Implementation and evaluation of a bony structure suppression software tool for chest X-ray imaging

Toinette-Lee Dixon
DXNTOI001

Submitted to the University of Cape Town
in partial fulfilment of the requirements for the degree of
Master of Science in Medicine, Biomedical Engineering.

Supervisor: Associate Professor Tania Douglas
Faculty of Health Science, University of Cape Town

Acknowledgements

I would like to acknowledge the following for their contribution:

- Assoc. Prof. Tania Douglas for her help and input
- Stef Steiner for his support and advice
- NRF/THRIP and Lodox Systems for funding
- Dr Bram van Ginneken, Laurens Hogeweg, Pragnya Maduskar and DIAG for their guidance
- Friends and family

Declaration

I, Toinette-Lee Dixon, hereby declare that the work on which this dissertation is based is my original work (except where acknowledgements indicate otherwise) and that neither the whole work nor any part of it has been, is being, or is to be submitted for another degree in this or any other university.

I empower the university to reproduce for the purpose of research either the whole or any portion of the contents in any manner whatsoever.

.....

Abstract

This project proposed to implement a bony structure suppression tool and analyse its effects on a texture-based classification algorithm in order to assist in the analysis of chest X-ray images.

The diagnosis of pulmonary tuberculosis (TB) often includes the evaluation of chest X-ray images, and the reliability of image interpretation depends upon the experience of the radiologist. Computer-aided diagnosis (CAD) may be used to increase the accuracy of diagnosis. Overlapping structures in chest X-ray images hinder the ability of lung texture analysis for CAD to detect abnormalities. This dissertation examines whether the performance of texture-based CAD tools may be improved by the suppression of bony structures, particularly of the ribs, in the chest region. The study was motivated by the difficulty of diagnosing childhood pulmonary TB, but rib suppression was performed and evaluated on both paediatric and adult images.

Rib suppression was divided into two stages: segmentation and subtraction. In segmentation, the ribs overlapping the lungs were automatically identified in the image. The segmentation was performed using two separate methods: active shape models (ASMs) and pixel classification (PC). This segmentation was then used to derive a model of the ribs. The rib models were built using existing methods that made use of principal component analysis (PCA). These models were subtracted from the images in order to suppress the ribs.

The ASM segmentation, when implemented as a semi-automated process, returned an overlap of 0.75 between algorithmically and manually segmented ribs, while PC segmentation returned an overlap of 0.61. Textural analysis on rib-suppressed images showed variable results with an improvement of 0.01 in the area under the ROC curve compared to textural analysis of the original images. Given the already high performance of the original CAD tool without rib suppression, the small improvement with the latter may not warrant its inclusion in the algorithm.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

> "The diagnosis of childhood intrathoracic tuberculosis depends on a constellation of symptoms, signs, and tuberculin skin test and chest radiograph findings. As a result, making a diagnosis is fraught with difficulty." (Gie, 2003, p. i)

The World Health Organisation estimated that a total of 1.4 million deaths occurred from TB globally in 2010 (WHO, 2011a). The rate of incidence is still increasing in South Africa, despite worldwide trends showing a decrease (WHO, 2011a). TB in adults is primarily diagnosed using sputum tests and chest X-ray images. TB in children, however, is more difficult to diagnose as sputum tests are often not possible and paediatric chest X-ray images are difficult to interpret (WHO, 2006).

In his MSc thesis, Mouton (2009) proposed a computer-aided diagnosis (CAD) tool to aid in the diagnosis of TB from chest X-ray images, specifically in paediatric cases. The tool initially performed pre-processing and segmentation on the chest radiographs, before using a texture-based algorithm to detect abnormalities. Pulmonary TB is diagnosed in the lung region of X-ray images, which is obscured by the overlapping ribcage. The effect of the ribs on Mouton's algorithm was decreased by the division of the lungs into smaller regions. These regions were identified in each image and their textures were analysed individually, incorporating the rib shadows as a local texture phenomenon of the lungs. Mouton did, however, indicate in his recommendations that there is a need for a more direct method to reduce the effects of background anatomy.

A bony structure suppression tool would provide a means of decreasing the effects of the overlapping ribcage. The rib shadows create false positives as well as false negatives with regard to the detection of abnormalities in chest X-ray images. False positives are created at rib crossings where the overlapping, pale rib-shadows create a bright spot in the image. False

negatives are created where the rib body obscures existing abnormalities, which may be closer in greyscale value to bone than to soft tissue. The false positives and false negatives generated by rib shadows have a negative impact on the performance of CAD.

Another more direct benefit of bony structure suppression is to radiologists. In a study by Samei et al. (1999), it was shown that anatomic noise is more limiting on the detection of circular nodules in chest X-ray images by human observers than quantum noise, which is present due to the nature of the imaging modality. Observers were able to detect smaller nodules in images with quantum noise than in images with anatomic noise.

In an attempt to improve the performance of the tool created by Mouton, this project implemented a rib suppression algorithm. The hypothesis to be tested was that Mouton's pulmonary CAD tool would perform better if overlapping bony structures, such as the ribcage, were removed from the chest X-ray images. The effect of the bony structure suppression algorithm was quantified through the resultant change in the output of a CAD tool that analyses lung texture.

1.1 Objectives

Analysis of chest radiographs is inhibited by the presence of bony structures. To aid the diagnosis of abnormalities in chest X-ray images, this project proposed to:

- Implement an appropriate rib suppression tool
- Analyse its effects on a texture-based classification algorithm

The second objective required the use of a texture-based classifier. As this project was an extension of the work performed by Mouton (2009), the same texture-based classifier was used. This tool had been created using MATLAB (MathWorks), which was also used here for bony structure suppression.

1.2 Thesis Outline

A review of literature relevant to this thesis is presented in Chapter 2. The image data used in the study is detailed in Chapter 3. The theoretical approaches used in the study are described in Chapter 4. The methods used to perform rib segmentation and suppression are described in Chapters 5 and 6 respectively. Some detail is given on Mouton's CAD tool and its implementation after rib suppression in Chapter 7. Given the methods described and justified in the preceding chapters, the results of the project are reported in Chapter 8 and discussed in Chapter 9.

Chapter 2

Literature Review

This project created a bony structure suppression tool. The effectiveness of the tool was determined by measuring its effect on a texture-based computer-aided diagnosis (CAD) algorithm. The tool is intended to be used in diagnosis of pulmonary TB in chest X-ray images. The fields of research involved were thus CAD (specifically texture analysis) and bony structure suppression. Separately, the two fields are well documented and articles describing studies can be found as early as the 1970's. Relevant studies in these areas are discussed in this section and information on the diagnosis of tuberculosis is provided as a background. Also described are the commercial products available that offer bone suppression.

2.1 Tuberculosis

Tuberculosis (TB) is an infectious disease that commonly affects the lungs. Few infected patients develop the active form of the disease. Those that do present with symptoms, transmit the disease through sneezing and coughing which creates airborne droplets of the bacteria, mycobacterium tuberculosis (MTB). Prolonged exposure to these airborne droplets can lead to infection. Common symptoms of TB are chest pain, prolonged cough, fever and night chills. The body is typically able to isolate the bacteria into small, walled-off nodules. These nodules often calcify and produce the characteristic shadow pattern seen in chest radiographs - this form of TB is described as post-primary (Fonseca-Santos, 2005). Post-primary TB is characterised by cavitation, the formation of a cavity due to degeneration in the nodules, and alveolar infiltration, or buildup of fluid in the alveoli (Gie, 2003).

Post-primary TB is uncommon in children (Harisinghani et al., 2000), children typically develop primary TB, where the hilar lymph glands enlarge (called hilar lymphadenopathy, see

Figure 2.1). Many complications can arise from an active infection. Enlargement of the lymph glands can trap the bronchus, possibly leading to a hyper-inflated lung or, in complete closure, a collapsed lobe. Should the infection spread to the pleura it will cause a pleural effusion - a buildup of fluid between the pleura. Young children are likely to develop unusual forms of TB compared to adults, partly caused by more compliant airways and a different immune response (Gie, 2003).

2.1.1 TB Diagnosis

TB in adults is diagnosed by tuberculin skin test, analysis of sputum and chest radiography. A sputum smear is the most common form of diagnosis worldwide (WHO, 2011a). Diagnosis in the paediatric case, in high income countries, relies on the following (Swaminathan and Rekha, 2010):

1. Contact with a TB infected individual
2. Positive tuberculin skin test
3. Abnormalities in a chest radiograph

In low income countries where the burden of TB is high, contact with a TB infected individual is difficult to determine due to lower detection rates and larger numbers of TB infected individuals (Swaminathan and Rekha, 2010). The tuberculin skin test can return false positives in patients with latent infections and false negatives in those with compromised immune systems, which could result from HIV infection, for example. Analysis of sputum in children is often unreliable (Marais and Madhukar, 2007). Diagnosis in children relies more heavily on the interpretation of chest X-ray images than in the adult case. Once diagnosed, the disease is treated with a course of antibiotics. Penetration of the nodules is difficult and treatment courses are usually several months in duration.

2.1.2 TB in South Africa

TB was first brought to South Africa by immigrants from Europe in the 17th century (Abdool Karim et al., 2009). Along with HIV, it was exacerbated in the poor living conditions and medical facilities of the mines during apartheid (Abdool Karim et al., 2009). In 1996, TB was declared a national emergency in South Africa (Weyer, 2007). Since then, many policies have been put in place to address the HIV/TB epidemic (Weyer, 2007). South Africa is the only country in Africa that has continuous surveillance of TB cases (WHO, 2010) and spends more, as a percentage of its gross domestic product, than any other African country on healthcare (Kleinert and Horton, 2009).

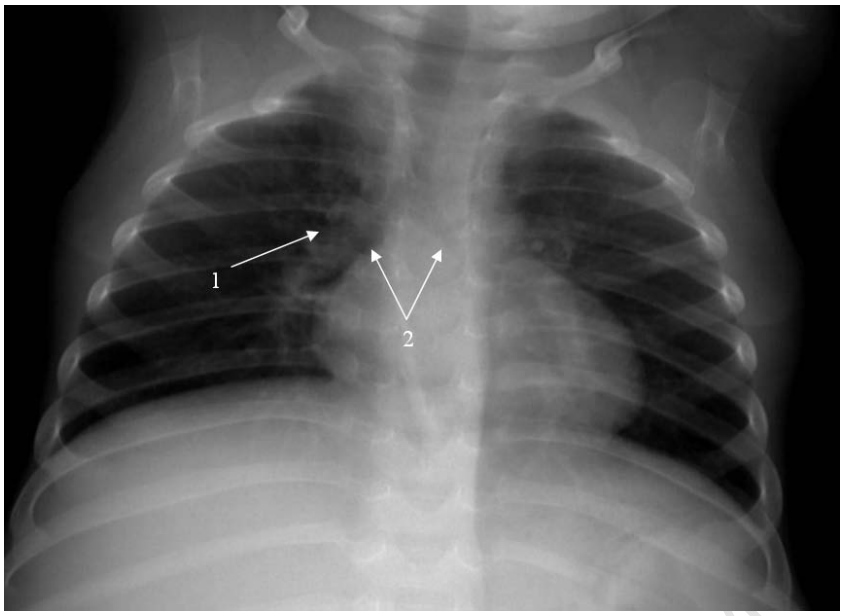


Figure 2.1: Paediatric image showing abnormality in the hilar region. The hilar region is labelled as 1 and the bronchus as 2.

Despite this, cure rates are around 50%, showing inadequate treatment (Weyer, 2007). Multi-drug resistant TB (MDR-TB) is a form of the disease that does not respond to first-line TB drugs and is more common amongst patients who have defaulted on treatment (WHO, 2010). Second-line TB treatments are significantly more expensive than first-line (WHO, 2010). As such, poor management of TB treatment will lead to a more expensive and more difficult to treat epidemic. The percentage of MDR-TB cases of all TB cases in South Africa is low compared to other countries such as China and Russia, but as countries in the African region have higher overall TB incidence rates, the number of MDR-TB cases is larger than that of more populous countries (WHO, 2010).

Extensively drug resistant tuberculosis or XDR-TB is MDR-TB that is further resistant to three or more second-line tuberculosis drugs. In 2006, an outbreak of XDR-TB was diagnosed in an area of KwaZulu-Natal with 53 patients. Fifty-two of the patients died with a median survival of 16 days (Gandhi et al., 2006). Of the 53 patients, 44 were tested for HIV and resulted positive. Two of the XDR-TB patients were healthcare workers, showing the direct negative impact of TB on the capacity of the healthcare workforce. Gandhi et al. (2006) indicate that it is likely the infections were nosocomial. By September 2007, 266 cases of XDR-TB had been diagnosed with a mortality rate of 84% in the same area (Friedland, 2008). Sixty percent of newly diagnosed TB patients in South Africa are tested HIV positive (Weyer, 2007). In studies in three eastern European countries a greater risk of MDR-TB has been shown in HIV positive individuals (WHO, 2010). Worldwide, however, there is insufficient data to support this correspondence as of yet.

HIV management brings together vulnerable and TB affected individuals: progress in HIV management in SA will be lost if similar progress is not made with TB (Weyer, 2007). MDR-

TB represents a risk that TB management will rapidly become more costly. The increased mortality rates lead to South Africans dying while economically active. Each TB patient puts a burden on their family and support network as they are unable to work (Bond et al., 2009). Nonetheless, the combined epidemic of TB and HIV has forced innovation in the healthcare system (Abdool Karim et al., 2009).

2.2 Computer-Aided Diagnosis

A computer-aided diagnosis tool for chest radiographs could provide significant benefits to paediatric TB diagnosis. These images are difficult to interpret, needing experienced viewers. Most of Sub-Saharan Africa falls in the critical shortage classification for health workers (WHO, 2011b). Sub-Saharan Africa also shows consistently high incidence rates for TB, as can be seen in Figure 2.2. CAD can assist radiologists by providing another opinion where specialists are in short supply and/or lacking in experience, and may increase the accuracy of diagnosis.

2.2.1 History

Computer-aided diagnosis using chest X-ray images is one of the oldest forms of CAD, first appearing in literature in 1964 (van Ginneken et al., 2001). Two articles were published that year which calculated the cardiothoracic ratio from scanned images automatically (Becker et al., 1964; Meyers et al., 1964). In these studies, 37 X-ray images were scanned into an image matrix of 320 by 500 and the vertical density was used to calculate the width of the chest and heart. At the time of the articles, the general hope was to create complete diagnostic programs that would remove the need of a doctor (van Ginneken et al., 2001). Since then, however, the research field has drifted towards the more realistic goal of aiding diagnosis rather than replacing it. Nonetheless, some researchers feel that CAD systems will surpass humans within the next 50 years (Doi, 2007).

CAD research owes a lot of its attention to its the success in the field of mammography (van Ginneken et al., 2001). Many screening programs using mammography have been performed, resulting in large databases of images to be scanned by radiologists. CAD is used to assist radiologists with processing these large amounts of data and success has been reported by several authors (Giger et al., 2000; Freer and Ulissey, 2001; Butler et al., 2004; Cupples et al., 2005). CAD is also used with other forms of medical imaging such as MRI, CT and even photographs of patients.

CAD would not have been practical without the inevitable development of digital radiography. Digital scanners allow a greater dynamic range to be detected, improving the quality of images (Chotas et al., 1999). It also allows CAD to be integrated into the viewing platform, along

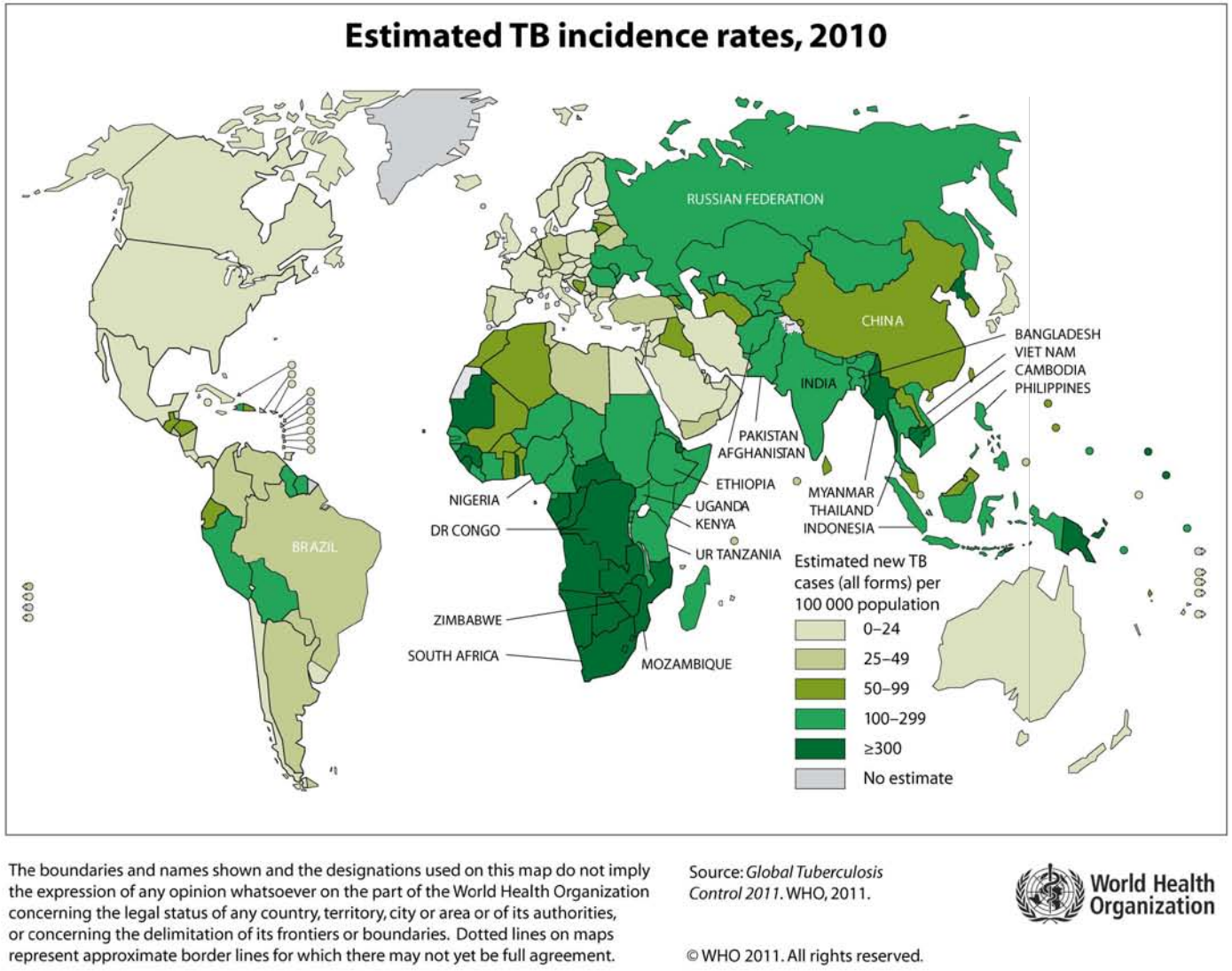


Figure 2.2: TB rate of incidence per country (WHO, 2011a).

with the picture archiving and communication system (PACS). PACS allows a consolidation of imaging and diagnosis data that greatly benefits both diagnosis and medical imaging research (Sutton, 2011).

2.2.2 Difficulties

Radiologists are able to detect most potential abnormalities in chest radiographs in the first few seconds. The difficulty lies in the decision-making regarding the lesions - de Hoop et al. (2010) determined that the main cause of disregarded abnormalities was not that they were not seen, but rather that they were incorrectly interpreted. In fact, Samulski et al. (2011) found that increasing the time spent searching the image rarely leads to the discovery of more abnormalities and possibly increases errors.

Samulski et al. (2011) investigated the presentation of CAD results to human radiograph viewers. It was found that an interactive method could possibly be more effective than a static presentation of the CAD results. The interactive method required the image viewer to query regions before showing the CAD abnormality score for those regions. It was shown that the interactive CAD did improve detection, performing better than radiologists with use of traditional CAD (Samulski et al., 2011). Of note is that regions obscured by other anatomy, such as the ribs, showed worse results as they were often not queried.

CAD tools perform well in finding nodules caused by lung cancer but have a high false positive rate. The problem of CAD tools remains the same as that of radiologists: not in spotting abnormalities, but in deciding whether they are indicative of pathology. Decreasing the false positive rate of CAD tools is key to improving their performance. Shah et al. (2003) showed that the majority of missed lung cancer lesions were located behind the ribs.

2.2.3 Classification

Classification, or specifically statistical classification, is the process of labelling instances of data as belonging to certain classes or types. This is a form of machine learning and the classes are user or data dictated. Normally the number of classes of data must be finite. Classification with an infinite amount of possible classes is considered regression (Loog et al., 2006). Each instance of data will have a number of properties that are used to compare data instances. Supervised learning bases the classification on previous examples of data that have already been classified. Unsupervised learning attempts to classify data based on clustering in the data (Hastie et al., 2009, p. 2).

There are a large number of methods available to perform classification: nearest-neighbours, linear discriminant analysis (LDA), support vector machine (SVM), neural networks and many

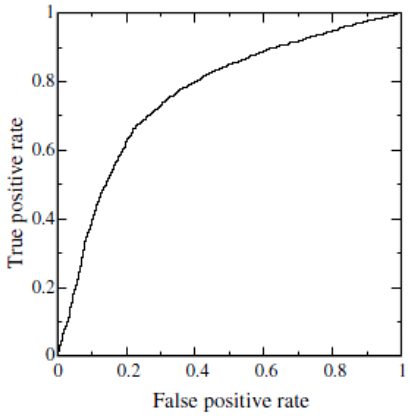


Figure 2.3: A typical ROC plot given by Fawcett (2006).

more. (Hastie et al., 2009). Thresholding is the most basic method that can be used for image classification. Nearest-neighbours is a simple method that classifies an instance as the class of the nearest training data instances. Nearest-neighbours is simple to implement and often outperforms more complex methods but can be unstable. It relies on the assumption that the data can be approximated by a locally constant function. LDA assumes the features of the data instance can be considered independent and be combined linearly to determine the class. This method is well suited for classes with Gaussian probability densities. SVM is suitable for data with a large number of dimensions or features that are highly correlated. SVM creates a hyperplane between the classes of training data in order to separate them. Neural networks consist of nodes or neurons that connect the input parameters to the output and whose functions are determined through training. Neural networks are quite powerful and can model non-linear functions. However they have a large number of parameters and can be difficult to train.

CAD attempts to classify images or areas thereof as normal (healthy) or abnormal (pathological). Mouton's (2009) CAD tool uses texture classification to analyse the chest X-ray images. Lung abnormalities change the appearance or texture of the lung tissue which is used as the basis of the classification.

2.2.4 Evaluating Classifiers

The ROC curve is a plot of the sensitivity of a classifier as, shown in Figure 2.3. It shows the true positive rate, the proportion of abnormal images correctly identified, against the false positive rate, the proportion of abnormal images classified normal. The area under the ROC curve, AUC or A_z value is commonly used to compare systems and the value itself gives the probability a random positive instance in the data will be scored higher than a random negative instance (Fawcett, 2006). A positive instance here is an image displaying abnormalities and a negative instance an image with no visible abnormalities.

Other methods used to quantify the results of classification include:

Sensitivity The ratio of true positives to the total number of positives in the database

Accuracy The ratio of correctly identified instances to total instances

Specificity The ratio of true negatives to the total number of negatives in the database

Care should be taken with these metrics as a classifier that always classifies instances as positive will achieve perfect sensitivity. Conversely always classifying instances as negative will result in perfect specificity. The output of a classifier should always be checked to ensure neither of these situations have occurred.

2.2.5 Texture Analysis

Classification of the texture of an image is often used by CAD systems to identify normal or abnormal tissue in a chest radiograph, where abnormality is caused by pathology of the lung tissue. The definition of texture used by Chen and Pau (1998, p. 208) is "a function of spatial variation in pixel intensities". A texture-based classifier then, in the context of this project, is an algorithm that uses this variation in pixel intensity in order to identify image regions that have a high likelihood of representing abnormal tissue in the lung.

A number of pre-processing steps are generally performed before texture classification. One commonly used step involves flipping one lung along its axis, matching the opposite lung in orientation, in order to allow textures from both lungs to be compared to a single training set. The inside of the lung region is also often reflected onto the outside of the lung border to reduce the effects of the surrounding tissue on filters acting around the periphery of the lung (Mouton, 2009; Arzhaeva et al., 2007).

Chest radiographs used for textural analysis should preferably have high resolution. Images of higher resolution provide more information about the tissue in a given area allowing more detailed analysis of the tissue represented. As explained by Castellano et al. (2004) - while texture analysis aims to provide more information about an image, it is limited by the resolution of the image.

The ribs were briefly mentioned by van Ginneken et al. (2002) as structures to which texture classifiers must be insensitive. As texture analysis attempts to analyse patterns in a specific structure, overlapping but unrelated patterns caused by background anatomic features would make this more difficult. Though this step may seem obvious, there have been few studies (Hogeweg et al., 2010b) that have directly shown the effect of "removing" the ribcage on textural analysis, although studies have circumvented the possible effects of the ribs by only analysing inter-rib regions (Powell et al., 1988).

2.3 Mouton's CAD System

Mouton's (2009) texture classifier was based on the work performed by van Ginneken et al. (2002). The four main steps taken in Mouton's software (as illustrated in Figure 2.4) were: lung segmentation, division of the lung region into smaller regions of interest (ROI), feature extraction and classification. Segmentation was performed first in order to identify the lung area of the image in which to analyse the image texture.

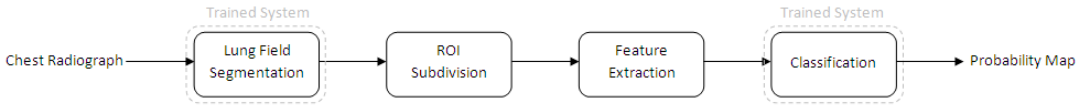


Figure 2.4: Flowchart showing sequence of steps taken by Mouton's algorithm (2009).

Lung segmentation was necessary to locate the focus area of the software and is a necessary step for most CAD tools. In texture analysis, it serves to locate the data that the algorithm needs to analyse. Van Ginneken et al. (2006) provided two possible definitions for the lung fields in a chest X-ray, firstly pixels that represent radiation that has passed through the lungs and secondly, pixels that represent radiation that has passed through the lungs but not through the "mediastinum, the heart, structures below the diaphragm, and the aorta". The first definition was not practical for lung segmentation in images as it is not always possible to identify radiation that has passed through the lungs where other structures dominate and so the second definition was the one usually taken in lung segmentation. In the context of this project, only the second definition was considered.

Lung segmentation was followed by division of the lung region into a number of smaller, overlapping areas or regions of interest (ROI). The ROI division was done in two ways: a small circular region and a big region approach. Van Ginneken et al. (2002) suggested that by using ROIs, anatomical variation was incorporated in the region classifier. It also increases the amount of information used from the training images.

The pre-processing steps of left-lung flipping and mirroring pixels around the lung border were performed prior to feature extraction. Flipping of the left lung allows the comparison of corresponding regions in each lung and assumes the lungs are identical. This effectively doubled the size of the training set. Given the small dataset, this was advantageous. Mirroring pixels around the border of the lungs provided a buffer for image filters that considered adjacent pixels, so only lung tissue would be considered.

Mouton (2009) investigated the texture of chest X-ray images by analysing the outputs of a bank of Gaussian filters, given the image as an input. The output of this bank of filters is a vector describing the local image texture. Thirty images were produced from the filters and the mean, standard deviation, skew and kurtosis of each image was used. The length of the feature vector was 124 (the original image was also included).

This vector was then used to perform classification which can be done with a number of different algorithms. Arzhaeva et al. (2007) made use of four classifiers: linear discriminant analysis (LDA), quadratic discriminant analysis (QDA), k-nearest-neighbours (k-NN) and a support vector machine (SVM). Both Mouton (2009) and van Ginneken et al. (2002) made use of a k-NN classifier which is one of the simpler, but still well performing, algorithms to implement. The k-NN classifier computes the distance-weighted average of k neighbours. This was not practical for the circular ROI approach due to the large number of regions and a LDA classifier was used instead. The result of the classifiers was a score in the range $[0, 1]$ for the area being considered.

This feature classification was not performed for the image as a whole, but for each of the regions. The output of this classification was presented in two ways: a probability map of the lung region and an overall image score. The probability map was created by giving each lung-region pixel the average probability for each of the regions containing the pixel. The image score was calculated by using a weighted multiplier to combine the region scores. The weights were the performance of the classifier in each region. Each region was classified by the texture classifier and the results compared to the radiologist annotations to calculate the A_z, or area under the ROC curve, value for each region. Furthermore, if a region performed below a certain threshold, it was excluded in the score calculations. The resulting scores indicated the posterior probability that the area described contained abnormalities.

A weakness of the system developed by Mouton was that the training data was neither histologically proven to be TB nor was follow-up data collection made. For this reason the "diagnosis" provided by the tool can only be said to be pulmonary abnormalities and not specifically TB. In low-incidence countries, a strong factor in the diagnosis of pulmonary TB is contact with a TB infected individual. In high-incidence populations, like South Africa, this is not a reliable or necessary factor. Given the difficulty in differentiating between TB and other lung pathologies, specifically HIV related (Gie, 2003), CAD for TB detection may require additional information, beyond the chest X-ray image.

The studies described in Section 2.2.2 were all intended to aid in the diagnosis of lung cancer. Typically small nodules were used for training and testing data, presumably as they are the most difficult to detect. Pulmonary TB in children tends to present with large abnormalities or lymphadenopathy. The study of Mouton (2009) performed well, with an area under the ROC curve of 0.982, and shows that the CAD tools translate well to suspected TB.

2.4 Bony Structure Suppression in Chest Radiographs

Van Ginneken et al. (2001) divided computer analysis of chest radiographs into three areas of research: general processing, segmentation and analysis. General processing includes image

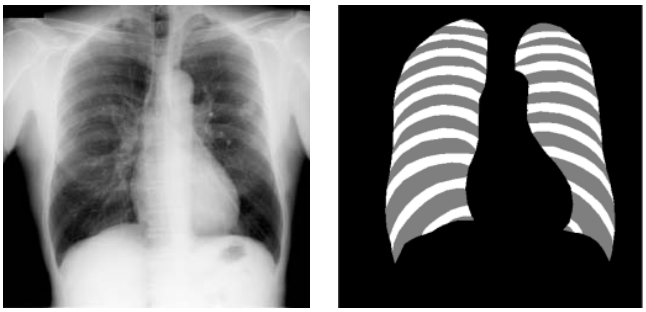


Figure 2.5: Images showing segmentation of ribs (Loog et al., 2006), where the image on the left is the original chest radiograph and the image on the right is the segmented image. Black is background area, grey is intercostal space and white is costal space.

enhancement and subtraction techniques, while analysis covers the majority of CAD techniques, such as nodule detection and texture analysis. Suppression of bony structures, such as the ribs or clavicles, often initially requires identifying (segmenting) the structures in chest radiographs. Thus, a large number of studies relevant to the identification and suppression of the ribs fall under segmentation, along with segmentation of the heart and lungs.

The general approach taken to decrease or analyse the effects of overlapping bony structures in lung analysis was to perform segmentation of the radiograph first (van Ginneken et al., 2001). This segmentation allows the areas of the image where bony structures overlap the lung region to be located (an example is shown in Figure 2.5). Enhancement techniques can then be applied to these regions to make the soft tissue more visible and bony structures less visible. The main culprit of anatomic background noise, in the lung region, is the ribcage.

The first article on rib segmentation is by Toriwaki et al. published in 1973 (Yue et al., 1995; van Ginneken and ter Haar Romeny, 2000). Since then many articles have been published on the subject. One of the more well known authors in rib segmentation, Doi, placed a patent (Doi and Katsuragawa, 1989) in the field in the late eighties and has contributed to a large number of papers in CAD. Van Ginneken published research in the area later (van Ginneken and ter Haar Romeny, 2000) and has been involved in a number of projects since.

Recently, a small number of studies have been performed that do not make use of rib segmentation in the suppression of background anatomic noise. One of the more well known examples was by Suzuki et al. (2006) in which an artificial neural network and dual-energy images were used. These studies are discussed in Section 2.4.4 and dual-energy images are discussed next in Section 2.4.1.

The field of bony structure suppression can be further broken down in a rough division suggested by Loog et al. (2006), where suppression techniques can be global or local. For example, approaches using geometric models such as parabola fitting techniques (Yue et al., 1995) are global, while a regression based technique (Loog et al., 2006) would be classed as

Figure 2.6: Example images from dual-energy subtraction, left: standard x-ray image, middle: bone image, right: soft tissue image (Uemura et al., 2005).

local. A local approach only makes use of local pixel information in the image being analysed to determine the algorithm output of that specific area.

2.4.1 Dual-Energy Imaging

Dual-energy imaging is a means of providing chest radiographs without rib shadows. It is a technique in which the X-ray beam projected through the subject's chest is simultaneously (single exposure) recorded by two receivers that record radiation of different characteristics, low energy and high energy (Bjorkholm, 1985). High energy radiation shows bone structures well while low energy is better for imaging soft tissue. These recorded images are then combined through subtraction to create images in which bone and soft tissue are preferentially emphasized (Armato, 2005). (This technique can also be performed by taking the two images at different times, referred to as double exposure.) The soft tissue image provided by this technique can be considered the ideal result for a bone suppression algorithm as the majority of the bones are not visible in these images (can be seen in Figure 2.6).

Dual-energy imaging requires specialized equipment which is not as widespread as standard chest X-raying equipment. Another disadvantage is increased radiation dose, for both single and double exposure, which excludes the technique as a viable option in the imaging of children, who are more susceptible than adults to the damage caused by ionising radiation. Children are also unlikely to remain still during the imaging process. Images taken at different times, in double exposure techniques, would therefore have to be registered to detect patient movement. For these reasons, dual-energy imaging is not attractive for paediatric imaging.

2.4.2 Rib Segmentation

Van Ginneken et al. (2001) performed a survey of over 150 articles published in the previous three decades in the field of computer analysis of chest radiographs. They describe the typical approach to rib segmentation as consisting of three stages:

- Selecting a geometrical model (such as ellipses or parabolas) and edge detection (pre-processing).
- Fitting of models to the detected edges.
- Selecting the most suitable candidate model (post-processing).

Pre-processing may also include steps such as enhancement and compensation for the nature of the imaging modality, as well as edge detection. Post-processing includes recognising which fitted models are clearly in error and selecting the most likely candidates from the remainder.

The ribs are mainly described by their edges and, as such, edge detection aids in their identification. In model fitting, the results of the edge detection are commonly fed into a modified Hough transform. A Hough transform takes points from the image plane and converts them to curves in the parameter plane. Points that are collinear in the image plane will create intersecting curves in the parameter plane (Duda and Hart, 1972). By detecting these intersections, lines or curves can be detected in the image. Yue et al. (1995) indicated that the first use of the Hough transform was to detect "sparse straight lines" in images. The Hough transform is commonly used to detect curves in images and was used by Wechsler and Sklansky (1977), Yue et al. (1995) and Vogelsang et al. (1998) in the detection of the ribs. A disadvantage of the Hough transform is its relatively high computational requirements (Yue et al., 1995) but is slowly becoming more competitive with improvements in technology and modifications to the methods (Fernandes and Oliveira, 2008).

The model fitting approach described above may be defined as a global approach (Loog et al., 2006). Another global approach to rib segmentation fitted a model of the ribcage, rather than individual ribs, to chest radiographs (van Ginneken and ter Haar Romeny, 2000). Principal component analysis (PCA) finds the principal modes of variation of a dataset and was used to reduce the number of model parameters by transforming the original parameters to a new set ordered by the explained variation. These parameters included representations for the height of the rib cage, the size of ribs relative to the intercostal spaces, asymmetry, differences in spaces between the first few ribs and lower ribs and the shape of individual ribs. The best result shown in this study, measured by the fraction of lung field correctly classified as ribs or intercostal spaces, was an accuracy of 0.882. The rib model used a set of two parabolas to define the upper and lower border of each rib but the choice of model was a compromise between accuracy and computational time.

An example of a local method for segmentation used iterated contextual pixel classification (ICPC) (Loog and van Ginneken, 2006). This is a supervised technique that achieves segmentation by iteratively reclassifying pixels, using their context, to maximize the posterior probability of the segmentation selection being correct. Accuracy was measured in comparison with segmentation performed by the author assisted by a radiologist and with a segmentation created

by a second observer, without the radiologist's assistance. The algorithm performed better than a simple pixel classification technique and the results, while not better than, were comparable to that of, the second observer.

The ICPC technique (Loog and van Ginneken, 2006) and the neural network approach taken by Suzuki et al. (2006), as described in Section 2.4.4, are the two main examples of supervised methods used in literature. Supervised methods require training images on which the algorithm bases its classifications. Training images can be from dual-energy imaging techniques or from images where the rib borders are outlined by a radiologist.

Active shape model (ASM) fitting is a supervised technique and has been employed in segmenting ribs in chest X-ray images. Ramachandran et al. (2002) made use of ASMs to locate the intercostal regions of the lungs. The segmentation was performed using four models: a left and right side model for a "short" lung and a left and right side model for a "long" lung. They found that two sets of models were necessary to cater for variation in the number of visible ribs. A long lung was considered to have 9 visible parenchymal regions and a short lung 8 regions. De Bruijne and Nielsen (2005) used an ASM to detect each rib individually. Both of these studies emphasized the difficulty of rib segmentation using ASMs due to the varying number of ribs visible in chest radiographs.

In the study performed by Sarkar and Chaudhuri (1998), the purpose was to quantify the characteristics of TB in chest radiographs in order to monitor recovery. Of the effects of TB on the lungs, only infiltration and cavitation were considered. The image was initially analysed by segmenting the ribcage and locating the clavicles (using the Hough Transform). A series of simple image processing steps, filtering and thresholding, was applied to the image in the pre-processing stage to obtain a binary image showing likely rib area. The intercostal distances were determined by analysis of the vertical profiles for rib curve parameters. Curve fitting was performed using a least-squares-based method. Dilation was then used along the detected ribs. Sarkar and Chaudhuri (1998) used this segmentation to perform rib suppression, using the average difference in pixel values between the rib and inter-rib area. Further segmentation was performed on the image to detect infiltration, to quantify its extent and thus the degree of recovery.

Summary of Segmentation Methods

Table 2.1 summarises the advantages and disadvantages of the segmentation methods described in Section 2.4. The suitability of the methods for this project is also mentioned. The advantages and disadvantages considered include computational complexity and need for training images.

Table 2.1: Summary of rib segmentation methods

Authors	Method	Advantages	Disadvantages	Suitability for this project
van Ginneken and ter Haar Romeny (2000)	Ribcage model fitted to image	Computational complexity contained in training, can infer the position of less visible ribs	Uses training images which are segmented, shows reduced accuracy near border of lung region	Can be used, but may show poor performance on paediatric images due to greater anatomical variation, requires segmented images
Loog and van Ginneken (2006)	Iterated contextual pixel classification	Allows for varying rib positions, does not force a pre-defined shape in segmentation	Computationally complex method, many internal variables to determine, likely to miss a rib	Can be used, but will be computationally expensive and will require initial lung segmentation
Sarkar and Chaudhuri (1998)	Parabola fitting and curve dilation	Allows for varying rib positions, relatively simple algorithm	Method is only tested on two images, assumes ribs have a parabolic shape	Can be used, but may have low border accuracy
de Bruijne and Nielsen (2005)	Active shape models used to detect ribs individually	Allows for a varying number of visible ribs, incorporates proximity of ribs to each other, increases training data per image	Does not make use of rib relationships, uses segmented training images	Can be used, but requires segmented images and will be computationally expensive
Ramachandran et al. (2002)	Active shape models used to detect parenchymal regions	Allows for a varying number of ribs, incorporates relationship of neighbouring ribs	Requires segmented training images and an indication of the number of ribs visible	Can be used, but requires segmented images and added information about the image
Yue et al. (1995)	Hough transform and active contour model	Makes use of anatomic knowledge	Computationally expensive method	Can be used, but will be computationally expensive

2.4.3 Rib Subtraction

In order to perform bony structure suppression, rib segmentation can be followed by further image processing to increase the visibility of the soft tissue under the rib shadows, or expressed differently, to suppress the ribs. Studies that have performed rib segmentation together with suppression of the ribs, include work by Vogelsang et al. (1998), Ahmed et al. (2007), Sarkar and Chaudhuri (1998) and Hogeweg et al. (2010b).

Vogelsang et al. (1998) made use of a "combined anatomical and grey-level model of the ribs" which is shown in Figure 2.7. Its exact derivation was not given. The model greyscale values were subtracted from the chest radiograph in the rib areas located by rib segmentation methods. The model was modified to include the estimated thickness of the rib at each point. The author reported success on the 154 images tested.

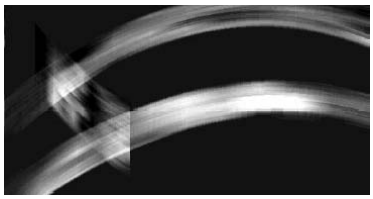


Figure 2.7: Greyscale model of the ribs used in bony structure suppression by Vogelsang et al. (1998).

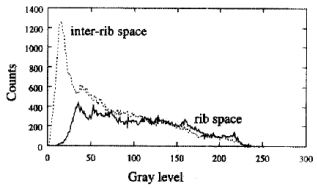


Figure 2.8: Histogram of the greyscale values of the lung region and rib area as shown by Sarkar and Chaudhuri (1998).

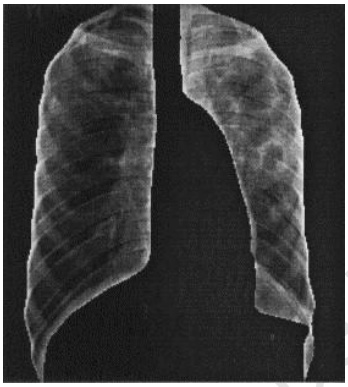


Figure 2.9: Rib suppressed image by Sarkar and Chaudhuri (1998).

In the study that made use of independent component analysis (ICA) (Ahmed et al., 2007), the only results provided were a few images. ICA separates signals that have been mixed linearly, given N observations of the mixed signal for N mixed signals. It was used to estimate the soft tissue of the chest radiograph without the overlapping bony structures. The segmented lung area was used as one observation and the segmented rib area as the second observation. The images show a slight improvement in soft tissue visibility in the costal areas.

Sarkar and Chaudhuri (1998) performed rib suppression using their initial rib segmentation, obtained through filtering, thresholding and curve fitting (Section 2.4.2). The rib suppression was performed by using the difference in the pixel magnitude frequencies. The histogram of the pixel magnitudes is shown in Figure 2.8 and it was assumed that the difference in profile between the rib and inter-rib area was due to the presence of the ribs. This difference in intensity was subtracted from the rib segmented area to perform rib subtraction. The result of this subtraction can be seen in Figure 2.9.

Hogeweg et al. (2010b) performed bony structure suppression on a chest radiograph using an initial rib segmentation. The segmentation was performed manually by connecting detected edges in the chest radiograph. The edge detection was not specified but could have been any simple method to detect a rapid change in greyscale value in the images. Profiles of the ribs were sampled perpendicularly to the direction of the ribs and were used to create a model for subtraction of the ribs. These samples started below the lower border of the ribs and extended

past the upper border. The model was created by using the two principal components from the PCA of the profiles and was normalised by using the greyscale values in the intercostal spaces. Using this method, each rib model was derived from the image on which the suppression was performed. An example model and model subtracted image is shown in Figure 2.10. A disadvantage, mentioned by the authors, is that the method is disturbed by large overlapping structures. This bony structure suppression was tested through a CAD texture-based system for the detection of TB and results showed an increase in the performance. This method resembles the earlier patent of Maton et al. (2009) which describes the same method in slightly more detail. This patent and others are described in Section 2.5.

The profiles sampled by Hogeweg et al. (2010b) extend beyond the borders of the rib shadow to incorporate data of the intercostal spaces for normalisation, which allows for small inaccuracies in the rib shadow segmentation. Such a method is practical as no computer segmentation method reaches ground truth accuracy in every application. The segmentation used, however, would have to be automated for it to be practical for large scale use - manual connection of detected edges would take too much time.

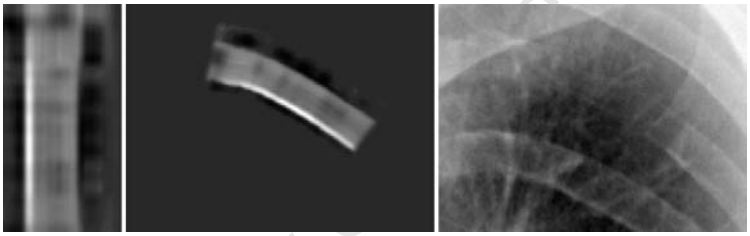


Figure 2.10: Results of rib suppression by Hogeweg et al. (2010b). The left image shows the modelled rib greyscale values, the middle image the model as fitted to the chest X-ray and the right image the rib suppressed area.

2.4.4 Other Methods

A few studies have performed bony structure suppression without performing rib segmentation. Examples of this are a study performed by Suzuki et al. (2006) using an artificial neural network and a regression based technique used by Loog et al. (2006). Such suppression techniques use only the original radiograph as an input.

The massive training artificial neural network, used by Suzuki et al. (2006), was trained using dual-energy images and made use of a multi-resolution scheme. While the network produced good results, a neural network is a complex method to implement and has a large number of internal variables that must be determined. Training the system using dual-energy images and applying it to standard X-ray images also introduces additional image processing requirements due to differing image properties across training and test image sets, for example the images

must be normalised to ensure the same range of pixel magnitudes. A database of 122 images was used where only 4 images were used for training, including the normal image. The output of the method was a bone tissue image which was then subtracted from the original radiograph. The results of the algorithm were measured using the normalized mean absolute error from the dual-energy images. A further 136 images from the Japanese Society of Radiological Technology (JSRT) database were tested where the contrast of nodules was improved in the simulated soft-tissue images.

Loog et al. (2006) used a regression technique that also used dual-energy images for training the system. In the study, the following steps were applied: pre-processing, dimensionality reduction and iterated regression. The regression analysed the features of the input image to estimate a corresponding soft tissue image and bone tissue image. The output of regression is a continuous range of values unlike classification which has a discrete outcome. Eight training images were used, each with their own set of dual-energy images. Thirty randomly selected images were chosen from JSRT for testing images. The final result was measured by thresholding the estimated bone image and comparing the results with manual segmentations. The subtraction of the estimated bone images from the original radiographs resulted in the suppressed images and was referred to as dual-energy faking.

2.5 Commercial Applications

A medical imaging software company, Riverain, offers commercially-available software for rib suppression. The tool, ClearRead Bone Suppression (Riverain) previously known as Softview, performs rib suppression on chest X-ray images. Riverain has two patent applications on object removal from images. The first patent (Maton et al., 2009) describes the same method seen in Hogeweg's publication (2010b). The second patent (Knapp, 2009) uses neural-network regression on image features to calculate bone value predictors and can be used to suppress bony structures. Li et al. (2011a) indicate that regression is used in ClearRead Bone Suppression. An example set of images showing the results of their bone suppression is given in Figure 2.11.

The effectiveness of Riverain's suppression was presented in a study (Li et al., 2011b) by the University of Chicago Medical Center. The study investigates bone suppression by presenting the suppressed images together with the original images to radiologists for the detection of lung nodules. The difference in AUC was found to be statistically significant, from 0.840 for the chest radiographs to 0.863 for chest radiographs with the bone suppression images. The tests were performed on a set 151 radiographs in the detection of subtle lung nodules. The average number of false positives detected increased from 18 to 25 with the use of the bone suppression images.

Li et al. (2011a) compared the effect of ClearRead Bone Suppression on radiologists' performance to that of dual-energy imaging. Ten observers were used in this study with 100 images,

50 normal, 50 abnormal. The observers all had experience with dual-energy images. There were significant improvements in the AUC values compared to standard radiographs. While dual-energy imaging showed the best results (0.894 to 0.945, P=.001), using bone suppression images was also an improvement over the standard radiograph alone (0.846 to 0.894, P<.001).

2.6 Conclusion

The diagnosis of pulmonary TB (PTB) in children is particularly reliant on chest X-ray images and could benefit from the use of CAD systems. Mouton et al. (2010) has shown that CAD systems can be used in the successful detection of paediatric lung abnormalities indicative of pulmonary TB. Rib suppression has been used to improve the performance of TB CAD systems (Hogeweg et al., 2010b) and was shown to improve the performance of radiologists directly (Li et al., 2011a). Given the suitability of CAD for paediatric PTB and the proven benefits of rib suppression, research should be performed to determine if rib suppression correspondingly benefits texture analysis of paediatric chest X-rays.

The algorithm would be expected to maintain a high resolution chest image and should not alter the visible borders of the lung. As background structures make the detection of patterns more difficult, it would be expected that bony structure suppression would improve the ability of a texture-based analyser. A fully automated method of rib segmentation is needed to make the method shown by Hogeweg et al. (2010b) practical to image analysis.

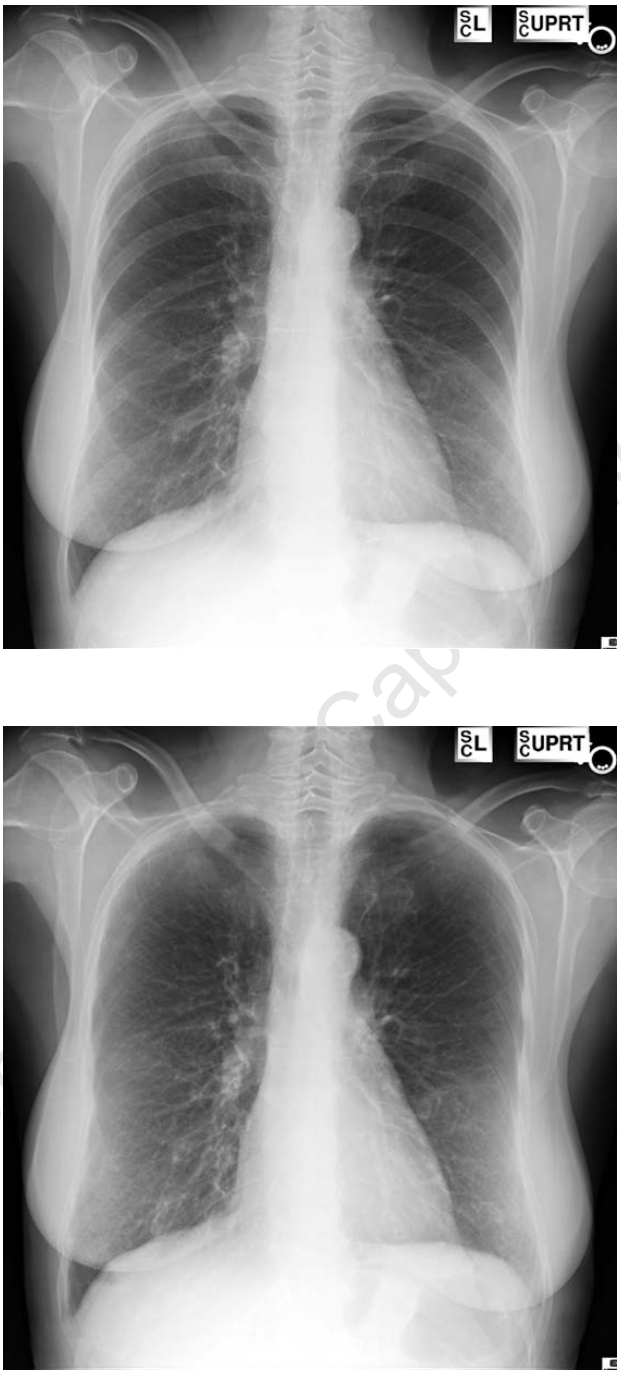


Figure 2.11: Radiograph (top) and bone suppressed image (bottom) provided by ClearRead Bone Suppression (Riverain). ClearRead performs bone suppression on the ribs overlapping the lungs. The soft tissue of the lungs is more clearly visible in the second image.

Chapter 3

Materials

Two datasets of chest X-ray images were used in this project - a set of 154 paediatric images and an adult dataset of 355 adult images. The paediatric dataset was the focus of the project while the adult images were included for comparison. These datasets are described in this chapter.

The paediatric images are antero-posterior (AP) in projection while the adult images are postero-anterior (PA). This indicates the direction of the X-ray beam. In AP imaging, the X-ray source is placed in front of the patient and the detector behind them; these positions are reversed for PA imaging. Chest X-ray images are traditionally PA in orientation and there is a slight difference to the appearance of the chest, depending on the view taken. The heart appears larger in AP images and the ribs are slightly different in shape and width (Novelline, 2004, pp. 15,82). Often AP patients are imaged lying down and so are unable to inhale as deeply as when standing. The ability to inhale is relevant as the level of air in the lungs affects the visibility of the anatomy.

3.1 Paediatric Dataset

The paediatric dataset used here consists of paediatric images obtained using the Lodox Statscan digital radiography machine located in the Trauma Unit of Red Cross Children's Hospital. There were 154 images in the dataset in total, consisting of two subsets from two different studies. All the images were taken at the Red Cross Children's Hospital. The patients were all in the age range of 0-5 years.

The Statscan machine is a digital X-ray machine that performs a full-body scan. The machine uses a linear slot-scanning technique that allows for high-quality images at a low radiation dosage (Pitcher et al., 2009). This low dose makes it preferable for paediatric imaging. The patient

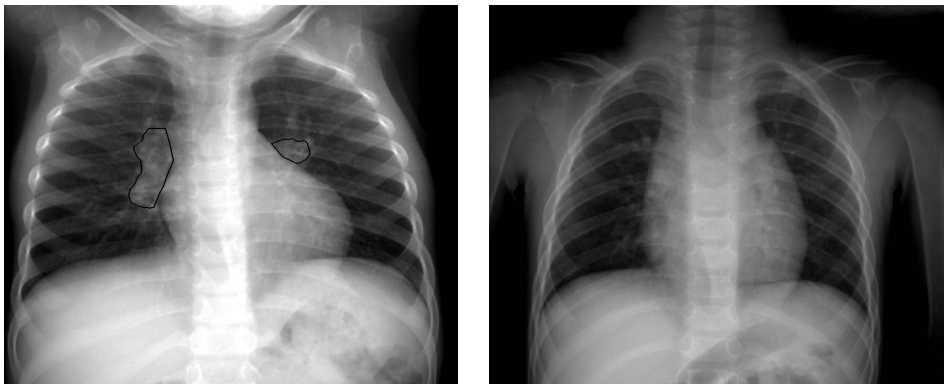


Figure 3.1: Example images from the paediatric dataset. Left: a typical image from the TB study, with the abnormalities outlined. Right: a typical image from the trauma study.

is traditionally in a supine position during imaging and the imaging angle can be rotated 100º about the mid-line. The images used here are AP views.

Images were assessed by a radiologist for signs of lung abnormality. Of the 154 images, 115 images were taken from a TB study of which 6 show no radiological signs of lung abnormalities. The remaining 39 images in the dataset were from trauma patients, all of which were normal (no observable lung abnormalities). In total, there were 45 normal images and 109 abnormal images. The images from the TB study and the trauma study were taken at differing spatial resolutions (4.16 lp/mm and 2.08 lp/mm). There was also a difference in pose between the trauma and TB images - in the TB study, the patients arms were raised while in the trauma study the patients' arms were left at their sides. In Figure 3.1, examples of the two subsets are shown. For use in this project, all the paediatric images were resampled to a resolution of 700x700 pixels, and cropped to show the chest area of the patient.

The abnormalities in the images have been outlined by a radiologist (as shown in Figure 3.1). The patients scanned during the TB study were suspected of having TB but there was no histological data for the images. Thus the abnormalities outlined in the images are not necessarily caused by pulmonary tuberculosis. As a result, the scores provided by the CAD system can only be considered in terms of lung abnormality.

In the survey by van Ginneken et al. (2001), the common distribution of pulmonary TB was concentrated in the apexes of the lungs. Analysis of the paediatric images revealed the distribution shown in Figure 3.2. The abnormalities appeared to be most common in the hilar region. The regions outlined as abnormal by the radiologist are stored as binary masks: the abnormal regions labelled 1 and normal regions 0. The distribution was determined by warping (using Mouton's methods) the abnormality masks to the mean lung shape and summing up the results. The result was an image where each pixel value is the number of images that show abnormality at that point.

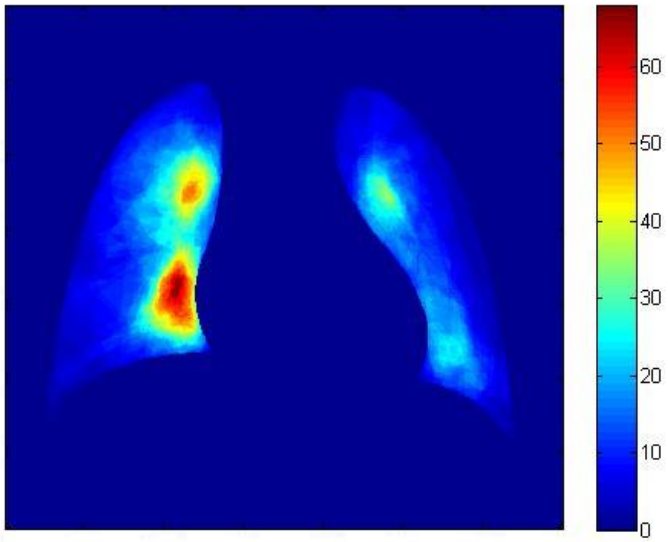


Figure 3.2: Abnormality distribution in the paediatric dataset.

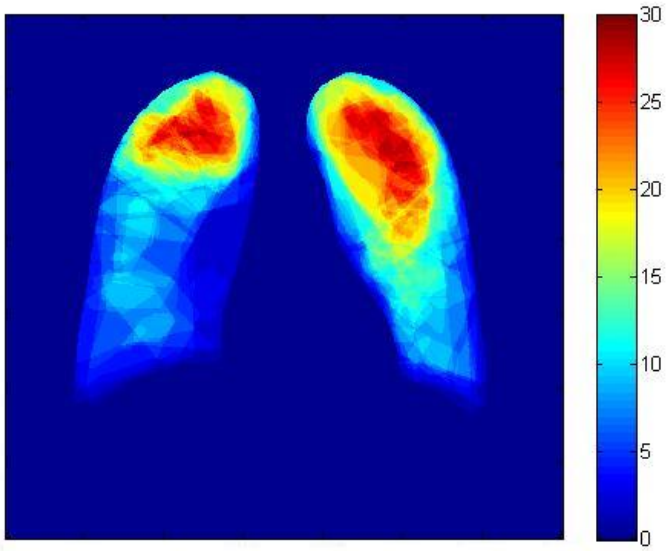


Figure 3.3: Abnormality distribution in the adult dataset.

3.2 Adult Dataset

A second dataset of adult chest X-ray images was provided by the Diagnostic Image Analysis Group (DIAG) of Radboud University Nijmegen Medical Centre. The dataset contained a total of 365 digital radiographs. The original images were 2048x2048 pixels (pixel size 0.25mm, Delft Imaging Systems). The images are PA in contrast to the AP view of the paediatric dataset. The training set consists of 216 images, 106 of which showed abnormalities. The remaining 149 images formed the test set, with 80 abnormal images. This dataset is also used by Hogeweg

et al. (2010a). The distribution of the the abnormalities can be seen in Figure 3.3 and was concentrated in the apexes of the lungs.

Chapter 4

Theory

The methods used to perform and analyse bony structure suppression are detailed in this chapter. A rib segmentation, modelling and subtraction approach was taken in this project. This could be seen as less elegant than a local approach, whose methods need only the original image as an input. Neural networks provide such a solution but require examples of images without ribs - like those provided by dual-energy imaging. Such images, however, are not desirable in the paediatric case, as explained in the previous section (Section 2.4.1). An additional benefit of separating the suppression from the segmentation is that it is a more modular approach that enables evaluation of the results of different components. Mouton's (2009) texture classifier was used here to evaluate the effect of the rib suppression implementation.

4.1 Rib Segmentation

Rib segmentation was performed in this project in preparation for rib suppression. The suppression methods used required the upper and lower borders of the ribs to be identified. Figure 4.1 shows the interaction of the segmentation and suppression methods.

Two methods were used for the segmentation of the ribs: active shape models (ASM) and pixel classification (PC). PC has been shown to be successful in locating the ribs in chest X-rays (Loog and van Ginneken, 2006) but can fail if there is extensive pathology which changes the appearance of normal structures. Mouton (2009) suggested that a shape based method such as ASM could be more robust to these changes. Ramachandran et al. (2002) used ASM to segment the intercostal and identify the ribs. PC is less strict with regard to the structure of the final result and can return improbable shapes, whereas ASM limits segmentations to probable shapes. PC and ASM have both been used to segment the ribs and were selected here as methods with contrasting properties.

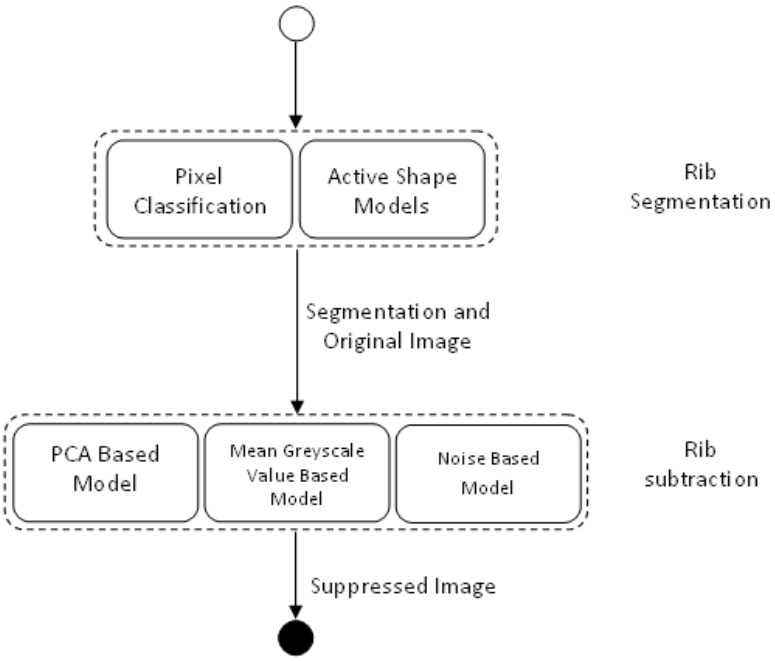


Figure 4.1: Flowchart showing rib segmentation and bony structure suppression method interaction. At the start of the method the images is segmented by one of the two segmentation methods. The segmentation is then provided, together with the radiograph, to three different suppression methods. The output of each of the suppressed methods is a radiograph with the ribs suppressed.

4.1.1 Active Shape Models

Active shape models (ASMs) are a statistical shape modelling method, first described by Cootes and Taylor (1992). Used for segmentation, an ASM is a deformable template matching model that is iteratively applied to an image. The final shape and position of the deformable model, given by the algorithm, is used to provide the image segmentation. It is a knowledge-based, supervised method that is general enough to be readily applied to more than one application.

ASM uses a statistical shape model built from training image segmentations. The training segmentations provide examples of the different shapes the segmented structure can take. The acceptable forms of shape deformation are summarised in a point distribution model (PDM). The pixel values in the training images are sampled along profiles perpendicular to the segmentation contours and their normalised derivative taken. This creates a sample set of the image gradient around the contour points. The mean shape is fitted to the test image and the corresponding profiles through the contours are sampled. The distance of these samples from the model sample set is used to determine the necessary model deformations to reduce this distance. The PDM is used to guide these deformations of the model and limit them to allowable shapes. The parameters that define the model are discussed below and summarised Table 4.1.

Procrustes Analysis

Shape is defined as what remains once the effects of scale, translation and rotation have been removed (Stegmann and Gomez, 2002). To convert the training segmentations to "shape space", before the PDM is determined, the effects of pose must be removed. This can be done using generalised Procrustes analysis (Stegmann and Gomez, 2002). The effect of translation is removed by aligning the centroids of each segmentation to the mean. The adjustment to scale required can be determined using Equation 4.1, where $(\bar{x}, \bar{y})$ is the centroid of the shape and (x_j, y_j) is a point on the contour. Rotation is removed by performing singular value decomposition (SVD) to optimally rotate one segmentation to another.

$$\mathbf{S}(\mathbf{x}) = \sqrt{\sum_{j=1}^{n}[(x_j - \bar{x})^2 + (y_j - \bar{y})^2} \tag{4.1}$$

$$\boldsymbol{x} = \frac{1}{N}\sum_{i=1}^{N}\mathbf{x_i} \tag{4.2}$$

The mean shape used is calculated iteratively. Initially, the mean shape is estimated. This estimate can be arbitrary, but in generalised Procrustes analysis is selected to be what is called the Procrustes mean (Equation 4.2). Procrustes analysis - removal of scale, translation and rotation - is performed using this mean estimate and the results are used to recalculate the mean shape (using Equation 4.2). This is iterated until there is no longer any change in the recalculated mean.

Principal Component Analysis

The point distribution model (PDM) aims to capture the underlying model of the training data. Towards this goal, in ASM, principal component analysis (PCA) is performed on the information in the training set. PCA finds the primary modes of variation to more effectively represent the data (Shlens, 2005).

The dataset can be considered a matrix $\mathbf{X}$, size $m \times n$, consisting of n training examples of contours with m points. The covariance of two variables, A and B, can be measured as in Equation 4.3, where a_i and b_i can are samples of A and B. A high value would indicate high covariance. The covariance matrix, $\mathbf{C}$, in Equation 4.4 represents the relationship between the contour points of the segmentations. The ijth element of the covariance matrix is the covariance between the ith contour point and the jth contour point over all the training examples. The diagonal terms of $\mathbf{C}$ are the variance of a contour point and the off-diagonal terms the covariance between different contour points (redundancy).

$$\sigma_{AB}^2 = \langle a_i b_i \rangle_i \tag{4.3}$$

$$\mathbf{C_X} = \frac{1}{n-1}\mathbf{X}\mathbf{X}^\mathbf{T} \tag{4.4}$$

In an ideal representation there is no redundancy between contour points, resulting in a diagonal covariance matrix. PCA removes redundancy by redefining the contour points using the eigenvectors of the covariance matrix. It is assumed that the probability distribution of the data is Gaussian or exponential. The eigenvectors of the covariance matrix define a linear, orthonormal basis of the segmentation contours that when the dataset is transformed to, diagonalises the covariance matrix (Shlens, 2005). In addition to diagonalising the covariance matrix, use of these eigenvectors orders the dataset by explained variance: the first component/mode accounts for the greatest amount of variation in the data.

As in Equation 4.5, a training instance $\boldsymbol{x}$ can be represented as the mean ($\bar{\boldsymbol{x}}$) and a difference component that is expressed as the product of the model parameters, $\boldsymbol{b}$, and the principal components or model dimensions, $\boldsymbol{P}$. The mean shape $\bar{\boldsymbol{x}}$ is the starting shape normally used by ASM. Segmentation using ASM determines the values of $\boldsymbol{b}$ during its search of an image.

$$\boldsymbol{x} = \bar{\boldsymbol{x}} + \boldsymbol{P}\boldsymbol{b} \tag{4.5}$$

Point Distribution Model and Allowable Shape Domain

The principal components or primary modes of variation of the data correspond to the greatest amount of variation in the appearance of the object of interest seen in the dataset. These principal components are used as the dimensions of allowable shape variation. The number of dimensions used depends on the amount of variation in the training set to be explained and is a parameter of the ASM segmentation.

The amount of variation (f) included in the model in this project was chosen to be 0.99. Mouton indicated for his project that adjusting the parameter between 0.95 and 0.99 had no significant impact on the performance of the ASM segmentation. As the same database and software (albeit a different target structure) was used here, this range was used. The upper bound was taken as paediatric images show a high amount of variation and the structure of interest is complex.

The limits of variation in the model for each individual component found by PCA are set to suit the requirements of the application and are given in units of the standard deviation and

k	Length of profile in points
n_s	Number of positions to evaluate
N_{max}	Maximum number of iterations before convergence
f	Fraction of total variance included in the training model
m	Bound in units of standard deviation on the the shape parameters
k_u	Number of points on profiles above landmark
k_l	Number of points on profiles below landmark
n_{su}	Number of positions to search above the landmark
n_{sl}	Number of positions to search below the landmark
L	Number of multi-resolution levels used

Table 4.1: ASM parameters

bounds the allowable shape domain (ASD). The limit is placed on the value $\boldsymbol{b}$ (Equation 4.5). This limit, m, stops the model from taking a deformation beyond that which was typically seen in the training set.

ASM can also be extended into multiple resolutions. This is easily implemented by creating a training model for each resolution level. The number of resolution levels used is represented as L. Convergence is determined to be when the model has reached a local optimum and further iterations no longer change the placement, or when a set number of iterations N_{max} have been reached.

4.1.2 Pixel Classification and Post-processing

Pixel classification (PC) determines which pixels belong to the ribs and which do not. The pixel classification used by van Ginneken et al. (2006) was also used here. The classification makes use of the following features: pixel greyscale value, position and outputs of Gaussian filters at various levels. The output of the classification is the calculated probability that the pixel belongs to the selected class.

Each pixel in an image is described using its set of associated features. Training images are used to provide example feature values typical for each class of pixel. Every pixel in the new testing image is classified using a k-nearest-neighbours (k-NN) classifier. The k-NN classifier considers the k most similar pixels from the training images and classifies the new pixel as the average class of the neighbours. Two classes were considered here: rib and non-rib pixels.

PC provides a class for every pixel in the image but does not necessarily ensure that the results form cohesive segments. To return a connected object, van Ginneken et al. (2006) first post-processed the results of PC by applying a blurring filter. The largest connected object was selected and any holes in the object filled. Loog and van Ginneken (2006) post-processed the PC with a further iterated form of PC that took the context of each pixel into account and reclassified pixels relative to other pixel classifications. The results showed good identification

of the ribs. Sarkar and Chaudhuri (1998) performed morphological filtering and thresholding to identify the ribs in chest radiographs. The results were post-processed by labelling and fitting curves to rib-like objects.

The requirement of the rib suppression method used in this project is contours representing the upper and lower borders of the ribs. The post-processing methods mentioned by van Ginneken et al. (2006) and Loog and van Ginneken (2006) do not assist in returning such contours and only enforce structure. Curve fitting, as performed by Sarkar and Chaudhuri (1998), both enforces a rib-like structure and easily returns contour points rather than a binary image: this was the approach taken here.

Morphological Filters

Sarkar and Chaudhuri (1998) and van Ginneken et al. (2006) used morphological operators in their chest X-ray image segmentation algorithms to modify binary images. Morphological filters or operators change the shape of structures in the images to which they are applied. The filter used is called the structuring element and can be used to emphasise or remove any particular structure. Sarkar and Chaudhuri (1998) used this type of filtering to remove small non-rib structures (erosion) while van Ginneken et al. (2006) used it to fill holes in segmented objects (dilation). Erosion changes the value of each pixel to the minimum of its neighbouring pixels and dilation to the maximum of the neighbouring pixels (Deserno, 2011, pp. 108-111). Opening - erosion and then dilation - disconnects neighbouring objects. Closing - dilation followed by erosion - connects neighbouring objects. These methods are used here on the binary images obtained from PC.

4.2 Rib Suppression

In order to remove the ribs, it is necessary to estimate a model of the ribs for subtraction. The difficulty lies in creating a model from a single image without any additional information. Ahmed et al. (2007) used the rib segmentation as a second observation of the image, making the assumption that the overlap of lung tissue and ribs is a linear mixing. Hogeweg et al. (2010b) created a model of the ribs using the assumption that the ribs are the dominant structure in the rib area. The model created by Vogelsang et al. (1998) was not specified and is therefore not useful to this project. Sarkar and Chaudhuri (1998) used the difference peaks in the pixel histogram. Of all these methods, only Hogeweg et al. (2010b) quantified their results, which they did in relation to the results of a CAD system in the detection of TB. As their method, when combined with texture classification for CAD, improved on texture analysis in the absence of the rib suppression, it was used in this project.

Hogeweg et al. (2010b) extracted the ribs from the image using a non-linear transformation: the rib border was considered the axis of the rib and used to sample the curved rib to return a rectangular matrix. This data was then approximated using the first few modes returned by principal component analysis (PCA) and corrected for the effects of the underlying soft tissue. The final result was transformed back to image co-ordinates and subtracted from the image to perform rib suppression.

4.2.1 Principal Component Analysis

PCA was introduced in Section 4.1.1 and can also be used to approximate the original data with the less common forms of variation removed. The principal components/modes, as described in Section 4.1.1, are the eigenvectors of the covariance matrix, returned in order of the amount variation in the dataset they explain. By reducing the number of modes used, the less common forms of variation in the dataset are removed and the main form of variation is preserved.

The use of PCA here to model the ribs was based on the assumption that the main form of variation in greyscale value in the rib area of the image is caused by the ribs. A visual inspection of a chest X-ray image will show that in most images, the rib is more dominant than lung tissue. Rib suppression proposes to separate the rib and lung tissue.

4.3 Texture Classification

Mouton's (2009) classifier, developed for paediatric chest radiographs, is used here as it is this CAD system that the rib suppression aims to improve. Details of the methods used are given in Section 2.3 and details of their implementation in Chapter 7. The performance of the system, shown by Mouton et al. (2010), is comparable to that of adult CAD systems developed by others and is suitable for the investigation of rib suppression in paediatric images.

Chapter 5

Implementation: Finding the Rib Borders

This chapter discusses the implementation of rib segmentation. Two different methods were used - active shape models (ASMs) and pixel classification (PC). Both methods were supervised and needed training examples of rib segmentation for the methods. These were supplied in the form of manual delineations of the ribs. These manual segmentations also provided a reference for measuring the results of segmentation.

5.1 Rib Contours

The visual rib in chest radiographs can be divided into three parts, starting from the spine the: posterior, lateral and anterior rib. The posterior portion of the rib in chest radiographs is seen as the almost-horizontal bars starting at the spine and crossing the lung area, to the sides of the chest. The lateral part of the ribs are the portion that run into the page and are often bright oval regions lateral to the lung areas. The anterior ribs run down diagonally from the lateral portions to meet with the sternum or costal cartilage - only the beginning of the anterior rib is usually visible in radiographs. These structures are shown in Figure 5.1.

The rib segmentation in this project identified the posterior ribs only. These are the most visible part of the rib in chest radiographs. Only the portions visible over the lung are considered - the rib suppression is being used to improve analysis of the lungs in radiographs. The visibility of the ribs outside of the lungs was variable and was often not segmentable at all, as such, ribs outside of the lungs were completely disregarded. The number of ribs visible within the lung area was variable. The superior and inferior boundaries of the posterior ribs will be referred to, from here on, as the upper border and lower border respectively. The upper and lower borders are closed by the lung boundaries.

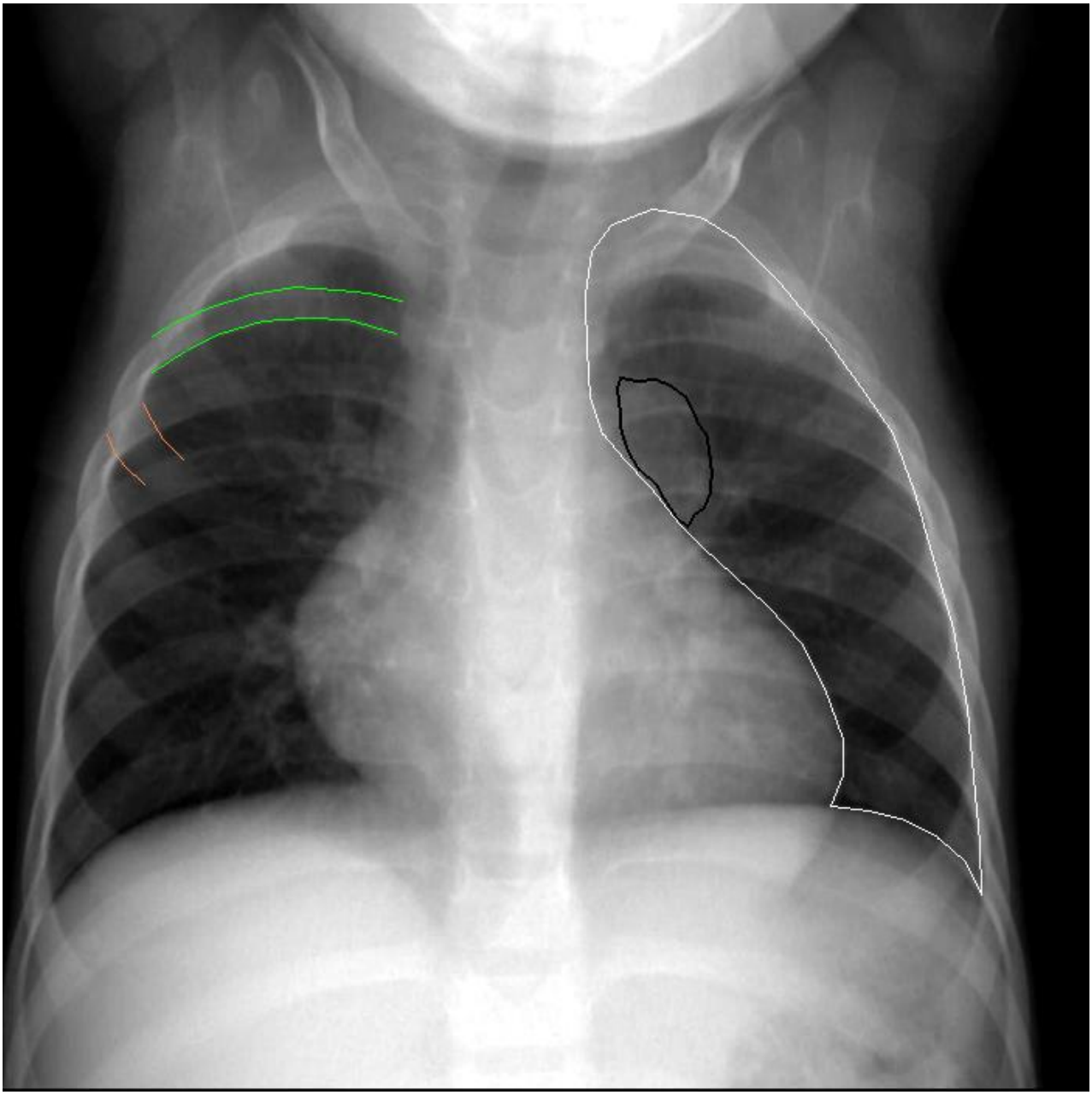


Figure 5.1: Structures of interest in a chest X-ray. The green outline shows the posterior 3rd rib on the right side of the patient. The part of the rib outlined in orange is the anterior portion of the 3rd rib. The lateral portion can be considered the part of the rib between these two sections. This is a typical representation of the lateral rib. The left lung is outlined in white. The first rib can be seen crossing the clavicles just above the visible lung while ribs 2-8 visibly overlap the lungs. Rib 8 is barely visible. The rest of the ribs lie below the diaphragm. This image also shows an abnormality in the left lung outlined in black.

The quantity of ribs visible in each image is dependent on the moment in the breathing cycle at which the radiograph was acquired. As the patients were all children and not always able to be instructed, the point in the breathing cycle, and thus the number of ribs over the lungs, accounted for the largest amount of variation in the images. The ASM method used here took only a set number of landmarks and therefore only a set number of ribs. The majority of the images in the dataset showed at least five ribs overlapping the lungs, so the training model was taken to have five ribs. Pixel classification does not have this limitation and can be trained with a variable amount of ribs.

The word "boundary" here must be taken in the context of medical imaging. Chest X-ray imaging is projective and all structures in a line from the radiation source are superimposed. This makes a visible border the point at which one tissue/organ becomes more visibly dominant than its neighbour, rather than a line defining the start of one structure and the end of another. In addition, borders can only be defined to the accuracy of the scanner. For example at 2.08 lp/mm, a pixel in the X-ray image represents 0.5 mm on the patient, so a border in the image is only accurate to 0.5 of a mm. The border here was considered the edge beyond which the structure of interest was indistinguishable from the surrounding tissue.

Manual segmentation of the paediatric images was performed by the author. Two sets of segmentation were provided - firstly a manual segmentation of the first five ribs visible fully overlapping the lung, from the lateral lung border to the medial lung border for both lungs. Secondly, segmentation of each rib that is visible fully overlapping the lung, from medial to lateral lung boundary. The segmentation of five ribs was used to train the segmentation methods while the segmentation of all the ribs (overlapping the lungs) was used for assessment of the segmentation methods (gold standard). These two segmentations are shown in Figure 5.2.

Both segmentation methods were trained with the same training segmentations to allow comparison of the methods on the same task - identifying the first five ribs over the lungs in a chest radiograph. It is common for the segmentation methods to identify other ribs besides the trained five ribs - ASMs can slip down the ribcage and PC will identify anything similar to the trained ribs. The results of these segmentations are not the trained ribs but are still correct rib segmentations. The inclusion of all the ribs overlapping the lungs in the gold standard allowed for the segmentation of ribs other than the first five overlapping the lungs.

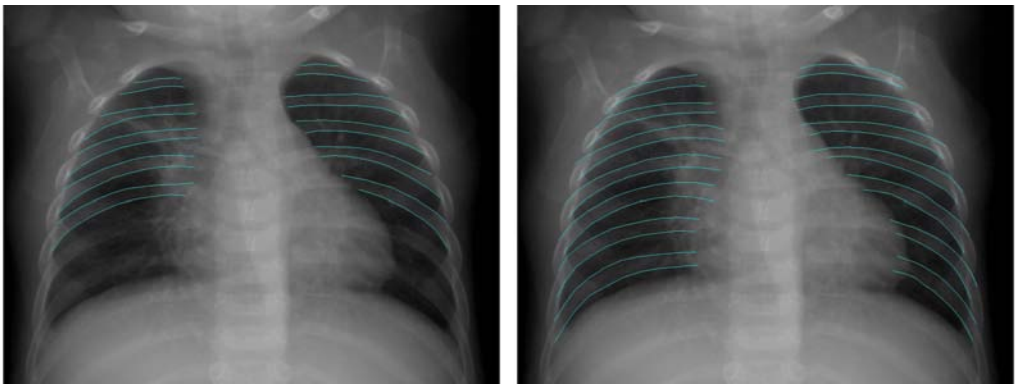


Figure 5.2: Two manual segmentations of the same radiograph. The left image shows a segmentation of the first five ribs. The right image shows the segmentation of all ribs overlapping the lungs.

5.2 Active Shape Models

The segmentation contours used by ASM are a set of points with some implied connectivity. These points can be either landmarks or boundary points - landmarks are typically corners or intersections of boundaries, while boundary points simply mark a boundary or sharp change in gradient. In segmentation, these landmarks and boundary points typically surround the area of interest.

Each rib in the manual segmentations was represented by an upper and lower border of ten points, making a total of 20 points per rib. In the literature, the ribs are often modelled as parabolas which theoretically can be defined by 3 points. Given that ribs are not perfectly described by parabolas, ten points was considered sufficient to allow for significant variation but low enough to allow for speedy ASM searching.

5.2.1 Storage and Implementation

The ASM implementation used in this project was the multi-resolution tool initially created by Hamarneh et al. (1998). The implementation was in MATLAB and was used and modified by Mouton to perform lung segmentation for his CAD tool. Hamarneh's method performs Procrustes analysis before creating the training model - scale, rotation and position were not incorporated into the training model. The tool's default starting position is the mean of all the example segmentations.

The ASM method assumes all the images are the same size. The profile length was given in pixels rather than relative distance. This makes the method the size of the profiles dependent on the resolution of the images. On the original images, 3 pixels might span the whole intercostal distance but only a fraction on a higher resolution image. All images were re-sized to a resolution

```
#######################################################
# Number of points:
140
# format: (path#)(type)(x rel.)(y rel.)(point#)(connects from)(connects to)
1     0     0.25486     0.1733      0     0     1
1     0     0.26875     0.16193     1     0     2
1     0     0.28333     0.15341     2     1     3
1     0     0.29861     0.14678     3     2     4
```

Figure 5.3: Header of an ASF File

of 700 by 700 pixels - by re-sizing the images any unpredictable effects of varying image sizes are removed. The number of profile points above and below each landmark can be specified separately (k_u and k_l respectively, see Table 4.1).

A few modifications were made to the ASM tool to make the processing of large datasets easier. The training stage was changed to allow the loading of stored training segmentations from file - previously the user was required to perform the training segmentations at run time. The application stage was also changed to perform ASM on a folder of images rather than a single image and to save all results to file. The file format used was active shape format.

Active shape format (ASF) (Stegmann, 2002) is an ASCII file format that can be used to store contours used for active shape/appearance models (ASMs/AAMs). The format stores the number of model points, the points themselves and certain attributes of the points - the path or contour, the point type, the previous contour point and the next contour point shown in Figure 5.3. This format was chosen as Hamarneh's code already provided for it and the file format was intended to be used with ASM. The point coordinates are stored as relative co-ordinates - the co-ordinates scaled to image size. This was practical as the same stored segmentation can be used for all scaled versions of the original image, independent of aspect ratio.

The number of positions to search, n_s, must also be selected. Hamarneh's implementation allows for the number of positions above n_{su} and below the contour n_{sl} to be specified separately. Selecting a large number of positions to search will search a larger area of the images but will take longer to perform. Setting n_s low will lead to a faster convergence but will only search in a small area around the starting position.

Points or landmarks in ASM do not need to be connected to form contours, closed or open. Practically, in implementation, the connectivity of points was used to calculate the gradient of the profiles. In calculating the gradient of the profile, the two points adjacent to the current point are used to determine the gradient at the point of interest. For open contours, the gradient at the last point of the contour is calculated using that point and the single adjacent point. The ribs were represented here as open contours.

5.2.2 Training Images and Search Parameters

The number of training images necessary is dependent on the nature of the images and the structure being segmented. The greater the variation seen in the data, the more training seg-

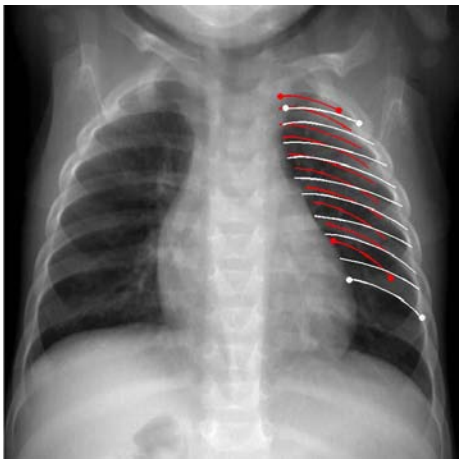


Figure 5.4: The result of initial placement. The red and white contours show the mean placement and manual placement respectively. The dots show the landmarks used to determine the warping from the mean placement (red) to the manual placement (white).

mentations are necessary. Five-fold cross-validation was chosen for selecting training and testing images. Cross-validation allows for a large ratio of images to be used for training without diminishing the amount of testing data. The dataset was divided into five folds with 31 images in each fold except for the last which had 30 images. Effort was made to stratify the datasets, maintaining similar ratios of each image type in each set to avoid any training set becoming biased to a specific type of data. Cross-validation uses one fold for testing and the remainder for training.

It was found that rib segmentation using ASM was very sensitive to initial placement. When the initial placement is far from the target structure in distance, scale or rotation the deformable model will converge on an improbable shape or on the anterior or lateral ribs. The mean model position was not sufficient to place the model close enough to prevent incorrect convergence. A semi-automated approach had to be taken to provide a realistic starting position.

To provide a starting position, the user was required to select four points over each lung, as shown in Figure 5.4. These points were the intersections of the upper border of the top rib with the lung border and the intersections of the lower border of the bottom rib with the lung border - the four corners of the model. The terminology top and bottom here refer to the top and bottom of the first five most visible ribs overlapping the lungs. These four points are then used to define an affine warp to adjust the mean shape model to the registration points. The fitting method was prevented from making further scale, translation and rotation (pose) adjustments from this point forward. The starting position fairly accurately places the rotation and, given the similarity of the ribs in correct and 180° rotated position, it was acceptable to remove further pose corrections to prevent the ASM latching onto the anterior ribs or rotating 180°.

Given the sensitivity of the model, the deviation from the mean shape was limited to one standard deviation, anything larger than this would cause at least one image in a set to converge

on an improbable shape. N_{max} defines the maximum number of iterations the ASM method will spend searching the image should the search not converge before that. As in Mouton's project, a maximum of 40 iterations was found to be adequate for the ASM implementation.

Inclusion of rotation, translation and scale into the training model, with regard to ASM, is a possibility for future testing. It is not considered here as paediatric images show a large amount variation in appearance. This is better modelled from a small dataset without considering the additional variation caused by pose.

5.3 Pixel Classification

As an alternative to segmentation with ASM, pixel classification (PC) was considered. PC approaches segmentation on a local basis and places each pixel in a class depending on its greyscale value and other such features. The pixel classification method was that described by van Ginneken et al. (2006). The features used were pixel spatial information, greyscale value and Gaussian filter outputs. Training was done on a leave-one-out basis and the rib areas in each image were highlighted using a binary mask of the ribs created using the segmentation contours used to train the ASM segmentation. The images were all resized to the same resolution. Training was performed with the manual segmentation of the first five ribs overlapping the lungs. The output of PC was an image where the value of each pixel represented a confidence value of that particular pixel belonging to rib area, shown in Figure 5.5.

The role of PC taken here was similar to that of Sarkar and Chaudhuri (1998) where the pre-processing step of image filtering is replaced by PC. There were two stages applied after PC: post-processing and model fitting. Morphological fitting and labelling was performed in post-processing. The model fitting was performed, as done by Sarkar and Chaudhuri (1998), by curve-fitting and dilating the result to the rib width. The result was the return of contours representing the upper and lower borders of the ribs.

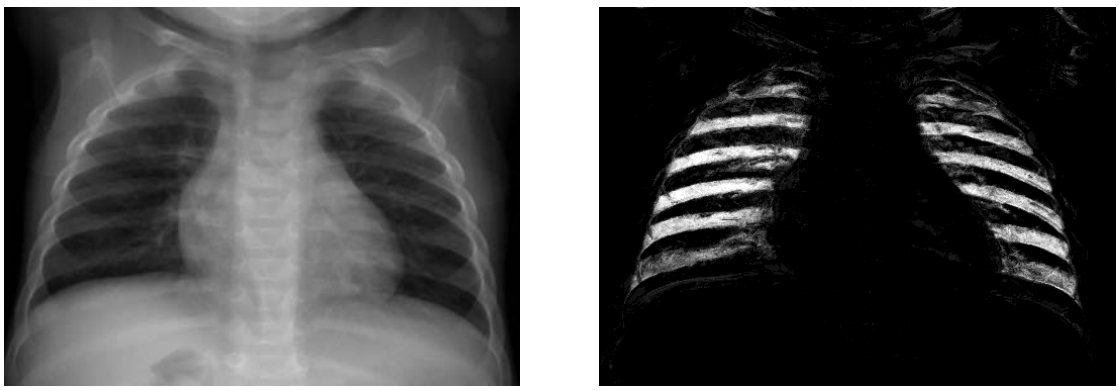
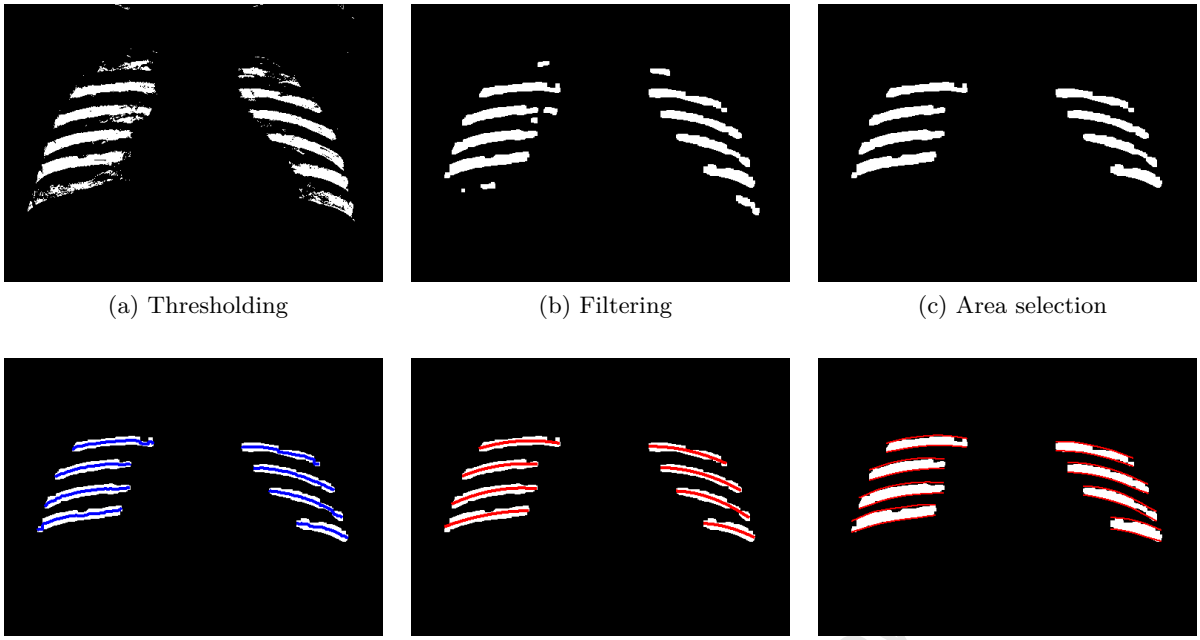


Figure 5.5: The figure on the left shows an example paediatric image and the image on the right is the result of pixel classification on this image. The lighter shades correspond to a higher confidence that that area describes the rib region.



(a) Thresholding (b) Filtering (c) Area selection

(d) Midline (blue) (e) Fitted parabola (red) (f) Upper and lower (red) estimation

Figure 5.6: Images highlighting the stages in PC post-processing (a-c) and curve fitting (d-f).

5.3.1 Post-Processing

The PC results, confidence values in the form of an image-sized matrix, were conservatively thresholded at a 1/3 of the maximum confidence level present in the image, shown in Figure 5.6a. This converted the continuous, range-bound results of PC into a binary segmentation. A opening-closing filter was applied to the resulting binary image to remove smaller islands while connecting larger clusters, shown in Figure 5.6b. All connected objects with an area greater than 0.25 of the area of the largest object, presumably a rib, were considered to be ribs and all other objects were removed from the binary image. By selecting objects greater than a specific area, rather than the N largest objects, a variable number of ribs could be found. The remaining objects were labelled, shown in Figure 5.6c. To summarise, the following steps were taken:

- The image was thresholded
- Morphological filtering was performed
- Objects in the image were filtered by area

5.3.2 Curve Fitting

The remaining connected objects in the image were labelled and assumed to be ribs. Labelling identifies connected objects in binary images. For each labelled structure, the midline, shown in Figure 5.6d, was found and a parabola was fitted to it using least-squares, shown in Figure

5.6e. Least-squares is a common estimation technique for fitting well represented data (Hastie et al., 2009, p. 11). The rib width was determined as the average distance between the upper and lower border of the labelled object. The upper and lower border of the ribs were created by adjusting the parabola up and down by the estimated rib width, as shown in Figure 5.6f. To summarise, the following steps were taken.

- Each remaining object was fitted with a parabola
- The rib width was estimated
- The lower and upper rib borders were estimated

The parabolas returned by the curve fitting were resampled to ten points per contour. This was the format of segmentation preferred for rib suppression. These contours were saved in ASF format as with the ASM segmentation, described in Section 5.2.1.

5.4 Quantitative Analysis

In order to measure the performance of rib segmentation, overlap was used. Overlap is calculated as the ratio of the intersection of the segmentation area and the ground truth segmentation area, to the ground truth area (Equation 5.1). This measure is sensitive to the size of the structure being segmented (Gerig et al., 2001).

$$Overlap = \frac{S \cap R}{R} \tag{5.1}$$

As there are multiple objects in this segmentation, a few specifications were given with the overlap measurement. If a segmentation is far from the "true" segmentation, it is very likely that one of the "ribs" provided by the segmentation would overlap with more than one of the reference ribs. For this reason, overlap was limited to one reference rib for each segmented rib. Also, given that ASM is designed to deal with a set number of contours and there is a variable number of ribs in each image, the number of ribs segmented would be lower than the number of ribs identified in the ground truth segmentation. For this reason, overlap was calculated only for the segmented ribs that overlapped with the ground truth segmentation.

Chapter 6

Implementation: Rib Suppression

Chest X-ray images are more difficult to interpret than chest computed tomography (CT) images because of the overlap of structures. In the "lung" area of the standard chest X-ray, in addition to the lungs, one sees the ribs, the spine, the mediastinum and the clavicles. Rib suppression here was performed in order to improve the visibility of the lung tissue over that of the background anatomy. As only the lung tissue was of interest, the rib suppression was performed only in regions where the ribs overlap the lungs.

The methods can be described in three steps (shown in Figure 6.1): rib acquisition, modelling and subtraction. The rib data was extracted from the image during rib acquisition, in which the rib segmentation was utilised. This data was then processed to create a model of the rib in the model derivation stage. In model subtraction, the model was fitted back to the original image coordinates and subtracted from the image. This was done for each rib in each image and no data was carried over from previous images or ribs. The result of this process, once it has been performed for each rib overlapping the lungs, is the rib suppressed image.

This project tests two rib suppression techniques. The primary method was based on the process described by Hogeweg et al. (2010b), which showed promising results in their study involving a texture classifier used for the detection of TB. The second method for rib suppression differs in the modelling stage - the first method made use of principal components while the second uses the mean value of the rib area. A comparison of these approaches was expected to show any advantages of using principal components. A third modelling method is used here but should not be considered rib suppression: noise modelling. The rib suppression methods are shown together with segmentation in Figure 4.1.

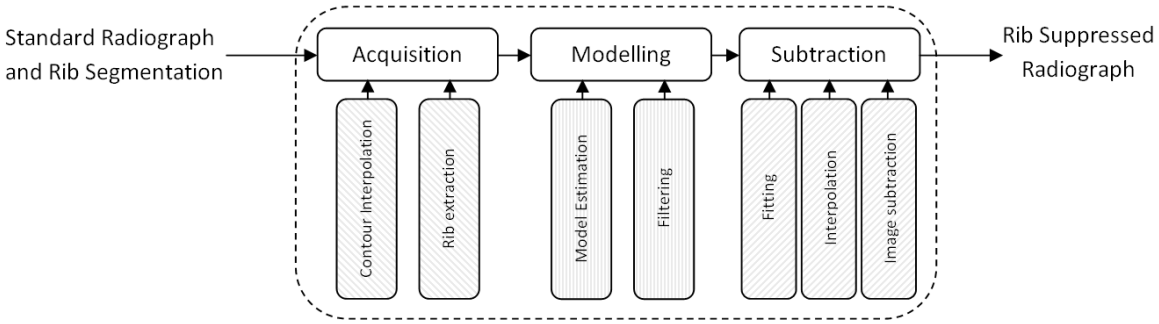


Figure 6.1: Flowchart of suppression methods

6.1 Rib Acquisition

Rib acquisition was performed using the rib segmentation of the image (described in Chapter 5). The segmentation was given as sets of contours of ten points. For a closer approximation of the rib, interpolation was used to obtain the maximum number of points - limited practically by the resolution of the image. The rib is sampled along lines running perpendicular to the rib border i.e. along cross-sections of the rib. In this project, an extracted rib is referred to as a stack.

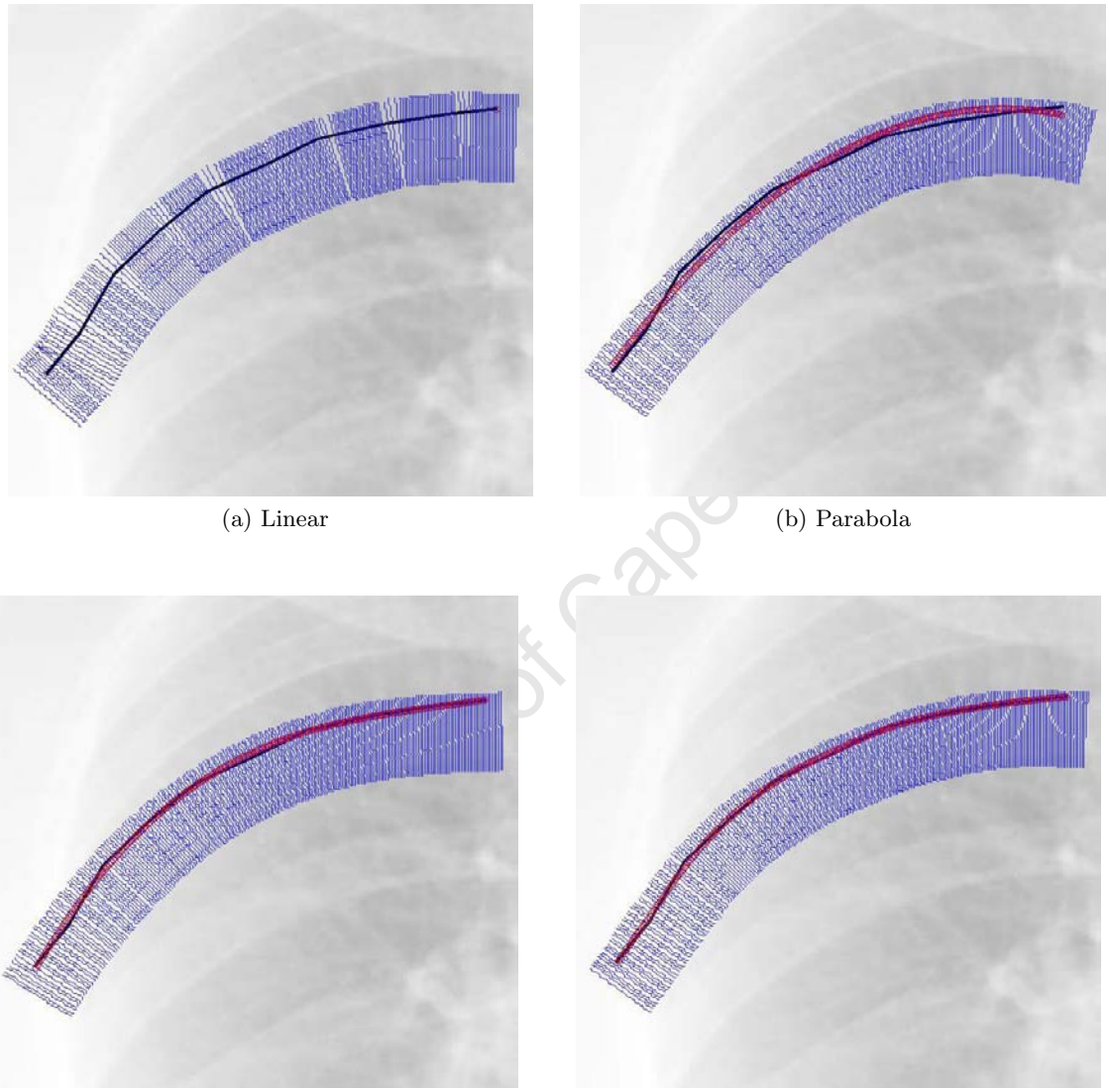
6.1.1 Rib Contour Interpolation

The interpolation of the rib contour had a significant impact on the rib model. Poor interpolation leads to warping of the extracted rib. This adds a form of artificial variation to the stack which may lead to degradation of the PCA model.

An interpolation using a fitted parabola (Figure 6.2b) performed well for cases in the paediatric dataset but failed on all of the adult images. Ribs in the adult database were steeper in orientation. Increasing the degree of the model (Figure 6.2c) corrected sufficiently for the majority of the paediatric case but not for the adult case. Cubic spline interpolation (Figure 6.2d) performed comparably on both the adult dataset and the paediatric dataset and was used in this project.

6.1.2 Rib Extraction

The rib area was extracted so as to take the variable shape of the rib and return a rectangular matrix. The dimensions of this matrix represented distance along the rib and distance across the rib. This was performed by sampling the rib area along lines running perpendicular to the boundary of the rib, which could be regarded as cross-sections, referred to here as profiles. These cross-sections or profiles were taken along the entire length of the rib segmentation, shown in Figure 6.2.



(a) Linear

(b) Parabola

(c) Polynomial

(d) Cubic spline

Figure 6.2: Profile sampling using various interpolation models on the rib contours, displayed on the chest X-ray. The interpolation is shown on a rib over the right lung. The blue lines across the ribs show the profiles/cross-sections used to extract the rib. The black line along the rib shows the result of linear interpolation (shown on its own in a). The red line (b-d) shows the interpolated rib contour.

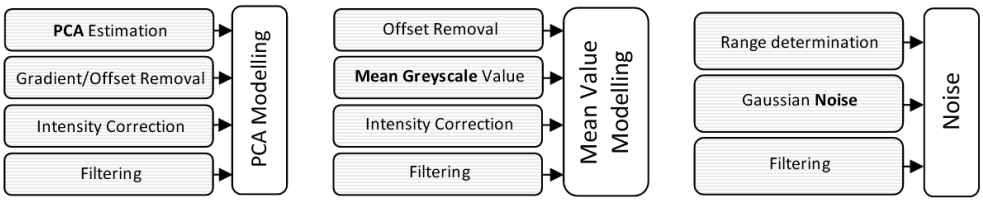


Figure 6.3: Detailed steps of the three methods used in the modelling stage of rib suppression.

The direction of each profile was perpendicular to the gradient of the rib contour at the point through which the profile passes. The gradient at that point was calculated by the difference in position of the adjacent points on the contour. The length of the profile was determined using the average rib width.

The rib width was calculated as the average of the minimum distance from each point on the contour to the opposite contour. The length of the profiles was set to a percentage more than this in order to include intercostal area in the sampling. The exact value was 50% more than the original width - the profiles extend beyond the upper and lower rib boundary equally by 25%. A single value was used for the width of the rib as principal component analysis requires data instances of a set size.

The profiles can be projected from the lower rib boundary up, the upper rib boundary down or outwards from a centre line of the rib. The approach taken will affect the length of the profiles as the calculation of rib width is not symmetrical. A projection from the lower rib boundary leads to a better suppression of the lower border while a projection from the upper boundary leads to better suppression of the upper border. These two methods are tested here.

6.2 Model Derivation

The extracted rib area was used to derive a model of the rib. Three methods were tested at this stage of rib suppression, modelling using: PCA, mean value and noise. PCA was the method used by Hogeweg et al. (2010b), mean greyscale was provided for comparison and noise modelling was included to determine whether the effect of rib suppression on texture analysis is random. The differences in the methods is highlighted in Figure 6.3.

6.2.1 Modelling using PCA

Principal component analysis (PCA) was performed on the extracted rib (shown in Figure 6.5a) in order to estimate a rib model. Each profile was taken as an instance of the data and three modes of variation were used to estimate each profile (shown in Figure 6.5b). Hogeweg et al. (2010b) used two modes of variation to estimate the ribs - no significant difference was found

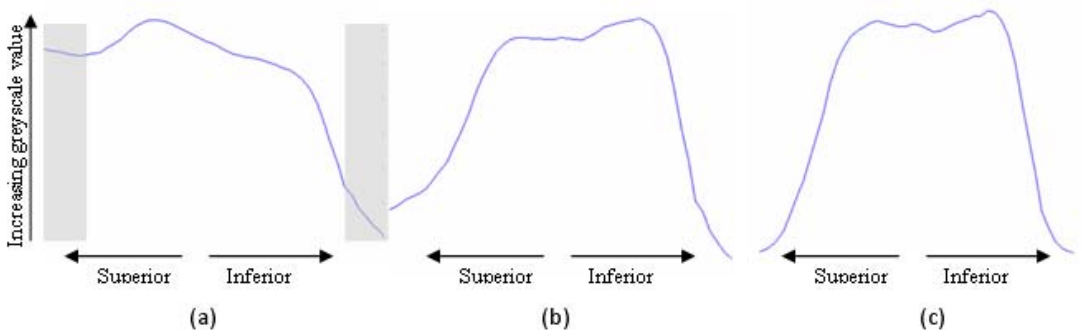


Figure 6.4: Example cross-sections of the rib at different scales. The signals show the average pixel value against the distance along the rib cross-section - intensity versus position along the cross-section. The left plot (a) is the average original cross-section, before modelling. The middle signal (b) shows the result of PCA modelling. The right image (c) shows the average final model after post-processing. In (a) the intercostal area is shown in grey. The scaling and offset of the signals are not shown.

here between two and three modes of variation. Three modes provided a smoothed version of the profile signal. The basis of this process is the assumption that the rib is the dominant visible structure in the segmented area, recognisable by PCA.

It was important that the data from the ribs in the extracted stack be separated from data from other anatomy, in particular the lung tissue. As mentioned in Section 6.1.2, the length of the profiles was taken to include the intercostal area. The rib data can be seen as the signal present in the rib area but not in the intercostal area. In Figure 6.4a. there is a difference in greyscale value between the leading and trailing edge of the signal. This difference or gradient starts in the intercostal area outside of the rib borders and cannot be attributed to the ribs, thus must be removed. The gradient in each profile is brought about by the soft tissue of the lung: possibly the anterior-posterior change in depth of the lung. To remove this, a line was fitted to the data and subtracted without the offset from the origin to straighten the profile.

The intercostal greyscale value was lower than that of the rib area but not zero, providing an offset. The offset can be determined as the distance from the x-axis to the lowest values of the cross-section. This was implemented by subtracting the average value of the lowest 10% of the profile as done by Hogeweg et al. (2010b). The profiles were corrected for any negative values. An example of a gradient and offset corrected model can be seen in Figure 6.5c.

Changes in the intensity of the rib in X-ray images is caused by the change in direction of the rib relative to the imaging device. Other changes in model brightness could be seen in the results of the model derivation at this stage that did not correspond to the direction of the rib. This was primarily caused by structures running over the rib in that area. To remove their effect on the model brightness, Hogeweg et al. (2010b) applied Gaussian smoothing to the model. This smoothed the model resulting in a more even appearance of model brightness.

As the correspondence between profiles depends on the accuracy of the rib segmentation

and interpolation, Gaussian smoothing was not used here to avoid influencing profiles with unmatched data from adjacent profiles. This was mostly avoided by Hogeweg et al. (2010b) through the use of manual segmentations. Instead, in this project, the profile intensities were scaled to fit a linear model (shown in Figure 6.5d). This was based on the assumption that the underlying density of the rib is fairly uniform.

The correction of offset did not zero the leading and trailing edges of each profile. Therefore, tapering the ends of each profile was necessary to reduce artefacts created when the model was subtracted from the image. This was performed by applying a Tukey window (or tapered cosine) (Hearn and Metcalfe, 1995, p. 122) where the ratio of taper to window was equal to the ratio of the intercostal area to rib area - 50/150 or 2/3. Similarly the lateral and medial edges of the model were tapered (to a ratio of 1/10). The Tukey window was used as it has a value of 1 for the body of the window, unlike most windowing functions, which allows the rib model to remain unaffected away from the edges.

While the PCA is not equipped to analyse profiles of varying sizes, the model can be post-processed using the known rib lengths. To compensate for varying rib width, the Tukey window applied to each profile was set to the width of the rib at that point. This was calculated as the distance from the starting contour, as in rib acquisition, to the closest point on the opposite contour. Signal outside of the Tukey window was set to zero. The results of the windowing can be seen in Figure 6.5e.

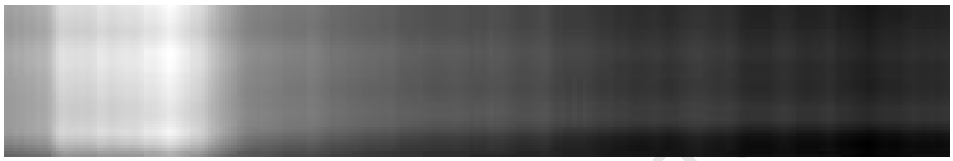
6.2.2 Modelling using the Mean

The most obvious difference between the rib and intercostal area of a lung is the difference in average greyscale value. To provide a simple method of rib suppression for comparison, a model to remove this difference was tested. To perform rib suppression using the mean greyscale value, the steps mentioned above used were: interpolating the rib contours, extracting the rib area using perpendicular profiles, gradient offset removal, intensity correction and Tukey windowing. The difference lay in the PCA modelling. Rather than using PCA and taking the first three modes of variation, the greyscale mean was taken of the stack.

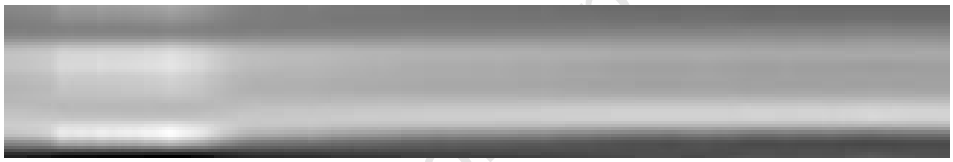
To derive the model, the average greyscale value of each profile was taken as the estimated rib value for that profile. The resulting model was intensity corrected as the mean value of each profile is also affected by background structures. The model was filtered to reduce subtraction artefacts as mentioned in the previous section.

The second significant difference between the greyscale method and the PCA method was in the length of the profiles. PCA expects data of equal dimensions and this was done by ensuring the profiles are all of equal length. With the greyscale method, this was not necessary. For this reason, variable rib width was used - each profile was the width of the rib at that point. The

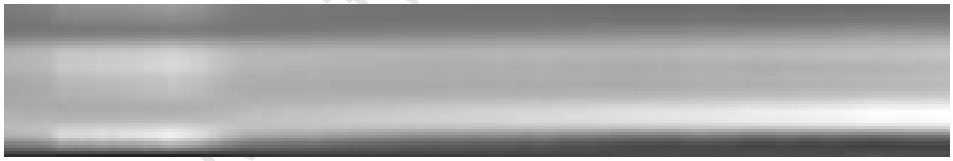
(a) Extracted rib area (maximum greyscale value of 154)



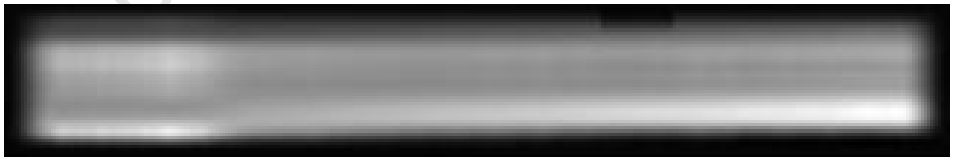
(b) First three modes of the stack (maximum greyscale value of 140)



(c) Gradient and offset corrected model (maximum greyscale value of 20)



(d) Profile intensity corrected model (maximum greyscale value of 19)



(e) Tukey windowed model (maximum greyscale value of 18)

Figure 6.5: The extracted rib and rib model at various stages of the method with the maximum greyscale value of the stack at that stage.

rib width was estimated as the distance from each point on the starting rib contour, as used in rib acquisition, to the closest point on the opposite contour. The rib width was limited to one standard deviation to avoid long profiles on the ends of the rib caused by mismatch in contour lengths.

Tukey windowing was performed individually on each profile due to the variation in length. An acceptable Tukey window ratio, signal to taper, was determined empirically as 0.6 with an additional leading zero signal portion of 10%. This best matched the model to the visible rib boundaries. A higher ratio lead to model subtraction in the intercostal area.

6.2.3 Modelling using Noise

A third rib model was included for comparison. In order to account for the possibility that the effect of the rib suppression using the previous rib modelling techniques increases the performance of the texture classifier through chance rather than any increased visibility, a modelling technique was used that intentionally did not attempt to model the rib. To this purpose, Gaussian noise was used. The noise was limited to the range of the PCA modelled rib to allow the "suppressed" rib area to still be comparable to the other modelling techniques. The model was also tapered with a Tukey window, as with the previous methods, to avoid subtraction artefacts. The expectation of the effects of this modelling technique was that it would decrease the performance of a texture classifier. No effect on the results using this method would indicate that the effect of rib suppression is possibly random.

6.2.4 Lung Abnormalities

It is important to consider how these methods will be affected by abnormalities in the lung area. There are three situations to consider: abnormalities larger than the width of the rib, those that are smaller than the rib width and those that span the length of the rib. The worst possible outcome would be decreased visibility of the abnormality in the image.

Abnormalities in the lung tissue wider than the width of the rib will affect both the intercostal area and the rib area. This would result in profiles whose greyscale value is increased in the intercostal and rib area. Any offset or gradient presented by the abnormality would be removed in the model processing. Any variation in shape of the signal is unlikely to be recognised as a significant form of variation along the entire rib length and will be excluded from the rib model as a minor form of variation. Therefore abnormalities wider than the rib would not be removed from the image and would remain as visible as the lung tissue.

Abnormalities smaller than the width of the rib would cause the rib profile to deviate from the standard profile shape and is unlikely to be recognised as a rib primary mode of variation.

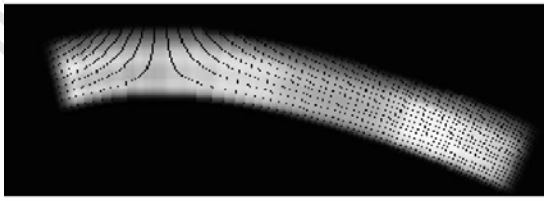
Consequently they will not be included in the rib model and will not be subtracted from the image. It is possible that the variation would cause an increase in scale of the rib model - in the profile intensity. This would be removed in profile intensity correction. Should this correction fail, the abnormality would be unnaturally darkened or brightened and its visibility would increase.

Should the abnormality span the width of the rib it would be visible regardless of rib suppression. Rib suppression would attempt to remove it, but unless it is the size, shape, in the position and of the consistency of a rib - a slow variation in signal - it would not be completely removed. Abnormalities with a shape of a rib could present a problem.

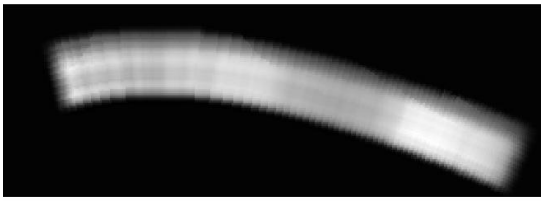
Abnormalities smaller or wider than a rib would not be affected. Abnormalities the shape of a rib would theoretically cause difficulties but are unlikely. In conclusion, theoretically the visibility of abnormalities will reliably not be decreased through the use of rib suppression.

6.3 Model Subtraction

To perform the rib subtraction, the model must be subtracted from the original image. The derived model is the size of the extracted rib. Therefore, at this stage, there is only model data along the profiles used in the original rib acquisition. The profiles, not being parallel, have gaps between them. The rib itself is not limited to the profiles and is a continuous structure in the image. To overcome this, the values between the profiles were interpolated. Figure 6.6 shows the model before and after interpolation.



(a) Fitted rib mask before interpolation

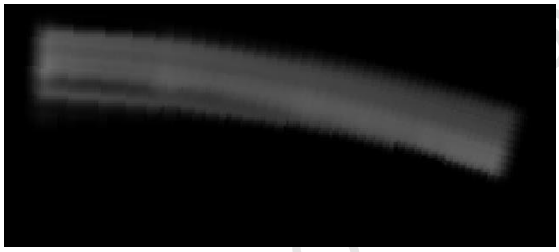


(b) Fitted rib mask after interpolation

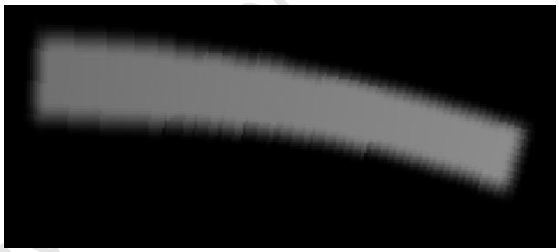
Figure 6.6: Images of a rib model fitted back to the image co-ordinates before (a) and after (b) the gaps are interpolated.

Standard methods of interpolation assume a linear transformation from one set of co-ordinates to another. The process of fitting the rib model back to the original profiles is not a linear system. Systems such as bi-cubic interpolation are not simple to apply to this situation. Interpolation as the average of k-nearest-neighbours overcomes this but presents a new problem of its own - often the neighbours are unknown values as well. The interpolation used was the average of the non-zero neighbours within a set distance. Once interpolation has been performed, the model is subtracted from greyscale values of the image. Figure 6.7 shows the rib model after interpolation has been performed for each of the three rib modelling techniques.

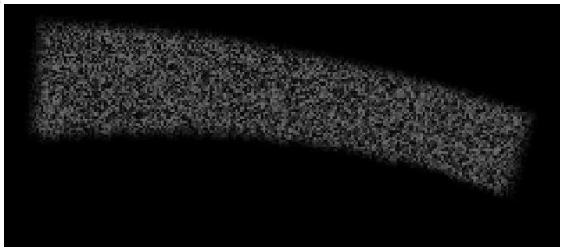
In each image, the steps described above were performed for each segmented rib. The resulting image was the rib suppressed image. The rib suppression steps were all performed at the original resolution of the images. This was to provide as much information as possible for PCA in order to build a more accurate model.



(a) PCA rib model



(b) Greyscale mean rib model



(c) Noise rib model

Figure 6.7: A single rib modelled with the three rib modelling techniques, ready for subtraction from the image.

Chapter 7

Implementation: Texture Analysis

The rib suppression developed in this project was an image pre-processing task for texture analysis. As an extension of the work done by Mouton (2009), his texture-based CAD tool was used to test the results. Mouton's texture analyser was trained and tested on standard chest X-ray images. Here rib suppression was performed on both training and testing images to determine whether the additional image processing improved the performance (both training and testing as similar images must be compared). An outline of the method used by Mouton was given in Section 2.2.5. This chapter describes the implementation of the texture-based classifier and the interaction of the classifier with the rib suppression.

Each of systems involved, the texture-based classifier and the rib suppression (rib segmentation included), operated on the chest radiograph. The bony structure suppression tool performed its function on the chest radiograph before the texture-based classifier was applied. Performance was evaluated with reference to the original texture-based classification analysis, where the rib suppression was not performed. This is illustrated in Figure 7.1.

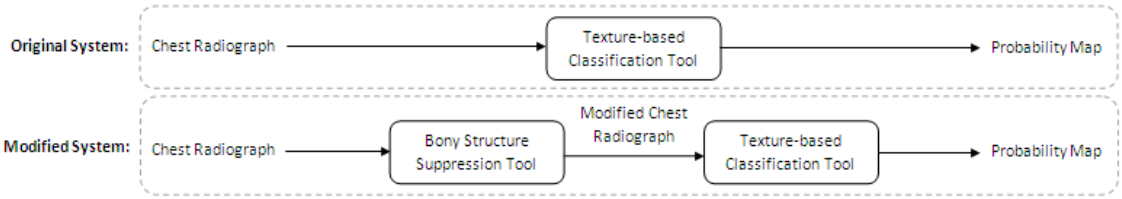


Figure 7.1: Texture classification with and without bony structure suppression.

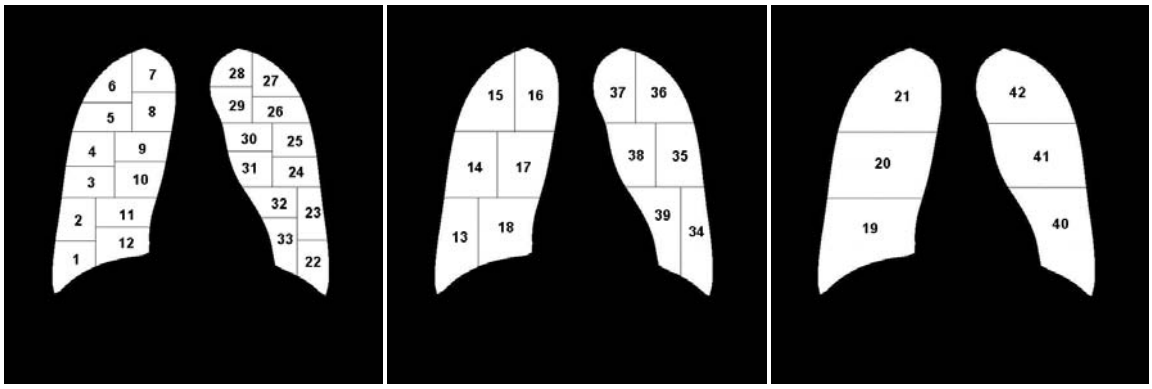


Figure 7.2: The lung was divided up into 24 smaller regions. These regions are considered individually and grouped into larger regions. The 24 regions shown in the left image are combined into the regions in the middle image and further merged into the regions in the right image.

7.1 Texture Analysis Parameters and Output

Mouton's texture classifier divided the lungs into smaller regions of equal area. Of the two types of regions that Mouton (2009) used, circular and big regions, only the big region approach was used here - circular regions were regarded by Mouton as "extremely" computationally expensive and would limit the amount of method variations that can be tested. Typical "big regions" are shown in Figure 7.2 with the numbering used in the implementation. The lungs were divided into 24 regions of equal area. These were combined with each other twice to create another 18 larger regions for a total of 42 regions.

Once the features for each region were derived, a k-NN classifier was used to reach a decision using the distance of the instance to the training examples. A training example was considered abnormal if the region overlapped with a radiologist-outlined area of abnormality - this was used to determine the final performance. All training and testing using Mouton's tool here on the paediatric dataset was performed on a leave-one-out basis. Leave-one-out testing removes one image from the set for testing and trains with the remaining images, for every image in the dataset.

There were two important parameters to consider with respect to the texture classifier, k and T_{A_z}. The number of neighbours, k, is considered in the k-NN decision stage when comparing image features to training images. The regions shown in Figure 7.2 are scored individually and are used to determine the image score. These scores represent the posterior probability that the region or image shows lung abnormalities. When the region scores are combined to determine the image scores, the scores are weighted by the recorded performance, the area under the ROC curve, of each particular region. If the performance is under a chosen threshold T_{A_z}, the weighting is zeroed and the region score is not included in the calculation.

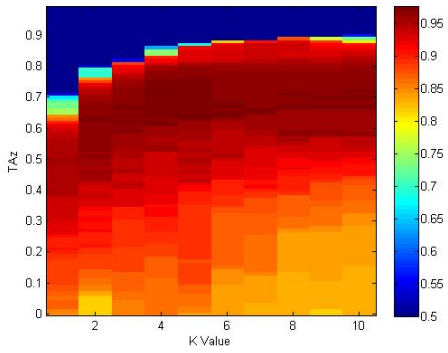


Figure 7.3: Map showing the area under the ROC curve for varying k and T_{A_z}. The optimal parameters correspond to the darkest red areas.

Figure 7.3 shows a map of the performance of the CAD tool at various k and T_{A_z} values. Mouton (2009) used the values $k = 9$ and $T_{A_z} = 0.8$ as they produced the best results. Given the drop off for T_{A_z} above 0.9 in Figure 7.3, a lower value was chosen here to reliably stay in the well-performing region. The values used by van Ginneken (2001) for the same texture analysis method falls in a well performing area in Figure 7.3 and so were used indiscriminately for all CAD testing in this project ($k = 7$ and $T_{A_z} = 0.60$).

The final result given for any image by the CAD system was a value in the range [0,1], giving the confidence of the tool that the image shows abnormality in the lung region. Zero indicates a certainty that the image was normal while 1 indicates it was abnormal. In a more practical setting this value could be thresholded for a binary result. The region scores can also be considered separately.

For the paediatric images, both the whole dataset and the TB-study subset are considered here. The whole dataset was considered to provide as much training information as possible. The TB study subset was included to determine the results from an unmixed dataset - with images with different scanning procedures there is always the chance of the classifier basing its classification on the image type and not from the image information. This is especially so when the abnormality distribution is biased towards a specific image type as it is in the paediatric database.

7.2 Observer Study

An observer study was performed on the paediatric image dataset. Observer studies show the performance of humans on the dataset. Two trained non-radiologists participated in the observer study done here.

The images in the dataset were randomly shuffled and presented to the observer. Each observer was then required to score the image out of 100 depending on the appearance of

abnormalities. A score of 0 would indicate that the image was clearly normal while a score of 100 indicated that the image shows obvious, severe abnormalities. This score, once scaled can be considered similar to the score in the range $[0, 1]$ provided by the CAD system.

The performance of the observers was calculated in the same way as that of the CAD system: in comparison to the original radiologists annotations. The scores were used to calculate the true and false positive rate for varying thresholds and the results were plotted as the ROC curve. The agreement of the observers was also calculated using kappa (McGinn et al., 2004). Kappa (also referred to as Cohen's kappa) attempts to calculate the agreement of observers beyond chance. It was determined as in Equation 7.1. An agreement of 0-0.4 would be considered slight to fair while above 0.6 would be substantial agreement (McGinn et al., 2004).

$$kappa = \frac{ObservedAgreement - ChanceAgreement}{ChanceAgreement} \tag{7.1}$$

Chapter 8

Results

The results of locating the ribs and subtracting them are presented in this chapter. The rib segmentations were evaluated using overlap with manual segmentations, used as the reference standard. The effect of rib suppression on lung texture classification was investigated. The unit of performance considered for the CAD tool was the area under the receiver operating characteristic (ROC) curve or A_z value. In addition, chest radiographs showing the rib suppression are given for the reader's assessment.

8.1 Rib Segmentation

The training process of the ASM segmentation method derived a statistical model of the ribs using the manual segmentations provided for training. The left and right sides of the ribcage were trained separately. Two ASM training models were derived during the training stage using the manually segmented images - one of the left ribs and one of the right ribs of the patient. Incorporating both sides into a single model only decreased the accuracy of the final segmentation and so the two sides were kept separate.

Segmentation methods could be combined by providing the output of one algorithm as the input to the other. For example, one combination was to use the output of the pixel classification (PC) method, the confidence image, as the input image to be segmented by the ASM. Using the PC and ASM in this way showed no improvement in the results. Using the lung segmentations to remove all non-lung area from the input image also showed no improvement on the ASM results. Therefore, the original radiographs were used as inputs to the ASM implementation.

It was also found that the ASM implementation failed to perform adequately without a manually selected starting point for the search algorithm. To provide this initial placement, the

user was asked to select four points on each side of the chest to place the mean rib model. These four points were selected to be the intersections of the top and bottom ribs with the boundary of the lungs. By doing this, the search algorithm starts in the vicinity of the expected position and was found to provide better performance.

A range of parameter combinations was tested with the ASM rib segmentation. The best performing parameters are shown below in Table 8.1 given the original images as the input and two separate training models. The parameters themselves were explained in Section 4.1.1. The model was found to iterate towards the ribs and then away from them, to outside the lung area if the number of positions searched above and below the ribs was too large. Thus the positions searched (n_{su} and n_{lu}) were limited to the values shown in the Table 8.1 to ensure the rib model stayed within the lungs within the maximum number of iterations, N_{max}.

Parameter	Value Chosen: Left	Value Chosen: Right
n_{su}/n_{lu}	6/6	6/7
N_{max}	40	
f	0.99	
m	1	
k_u	3	
k_l	3	
L	4	

Table 8.1: Suitable parameters found empirically for ASM segmentation of the ribs.

8.1.1 Overlap

The overlap value was calculated, as described in Section 5.4, to compare the results of the segmentation methods with the manual segmentation. Two separate models were trained and used for the left and right ribs in each image with ASM and so the overlap results are given separately for the left and right ribs. The results of the segmentation can be seen in Table 8.2.

Method	Left Overlap	Right Overlap	Average
Intra-observer	0.839 ±0.042	0.858 ±0.042	0.849
ASM	0.734 ±0.109	0.769 ±0.135	0.751
Initial placement	0.710 ±0.105	0.709 ±0.106	0.709
PC	0.626 ±0.146	0.603 ±0.127	0.615
PC-ASM Hybrid	0.351 ±0.276	0.279 ±0.223	0.315

Table 8.2: Overlap results for rib segmentation shown in order of performance. These values consider overlap of the manual segmentation of all ribs in the lung region with algorithm segmentations. "Intra-observer" is the overlap with the 5-rib training segmentations. "ASM" shows the results of active shape model segmentation. "Initial placement" shows the overlap with the user's placement of the ASM model. "PC" shows the results of pixel classification segmentation, which is trained with 5 ribs but returns a variable number of ribs. "PC-ASM" hybrid refers to the overlap results of combined PC and ASM segmentation

The PC-ASM segmentation trained and tested the ASM on the confidence images from pixel classification. This hybrid system returned the worst results of all the segmentation methods. The PC segmentation returned a better overlap - almost double that of the hybrid system. The initial placement segmentation overlap value shows the accuracy of the starting position of the ASM segmentation, i.e. the placement of the ASM model by the user before the ASM fits it to the image. The ASM segmentation improved upon this starting position, giving the best segmentation results.

The intra-observer overlap shows the overlap between the manual segmentation of five ribs, which constitutes the training segmentation for ASM, and the manual segmentation of all lung-overlapping ribs which constitutes the gold standard for comparing algorithm segmentations (see Section 5.1). These two sets of segmentation were performed separately and the overlap can be considered the intra-observer overlap. This comparison of manual segmentations returned the best overlap results.

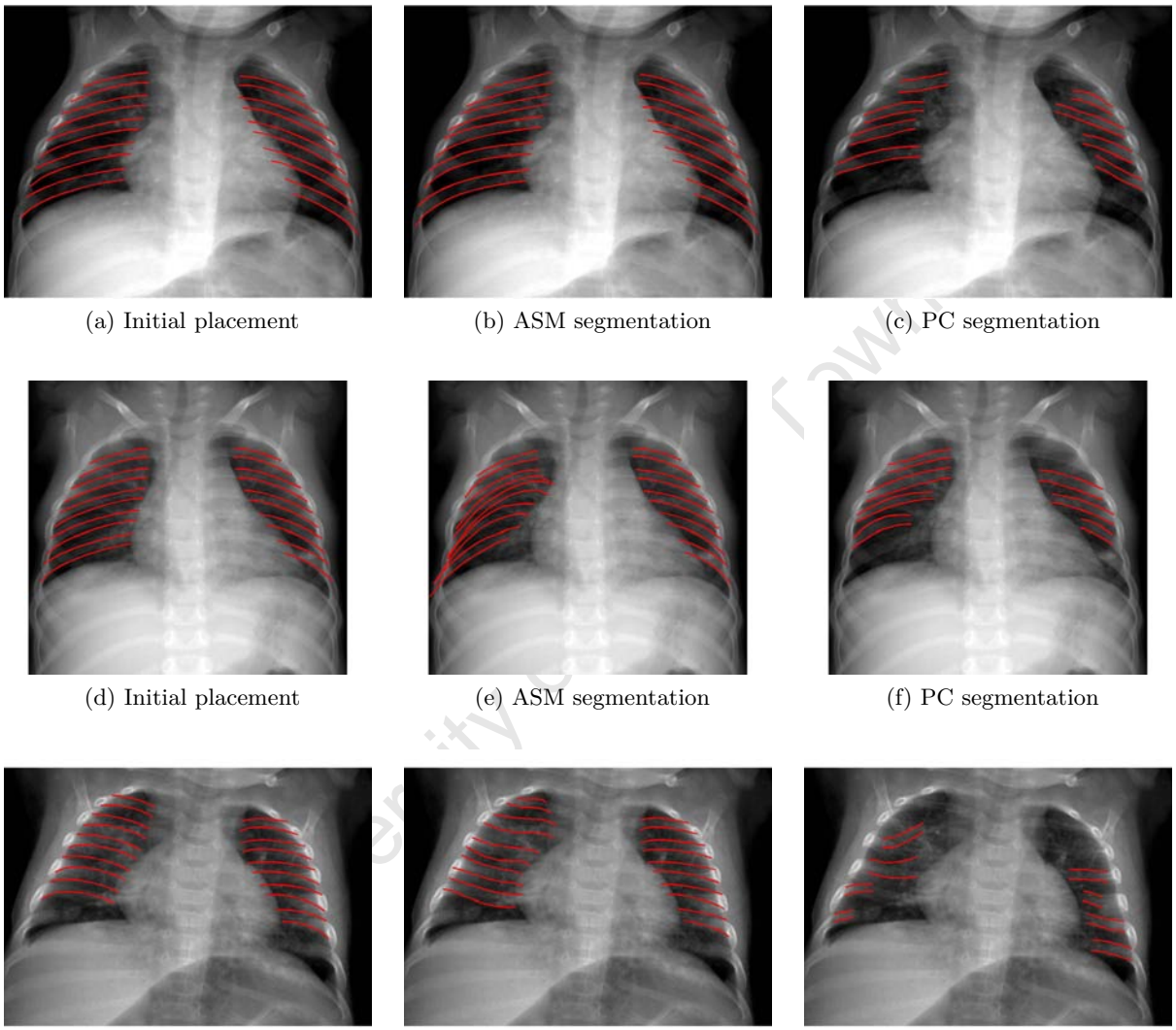
8.1.2 Visual Results

Visual examples of the rib segmentations can be seen in Figure 8.1. The results of the left and right segmentation are displayed together on the same image, even though each side was segmented independently. Each column shows a type of segmentation and each row a different example image. All three images were identified as containing abnormalities.

The first row is a typical example of segmentation where the ASM improves slightly upon the initial placement. The second row shows a segmentation failure over the right lung where two rib segmentations were adjusted away from the ribs rather than towards them. This occurred

more commonly on the right-side of the patient. The third example shows an image where the rib model did not match the shape of the ribs in the image. The ASM recovered well from this and moved towards the ribs.

In the PC segmentation images it can be seen that PC returns a variable number of ribs and typically less than the five found by ASM. While the contours match the ribs quite well, the segmentation contours tend to be offset slightly from the true border. In addition, the PC segmentation results typically lie outside the areas of abnormality while the ASM remains unaffected



(a) Initial placement (b) ASM segmentation (c) PC segmentation

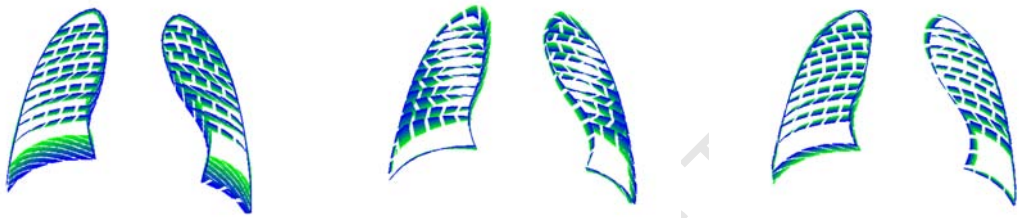
(d) Initial placement (e) ASM segmentation (f) PC segmentation

(g) Initial placement (h) ASM segmentation (i) PC segmentation

Figure 8.1: Segmentation results of the ribs. The final segmentation is shown by the red contours. The left column shows the manual placement of the rib model, the second column the results of ASM and the third column the PC segmentation results. The first row shows a typical image segmentation, the second a poor ASM segmentation result and the third an image where the initial rib model does not fit the ribs well.

8.1.3 Modes of Variation

Generation of the ASM model includes determining the allowable shape domain. Once it was generated, the modes of variation of the model could be viewed. Figure 8.2 shows the first three modes of variation of the combined lung and rib training model. A combined rib and lung model was tested initially but not used for the final rib segmentation. It is provided here to show the typical variation of the ribs and lungs. The primary mode of variation showed little movement of the ribs but significant change in position of the diaphragm. The second mode of variation found in the data showed the change in the angle of the ribs. The third mode of variation was the rotation of the patient shown by the asymmetry of the diaphragm.



(a) First mode (b) Second mode (c) Third mode

Figure 8.2: First three modes of variation of the combined model of the ribs and lungs. The change in shape is shown as a change in colour, from green to blue. Each mode was varied by 0.5 of the standard deviation.

8.2 Texture Analysis

This section presents the results of texture analysis without rib suppression. Observer studies on the paediatric dataset were performed to establish the performance of human observers in identifying pathology in the images. The results of Mouton's classifier are also presented in this section for comparison.

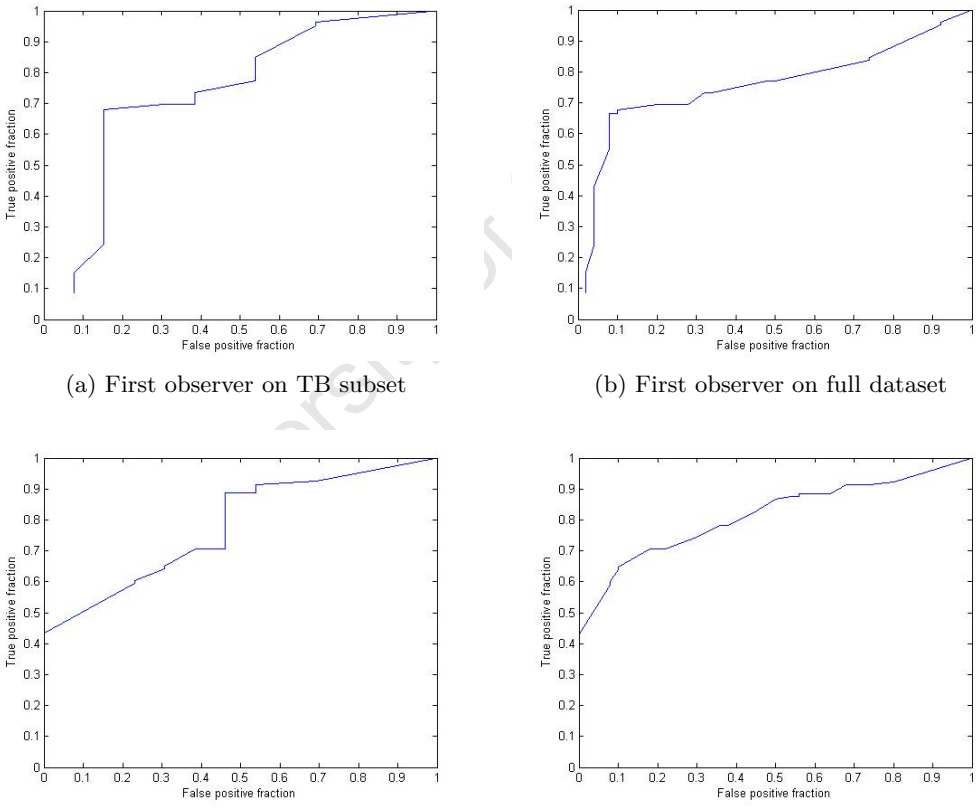
8.2.1 Paediatric Dataset

Two observers performed diagnoses on the paediatric dataset. Their results for the TB study subset and for the complete dataset were treated separately. The observers were trained non-radiologists with experience in analysing chest radiographs. For each image the observer scored the image according to the severity of the abnormalities, if any, that were present. Their performance was evaluated using the manual annotations given by the original radiologist (Mouton et al., 2010). The observers' scores are documented below in Table 8.3. The ROC curves corresponding to these AUC values are plotted in Figure 8.3. Mouton's classifier returned an $A_z = 0.778$ on the TB study and an $A_z = 0.953$ on the full paediatric dataset.

Observers scored the images with a value between 0 and 100. In order to score agreement, this range was divided into smaller ranges, for example 0-50 and 51-100, and if the two observers' scores fell in the same range, it was registered as an agreement. For a division of the range into two sections - essentially a yes/no classification - the agreement was $kappa = 0.383$. For five ranges of score the agreement was $kappa = 0.180$. The agreement for two ranges is fair and the agreement for five ranges is slight.

Observer	TB Study	TB and Trauma Study
Observer 1	0.732	0.759
Observer 2	0.779	0.815
CAD system	0.778	0.953

Table 8.3: The area under the ROC curves for the observer studies on the paediatric dataset. The TB study results are for the TB study images alone and TB and Trauma are for the full set of paediatric images.



(a) First observer on TB subset

(b) First observer on full dataset

(c) Second observer on TB subset

(d) Second observer on full dataset

Figure 8.3: ROC plots of the results of the observer study: for TB dataset and the whole dataset for each observer.

8.2.2 Adult Dataset

The CAD developed for and tested on the paediatric dataset by Mouton (2009) was run on the adult dataset. His lung segmentation tool was used to provide lung segmentations for the classifier. It is semi-automatic and a starting point had to be selected for each image by clicking the apexes of the lungs. Both the classifier and ASM method were trained on 216 adult chest images.

The training provided both texture examples for the classifier and the weights for the combination of region scores into an image score (as described in Section 2.3). The remaining 149 images were used to test the classifier. The results were a value of 0.83 for the area under the ROC curve.

8.3 Suppression

The results of the suppression are described below. As the suppression was performed in a context in which there are no ideal results for comparison, for example soft-tissue images provided by dual-energy imaging, the success of the suppression could only be considered with regard to textural analysis. Visual examples are also provided.

8.3.1 Visual Appearance

Figure 8.4 shows example before- and after-images of rib suppression from the paediatric dataset using manual segmentations. There is a noticeable decrease in the visibility of the ribs. The original ribs can still be seen in the suppressed image. Larger versions of PCA suppression and greyscale suppression are shown in Figure 8.5 for closer examination. Figure 8.6 shows an example of suppression on an image from the adult image set. Again the ribs are still discernible in the final image. Rib suppression on the adult images was more pronounced than in the paediatric images.

8.3.2 CAD Results

The CAD tool results for images suppressed using the manual segmentation, PC segmentation and ASM segmentation are presented in Table 8.4. The texture classifier was trained and tested using the suppressed images. Suppression results are shown for the following segmentations:

- Initial placement of the ASM model: starting position provided by the user for the ASM segmentation, but suppression using the placement without ASM segmentation

(a) Original image (b) PCA suppressed (c) Mean greyscale suppressed

(d) Original image (e) PCA suppressed (f) Mean greyscale suppressed

(g) Original image (h) PCA suppressed (i) Mean greyscale suppressed

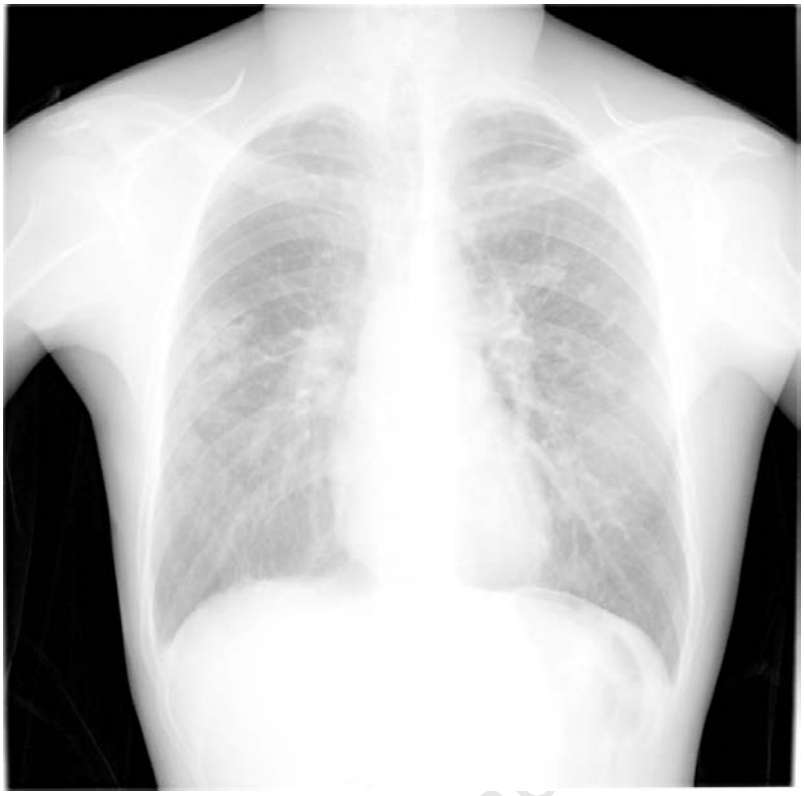
(j) Original image (k) PCA suppressed (l) Mean greyscale suppressed

Figure 8.4: Examples of suppressed images. The images on the left show the original unsuppressed images. The images on the right show the corresponding rib suppressed images. The suppression was applied to the posterior ribs overlapping the lungs using the manual (5-rib) segmentation of the ribs. The suppression uses a projection up from the lower rib (see section 6.1.2)

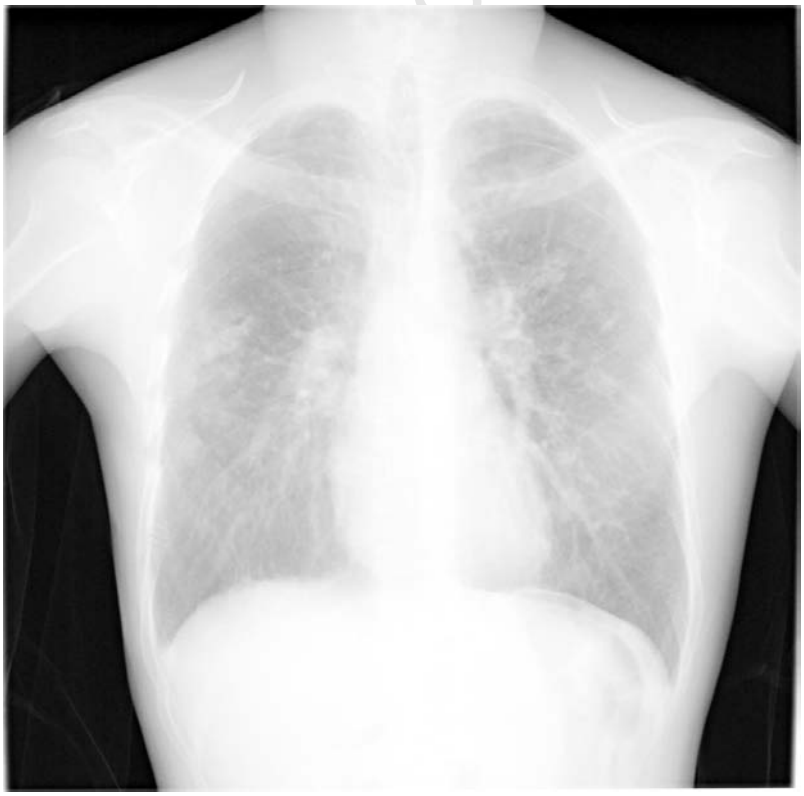
(a) Rib suppressed image using PCA modelling

(b) Rib suppressed image using greyscale modelling

Figure 8.5: The top image shows a PCA rib suppressed image. A projection up from the lower rib (see section 6.1.2) was taken and a manual segmentation of five ribs was used. The bottom image shows the same image suppressed using a mean greyscale model with the same manual segmentation.



(a) Original chest X-ray image



(b) Rib suppressed chest X-ray image

Figure 8.6: The top image shows an adult chest X-ray image and the bottom the corresponding PCA rib suppressed image. The rib extraction was by projection up from the lower rib (see Section 6.1.2) and a manual segmentation of all the ribs was used.

- ASM segmentation using initial manual placement of the model
- Manual segmentation of five ribs: the segmentation of the first five ribs overlapping the lungs provided for ASM training
- Manual segmentation of all ribs
- Pixel classification: the segmentation returned by PC segmentation

The following rib modelling methods were tested for rib suppression:

- PCA: modelling based on the first three modes of variation given by PCA
- Greyscale: modelling based on the average greyscale value of the rib area
- Noise: modelling using greyscale noise

Results are provided for testing on the whole paediatric dataset, as well as just the TB study subset. The TB study subset was included to provide an example of the results on a set of images taken with one set of imaging parameters. Considering the TB study images only, there was a total of 4788 sub-regions of which 1683 were abnormal, so while there are only 6 normal images, at a regional level the ratio of abnormal regions to normal regions was more balanced. Nonetheless the results for the TB study subset are lower than those of the complete dataset.

In Table 8.3a the results of texture analysis before the inclusion of rib suppression are shown. All values in Tables 8.3b and 8.3c are given with an arrow to indicate whether the value was greater than or less than this original texture analysis. At a glance, the effect of rib suppression was dependent on the segmentation, modelling technique used and dataset. Suppression typically resulted in an improvement in texture analysis on the full dataset and a deterioration in the TB study subset. The greatest improvement, on the full dataset, was shown by suppression using greyscale modelling and manual segmentation of all the ribs, using projection from the bottom rib. The greatest deterioration, on the full dataset, was shown by suppression using noise modelling.

Figure 8.7 shows the results averaged over the lower and upper rib projection methods (see Section 6.1.2) and the different datasets. The best performing segmentation method ASM and the manual segmentations are considered. It can be seen that using PCA modelling returns the best results and noise modelling the worst, on average. For the different segmentations, PCA was greatly affected by the segmentation used. The average AUC value of the texture classification without rib suppression was 0.866. The only method that performs better than this on average in the graph, was rib suppression using PCA modelling and manual segmentation of five ribs.

For greyscale and noise modelling methods a pattern can be seen: ASM segmentation returns the best results and manual segmentation of all the ribs the worst. This may be because

(a) AUC values for texture analysis without rib suppression

Method	TB Study	All Paediatric Images
Original texture classification	0.778	0.953

(b) AUC values for suppression using projection from the lower rib

Method	Segmentation	TB Study	All Paediatric Images
PCA	ASM initial placement	0.706 ▼	0.950 ▼
PCA	ASM final result	0.771 ▼	0.950 ▼
PCA	Manual of five ribs	0.781 ▲	0.961 ▲
PCA	Manual of all ribs	0.747 ▼	0.961 ▲
PCA	Pixel classification	0.711 ▼	0.948 ▲
Greyscale	Manual of five ribs	0.699 ▼	0.600 ▲
Greyscale	ASM final result	0.747 ▼	0.952 ▼
Greyscale	Manual of all ribs	0.735 ▼	0.962 ▲
Noise	Manual of five ribs	0.726 ▼	0.952 ▼
Noise	Manual of all ribs	0.699 ▼	0.954 ▲
Noise	ASM final result	0.763 ▼	0.956 ▲

(c) AUC values for suppression using projection from the upper rib

Method	Segmentation	TB Study	All Paediatric Images
PCA	ASM initial placement	0.808 ▲	0.961 ▲
PCA	ASM final result	0.741 ▼	0.951 ▼
PCA	Manual of five ribs	0.781 ▲	0.961 ▲
PCA	Manual of all ribs	0.747 ▼	0.963 ▲
PCA	Pixel classification	0.756 ▼	0.945 ▼
Greyscale	Manual of five ribs	0.767 ▼	0.962 ▲
Greyscale	ASM final result	0.781 ▲	0.963 ▲
Greyscale	Manual of all ribs	0.712 ▼	0.970 ▲
Noise	Manual of five ribs	0.744 ▼	0.944 ▼
Noise	Manual of all ribs	0.676 ▼	0.955 ▲
Noise	ASM final result	0.765 ▼	0.908 ▼

Table 8.4: The AUC results for the texture analysis on rib suppressed images. The difference between projection from the upper or lower rib is detailed in Section 6.1.2. Each value is given is given an arrow to indicate an increase or decrease from the original text classifier results shown in a.

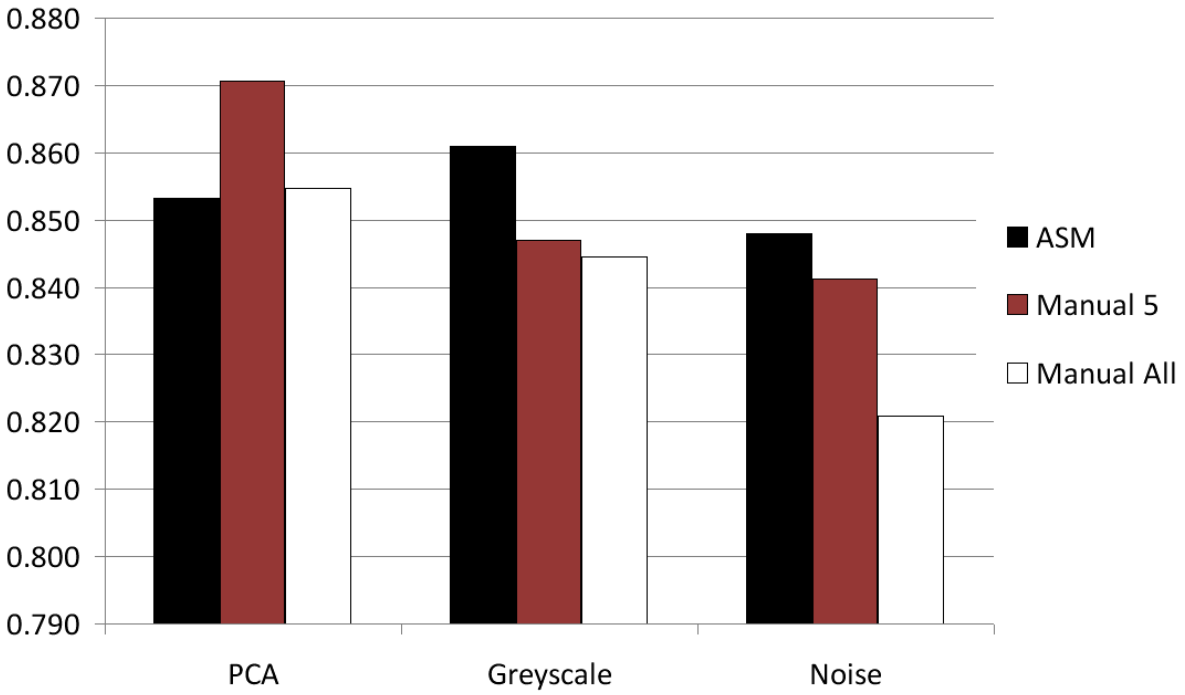


Figure 8.7: Graph showing the average AUC value for suppression for each modelling type, manual segmentation and segmentation by the best performing method - ASM using the user initial placement.

of the particular interaction between the segmentation and modelling methods, influenced by greyscale transitions at the rib borders. The noise modelling results indicate that the effect of rib suppression on texture classification is not random.

Chapter 9

Conclusions and Recommendations

In this chapter the results of the project are discussed. Possible improvements in each of the tasks - rib segmentation, texture classification and suppression - are considered. For rib segmentation, the overlap results and the final convergence are considered. For texture classification, the results of the observer study and the performance of the texture classifier on the adult dataset are examined. With regard to suppression, the results of the texture classifier with the use of rib suppression are considered.

9.1 Rib Segmentation

The review of chest radiograph segmentation by van Ginneken et al. (2006) shows that the degree of overlap of segmentation algorithm results with the gold standard varies depending on the structure being segmented. Automatic lung segmentations return high overlap values of 0.8 and above, whereas the clavicles return 0.7-0.8 at best. The clavicles are closer in structure to the ribs, being long, thin and defined by their edges. The ASM results here are comparable to the overlap results for clavicle segmentation, both being in the 0.7-0.8 range, but lower than the intra-observer overlap

While the ASM segmentation output approached the shape of the ribs, it did not converge, but relied on a maximum number of iterations as a stopping criterion. A possible reason for the failure to converge is insufficiently unique landmarks, given the appearance of the rib is very similar when rotated 180° and that if the segmentation results were to shift up or down a rib the results could still be a successful segmentation. A more extensive selection of landmarks may improve the ASM segmentation results.

Another problem could be the size of the database. The larger the database the better the training model, provided the dataset is representative of the actual population. The more variation typically present in the population, the more images are necessary. Paediatric images display a lot of variation in appearance. Further work should include determining if a larger training dataset would aid the ASM segmentation.

Equation 9.1 was given by van Ginneken (2001, p. 103), as the upper limit in the number of image pixels within which an ASM implementation, making use of more than one resolution of an image, is likely to converge on the correct structure. Given that N_{max} was 40, this safely puts the range in number of pixels beyond any current probable image size.

$$Range = kn_s 2^{N_{max}-1}\ pixels \tag{9.1}$$

k = Length of profile in points

n_s = Number of positions to search around the landmark

N_{max} = maximum number of iterations allowed

An ASM model can break down when the segmentation object is susceptible to bending or rotation (Cootes et al., 1995). Ribs are subject to some rotation in the chest radiograph. Imaging technique can also allow for some rotation of the patient. The first three modes of variation of the rib-lung model included such rotation. When searching an image for the structure to be segmented, the parameters of the shape model are adjusted linearly in the allowable shape domain. Rotation is a situation where a linear model can break down (Cootes et al., 1995) and allow the generation of unrealistic shapes. A non-linear model should be considered for possible improvement upon the results here.

To extend the ASM to segment a variable number of ribs, a complete ribcage could be modelled and the contour points in the model weighted. This weighting can be put to zero in case of non-visible ribs. This would require a large number of images with well visible ribs. The anterior portion of the ribs could also be included to reduce model rotation mistakes, but would limit the number of images suitable for training as their visibility is particularly low.

The images used in this project were used in JPEG format. The original images were in DICOM format and there may have been a loss in information in the conversion. The methods used here would benefit from images with more detail. The rib suppression appeared more effective on the higher resolution images in the adult dataset. The methods used here may show improvement when applied to original DICOM images.

9.2 Texture Classification

The performance of the texture classifier on the dataset was higher than that of the observers. A typical result one might expect from observers without CAD is around $A_z = 0.8$ (Abe et al., 2003) given the typical large size of the abnormalities. The texture classifier scored $A_z = 0.9529$ on the combined dataset, well above that of the observers. The lower performance could be due to reader inexperience with paediatric images.

The agreement between the observers was also fair to slight. Agreement between readers of chest radiographs is typically low but particularly low agreement can be attributed to poor experience or a difficult dataset. Including observers with more experience in paediatric images may lead to a better reference of diagnostic sensitivity to which to compare CAD results.

The results of the classifier on the adult dataset was closer to that of other classifiers. The classifier scored $A_z = 0.83$. On the same dataset, the textural classifier Hogeweg et al. (2010a) returned a result of 0.65. When combined with analysis of the lung shape - whether the shape of the lung was abnormal or not - their AUC increased to 0.85. It is important to note that different lung segmentations were used by Hogeweg et al. (2010a). The lung segmentation method used here was the ASM used by Mouton. Comparing the lung segmentations with those used by Hogeweg et al., it was seen that Mouton's ASM segmentation included more image area in the segmentation. This includes more abnormal lung area in the lung analysis and is comparable to analysing the lung shape.

A possible explanation for the high scores of the classifier on the paediatric dataset is that the classifier was able to recognise the difference in scanning resolution, despite the efforts to remove the differences. The abnormal images were all scanned at 4.16 lp/mm while the normal images, from the trauma study, were scanned at 2.08 lp/mm. The image resampling should have removed this but while two imaging environments were used, the possibility remains.

In this project the rib suppressed image was given as an input to the texture-based classifier. Another possible arrangement is supplying the suppressed image in conjunction with the original image when the features are calculated and compared. Considering both original and suppressed images in texture analysis instead of only the suppressed image could provide robustness.

9.3 Suppression

In Figure 8.7 greyscale and noise suppression showed a decrease in performance when manual (and presumably more accurate) segmentations were used. It would be expected that better segmentations would lead to better suppression and thus texture classification. Noise-based rib modelling, however, does not attempt to model the ribs and greyscale modelling was included

for comparison. PCA based rib suppression did not follow the same trend which means it is a candidate for successful rib suppression.

The effect of rib suppression varied with the segmentation used. Manual segmentation of five ribs showed a greater improvement in texture classification than the ASM segmentation of five ribs, in PCA based suppression. The manual segmentation of all the ribs over the lungs, however, performed worse than segmentation of five ribs on average. Given that there can be inter- and intra-observer variation in segmentation, for future work, more manual segmentations should be considered to reduce the effect of observer variation. The aggregation of multiple segmentations is suggested by Gerig et al. (2001). This approach should be investigated.

The starting contour of the suppression had a visible impact on the rib suppression. If the suppression was performed starting on the lower border, the lower border was less visible in the suppressed image. The upper border was less visible in the images suppressed starting on the upper border but these images showed a more visible lower border. The effect on the final AUC was not clear and both methods should still be considered in further work. Other possible methods include performing rib suppression twice, once from each border or projecting out from a midline contour.

The performance of the classifier used in this project, as seen in the original project by Mouton (2009), was already very high. Given this and the fact that the effect of the suppression was variable and small, there was insufficient data to reach a conclusion on the impact of the rib suppression. The slight increase may be random and not warrant inclusion in a CAD system. Further testing on different datasets may reveal which areas/types of image benefit from rib suppression. The performance of the CAD system may be expected to improve with more training examples of both normal and abnormal cases

The rib areas were extracted from images using a set rib width, as PCA accepts data instances with the same number of features. Realistically, both visually and physically, the width of the rib varies. The rib models did not show this variation. It is possible to model the variation by extracting the rib profiles at the estimated rib width at that point and resampling it to have the same number of points as the largest profile. Future work should include implementing such a scheme to better model the underlying rib.

The methods used here to compensate for changes in model intensity along the length of the rib should be further investigated. In the Riverain patent (Maton et al., 2009), it was stated that "ribs are examples of long, approximately linear structures." The intensity correction used here also assumed linearity. The suitability of this approach should be investigated and confirmed.

9.4 Summarised Project Contribution

This project investigated an implementation of rib suppression for aiding texture-based CAD in paediatric chest X-rays. Segmentation and suppression of the ribs was performed separately. The suppression was based on the work of Hogeweg et al. (2010b) who implemented and tested rib suppression in relation to texture-based CAD for TB on adult images. Segmentation using PC and ASM of five ribs was tested here where Hogeweg et al. (2010b) used manual segmentation only. In addition to the PCA method of Hogeweg et al. (2010b), greyscale and noise modelling were tested. The texture-based CAD system developed by Mouton (2009) for paediatric images was utilised for testing purposes.

The results can be divided into three sections: segmentation of the ribs, use of a texture based CAD system and the effect of rib suppression on the texture classification. Of the rib segmentation methods investigated, ASM performed the best segmentation of the ribs in the paediatric images. The ASM rib segmentation was comparable to similar segmentation in literature but will need further work in order to achieve results comparable to human inter-observer overlap.

The CAD system developed by Mouton (2009) was investigated. It was determined that the performance of the classifier on the paediatric dataset was greater than that of non-radiologists. In addition, its performance on an adult database was comparable to that of other CAD systems.

The results of the rib suppression methods investigated here showed that greyscale modelling was not a useful form of rib suppression. As was expected, noise-based modelling also proved to be unacceptable, but helped rule out a random effect of rib suppression on texture classification. PCA based rib suppression did show an improvement in texture classification in paediatric CAD but was sensitive to method and segmentation. More research will be needed to determine if the improvement warrants inclusion of rib suppression in the CAD system.

References

Abdool Karim, S., Churchyard, G. J., Abdool Karim, Q., and Lawn, S. D. (2009). HIV infection and tuberculosis in South Africa: An urgent need escalate the public health response. *Lancet*, 374:921–933.

Abe, H., MacMahon, H., Engelmann, R., Li, Q., Shiraishi, J., Katsuragawa, S., Aoyama, M., Ishida, T., Ashizawa, K., Metz, C. E., and Doi, K. (2003). Computer-aided diagnosis in chest radiography: Results of large-scale observer tests at the 1996-2001 RSNA Scientific Assemblies. *RadioGraphics*, 23:255–265.

Ahmed, B., Rasheed, T., Khan, M. A. U., Rashid, A., and Ahmad, S. (2007). Rib suppression in chest radiographs using ICA algorithm. *Information Technology Journal*, 6(7):1085–1089.

Armato, S. G. (2005). How digital capture is transforming chest radiography: Dual energy subtraction chest imaging. Retrieved on April 2, 2012, from http://www.imagingeconomics.com/issues/articles/2005-05_03.asp.

Arzhaeva, Y., Prokop, M., Tax, D. M., De Jong, P. A., Schaefer-Prokop, C. M., and van Ginneken, B. M. (2007). Computer-aided detection of interstitial abnormalities in chest radiographs using a reference standard based on computer tomography. *Medical Physics*, 34(12):4798–4809.

Becker, H. C., Nettleton, W. J. J., Meyers, P. H., Sweeney, J. W., and Nice, C. M. J. (1964). Digital computer determination of a medical diagnostic index directly from chest x-ray images. *IEEE Transactions on Bio-Medical Engineering*, 11(3):67–72.

Bjorkholm, P. J. (1985). Dual energy imaging. U.S. Patent No. *4,511,799*. Washington, DC: U.S. Patent and Trademark Office.

Bond, V., Chileshe, M., Sullivan, C., and Magazi, B. (2009). The converging impact of Tuberculosis, HIV/AIDS, and food insecurity in Zambia and South Africa. Technical report, RENEWAL.

Butler, S. A., Gabbay, R. J., Kass, D. A., Siedler, D. E., O'Shaughnessy, K. F., and Castellino, R. A. (2004). Computer-aided detection in diagnostic mammography: Detection of clinically unsuspected cancers. *American Journal of Roentgenology*, 183(5):1511–1515.

Castellano, G., Bonilha, L., Li, L. M., and Cendes, F. (2004). Texture analysis of medical images. *Clinical Radiology*, 59:1061–1069.

Chen, C. H. and Pau, L. F. (1998). *The handbook of pattern recognition and computer vision.* World Scientific Publishing Company, 2nd edition.

Chotas, H. G., Dobbins, J. T., and Ravin, C. E. (1999). Principles of digital radiography with large-area, electronically readable detectors: A review of the basics. *Radiology*, 210:595–599.

Cootes, T. F. and Taylor, C. J. (1992). Active shape models - 'smart snakes'. In *British Machine Vision Conference*, pages 266–275. Springer-Verlag.

Cootes, T. F., Taylor, C. J., Cooper, D. H., and Graham, J. (1995). Active shape models - their training and application. *Computer Vision and Image Understanding*, 61(1):38–59.

Cupples, T. E., Cunningham, J. E., and Reynolds, J. C. (2005). Impact of computer-aided detection in a regional screening mammography program. *American Journal of Roentgenology*, 185(4):944–950.

de Bruijne, M. and Nielsen, M. (2005). Multi-object segmentation using shape particles. In *Proceedings of the 19th international conference on Information Processing in Medical Imaging*, volume 3565 of *IPMI'05*, pages 762–773, Berlin, Heidelberg. Springer-Verlag.

de Hoop, B., De Boo, D. W., Gietema, H., A., van Hoorn, F., Mearadji, B., Schijf, L., van Ginneken, B., Prokop, M., and Schaefer-Prokop, C. (2010). Computer-aided detection of lung cancer on chest radiographs: Effects on observer performance. *Radiology*, 257:532–540.

Deserno, T. M. (2011). *Biomedical image processing.* Biological and Medical Physics, Biomedical Engineering. Springer.

Doi, K. (2007). Computer-aided diagnosis in medical imaging: Historical review, current status and future potential. *Computerized Medical Imaging and Graphics*, 31:198–211.

Doi, K. and Katsuragawa, S. (1989). Method and system for localization of inter-rib spaces and automated lung texture analysis in digital chest radiographs. U.S. Patent No. *4,851,984*. Washington, DC: U.S. Patent and Trademark Office.

Duda, R. O. and Hart, P. E. (1972). Use of the Hough transformation to detect lines and curves in pictures. *Communications of the ACM*, 15(1):11–15.

Fawcett, T. (2006). An introduction to ROC analysis. *Pattern Recognition Letters*, 27(8):861–874.

Fernandes, L. A. F. and Oliveira, M. M. (2008). Real-time line detection through an improved Hough transform. *Pattern Recognition*, 41(1):299–314.

Fonseca-Santos, J. (2005). Tuberculosis in children. *European Journal of Radiology*, 55(2):202–208.

Freer, T. W. and Ulissey, M. J. (2001). Screening mammography with computer-aided detection: Prospective study of 12,860 patients in a community breast center. *Radiology*, 220:781–786.

Friedland, G. (2008). Confronting the catastrophe of M/XDR-TB. In *15th Conference on Retroviruses and Opportunistic Infections*, Boston, MA, US. Abstract 112.

Gandhi, N. R., Moll, A., Sturm, A. W., Pawinski, R., Govender, T., Lalloo, U., Zeller, K., Andrews, J., and Friedland, G. (2006). Extensively drug-resistant tuberculosis as a cause of death in patients co-infected with tuberculosis and HIV in a rural area of South Africa. *Lancet*, 368:1575–1580.

Gerig, G., Jomier, M., and Chakos, M. (2001). Improving 3D object segmentation. In *Medical Image Computing and Computer-Assisted Intervention*, volume 2208 of *Lecture Notes in Computer Science*, pages 516–523.

Gie, R. (2003). Diagnostic atlas of intrathoracic tuberculosis in children. International Union Against Tuberculosis and Lung Disease, 68 Boulevard Saint-Michel, 75006 Paris, France.

Giger, M. L., Huo, Z., Kupinski, M. A., and Vyborny, C. J. (2000). Computer-aided diagnosis in mammography. In Sonka, M. and Fitzpatrick, J., editors, *Handbook of Medical Imaging, 2: Medical Image Processing and Analysis*, pages 915–1004, Bellingham, Wa. SPIE Press.

Hamarneh, G., Abu-Gharbieh, R., and Gustavsson, T. (1998). Active shape models - Part I: Modeling shape and gray level variations. In *Proceedings of the Swedish Symposium on Image Analysis*.

Harisinghani, M. G., McLoud, T. C., Shepard, J. O., Ko, J. P., Shroff, M. M., and Mueller, P. R. (2000). Tuberculosis from head to toe. *RadioGraphics*, 20:449–470.

Hastie, T., Tibshirani, R., and Frienman, J. (2009). *The elements of statistical learning: Data mining, inference and prediction*. Springer-Verlag, 2nd edition.

Hearn, G. E. and Metcalfe, A. V. (1995). *Spectral analysis in engineering: Concepts and cases*. Butterworth-Heinemann.

Hogeweg, L., Mol, C., de Jong, P., Dawson, R., Ayles, H., and van Ginneken, B. (2010a). Fusion of local and global detection systems to detect tuberculosis in chest radiographs. In *Medical Image Computing and Computer-Assisted Intervention*, volume 6363 of *Lecture Notes in Computer Science*, pages 650–657.

Hogeweg, L. E., Mol, C., de Jong, P. A., and van Ginneken, B. (2010b). Rib suppression in chest radiographs to improve classification of textural abnormalities. In *Medical Imaging*, volume 7624 of *Proceedings of SPIE*, pages 76240Y1–76240Y6.

Kleinert, S. and Horton, R. (2009). South Africa's health: Departing for a better future? *Lancet*, 374:759–760.

Knapp, J., W. S. (2009). Feature based neural network regression for feature suppression. U.S. Patent No. *2009/0290779 A1*. Washington, DC: U.S. Patent and Trademark Office.

Li, F., Engelmann, R., Pesce, L. L., Doi, K., Metz, C. E., and MacMahon, H. (2011a). Small lung cancers: Improved detection by use of bone suppression imaging - comparison with dual-energy subtraction chest radiography. *Radiology*, 261(3):937–949.

Li, F., Hara, T., Shiraishi, J., Engelmann, R., M. H., and Doi, K. (2011b). Improved detection of subtle lung nodules by use of chest radiographs with bone suppression imaging: Receiver operating characteristic analysis with and without localization. *American Journal of Roentgenology*, 196(5):W535–W541.

Loog, M. and van Ginneken, B. (2006). Segmentation of the posterior ribs in chest radiographs using iterated contextual pixel classification. *IEEE Transactions on Medical Imaging*, 25:602–611.

Loog, M., van Ginneken, B., and Schilham, A. M. R. (2006). Filter learning: Application to suppression of bony structures from chest radiographs. *Medical Image Analysis*, 10:826–840.

Marais, B. J. and Madhukar, P. (2007). New approaches and emerging technologies in the diagnosis of childhood tuberculosis. *Paediatric Respiratory Reviews*, 8:124–133.

MathWorks. Matlab. Retrieved on April 2, 2012, from http://www.mathworks.com/products/matlab/.

Maton, P., Worrell, S. T., and Kakumanu, P. (2009). Object removal from images. U.S. Patent No. *2009/0060372 A1*. Washington, DC: U.S. Patent and Trademark Office.

McGinn, T., Wyer, P. C., Newman, T. B., Keitz, S., Leipzig, R., and Guyatt, G. (2004). Tips for learners of evidence-based medicine: 3. Measures of observer variability (kappa statistic). *Canadian Medical Association Journal*, 171(11):1369–1373.

Meyers, P. H., Nice, C. M. J., Becker, H. C., Nettleton, W. J. Jr., S. J. W., and Meckstroth, G. R. (1964). Automated computer analysis of radiographic images. *Radiology*, 83:1029–1034.

Mouton, A. (2009). Computer-aided diagnosis of tuberculosis in paediatric chest x-rays using local textural analysis. Master's thesis, University of Cape Town.

Mouton, A., Douglas, T. S., and Pitcher, R. D. (2010). Computer-aided detection of pulmonary pathology in paediatric chest radiographs. In *Lecture Notes in Computer Science*, volume 6363, pages 619–625.

Novelline, R. A., editor (2004). *Squire's fundamentals of radiology.* Harvard University Press, Cambridge, MA, 6th edition.

Pitcher, R. D., Wilde, J. C., Douglas, T. S., and van As, A. B. (2009). The use of the Statscan digital X-ray unit in paediatric polytrauma. *Paediatric Radiology*, 39:433–437.

Powell, G. F., Doi, K., and Katsuragawa, S. (1988). Localization of inter-rib spaces for lung texture analysis and computer-aided diagnosis in digital chest images. *Medical Physics*, 15(4):581–587.

Ramachandran, J., Pattichis, M. S., and Soliz, P. (2002). Pre-classification of chest radiographs for improved active shape model segmentation of ribs. In *5th IEEE Southwest Symposium on Image Analysus and Interpretation*, pages 188–192, Santa Fe, New Mexico.

Riverain. ClearRead Bone Suppression. Retrieved on April 2, 2012, from http://www.riverainmedical.com/ClearRead-bone-suppression-product-overview.html.

Samei, E., Flynn, M. J., and Eyler, R. E. (1999). Detection of subtle lung nodules: Relative influence of quantum and anatomic noise on chest radiographs. *Radiology*, 213:727–734.

Samulski, M. R. M., Snoeren, P. R., van Ginneken, B., and Hogeweg, L., K. N. (2011). Computer-aided detection as a decision assistant in chest radiography. In *Medical Imaging*, volume 7966 of *Proceedings of SPIE*, pages 796614–1–796614–6.

Sarkar, S. and Chaudhuri, S. (1998). Evaluation and progression analysis of pulmonary tuberculosis from digital chest radiographs. *Computerized Medical Imaging and Graphics*, 22:145–155.

Shah, S. K., Austin, J. H. M. M., White, C. S., Patel, P., Haramati, L. B., Pearson, G. D. N., Shiau, M. C., and Berkmen, Y. M. (2003). Missed non-small cell lung cancer: Radiographic findings of potentially resectable lesions evident only in retrospect. *Radiology*, 226:235–241.

Shlens, J. (2005). A tutorial on principal component analysis. Retrieved on April 2, 2012, from http://www.brainmapping.org/NITP/PNA/Readings/pca.pdf.

Stegmann, M. B. (2002). Analysis and segmentation of face images using point annotations and linear subspace techniques. Technical report, Informatics and Mathematical Modelling, Technical University of Denmark, DTU, Richard Petersens Plads, Building 321, DK-2800 Kgs. Lyngby.

Stegmann, M. B. and Gomez, D. D. (2002). A brief introduction to statistical shape analysis. Technical report, University of Denmark, DTU. Retrieved on April 2, 2012, from http://www2.imm.dtu.dk/pubdb/views/publication_details.php?id=403.

Sutton, L. N. (2011). PACS and diagnostic imaging service delivery - a UK perspective. *European Journal of Radiology*, 78:243–249.

Suzuki, K., Abe, H., MacMahon, H., and Doi, K. (2006). Image-processing technique for suppressing ribs in chest radiographs by means of massive training artificial neural network (MTANN). *IEEE Transactions on Medical Imaging*, 25(4):406–416.

Swaminathan, S. and Rekha, B. (2010). Pediatric tuberculosis: Global overview and challenges. *Clinical Infectious Diseases*, 50(S3):S184–S194.

Uemura, M., Miyagawa, M., Yasuhara, Y., M. T., Ikura, H., S. K., Tagashira, H., Arakawa, K., and Mochizuki, T. (2005). Clinical evaluation of pulmonary nodules with dual-exposure dual-energy subtraction chest radiography. *Radiation Medicine*, 23(6):391–397.

van Ginneken, B. (2001). *Computer-aided diagnosis in chest radiography*. PhD thesis, Utrecht University.

van Ginneken, B., Katsuragawa, S., ter Haar Romeny, B. M., Doi, K., and Viergever, M. A. (2002). Automatic detection of abnormalities in chest radiographs using local texture analysis. *IEEE Transactions on Medical Imaging*, 21(2):139–149.

van Ginneken, B., Stegmann, M. B., and Loog, M. (2006). Segmentation of anatomical structures in chest radiographs using supervised methods: A comparative study on a public database. *Medical Image Analysis*, 10:19–40.

van Ginneken, B. and ter Haar Romeny, B. M. (2000). Automatic delineation of ribs in frontal chest radiographs. In *Medical Imaging*, volume 3979 of *Proceedings of SPIE*, pages 825–836.

van Ginneken, B., ter Haar Romeny, B. M., and Viergever, M. A. (2001). Computer-aided diagnosis in chest radiography: A survey. *IEEE Transactions on Medical Imaging*, 20(12):1228–1241.

Vogelsang, F., Weiler, F., Dahmen, J., Kilbinger, M., Wein, B., and Gunther, R. W. (1998). Detection and compensation of rib structures in chest radiographs for diagnose assistance. In *Medical Imaging*, volume 3338 of *Proceedings of SPIE*, pages 774–785.

Wechsler, H. and Sklansky, J. (1977). Finding the rib cage in chest radiographs. *Pattern Recognition*, 9:21–30.

Weyer, K. (2007). Case study: South Africa. Bulletin of the World Health Organization.

WHO (2006). Tuberculosis - Factsheets - TB and children. Retrieved on April 2, 2012, from http://www.searo.who.int/en/Section10/Section2097/Section2106_10681.htm.

WHO (2010). Multidrug and extensively drug-resistant TB (M/XDR-TB): 2010 global report on suveillance and response. Retrieved on April 2, 2012, from http://whqlibdoc.who.int/publications/2010/9789241599191_eng.pdf.

WHO (2011a). Who report 2011: Global tuberculosis control. Retrieved on April 2, 2012, from http://whqlibdoc.who.int/publications/2011/9789241564380_eng.pdf.

WHO (2011b). World health statistics. Retrieved on April 2, 2012, from http://www.who.int/gho/publications/world_health_statistics/en/index.html.

Yue, Z., Goshtasby, A., and Ackerman, L. V. (1995). Automatic detection of rib borders in chest radiographs. *IEEE Transactions on Medical Imaging*, 14(3):525–536.