Rib Suppression for Enhancing Frontal Chest Radiographs Using Independent Component Analysis

Bilal Ahmed[1], Tahir Rasheed[1], Mohammed A.U. Khan[2], Seong Jin Cho[1], Sungyoung Lee[1], and Tae-Seong Kim[3]*

[1] Department of Computer Engineering, Kyung Hee University, South Korea
{bilal, tahir, sjcho, sylee}@oslab.khu.ac.kr
[2] Comsats Institute of Technology, Abbotabad, Pakistan
mohammad_a_khan@yahoo.com
[3] Department of Biomedical Engineering, Kyung Hee University, South Korea
tskim@khu.ac.kr

Abstract. Chest radiographs play an important role in the diagnosis of lung cancer. Detection of pulmonary nodules in chest radiographs forms the basis of early detection. Due to its sparse bone structure and overlapping of the nodule with ribs and clavicles the nodule is hard to detect in conventional chest radiographs. We present a technique based on Independent Component Analysis (ICA) for the suppression of posterior ribs and clavicles which will enhance the visibility of the nodule and aid the radiologist in the diagnosis process.

1 Introduction

Chest X-rays play an important role in the diagnosis of lung cancer. According to the latest statistics provided by the American Cancer Society, lung cancer is estimated to produce 174,470 new cases in 2006, accounting for about 12 percent of all cancer diagnoses. Lung cancer is the most common cancer related death in both men and women. An estimated 162,460 deaths, accounting for about 29 percent of all cancer related deaths, are expected to occur in 2006. Early detection of lung cancer is the most promising strategy to enhance a patients' chances of survival. Early detection can be achieved in a population screening: the most common screenings for lung cancer make use of Chest Radiography, or low radiation dose Computer Tomography (CT) scans. Despite the development of advanced radiological exams such as CT the conventional Chest X-ray remains the most common tool for the diagnosis of lung cancer. The main reason behind this being the fact that CT and helical CT exams expose the patient to
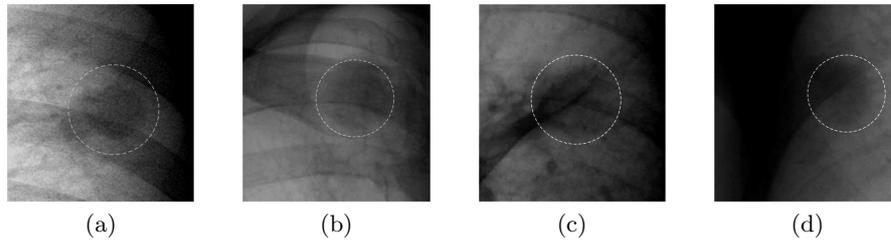
Fig. 1: Effect of location on nodule detection: (a) Nodule with overlapping posterior and anterior ribs. (b) Overlapping clavicle. (c) Overlapping rib and blood vessels. (d) Nodule lying in the hilum area of the radiograph. (Contrast enhanced images)

a higher dose of radiation, estimated to be about 100 times higher than that for a conventional chest X-ray.

Due to the heavy use of conventional chest x-rays for early detection, there is a need for enhancing its diagnostic value. In a recent study carried out to assess various reasons for litigation against physicians, it was revealed that failure to diagnose lung cancer accounted for 80 percent of the cases [13]. For these reasons, there has been a particular interest for the development of computer aided diagnostic (CAD) systems that can serve as a follow-up reader, paying attention to the suspicious regions in the radiograph that then have to be examined by a radiologist. Earlier work on CAD systems for automated nodule detection in chest radiographs was reported in [7]. The process for nodule detection employed multiple gray-level threshholding of the difference image (which corresponds to the subtraction of a nodule-enhanced image and a nodule-suppressed image) and then classification. The system resulted in a large number of false positives which were eliminated by adaptive rule-based tests and an artificial neural network (ANN). Deus Technologies received FDA pre-market approval for its RapidScreen CAD system in July 2001. Its intended use is to identify and mark regions of interest on digital or digitized frontal chest radiographs. The authors in [8] evaluated the usefulness of CAD that incorporated temporal subtraction for the detection of solitary pulmonary nodules on chest radiographs. The authors in [9] concluded that the accuracy of radiologists in the detection of some extremely subtle solitary pulmonary nodules can be improved significantly when the sensitivity of a computer-aided diagnosis scheme can be made extremely high. The authors noted however, that all of the six radiologists failed to identify some nodules (about 10 percent), even with the correct computer output.

The analysis of existing CAD systems revealed that the detectability of the nodule is largely dependent on its location in the chest radiograph. The problematic areas for nodule detection can be categorized as follows: 1) a nodule overlaps completely with an anterior rib and partly with the posterior rib, as shown in Figure 1. The ribs need to be suppressed for identifying the true shape of the nodule; 2) a nodule hidden in the hilum area (the area surrounding the

lung cavity which has high illumination) is hard to detect due to poor contrast as shown in Figure 1; 3) nodule overlaps with the clavicles as shown in Figure 1. Therefore, the suppression of ribs and clavicles in chest radiographs would be potentially useful for improving the detection accuracy of a given CAD system. An attempt was made in [2] where the rib cage was suppressed by means of Massive Training Artificial Neural Network (MTANN). A drawback of the system is the need of a dual-energy bone image in which the ribs are separated from the soft-tissue at the initial training phase. The technique was found to be sensitive to the noise levels, due to the subtraction process.

From a pattern recognition point of view, the rib-cage can be considered as a separate structure that needs to suppressed independently from the rest of the radiograph. We formulate the problem of chest radiograph enhancement considering rib-cage as one class and rest of the image as another separate class. In order to discriminate between the rib-cage class and the other we need data-dependant basis functions that naturally align themselves with the rib structure and remain independent from the other class.

Independent Component Analysis (ICA) has been used quite recently for enhancing EEG signals [10]. It has been shown that ICA can remove undesired artifacts from the EEG signal without causing any substantial damage to the original source signal. The main motivation behind such applications of ICA is the reasonable assumption that noise unrelated to the source is independent from the event-related signals, and hence can be separated. For image data ICA basis work as localized edge filters [5]. In a chest radiograph the ribcage constitutes a sparse bony structure and thus contributing significantly to the overall edge count. ICA can be used to find basis functions which represent this structure and its appropriate characteristics. Our work includes the creation of ICA basis from an enhanced chest radiograph, clustering these basis and the reconstruction of the chest radiograph with the help of non-edge basis. These non-edge basis would correspond to the non-rib component and the exclusion of edge basis would suppress the bone-structure and other undesired edges in the original image.

2   Image Enhancement

Chest radiographs contain non-uniform illumination due to the acquisition apparatus and their complex structure. Non-uniform illumination can distort edges of ribs and suppress the detail of the texture. Due to this non-uniform illumination pattern the nodule might get partially obscured and loose it basic characteristics. The ribs and the blood vessels can be viewed as potential edges, and in order to clearly demarcate them it is desirable that they have high detail in the image. Removing the non-uniform illumination pattern and making the image homogeneous would enhance the texture detail and the rib edges.

Conventional methods of homogenizing image such as homomorphic filtering assume an illumination and reflectance model which considers the illumination as a multiplicative component. We have adopted the technique of [1] for normalizing
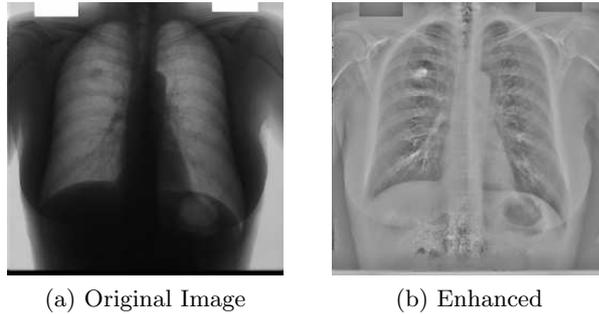
(a) Original Image   (b) Enhanced

Fig. 2: Enhancing the original chest radiograph

the image's intensities locally. The technique for normalization achieves results which enhance the overall image contrast and strengthen the edges.

$$I_N = \frac{I - I_{LP}}{\sqrt{(I^2)_{LP} - (I_{LP})^2}} \tag{1}$$

where 'LP' simply denotes Gaussian blurring of appropriate scale. This local normalization can be considered as making the image zero mean and unit variance. Figure 2 shows the result of this local normalization on a conventional chest radiograph.

3   Independent Component Analysis

ICA performs a blind source separation (BSS), assuming linear mixing of the sources. ICA generally uses techniques involving higher-order statistics. Several different implementations of ICA can be found in the literature [4]. We will not discuss those implementations here and restrict ourselves to the topic. Let us denote the time varying observed signal by

$$\mathbf{x} = (x_1 x_2 .... x_n)^T$$

and the source signal consisting of independent components by

$$\mathbf{s} = (s_1 s_2 .... s_m)^T$$

The linear ICA assumes that the signal x is a linear mixture of the independent components,

$$\mathbf{x} = \mathbf{A} \cdot \mathbf{s} \tag{2}$$

where the matrix A of size $m \times n$ represents linear memory less mixing channels. It is often assumed that $n = m$ (complete ICA) for simplicity, which can be relaxed without any loss of generality. A common preprocessing step is to zero the mean of the data and then apply a linear "whitening" transform so that the
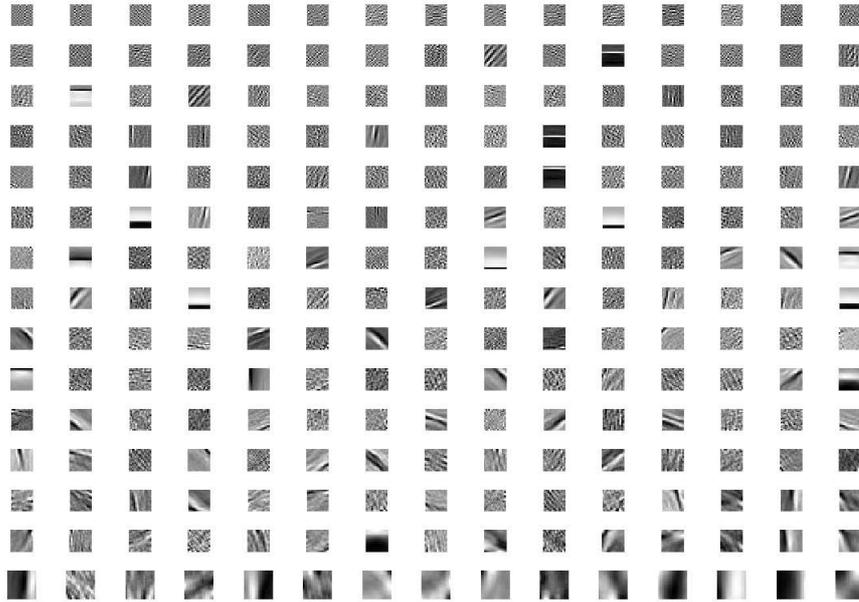
Fig. 3: ICA basis for a chest radiograph

data has unit variance and is uncorrelated. The algorithms must find a separating or de-mixing matrix such that

$$\mathbf{s} = \mathbf{W} \cdot \mathbf{x} \tag{3}$$

where ($\mathbf{W}$) is the de-mixing matrix. ICA has some limitations which can be summarized as: 1) neither energies nor signs of the ICs can be calculated; 2) there is no ordering of the found ICs. Higher-order statistics have been adopted to estimate independent sources in ICA, like kurtosis (the fourth order cumulant) and the negentropy.

3.1 Application to Chest Radiographs

Chest radiographs can be viewed as a mixture of two components the bone-structure and the soft-tissue. From an image processing point of view the first component is mostly composed of edges (due to the sparse bony structure). Suppressing these edges, without affecting the other information in the radiograph would enhance the remaining information pertaining to the second component.

ICA features for natural images have been described in literature as being Gabor filters [3][4]. Application of ICA on medical images such as MRI scans of the skull showed that ICA produced basis which could be more efficiently described as step functions, mainly due to the presence of a large number of flat

intensity values with rapid transitions [3]. In our case the complete set of basis found for a single nodule-containing radiograph are shown in Figure 3. As can be seen the basis are a combination of Gabor filters and step functions(depicting the rapid transition of intensity values). These basis can be best described as being edge filters [5]. Analysis of these basis reveal the fact that there are inherently three major classes in the chest radiograph: namely, rib-cage and blood vessel edges, noise, and the background (containing the hilum, organ shadows and soft-tissue).

Image reconstruction using only a subset of these basis would result in an image in which the all the information related to the left-out basis would be suppressed. For this specific application we would like to leave out the basis that contain the information pertaining mainly to the edges. Simple unsupervised clustering methods such as the k-means clustering algorithm can be employed for the classification of these basis into their respective classes.A distance or similarity metric is needed for describing each class. This distance metric can be obtained from simple kurtosis values of the obtained basis. Image reconstruction from the selected basis would involve making the other basis vectors zero. Let C_1 and C_2 be two classes obtained and the corresponding basis matrix is $\mathbf{A}_1$ and $\mathbf{A}_2$ having the same dimensions as that of the original basis matrix $\mathbf{A}$, but containing zero vectors in place of non-class basis. The image reconstruction for n classes can be given as,

$$\mathbf{x}_i = \mathbf{A}_i \cdot \mathbf{s} \qquad i = 1, 2....n \tag{4}$$

thus obtaining a set of n images corresponding to the n classes.

4   Experiment and Results

The chest radiographs were taken from the JSRT database [11]. The images are digitized to 12 bits posterior-anterior chest radiographs, scanned at a resolution of 2048 x 2048 pixels; the size of one pixel is 0.175 x 0.175 mm ^{2}. The database contains 93 normal cases and 154 cases of proven lung nodule. Diameters and the positions of the nodules are provided along with the images. The nodule diameters range from 5-60 mm, and are located throughout the lung field, and their intensities vary from nearly invisible to very bright. The nodules present in the database are representatives of the problems we delineated in the introduction.

4.1   Preprocessing

The images were first enhanced according to the method discussed in Section 2. The images were down-sampled from 2048 x 2048 to 256 x 256 pixels. Overlapping 15 x 15 blocks of the image were taken as the input to the ICA algorithm.

4.2   Results

The data was initially made zero mean and whitened. The FastICA [4] algorithm was used with the tanh(hyperbolic tangent) non-linearity relating to the Super-Gaussian source distributions. Note that we did not apply PCA reduction at any

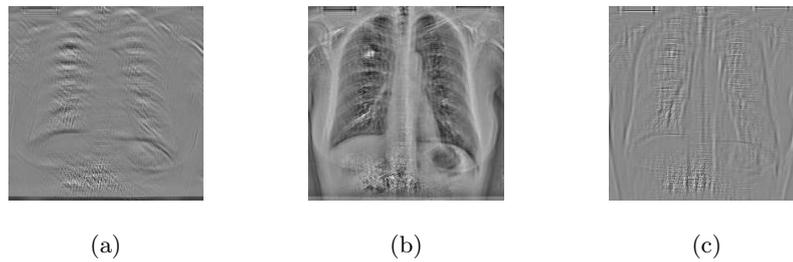(a)   (b)   (c)

Fig. 4: Three classes generated from the clustering of ICA basis.(a) and (b) depict the images corresponding to the bone and noise components whereas (c) depicts the image corresponding to the non-edge basis.

stage to preserve the texture of the overall chest radiograph. Next we show the ICA basis functions obtained and discuss their relevance to our chest radiograph analysis.

Figure 3 shows the resulting basis vectors– (i.e., columns of the A matrix). FastICA was used for learning the structure inherently present in the candidate image. The structure contained in the blocks can be best analyzed by inspecting the linear basis extracted through the algorithm. These basis are used to reconstruct the original image along with the help of the independent components. As can be seen, the resulting basis vectors are localized, oriented and have multiple scales. Second, most of the ICA basis vectors include a small global step-like grayscale change. However, majority of the basis vectors can be conveniently associated with the directional feature associated closely with the ribcage. We do find some small high frequency edges in the basis that seem to be linked with small tissue like structure in the chest radiograph.

The basis so obtained were then clustered according to the k-means algorithm with the clustering feature being the kurtosis of the basis images. The results

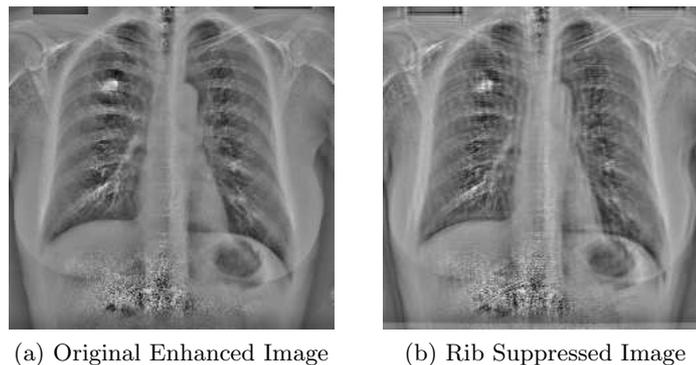(a) Original Enhanced Image   (b) Rib Suppressed Image

Fig. 5: Effect of rib-suppression, the ribs are suppressed but the nodule remains intact.

of the clustering can be seen from Figure 4. The final rib-suppressed image is shown in Figure 5 along with the original enhanced image initially given to ICA. As can be seen that other structures as well as the ribs have been suppressed, with the original nodule left intact in the due course.

The effect of the suppression from a pattern recognition point of view can be characterized as increasing the Gaussianity of the nodule. Pulmonary chest
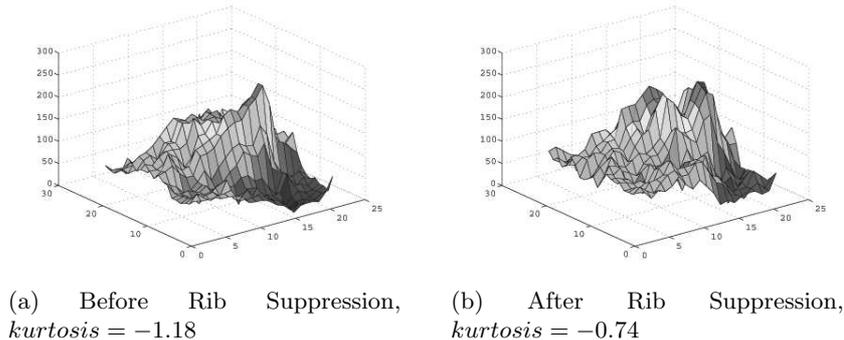

(a) Before Rib Suppression, $kurtosis = -1.18$

(b) After Rib Suppression, $kurtosis = -0.74$

Fig. 6: Gaussianity of the nodule before and after rib-suppression.*(Excess Kurtosis is shown(kurtosis-3))*

nodules have been characterized as Gaussian shapes in [1]. The overlapping of the ribs and other bony structure tends to distort this Gaussianity as can be seen from Figure 6. The removal of these overlapping structures increases the Gaussianity of the underlying nodule. Simple blob detection techniques [6] can then be employed for the detection of these nodules with more accuracy [1]. An illustration of the enhanced Gaussianity is given in Figure 6.

5   Conclusion and Future Work

We have demonstrated that the suppression of ribs and clavicles can be efficiently achieved through the use of Independent Component Analysis. The removal of these artifacts result in the enhancement of the nodule. This increase in the overall Gaussianity of the nodule would enable a standard multi-scale blob-detector[6] to detect a nodule with more accuracy [1].

In our future work, we would like to apply constrained-ICA [12] for the suppression of ribs and removing the blood vessels which hinder the efficient detection of nodules. We would also like to assess the performance of a standard classifier for obtaining the number of false positives produced as a result of applying ICA for radiograph enhancement.

References

1. Schilham, Arnold M.R. and Bram Van Ginneken and Loog, Marco: A computer-aided diagnosis system for detection of lung nodules in chest radiographs with an evaluation on a public database. Medical Image Analysis, Vol. 10. Elsevier (2006) 247–258
2. Suzuki, Kenji and Abe, Hiroyuki and MacMahon, Heber and Doi, Kunio: Image-Processing Technique for Suppressing Ribs in Chest Radiographs by Means of Massive Training Artificial Neural Network (MTANN). IEEE transactions on Medical Imaging, Vol. 25(4). IEEE (2006)
3. Inki, Mika: ICA features of image data in one, two and three dimensions. Proceedings of Independent Component Analysis (ICA)-2003. IEEE (2003) 861–866
4. Hyvarinen, A. and Karhunen, J. and Oja, E.: Independent Component Analysis. Wiley Interscience (2001)
5. Bell, Anthony J. and Sejnowski, Terrence J.: The 'Independent Components' of natural images are edge filters. Vision Research (1997) 3327-3338
6. Lindeberg, Tony: Feature Detection with automatic scale selection. International Journal of Computer Vision, Vol. 30 (1998) 79–116
7. XU, X.W. and Doi, K. and Kobayashi, T. and MacMahon, H. and Giger M.L.: Development of an improved CAD scheme for automated detection of lung nodules in digital chest images. Med Phys., Vol. 24 (1997) 1395–1403
8. Johkoh, T. and Kozuka, T. and Tomiyama, N. and Hamada, S. and Honda, O. and Mihara, N. and Koyama, M.: Temporal subtraction for detection of solitary pulmonary nodules on chest radiographs: evaluation of a commercially available computer-aided diagnosis system. Radiology, Vol. 223 (2002) 806–811
9. Shiraishi, J. and Abe, H. and Engelmann, R. and Doi, K.: Effect of high sensitivity in a computerized scheme for detecting extremely subtle solitary pulmonary nodules in chest radiographs: observer performance study. Acad. Radiology, Vol. 10. (2003) 1302–1311
10. Lange, D.H. and Pratt, H. and Inbar, G.F.: Modeling and Estimation of single evoked brain potential components. IEEE transactions on BioMedical Engineering, Vol. 44. IEEE (1997)
11. Shiraishi, J. and Katsuragawa, S. and Ikezoe, J. and Matsumoto, T. and Kobayashi T. and Komatsu, K. and Matsui, M. and Fujita, H. and Kodera, Y. and Doi, K.: Development of a digital image database for chest radiographs with and without a lung nodule: receiver operating characteristic analysis of radiologists' detection of pulmonary nodules. American Journal of Roentgenology, Vol. 174 (2000)
12. Lu, Wei and Rajapakse, Jagath C.: Approach and Applications of Constrained ICA. IEEE transactions on Neural Networks, Vol. 16(1). IEEE (2005)
13. McLean, Thomas R.: Why Do Physicians Who Treat Lung Cancer Get Sued? Chest, Vol. 126. American College of Chest Physicians (2004)