, pp. 782–787 (1988)
[4] Shaker, et al., "Identification of patients with chronic obstructive pulmonary disease (COPD) by measurement of plasma biomarkers," The Clinical Respiratory Journal 2(1), 17–25 (2008)
[5] R Uppaluri, G McLennan, P Enright, and E A Hoffman, "Adaptive multiple feature method (AMFM) for the early detection of parenchymal pathology in smoking population," in Proc. SPIE Conf. Medical Imaging, vol. 3337, pp. 8-13 (1998).
[6] Y. Xu, M. Sonka, G. McLennan, J. Guo, and E. A. Hoffman, "MDCT based 3-D texture classification of emphysema and early smoking related lung pathologies," IEEE Trans. Med. Imag., vol. 25, no. 4, pp. 464–475, (2006).
[7] Sørensen, L., Shaker, S.B., de Bruijne, M., "Quantitative Analysis of Pulmonary Emphysema Using Local Binary Patterns." IEEE Trans. Med. Imag. 29(2), 559–569, (2010).

[8] Gangeh, M.J., Sørensen, L., Shaker, S.B., Kamel, M.S., de Bruijne, M., Loog, M., "A Texton-Based Approach for the Classification of Lung Parenchyma in CT Images." In MICCAI 2010, LNCS, vol. 6363, pp. 596--603. Springer, Heidelberg (2010).

[9] M. Zhang, X. Zhou, S. Goshima, H. Chen, C. Muramatsu, T. Hara, R. Yokoyama, M. Kanematsu and H. Fujita, "An application to pulmonary emphysema classification based on model of texton learning by sparse representation", SPIE medical imaging conference (2012)

[10] A. Bruckstein, D. Donoho, and M. Elad, "From sparse solutions of systems of equations to sparse modeling of signals and images", SIAM. Review, vol. 51, no. 3, pp. 34–81(2009).

[11] M. Aharon, M. Elad, and A. Bruckstein, "The K-SVD: an algorithm for designing of overcomplete dictionaries for sparse representation", IEEE Trans. Signal Processing, vol. 54, no. 12, pp. 4311–4322(2006).

[12] J. Yang, K. Yu, and T. Huang, "Efficient highly over-complete sparse coding using a mixture model", in ECCV, (2010).

[13] J. Wright, A. Yang, A. Ganesh, S. Sastry, and Y. Ma, "Robust face recognition via sparse representation", IEEE Trans. PAMI, vol. 31, no. 2, pp. 210–227, 2009. E. Candes, and D. Donoho, "Continuous curvelet transform: I. Resolution of the wavefront set", Applied and Computational Harmonic Analysis, vol. 19, no. 5, pp. 162–197 (2003).

[14] Xie, J., Zhang, L., You, J., Zhang, D., "Texture Classification via Patch-Based Sparse Texton Learning," In: Int'l Conf. on Image Processing (ICIP), Hong Kong, pp. 2737–2740 (2010)

[15] E. Candes, and D. Donoho, "Continuous curvelet transform: II. Discretization and frames", Applied and Computational Harmonic Analysis, vol. 19, no. 5, pp. 198–222 (2003).

[16] M. N. Do, and M. Vetterli, "The contourlet transform: an efficient directional multiresolution image representation", IEEE Trans. IP, vol. 14, no. 12, pp. 2091-2106 (2005).

[17] E. Lepennec, and S. Mallat, "Discrete bandelets with geometric orthogonal filters", in ICIP, (2005).

[18] H. Lee, A. Battle, R. Raina, and A. Ng, "Efficient sparse coding algorithms", in NIPS, pp. 801-808 (2006).

[19] S. Mallat, and S. Zhang, "Matching pursuits with time-frequency dictionaries", IEEE Trans.SP, vol. 41, no. 12, pp. 3397–3415(1993).

[20] J. Tropp, "Greed is good: algorithmic results for sparse approximation", IEEE Trans. Information Theory, vol. 50, no. 10, pp. 2231–2242(2004).