Journal of Digital Imaging

An Automatic Computer-Aided Detection Scheme for Pneumoconiosis on Digital Chest Radiographs

Peichun Yu,[1] Hao Xu,[2] Ying Zhu,[1] Chao Yang,[2] Xiwen Sun,[3] and Jun Zhao[1]

This paper presents an automatic computer-aided detection scheme on digital chest radiographs to detect pneumoconiosis. Firstly, the lung fields are segmented from a digital chest X-ray image by using the active shape model method. Then, the lung fields are subdivided into six non-overlapping regions, according to Chinese diagnosis criteria of pneumoconiosis. The multi-scale difference filter bank is applied to the chest image to enhance the details of the small opacities, and the texture features are calculated from each region of the original and the processed images, respectively. After extracting the most relevant ones from the feature sets, support vector machine classifiers are utilized to separate the samples into the normal and the abnormal sets. Finally, the final classification is performed by the chest-based report-out and the classification probability values of six regions. Experiments are conducted on randomly selected images from our chest database. Both the training and the testing sets have 300 normal and 125 pneumoconiosis cases. In the training phase, training models and weighting factors for each region are derived. We evaluate the scheme using the full feature vectors or the selected feature vectors of the testing set. The results show that the classification performances are high. Compared with the previous methods, our fully automated scheme has a higher accuracy and a more convenient interaction. The scheme is very helpful to mass screening of pneumoconiosis in clinic.

KEY WORDS: Pneumoconiosis, digital radiography, computer-aided detection (CAD), active shape model (ASM), texture analysis, support vector machine (SVM)

INTRODUCTION

Pneumoconiosis is a major occupational disease caused by inhaled free silica, asbestos, mixed dust, coal, beryllium, cobalt, and so on. Mass chest screening is a key to the early detection and cure of pneumoconiosis. Currently, chest radiograph is the most useful tool for diagnosis of pneumoconiosis. To facilitate mass chest screening, the International Labor Organization (ILO) developed the diagnosis guidance[1,2] to standardize the classification of pneumoconiosis. In the ILO classification guidance, each lung field is divided into three zones: upper, middle, and lower zone. The profusion level of small opacities can reflect the degree of pneumoconiosis. The profusion level of small opacities in each area is scored based on the observations of a trained radiologist. The radiologist assesses the profusion of small opacities by comparing with ILO standard radiographs and gives a discrete score from 0/−, 0/0, 0/1, 1/0, 1/1, 1/2, 2/1, 2/2, 2/3, 3/2, 3/3, and 3/+, corresponding to the twelve possible categories. In this paper, the ground truth database is built up according to Chinese diagnosis criteria of pneumoconiosis. Hodous et al.[3] made a comparison of pneumoconiosis interpretation between Chinese and American readers and classifications. Some differences exist between Chinese and American interpretations. For small opacity profusion, the Chinese readers tended to read slightly more diseases than the American readers. However, there was a general agreement between the Chinese and American interpretations, which suit our scheme for use in epidemiologic research.

[1]*From the Department of Biomedical Engineering, School of Life Sciences and Biotechnology, Shanghai Jiao Tong University, NO.800, Dongchuan Road, Shanghai 200240, China.*

[2]*From the Imaging Technologies Lab, GE Global Research, Shanghai 201203, China.*

[3]*From the Shanghai Pulmonary Hospital, Shanghai 200433, China.*

Correspondence to: Jun Zhao, Department of Biomedical Engineering, School of Life Sciences and Biotechnology, Shanghai Jiao Tong University, NO.800, Dongchuan Road, Shanghai 200240, China; tel: +86-13621706907; fax: +86-021-34204872; e-mail: junzhao@sjtu.edu.cn

Copyright © 2010 by Society for Imaging Informatics in Medicine

Online publication 20 February 2010
doi: 10.1007/s10278-010-9276-7

Digital radiograph is a prevalent approach for lung disease diagnosis, especially in mass screening. A CAD scheme is necessary to reduce workload and improve the workflow in mass chest screening. It is also useful in the training program for young radiologists. One promising application is the integration of CAD system with PACS environment.

In the past two decades, many researchers have dedicated in the development of CAD in various medical fields including chest radiography and mammography.[4–7] Three main research projects of CAD scheme have been investigated:[7] quantitative analysis of lesions involved in vascular imaging; detection of lung nodules and interstitial lung diseases in chest radiographs; and detection of clustered microcalcifications in mammograms.

Generally, a complete CAD scheme for chest consists of three main steps: segmentation, feature extraction and feature classification. Segmentation is a necessary prerequisite process for quantitative analysis applications. In this stage, the object in the image should be delineated accurately. Feature extraction is the second process of describing the image feature using geometrical feature, texture feature, intensity, location, and so on. The third process is feature classification, where various classifiers, like k nearest neighbors (k-NNs), random forest or support vector machines (SVM), could be employed to discriminate normal samples from abnormal.

Our methodology is engaged in the development of CAD scheme for pneumoconiosis. The studies of investigating CAD for pneumoconiosis date back to the 1970s,[8–11] with a recent revival of interest in the late 1990s.[12–18] However, in practice, they all face a considerably serious limitation: based on manually selected regions of interest (ROI), the previous CAD schemes were constrained to be applicable to the clinical use. Kondo and Kouda[12] proposed an automatic method for the detection of rounded pneumoconiosis opacities. Nevertheless, their method is on the basis of manual lung field subdivision and then these subdivided regions were quarried for the automatic analysis using neural network.

To overcome this critical difficulty, a fully automatic strategy is presented here. Our work is aimed at finding abnormal signs of a diffuse textural nature encountered in mass chest screening against pneumoconiosis automatically. The main contributions include the subdivision of lung fields into six zones according to the ILO classification guidance, the combination of multi-scale difference filter bank with histogram and co-occurrence matrices to extract discriminatory features from each zone, the utility of support vector machine as region level classifier and the employment of chest level classifier to incorporate six regions' prediction results into the final classification.

The scheme is initialized with the lung field segmentation using the active shape model (ASM), a general image segmentation algorithm. Next, the segmented lung fields are subdivided into smaller regions. Then, a multi-scale difference filter bank is used to enhance subtle abnormalities in pneumoconiosis chest radiograph. On the original and processed lungs, the histogram and the co-occurrence matrices features are calculated from each region. Based on all the features calculated, support vector machine (SVM), a promising classifier, is employed to extract the relevant discriminatory information through a leave-one-out cross-validation method. Finally, a chest level classifier is used to integrate six regions' prediction results to obtain the final classification.

METHODS

Lung Segmentation

Active Shape Model Segmentation

The aim of segmentation is to extract corresponding regions within the lung fields in order to avoid the background noise. Lung field segmentation on posterior-anterior (PA) chest radiographs has attracted a number of researchers. There are rule based schemes,[19,20] Markov random field modeling,[21] ASM method with optimal features[22] and the detection of oriented edges together with active contours models.[23] As the efficiency and accuracy are required in the segmentation process, ASM algorithm is employed to delineate the lung fields in our work.

The ASM algorithm could be mainly divided into two steps, training and segmentation. The main steps of ASM algorithm[24,25] are briefly summarized below. Figure 1 shows the result of a lung field segmentation with ASM algorithm.

(Training/model building)
Begin

Step 1: *Label the lung contour landmarks on the training images.*

Step 2: *Align the manual landmark shapes by scaling, rotating and translating.*
Step 3: *Capture the statistics of a set of aligned shapes.*
Step 4: *Construct the shape model.*

End

(Segmentation/lung contour delineation)
Begin

Step 1: *Initialize the process with the average shape obtained in training phase.*
Step 2: *For each landmark in the initial contour, search the optimal point along the direction perpendicular to the contour.*
Step 3: *Find the best fitted lung field shape and update the parameters and the shape model.*
Step 4: *Repeat Step2 and Step3 until the process converges.*

End

Subdividing the Lung Fields into Six Zones

According to the domain knowledge of pneumoconiosis diagnosis, the texture features are extracted from multiple regions to better capture the features of the diffuse abnormalities. On the basis of the ILO guidance and radiologists' diagnostic measurement, six non-overlapping regions are subdivided in the lung fields. Each region will be analyzed with a separate classifier that is built by the features extracted solely from this region. The goal is to capture the feature of abnormal variations in that particular part; in principle, this will reduce the influence of superimposed anatomical structures, which tend to make the abnormalities difficult to distinguish.

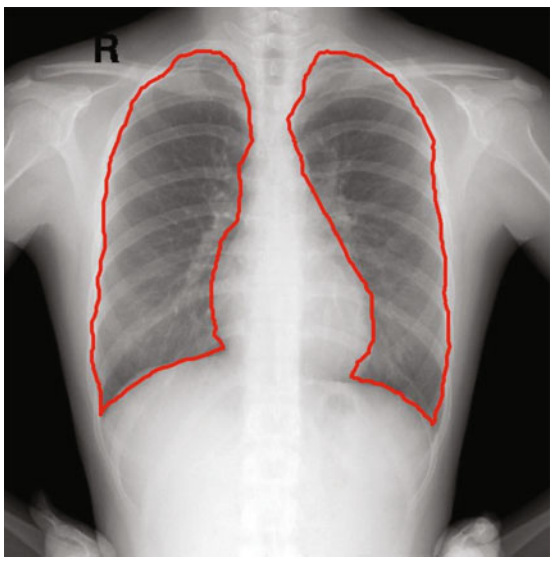

Fig. 1. The results of lung field segmentation using ASM algorithm. The *first row images* are the segmentation results of pneumoconiosis images. The *second row images* are the segmentation results of normal images.

The six subdivision zones are decided as follow. Given a lung field mask image, each lung field is subdivided into three parts—the upper, middle and lower part. The subdivision of the right lung field is done by calculating the apex and diaphragm position, and computing the horizontal lines that divide the field into three parts with equal height. However, as the large variability of the shapes of the left lung leads to difficulty in obtaining diaphragm position, the left lung field subdivision differs from the right lung field. The left lung field is subdivided by calculating the apex and bottom position, and computing the horizontal lines that divide the field into three parts with equal height, see Figure 2, where R_i denotes a non-overlapping region.

Feature Extraction

Feature extraction in a CAD scheme is to find effective measures to represent the abnormalities appeared in medical image. Generally, chest X-ray pneumoconiosis contains diffuse abnormalities. For the detection of diffuse abnormalities, a large number of researches had proved the effectiveness of the textural analysis method.[26–28] The whole feature extraction operation consists of two phases: image filtering and feature calculation. The image filtering is the initial level preprocessing for the enhancement of abnormalities. The feature calculation is the second level stage of extracting abnormalities in the object image. The efficient integration of the initial level image filtering with the second level feature calculation has remained a challenge. Innovative solutions to this problem will advance the field of feature extraction, even the final image classification.

Preprocessing by Difference Filtering

In this paper, we investigate the chest X-ray pneumoconiosis where abnormalities are of various sizes, of various degrees of opacities, rounded or irregular. The image enhancement preprocessing is used to highlight small opacities. To this end,

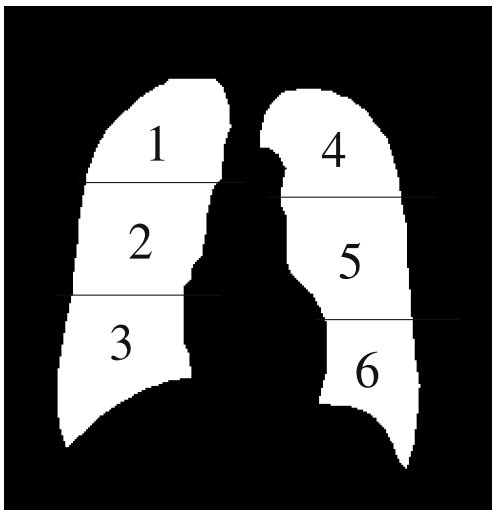

Fig. 2. The lung fields are subdivided into six regions. Respectively, region (1–6) denote upper right, middle right, lower right, upper left, middle left, and lower left of lung field. The shape of the lung fields illustrated here is obtained from ground truth database that is manually specified by an experienced radiologist.

Ginneken[27] used a multi-scale Gaussian filter bank to preprocess image. Similarly, in order to search the significant histogram features and the co-occurrence features, a multi-scale difference filter bank $L_n^{\theta}(x,y,d)$ is developed to improve the image contrast, where n, θ, and d denote the order of difference, the orientation in which the difference is computed and the difference scale, respectively. Let $I(x,y)$ denotes image gray value. The first and second order difference filter bank are defined as

$$L_1^{\theta}(d) = f_x \cdot \cos\theta + f_y \cdot \sin\theta, \tag{1}$$

and

$$L_2^{\theta}(d) = f_{xx} \cdot \cos^2\theta + f_{xy} \cdot \cos\theta \cdot \sin\theta + f_{yy} \sin^2\theta, \tag{2}$$

where f_x and f_y represent the first order differences, f_{xx}, f_{xy} and f_{yy} represent the second order differences:

$$f_x = |I(x+d,y) - I(x-d,y)|, \tag{3}$$

$$f_y = |I(x,y+d) - I(x,y-d)|, \tag{4}$$

$$f_{xx} = |f_x(x+d,y) - f_x(x-d,y)|, \tag{5}$$

$$f_{yy} = |f_y(x,y+d) - f_y(x,y-d)|, \tag{6}$$

$$f_{xy} = |f_x(x,y+d) - f_x(x,y-d)|. \tag{7}$$

Feature Calculation

The histogram and the co-occurrence matrices are useful measures for image feature.[8–11,27–31]

The histogram features[19] (mean, M_1; standard deviation, M_2; skew, M_3; kurtosis, M_4; energy, EG; entropy, EP) are computed as

$$M_1 = \sum i \cdot h(i), \tag{8}$$

$$M_2 = \sqrt{\sum (i - M_1)^2 \cdot h(i)}, \tag{9}$$

$$M_3 = \frac{1}{M_2^3} \sum (i - M_1)^3 \cdot h(i), \tag{10}$$

$$M_4 = \frac{1}{M_2^4} \sum (i - M_1)^4 \cdot h(i) - 3, \tag{11}$$

$$EG = \sum h(i)^2, \tag{12}$$

$$EP = -\sum h(i) \cdot \log h(i), \tag{13}$$

where $h(i)$ denotes the histogram of the intensity i.

The co-occurrence matrices features[19] (energy, EG; entropy, EP; local homogeneity, LH; correlation, CO; and inertia, IN) are computed as

$$EG = \sum\sum H(i_1,i_2)^2, \tag{14}$$

$$EP = -\sum\sum H(i_1,i_2) \log(H(i_1,i_2)), \tag{15}$$

$$LH = \sum\sum \frac{H(i_1,i_2)}{1 + (i_1 - i_2)^2}, \tag{16}$$

$$CO = \sum\sum \frac{(i_1 - \mu_{i_1})(i_2 - \mu_{i_2})H(i_1,i_2)}{\sigma_{i_1}\sigma_{i_2}}, \tag{17}$$

$$IN = \sum\sum (i_1 - i_2)^2 H(i_1,i_2)^2, \tag{18}$$

where $H(i_1,i_2)$ denotes the two-dimensional histogram of the two intensities i_1 and i_2, μ and σ denote the first and second moment around the mean of the joint distribution.

Feature Normalization

Since significant difference exists among features, normalizing the features provides authentic information for the classifier. The normalization method we used here is to compute the maximum and minimum data in the feature vector set and map them onto the same reference gray intervals. The computation can be carried out by

$$X_{\text{normal}} = \frac{2X - (X_{\max} + X_{\min})}{X_{\max} - X_{\min}}, \qquad (19)$$

where X_{normal}, X, $X_{\max}$, and $X_{\min}$ represent the normalized data, the raw data, the maximum data and the minimum data, respectively.

Feature Selection

Feature selection[32] is selecting a subset of relevant features for building robust learning models. Through removing most irrelevant and redundant features from the data, feature selection improves the performance of learning models by enhancing generalization capability, improving model interpretability and saving time cost. Moreover, the feature selection helps people to acquire better understanding about the data sets by telling them which are the key features and how they are related with each other. Many standard data analysis software systems are utilized for feature selection, such as MATLAB, Orange, WEKA (Waikato Environment for Knowledge Analysis[32]), etc.

In this paper, feature selection is accomplished by a ranker used in conjunction with information gain attribute evaluator which is available on WEKA.[32] WEKA is an open source machine learning software developed by the university of Waikato, New Zealand, available on the web at http://www.cs.waikato.ac.nz/ml/weka/. WEKA contains a collection of algorithms for data analysis and predictive modeling, together with graphical user interfaces for easy access to these algorithms. It supports several data mining tasks, such as data preprocessing, classification, regression and feature selection. Feature selection methods in WEKA consist of two steps: an attribute evaluation and a searching method. The attribute evaluation method we chose is the information gain attribute evaluator that evaluates the individual attribute by measuring the feature information gain with respect to the class. We integrate this evaluation method with the ranker searching method, a ranking scheme for individual attribute. It sorts attributes by their individual evaluations.[32]

Classification

In this paper, the classification is made up of two processes: the region level classification and the chest level classification.

Region Level Classification

The technique of classifier design has been evolved rapidly in the past two decades. Arya and Mount[33] developed a fast k-NNs algorithm, wherein a probability measure of a region is computed using weighted voting among the k-NNs. More recently, SVM has been used for various classification tasks in pattern recognition field, which is a promising classification technique developed by Vapnik et al.[34] Based on the statistical learning theory and the structural risk minimization,[35] the SVM obtains the best generalization ability through achieving the optimal balance between the complexity of model and the learning ability. Random forest[36] is another widely used classifier. The classifier utilizes an ensemble of unpruned classification trees.[37–39] Each classification tree is created by using bootstrap samples of the training data and the random feature selection in tree induction. Estimation is made by integrating (majority voting or averaging) the estimations of the ensemble.

Our paper proceeds with the region level classification using the SVM classifier. The key step of the SVM algorithm is to map the input vector onto a higher dimension with a selected kernel and construct an optimal linear separating hyperplane for the separable samples, as is shown in Figure 3.

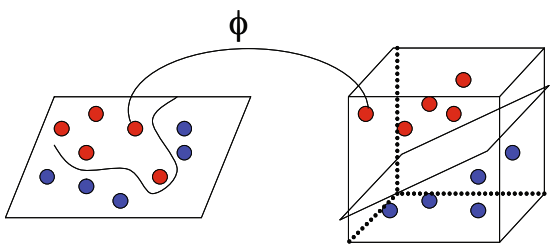

Fig. 3. The input vector is mapped onto a higher dimension with kernel ϕ. The optimal linear separating hyperplane is constructed for the separable samples. The *red circles* denote class A. The *blue circles* denote class B.

Usually, the SVM facilitates high dimensional data to transform nonlinear cases into linear separable cases. However, when it comes to a large number of training data and multi-classification cases, the SVM performs undesirably. Herein, we use the SVM to deal with two categories cases and small training data to make great use of its classification performance. The key steps of the SVM are summarized below.

Step 1: Choose a kernel and a penalty factor C for the nonlinear transformation.
Step 2: Input the feature vector set, and map them onto a higher dimension.
Step 3: Search the optimal linear separating hyperplane in a higher dimension.

The feature extraction yields a feature vector for each subdivided region in each image. These feature vectors are used by the classifier to estimate the abnormalities of each region. For this purpose, a training dataset is constructed for each region. The K folds cross-validation method with the leave-one-out is used to obtain the SVM parameters. For each region, the received operating characteristic (ROC) analysis[40] could be performed. A ROC curve plots the true positive fraction as a function of the false positive fraction. The area under the ROC curve (AUC) value is denoted as A_v. It demonstrates how reliable the classification can be performed for the region under consideration. A value of A_v=1 corresponds to perfect classification, A_v=0.5 represents random guessing. The AUC value is an important parameter for the subsequent chest level classification as the weighting factor.

Chest Level Classification

The prediction result per region may be an essential final result for a CAD scheme. Regions with a high abnormality score are outlined when a radiologist read the image. Furthermore, a chest level classifier is employed to integrate the probabilities of the six regions to obtain the classification result for the whole image. The area under the ROC curve for each region derived in the training phase is used as an indicator of the classification performance of the scheme for each region. In this paper, we incorporate the regional scores into a stand-alone result via two methods: the weighted voting and the weighted multiplying.

The weighted voting is a useful measure widely used. Herein, the weighting factors are calculated for the six regions. Each region votes for the region to be normal or abnormal. The final probability of an image being abnormal based on the six regions is given by

$$\Pr ob = \sum_{i=1}^{N} W_i P_i, \tag{20}$$

where Pr ob denotes the final probability, P_i, which ranges from 0 to 1, denotes the prediction result of the ith region, N is the number of regions, herein N=6, and W_i is the weighting factor for region i determined by the AUC value of each region obtained in the training phase given by

$$W_i = \frac{A_{v_i}}{\sum_{i=1}^{N} A_{v_i}}. \tag{21}$$

Weighted multiplying was proposed by Ginneken,[26] for which the reason is that the image is abnormal if any of the regions is abnormal. The final probability of an image being abnormal based on six regions is given by

$$\Pr ob = 1 - \prod_{i=1}^{N} (1 - f_i P_i), \tag{22}$$

where Pr ob denotes the final probability, P_i, which ranges from 0 to 1, denotes the prediction result of the ith region, N is the number of regions, herein N=6, and f_i is the weighting factor for region i determined by the AUC value of each region obtained in the training phase given by

$$f_i = \mathrm{MAX}\left(\frac{A_{v_i} - T_{A_v}}{1 - T_{A_v}}, 0\right). \tag{23}$$

The threshold T_{A_v} determines the minimum value that A_{v_i} must have in order to be taken into account in the classification of the complete image.

Database of Digital Chest X-rays for Training and Testing

A total of 850 PA chest radiographs were used to evaluate the scheme, which were collected from Shanghai Pulmonary Hospital. The digital chest radiographs were acquired from GE Revolution XQi, with a matrix of 2022*2022 and 14 bit depth

in the DICOM format. The radiographs consist of 600 normal cases and 250 abnormal cases (the age of the patient ranges from 18 to 70 years. The abnormal database includes those with pneumoconiosis ranging from stage 0+ to stage III. The ground truth is the diagnosis conclusion by three qualified radiologists from Shanghai Pulmonary Hospital. 850 images were randomly divided into a training set and a testing set, each consists of 300 normal cases and 125 abnormal cases, about 30% of which were abnormal, as is defined in Table 1. Note that this percentage of abnormal cases is much higher than that is encountered in realistic screening application.

Experiments are conducted with these 850 chest images, and the normal and abnormal categories selected for this study are defined under the supervision of three radiologists as is illustrated in Table 1.

We use the first and second order difference filter bank with a given orientation $\theta \in \{0,45,90,135\}$ and a given scale $d \in \{1,2\}$. This filter is applied to both the histogram feature calculation and the co-occurrence matrices feature calculation. The grey level co-occurrence matrices are constructed for a given direction $\varphi \in \{0,45,90,135\}$ and a given distance $D = 1$. Moreover, we compute the features from the original image without any filtering. Therefore, a feature database is created through extracting 427 features such as the histogram features (mean, standard deviation, skew, kurtosis, energy, entropy), the co-occurrence matrices features (energy, entropy, local homogeneity, correlation and inertia).

In order to evaluate the effectiveness of the feature selection skill, the feature selection approach in WEKA is used, separately for each region. Features ranking top 20 of each region are selected. Our experiments were conducted on two feature vectors: the full feature vector consisting of 427 features and the selected feature vector consisting of 20 selected features for each region.

In this paper, we use SVM (support vector machine) classifier for feature training and classification. Note that the choice of a kernel and a penalty factor C is crucial for the SVM. We choose the radial basis function (RBF) kernel since it has better performance than the linear and polynomial kernel at most situations. Let Γ be the parameter of the RBF kernel. The value of C and Γ is determined for each region in the training phase.

The evaluation specificity consists of accuracy (*Acc*), sensitivity (*Ses*), specificity (*Spe*), and average of sensitivity and specificity (*SSAve*). They are defined as follows:[6]

$$Acc = (TP + TN)/(TP + FN + TN + FP), \quad (24)$$

$$Ses = TP/(TP + FN), \quad (25)$$

$$Spe = TN/(TN + FP), \quad (26)$$

$$SSAve = (Ses + Spe)/2, \quad (27)$$

where *TP*, *FN*, *FP*, and *TN* denote the number of abnormal samples classified as abnormal, the number of abnormal samples classified as normal, the number of normal samples classified as abnormal and the number of normal samples classified as normal, respectively.

TEST RESULTS

Results of Region Level Classification

As the first step, optimal parameters are requisite via testing multiple pairs of penalty factor C and Γ, wherein $\lg C \in \{-11,-9,-7,-5,-3,-1,1,3,5,7,9,11\}$ and $\lg\Gamma \in \{-11,-9,-7,-5,-3,-1,1,3,5,7,9,11\}$; lg denotes logarithm to base 10. 4 folds cross-validation method with leave-one-out is used. The optimal parameters of SVM for each region on two feature vectors are derived according to the maximum value of *SSAve*. As is shown in Table 2, on the basis of the value of *SSAve*, the best training performance of the six regions on the two feature vectors are obtained separately. Mean-

Table 1. The Definition of Opacity Profusion Level for Normal and Abnormal Categories. The Definition of Experimental Database. Training Set Consists of 300 Normal Cases and 125 Abnormal Cases. Testing Set Consists of 300 Normal Cases and 125 Abnormal Cases

Major category	Minor categories	Training set (Case #)	Testing set (Case #)
Normal	0/−, 0/0	300	300
Abnormal	0/1, 1/0, 1/1, 1/2, 2/1, 2/2, 2/3, 3/2, 3/3, 3/+	125	125

Table 2. Classification Results of Training Set on Two Feature Vectors for Each Region Using Leave-One-Out Cross-Validation Method. The Optimal Parameters of SVM are Obtained According to the Maximum Value of *SSAve*. The AUC Value is Obtained Using ROC Analysis

Feature Vector	Region #	lg$C$	lg Γ	SSAve	Acc	Ses	Spe	AUC value
Full	R1	9	−11	0.912	0.908	0.920	0.903	0.998
Selected	R1	11	5	0.878	0.882	0.862	0.894	0.968
Full	R2	7	−9	0.923	0.925	0.920	0.926	0.999
Selected	R2	11	−7	0.879	0.889	0.851	0.907	0.970
Full	R3	9	−9	0.928	0.922	0.944	0.913	0.999
Selected	R3	7	−5	0.884	0.892	0.862	0.905	0.971
Full	R4	7	−9	0.895	0.892	0.904	0.887	0.998
Selected	R4	11	−5	0.873	0.875	0.866	0.879	0.988
Full	R5	11	−11	0.924	0.924	0.921	0.927	0.998
Selected	R5	11	−3	0.889	0.880	0.911	0.867	0.998
Full	R6	9	−11	0.896	0.892	0.905	0.885	0.993
Selected	R6	11	−5	0.878	0.875	0.886	0.870	0.979

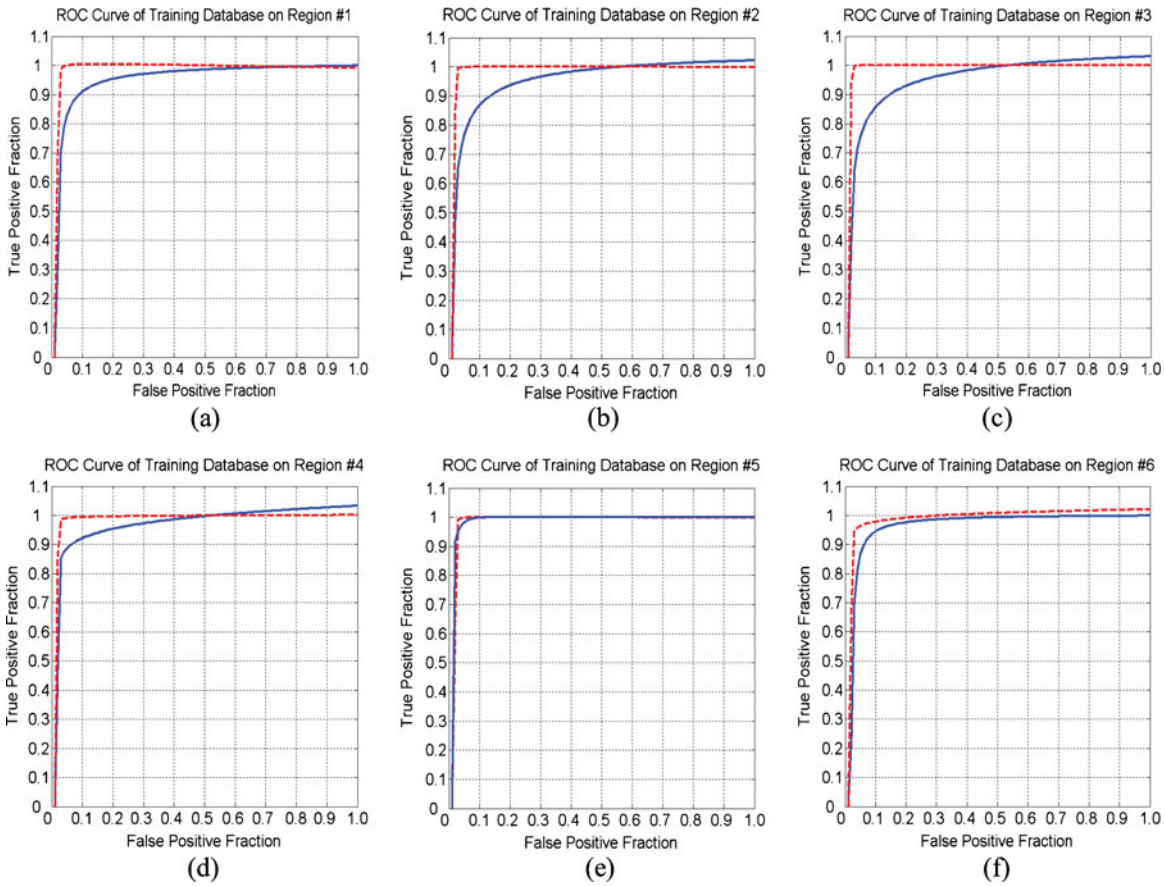

Fig. 4. ROC curves for the training set for the six regions on two feature vectors. The *red dashed line* represents ROC curve on full feature vector. And the *blue solid line* represents ROC curve on selected feature vector. Respectively, the area under the curve on full and selected feature vectors are a 0.998, 0.968 for region #1; b 0.999, 0.970 for region #2; c 0.999, 0.971 for region #3; d 0.998, 0.988 region #4; e 0.997, 0.998 for region #5; f 0.993, 0.979 for region #6.

while, the cross-validation classification results of the training set for each region are given. The sensitivity, one of the most important specificities, of the six regions for the full feature vector are over 90%, which provide essential information for further validation on the testing set. The overall performance of the cross-validation on the training set shows great potential as all the evaluation specificities are over 85%. Then, the ROC curve analysis is used to evaluate the training performance of each region. Figure 4 plots the ROC curves for each region on the two feature vectors. The red dashed line represents ROC curve on the full feature vector. And the blue solid line represents ROC curve on the selected feature vector. Correspondingly, Table 2 lists the AUC values for each region. As is demonstrated in the figure and the table, the AUC values for the six regions on the two feature vectors are over 0.97, in which some values approach one, the perfect result. Therefore, we believe that the training results are reliable and the optimal training models are built up for further testing sets. Next, we use the training parameters and weighting factors derived from the ROC analysis to classify the testing database using the chest level classifier.

Results of Chest Level Classification

The result of chest classification is derived via the integration of the six regions prediction results with the chest level classifier. Here we utilize two different classifiers to validate how the design of chest level classifier affects the final performance. Threshold should be set to decide the final classification of the chest radiographs, normal or abnormal. The threshold we used here is obtained in the training phase by tuning. The optimal thresholds on the two feature vectors from the training set using two different chest level classifiers are derived according to the maximum value of *SSAve*. The tuning results are given out in Table 3. Table 3 illustrates the variability of threshold value for different classifiers. Let T_v denotes the threshold value.

$$\text{Normal cases,} \begin{cases} TN, \Pr ob < T_v \\ FP, \Pr ob \geq T_v \end{cases}, \tag{28}$$

$$\text{Abnormal cases,} \begin{cases} FN, \Pr ob < T_v \\ TP, \Pr ob \geq T_v \end{cases}. \tag{29}$$

Table 4 lists the final chest classification results of the testing database on the two feature vectors using two different chest level classifiers. Correspondently, the ROC curves are plotted in Figure 5 and the AUC values are also listed in Table 4. The results for the testing set are similar with the cross-validation results on the training set. First of all, the sensitivity on the two feature vectors using two different chest level classifiers is higher than 90%. For the AUC value, the performance is over 0.948, still very good. For the full feature vector, the accuracy is around 89.2%, for the selected feature vector, the accuracy is around 87.8%.

DISCUSSION

In this section, we discuss whether some components of the system affect the final classification performance.

Lung field segmentation is a crucially important stage before the calculation of the histogram and co-occurrence matrices features, because it removes the effect of body and other noise background and thus makes it possible to distinguish the normal lung texture from abnormalities.

Another important process is the multi-scale difference filtering, with which discriminatory features could be extracted; these features appear to be related closely to small opacities and profusions in pneumoconiosis chest.

In addition, feature selection is helpful to save time cost in our scheme. With a total of 427 texture

Table 3. Tuning Results of Threshold on Two Feature Vectors from Training Set Using Two Different Chest Level Classifiers. The Optimal Threshold is Derived According to the Maximum Value of *SSAve*

Feature Vector	Chest Level Classifier	T_v	SSAve	Acc	Ses	Spe
Full	Voting	0.26	0.995	0.993	1.000	0.990
Full	Multiplying	0.89	0.995	0.993	1.000	0.990
Selected	Voting	0.24	0.953	0.936	0.992	0.913
Selected	Multiplying	0.82	0.955	0.939	0.992	0.917

Table 4. The Final Chest Classification Results of Testing Database on Two Feature Vectors Using Two Different Chest Level Classifiers. The AUC Value is Obtained Using ROC Analysis

Feature Vector	Chest Level Classifier	Acc	Ses	Spe	A_v
Full	Voting	0.889	0.920	0.877	0.978
Full	Multiplying	0.892	0.920	0.880	0.967
Selected	Voting	0.878	0.912	0.863	0.958
Selected	Multiplying	0.877	0.904	0.867	0.948

features calculated, we apply the ranker selection method to extract the ranking top 20 features for each region. From Figure 5 and Table 4, the full feature vector yields slightly better performance over the selected feature vector. Nevertheless, the utility of feature selection do save time cost greatly. For the full feature vector, to process a chest image, it needs averaged 261.78 s in total on a DELL OPTIPLEX 760 PC. While for the selected feature vector, it only needs averaged 45.66 s. This is meaningful for clinical application in mass screening. It is also interesting to note that approximately 90% of the selected features are co-occurrence matrices features. Obviously, the co-occurrence matrices features are more discriminative over histogram features for pneumoconiosis abnormalities detection. In addition, among the selected co-occurrence matrices features, over 95% of them are energy, entropy and correlation

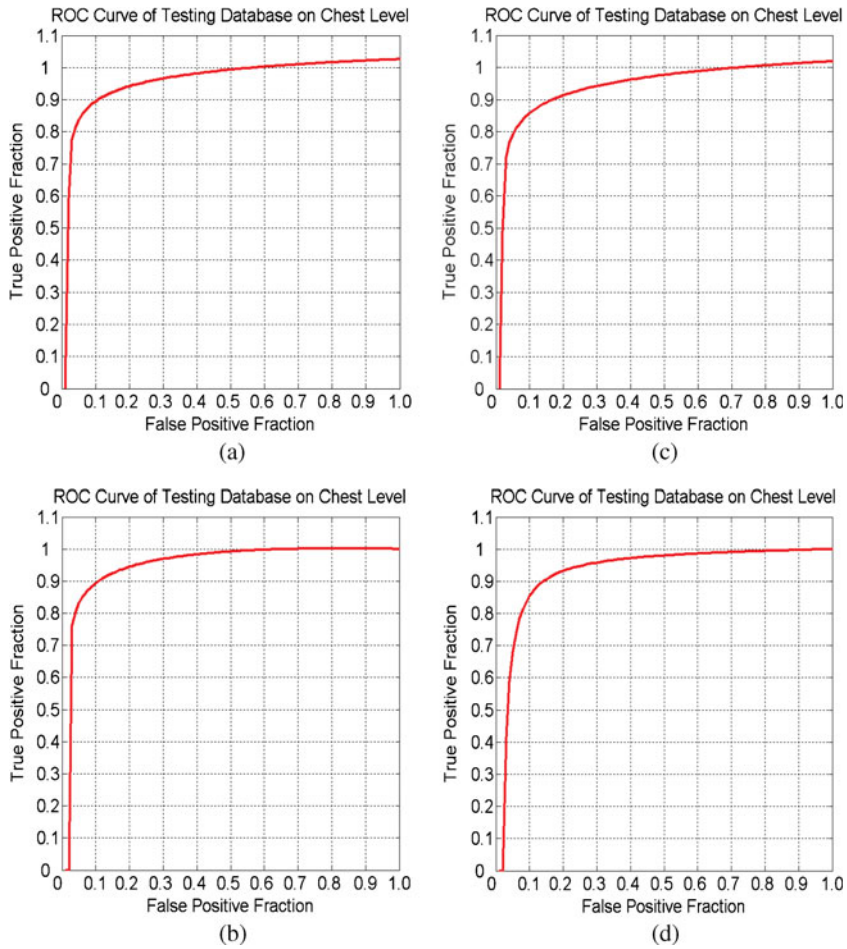

Fig. 5. ROC curves for the testing database on chest image level. a Full feature vector using weighted voting classifier, the area under the curve is 0.978. b Full feature vector using weighted multiplying classifier, the area under the curve is 0.967. c Selected feature vector using weighted voting classifier, the area under the curve is 0.958. d Selected feature vector using weighted multiplying classifier, the area under the curve is 0.948.

features. This means these three features contribute the most to the final classification performance.

For the component of chest level classifier, we compared two different classifiers. Table 4 indicates that both classifiers almost resulted in identical results. Also, from Figure 5 and Table 4, the ROC curve derived from two classifiers nearly overlaps. And the AUC value of the voting classifier only has 1% better performance than the multiplying classifier. Therefore, we can say that the choice of chest level classifier does not improve the final classification result significantly although it may well be that another chest level classifier would yield slightly better results.

Our results are comparable to, even better than, those reported by Soliz[16] and co-workers. Our scheme yields the 89.2% accuracy on the full feature vector. Even on the selected feature vector, the 87.8% accuracy is obtained, while Soliz et al. reported the 86.8% accuracy. They used a combined scheme for the detection of pneumoconiosis disease. First of all, the ROIs they defined were manually labeled by the pulmonologist with ground truth. Their method is not suitable to the mass chest screening in clinic use. Secondly, the selected ROIs were characterized using the method based on sum and difference histogram, and co-occurrence matrices statistics parameters. With such a small set of features, they are not able to find full information of features representing the abnormalities. Thirdly, the laterally primed adaptive resonance theory neural network was implemented by Soliz et al. as a means for clustering and classifying complex ROI feature vectors. The SVM classifier used in our work obtains the best generalization ability through achieving the optimal balance between the complexity of model and the learning ability that might outperform the neural classifier used by Soliz et al. Fourthly, the database used for their study consisted of 212 radiographs, while our database consisted of 850 radiographs with wider age range.

CONCLUSION

We have presented an automatic scheme for the pneumoconiosis abnormality detection in digital chest radiographs. The whole scheme begins with the lung field segmentation, followed by the feature extraction based on the multi-scale difference filtering histogram and the co-occurrence matrices to represent opacities in chest X-ray. Finally, the SVM classifier is employed to characterize the feature vector and the chest level classifier is used to classify the whole chest image.

Experiments were performed on both the training set and the testing set to classify the pneumoconiosis cases from the normal cases. In the training phase, training models and weighting factors for each region are obtained. We evaluate our scheme on the testing set with the full feature vector or the selected feature vector by using two different chest level classifiers. The evaluation result on the full feature vector is the sensitivity 92.0%, the specificity 87.7%, the accuracy 88.9% and the AUC value is 0.978, where the threshold value is 0.26. The evaluation result on the selected feature vector is the sensitivity 91.2%, the specificity 86.3%, the accuracy 87.8% and the AUC value is 0.958, where the threshold value is 0.24, as is demonstrated in Figure 5 and Table 4. Compared with previous researches,[7–17] our fully automated scheme has higher classification performance and more convenient interaction, which is helpful for mass screening in clinic.

According to the classification performance evaluated in the testing set, we believe that this scheme may be helpful to radiologists who are using CAD output for mass chest screening, diagnostic image interpretation and differentiating the pneumoconiosis cases from the normal cases.

ACKNOWLEDGEMENTS

This work is supported in part by the Science and Technology Commission of Shanghai Municipality, China (Grant No.074107022), the National High Technology Research and Development Program of China (863 Program) (2007AA02Z452), National Natural Science Foundation of China (30570511 and 30770589) and National Basic Research Program of China (2010CB834300).

REFERENCES

1. International Labor Organization (ILO): Guidelines for the use of the ILO international classification of radiographs of pneumoconioses. Occupational Safety and Health Series, No. 22 (Rev.). International Labor Office, Geneva Switzerland, 1980

2. Hering KG, Jacobsen M, Bosch-Galetke E: Further development of the International Pneumoconiosis Classification—from ILO 1980 to ILO 2000 and to ILO 2000/German Federal Republic version. Pneumologie 57(10):576–584, 2003

3. Hodous TK, Chen RA, Kinsley KB: A comparison of pneumoconiosis interpretation between Chinese and American reader and classifications. J Tongji Med Univ 11(4):225–229, 1991

4. Ginneken BV, Romeny BM, Viergever MA: Computer-aided diagnosis in chest radiography: a survey. IEEE Trans Med Imag 20(12):1228–1241, 2001

5. Elter M, Horsch A: CADx of mammographic masses and lustered microcalcifications: a review. Med Phys 36(6):2052–2068, 2009

6. Giger ML, Chan HP, Boone J: Anniversary paper: history and status of CAD and quantitative image analysis: the role of medical physics and AAPM. Med Phys 35(12):5799–5820, 2008

7. Doi K: Computer-aided diagnosis in medical imaging: historical review, current status and future potential. Comput Med Imaging Graph 31:198–211, 2007

8. Kruger RP, Thompson WB, Turner AF: Computer diagnosis of pneumoconiosis. IEEE Trans Syst Man Cybern 4(1):40–49, 1974

9. Ledley RS, Huang HK, Rotolo LS: A texture analysis method in classification of coal workers' pneumoconiosis. Comput Biol Med 5(1–2):53–67, 1974

10. Savol AM, Li CC, Hoy RJ: Computer-aided recognition of small rounded pneumoconiosis opacities in chest X-rays. IEEE Trans Pattern Anal Mach Intell 2(5):479–482, 1980

11. Hall EL, Crawford WO, Roberts FE: Computer classification of pneumoconiosis from radiographs of coal workers. IEEE Trans Biomed Eng 22(6):518–527, 1975

12. Kobatake H, Ohishi K, Miyamichi J: Automatic diagnosis of pneumoconiosis by texture analysis of chest X-ray images. IEEE ICASSP 12:610–613, 1987

13. Ugurlu Y, Ohkura K, Obi T: Detection of increasing profusion of opacities from a sequence of personal chest radiographs. IEEE Int Conf Image Proc 3:402–406, 1999

14. Huang Z, Yu D, Zhao J: Application of neural networks with linear and nonlinear weights in occupational disease incidence forecast. IEEE Asia-Pacific Conference on Circuits and Systems 383–386, 2000

15. Kondo H, Kouda T: Detection of pneumoconiosis rounded opacities using neural network. Joint 9th IFSA World Congress and 20th NAFIPS International Conference 3: 1581–1585, 2001

16. Kondo H, Kouda T: Computer-aided diagnosis for pneumoconiosis using neural network. IEEE Symposium on Computer-Based Systems 467–472, 2001

17. Soliz P, Pattichis MS, Ramachandran J. Computer-assisted diagnosis of chest radiographs for pneumoconioses. Proceedings of SPIE 667–675, 2001

18. Pattichis MS, Pattichis CS, Christodoulou CI: A screening system for the assessment of opacity profusion in chest radiographs of miners with pneumoconiosis. Fifth IEEE Southwest Symposium on Image Analysis and Interpretation 130–133, 2002

19. Ginneken BV: Computer-aided diagnosis in chest radiography. Med Phys 28(6):1144–1150, 2001

20. Li L: Improved method for automatic identification of lung regions on chest radiographs. Acad Radiol 8(7):629–638, 2001

21. Vittitoe NF, Vargas-Voracek R, Floyd CE: Identification of lung regions in chest radiographs using Markov random field modeling. Med Phys 25(6):976–985, 1998

22. Ginneken BV, Frangi AF, Staal JJ: Active shape model segmentation with optimal features. IEEE Trans Med Imag 21(8):924–933, 2002

23. Iglesias I, Souto M, Alegria AM: Lung segmentation on postero-anterior digital chest radiographs using active contours. Lect Notes Comput Sci 3138:538–546, 2004

24. Cootes TF, Taylor CJ, Cooper DH: Active shape models-their training and application. Comput Vis Image Underst 61(1):38–59, 1995

25. Cootes TF, Taylor CJ: Statistical models of appearance for computer vision. Wolfson Image Analysis Unit. University of Manchester, Manchester, UK, 1999. Tech. Rep

26. Katsuragawa S, Doi K: Computer-aided diagnosis in chest radiography. Comput Med Imaging Graph 31:212–222, 2007

27. Ginneken BV, Katsuragawa S, Romeny BM: Automatic detection of abnormalities in chest radiographs using local texture analysis. IEEE Trans Med Imag 21(2):139–148, 2002

28. Sutton RN, Hall EL: Texture measures for automatic classification of pulmonary disease. IEEE Trans Comput 21:667–676, 1972

29. Jagoe JR, Paton KA: Reading chest radiographs for pneumoconiosis by computer. Br J Ind Med 32:367–372, 1975

30. Tully RJ, Connors RW, Harlow CA: Toward computer analysis of pulmonary infiltration. Investigat Radiol 13:198–305, 1978

31. Jagoe JR: Gradient pattern coding-an application to the measurement of pneumoconiosis in chest X-rays. Comput Biomed Res 12:1–15, 1979

32. Witten IH: Data mining: practical machine learning tools and techniques. Morgan Kaufmann, San Francisco, 2005, pp 369–424

33. Arya S, Mount DM: Approximate nearest neighbor queries in fixed dimensions. Proc. 4th ACM-ALAM Symp. Discrete Algorithms 271–280, 1993

34. Vapnik VN: The nature of statistical learning theory. Springer, New York, 199, pp 4–80

35. Burges JC: A tutorial on support vector machines for pattern recognition. Data Min Knowl Disc 2:121–167, 1998

36. Breiman L: Random forests. Mach Learn 45:5–32, 2001

37. Breiman L, Friedman J, Olshen R: Classification and regression trees. Chapman & Hall, New York, 1984

38. Ripley BD: Pattern recognition and neural networks. Cambridge University Press, Cambridge, 1996

39. Hastie T, Tibshirani R, Friedman J: The elements of statistical learning. Springer, New York, 2001

40. Park SH, Goo JM, Jo CH: Receiver operating characteristic (ROC) curve: practical review for radiologitst. Korean J Radiol 5(1):11–18, 2004