COMPUTED TOMOGRAPHY IMAGE ANALYSIS FOR THE DETECTION OF OBSTRUCTIVE LUNG DISEASES

François Chabat

June 2000

A thesis submitted for the degree of Doctor of Philosophy of the University of London
and for the Diploma of Membership of Imperial College

National Heart and Lung Institute
Imperial College of Science, Technology and Medicine
University of London
London SW3 6LY

ABSTRACT

Damage to the small airways resulting from direct lung injury or associated with many systemic disorders is not easy to identify. Non-invasive techniques such as chest radiography or conventional tests of lung function often cannot reveal the pathology. On Computed Tomography (CT) images, the signs suggesting the presence of obstructive airways disease are subtle, and inter- and intra-observer variability can be considerable. The goal of this research was to implement a system for the automated analysis of CT data of the lungs. Its function is to help clinicians establish a confident assessment of specific obstructive airways diseases and increase the precision of investigation of structure/function relationships. To help resolve the ambiguities of the CT scans, the main objectives of our system were to provide a functional description of the raster images, extract semi-quantitative measurements of the extent of obstructive airways disease and propose a clinical diagnosis aid using *a priori* knowledge of CT image features of the diseased lungs. The diagnostic process presented in this thesis involves the extraction and analysis of multiple findings. Several novel low-level computer vision feature extractors and image processing algorithms were developed for extracting the extent of the hypo-attenuated areas, textural characterisation of the lung parenchyma, and morphological description of the bronchi. The fusion of the results of these extractors was achieved with a probabilistic network combining *a priori* knowledge of lung pathology. Creating a CT lung phantom allowed for the initial validation of the proposed methods. Performance of the techniques was then assessed with clinical trials involving other diagnostic tests and expert chest radiologists. The results of the proposed system for diagnostic decision-support demonstrated the feasibility and importance of information fusion in medical image interpretation.

To my father

ACKNOWLEDGEMENTS

I wish to express my gratitude to Professor David M Hansell and Dr Guang-Zhong Yang who supervised this research. Their advice, encouragement, and criticisms proved invaluable for the realisation of this work. I wish to thank them for their teaching and support, and for giving me the opportunity to work with them.

I am indebted to Imatron, Inc. (San Francisco, CA), for providing me with financial support during the whole length of this project. I particularly thank Mr David King for facilitating this funding.

I wish to thank Dr Michael Rubens, Director of Imaging at the Royal Brompton Hospital, for providing unhindered access to all the facilities of the Radiology Department.

I am thankful to Ms Clare Charrier, Ms Permi Jhooti, Mr Steven Collins and all my colleagues at the Cardiovascular Magnetic Resonance Unit and Computed Tomography Department, for helping me with my work at the Royal Brompton Hospital.

Thanks to Mathieu, Eyal and my friends. Thanks to Kevin.

I wish to thank my sister Emmanuelle and my father Benjamin for their kind and unconditional support.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1

INTRODUCTION

1.1 INTRODUCTION TO OBSTRUCTIVE LUNG DISEASES

1.1.1 Anatomy of the normal lung

The essential function of the lung is the exchange of oxygen and carbon dioxide between the blood and the atmosphere, by a process of molecular diffusion across an alveolar membrane. The lungs are divided into lobes – three on the right and two on the left – separated by invaginations of the pleural space. Each lobe is further subdivided by incomplete fibrous septa. Individual lobules within each segment are outlined by smaller incomplete septa. Recently, the anatomical unit of the *secondary pulmonary lobule* has found favour with radiologists because it is clearly recognisable in diseased states on high-resolution computed tomography. The acinus, a portion of the lung supplied by the terminal bronchioles, is regarded by some authorities as the functional unit of the lungs. Passive gas exchange takes place in the alveoli of the acini. The branching of the conducting airways follows a hierarchical pattern. The trachea divides into bronchi, lobar bronchi, then segmental bronchi. Further divisions take place in an uneven dichotomous fashion. Smaller airways without cartilage in their walls are referred to as bronchioles. The bronchioles are accompanied by thin-walled pulmonary arterioles. Figure (1.1) provides a schematic representation of the anatomy of a group of secondary pulmonary lobules, and shows how they are depicted on high-resolution computed tomography. A photomicrographic example of secondary pulmonary lobule is also illustrated in Figure (1.2a). The term *small airways* is generally used when referring to those airways with a diameter less than 2-3 mm, and describes a concept rather than a specific anatomic group of airways.

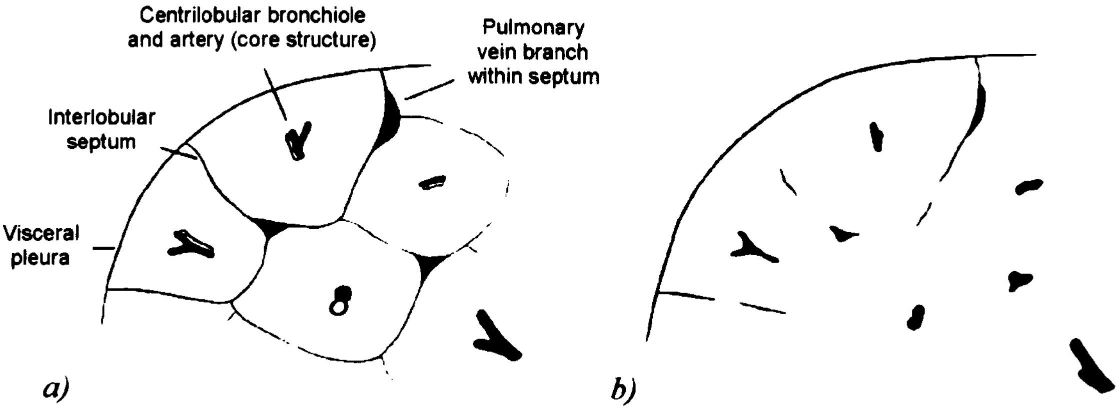


Figure 1.1. *a)* Anatomy of a group of second pulmonary lobules. *b)* Features of pulmonary lobules identifiable on High-Resolution Computed Tomography (HRCT). (Reproduced with permission from *Imaging of Diseases of the Chest*, P. Armstrong, A.G. Wilson, P. Dee, D.M. Hansell, Third edition, London: Mosby, 2000.)

1.1.2 Pathology of obstructive lung diseases

Obstructive lung disease includes many pathological conditions characterised by airflow obstruction. Four main categories can be distinguished: emphysema, small airways disease, chronic bronchitis, and asthma.

Emphysema is defined as a condition of the lungs characterised by "abnormal, permanent enlargement of air spaces distal to the terminal bronchioles, accompanied by destruction of their walls" [1], with no obvious fibrosis [2]. In the case of panlobular emphysema, the entire acinus is affected uniformly, with destruction and distortion of airspaces within lobules and acini, as illustrated in the photomicrograph of Figure (1.2b). In the case of centrilobular emphysema, abnormal air spaces are situated in the central portions of secondary lobules, surrounded by normal parenchyma, as shown in the photomicrograph of Figure (1.2c). The two types of emphysema may coexist in one patient.

The concept of small airways disease is based on the work of Hogg *et al* [3] who identified a site of airflow resistance in the peripheral airways. Constrictive obliterative bronchiolitis is a non-specific pathologic finding that may be primary or secondary to a number of disorders. It has been documented in connective tissue diseases, after toxic fume inhalation or viral infection and as a manifestation of chronic rejection in allograft transplantation [4]. An example of bronchiolar obliteration is illustrated in Figure (1.2d).

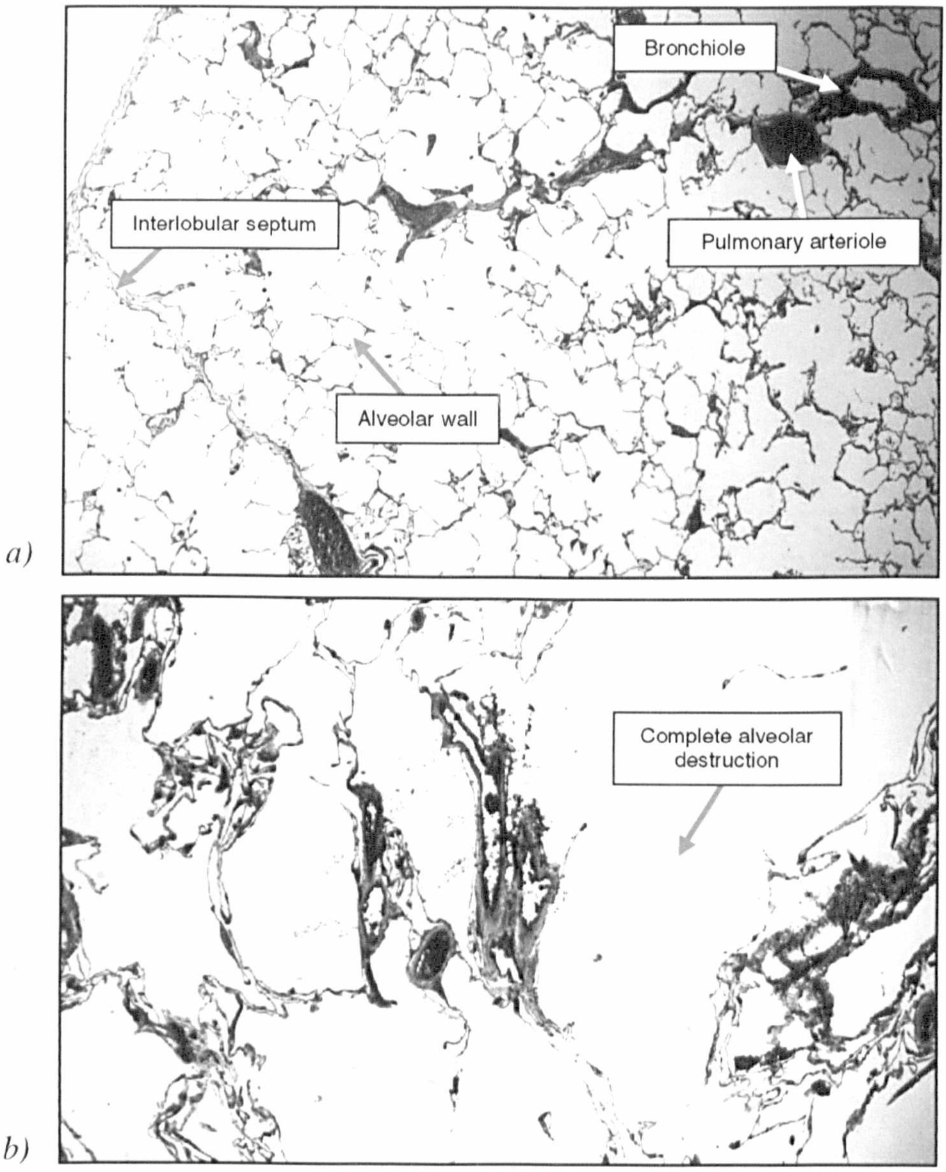


Figure 1.2. Photomicrography of normal lung and three obstructive lung diseases. The illustrations were not produced at the same magnification. *a)* Normal lung. The alveoli are organised in a regular pattern. A bronchiole and its adjacent pulmonary arteriole, running perpendicular to the section, can be observed. *b)* Panlobular emphysema. The disease results in complete alveolar destruction.

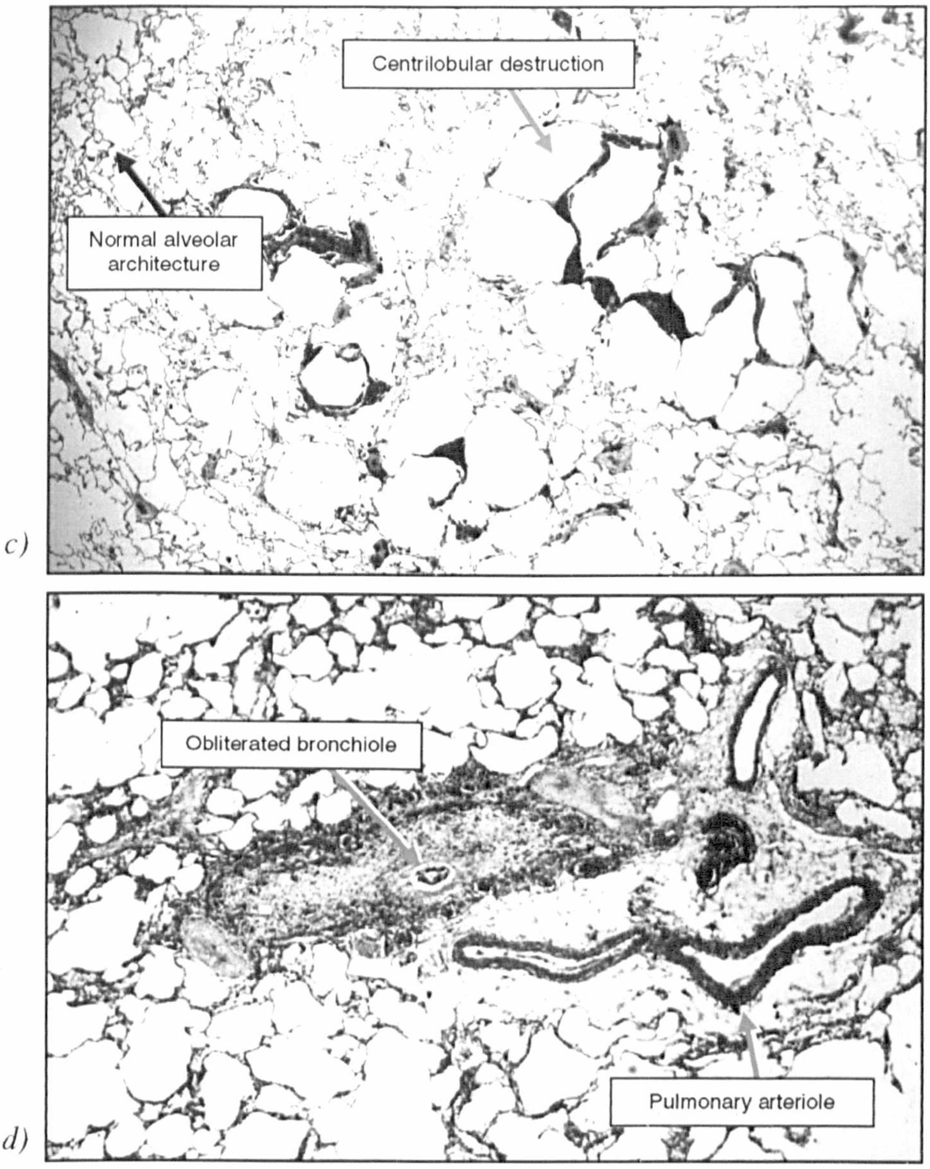


Figure 1.2 ***(continued).*** *c)* Centrilobular emphysema. Areas of centrilobular destruction appear adjacent with areas of normal alveolar architecture. *d)* Constrictive obliterative bronchiolitis. The lumen of a bronchiole running perpendicular to the section appears obstructed.

The diagnosis of obstructive lung disease is important as each condition has a different prognosis. For emphysematous patients, it also gives weight to smoking cessation advice and reliable quantification of the extent of the disease has the potential to chart disease progression and the effects of treatment, for example gene therapy.

The conditions studied in this thesis are mainly emphysema and small airways disease. Asthma and chronic bronchitis will not be addressed in this thesis, as they are usually diagnosed without resorting to imaging techniques. The presence of chronic bronchitis, characterised by hypersecretion of mucus by the respiratory tract [5], is established on the basis of a clinical diagnosis (stereotypical history of recurrent productive cough [6]). Similarly, the diagnosis of asthma, characterised by reversible intermittent airways constriction, known as *bronchospasm*, can be established simply from the patient's clinical history and functional response to bronchodilator treatment.

1.1.3 Detection of obstructive lung diseases

Pulmonary function testing, such as spirometry, which quantifies expiratory air flow rate is of diagnostic value for the detection of obstructive lung diseases. The measure of the forced expiratory volume after 1 s (FEV_1), forced vital capacity (FVC), diffusion capacity of the lung for carbon monoxide (DL_{CO}), are also used for the detection of airways diseases. For instance, abnormally low values of FEV_1, FEV_1/FVC and DL_{CO} are typical in emphysema. Values are regarded as abnormal when they are outside the 95 % range of people with same age, height and sex. However, pulmonary function tests are insensitive to mild disease, cannot reproducibly and reliably measure the extent of emphysema [7], and cannot distinguish between the different causes of obstructive lung diseases.

Although chest radiography is a valuable modality for the diagnosis of pulmonary diseases, it is insensitive for all but the most advanced cases of emphysema. Thurlbeck *et al* [8] demonstrated by post-mortem examination of the lungs of 696 patients, that only 41 % of emphysematous cases were diagnosed correctly on chest radiography. Similarly, the manifestations of small airways disease on chest radiography, as reviewed by McLoud *et al* [9], are often too subtle to be distinguished from the normal state.

The emergence of computed tomography and particularly high-resolution computed tomography has allowed the improved demonstration of parenchymal lung architecture. These modalities have become the most accurate imaging techniques for the detection of obstructive lung diseases. Nevertheless, the accurate quantification and differentiation of the main forms of obstructive lung diseases identifiable on HRCT remains challenging, as evidenced by the degree of observer variation in several studies [10-12].

1.2 COMPUTED TOMOGRAPHY OF THE LUNGS

1.2.1 Computed Tomography

Computed Tomography (CT) is an imaging modality that produces two-dimensional representations of the linear X-ray attenuation coefficients through a narrow planar cross-section [13]. Data for CT is acquired using a rotating X-ray source. Sensors measure the decrease in the intensity of the emitted X-ray beam after it has crossed the object scanned. For every angle θ of the rotation of the X-ray source around the imaged object, sensors measure a one-dimensional profile of the X-ray attenuation by the object, as illustrated in Figure (1.3). The set of profiles obtained for all θ provides the back-projection of the attenuation. It is the possible to reconstruct from this data a cross-sectional image of the object, by using inverse Radon transformation. The grey-scale intensity at every pixel on the reconstructed image represents the CT number at the co-ordinates of the pixel, of a volume of thickness L, L being the width, or collimation, of the X-ray beam. Given the measured linear attenuation coefficient μ of a structure, the corresponding CT number n, in Hounsfield Unit (H.U.), is given by:

$$n = K \frac{\mu - \mu_{water}}{\mu_{water} - \mu_{air}} \qquad \textbf{(1.1)}$$

where μ_{water} and μ_{air} are the linear attenuation coefficients of water and air respectively, and K is a constant. The constant K is normally chosen equal to 1000. Thus, CT number of air is -1000, and CT number of water is 0. When imaging inhomogeneous objects, CT numbers represent the average attenuation of the structures present at every voxel. This phenomenon, referred to as partial volume effect, produces CT numbers that are difficult to interpret at the interface of anatomical structures with different CT densities [14]. For displaying CT images, two parameters are selected, window level l, and window width w. CT numbers are represented on a grey-scale, varying linearly from black to white for a range of values from $l-(w/2)$ to $l+(w/2)$, as shown in Figure (1.4). The values of l and w are chosen according to the typical CT densities of the structures to examine.

Artefacts observed on CT can be due to random noise, the geometry of the scanner (size, discretisation), non-monochromatic X-rays, X-ray intensity fluctuations and object motion. The impact of the latter artefact can be minimised by using fast acquisition scanners, notably Electron-Beam Computed Tomography (EBCT) scanners.

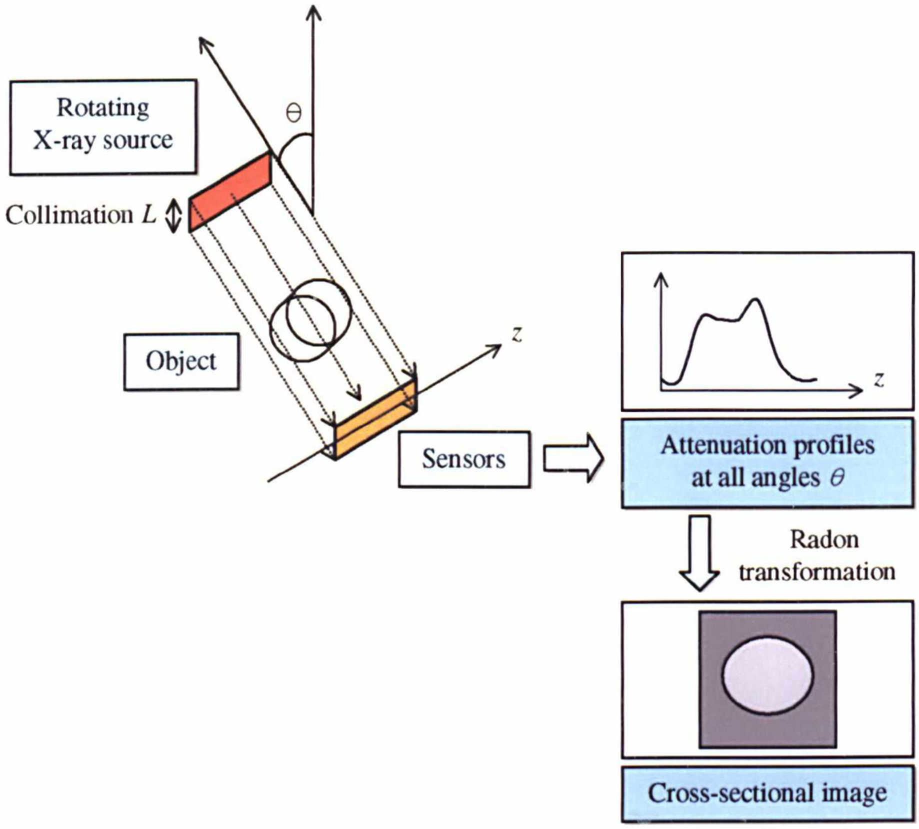


Figure 1.3. Schematic representation of data acquisition with CT. Sensors measure the decrease of intensity of an X-ray beam emitted by a rotating source. With the one-dimensional attenuation profiles thus obtained at all angles *θ*, a cross-sectional image of a slice with a width *L* of the object can be reconstructed.

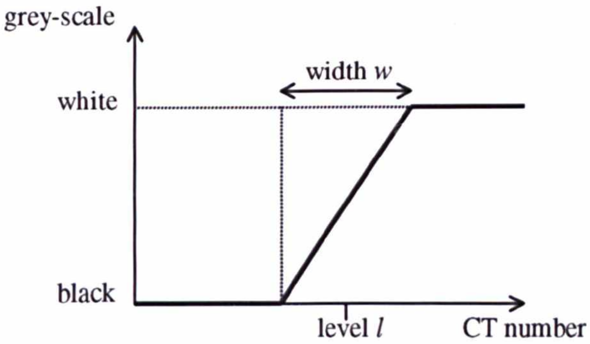


Figure 1.4. Definition of window level *l* and window width *w*. CT numbers are displayed on a linear grey-scale ranging from black for values of *l - (w/2)* and less, to white for values of *l + (w/2)* and more.

The X-rays used in EBCT are not emitted by a mechanically rotating X-ray source, but by tungsten targets hit by an electron beam. The electron beam being focused and deflected by electro-magnetic coils, there are no moving parts in the system, and exposure time can be reduced to 50 ms (as opposed to 500 ms on conventional mechanical CT scanners). A schematic illustration of the EBCT scanner used for this study is given in Figure (1.5).

Since the mid-1980s, technical hardware improvements have made it possible to increase the spatial resolution of computed tomography, culminating in the emergence of High-Resolution Computed Tomography (HRCT). HRCT is characterised by thin scan collimation, shortest possible scan time, high-spatial-frequency reconstruction algorithm (also referred to as 'sharp', 'edge-enhancing', or 'bone'), and targeted reconstruction on a selected small field of view of particular interest.

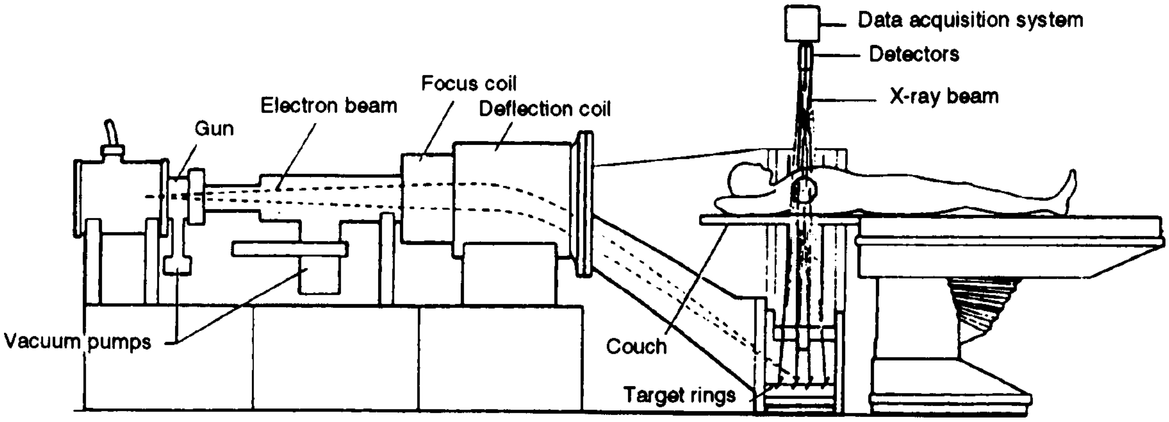


Figure 1.5. Schematic illustration of an Electron Beam Computed Tomography (EBCT) scanner (courtesy of Imatron, Inc., San Francisco, CA). An electron beam, focused and deflected by electro-magnetic coils, hits tungsten targets rings. The targets then emit X-ray beams, whose attenuation after they have crossed the patient is measured by detectors. The absence of moving mechanical parts allows for ultrafast data acquisition.

1.2.2 HRCT findings of obstructive lung diseases.

HRCT has been demonstrated to be the most sensitive modality for the *in vivo* detection of obstructive lung diseases. It also allows for subjective assessment of the extent and severity of these conditions. Differential diagnosis is achieved by taking into account many visual features: hypo-attenuation of the lung parenchyma, bronchial abnormalities, and characteristics of the parenchymal texture.

1.2.2.1 Hypo-attenuated lung

A cardinal CT finding of obstructive lung diseases is the presence of areas of hypo-attenuation of the lung parenchyma [4,15]. In the case of emphysema, the hypo-attenuated areas are a direct consequence of lung destruction; the lung parenchyma is replaced by air. In airways centred disease, there is pathological air-trapping resulting in under-ventilation and, by physiological reflex, under-perfusion. Diseased areas appear abnormally hypo-attenuated, as compared to adjacent areas of normal lung. This pattern is referred to as mosaic attenuation. As an example, Figure (1.6) shows three CT images of lungs with different degrees of severity of small airways disease. Hypo-attenuated areas in the lung parenchyma are obvious in Figure (1.6a). However, it becomes harder to detect them on mild and asymptomatic cases, as shown in Figures (1.6b) and (1.6c) respectively. The resulting density differences may be subtle and the hypo-attenuated areas may be poorly marginated. The detection of hypo-attenuated lung is further complicated by the fact that the observed differences in the lung parenchyma may not be pathological. In fact, a linear increase in opacification from the non-dependent to the dependent lung is observed in normal lung. The physiological mechanism of this density gradient has been ascribed to gravity-dependent perfusion [16-19] and/or relative atelectasis of the dependent lung [20-22]. On CT images, this gradient can be observed in most cases: the dependent part of the parenchyma appears denser than the ventral part, as illustrated by the scan of a normal subject in Figure (1.7a). The density histograms of two regions of interest, one in the dependent lung and one in the non-dependent lung, as given in Figure (1.7b), demonstrate the shift in CT densities.

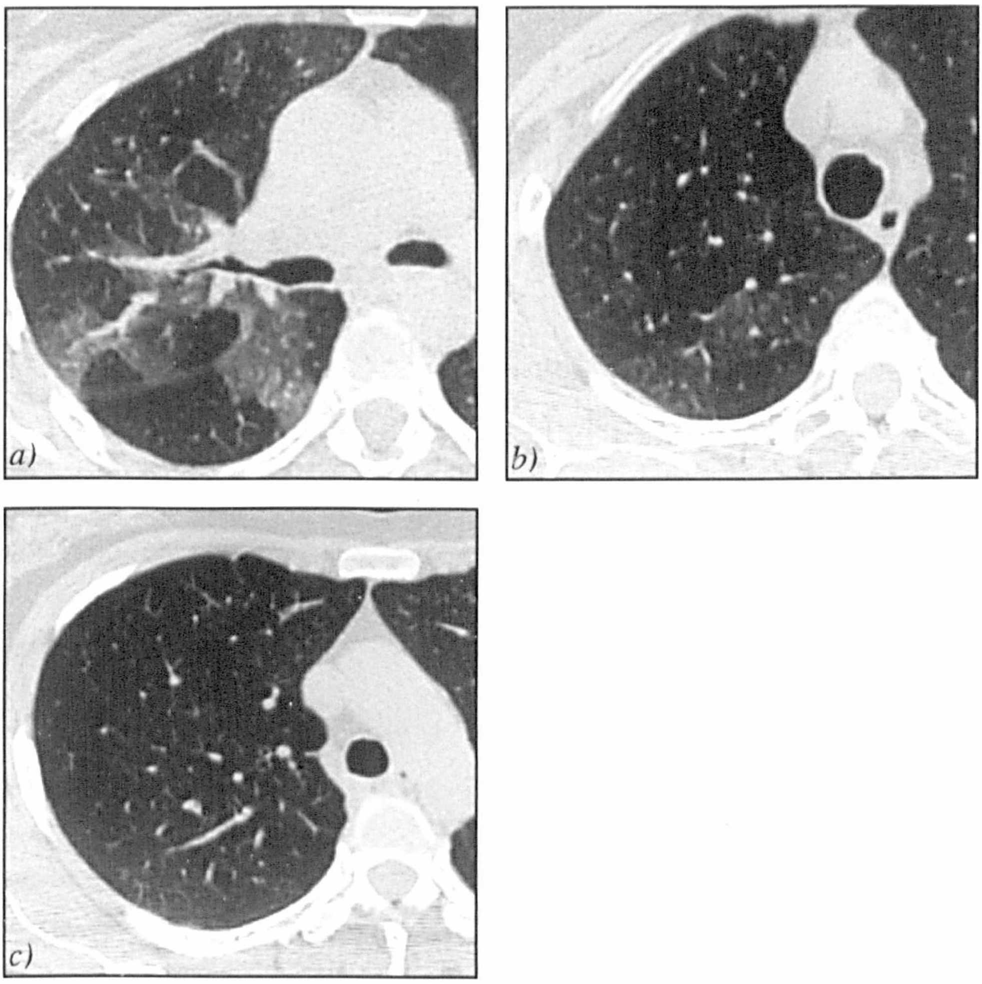


Figure 1.6. Three examples of CT scans of lungs with constrictive obliterative bronchiolitis of different severity. *a)* Severe case where the hypo-attenuated areas can be easily detected. *b)* Mild case where the intensity differences are more subtle. *c)* Pre-symptomatic case where the identification of the areas of hypo-attenuation is difficult, even for an experienced observer

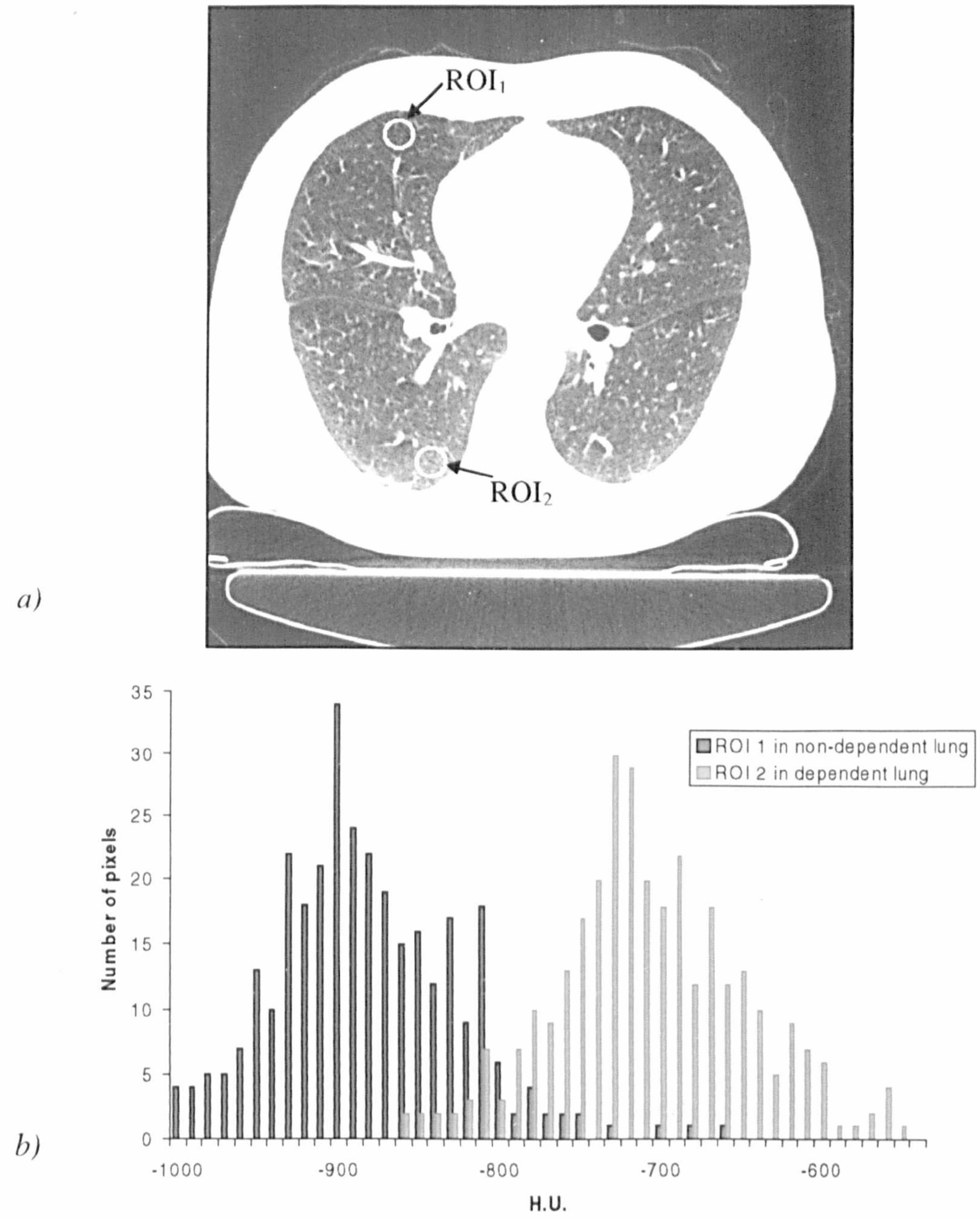


Figure 1.7. *a)* HRCT scan of a normal subject (window settings: level = -800 H.U., width = 1000 H.U.). Due to gravity, the dependent parts of the lungs appear denser than the non-dependent parts. The density differences in the lung parenchyma are not related to obstructive lung diseases. Two circular regions of interest (ROIs) are highlighted: ROI_1 in the non-dependent lung and ROI_2 in the dependent lung. *b)* Intensity histograms of the two ROIs. The mean density of ROI_1 is -877.9 H.U., whereas the mean density of ROI_2 is -703.5 H.U. The shift in CT density, due to gravity, between dependent and non-dependent lung appears clearly on the histograms.

1.2.2.2 Bronchial abnormalities

Mosaic attenuation pattern on thin-section CT scans may be caused by obstructive lung diseases, but may also be a manifestation of vascular occlusion or infiltration of the lungs [23]. A confident diagnosis of airways disease therefore involves the identification of ancillary signs, such as the presence of bronchial abnormalities [24,25]. Each bronchus runs alongside a pulmonary artery, the two structures comprising a bronchovascular bundle. The relative sizes of the bronchus, the bronchial wall and the adjacent vessel gives an indication of the state of the bronchi. In normal lungs, the diameter of the pulmonary arteries is equal to or slightly larger than the diameter of the bronchi they accompany. Along a short distance, bronchi will appear as cylindrical rather than tapering structures. They are filled with air and thus the luminal CT density is of the value of -1000 H.U. The bronchial wall consists of tissues of higher density (approximately 100 H.U.). On CT images, near perpendicular cross-sections of bronchi therefore appear as elliptical high-attenuation rings superimposed on the lung parenchyma, which has a low average CT density (typically lesser than -800 H.U.). Figure (1.8a) provides an example of HRCT scan of the lungs, with the conspicuous near-perpendicular cross-section of a major airway, which can be identified as a bright near-circular ring. Measuring bronchial dilatation and bronchial wall thickening provides valuable diagnostic information. However, the size of a bronchus is dependent on its position in the bronchial tree and there is no absolute reference of what the size of a normal bronchus should be. The diameter of the accompanying artery therefore provides a relative yardstick against which the bronchial diameter can be compared. The ratio of diameters of the two components of the bronchovascular bundle shown in Figure (1.8b) is close to 1, and is characteristic of a normal bronchus. Conversely, from the HRCT scan shown in Figure (1.8c), the ratio of bronchovascular diameters for the bundle highlighted in Figure (1.8d) is greater than 2, and is a sign of severe bronchial dilatation. It needs to be recognised, however, that the sign of bronchial dilatation does not infallibly reflect airways disease. For example, in patients with chronic occlusive vascular disease, dilatation of the airways is secondary to reduced perfusion and, in this context, does not reflect a primary airways abnormality [26]. Ideally, the bronchi used for visual assessment should be perpendicular to the plane of acquisition and appear circular. However, radiologists often need to evaluate the elliptical patterns of the many near-perpendicular bronchi, as shown in Figure (1.8c).

1.2.2.3 Parenchymal texture

The textural appearance of the parenchyma on CT provides essential information for the differential diagnosis of obstructive lung disease. On HRCT, the abnormal airspaces of emphysema resulting from destruction of alveolar walls of the distal airspaces produce areas of hypo-attenuation that often merge imperceptibly. The CT appearances of panlobular and centrilobular emphysema differ. Whereas panlobular emphysema results in areas of uniform hypo-attenuated lung, as shown in Figure (1.9a), centrilobular emphysema produces roughly circular areas of lung destruction that superficially resemble cysts, *i.e.*, air-containing spaces sometimes with a thin definable wall, as shown in a severe case in Figure (1.9b). The conspicuity of the pattern is dependent on the severity of the disease. For a mild case, as shown in Figure (1.9c), the characteristic textural appearance is still visible but less obvious than on a severe case. The air-trapping and under-perfusion caused by constrictive obliterative bronchiolitis result in homogeneous hypo-attenuated lung, as shown in Figure (1.9d). As a comparison, Figure (1.9e) shows the CT appearance of normal lung parenchyma. Figure (1.9) also shows enlarged views of the areas characteristic of the diseases.

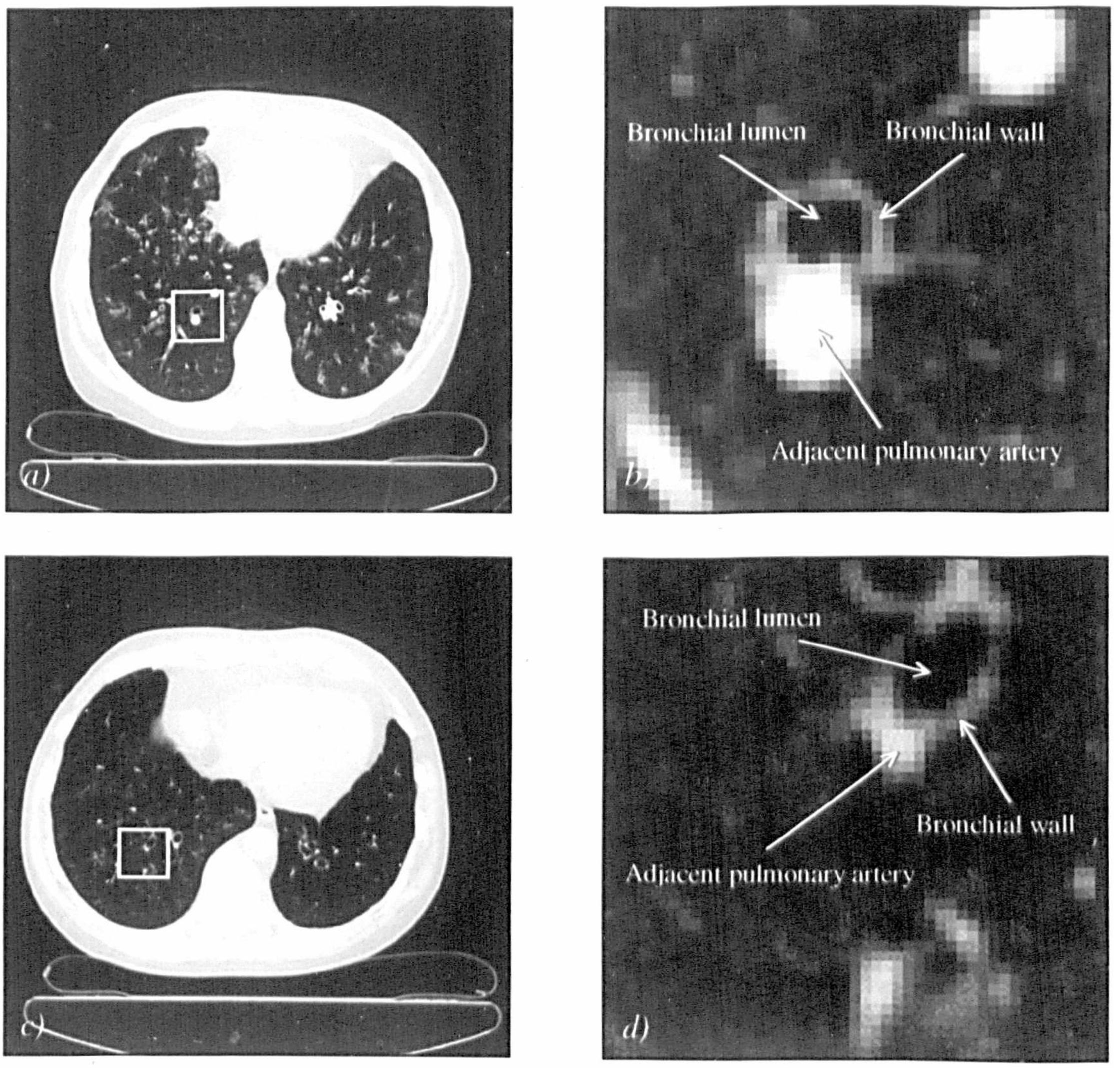


Figure 1.8. *a)* HRCT of the chest (window settings: level = -500 H.U., width = 1500 H.U.). A major bronchovascular bundle nearly perpendicular to the plane of acquisition is highlighted with a white square. *b)* Enlarged view of the bronchus highlighted in Figure (1.8a). The bronchial wall appears as a bright elliptical ring. Detecting bronchial abnormalities is achieved by comparing the radius and thickness of the bronchus to the radius of the adjacent pulmonary artery. Here, the radius of the bronchus is roughly equal to the radius of the artery, demonstrating normal bronchial morphology. Using this criterion, other bronchi on this section can be seen to be abnormally dilated. *c)* HRCT of the chest showing severe bronchial abnormalities. *d)* Enlarged view of the bronchus highlighted in Figure (1.8c). The bronchus does not run exactly perpendicular to the plane of acquisition, and appears elliptical. Nevertheless, it may be assessed by an experienced radiologist for diagnostic purposes. The ratio of bronchovascular diameters is greater than 2, indicating bronchiectasis.

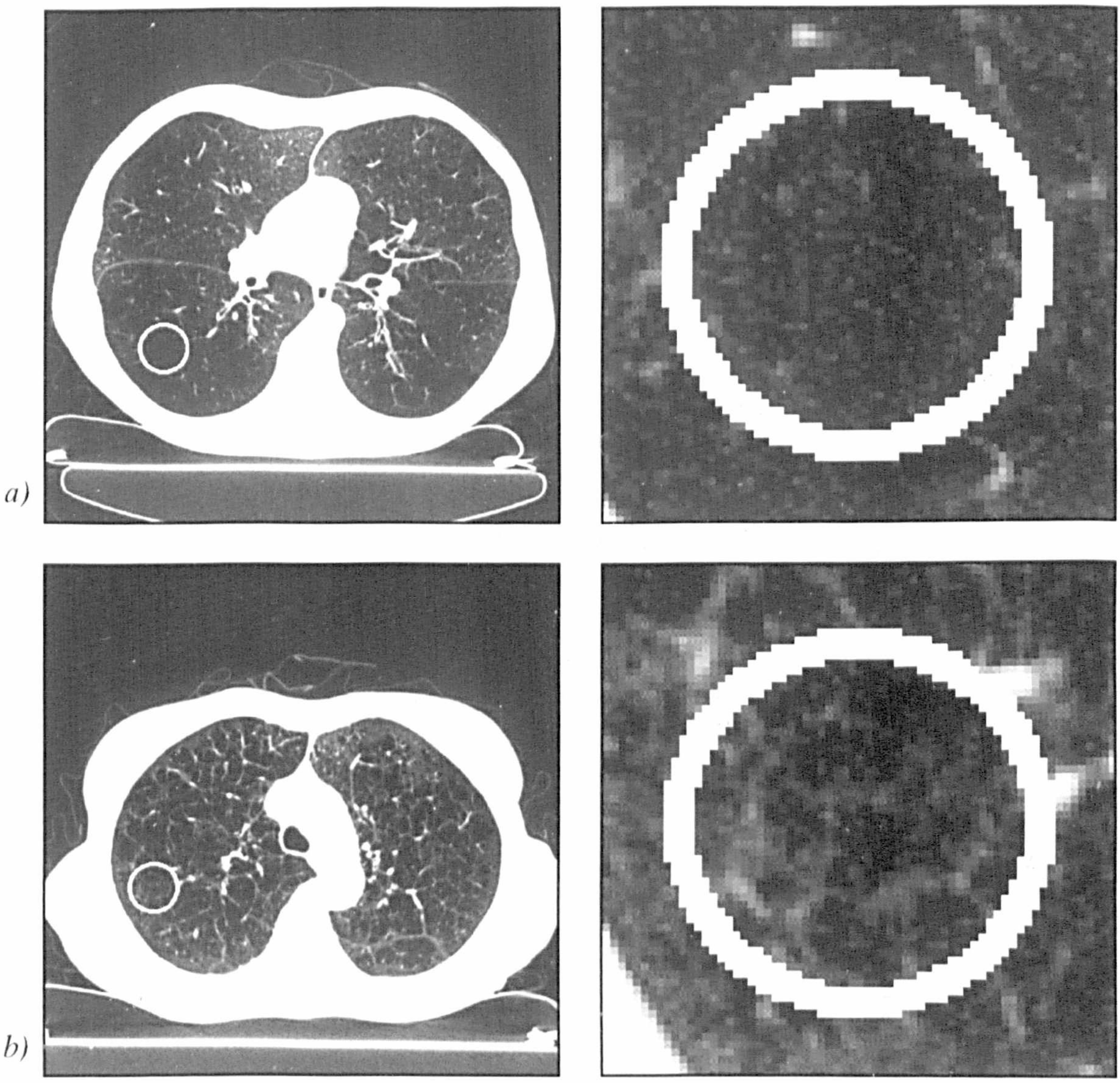


Figure 1.9. Five examples of high-resolution CT scans of the chest (window settings: level = -500 H.U., width = 1000 H.U.). On each image, the circle highlights a region of interest (ROI) that is characteristic of a particular condition. An enlarged view of the ROI is also shown on the right of each image. *a)* Panlobular emphysema: destruction of the lung parenchyma results in homogeneous hypo-attenuated areas. *b)* Severe centrilobular emphysema: the CT appearance of the disease superficially resembles air cysts.

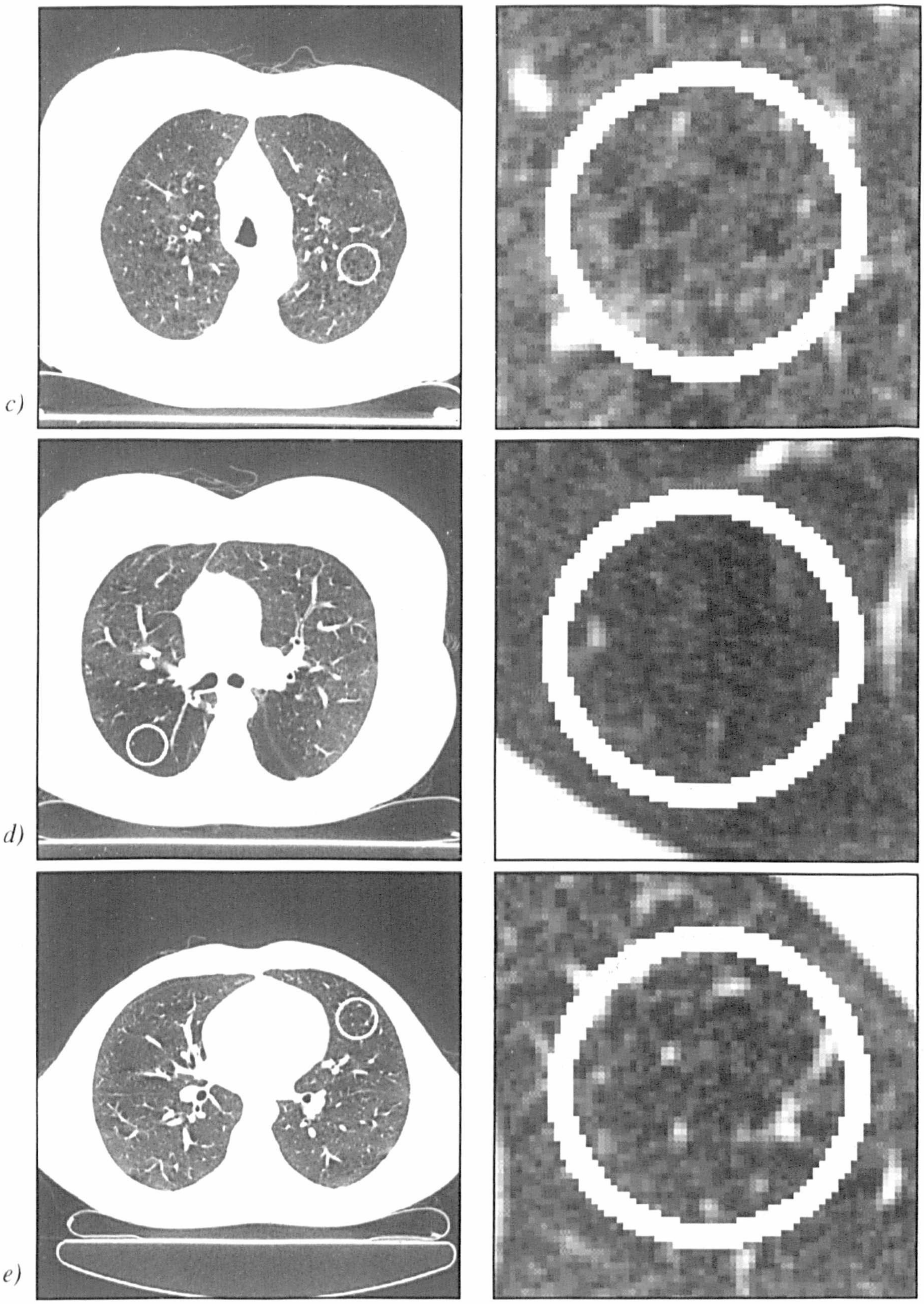


Figure 1.9 ***(continued).*** *c)* Mild centrilobular emphysema: the characteristic appearance of centrilobular emphysema is visible but less obvious then in Figure (1.9b). *d)* Constrictive obliterative bronchiolitis results in hypo-attenuated but homogeneous lung. *e)* Normal subject: the mean CT density of the lung parenchyma is higher than in areas affected by obstructive diseases.

1.3 RESEARCH OBJECTIVES

The purpose of this research is to design a system for the automated analysis of CT data of the lungs. It is aimed to help clinicians establish a confident diagnosis of obstructive lung diseases (emphysema, small airways disease) and increase the precision of investigation of structure/function relationships. This goal will be achieved by implementing low-level computer vision algorithms designed to reliably extract image features identified by radiologists as relevant to the diagnostic process. By combining these features through a probabilistic network incorporating *a priori* knowledge of the CT images of the diseased lungs and that allows for data fusion, we expect to achieve a comprehensive diagnostic conclusion. Given an HRCT image of the chest, the system presented in this thesis aims to:

- reliably extract from the raster image semi-quantitative measurements relative to obstructive airways diseases, including the extent of hypo-attenuated lung and the presence and severity of bronchial abnormalities,
- draw the attention of the clinician to regions of interest: pathologies (panlobular emphysema, centrilobular emphysema, constrictive obliterative bronchiolitis) associated with particular areas of the parenchyma, abnormal bronchi,
- on the basis of the output of the feature extractors presented, suggest a clinical diagnosis.

1.4 OUTLINE OF THE THESIS

The following chapters of this thesis are organised as follows. Chapter 2 presents a review of the image processing techniques related to the CT features of obstructive lung diseases. A brief survey of the existing methods tackling the issues addressed in this thesis is presented. This is followed by a general review of automated decision-support in medical imaging. It describes the problems relative to expert systems applied to medical diagnosis, with a particular emphasis on the formalisms employed to model the probabilistic nature of the diagnostic process.

In Chapter 3, we present a method for the automated segmentation of the main anatomical structures on CT of the chest based on thresholding, mathematical morphology and region-growing. To correct for the gravity-dependent density gradient associated with lung CT, a two-staged image processing technique is introduced. Based on intensity histogram deconvolution and intensity histogram energy optimisation, the algorithm offers a general and non-parametric means of gradient-correction. This is followed by adaptive clustering for automatically detecting and quantifying the areas of hypo-attenuated lung. Initial validation of the technique was achieved by creating a CT lung phantom. A patient study was also undertaken to measure the correlation of the technique with pulmonary function tests as well as subjective assessment by two radiologists.

Chapter 4 details an algorithm for the automated detection and measurement of bronchial abnormalities on CT. Based on the geometrical properties of elliptical patterns in local polar co-ordinates, a novel ERS transform is introduced for identifying the bronchi near-perpendicular to the plane of acquisition. By comparing the size of each bronchus to the size of the adjacent pulmonary vessel, the algorithm provides semi-quantitative measurements of bronchial dilatation and bronchial wall thickening.

A classifier for the differentiation of three types of hypo-attenuated lung (centrilobular emphysema, panlobular emphysema, constrictive obliterative bronchiolitis) and normal parenchyma is presented in Chapter 5. Visual characteristics of each condition are encapsulated in the form of textural descriptors. Supervised learning from a set of regions of interests chosen by an experienced observer allowed for the implementation of a Bayesian classifier, which was subsequently validated with a set of 220 CT images acquired from 44 patients.

In Chapter 6, we present how data fusion with a Bayesian network allows for the combination of the different feature extractors presented. A probabilistic model is employed to express the relationships between salient image features, as extracted by the low-level algorithms, for reaching diagnostic conclusion. This high-level clinical knowledge is encapsulated through supervised learning. The numerical expression of the causal relationships, represented by matrices of conditional probabilities, is derived from a training set of images assessed by an experienced observer. Performance of the Bayesian network was validated with patient data.

Finally, Chapter 7 summarises the main contributions of the work presented in this thesis and outlines potential future developments.

CHAPTER 2

CT LUNG IMAGE UNDERSTANDING

2.1 ISSUES IN CT IMAGE UNDERSTANDING

Clinical diagnosis from CT lung images is a subjective process. This approach can entail considerable inter- and intra-observer variability [10-12]. This leads to lack of precision in both diagnosis and any investigation of structure/function relationships. For instance, assessment of the natural or treated history of emphysema requires reproducible measurements over time of the extent of diseased lung, which cannot always be reliably obtained by subjective assessment. Some techniques have been proposed for enhanced visualisation of lung CT images [12,27,28], but detection and quantification of the findings remain largely subjective. The development of dedicated image processing algorithms for the reproducible assessment of CT lung images would alleviate this problem.

There are, however, several difficulties associated with the automated processing of CT images. CT numbers, as absolute values characteristic of particular tissues, have been demonstrated to be unreliable [29]. This is due to the fact that CT numbers are normally dependent on the scanner used for data acquisition. For each scanner, regular calibration is needed to ensure the reproducibility of the measurements. Partial volume effect also introduces problems by producing CT numbers that do not reflect the density of the structures represented within a voxel. Furthermore, the variability inherent to biological systems also represents major challenges. For instance, the CT density of hypo-attenuated lung in one patient may be equal to the CT density of normal lung in another. Therefore, CT image processing algorithms cannot always rely on absolute values for making meaningful measurements.

For achieving a comprehensive assessment of the lungs there are several inter-related image features that must be taken into consideration. On a CT image of the lungs, an experienced observer takes into account the size of the pulmonary vessels, the texture and the density of the lung parenchyma, and the morphology of the airways. Assimilation of the presence and distribution of these findings allows a radiologist to diagnose a specific or generic lung disease. It is therefore necessary to implement different automated feature extractors to represent each of these visual processes. It should be emphasised that it is important that multiple extractors are employed to complement each other. Each finding considered on its own is often ambiguous and does not provide enough information to reach a diagnostic conclusion. As an example, Figure (2.1a) illustrates a cross-section of the lungs acquired with HRCT. Automated segmentation of the major anatomical structures, with a technique presented in this thesis, suggests that blood perfusion might be abnormally low in some areas. In fact, the macroscopic pulmonary vessels as shown in red in Figure (2.1b) are abnormally sparse, especially in the left lung (shown on the right of the image). This CT finding is corroborated by the automated segmentation of the areas of hypo-attenuation, the "dark patches" within the lung parenchyma. The classifier based on CT densities presented in Chapter 3 reveals the inhomogeneities referred to as mosaic attenuation pattern [30]. The areas classified as under-attenuated are displayed in dark blue in Figure (2.1c) and can be consistently quantified for the purpose of follow-up studies of the extent of the disease.

The finding of mosaic attenuation may be ascribable to vascular or airways disease, and the analysis of parenchymal CT densities offers no specific means of distinguishing between the two types of pathologies. However, by including another finding, such as the state of the bronchi on the image, it is possible to resolve this ambiguity [23]. The major perpendicular bronchovascular bundles automatically detected by ERS transform, presented in Chapter 4, are shown in Figure (2.1d) [31]. Each bronchial lumen is shown as light blue, bronchial wall as dark blue and adjacent pulmonary artery as red. The ratios of broncho-arterial diameters are significantly greater than one, demonstrating abnormal bronchial dilatation [32]. It is therefore possible to conclude that the mosaic attenuation pattern identified in Figure (2.1c) is likely to be a manifestation of airways disease rather than vascular disease.

2.2 EXTRACTION OF SALIENT VISUAL FEATURES ON LUNG CT

The automated detection of obstructive lung diseases on CT requires the implementation of specific image feature extractors. These extractors must provide a numerical representation of the salient visual features presented in Chapter 1, and are of diagnostic value for the assessment of airways diseases. The following aspects must be included:

- correction of the density gradient and segmentation of the areas of hypo-attenuation of the parenchyma,
- detection of the major bronchi and measurement of bronchial abnormalities,
- texture analysis of the parenchyma.

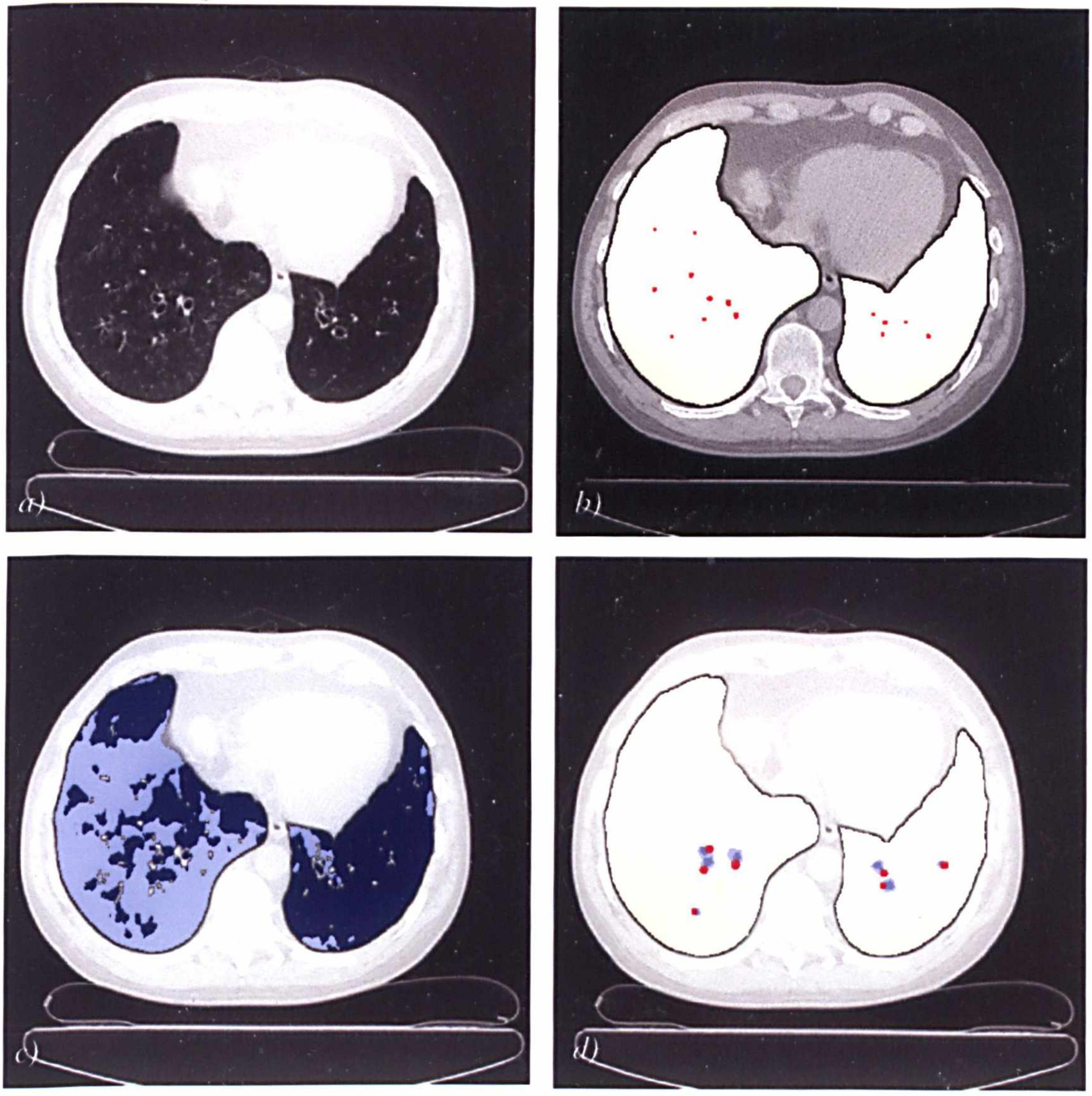


Figure 2.1. Example of the application of several features extractors to an image, in order to reach diagnostic conclusion. *a)* Cross-sectional image of the chest acquired with High-Resolution Computed Tomography (HRCT). *b)* Automated segmentation of the macroscopic pulmonary vessels, shown in red. The scarcity of vessels, particularly in the left lung (shown on the *right* of the image) suggests that perfusion is abnormally low. *c)* The automated segmentation of the areas of hypo-attenuation of the lung parenchyma (shown in dark blue) supports this hypothesis. The classifier based on CT densities reveals local inhomogeneities, referred to as mosaic attenuation pattern, and identifies 49 % of the lungs as hypo-attenuated. This pattern may be caused by an airways or vascular disease. *d)* Automated detection of the major perpendicular bronchovascular bundles. Each bronchus is displayed as blue and its accompanying pulmonary artery as red. The abnormally high ratio of bronchial/arterial diameters is evidence of severe bronchial dilatation and supports the diagnostic hypothesis of an airways disease as opposed to a vascular cause to the mosaic attenuation pattern.

2.2.1 Gradient-correction

As described in Chapter 1, CT lung images are usually affected by a linear density gradient due to gravity. The density variations this gradient generates have little impact on visual assessment by experienced radiologists, who are able to recognise and discount this phenomenon. However, this "overlay" may prove problematic for inexperienced observers, and affect the performance of automatic segmentation techniques, which assume homogeneity of density within each class. In order to achieve automated segmentation of the hypo-attenuated areas of the parenchyma, it is therefore necessary to implement a method for the correction of density non-uniformity. There are no reports in the literature of such a technique being applied to CT images. This may be due to the fact that the gradient observed on lung CT scans cannot be ascribed to a general limitation of the imaging modality, but to the normal physiology of the lungs. However, algorithms for the correction of intensity non-uniformity have been proposed for Magnetic Resonance (MR) images. Indeed, the physical constraints of data acquisition with MR are such that a multiplicative non-uniformity field affects most MR images. In addition to the modified acquisition protocols that cannot be adapted to other imaging modalities [33,34], several data-driven post-processing methods have been proposed. The most popular approach, *homomorphic filtering*, assumes that the frequency content of the non-uniformity field is lower than that of the anatomy [35-37]. To eliminate the low frequency components, *tissue intensity based models* [38-40] make assumptions about the value of certain classes of intensity on the image, in order to estimate the uniformity field that affects them. The main limitation of this approach is the need to determine a correct tissue intensity model for every pulse sequence (specific MR acquisition protocol) and perhaps every scanner. An alternative assumption is that of *spatial homogeneity*. Based on *LCJ* (an acronym for Liou, Chiu and Jain) segmentation of the data [41], a smooth field can be globally fit allowing for a constant offset in each region to reflect its tissue type [42]. However, the technique is heavily dependent on successful segmentation and cannot be used prospectively for the identification of anatomical structures. A third type of modelling is that of a *field model* [43]. It relies on the assumption that non-uniformity blurs the histogram of the data in a way that it can be identified, quantified and removed. The method described in Chapter 3 is based on this assumption.

2.2.2 Segmentation of hypo-attenuated lung

In spite of substantial inter- and intra-observer variability in the subjective assessment of mosaic attenuation [10-12,28,44], methods for the automated detection of hypo-attenuated lung have been confined to relatively crude studies of density mask quantification of patients with emphysema [45-47]. However, image segmentation, the process of separating the components of an image into subsets that correspond to the physical objects in the scene, has long been a field of investigation in image processing. Existing techniques can be divided into two major categories: edge-based approaches and region-based approaches. Edge-based techniques detect local discontinuities before grouping them into connected boundaries. They make use of edge detectors such as the Roberts and Sobel detector or the Laplacian of the Gaussian. The Roberts and Sobel operator, a simple discrete implementation of the computation of the image first derivatives, has a low computational cost since it is based on convolution with a small mask (typically 3 by 3 pixels) containing only integer values. However, it only handles images satisfactorily which have low noise and little texture. Noise-sensitivity can be minimised by smoothing the image prior to edge-detection, as in the case of the Laplacian of the Gaussian operator, based on the detection of zero-crossing of the second derivatives [48]. Applied to an image *f(x, y)*. the Laplacian of the Gaussian is given by the following expression:

$$\text{Laplacian of the Gaussian} = \left\{\nabla^2 N(x,y,\sigma)\right\} * f(x,y) \tag{2.1}$$

where '*' denotes convolution, and $N(x, y, \sigma)$ is the normal distribution defined by:

$$N(x,y,\sigma) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-(x^2+y^2)/2\sigma^2} \tag{2.2}$$

Computationally, zero-crossings can be identified more easily and with greater accuracy that the extrema of the first derivatives. Also, the Laplacian of the Gaussian operator is in keeping with the conclusions that Marr reached in the 1970s when studying the neurophysiological principles of human vision [49]: edge-detection by the human eye seems to be based on similar operations. Surface-fitting of local analytical functions, as with the Hueckel operator, also allows for a better management of image noise [50,51]. Canny proposed a procedure for the design of edge detection for arbitrary edge profiles based on the specification of detection and localisation criteria [52]. Complete contours can then be obtained from disconnected edges by performing edge following [53,54], possibly by dynamic programming, as proposed by Ballard [55]. High-level knowledge can be incorporated in the search of contours by employing *snakes*, or energy-minimising splines guided by external constraints forces that deform them toward salient image features such as edges. Snakes were introduced by Kass *et al* [56]. For a given parameterisation *f(s)* of a contour, the generic definition of its energy E_{snake} is given by the following expression:

$$E_{snake} = \int \left(E_i(s) + E_e(s) + E_c(s)\right) f(s)\,ds \tag{2.3}$$

where $E_i(s)$ represents the internal spline energy caused by the deformation of the contour, $E_e(s)$ measures the attraction of image features such as salient edges, and $E_c(s)$ measures external constraints such as prior knowledge of the expected geometry of the contour. An expression of the internal energy $E_i(s)$ that represents a constraint on the smoothness of the contour is given by:

$$E_i(s) = \alpha(s)\left|\frac{df}{ds}\right|^2 + \beta(s)\left|\frac{d^2 f}{ds^2}\right|^2 \tag{2.4}$$

where $\alpha(s)$ and $\beta(s)$ measure the elasticity and stiffness of the snake respectively. Snakes cannot identify contours without supervision, but they can refine precisely a contour, starting from an initial approximation given by a user, or devised by some boundary detector. The active contour model has been extensively used in medical imaging for delineating, for examples, cardiovascular borders on MR images [57,58].

However, edge-based approaches are poorly suited for the detection of hypo-attenuated lung on CT, the boundaries between adjacent areas being subtle and loosely marginated. Alternatively, region-based approaches identify the areas of an image that satisfy some uniformity criterion, which in turn give the boundaries [59,60]. Region-growing algorithms iteratively merge pixels satisfying a homogeneity criterion. Starting from a 'seed', connected neighbours are explored and incorporated to the segmented region if some Boolean function measuring local uniformity is made true by the resulting region. The method is inherently sequential which can be a limiting factor, and the result of the segmentation depends on the search strategy since merging two regions modifies the measured property. The method is also prone to unwanted region merge errors, and *split-and-merge* techniques, in which both region-splitting and merging processes may take place, were introduced to overcome this limitation [61]. The image is stored in a segmentation tree, referred to as quadtree, in which each node represents a square region and its corresponding search property. Adjacent nodes satisfying a uniformity criterion are merged and replaced by their parent. Similarly, non-uniform nodes are split into several children. Another simple technique of region segmentation is by thresholding [62,63]. The thresholds used can be determined by examining the histogram of intensity values [64]: if clear peaks are apparent, then the valleys between the peaks make appropriate cut-off values. By modelling intensity distributions by Gaussian distributions, discrimination between two areas denoted i and j can be performed if Fischer's criterion is verified:

$$\frac{|\mu_i - \mu_j|}{\sqrt{v_i^2 + v_j^2}} > \lambda \tag{2.5}$$

where μ_i and μ_j are the mean of regions i and j respectively, v_i and v_j represent their variance, and λ is a threshold. When clear peak separation cannot be derived, other techniques must be considered to alleviate the difficulty. For images with well-defined edges, a histogram can be created considering only pixels with a large Laplacian magnitude, *i.e.*, situated at each side of the boundaries [65]. The resulting histogram exhibits distinct peaks with similar number of pixels. Giving stronger weights to the pixels near boundaries is also the principle behind the works of Milgram *et al* [66], Panda *et al* [67], and Watanabe [68], that incorporate gradient information in the intensity histogram. Similarly, region-growing techniques can be combined with edge detection algorithms [69]. In the absence of well-defined edges, the threshold selection technique presented in Chapter 3 minimises a cost function between the original image and the resulting segmented image. Assuming that the parenchyma can be modelled by three classes of homogeneous density, the algorithm searches for the three class-specific CT numbers and the corresponding image segmentation that minimises the distance to the original image.

2.2.3 Detection of bronchial abnormalities

Thus far, there has been little reported work on the automated identification of bronchi on CT images of the lungs, but semi-automated techniques requiring user interaction have been described. Seneterre *et al* presented a method based on user-dependent greyscale thresholding [70]. Manual delineation of bronchial contours on enlarged CT images and the use of digital calipers to obtain accurate measurements were performed by McNamara *et al* [71] and Desai *et al* [72]. A similar method was employed by Schulte *et al* [73] for the assessment of extra-hepatic bile ducts on CT. The absence of a fully computerised technique may be due to the difficulty for automated methods to match the subjective assessment made by experienced radiologists. Indeed, it has been demonstrated that inter-observer agreement for the detection of bronchial wall thickening, for instance, is consistent. The kappa statistic, which measures agreement between two observers using a nominal scale, typically has a value of $\kappa = 0.6$ [74], which is often regarded as "clinically acceptable" (a value of 0 denoting random agreement and a value of 1 denoting perfect agreement [75,76]). To incorporate contextual information into a diagnostic aid system based firstly on the analysis of density differences, it is necessary to design an automated technique for the detection of bronchi and the assessment of bronchial abnormalities.

As described in Chapter 1, major near-perpendicular bronchi on CT images appear as elliptical rings. Methods in computer vision have been proposed for the recognition of such patterns. Template matching [77] is the most common approach, but searching a space with 6 degrees of freedom is computationally expensive. In fact, an elliptical ring is defined by 6 parameters: co-ordinates x_c and y_c, orientation θ, long axis r_1, short axis r_2, and thickness t, as illustrated in Figure (2.2).

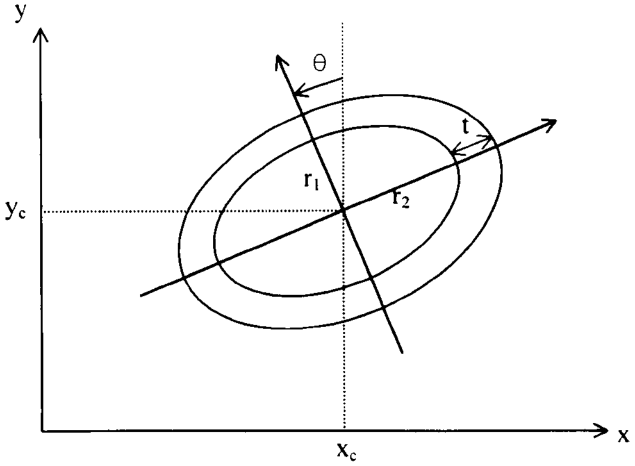


Figure 2.2. Representation of the parameters defining an elliptical ring: co-ordinates x_c and y_c, orientation θ, short axis r_1, long axis r_2, and thickness t. Six degrees of freedom make template matching and Hough transform for the detection of elliptical rings computationally expensive.

Another general method for the recognition of geometrical shapes is the Hough transform, which was proposed by Paul Hough in 1962 and later introduced to the image processing community by Rosenfeld [78]. The Hough transform is based on mapping feature points in the Cartesian space to points in a chosen parameter space. The parameter space is defined so as to represent a correct parameterisation of the shape to recognise. For instance, a Cartesian formulation of a straight line is:

$$y - \alpha x - \beta = 0 \tag{2.6}$$

where α and β are the parameters defining the line. Every feature point in an image (as identified by an edge detector, for instance) can be seen as belonging to a family of straight lines that can be represented in the (α, β) parameter space. To each feature point (x_i, y_i) in the Cartesian space corresponds one line in the parameter space, given by Equation (2.6). If the feature points are aligned, the lines in parameter space intersect at a point (α_0, β_0) that gives a parameterisation of the feature line in Cartesian space. However, the lines in parameter space may not intersect exactly, and the most suitable value (α_0, β_0) is found by building a histogram in parameter space and extracting its peak. A problem may arise when several lines are present in the Cartesian space. The secondary peaks, or side lobes, they create may combine their influence and result in false histogram peaks. Also, the choice of the parameterisation may introduce some bias in the histogram. Nevertheless, the Hough transform is a general method, suitable for a parallel implementation, and that can detect slightly deformed and partially occluded shapes [79]. It has been used in a wide variety of applications [80]. However, when dealing with patterns with a large degree of freedom (typically three or more), the requirements in term of storage capacity and computational power are considerable. For instance, the complexity of the Hough transform applied to the detection of elliptical rings is $0(N^4)$ [77]. Should each parameter be quantified on n levels, then the histogram requires n^6 points, *i.e.*, 1024 Mbytes of storage for $n=32$. A computational strategy to deal with a high number of parameters is to iteratively increase the resolution of the parameter space. This adaptive method looks for the histogram peak at low resolution then iteratively focuses on the vicinity of the peak for a more accurate estimate. However, the correct histogram peak may be hidden by combination of side lobes. The detection of ellipses with Hough transform can also be simplified by taking into account some geometrical properties allowing for the decomposition of the problem in two: estimation of the centre of the ellipse followed by estimation of the orientation and axes. The method presented in Chapter 4 identifies elliptical patterns in one-pass, at a lower cost than the Hough transform.

2.2.4 Texture analysis

The textural characteristics of the lung parenchyma on CT provide essential information for the differential diagnosis of obstructive and other diffuse lung diseases. Although there is no universally accepted definition, texture usually refers to a structure composed of a large number of similar elements or patterns, organised in a stochastic or deterministic fashion. Several characteristics are generally chosen to describe textures: *uniformity, density, coarseness, roughness, regularity, intensity* and *directionality*. Mathematical descriptors capturing textural features have been developed for a wide range of image understanding applications [77], such as the segmentation of satellite imagery [81]. Studies on human perception suggest that texture discrimination is based on the differentiation of second-order statistics [82]. Haralick describes how such characteristics can be extracted from grey-level co-occurrence matrices [83,84]. Co-occurrence matrices are based on the estimation of the second-order joint probability density functions $P(i,j \mid d, \theta)$, which represent the probability of going from grey level i to grey level j given the inter-sample spacing d and the direction θ. Other texture descriptors have been proposed: grey level run-length methods [85] measure the properties of linearly adjacent sets of pixels having the same grey level; fractal dimension corresponds closely to the intuitive notion of *roughness* [86]; Fourier spectrum analysis extracts the sum of energy contained in some defined regions of the spectrum, which are correlated with texture coarseness [87].

Application of texture analysis to lung imaging was tested as early as the 1970s [88]. There have been examples of automated quantification of emphysema on chest radiographs proposed by Buteau *et al* [89], but CT has been demonstrated to be a superior imaging modality for the detection of alterations of the lung parenchyma [15]. Uppaluri *et al* have shown that high sensitivity (94.8 % on average) could be achieved with a binary textural classifier distinguishing between emphysematous and normal lung on CT [90]. Other binary classifiers, were proposed for the textural differentiation of normal versus abnormal lung for cases of interstitial diseases by Katsuragawa *et al* [91] and ground-glass opacities by Heitmann *et al* [92] and Shimizu *et al* [93]. An example of textural classifier, proposed by Delorme *et al* [94] handling more than two classes showed limitations because it was not applied to segmented images. There are no reported examples of classifiers applied to segmented images, or designed to distinguish between different types of hypo-attenuated lung such as cases of constrictive obliterative bronchiolitis, panlobular emphysema and centrilobular emphysema, as well as normal lung. Chapter 5 presents explores the efficacy of such a classifier.

2.3 DECISION-SUPPORT IN MEDICAL IMAGING

The outputs of the image feature extractors presented in this thesis need to be combined in a high-level framework in order to reach diagnostic conclusion. An overview of the methods developed for the implementation of similar decision-support systems, in medicine in general, and in medical imaging in particular, is given in this section.

2.3.1 Medical expert systems

The first Artificial Intelligence systems for medical diagnosis started to emerge as early as the 1950s [95]. To model the relationships between *symptoms* and *diseases*, explicit domain knowledge was incorporated in expert systems. Such systems use inference strategies to apply knowledge to the available data (which is often noisy and incomplete), through heuristic reasoning. As with all expert systems, the design of medical diagnostic systems raises essential questions:

- how to find a suitable representation of the observed data
- how to define a suitable representation of the knowledge of the domain
- how to obtain measurements that could validate a diagnostic hypothesis
- how to handle inaccuracy and uncertainty of the observed data

Early systems adopted a sequential data-driven strategy to perform heuristic classification. Knowledge was encapsulated in production rules in the form of "*if* ... *then* ..." statements as in the MYCIN system [96]. Rules are a simple and modular tool, but their formalism is restrictive, and a great number of rules are needed to deal with real-life problems. Frames [97], another traditional AI technique, were also applied, as they are better suited to represent "descriptive knowledge" (an important issue in medical diagnosis), but revealed similar limitations to production rules. More complex inference strategies involving both data-driven and goal-driven strategies were then developed. Blackboard systems for instance, work through an iterative process of hypotheses and tests [98]. These different forms of knowledge representation and control strategies can be regarded as theoretically equivalent, offering different solutions along the power/generality trade-off. Because of the potential complexity of medical diagnostic, revealed by the limitations of the general problem solvers developed in the 1960s and 1970s, medical expert systems developed in recent years tend to focus on narrow domain-specific problems.

The investigation of medical diagnostic systems revealed three main specific problems. Firstly, knowledge engineering proved to be a task that is not trivial. The term *knowledge engineering* refers to the process during which the knowledge of domain experts is made explicit in order to be represented in an expert system. Problems arise even before any attempt to chose a computationally acceptable formalism. It is often difficult for experts to identify precisely the logical mechanisms they use, or to expound the unsaid assumptions that they rely on [99]. Medical practitioners find it hard to analyse their intellectual diagnostic process which is the product of both academic training and clinical experience. The cultural differences between clinicians and computer scientists also increase the risk of ambiguous communication. File and Dugard give the example of the elaboration of an expert system for the diagnosis of different thyroid disorders [100]. It appeared that domain knowledge was not correctly encapsulated because the knowledge engineers had wrongly equated the expert's term *diagnosis* with *outcome*. To overcome this type of pitfall, generic methodologies were developed for capturing expert knowledge and problem requirements within a standard framework which promotes unambiguous communication between the members of a project team. For instance, CommonKADS, a successor to the KADS technology, consists of a collection of structured methods for building knowledge-based systems. It produces systems in a manner which is both systematic and well documented [101]. Automated support, in the form of "knowledge engineering workbenches", can be useful when going through the very detailed steps of the methodology [102]. In the particular case of medical systems, the PROforma approach developed a formalism for capturing general medical knowledge and clinical procedures. It is both a *formal specification language* in the sense used by software engineers, to ensure clear, logical, and unambiguous design, and a *knowledge representation language* which seems natural to medical professionals [103,104].

Secondly, a source of difficulty in medical diagnosis systems is the common occurrence of multiple disorders. Many patients have concurrent, potentially inter-related, diseases. Not only does this increase the complexity of the diagnostic process by expanding the search space (combinations of hypotheses need to be tested), but also it introduces more noise in the observed data. Indeed, the symptoms of a disease may hide the symptoms of another condition. Deep causal reasoning attempts to overcome this problem by incorporating explicit representation of the mechanisms linking diseases to symptoms at different levels [105]. The ABEL system, for instance, employs a five-level pathophysiological model, from the clinical state of the patient down to biomedical understanding of the mechanisms of acid/base physiology [106]. The main limitation of this approach, however, is the absence of sufficiently detailed and accurate pathophysiological models in most medical domains.

Finally, expert systems need to model uncertainty, in order to cope with noise, incomplete and inaccurate data, and intrinsic variability. This is particularly important in a domain like medicine where unlike, say, electronics, knowledge cannot be modelled with causal relations. Early diagnostic systems relied on idiosyncratic empirical methods to model uncertainty, like the “Confirmation Theory” used in MYCIN. First attempts to use mathematically sound models relied on the theory of classical probabilities. Nevertheless, the restrictions for applying classical probabilities are strict: data independence, exhaustive set of mutually exclusive hypotheses. To model the lack of knowledge, and the gaining of knowledge through experimental results, Dempster-Shafer’s theory of evidence was also proposed [107]. However, the mathematical assumptions the theory of evidence relies on are as strict as for classical probabilities, and implementation incurs increased computational cost. A more straightforward implementation of a model of uncertainty can be achieved through fuzzy logic. Fuzzy logic allows for an intuitive mapping of ambiguous terms used by domain experts, like 'large' and 'small'. But there is no standard mathematical formulation of the fuzzy logic operators, and different definitions may yield to very different behaviours of an expert system. As opposed to these numerical methods, some authors suggested the use of a symbolic representation of uncertainty. Symbolic approaches keep track explicitly of the sources of uncertainty, and in some cases may allow recovery from errors [97]. The model of endorsements [108] stores the arguments in favour and against each hypothesis. But weighting endorsements in order to compare alternatives proved to be a difficult task.

All these difficulties stressed the limitations of traditional AI techniques applied to medical diagnosis. However, Bayesian probabilities have commonly been regarded as a suitable formulation with a sound mathematical justification, for the modelling of medical diagnosis. Conditional probabilities offer a good paradigm for representing the relationships between 'symptoms' and 'diseases'. However, it has long been considered that the implementation of a scheme based on Bayesian probabilities was impractical. The number of estimated joint probabilities required in a medical diagnostic problem was too great [109]. In the 1980s however, the work of Cooper and Pearl presented a new formulation of the theory, based on the propagation and updating of belief through a network of propositions [110]. In medical expert systems, nodes of the network represent the observed data or pathophysiological states and hypotheses, and a conditional probability is associated with each link between the nodes. This new approach has had a considerable impact in a wide range of AI applications. Bayesian networks are a general and manageable tool, with a strong theoretical justification, while allowing for deep causal reasoning. Therefore, they have become, since the late 1980s, a popular framework for medical expert systems such as the PATHFINDER system for the diagnosis of lymph-node diseases [111] or the DIAVAL system for echocardiography [112]. Existing programs developed with other formalisms such as Internist-1/QMR [113] or HEPAR [114] have been converted to Bayesian networks. Empirical studies showed that Bayesian networks performed better than other approaches to expert systems [115] while patient studies showed that satisfactory results could be obtained in a clinical practice. For instance, the MUNIN system for the diagnosis of neuromuscular disorders was demonstrated to perform "at the same level as an experienced neurophysiologist" [116]. A simplified chronology of the main trends in the design of medical decision-support systems is given in Figure (2.3).

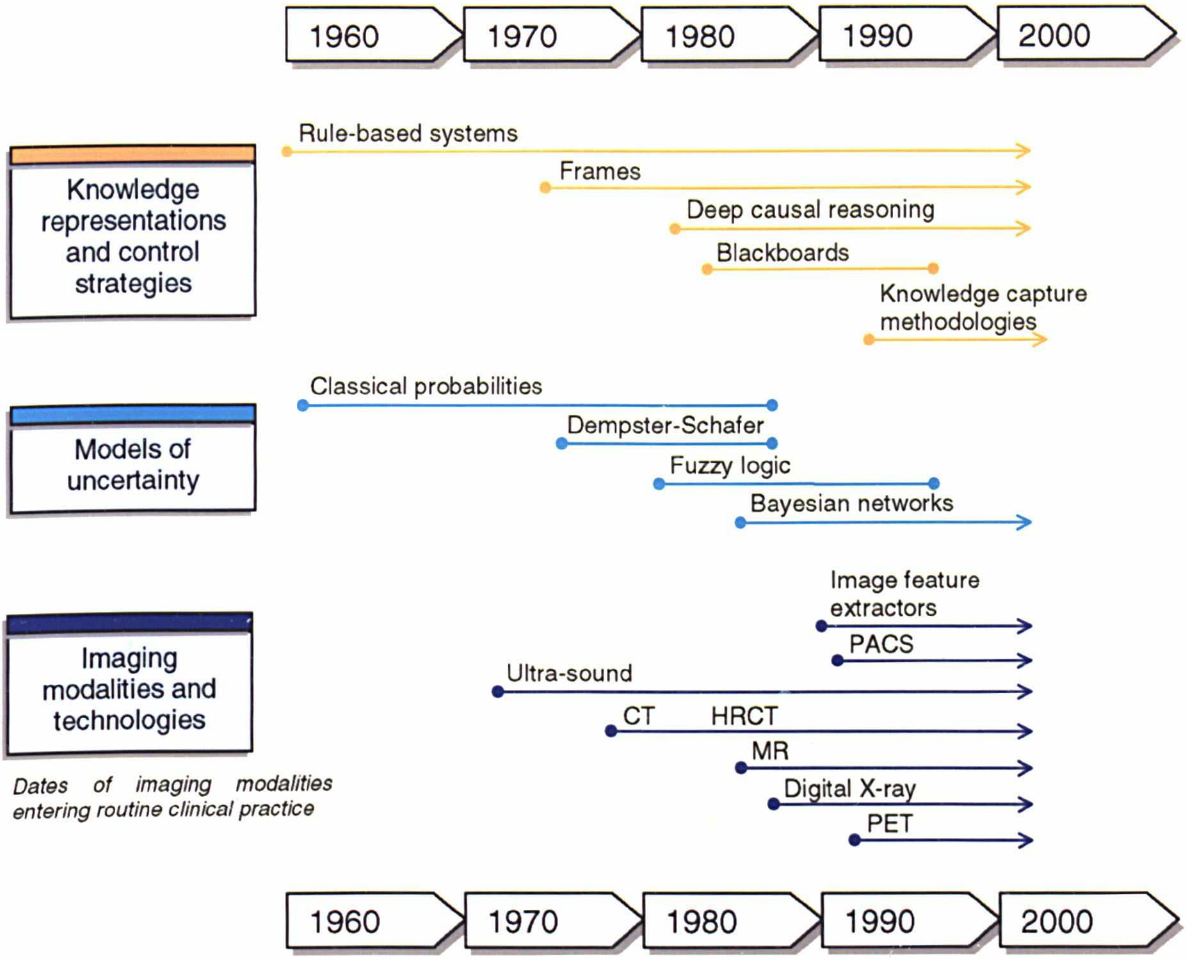


Figure 2.3. Chronology of trends in medical decision-support systems. The most popular forms of knowledge representations and control strategies are represented, as well as the formalisms employed to handle uncertainty. The introduction of the main digital imaging technologies is also shown.

2.3.2 Decision-support systems in medical imaging

Despite applications of AI techniques to a wide range of clinical activities (histopathology [117], infectious diseases [118], internal medicine [119], psychiatry

[120]), few expert systems have been implemented in the field of medical imaging. In most cases, expert systems in radiology rely on subjective findings established by a human observer: the Radiology Image Interpretation System (RIIS) for the diagnosis of focal bone abnormalities [121], the DIAVAL system for echocardiography [112], the Skeletal Dysplasias Diagnostician (SDD) [122]. Therefore, the general methods of AI in medicine can be applied. Unsurprisingly, it has been shown on such a system (the production rules-based RIIS) that the quality of the computerised image interpretation is user-dependent: a less valid conclusion will be derived from the findings of an inexperienced observer.

One of the reasons for the scarcity of imaging expert systems is the difficulty inherent in high-level vision. The data acquired from medical scanners can be noisy and ambiguous. Segmentation into areas representing anatomical or functional units is often a complex and ill-defined problem [123]. Dealing with this uncertain, imprecise information, known as *evidential information* [124], increases the complexity of expert systems. Furthermore, each imaging modality and each type of pathology requires the development of dedicated low-level feature extractors. Although standard computer vision techniques may be used (*e.g.* template matching, Hough transform, region growing, etc) *ad hoc* tools need to be designed for most applications. These difficulties, combined with the computational cost associated with computer vision applications have prevented so far the development of fully automated imaging decision-support systems.

Nevertheless, the potential benefits make it tempting to aim at designing expert systems using the digital images provided by the various modalities. Relying solely on the objective acquired data allows for greater reproducibility of the diagnosis aid. By producing a symbolic representation of the images, imaging expert systems also enable new applications like the automated generation of anatomical and functional atlases [125], or content-driven image database retrieval [126]. Such applications are becoming increasingly relevant, with the rapid generalisation of digital archiving and networking of medical images through Picture Archiving and Communication Systems (PACS). Automated indexing of the large image databases currently available to most healthcare institutions would allow for a more efficient use of them. Content-driven image retrieval would be useful, for instance, in research and epidemiological studies that examine large populations with a given pathology. Also, the abundance of information derived from the recent multiplication of imaging modalities (Computed Tomography (CT), Magnetic Resonance (MR), Positron Emission Tomography (PET), digital X-ray, ultra-sound scans, as shown in Figure (2.3)) highlights the need for imaging decision-support systems [127]. Such programs should improve diagnostic accuracy and overall reproducibility by providing a second opinion: they should compute objective measurements of normal and abnormal patterns and draw the attention of radiologists to areas of potential abnormalities [128]. For instance, it was demonstrated by studying eye movements that 30 % of missed lung nodules on chest radiographs were due to the clinician not looking at the area of the lesion on the image [129]. Imaging expert systems should help decrease this number of false negatives. These potential benefits are summarised in Figure (2.4) that represents the schematic architecture of a medical imaging expert system.

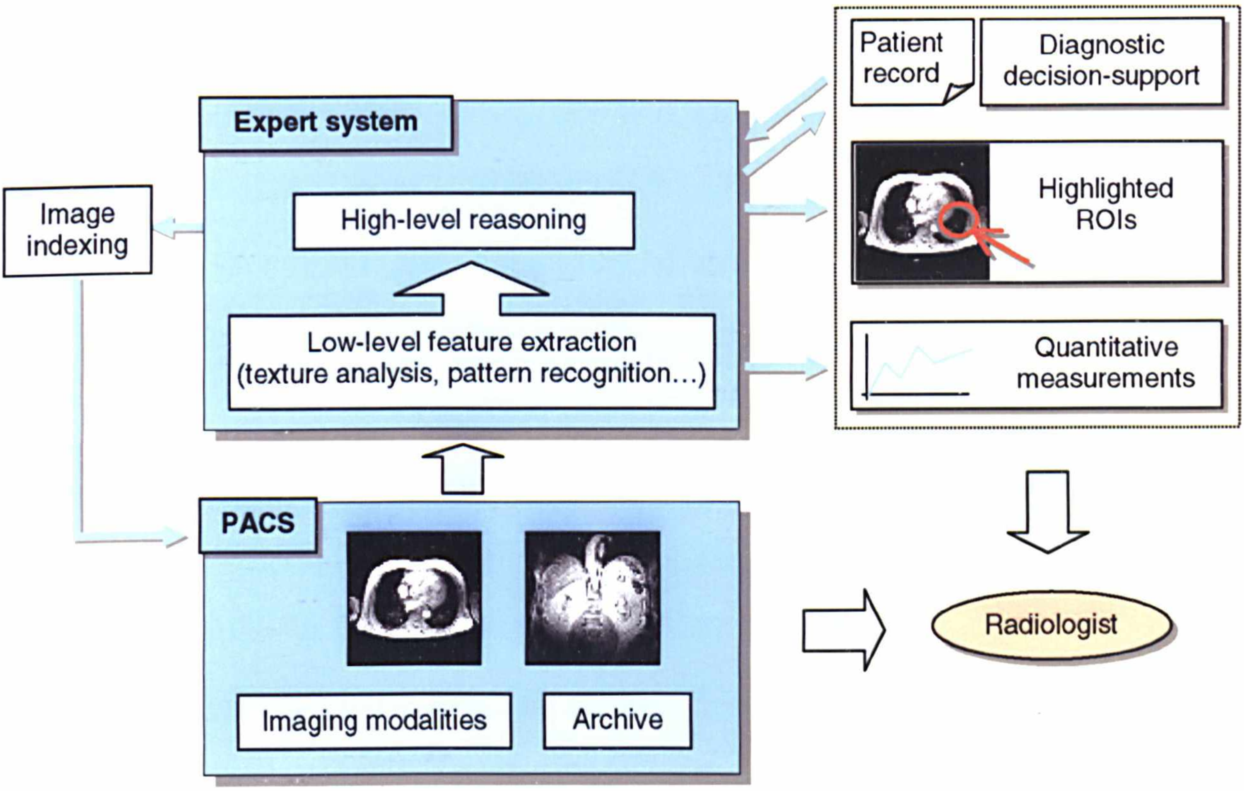


Figure 2.4. Architecture of a medical imaging expert system. Images recently acquired with all modalities as well as archived images are made available through PACS. Low-level features extractors applied to the images provide the radiologist with objective measurements. They also serve as an input for the high-level reasoning process that incorporates general clinical knowledge and data derived from the patient record. As an output, the expert system highlights the regions of potential abnormalities, suggests diagnostic hypotheses, and assign to the images in the PACS database a symbolic representation of their contents.

It is becoming feasible to investigate the issue now that the required technology behind it has become readily available. Medical images are now commonly stored in a digital inter-connected format. DICOM has emerged as an international standard, recognised by most hardware and software manufacturers, for the storage and transmission of medical images acquired with all modalities [130]. Network infrastructures and the Internet allow for the joint development of technologies and the sharing of clinical image databases through different healthcare institutions across the globe. The past decade has seen dramatic hardware improvements, the exponential growth of memory and storage size, processor clock speed, and bandwidth. The computing power needed to deal with the bulk of data encountered in vision systems is no longer prohibitively expensive.

Consequently, in recent years, there has been increased interest in applying computer vision in medicine. Low-level feature extractors applied to medical images have been extensively researched and tested. For instance, linear filtering and local adaptive thresholding for the automated detection of clustered micro-calcifications on mammograms were demonstrated to improve the radiologist's accuracy [131]. Other applications, presented by Doi *et al* [131], include the computerised detection of pulmonary nodules, the computerised analysis of heart size and the characterisation of interstitial disease on chest radiographs, or the automated tracking of vessels on angiographic images. Texture analysis was applied to ultrasound scans, radiographs and CT images proved successful in some instances: a regional accuracy of 97.3 % was reported for the computerised detection of emphysema on CT images [90]. However, in complex clinical situations, such approaches revealed their limitations. The accuracy of a texture analyser designed for the detection a whole range of CT findings, and not just two classes of images, drops to 70 % [94]. To overcome this difficulty, the need to incorporate clinical knowledge encapsulated in high-level rules was emphasised [92]. The low-level feature extractors developed in recent years showed promising results, but their limitations demonstrated that they needed to be combined in future imaging decision-support systems.

Multiplying the inputs increases the quantity of information the expert system can use to support diagnostic hypotheses. For instance, both texture analysis and geometric pattern measurements can be used for the detection of lung infiltrates on chest radiographs [91]. However, this raises the issue of combining classifiers and resolving potential conflicts. The different low-level feature extractors may use different formalisms: some may output numerical results (the size of the bronchi, for instance), some may output symbolic results ("density of the lung is abnormally low"), possibly weighted by different measures of confidence. Kittler *et al* presented a general theoretical framework based on Bayesian probabilities for combining classifiers [132] which is appropriate for this task. Theoretical considerations about Bayesian networks, namely that instantiated data must be conditionally independent of its parents, led to a new methodological approach to building decision structures for vision [133]. These results provided the basis for the choice of Bayesian formalism for the implementation of the diagnostic aid system presented in Chapter 6.

CHAPTER 3

IMAGE SEGMENTATION AND GRADIENT-CORRECTION

3.1 INTRODUCTION

A key feature of obstructive lung disease on CT images is the presence of hypo-attenuated areas in the parenchyma. These areas may be poorly marginated and regional density differences between normal and diseased lung parenchyma are often small and may be perturbed by the superimposed normal gravity-dependent density gradient. Even for experienced observers, inter- and intra-observer variability in the assessment of mosaic attenuation can be substantial [10-12,28,44]. In most situations, segmentation of the diseased lung parenchyma cannot be achieved by thresholding due to the subtlety of density differences. Furthermore, the definition of the threshold requires subjective user interaction and can be patient dependent [29]. The superimposition of a linear vertical density gradient, due to gravity effect as explained in Chapter 2, also introduces density differences in the lung parenchyma, not related to the pathological pattern of mosaic attenuation. This phenomenon can sometimes lead to misdiagnosis by inexperienced radiologists. An illustration of this density gradient in normal lungs is given in Figure (3.1), where Figure (3.1a) shows posterior lungs appearing denser than anterior lungs due to the gravity effect. Figure (3.1b) illustrates the average density of the parenchyma, as measured along the vertical axis of the image. A linear density increase is evident and needs to be corrected for prior to identifying and quantifying areas of hypo-attenuation.

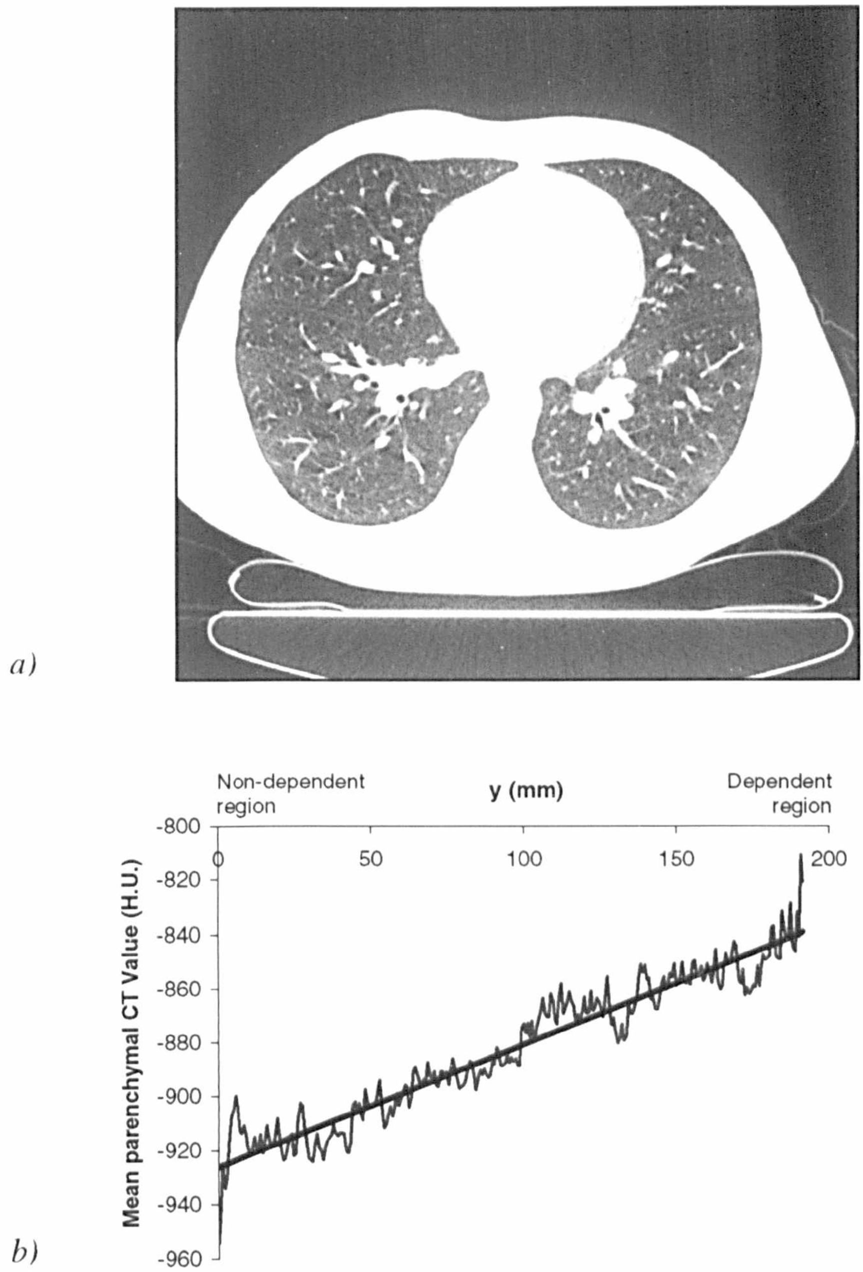


Figure 3.1. *a)* Example of CT scan of normal lungs with visible vertical density gradient (window settings: level = -800 H.U./ width = 1000 H.U.). Due to gravity, the posterior part of the lung appears denser than the anterior part. These density differences in the parenchyma are not related to small airways disease. *b)* Measure of the mean density of the parenchyma along horizontal lines, from anterior lungs to posterior lungs. A density increase can be observed, and modelled by a linear gradient. In this case, the slope of the regression is 0.46 H.U./mm.

The purpose of this chapter is to develop and evaluate a fully automated method for the reproducible quantification of the hypo-attenuated areas of the lung parenchyma. To achieve this goal, the following three major issues are addressed:

- automatic segmentation of the anatomical structures in order to isolate pixels of the lung parenchyma,
- correction of the density gradient caused by gravity and/or atelectasis of the lungs,
- determination of the hypo-attenuated areas of the lung parenchyma through adaptive clustering.

Figure (3.2) schematically outlines the main structure of the technique. The proposed method was validated with a CT lung phantom, and a comparative study was undertaken in which the results of the automated algorithm were correlated with measurements made by two experienced radiologists on 23 normal subjects and patients with constrictive obliterative bronchiolitis.

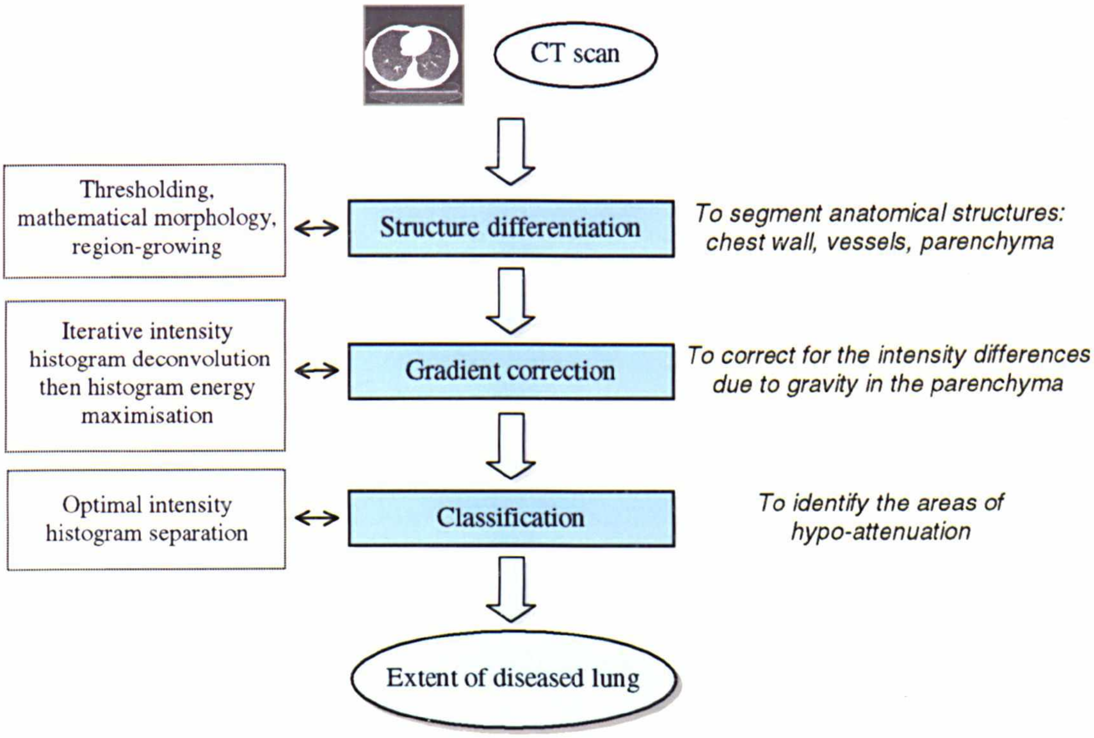


Figure 3.2. General architecture of the proposed algorithm. First, structure differentiation is applied to the acquired CT lung image. Pixels belonging to the lung parenchyma are automatically distinguished from the chest wall and macroscopic pulmonary vessels. Then gradient-correction is applied to compensate for the intensity differences in the lung parenchyma due to gravity. Finally, a classification algorithm is applied that identifies and measures the areas of hypo-attenuation corresponding to diseased lung. The boxes on the left indicate the methods and mathematical tools used at each step of the algorithm.

3.2 METHODS

3.2.1 Segmentation of anatomical structures

Prior to delineating hypo-attenuated areas of the lung, pixels belonging to the lung parenchyma have to be identified. Because of superimposition of pulmonary vessels and partial volume effect, it is difficult in practice to rely on simple thresholding. To overcome this problem, a hybrid approach was adopted in this study. The chest wall, which has a very high density value is thresholded out by using an empirically selected CT value of $T_{chest} = -250$ H.U. Because of the large difference in CT attenuation between the soft tissue density of the chest wall (soft tissue $\approx$ 0-100 H.U.) and lung parenchyma ($\approx$ -800 H.U.), a threshold selected empirically is adequate in spite of variation between subjects. To account for the partial volume effect at the interface between the lung parenchyma and the chest wall, a morphological dilation operator with a square flat-top structuring element with a size of 5 pixels is applied to remove pixels located within these regions [134]. This creates a regional mask through which the largest connected component is extracted. This represents the chest wall, and the two inner regions delineated the boundary of the lung parenchyma. The steps described above are illustrated in Figure (3.3), in which (a) is the original image, (b) is the image after thresholding, (c) is the result after dilation, and (d) is the derived mask for the chest wall. Based on the inner regions of (d), the areas covered by the lungs can be extracted from the original image.

For the correction of the density gradient, it is also necessary to remove the superimposed macroscopic blood vessels. Given the typical CT value of aerated lung of -800 H.U. [135], an empirically selected thresholding value of $T_{vessels} = -750$ H.U.

combined with a morphological dilation operator is used to create a vessel mask for removing the superimposed blood vessels. A 3-pixel flat-top square structuring element is used for the morphological operator. Figure (3.4) illustrates the intermediate results involved in this process, where (a) is the initial thresholded blood vessels, (b) the vessel mask created by morphological dilation, and (c) the final extracted lung parenchyma free of other structures. Finally, morphological erosion is applied to eliminate blood vessels with a size of magnitude of one pixel. Thus, the set of pixels selected for classification does not represent any anatomical detail with a size matching the spatial resolution of the scanner, but only the average CT attenuation of the finest lung structures.

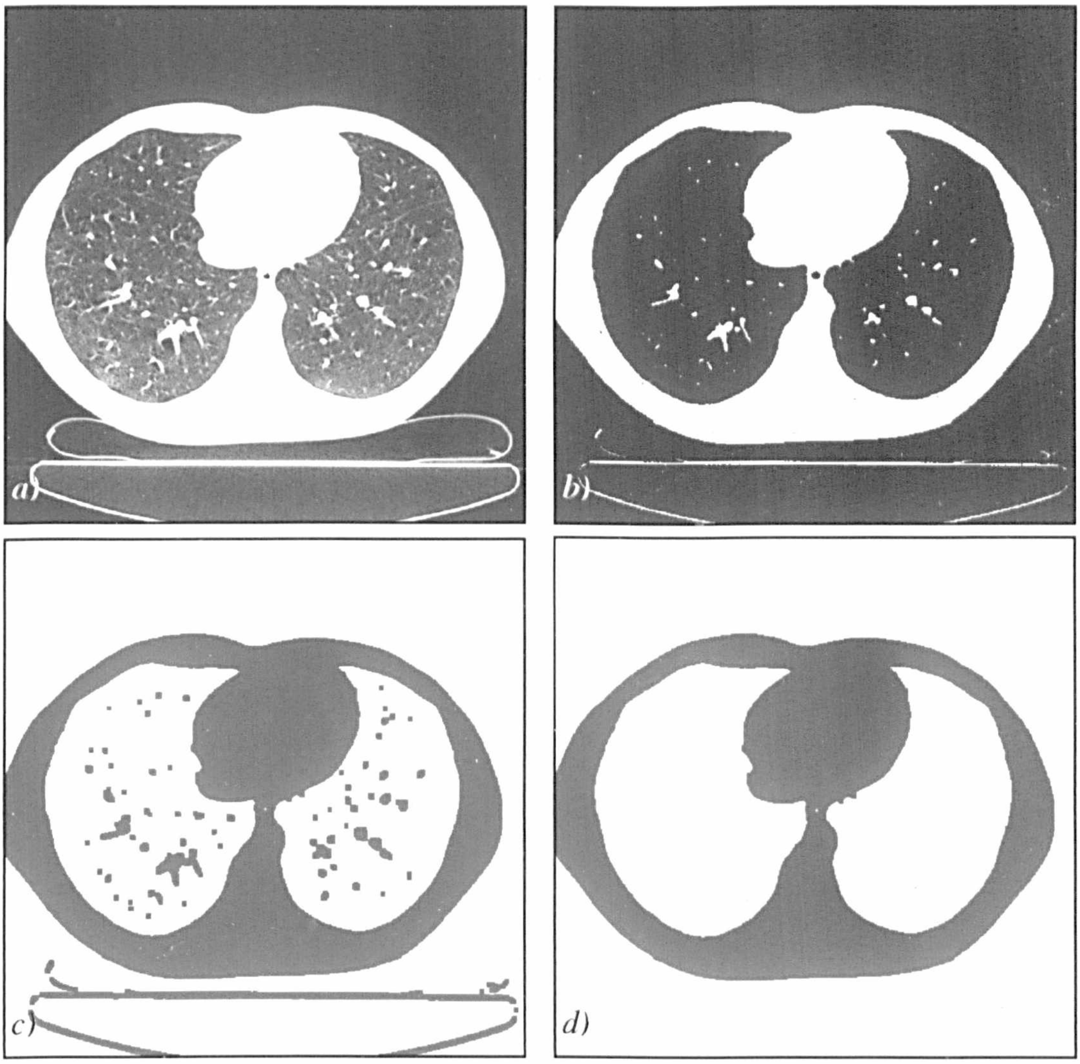


Figure 3.3. Segmentation of the chest wall. *a)* Original image. *b)* Thresholding at T_{chest} retains the chest and some macroscopic pulmonary vessels *c)* After dilation, a mask is created. *d)* The largest connected component gives the mask for the chest.

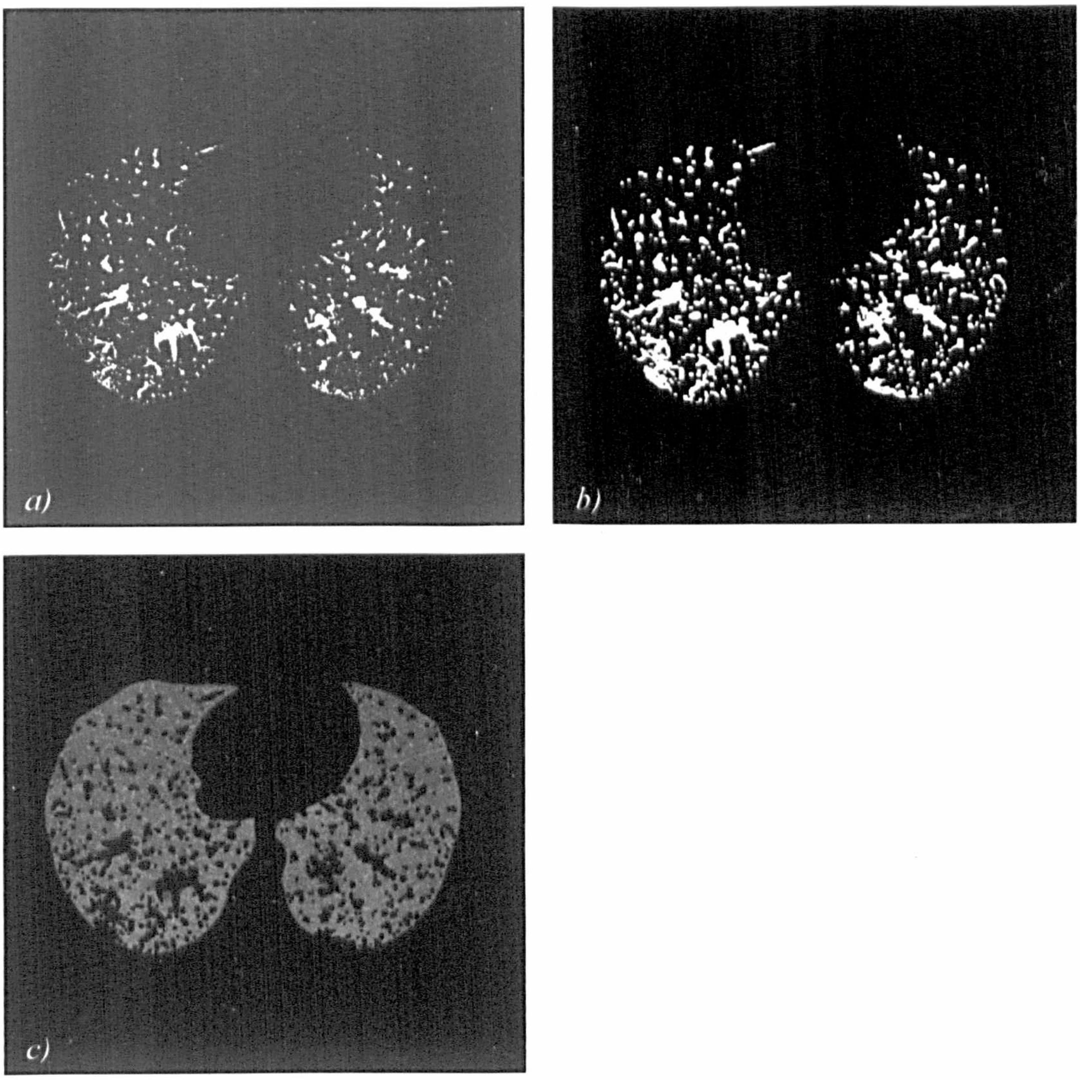


Figure 3.4. The steps involved in the segmentation of the macroscopic blood vessels. *a)* Thresholding at T_{vessels}. *b)* Applying morphological dilation creates a mask that takes partial volume effect into account. *c)* Lung parenchyma without any other anatomical structures superimposed.

3.2.2 Gradient-correction by histogram deconvolution

To correct for the density gradient due to the gravity effect, an algorithm based on intensity distribution is then applied to pixels of the lung parenchyma. It relies on the assumption that an idealised CT lung image is superimposed with a linear vertical density gradient. Statistically, the effect of this linear gradient broadens the intensity histogram of the image. By iteratively deconvoluting the intensity histogram with a smoothing kernel, it is possible to recover the signal with maximum frequency content [43].

Analytically, the acquired image $I_v(x)$ can be modelled as the superimposition of a theoretical CT value $I_u(x)$ and a vertical linear gradient $f(x)$.

$$I_v(\mathbf{x}) = I_u(\mathbf{x}) + f(\mathbf{x}) \tag{3.1}$$

Let U, V, and F be the probability densities of I_u, I_v, and f respectively. On 2D images, U, V, and F can be approximated by the intensity histograms of I_u, I_v, and f. Assuming that I_u and f are uncorrelated random variables, the probability density of the sum $I_u(\mathbf{x}) + f(\mathbf{x})$ is equivalent to the convolution of U and F [136]. This is due to the fact that the joint probability $P(I_v=v, I_u=u)$ distribution of I_v and I_u can be computed by using the Jacobian J for the change of variable from (f, u) to (v, u) [137]:

$$\begin{aligned} P(v,u) &= F(v-u)U(u)J \\ &= F(v-u)U(u)\begin{vmatrix} 1 & 1 \\ 0 & 1 \end{vmatrix} \\ &= F(v-u)U(u) \end{aligned} \tag{3.2}$$

Therefore, the distribution $V(v)$ can be derived as:

$$\begin{aligned} V(v) &= \int_{-\infty}^{+\infty} P(v,u)du = \int_{-\infty}^{+\infty} F(v-u)U(u)du \\ &= U(v) * F(v) \end{aligned} \tag{3.3}$$

Equation (3.3) indicates that given distributions F and V, U can be recovered by deconvolution. Based on Fourier theorem, U can be approximated using:

$$\widetilde{U} \approx \frac{\widetilde{F}^*}{\left|\widetilde{F}\right|^2 + Z^2}\widetilde{V} \tag{3.4}$$

where $\widetilde{U}$, $\widetilde{V}$ and $\widetilde{F}$ are the Fourier transforms of U, V and F respectively, '*' denotes complex conjugate, and Z is a small constant term introduced for numerical stability. Figure (3.5) shows the distribution V measured from a CT phantom image and the recovered distribution U. It is evident that the distribution V is broadened due to the superimposed density gradient. After deconvolution, U exhibits multiple distinct peaks, corresponding to different classes of density on the image. From this estimate of U, the expected value of I_u, given a measurement v of I_v, can be computed. Subtracting this expected value $E_u(v)$ from a measurement $I_v = v$ gives an estimate $\hat{f}$ of the density gradient f:

$$\hat{f}(v) = v - E_u(v) \tag{3.5}$$

Since f is assumed to be a linear vertical gradient, it can then be modelled by $\hat{f}_{fit}$, the least-mean square fit of $\hat{f}$ described by:

$$\hat{f}_{fit}(x, y) = \alpha y + \beta \tag{3.6}$$

In our implementation of the algorithm, V is derived from the actual CT image I_v, and the distribution F used for deconvolution in Equations (3.4) is a Gaussian distribution. Gradient-correction is achieved by subtracting the field $\hat{f}_{fit}$ (estimated with Equations (3.5) and (3.6)) from the original image I_v.

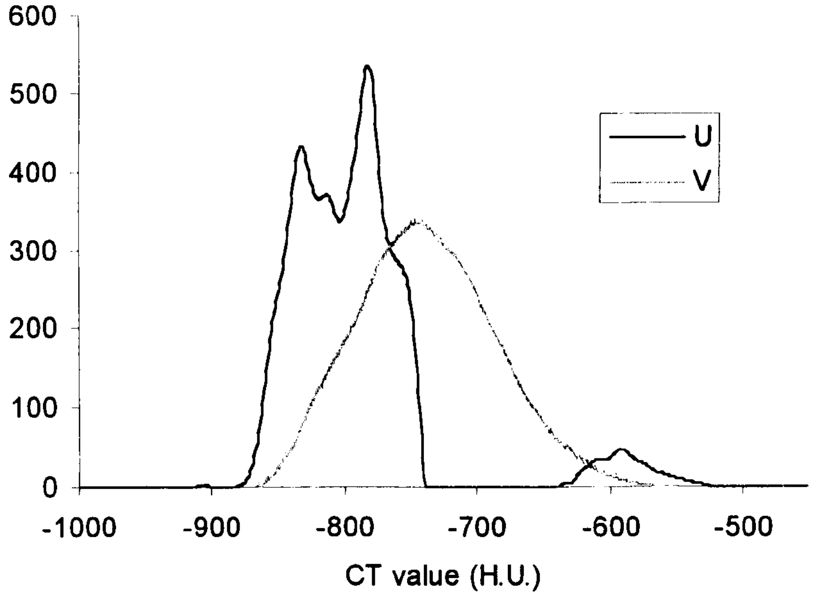


Figure 3.5. Probability distributions before and after deconvolution. The distribution *V* was computed from a CT phantom image. It is unimodal and blurred because of the gradient superimposed to the original signal. After deconvolution by a smoothing Gaussian *F*, the distribution *U* is recovered. It shows multiple distinct sharper peaks, corresponding to different homogeneous classes of intensity.

In practice, the density gradient cannot be precisely estimated with one iteration of the algorithm, due to noise and discretisation errors when computing probability distributions, and imperfect modelling of the deconvolution filter *F*. Several iterations are therefore needed for the algorithm to converge. In our numerical implementation, the signal I_v is defined by a 512x512 matrix. Histograms are computed with one bin per Hounsfield Unit (H.U.). The optimal standard deviation of the Gaussian distribution *F* to use was estimated with the phantom study described below. The numerical constant *Z* in Equation (3.4) is selected to be 1. The algorithm is considered to have converged when the slope α of the gradient $\hat{f}_{fit}$ estimated at the n^{th} iteration is smaller than 1/512, which is equivalent to density variations of less than 1 H.U. across the entire image.

3.2.3 Refining gradient-correction by histogram energy maximisation

The gradient-correction algorithm by histogram deconvolution is general and non-parametric. However, careful assessment reveals some numerical limitations and histogram deconvolution slightly underestimates the slope of the gradient, as shown in the following section. However, it is possible to overcome this problem by using histogram energy optimisation.

We assume that gradient deconvolution provides a good first approximation of the slope of the density gradient. This step of correction ensures that the density distributions of all classes do not overlap. The validity of this assumption will be demonstrated in section (3.4.2.3). The histogram of the image obtained after iterative deconvolution of the original histogram should exhibit several distinct peaks corresponding to different classes of intensity values. The estimate of the slope α of the gradient can be fine-tuned by searching in the vicinity of α for the slope that maximises the energy of the intensity histogram of the image. In fact, the higher this energy is the narrower the peaks of the histogram are. Let $I_{v,\beta}(x, y)$ be the function obtained by subtracting a vertical gradient with a slope β from the original image $I_v(x, y)$:

$$I_{v,\beta}(x,y) = I_v(x,y) - \beta y \tag{3.7}$$

The corresponding intensity histogram $H(I_{v,\beta})$ is given by the value of each bin b_j, b_j being the number of points (x, y) on the image for which $I_{v,\beta}(x, y)$ is equal to j. The histogram energy $E(I_{v,\beta})$ is computed as follows:

$$E(I_{v,\beta}) = \sum_j b_j^2 \tag{3.8}$$

Let α be the slope of the gradient estimated by histogram deconvolution. The value of β, found in the neighbourhood of α, that maximises $E(I_{v,\beta})$ allows for the recovery of the image with the sharpest histogram peaks, *i.e.*, the optimal gradient-corrected image. A mathematical proof is given in Appendix A.

3.2.4 Classification

After intensity gradient correction, the pixels within the lung parenchyma are classified into three classes, each representing under-attenuated, normal, and hyper-attenuated areas of the parenchyma respectively. The devised classification algorithm makes no reference to a pre-defined absolute index of CT densities. This is desirable in that the absolute CT value of under-attenuated lung on different patients may be very different.

Let $\{P_i\}_{0 \le i < N}$ be the set of pixels to be classified where N is the number of pixels. Each pixel P_i is represented on the image by a CT value of C_i. By defining $C^*_{L_i}$ as the intrinsic class intensity value of pixel P_i if label L_i is assumed, the proposed classification process is achieved by minimising the following objective function:

$$E = \sum_{0 \le i < N} \left| C_i - C^*_{L_i} \right| \tag{3.9}$$

The definition of this function is based on the following reasons. For an intensity pattern that is piecewise constant, the pixel value of P_i should not deviate substantially from its intrinsic class intensity value C_j^* if the image is segmented correctly. Therefore, for each class, the value of $\sum_{P_i \in class(j)} \left| C_i - C_j^* \right|$ will become minimum. When all the classes are taken into consideration, this leads to $\sum_j \sum_{P_i \in class(j)} \left| C_i - C_j^* \right|$ being minimum, and this function is equivalent to E defined in Equation (3.9). For the purpose of segmenting the lung parenchyma into under- *(u)*, normal- *(n)*, and hyper-attenuated *(h)* regions, the labelling set is defined as $\{u, n, h\}$ and $j, L_i \in \{u, n, h\}$. The intrinsic intensity value of class j is defined as the mean intensity value of P_i for all $P_i \in class(j)$. In Equation (3.9), C_j^* are unknown quantities and the classification of the lung parenchyma is performed by finding C_j^* such that E is minimised. Since we know that $C_u^* < C_n^* < C_h^*$, numerically the problem formulated by Equation (3.9) can be solved by considering the histogram b_ξ derived after intensity gradient correction. We define two cut-off values ξ_1 and ξ_2 and for each pixel P_i the selection of the labelling is determined by:

$$L_i = \begin{cases} u & if\, C_i < \xi_1 \\ n & if\, \xi_1 \le C_i < \xi_2 \\ h & if\, \xi_2 \le C_i \end{cases} \tag{3.10}$$

Accordingly, Equation (3.9) can be rewritten as:

$$E = \sum_{C_i < \xi_1} \left| C_i - C_u^* \right| + \sum_{\xi_1 \le C_i < \xi_2} \left| C_i - C_n^* \right| + \sum_{\xi_2 \le C_i} \left| C_i - C_h^* \right| \tag{3.11}$$

and the values of values ξ_1 and ξ_2 are sought by an iterative search. Within each iteration the intrinsic class intensity values are updated as:

$$C_u^* = \underset{C_i < \xi_1}{mean} \{ C_i \}; C_n^* = \underset{\xi_1 \le C_i < \xi_2}{mean} \{ C_i \}; C_h^* = \underset{\xi_2 \le C_i}{mean} \{ C_i \} \tag{3.12}$$

This process incurs a computational cost of $O(N^2/2)$ and the extent of under-attenuated areas in the final segmented image is expressed as the percentage of the area of the lung labelled as *(u)*. Figure (3.6) shows an example of processed CT scan. By applying the classification algorithm to Figure (3.6a), the segmented image shown on Figure (3.6b) is obtained. The areas classified as hypo-attenuated are shown as the darkest grey shade. This result was validated by that of experienced radiologists. Figure (3.6c) shows the segmentation obtained when classification is performed directly without gradient-correction. Most of the anterior lung is then wrongly identified as diseased, demonstrating the importance of the correction for inhomogeneous perfusion.

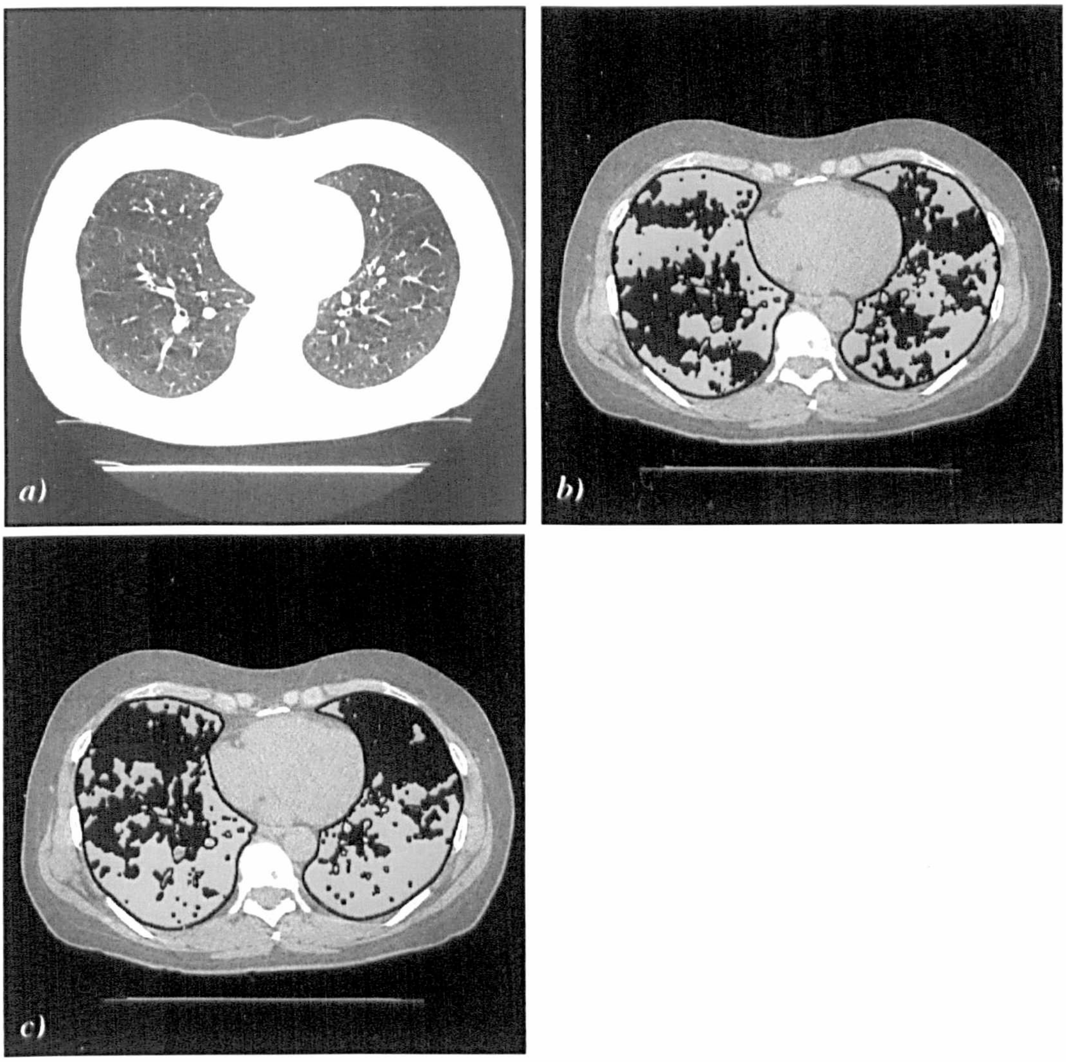


Figure 3.6. The effect of gradient-correction on the result of classification. *a)* Original CT scan. The anterior part of the parenchyma appears darker because of gravity effect. *b)* Segmented image after gradient-correction. The correct hypo-attenuated areas are identified (shown as the darkest grey shade). *c)* Segmented image without gradient-correction. The areas labelled as hypo-attenuated include most of the anterior lung.

3.2.5 Phantom evaluation

The accuracy of the approximations made by using a deconvolution filter in the Fourier domain was estimated by applying convolution/deconvolution to a set of synthetic images. In order to quantitatively evaluate the proposed algorithm, a CT lung phantom was constructed. The main features of the parenchyma that the phantom modelled were the texture and the density differences of the normal and diseased lung. This was achieved by employing polyethylene (PE) foams of various densities. A detailed description and justification of the lung phantom are given in Appendix B. Figure (3.7) shows an example of CT scan of the lung phantom, and the segmented image. By using PE foams of different densities, patches of different CT attenuation can be obtained, as shown in Figure (3.7a). Robustness to noise of the technique was estimated using several scanning protocols with increasing associated levels of noise. The accuracy of the gradient-correction was measured by superimposing linear gradients of various intensities on cross-sectional images of the phantom. The slope of the gradients applied varied from 0 H.U./mm to 1 H.U./mm. Gaussian distributions of various sizes were tested, so as to practically determine the optimal deconvolution kernel to be used. For a set of 6 phantom images with gradient slopes ranging from 0 H.U./mm to 1 H.U./mm, the average error and the speed of convergence of the gradient-correction algorithm were measured, for Gaussian distributions with standard deviation between 25 and 200 bins.

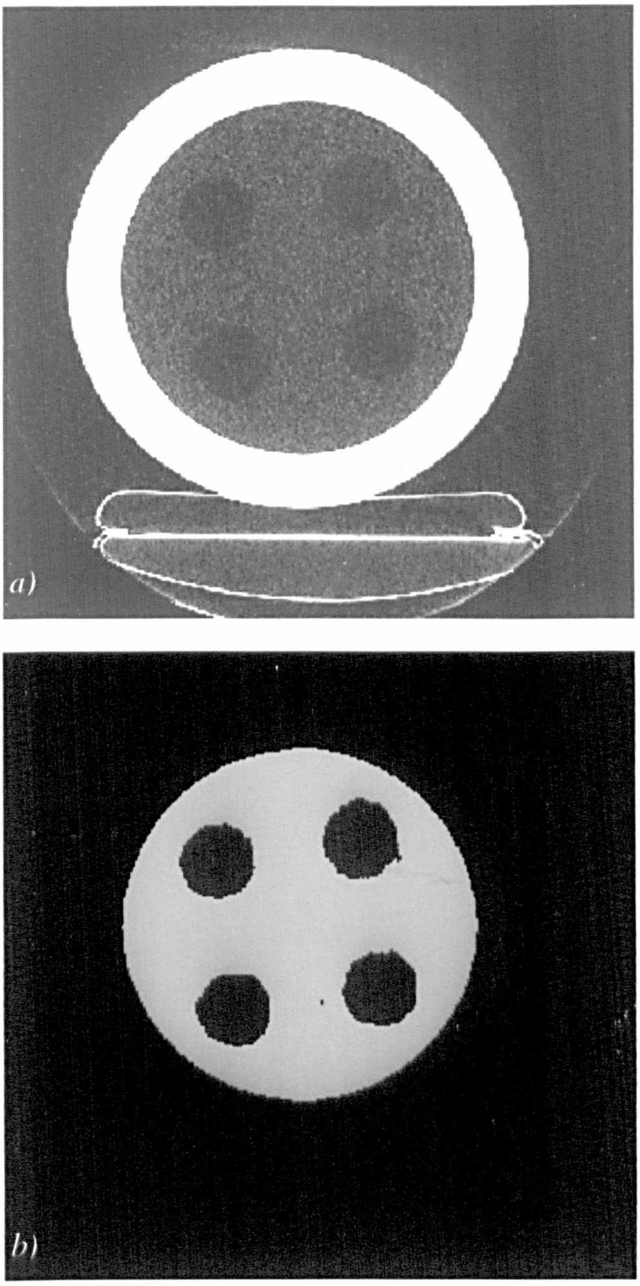


Figure 3.7. *a)* Example of CT scan of lung phantom with four hypo-attenuated patches of different densities. *b)* The result of region segmentation by using the proposed technique.

3.2.6 Patient studies

To assess the value of the proposed technique in a clinical context, the algorithm was applied to a group of 15 patients with suspected or known constrictive obliterative bronchiolitis and 8 normal subjects, *i.e.*, subjects with normal pulmonary function (forced expiratory volume in 1 s (FEV_1), forced vital capacity (FVC), total lung capacity (TLC), residual volume (RV), and gas transfer adjusted for accessible volume (K_{CO}) all within 20 % of predicted values). Images were acquired using a High-Resolution CT protocol (1.5 mm beam collimation, sharp kernel reconstruction algorithm, on an Imatron C-150-L ultrafast scanner, Imatron, Inc., San Francisco, CA). The normal subjects underwent CT as part of a separate clinical trial, approved by the local Ethics Committee; the remaining studies were undertaken for clinical indications. The patients with constrictive obliterative bronchiolitis have previously been included in a study of the structure/function relationship of this disease [4]. The results were compared with those from two experienced observers. Images were acquired from the apex to the base of the lungs, and one slice at a pre-defined anatomical position (midway between the pulmonary venous confluence and one centimetre above the right hemi-diaphragm) was retained in each study. The images were scored independently by two experienced radiologists who subjectively estimated the percentage (to the nearest 5 %) of abnormal lung of hypo-attenuation on each image (window settings: level = -800 H.U.; width = 1000 H.U.). The automated quantification algorithm was applied, and the average density difference, denoted as δ, between the areas labelled as normal and the areas labelled as hypo-attenuated was measured.

3.3 RESULTS

3.3.1 Phantom Study

By using the standard HRCT protocol, the segmentation error accounted for 0.27 % of all pixels being misclassified, for density differences between the background and the areas of hypo-attenuation of at least 30 H.U. For a density difference of 6 H.U., misclassification reached more than 30 % of all pixels. With the phantom presenting density differences of 30 H.U., using various HRCT scanning protocols, including the available one with the highest associated level of noise (standard deviation of homogeneous air: 20.45 H.U.), the segmentation error never exceeded 0.52 % of all pixels being misclassified. Figure (3.8) shows the error made by the gradient-correction algorithm, as well as its speed of convergence, for deconvolution Gaussian distributions of various standard deviations. These values were measured with a set of 6 phantom images. Figure (3.8) shows that with increasingly large Gaussian distributions, accuracy of the results decreases but speed of convergence increases. As a trade-off, a standard deviation of 100 was retained for the patient study. Figure (3.9) shows the slope of the vertical density gradient estimated by histogram deconvolution on several phantom images. Error was on average 6.7 % of the theoretical slope.

The limitations of the approximate deconvolution process in the Fourier domain, as described in Equation (3.4), are illustrated in Figure (3.10). Using a Gaussian filter F, applying convolution then deconvolution to a wide distribution U allows for the recovery of a good approximation U' of U, as shown in Figure (3.10a). However, a much poorer approximation is obtained when starting from a narrow distribution, as shown in Figure (3.10b).

The limitation of the histogram deconvolution technique can be overcome by subsequently applying the histogram energy maximisation algorithm. Figure (3.11) shows the slope of the vertical density gradient recovered on several phantom images, after refining with histogram maximisation the initial estimate. Error was found to be 0.7 % on average.

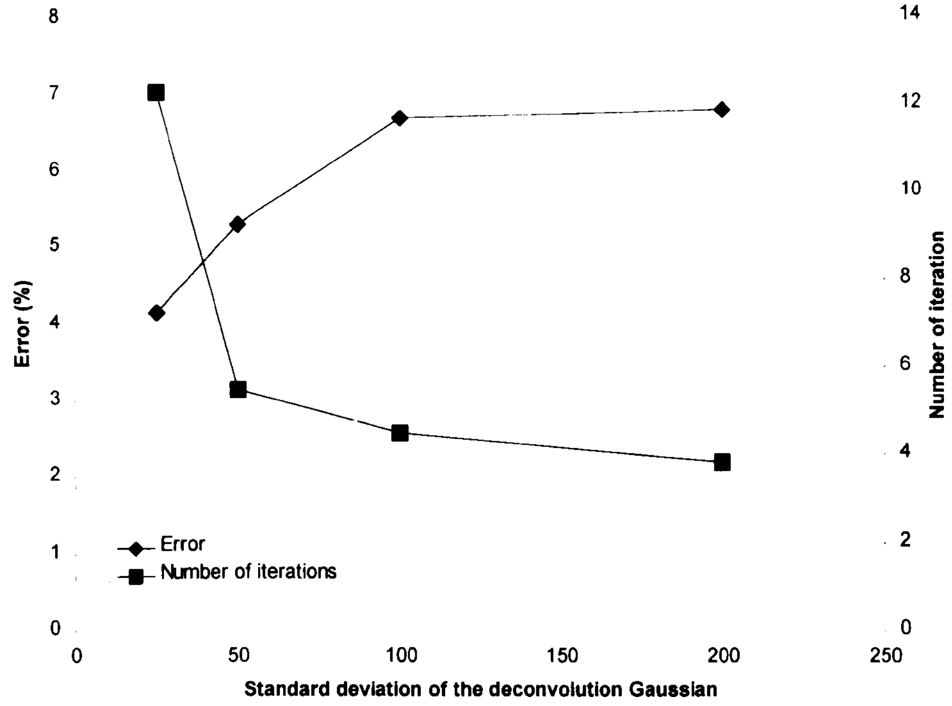


Figure 3.8. Influence of size of the Gaussian distribution F in the gradient-correction algorithm. When the standard deviation of F increases, the number of iterations required for the algorithm to converge decreases, but the average error made when estimating the gradient slope on a set of 6 phantom images increases. Determining the optimal size of F is a trade-off between accuracy and speed of convergence.

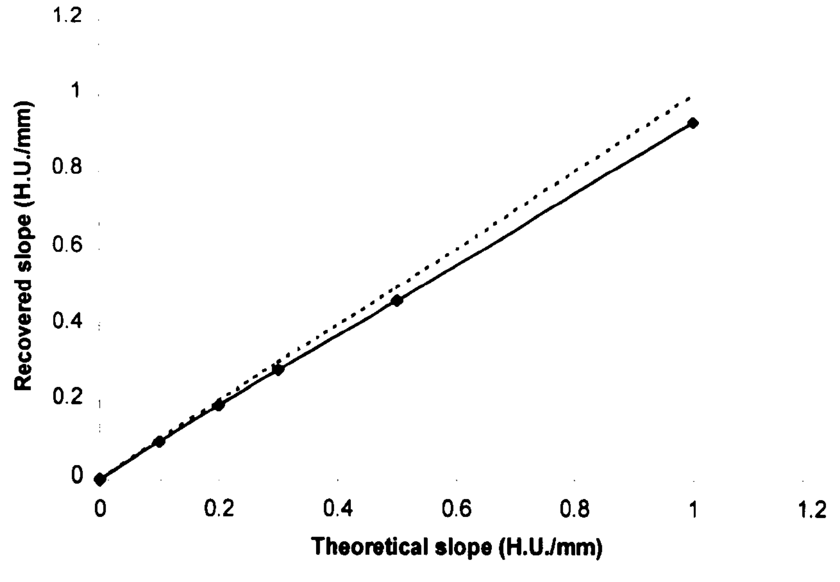


Figure 3.9. The slope of the vertical intensity gradient, as estimated by the proposed histogram deconvolution algorithm applied to several phantom CT images with different superimposed gradients. The error made by the algorithm is on average 6.7 % of the theoretical slope.

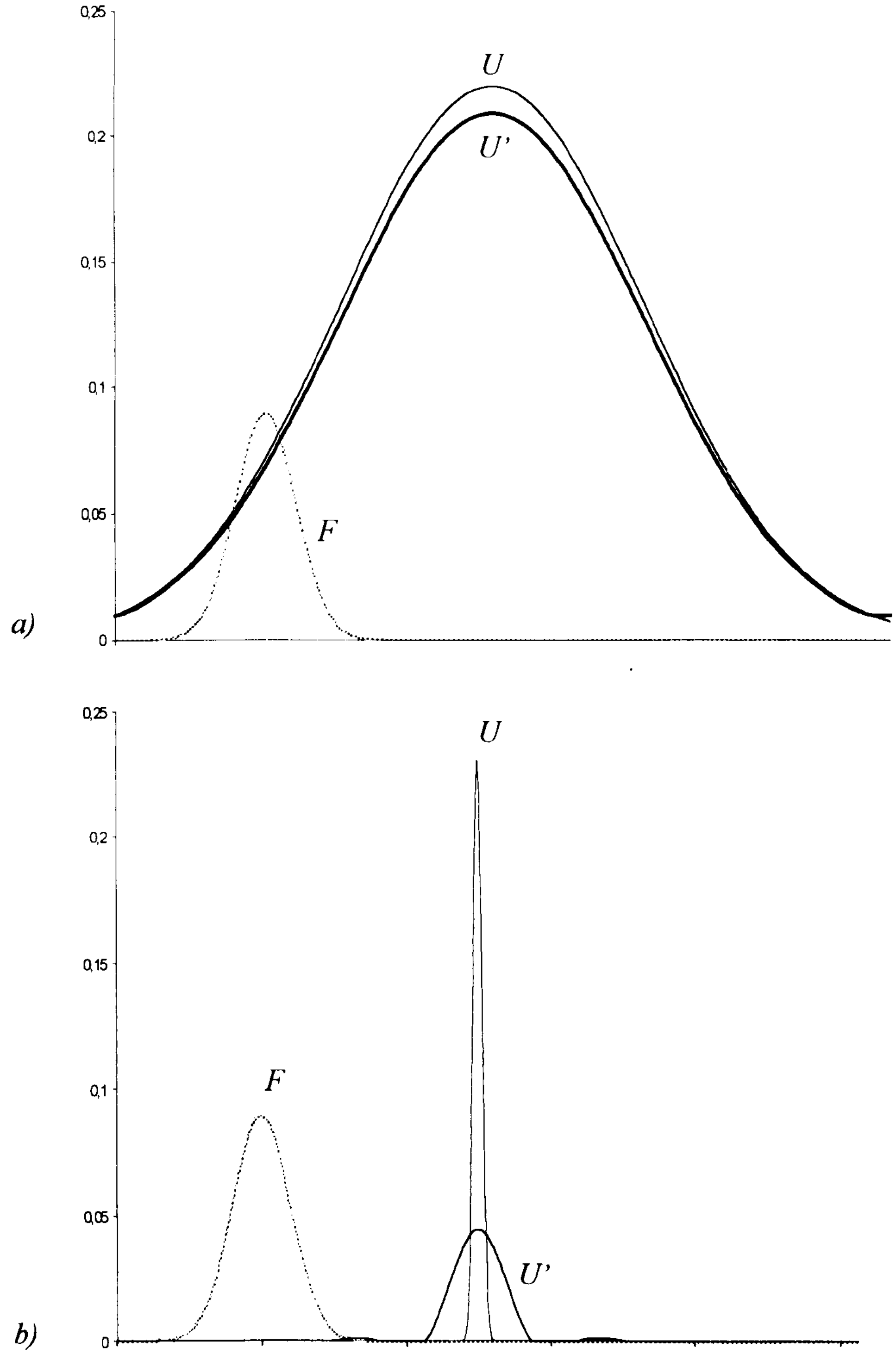


Figure 3.10. *a)* Applying convolution then deconvolution to a wide distribution *U*, using a Gaussian filter *F*, allows for the recovery of a good approximation *U'* of the original wide distribution. *b)* However, when applying the same process to a narrow distribution *U*, the recovered distribution *U'* is much wider. The approximations of the deconvolution filter handle poorly narrow distributions.

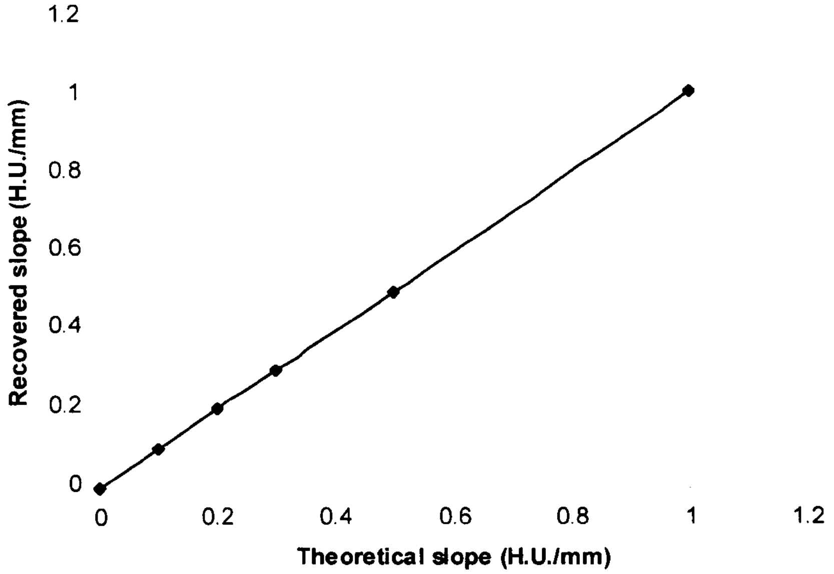


Figure 3.11. The slope of the vertical intensity gradient, as estimated by histogram energy maximisation, after the histogram deconvolution algorithm has been applied to several phantom CT images with different superimposed gradients. The error made by the algorithm is on average 0.7 % of the theoretical slope.

3.3.2 Patient Studies

For the described group of 23 subjects, the weighted Kappa statistic (κ_w) [76] measuring inter-observer variability for the two expert radiologists, was equal to 0.49, suggesting moderate inter-observer agreement. To assess the accuracy of the classification technique, the average of the two scores (denoted S_{ave}) was retained for comparison with the automated scoring (denoted S_{auto}). For patients with constrictive obliterative bronchiolitis the difference between S_{ave} and S_{auto} is on average 8.2 %. Figure (3.12) illustrates the difference between S_{ave} and S_{auto} for these 15 images. For normal subjects, the difference between S_{ave} and S_{auto} is on average 25.8 %. The density difference δ between areas labelled as normal and areas labelled as under-attenuated is on average 48.7 H.U. (standard deviation $\sigma = 1.5$) for normal subjects and 83.0 H.U ($\sigma = 10.4$) for patients with constrictive obliterative bronchiolitis, as illustrated in Figure (3.13). The gradient-correction algorithm converged in 4.08 iterations on average (minimum: 2 iterations, maximum: 5 iterations). Figure (3.14) shows an example of the speed of convergence of the proposed technique. The slope of the gradient was on average 0.37 H.U./mm for normal subjects and 0.21 H.U./mm for patients with constrictive obliterative bronchiolitis.

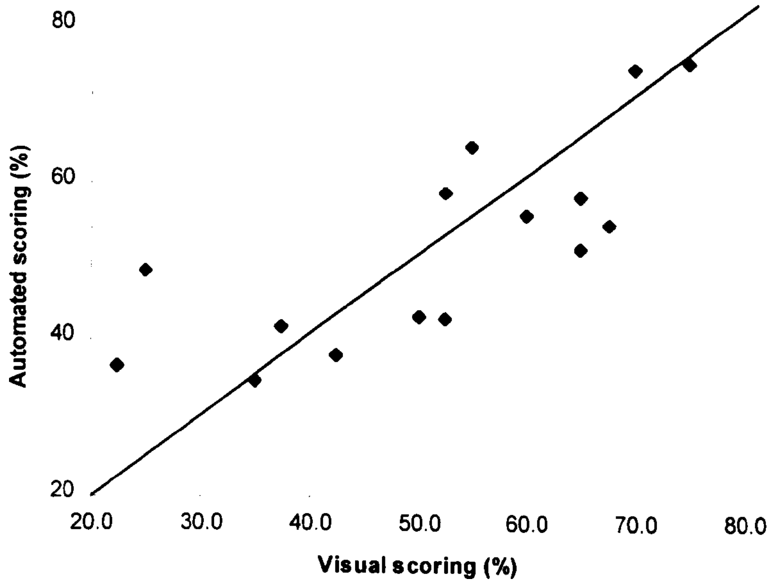


Figure 3.12. Comparison between visual scoring (S_{ave}) and automated scoring (S_{auto}) when measuring the extent of under-attenuated lung on 15 scans of patients with constrictive obliterative bronchiolitis. The difference between S_{ave} and S_{auto} is on average 8.2 %.

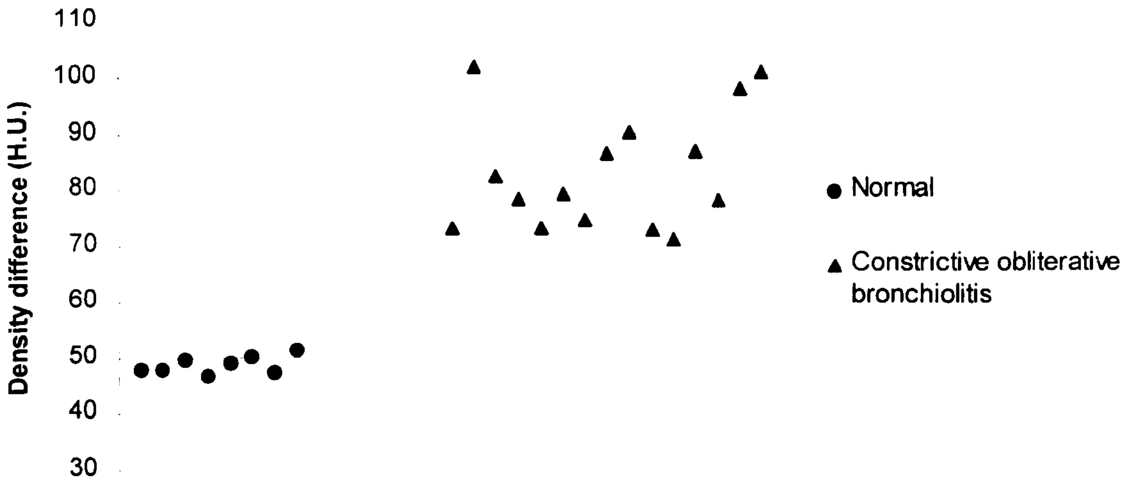


Figure 3.13. Mean density differences δ between areas labelled as 'normal' and areas labelled as 'hypo-attenuated' for 8 normal subjects and 15 patients with constrictive obliterative bronchiolitis. The density differences in the diseased lung parenchyma are consistently greater than in the normal lung.

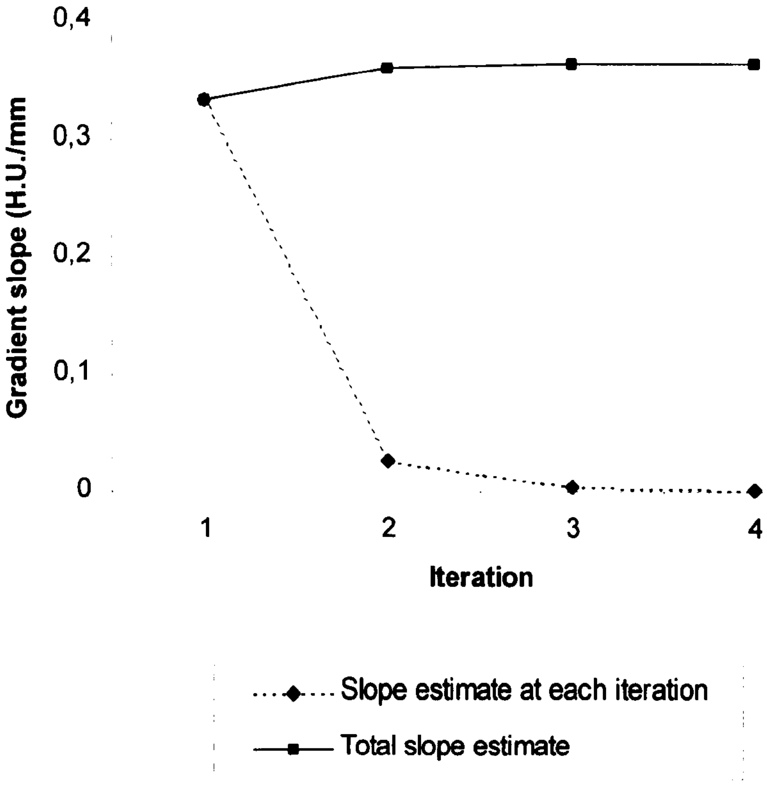


Figure 3.14. The convergence of the gradient-correction algorithm applied to a CT scan of the lungs with suspected constrictive obliterative bronchiolitis. After one iteration, 91.4 % of the final estimate was recovered. The algorithm converged in 4 iterations.

3.4 DISCUSSION

In this chapter, a fully automated technique for the segmentation of areas of hypo-attenuation of the lung parenchyma has been described. Without any human interaction, it identifies the different anatomical structures visible on a cross-sectional chest image. It then automatically corrects for the density-gradient due to gravity, by using a non-parametric mathematical model. Such a correction ensures that subsequent classification based on the CT values is accurate. Both phantom and patient studies demonstrated the validity of this approach.

3.4.1 Image segmentation

A fully automated algorithm performs segmentation of the chest, trachea, macroscopic blood vessels, and lung parenchyma. Thresholding and region-growing are simple tools well-suited for the segmentation of organs with non-overlapping distributions of CT densities, such as the chest wall and the lung parenchyma. These techniques, however, may fail at the interface between different anatomical structures. Because of partial volume effect, the voxels at the interface of the chest wall at the parenchyma, for instance, may be assigned any CT value between –900 H.U. (aerated lung) and 700 H.U. (rib). The use of morphological operators allows for the correct segmentation of such interfaces. The operation of dilation, for instance, expands the region identified as chest-wall by thresholding and region growing, so as to incorporate the interface with the parenchyma. It can then be safely assumed that the remaining pixels represent only the lung parenchyma. The structuring elements, which define the spatial range of action of morphological operators, can be chosen according to the parameters of the scanning acquisition. Given the size of the pixels in the x, y, and z directions, one can estimate the number of voxels affected by the partial volume effect at the interface between two different organs, and can therefore estimate the size of the required structuring element. The fact that no user interaction is required ensures the reproducibility of the segmentation method. Admittedly, the proposed technique is *ad hoc* by using empirically selected parameters; however, it proved to be effective for performing unsupervised anatomical segmentation on a large set of images.

3.4.2 Gradient-correction

3.4.2.1 Gradient-correction by histogram deconvolution

On scans of a phantom emulating the parenchyma and its density differences, the histogram deconvolution algorithm was shown to be able to deal with gradients 3 times more intense than those seen typically on scans of normal human lungs. Its accuracy was on average 93.3 % (95.9 % when using a narrow Gaussian distribution F). It is worth emphasising that the assumption that I_u and f represent uncorrelated variables may not be true. On CT images, the different classes of density tend to cluster to form patches. Therefore, the effect of a superimposed gradient depends on the spatial distribution of the images and the histogram-convolution only approximates the influence of the gravity-dependent density gradient on CT images. Another reason that may explain the average error made by the gradient-correction algorithm, is the approximate nature of the deconvolution filter used. One of the reasons why the algorithm converges too early is the inability of the process to deal with narrow distributions, as shown in Figure (3.10).

3.4.2.2 Refining gradient-correction by histogram energy maximisation

In spite of its limitations, the histogram deconvolution algorithm provides a good initial estimate of the slope of the gradient. This estimate can subsequently be fine-tuned by histogram energy maximisation. The accuracy of the method is high, greater than 99.3 % on average, as tested on a set of phantom images.

3.4.2.3 Validity of the assumptions

The histogram deconvolution algorithm is general and provides a first approximation of the slope to recover. A precise estimate of the slope can then be obtained by histogram energy maximisation. The latter technique relies on the assumption that the density distributions for the different classes do not overlap. It can be verified numerically that this condition holds for lung CT images. It may be assumed that there are two homogeneous classes, with a density difference δ_{class}. They are superimposed with a vertical gradient with a slope α. After histogram deconvolution, the remaining gradient has a slope α'. The maximum distance between two points belonging to a same class, projected onto the y axis, is denoted Y_{class}. Therefore, the remaining vertical gradient introduces density differences within each class that are at most equal to $\delta'_{in\text{-}class}$ given by:

$$\delta'_{in-class} = Y_{\max} \cdot \alpha' \qquad \textbf{(3.11)}$$

The density distributions of the two classes do not overlap if the effect of the remaining gradient is smaller than the original inter-class density difference, *i.e.*, if $\delta_{in\text{-}class} < \delta_{class}$. The maximum gradient slope encountered in this study was 0.61 H.U./mm. We can assume the maximum slope α_{max} to be 1 H.U./mm. This is equivalent to $\alpha_{max} = 0.68$ H.U./pixel when reconstructing the largest field of view used on lung CT. After gradient-correction by histogram deconvolution, the remaining gradient slope was found to be at most 10 % of the original gradient slope. Histogram energy optimisation is therefore applied to images with a slope lesser than $\alpha'_{max} = 0.068$ H.U./pixel. On a 512*512 lung CT image, the largest y dimension Y_{class} of a class of pixels is less than

$Y_{max} = 512$. Therefore, the density difference $\delta'_{in\text{-}class}$ between two points belonging to the same class superimposed with the remaining gradient verifies:

$$\begin{aligned}\delta'_{in-class} &= Y_{\max} \cdot \alpha'_{\max} \\ &\leq 512 * 0.068 = 34.8 \text{ H.U.}\end{aligned} \qquad \textbf{(3.12)}$$

In this study, the minimum inter-class density difference δ_{class} between areas of normal and diseased lung was found to be 70 H.U. We therefore have:

$$\delta'_{in-class} \leq \delta_{class} \qquad \textbf{(3.13)}$$

Consequently, the assumption that density distributions between classes do not overlap after histogram deconvolution is verified. It is legitimate to subsequently apply the histogram energy maximisation algorithm. It is worth noting that prior to histogram deconvolution, the class separation assumption does not necessarily hold. The two steps of the gradient-correction technique are therefore necessary: histogram deconvolution provides a first approximation of the slope of the gradient. Correcting for it ensures that intensity classes no longer overlap, and therefore allows for the use of histogram energy maximisation to obtain a more precise estimate of the slope. A schematic illustration is given in Figure (3.15).

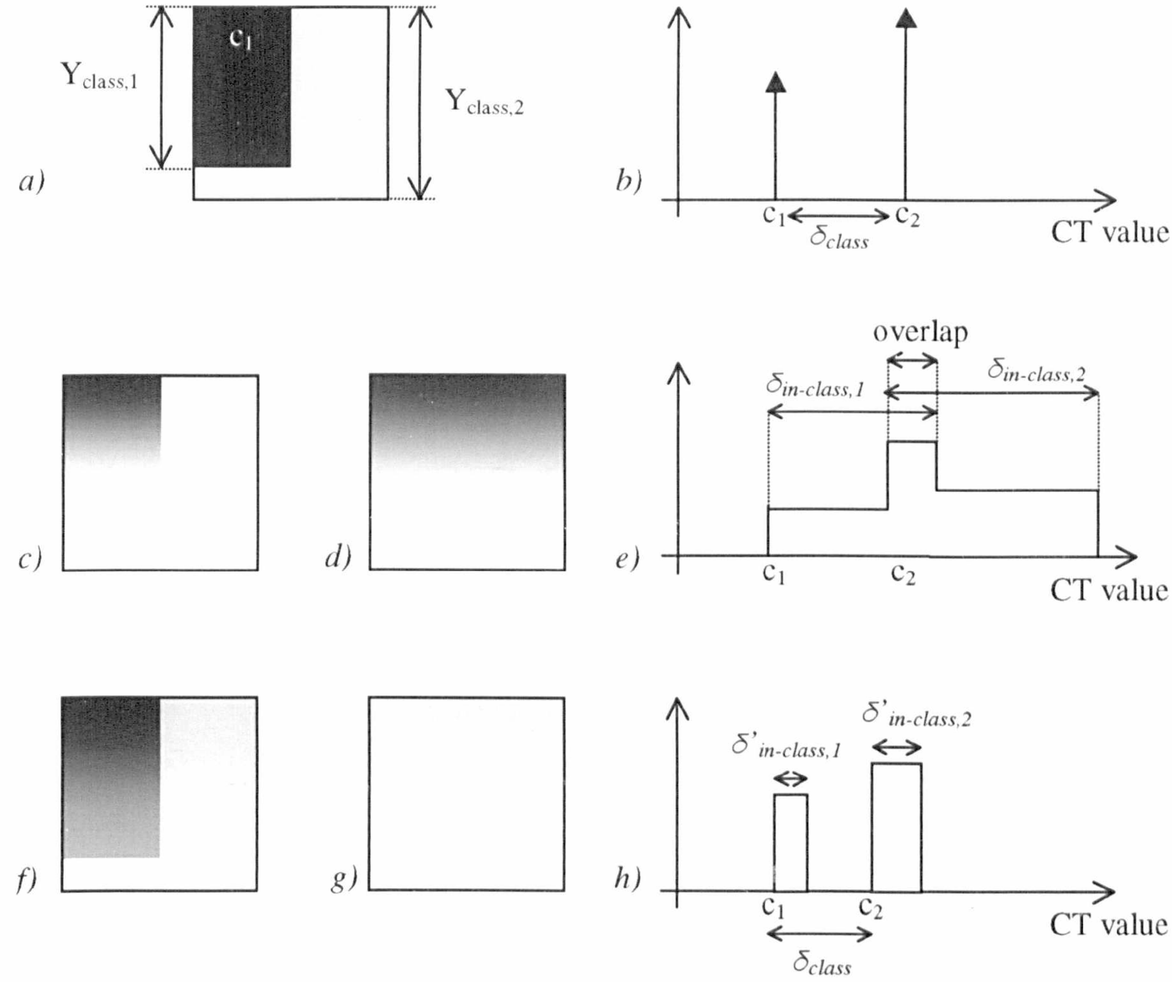


Figure 3.15. Schematic illustration of the gradient-correction technique. *a)* Original image showing two homogeneous classes of intensity c_1 and c_2. The largest vertical dimension for each class i is denoted $Y_{class,i}$. *b)* Intensity histogram of the original image showing two peaks at c_1 and c_2 separated by δ_{class}. *c)* Original image superimposed with strong vertical intensity gradient with a slope α, shown in Figure (3.15d) *d)* Strong vertical intensity gradient. *e)* Intensity histogram for Figure (3.15c). The effect of the gradient is to spread the values of each class on a range given by $\delta_{in\text{-}class,i} = Y_{class,i}.\alpha$. For a strong gradient such as the one shown in Figure (3.15d), this effect is such that the two intensity classes overlap. Gradient-correction cannot be achieved by histogram energy maximisation, but by histogram deconvolution. *f)* The result of histogram deconvolution applied to Figure (3.15c). *g)* A small remaining intensity gradient with a slope α' can still be observed. *h)* Histogram for Figure (3.15f). The effect of the remaining gradient does not let intensity classes overlap. It is possible to apply the histogram energy maximisation technique to remove the small intensity gradient.

Without a precise mathematical model of the image to recover, linear gradient-correction is difficult. The proposed two-staged algorithm, based on histogram deconvolution and energy maximisation, is general, non-parametric, and provides an accurate estimate of the density gradient.

3.4.3 Classification

The phantom studies demonstrated the robustness of the classification algorithm in the presence of image noise. The algorithm only failed for the segmentation of classes with a mean intensity difference of 6 H.U. The intensity distributions of such classes overlap greatly, and their distinction based on CT value, becomes extremely difficult. The clinical evaluation confirmed that the proposed automated method accurately quantifies the extent of diseased lungs on CT images for patients with constrictive obliterative bronchiolitis. As a comparison, a fuzzy C-means algorithm [138] based on similar assumptions was also applied, but it gave less satisfactory results (average error: 24.0 %). It must be noted, however, that the proposed clustering algorithm over-estimates the areas of hypo-attenuation on scans of normal lungs. This is due to the small inhomogeneities of the normal parenchyma. Density differences of 50 H.U., as typically identified by the algorithm applied to normal lung images, may not be visible to the human eye on window settings normally used in clinical practice. Nevertheless, the mean density difference between the two classes segmented by the algorithm may provide a robust discriminant for distinguishing between normal and abnormal scans. As seen in Figure (3.13), this difference is consistently and significantly greater for patients with constrictive obliterative bronchiolitis. This finding can be incorporated in the decision-support system presented in Chapter 6.

The proposed technique is mainly based on density distributions. Without doubt, further improvements in the diagnostic discrimination of this technique can be achieved by incorporating other CT signs and features such as the texture of the lung parenchyma and state of the bronchi. The integration of these features in a high-level framework with *a priori* clinical knowledge is likely to enhance the reliability of the method.

3.5 CONCLUSION

This chapter presents a fully automatic method for the analysis of CT lung images of patients with suspected airways disease. The technique performs segmentation of the anatomical structures, corrects for the density gradient due to gravity, and classifies the areas of hypo-attenuation of the parenchyma. The phantom study demonstrated the robustness of the technique against image noise, and its ability to accurately remove the linear density gradient observed in the lungs. Clinical testing showed that it is able to reliably measure the extent of the areas of hypo-attenuation on images of patients with small airways disease.

CHAPTER 4

ERS TRANSFORM FOR THE DETECTION OF BRONCHIAL ABNORMALITIES

4.1 INTRODUCTION

The most important indirect sign for the detection of obstructive lung diseases on HRCT is the identification of areas of hypo-attenuation in the lung parenchyma. Based on the analysis of global density distribution, methods for the quantification of the extent of hypo-attenuated areas of the lungs have been described in Chapter 3 [30,139]. Techniques for enhanced visualisation of the density differences have also been presented [12,27,28,44]. The inherent problem with these approaches is that contextual information is not taken into account. Indeed, areas of hypo-attenuation in the lung parenchyma cannot be consistently and confidently labelled as abnormal without considering ancillary findings, such as the morphology of the airways. Several low level feature extractors need to be combined in a high-level system incorporating *a priori* knowledge to achieve a computerised diagnostic aid. Within this framework, essential information can be derived from the morphology of the airways by measuring bronchial dilatation and wall thickening. More specifically, the presence of density differences in the lung parenchyma (known as mosaic attenuation pattern) can be caused by different pathologies: interstitial lung disease, vascular abnormalities, or airway disease [23].

Among other discriminators, the identification of bronchial wall dilatation on inspiratory scans suggests an airway disease rather than a vascular cause. This example stresses that there are ambiguities associated with the isolated consideration of intensity differences. To resolve this problem and derive clinically meaningful conclusions, an algorithm for the automated detection of bronchial abnormalities is necessary. Such a technique can also be applied to the automated assessment and quantification of other lung pathologies, for example asthma [11,32] or allograft transplant rejection [25,140].

On cross-sectional images, bronchi running near perpendicular to the image plane appear as high-attenuation near circular rings. This chapter presents a novel technique for the recognition of elliptical rings, based on their geometric properties. An Edge-Radius-Symmetry (ERS) transform is used for the analysis of gradient maxima and minima in local polar co-ordinates. Based on the intensity (measured by function E), radial distribution (measured by function R), and symmetry (measured by function S) of these maxima and minima, the pixels are ranked to provide a sorted list of the most likely positions of dominant elliptical rings. The algorithm was evaluated on synthetic data and a group of nine patients showing different lung pathologies.

4.2 AUTOMATED PATTERN IDENTIFICATION

Template matching and Hough transform are the most common techniques for the detection of circular patterns. Although they are suitable for the current application of detecting bronchial wall, they can be computationally expensive due to the high number of parameters required to define a ring-shaped structure. These parameters include position, size, orientation, ratio of the short and long axes, and width of the ring. The five parameters involved for detecting an ellipse, for instance, can make a standard

Hough transform impractical in terms of computational and memory demands. It has been shown that even adaptive and two-pass methods remain time consuming [79].

4.2.1 ERS Transform

The ERS transform introduced in this chapter is based on the analysis of the distribution of gradient maxima and minima in the neighbourhood of each pixel. Ellipses and rings are patterns symmetrical relative to their centroids. In polar co-ordinates, the expression of the intensity and spatial distribution of the most significant edges has characteristic properties of symmetry and uniformity. The ERS transform is based on the analysis of edge distribution in local polar co-ordinates. In the case of CT images, given the CT densities of air, bronchial wall, and lung parenchyma, the intensity gradient should show local maxima and minima at the points on the inner border (*i.e.* at the interface between airway lumen and bronchial wall) and outer border (*i.e.* at the interface between bronchial wall and lung parenchyma) of a bronchus respectively.

Let $I(x, y)$ be the original image, we consider the local polar co-ordinates (r, θ) centred at (x_0, y_0), and a neighbourhood Ω of radius r_0. The maximum of the intensity gradient within Ω along direction θ is defined as $e_1(\theta)$:

$$e_1(\theta) = \max_{0<r<r_0}\left(\frac{\partial I}{\partial r}(r,\theta)\right) \tag{4.1}$$

with its corresponding radius defined as $r_1(\theta)$. The size r_0 of the neighbourhood Ω is to be defined according to the resolution of the image and the typical size of the patterns to identify. Similarly, the minimum of the intensity gradient along direction θ within the neighbourhood Ω is defined as $e_2(\theta)$:

$$e_2(\theta) = \min_{0<r<r_0}\left(\frac{\partial I}{\partial r}(r,\theta)\right) \tag{4.2}$$

and its radius is denoted as $r_2(\theta)$. This generates four functions in relation to θ: $e_1(\theta)$, $e_2(\theta)$, $r_1(\theta)$, and $r_2(\theta)$ which describe the radial distribution of edges around point (x_0, y_0). Figure (4.1) gives a schematic representation of the definition of $e_1(\theta)$, $e_2(\theta)$, $r_1(\theta)$, and $r_2(\theta)$ for a pixel (x_0, y_0) located inside an elliptical ring, shown in Figure (4.1a). By searching along direction d_θ, the gradient maximum $e_1(\theta)$, shown in the gradient profile in Figure (4.1b), is found at the inner border of the ring. Similarly, the gradient minimum $e_2(\theta)$ is found at the outer border of the ring. The positions $r_1(\theta)$ and $r_2(\theta)$ of these maximum and minimum, relative to (x_0, y_0), are the intersections of the ring along direction d_θ.

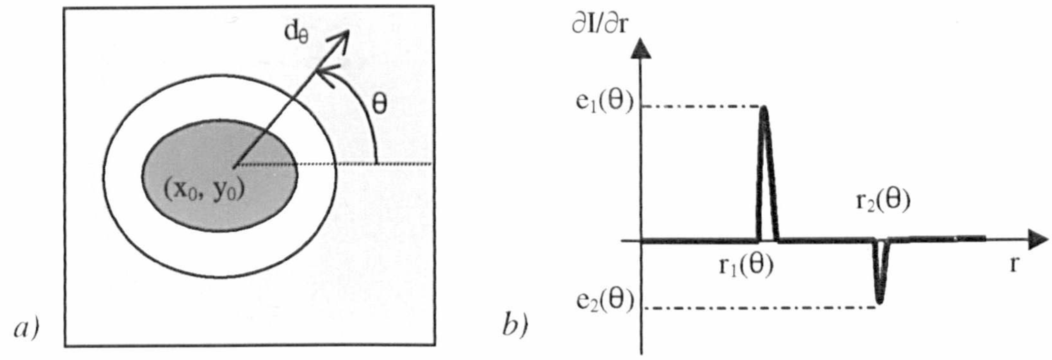


Figure 4.1. *a)* Schematic representation of the function $I(x, y)$ for an elliptical ring. The darkest areas represent the lowest intensities. At pixel (x_0, y_0), the maximum and minimum of the intensity gradient are searched along each direction d_θ. *b)* The profile of the intensity gradient along direction d_θ. The distance to the pixel (x_0, y_0) is denoted r, and $\partial I/\partial r$ measures the intensity gradient. At the inner border of the ring, a sharp intensity increase results in a maximum peak $e_1(\theta)$ at a distance $r_1(\theta)$. Further, at the outer border of the ring, a sharp intensity decrease results in a minimum peak $e_2(\theta)$ at a distance $r_2(\theta)$.

At every pixel *(x, y)*, three quantities *E*, *R*, and *S* are then derived from the functions $e_1(\theta)$, $e_2(\theta)$, $r_1(\theta)$, and $r_2(\theta)$.

- The value *E(x, y)* measures the strength of the edges found in the neighbourhood of *(x, y)*. It is defined as:

$$E(x,y) = \frac{1}{\underset{0\le\theta<2\pi}{mean}(|e_1(\theta)|) + \underset{0\le\theta<2\pi}{mean}(|e_2(\theta)|)} \tag{4.3}$$

where:

$$\underset{0<\theta<2\pi}{mean}(f(\theta)) = \frac{1}{2\pi}\int_0^{2\pi} f(\theta)d\theta \tag{4.4}$$

For a well-defined ring with sharp edges, the values of $e_1(\theta)$ are highly positive and the values of $e_2(\theta)$ are highly negative, thus contributing to a small value of *E(x, y)*.

- The value *R(x, y)* measures the uniformity of the radial distribution of the edges, which is:

$$R(x,y) = \frac{\sigma(r_1(\theta))}{\underset{0\le\theta<2\pi}{mean}(r_1(\theta))} + \frac{\sigma(r_2(\theta))}{\underset{0\le\theta<2\pi}{mean}(r_2(\theta))} \tag{4.5}$$

where $\sigma(f(\theta))$ is the average deviation of $f(\theta)$ defined as:

$$\sigma(f(\theta)) = \frac{1}{2\pi}\int_0^{2\pi} \left| f(\theta) - \underset{0\le\theta<2\pi}{mean}(f(\theta)) \right| d\theta \tag{4.6}$$

For a perfectly circular ring centred at *(x, y)*, the values of $r_1(\theta)$ are all identical, and so are the values of $r_2(\theta)$. The corresponding value of *R(x, y)* is zero.

- The value *S(x, y)* measures the symmetry of the strength of the edges. It is defined as:

$$S(x,y) = mean(\delta_1(\theta)) + mean(\delta_2(\theta)) \tag{4.7}$$

where:

$$\delta_j(\theta) = \frac{\max(e_j(\theta), e_j(\theta+\pi))}{\min(e_j(\theta), e_j(\theta+\pi))} - 1 \tag{4.8}$$

The value of $\delta_j(\theta)$ measures the difference of contrast between two opposing edges in relation to *(x, y)*. For a pattern that is perfectly symmetrical in relation to *(x, y)*, the function $\delta_j(\theta)$ is always nil, yielding to a minimum value *S(x, y)* of zero.

If *(x, y)* denotes the location of a well-defined elliptical ring, the values *E(x, y)*, *R(x, y)* and *S(x, y)* should therefore be small. The ERS transform is performed by combining numerically these three measures. Each pixel is given a rank n_E reflecting its position in the list of pixels sorted according to *E(x, y)*. The pixel with rank $n_E = 1$ represents the location with the lowest value of *E(x, y)* (*i.e.* the sharpest edges) found on the image. Similarly, each pixel is given a rank n_R and a rank n_S according to its value of *R(x, y)* and *S(x, y)* respectively. The three ranks are then combined to provide a list of pixels sorted according to the value of *n* defined as:

$$n = n_E + n_R + n_S \tag{4.9}$$

The pixel with the highest rank *n* represents the position of the most dominant elliptical ring within the image. This ranking mechanism ensures that there is no need to normalise and weight the three quantities *E(x, y)*, *R(x, y)* and *S(x, y)*. Since pixels are sorted according to each of the criteria independently, the final ranking does not depend on how rapidly the numerical expressions of *E*, *R*, and *S* increase.

4.2.2 Discrete approximation

Since the typical size of the elliptical patterns to be identified in CT is less than 12 pixels, it is possible to derive a precise approximation of them by considering only 8 points, taken along 8 principal directions, rather than considering all points of an ellipse. Figure (4.2) illustrates the eight chosen directions $\{d_i\}_{1 \le i \le 8}$ and how they intersect with an ellipse.

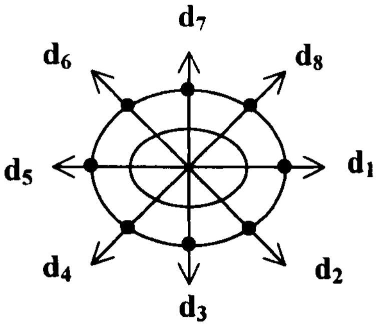


Figure 4.2. The directions $\{d_i\}_{1 \le i \le 8}$. At a small scale, an approximate description of an ellipse can be obtained by retaining the eight pixels at the intersection of the ellipse and the eight principal directions.

Edge information is derived from the original image by pre-computing the gradient along directions d_1, d_2, d_3, and d_4. The intensity gradient maxima and minima along directions $\{d_i\}_{1 \le i \le 8}$ are represented by 8-dimensional vectors $\boldsymbol{e_1} = (e_{1,i})_{1 \le i \le 8}$ and $\boldsymbol{e_2} = (e_{2,i})_{1 \le i \le 8}$ respectively, whereas the radial distribution of the edges are represented by vectors $\boldsymbol{r_1} = (r_{1,i})_{1 \le i \le 8}$ and $\boldsymbol{r_2} = (r_{2,i})_{1 \le i \le 8}$. The definitions of $E(x, y)$ and $R(x, y)$ are the discrete equivalent of those given in Equations (4.3) and (4.5). The measure of the symmetry of the strength of the edges $S(x, y)$, however, is simplified as:

$$S(x,y) = mean(\delta_{1,1}, \delta_{1,2}, \delta_{1,3}, \delta_{1,4}) + mean(\delta_{2,1}, \delta_{2,2}, \delta_{2,3}, \delta_{2,4}) \qquad \textbf{(4.10)}$$

where:

$$\delta_{j,k} = \frac{\max(e_{j,k}, e_{j,k+4})}{\min(e_{j,k}, e_{j,k+4})} - 1 \tag{4.11}$$

In the discrete implementation of the ERS transform, the ranking procedure involved is done using the quicksort algorithm [141]. On average, the computational cost of the discrete implementation of the ERS transform is $O(N \log N, r_0)$, with N being the number of pixels, and r_0 being the radius of the neighbourhood searched at each point.

As an example, Figure (4.3) demonstrates the pixels ranked according to E, R, and S for a synthetic image. Figure (4.3a) represents an image with elliptical rings and non-elliptical patterns with different sizes. The arrow highlights incomplete rings and partly occluded patterns. The SNR of this image is 44 dB, which is close to that of a typical HRCT scan. The locations surrounded by significant edges have high ranks n_E as shown in Figure (4.3b), signified by the brightness of the pixels. The results for $R(x, y)$ (uniformity) and $S(x, y)$ (symmetry) are illustrated in Figures (4.3c) and (4.3d) respectively. The incomplete patterns highlighted by the arrow on Figure (4.3a) are penalised by low ranks n_S, as shown in Figure (4.3d).

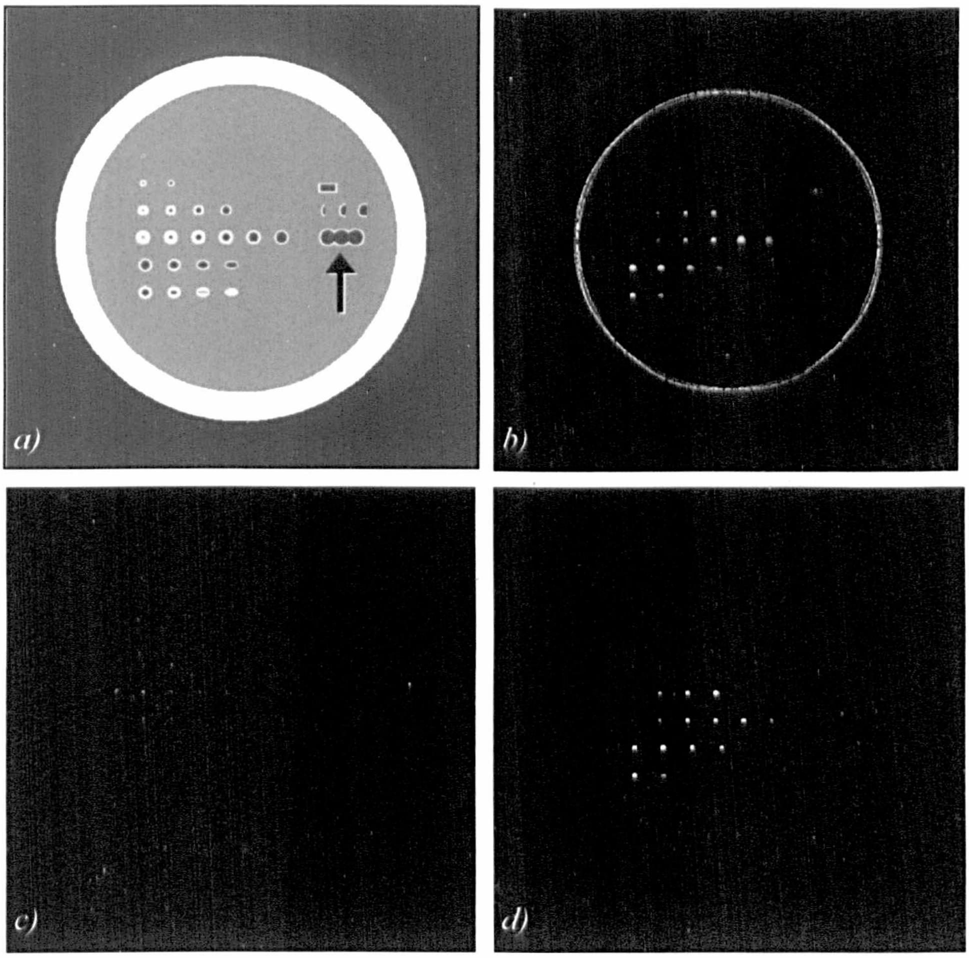


Figure 4.3. *a)* Original synthetic image showing various elliptical and non-elliptical patterns with different sizes. The arrow highlights incomplete and partly occluded patterns. Gaussian noise was superimposed to the image (SNR = 44dB). *b)* Ranks n_E derived from $E(x, y)$ (displayed only in the large inner circle). Pixels surrounded by strong edges have high ranks n_E and appear brighter. *c)* Ranks n_R derived from $R(x, y)$. The centres of the elliptical patterns are highlighted. *d)* Ranks n_S derived from $S(x, y)$. The partly occluded patterns have poor edge strength symmetry and appear darker.

4.2.3 Template matching

Once a bronchus has been identified with ERS transform, the adjacent pulmonary vessel, used primarily as a reference for assessing bronchial dilatation but also for gauging the degree of wall thickening, can be found by template matching in the neighbourhood of the bronchial wall. The tight constraints on the potential location and size of the vessel make the search space tractable and template matching a computationally acceptable solution. A blood vessel is modelled by a disc of CT density characteristic of blood surrounded by a circular ring of CT density characteristic of lung parenchyma. The position (x_V, y_V) and radius r_V of the model that best fits the image, in the neighbourhood of a bronchus, give a description of the accompanying pulmonary vessel. Figure (4.4), as an example, shows the segmentation of the bronchovascular bundle highlighted in Figure (1.8a).

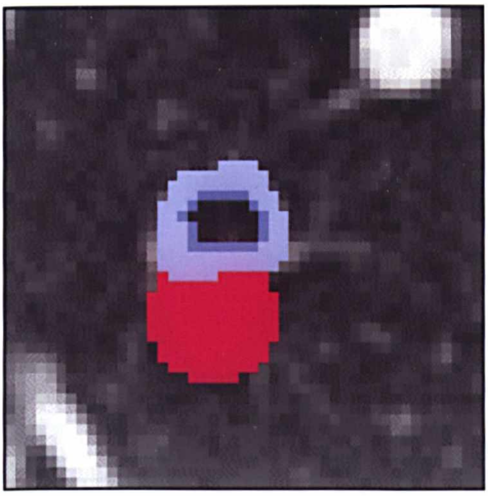


Figure 4.4. Segmentation of the bronchial wall (shown in blue) with ERS transform, and subsequent segmentation of the accompanying artery (shown in red) with template matching, for the airway shown in Figure (1.8a). Automated measurements of bronchial dilatation and bronchial wall thickening can be derived.

4.2.4 Measurement of bronchovascular abnormalities

In the absence of absolute reference values for the dimension of normal bronchus at a given point, detecting bronchial abnormalities is achieved by comparing the size of bronchus to the size of its accompanying artery. Numerically, a measure of bronchial dilatation is given by the ratio of bronchovascular diameters:

$$d = \frac{r_B}{r_V} \tag{4.12}$$

where r_B measures the external radius of a bronchus and r_V the radius of the accompanying vessel. In clinical studies, the severity of bronchial dilatation on CT is commonly scored on a 4-point scale [74]:

	$d \leq 1$	$1 < d \leq 2$	$2 < d \leq 3$	$d > 3$
Dilatation	0 (Normal)	1	2	3

Similarly, a measure of bronchial wall thickening is given by:

$$t = \frac{r_B - r_L}{r_V} \tag{4.13}$$

where r_L measures the radius of the lumen (internal radius) of the bronchus. Severity of bronchial wall thickening is scored on a 4-point scale, adapted from Reiff *et al* [74]:

	$t \leq 0.4$	$0.4 < t \leq 0.6$	$0.6 < t \leq 1$	$t > 1$
Wall thickening	0	1	2	3

These scales can also be simplified to two-point scales scoring the absence or presence of abnormal bronchial dilatation and thickening [24,25,140,142]. Indeed, merely detecting the presence of bronchial dilatation is a reasonably specific and sensitive sign for distinguishing, for instance, cases of obliterative bronchiolitis from normal subjects [24].

4.2.5 Patient study and radiologist scoring

To pilot the value of the method in a clinical context, the ERS transform was applied to CT scans of the lungs, which were also scored by an experienced radiologist. For 9 patients presenting suspected airways disease, 2 HRCT scans (1.5 mm collimation, sharp kernel reconstruction algorithm, as described in Chapter 3) were selected from pre-defined anatomical levels in the upper and lower zones of the lungs respectively, the areas of the lungs where major bronchi are most likely to run perpendicular to the plane of acquisition. The radiologist identified on each of these 18 cross-sectional images (viewed with a window width of 1500 H.U. and a level of -500 H.U [143]) the positions of the bronchi regarded as suitable for scrutiny for diagnostic purposes. Dilatation and thickening of each bronchus, on the 4-point scales described above, were also scored. Because of the documented high inter-observer agreement for the assessment of bronchi on HRCT scans [11], a single observer was considered appropriate.

4.3 RESULTS

4.3.1 Noise sensitivity

Since the proposed method is effectively based on the analysis of local maxima and minima of the first derivatives of the image intensity, it is necessary to evaluate its sensitivity to noise. A synthetic image, shown in Figure (4.3a), displaying various elliptical rings with different sizes was created to evaluate the noise-sensitivity of the edge-based technique. The intensity of the patterns is typical of the structures seen on HRCT scans of the lungs. The added Gaussian noise is similar to the noise inherent to the imaging modality. Within the first 100 locations determined by the algorithm, no false positive was detected, and all rings were identified. Measuring how the method copes with added Gaussian noise of higher variance shows that, for a signal-to-noise ratio (SNR) equal to or lower than 14 dB, no false positive was found amongst the first 100 results. With an SNR equal to 12 dB, 3 % of false negatives and 4 % of false positives occurred in the first 100 results.

4.3.2 Patient studies

On a set of 18 cross-sectional images, the radiologist found 123 bronchi significant in a clinical evaluation (minimum: 2 bronchi per image, maximum: 17 bronchi, average: 6.8 bronchi). The ERS transform was also applied and, for each scan, the list of the N first positions estimated by the algorithm was compared to the N positions determined by the human observer (N being case-dependent, and fixed for each case by the radiologist). Each location found both in the N-long list established by the radiologist and the N-long list established by the automated method was regarded as a case of agreement. Agreement with the ERS transform was reached in 76 instances, giving a rate of agreement of 62 %. Figure (4.5) shows an example of the compared results found by the human observer, as displayed by the crosses on Figure (4.5a), and the proposed algorithm, as displayed by the crosses on Figure (4.5b). In this case, agreement was reached in 9 out of 10 instances.

Without prior knowledge of the number of bronchi scored by the radiologist, the value of N should be fixed arbitrarily. The higher the value of N is, the greater the number b_l of bronchi correctly found is, providing more information for clinical assessment. However, when N increases, the rate of agreement a_l with the radiologists decreases, as shown by the solid lines in Figure (4.6). A trade-off can be obtained with a value of

$N = 4$. Thus, agreement rate a_1 is 75 %, and the technique correctly identifies $b_1 = 3$ clinically significant bronchi on average on every image.

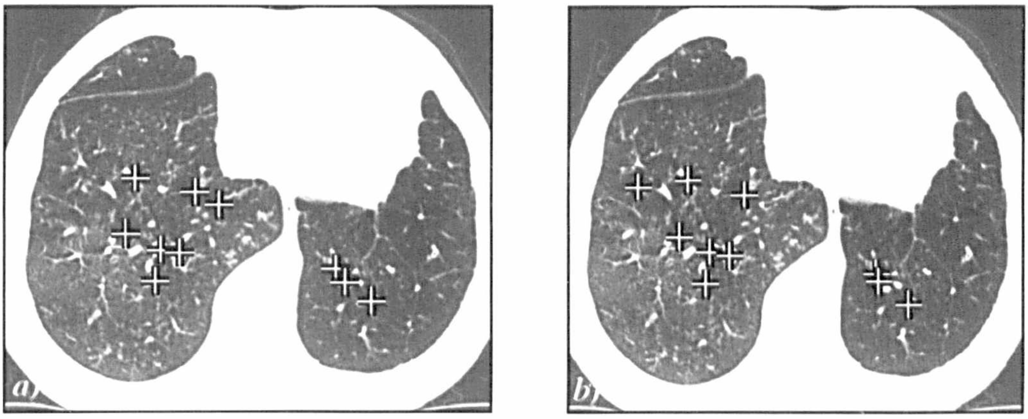


Figure 4.5. *a)* HRCT scan of the lower zone of the lungs. The most significant bronchi, as identified by an experienced radiologist, are marked with a cross. *b)* Bronchi found by the ERS transform on the same image. The automated method matches the results of the human observer in 9 out of 10 instances.

Because of high-level clinical knowledge, the radiologist may have discarded visible bronchi that he regarded as diagnostically irrelevant, resulting in cases of disagreement with the ERS transform. It is therefore important to assess the capacity of the proposed low-level feature extractor to detect elliptical rings representing bronchi, irrespective of their diagnostic value. The radiologist was asked to validate, *a posteriori*, which positions found by ERS transform actually represent near-perpendicular airways. As shown by the dotted line in Figure (4.6), the rate of agreement a_2 decreases while N increases, but is globally higher than agreement a_1. For a value of $N = 1$, agreement a_2 is 100 %, and for a value of $N = 4$, agreement a_2 is still greater than 83 %.

Template matching and automated bronchovascular measurement were performed at the locations given by the radiologist. Agreement between the radiologist and the automated technique for scoring these bronchi can be measured with the kappa statistic. When scoring the severity of bronchial dilatation, the value of the weighted kappa was $\kappa_{d,w} = 0.55$; when scoring the presence of abnormal bronchial dilatation the value of the kappa was $\kappa_d = 0.65$. Similarly, for the assessment of the severity of bronchial wall thickening, the weighted kappa is $\kappa_{t,w} = 0.51$ and the kappa measuring agreement in the detection of the presence of abnormal bronchial wall thickening was $\kappa_t = 0.58$. Similar results were obtained at the bronchial locations determined by the ERS transform, using a value $N = 4$. The kappa statistics were then: $\kappa'_{d,w} = 0.53$, $\kappa'_d = 0.63$, $\kappa'_{t,w} = 0.50$, $\kappa'_t = 0.55$. The automated technique and the radiologist agreed on the presence of abnormal bronchial dilatation in 81.6 % of bronchovascular bundles, and on the presence of abnormal bronchial wall thickening in 77.0 % of bronchovascular bundles.

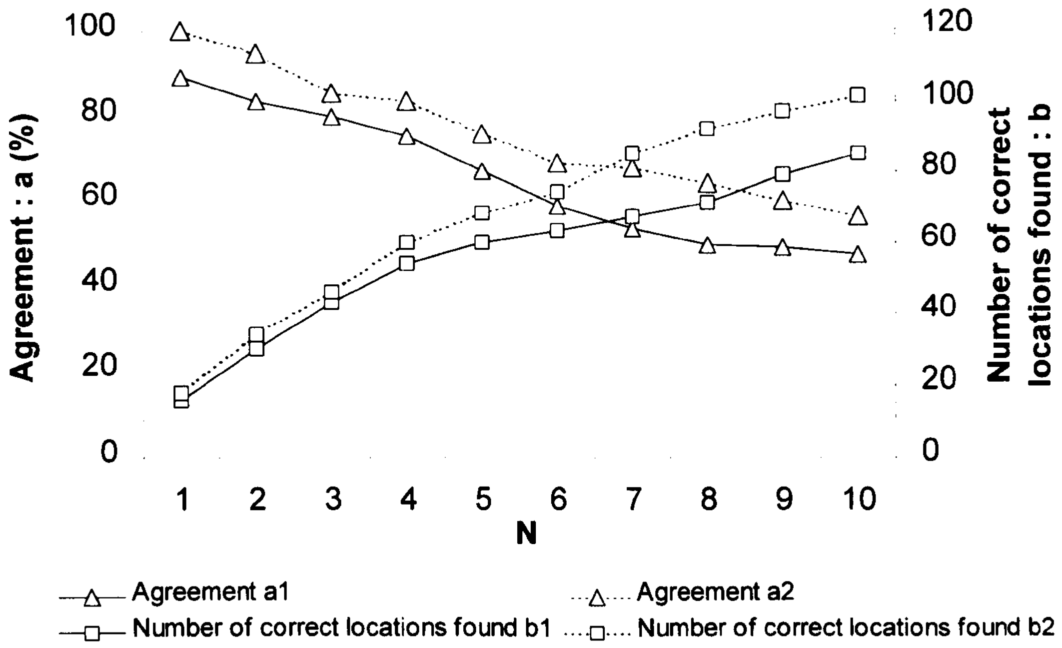


Figure 4.6. The agreement of the ERS transform for the detection of bronchi depends on the number *N* of locations retained for comparison with the scoring of the radiologist. This graph displays the rate of agreement for several values of *N*, when applying the algorithm to the 18 images of the patient study. For a small value of *N*, the rate of agreement is high, but little information can be derived from such a small number *b* of bronchi. For a high value of *N*, a large number *b* of bronchi are available for assessment, but the rate of agreement decreases. The solid lines show the values of the agreement rate a_1 and the number of correct locations b_1 found by ERS transform, when compared to the set of bronchi selected for their clinical value by a radiologist. The dotted lines show agreement a_2 and b_2 between the proposed algorithm and the radiologist, when no high-level clinical knowledge is taken into account. For a value of $N = 1$, the radiologist agreed that all locations found by ERS transform represent a near-perpendicular airway. Agreement a_2 decreases when *N* increases, but for a value of $N = 4$, the radiologist still agreed that more than 83 % of locations found by the proposed algorithm were indeed bronchi.

4.4 DISCUSSION

Because the elliptical rings being sought are defined by 6 parameters, the computational cost of methods like template matching and Hough transform (complexity $0(N^4)$) is considerable [77]. The implementation of the ERS transform compares favourably with these techniques in terms of computational cost.

The agreement rate of 62 % in a clinical context is satisfactory for a low-level feature extractor such as the ERS transform. The patient study showed that the bronchi identified by the algorithm possessed the expected geometric properties. When the human observer did not primarily choose them, it was probably due to high-level clinical knowledge. For example, the structures identified within consolidated lung or at a level near bifurcation are indeed bronchi, but are discarded by an experienced observer because the reliable interpretation of airways in these situations is problematic. When omitting such radiological expertise, which is out of the scope of the proposed algorithm, agreement was shown to be greater than 83 % for a value of $N = 4$. Most cases of disagreement were then due to the "twinkling star" artefact [144], a focal low attenuation artefact immediately adjacent to vessels that can mislead inexperienced observers. Nevertheless, using the ERS transform, detection of bronchi is possible in a clinical situation. In a computerised diagnosis-aid system, multiple feature extractors are used. Their outputs are combined in a statistical high-level framework incorporating *a priori* knowledge, in order to reach valid clinical conclusions. In this context, the agreement rate of 62 % between the proposed method and a human observer suggests satisfactory statistical significance. In the absence of a prior knowledge of the potential total number of bronchi that could be detected, it would be possible to retain the 4 first bronchi identified by ERS transform on every image. The agreement rate then becomes

75 % and clearly indicates the potential of the method in a lung CT expert system. The purpose of the algorithm used in this context is not to sample all visible bronchi, but to locate and derive diagnostic information from a significant set of bronchi.

When scoring the presence and severity of bronchial abnormalities, the kappa statistics (all greater than 0.5) suggested moderate to good agreement with the radiologist. Such inter-observer variability is similar to that encountered between experienced observers [11,24].

4.5 CONCLUSION

In this chapter, a novel technique for the detection of elliptical rings on HRCT images was presented. The ERS transform is based on edge analysis within local polar co-ordinates, and relies on the geometric properties of the patterns to be identified. It was shown to perform robustly in the presence of noise. The detection of bronchi on HRCT scans of the lungs was also evaluated. Comparison with the assessment made by a highly experienced human observer showed that the agreement for the detection of the location of bronchi was satisfactory. It indicates the potential value of incorporating the proposed technique in a medical imaging expert system. For automatically scoring the presence and severity of bronchial abnormalities, the agreement between the proposed method and an experienced observer was found to be comparable to the variability measured between experienced human observers.

CHAPTER 5

TEXTURE CLASSIFICATION OF THE LUNG PARENCHYMA

5.1 INTRODUCTION

The main HRCT characteristic used for the detection of obstructive lung diseases is the presence of areas of abnormally low attenuation in the lung parenchyma, which can be detected automatically, as shown in Chapter 3 [30,139]. The automated analysis of bronchial morphology with the ERS transform presented in Chapter 4 [31,145], is also of diagnostic value. However, these findings do not always permit confident diagnosis and the correct separation between different causes of obstructive lung disease. To address this problem, it is necessary to take into account the textural appearances of the lung parenchyma. The aim of this chapter is to present a fully automated method for the differentiation of centrilobular emphysema, panlobular emphysema, constrictive obliterative bronchiolitis, and normal lungs, on the basis of HRCT texture features alone. The proposed method involves an automated segmentation of the main anatomical structures - including the chest wall, lung parenchyma and macroscopic pulmonary vessels - followed by Bayesian classification. The technique was evaluated with images acquired from normal subjects and patients showing obstructive lung diseases.

5.2 MATERIALS AND METHODS

Automated texture classification involves a process of supervised learning of the textural characteristics of each class. The samples used for supervised learning were obtained from a set of CT images of normal subjects ($n = 11$), patients with panlobular emphysema ($n = 7$), centrilobular emphysema ($n = 11$) and constrictive obliterative bronchiolitis ($n = 15$). Images were acquired using a standard HRCT protocol, as described in Chapter 3. The visual characteristics for each of the four classes of images were illustrated in Chapter 1.

Automated segmentation of the main anatomical structures was performed to remove the macroscopic pulmonary vessels of the lung parenchyma, with the structure-filtering operator based on mathematical morphology presented in Chapter 3 [30]. This step is important in that the macroscopic structures, such as macroscopic pulmonary vessels, have a typical size that matches that of the ROIs, and a significant statistical description cannot be obtained by textural extractors. Performing automated anatomical segmentation allowed for the textural analysis of the finest structures of the lung parenchyma. For each case, an experienced radiologist then selected 4 circular ROIs (radius = 22 pixels), each categorised as showing textural characteristics typical of one of the three pathologies or normality, at five pre-defined anatomical levels (origin of the great vessels, tracheal carina, pulmonary venous confluence, one centimetre above the right hemi-diaphragm, and midway between the two previous cross-sections). This provided a total number of $n_{train} = 880$ ROIs for training. On an adjacent non-contiguous section, the radiologist also selected 4 ROIs to provide another $n_{test} = 880$ different ROIs for testing. For each ROI, a statistical descriptor was derived in the form of an N-dimensional vector v. Each vector v contained the values of $N = 13$ textural features chosen to describe the CT densities in the ROI. These features included n^{th}-order statistics of the distribution of the CT values: mean, standard deviation, skewness and kurtosis.

Other features describing spatial dependence of greyscale distributions were derived from the set of co-occurrence matrices computed at each ROI [83]. Each ROI was approximated on $p = 16$ level greyscale. A set of 20 matrices $\{C_{d,l}\}_{1 \le d \le 4,\ 1 \le l \le 5}$ was then derived, representing the co-occurrence of grey levels along 4 directions d, at distances ranging from $l = 1$ pixel to $l = 5$ pixels. They characterise the spatial relationships of grey levels in textural patterns, in a way that is invariant under monotonic grey-level transformations. To reduce the dimensionality of the feature vector $\boldsymbol{v}$, five scalar measurements were extracted from each co-occurrence matrix: *energy*, *entropy*, *maximum*, *contrast* and *homogeneity*. From a matrix $\mathbf{C} = (C_{i,j})_{1 \le i \le 4,\ 1 \le j \le 4}$, they are computed as follows [83]:

$$\text{energy}(\mathbf{C}) = \sum_{i,j} C_{i,j}^2 \tag{5.1}$$

$$\text{entropy}(\mathbf{C}) = -\sum_{i,j} C_{i,j} \log C_{i,j} \tag{5.2}$$

$$\text{maximum}(\mathbf{C}) = \max_{i,j} C_{i,j} \tag{5.3}$$

$$\text{contrast}(\mathbf{C}) = \sum_{i,j} |i-j|^{\alpha} C_{i,j}^{\beta} \quad \text{with } \alpha = 1, \beta = 1 \tag{5.4}$$

$$\text{homogeneity}(\mathbf{C}) = \sum_{i,j} \frac{C_{i,j}}{1+|i-j|} \tag{5.5}$$

The mean value of these 5 parameters, over the 20 co-occurrence matrices, were incorporated in the feature vector $\boldsymbol{v}$ describing a ROI.

Since co-occurrence matrices do not capture the shape aspects of the grey level primitives, run-length parameters were also computed at each ROI, providing another set of features to be included in $\boldsymbol{v}$. A primitive is a maximum contiguous set of constant grey level pixels located in a line [85]. For each ROI approximated on a $p = 16$ level greyscale, the number $B(a, r)$ represents the total number of primitives of the length r and grey-level a. Textural description can be summarised by deriving short primitive emphasis *spe*, long primitive emphasis *lpe*, grey-level uniformity *glu* and primitive length uniformity *plu*. Short primitive emphasis *spe* measures the predominance of short primitives in a textural pattern. A high *spe* value denotes a pattern consisting mainly of short lines with a constant grey-level. Short primitive emphasis can be computed with the following formula which gives higher weight to the shorter primitives:

$$spe = \frac{1}{B_{tot}} \sum_{a=1}^{p} \sum_{r=1}^{r_{max}} \frac{B(a,r)}{r^2} \qquad \textbf{(5.6)}$$

where r_{max} denotes the maximum primitive length in the ROI, and B_{tot} is the total number of runs:

$$B_{tot} = \sum_{a=1}^{p} \sum_{r=1}^{r_{max}} B(a,r) \qquad \textbf{(5.7)}$$

Similarly, long primitive emphasis *lpe* measures the predominance of long primitives in a textural pattern, and its value is given by:

$$lpe = \frac{1}{B_{tot}} \sum_{a=1}^{p} \sum_{r=1}^{r_{max}} B(a,r) r^2 \qquad \textbf{(5.8)}$$

Grey-level uniformity *glu* measures grey-level dispersion of the primitives. A high *glu* value denotes a textural pattern where primitives belong to a small number of grey levels, as in a chequerboard for instance. Grey-level uniformity is computed as follows:

$$glu = \frac{1}{B_{tot}} \sum_{a=1}^{p} \left(\sum_{r=1}^{r_{max}} B(a,r) \right)^2 \tag{5.9}$$

Primitive length uniformity *plu* measures the similarity in length of the primitives. A high *plu* value denotes a pattern where all primitives have a roughly equal length. Primitive length uniformity is computed as follows:

$$plu = \frac{1}{B_{tot}} \sum_{r=1}^{r_{max}} \left(\sum_{a=1}^{p} B(a,r) \right)^2 \tag{5.10}$$

In summary, each ROI was characterised by a 13-dimensional vector ***v*** containing the values given in Table (5.1). The validity of these descriptors for capturing textural information has been demonstrated in a wide range of image understanding applications [146]. The use of fractal dimensions has also been suggested [93], but was discarded here because of its computational cost.

n^{th} order statistical moments	• mean • standard deviation • skewness • kurtosis
Mean values extracted from 20 co-occurrence matrices	• energy • entropy • maximum • contrast • homogeneity
Run-length parameters	• short primitive emphasis • long primitive emphasis • grey-level uniformity • primitive length uniformity

Table 5.1. Summary of the 13 textural descriptors representing each ROI.

To perform automated discrimination of the samples, a Bayesian classifier was implemented since conditional probabilities provide a suitable framework for handling uncertainty in medical decision-support systems. Bayesian probabilities could also allow for combining the output of the classifier with the output of other low-level image feature extractors [132], in order to refine the diagnostic process. The 4 classes of samples were labelled as k = 0, 1, 2, 3 (denoting samples from cases of, respectively, centrilobular emphysema, panlobular emphysema, constrictive obliterative bronchiolitis and normal subjects). Numerical description of the classes was contained in the 4 sets of training samples $\{S_k\}_{0 \leq k \leq 3}$. Each set S_k contained n_k vectors v_i. Based on first-order and second-order statistical moments of each set S_k, a pattern classifier could be established. In this study, it is assumed that the distribution of feature vectors in each class could be modelled by an N-dimensional normal distribution. This is a reasonable assumption for modelling a stochastic process that shows some subjective unifying property, but with no underlying theoretical model. It has been demonstrated that Bayesian classifiers using such a probabilistic model were robust, even when the assumption on class-specific probabilities being normally distributed was considerably violated [147]. Each N-dimensional normal distribution $\Psi_k(v)$ can be written as:

$$\Psi_k(\mathbf{v}) = \frac{1}{\sqrt{(2\pi)^N \det \mathbf{K}_k}} \exp\left[-\frac{1}{2}(\mathbf{v} - \boldsymbol{\mu}_k)^T \mathbf{K}_k^{-1} (\mathbf{v} - \boldsymbol{\mu}_k) \right] \qquad \textbf{(5.11)}$$

where μ_k represents the mean of vectors in class k:

$$\boldsymbol{\mu}_k = \frac{1}{n_k} \sum_{\mathbf{v}_i \in S_k} \mathbf{v}_i \qquad \textbf{(5.12)}$$

and $\boldsymbol{K}_k$ is the covariance matrix of vectors in class k:

$$\mathbf{K}_k = \frac{1}{n_k - 1} \sum_{\mathbf{v}_i \in S_k} (\mathbf{v}_i - \boldsymbol{\mu}_k)(\mathbf{v}_i - \boldsymbol{\mu}_k)^T \tag{5.13}$$

Each distribution is unimodal, exhibiting one single peak at μ_k, and decays in all directions according to the quadratic form:

$$Q_k(\mathbf{v}) = (\mathbf{v} - \boldsymbol{\mu}_k)^T \mathbf{K}_k^{-1} (\mathbf{v} - \boldsymbol{\mu}_k) \tag{5.14}$$

The mean vectors μ_k and the symmetric positive definite covariance matrices $\boldsymbol{K}_k$ completely define the distributions Ψ_k, and could be derived from the sets of samples $\{S_k\}_{0 \leq k \leq 3}$, using Equations (5.12) and (5.13). The inverse and the determinant of each matrix $\boldsymbol{K}_k$ were obtained through LU-decomposition [148]. Following the assumption that the distribution of vectors in each class is normal, a Bayes classifier could be implemented, with the k class-specific probabilities being equal to the distribution $\Psi_k(\boldsymbol{v})$:

$$\text{For all } k\text{: } prob(\mathbf{v}|k) = \Psi_k(\mathbf{v}) \tag{5.15}$$

It is important to stress that the *a priori* probability for each class could not be computed from the sets of samples $\{S_k\}_{0 \leq k \leq 3}$, using the following relationship:

$$\text{For all } k\text{: } prob(k) = \frac{n_k}{\sum_k n_k} \tag{5.16}$$

As a matter of fact, Equation (5.16) relies on the assumption that the cardinal of each of the training sets reflects the prevalence of each of the 4 classes of disease. However, there is no evidence supporting such an assumption. Here it was assumed that all *a priori* probabilities were equal:

$$\text{For all } k\text{: } prob(k) = \frac{1}{4} \tag{5.17}$$

Once the vectors μ_k and the matrices $\boldsymbol{K}_k$ had been defined using the training sets $\{S_k\}_{0 \leq k \leq 3}$, each new testing samples $\boldsymbol{v}$ could be classified as belonging to class k_0, using the minimum error rate decision from the Bayes decision rule [149]:

$$\text{Decide for } k_0 : \text{if } prob(k_0|\mathbf{v}) = \max_k\left(prob(k|\mathbf{v})\right) \tag{5.18}$$

where:

$$\begin{aligned} prob(k|\mathbf{v}) &= \frac{prob(\mathbf{v}|k)prob(k)}{prob(\mathbf{v})} \\ &= \frac{\Psi_k(\mathbf{v})prob(k)}{\sum_k \Psi_k(\mathbf{v})prob(k)} \end{aligned} \tag{5.19}$$

In the case of equal *a priori* probabilities for all classes, as shown in Equation (5.17), Equation (5.19) can be simplified as:

$$prob(k|\mathbf{v}) = \frac{\Psi_k(\mathbf{v})}{\sum_k \Psi_k(\mathbf{v})} \tag{5.20}$$

The *a posteriori* probability $prob(k_0|v)$ gave a measure of the confidence of the classifier in assigning the label k_0 to a sample $\boldsymbol{v}$. To model the lack of confidence of the classifier, samples could be rejected and remain unclassified if their *a posteriori* probability was lesser than a predefined threshold $p_{threshold}$:

$$\text{Reject } v \text{ if} : prob(k_0|\mathbf{v}) < p_{threshold} \tag{5.21}$$

The value of $p_{threshold}$ could be selected so as to maximise the sensitivity and specificity of the classifier used as a diagnostic test, while retaining a sufficient number of reclassified samples. Since it was not feasible to obtain histopathologic confirmation of the dominant disease process in each of the ROIs, the diagnostic assigned by the experienced radiologist for each ROI was regarded as a "gold-standard" for the purpose of this study.

5.3 RESULTS

When setting no probability threshold $p_{threshold}$, all 880 ROIs of the testing set were classified, irrespective of the confidence of the classifier. A total of 619 ROIs were then labelled correctly, giving an overall sensitivity of 70.3 % and overall specificity of 89.3 %. When setting a value of $p_{threshold} = p_1$ such that half the samples (440 ROIs) were confidently reclassified, and half the samples were rejected, 350 ROIs were correctly classified, giving an overall sensitivity of 79.5 % and overall specificity of 93.2 %. Sensitivity, specificity and positive predictive value (PPV) for each of the four classes are given in Table (5.2). With a threshold $p_{threshold} = p_2$ such that a fourth of samples are reclassified, 200 out of 220 ROIs were labelled correctly, giving overall sensitivity and specificity of, respectively, 90.9 % and 96.7 %, as shown in Figure (5.1) which gives the overall sensitivity and specificity for several values of $p_{threshold}$. Globally, setting a higher value of $p_{threshold}$ improves the accuracy of the algorithm, but leaves a larger number of samples unclassified. The sensitivity and specificity for each class for several values of $p_{threshold}$ are shown in Figure (5.2).

	Sensitivity	Specificity	PPV
Centrilobular emphysema	78.7 %	97.3 %	82.8 %
Panlobular emphysema	61.2 %	97.4 %	85.3 %
Constrictive obliterative bronchiolitis	73.0 %	91.2 %	79.4 %
Normal	96.6 %	84.8 %	77.0 %

Table 5.2. Class-specific sensitivity, specificity and positive predictive value (PPV).

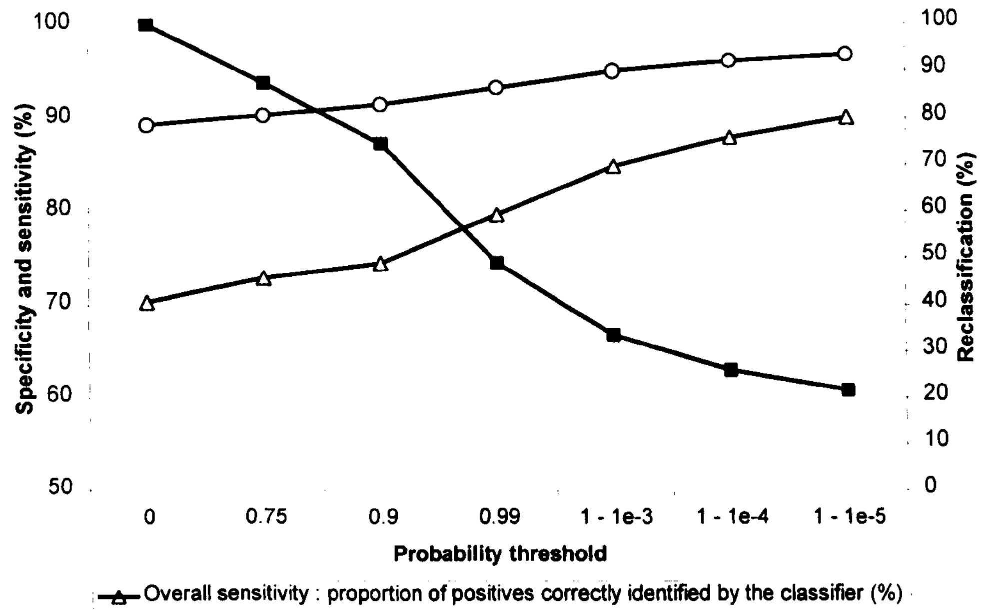


Figure 5.1. Relationship between the threshold $p_{threshold}$ on *a posteriori* probabilities and the overall sensitivity and specificity of the classifier. As $p_{threshold}$ increases, and only samples with increasing confidence are classified, the overall sensitivity and specificity of the technique improve. However, a larger proportion of samples are then rejected and remain unclassified. A trade-off can be obtained with a value of $p_{threshold}$ such that half the samples are confidently reclassified, for instance. Corresponding overall sensitivity and overall specificity are respectively 79.6 % and 93.2 %.

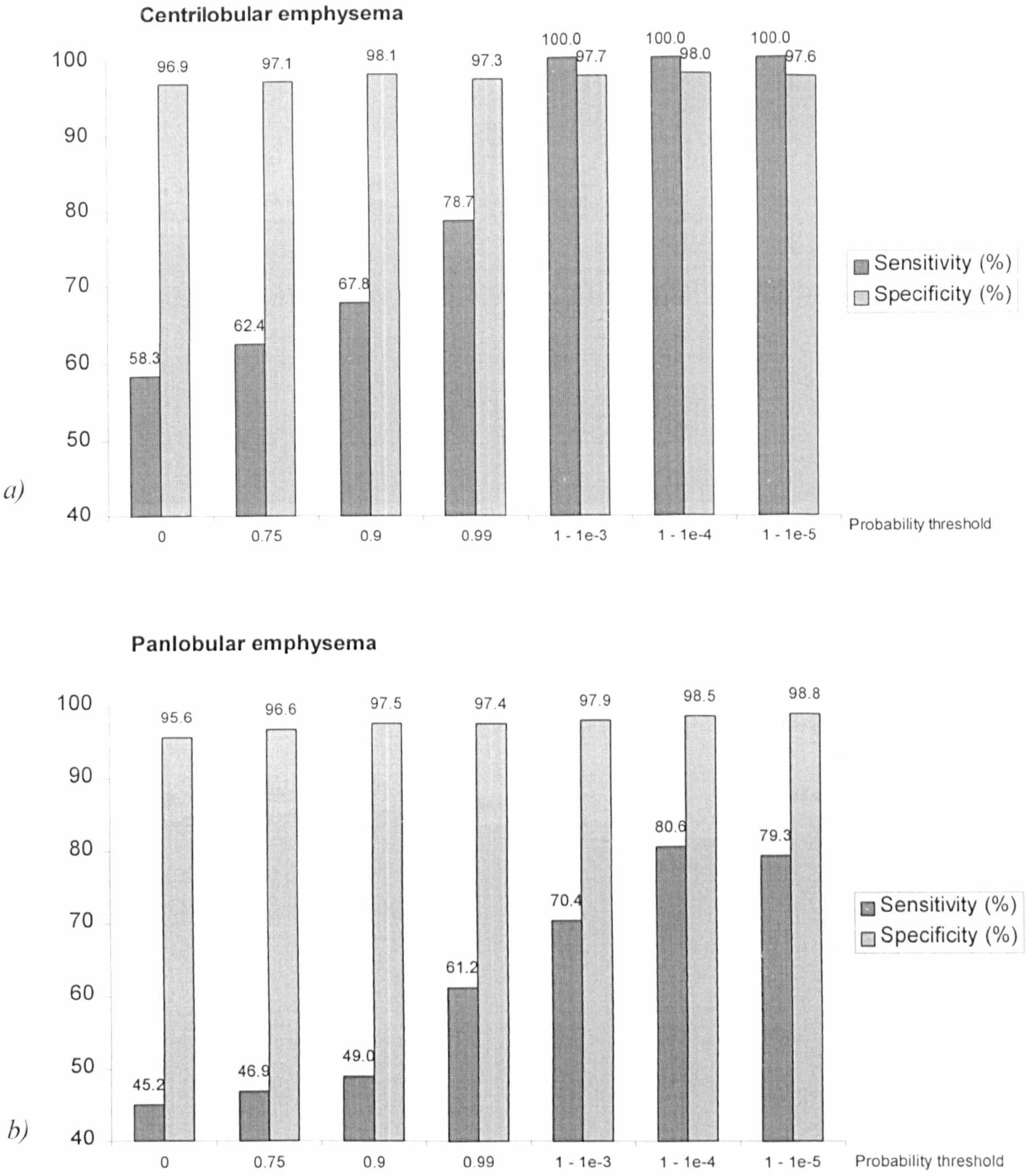
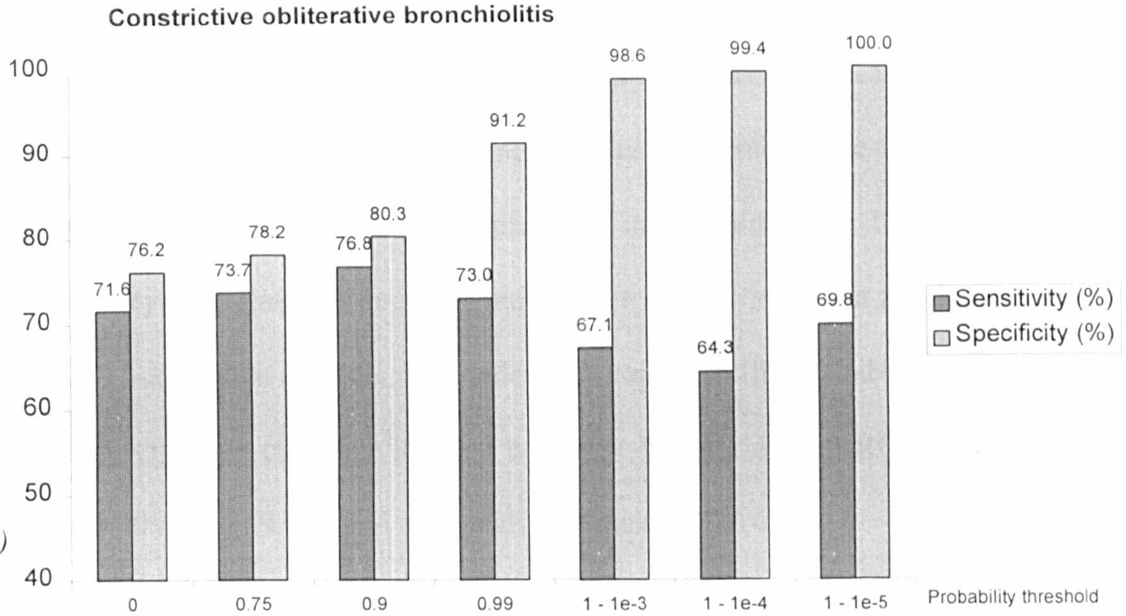


Figure 5.2. *a)* Class-specific sensitivity and specificity of centrilobular emphysema, for several values of $p_{threshold}$. *b)* Class-specific sensitivity and specificity of panlobular emphysema, for several values of $p_{threshold}$.

c)



d)

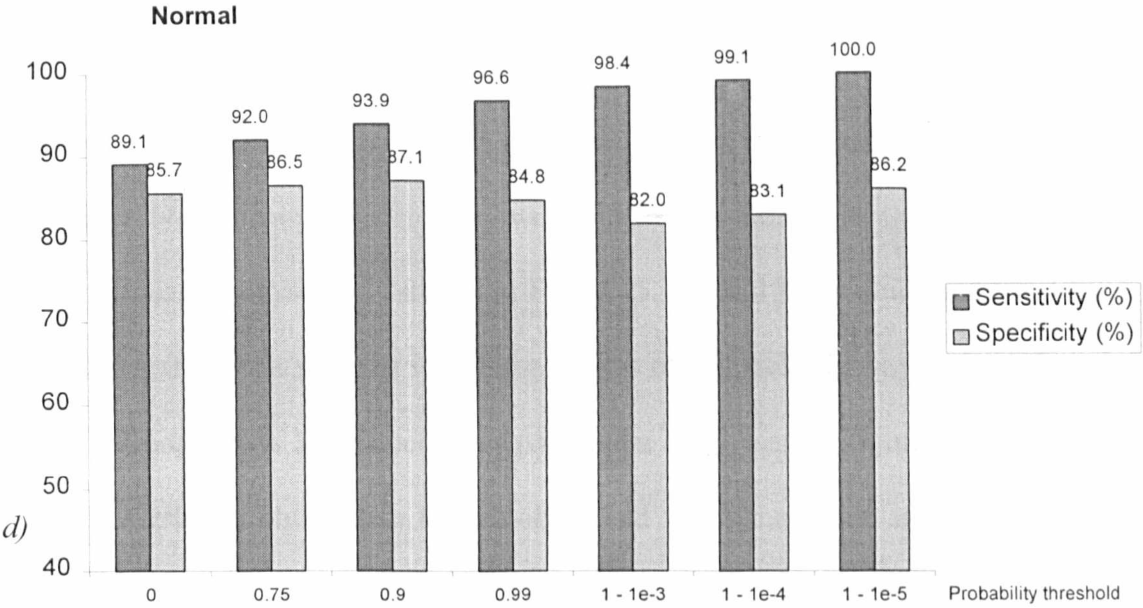


Figure 5.2 ***(continued).*** *c)* Class-specific sensitivity and specificity of constrictive obliterative bronchiolitis, for several values of $p_{threshold}$. *d)* Class-specific sensitivity and specificity of normals, for several values of $p_{threshold}$. The best accuracy is obtained with samples from normal cases and cases of centrilobular emphysema. Lesser sensitivity is obtained with cases of panlobular emphysema and constrictive obliterative bronchiolitis, which are conditions with similar textural appearances.

To evaluate the performance of the proposed technique, comparison with a minimum-distance classifier was undertaken. A minimum distance classifier measures the distance of a tested sample to each of the training samples. It classifies the tested samples as belonging to the class of the closest sample found in the training set. The key difference between a Bayesian classifier and minimum-distance classifier is that the latter does not rely on any assumption regarding the underlying class-specific probability distributions. A minimum-distance classifier labelled correctly 438 out of 880 ROIs, achieving an overall sensitivity of 49.8 % and an overall specificity of 76.6 %.

As an illustration, Figure (5.3) shows how the proposed Bayesian classifier can be applied, not only to user-selected ROIs, but also to the whole parenchyma for textural segmentation. A normal case is shown in Figure (5.3a), and the result of the classification is shown in Figure (5.3b). All the samples confidently classified, with *a posteriori* probabilities greater than p_l, were labelled as normal. A case showing signs of centrilobular emphysema is shown in Figure (5.3c), and the segmentation in Figure (5.3d) shows that most samples were classified as centrilobular emphysema, with some areas of homogeneous lung labelled as panlobular emphysema. Figure (5.3e) shows a case of constrictive obliterative bronchiolitis, and its segmentation is given in Figure (5.3f). An example of panlobular emphysema is shown in Figure (5.3g), and the corresponding segmentation is shown in Figure (5.3h).

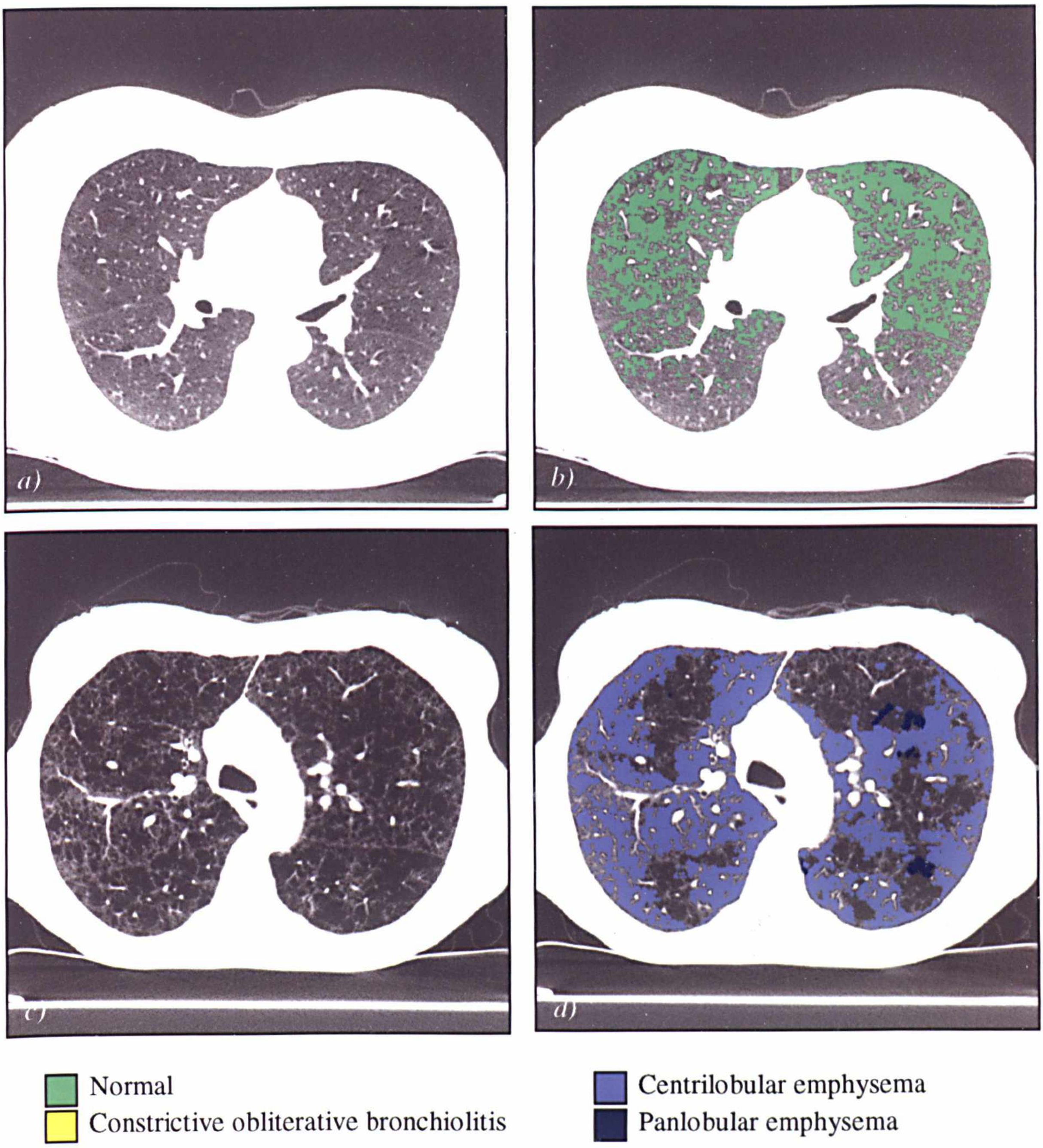


Figure 5.3. Application of the classifier to the segmentation of the lung parenchyma. At pixels where the confidence of the classifier is low (smaller than the p_l threshold) no label is assigned. *a)* Scan of a normal subject (window level = -800 H.U. / width = 1000 H.U.) *b)* Automated classification. The samples confidently classified are labelled as normal. *c)* Scan of a subject showing signs of centrilobular emphysema. *d)* Most of the classified samples are labelled as centrilobular emphysema. Small areas of homogeneous hypo-attenuated lung are classified as panlobular emphysema.

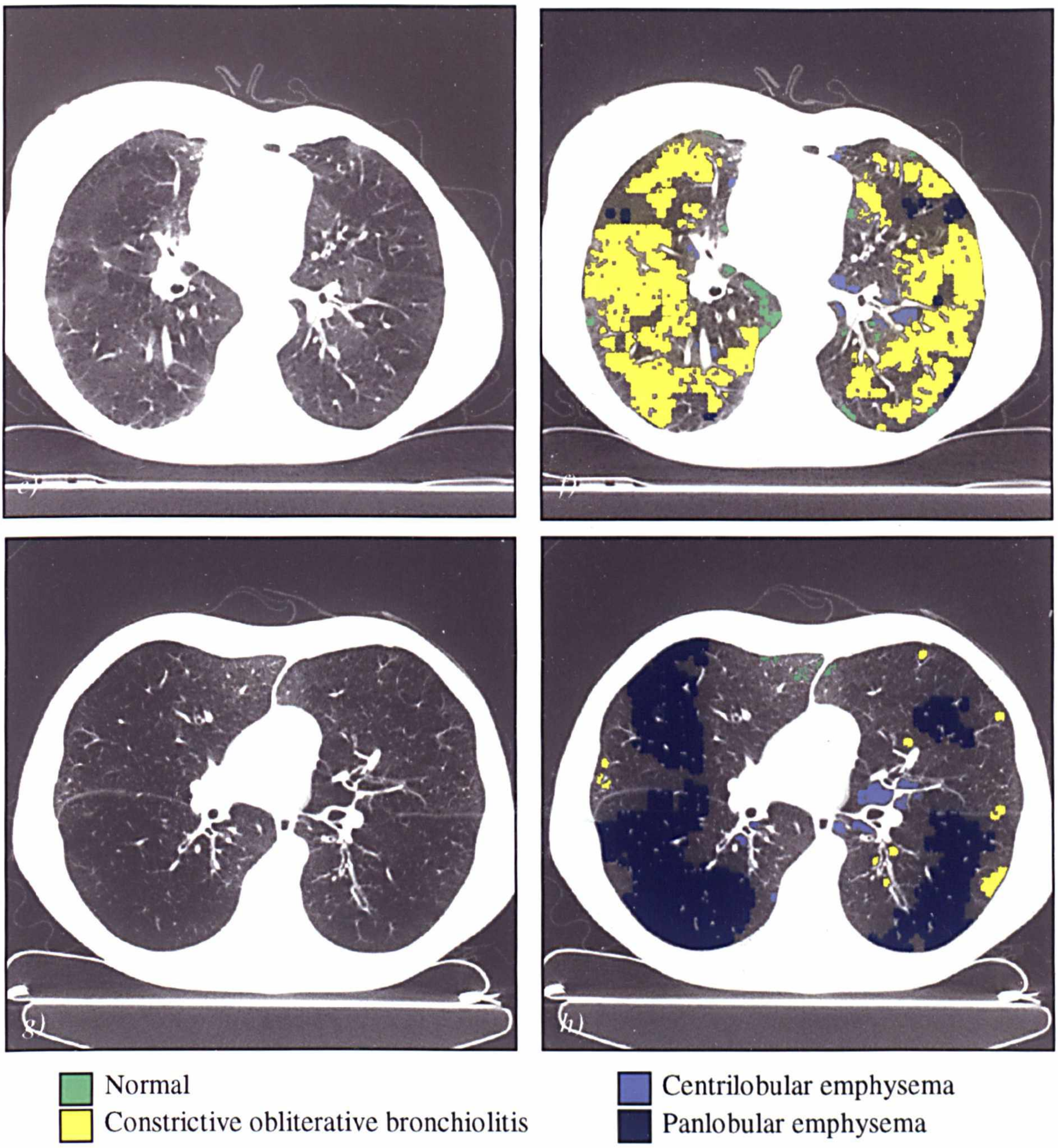


Figure 5.3 ***(continued)***. *e)* Scan of a patient with constrictive obliterative bronchiolitis. *f)* Most of the parenchyma is labelled as constrictive obliterative bronchiolitis. High-density structures adjacent with hypo-attenuated lung results in the classifier identifying the texture as centrilobular emphysema, especially in the vicinity of the major bronchi. *g)* Scan of a patient with panlobular emphysema. *h)* Most of the parenchyma is labelled as panlobular emphysema. As in Figure (5.3f), some areas are labelled as centrilobular emphysema near the major bronchi.

5.4 DISCUSSION

The method described in this chapter for the textural differentiation of obstructive airways diseases distinguishes between normality and three types of obstructive lung diseases with considerable accuracy. It should be emphasised that the proposed classifier is based only on the consideration of parenchymal texture. Other CT findings normally used for differential diagnosis (such as the regional distribution of the disease, the state of the airways, and the size and disposition of the pulmonary vessels) were not incorporated in this feature extractor. Nevertheless, the method presented achieves a high level of discrimination.

There have been previous examples of the automated quantification of emphysema on chest radiography [89], but HRCT is clearly a superior imaging modality for the detection of subtle alterations of the lung parenchyma architecture [15]. Textural analysis of CT images has shown that high sensitivity (94.8 % on average) can be achieved with a binary classifier for the differentiation between emphysematous and normal lung [90]. Other binary classifiers have been proposed for the textural differentiation of normal versus abnormal lung for cases of interstitial diseases [91] and ground-glass opacification of the lung [92,93]. An example of textural classifier handling more than two classes showed limitations because, unlike the classifier presented in this chapter, it was not applied to images on which the macroscopic anatomical structures have been segmented [94]. To date, there have been no reported examples of classifiers designed to distinguish between various causes of hypo-attenuated lung and normal lung.

In the present study, training and testing of the classifier were based on the visual classification made by an experienced radiologist. The reliability of subjective visual assessment for this classification is open to question. However, other methods such as pulmonary function tests do not reliably distinguish between different pathologic causes of obstructive lung diseases. Lung biopsy is not performed routinely, either because patients have such severe disease that surgery is precluded, or because patients have such mild disease that an invasive procedure is not warranted. Furthermore, there are regional differences in the predominant pathology, such that in patients with emphysema, areas of centrilobular and panlobular may coexist in adjacent parts of the lungs. Despite its limitations, CT evaluation by an experienced radiologist is, in this context, a reasonable "gold-standard". With this *caveat*, the proposed classifier achieves high sensitivity and specificity (of the order of 80 % and 90 % respectively). Also, evaluation of the classifier with the training set demonstrated that the proposed technique allows for correct encapsulation and discrimination of the textural information contained in the training set. Irrespective of the diagnostic reliability of subjective assessment, it can be claimed that the method classifies ROIs in a way that is consistent with the experienced observer who provided the training data.

The assumption that feature vectors in each class are normally distributed can also be questioned. However, performance comparison with a minimum distance classifier that makes no assumption about the model of the probability densities showed that the proposed method is less sensitive to sampling noise.

In this study, the largest number of cases of misclassification resulted from confusion between cases of panlobular emphysema and constrictive obliterative bronchiolitis, the classes with poorest sensitivity, as shown in Figure (5.2). This is not surprising, given the similar visual appearances of the lung parenchyma on CT for these two conditions. Nevertheless, the general diagnostic value of the proposed classifier was demonstrated, with high overall sensitivity and specificity.

The size of the ROIs (radius of 22 pixels) was established in collaboration with an experienced radiologist who was asked to select ROIs characteristic of the CT appearance of the four pathologies. An example of textural classifier applied to CT images relied on smaller ROIs (5x5 pixels [94]). However, in that situation, the purpose of texture-based image analysis was to perform full anatomic segmentation, and the size of the ROIs needed to match the typical size of anatomical structures such as macroscopic pulmonary vessels or bronchial lumina. Anatomical segmentation had already been performed on the images analysed by our classifier, and the ROIs encompassed only the lung parenchyma.

Thresholding *a posteriori* probabilities allows for tuning the sensitivity and specificity of the classifier according to the relative cost of false positives and false negatives. A probabilistic model is well suited for integration in a fully automated decision-support system. Automatically selecting ROIs necessarily produces samples that belong to none of the four categories the classifier was trained for, and it is then important that the low confidence of the classifier can be measured. Bayesian probabilities offer a good paradigm for measuring such uncertainties.

5.5 CONCLUSION

We have demonstrated that textural distinction between several pathologies causing hypo-attenuation of the lung parenchyma is feasible. A description of the spatial distribution of CT values with statistical moments, co-occurrence matrices and run-length parameters seems to provide sufficient information for a pattern classifier. Under the assumption that the distribution of feature vectors for each class is normal, Bayes decision rule allows for the implementation of a classifier with high sensitivity and specificity. By selecting a suitable threshold on the *a posteriori* probabilities, classification of samples characteristic of the CT appearance of three related lung pathologies and normality can be achieved with a sensitivity and specificity greater than 80 % and 90 % respectively. The accuracy of the method is good, suggesting its value as one of the main CT feature extractors for the automated detection of obstructive airways diseases.

CHAPTER 6

PROBABILISTIC NETWORK FOR DECISION-SUPPORT

6.1 INTRODUCTION

The low-level algorithms presented in Chapters 3, 4 and 5 extract quantitative and semi-quantitative measurements of HRCT features of obstructive lung diseases (extent of hypo-attenuated lung, severity of bronchial dilatation and bronchial wall thickening) which can be used to draw the attention of the clinician to specific areas of interest (*i.e.*, areas of hypo-attenuation, abnormal parenchymal texture [150], abnormal bronchial morphology) and so aid the diagnostic process. The use of each technique alone has obvious limitations. The purpose of this chapter is to combine several classifiers in order to resolve ambiguities, and improve the overall accuracy of the features extractors for the detection of abnormal findings. To allow for both probabilistic reasoning in medical decision-support [151], as reviewed in Chapter 2, and for data fusion in computer vision [133], probabilistic networks have been found to offer a suitable framework. Through supervised learning, *a priori* high-level knowledge can be systematically encapsulated in the form of conditional probabilities. This chapter describes how the output of the image processing algorithms presented in Chapter 2, 3 and 4 can be combined through a probabilistic network, in order to reach a diagnostic conclusion.

6.2 PROBABILISTIC NETWORKS

One of the first methods based on the propagation of uncertain knowledge through an inference network was developed by Duda *et al*, for the *Prospector* expert system for geological exploration [152]. An inference network represents a set of collection rules as a graph structure. Each node represents a proposition, and each arc connecting the nodes represents an inference rule. The formalism presented in *Prospector* is restricted: the network structure needs to represent a directed tree, there cannot be multiple paths from an evidence node to other nodes or hypotheses. Also, it requires the definition of interpolation functions, with an *ad hoc* design tailored for every application. Pearl introduced a more solid theoretical framework, in the form of *Bayesian networks* (also referred to as *probabilistic networks*, *causal networks*, or *belief networks*), a graphical representation of probabilistic dependencies for probabilistic reasoning in expert systems. Bayesian networks are directed acyclic graphs (DAGs) in which the nodes represent variables, the arc signify the existence of causal direct influences between the linked variables, and the strength of these influences are expressed by forward conditional probabilities. Traditionally, expert systems, such as MYCIN [96], tend to encapsulate knowledge in a form that emulates a flow of mental inference: from evidence to hypothesis. Conversely, the arrows in Bayesian networks point from causes to effects, thus denoting a flow of constraints attributed to the physical world. This convention reflects the fact that people often prefer to encode experiential knowledge in a causal schemata [153]. The topology of a Bayesian network represents the dependency relationships between the different variables, and in particular, conditional independence of two variables given another variable. By definition, A is conditionally independent from B, given C, if:

$$P(A \mid B,C) = P(A \mid C) \qquad \textbf{(6.1)}$$

In the graphical representation of dependencies, the node C separates A from B: removing C from the network would make A and B disconnected. This can be generalised to subsets of nodes $\boldsymbol{A}$, $\boldsymbol{B}$ and $\boldsymbol{C}$, to give the following definition: $\boldsymbol{A}$ is independent from $\boldsymbol{B}$, given $\boldsymbol{C}$, if there is no path between any node in $\boldsymbol{A}$ and any node in $\boldsymbol{B}$ that satisfies:

1. Every node with converging arrows is in or has a descendant in $\boldsymbol{C}$
2. All other nodes are outside $\boldsymbol{C}$

Constructing a knowledge base (KB) for visual recognition with Bayesian networks is a two-fold process. Knowledge-acquisition, in expert systems terminology, is achieved through two phases of learning, or inductive reasoning. Firstly, the topology of the network, that expresses causal relationships between the variables, needs to be defined. This process is referred to as *structure learning*. Secondly, the prior probabilities and conditional probabilities that quantify the causal relationships need to be estimated. This process is referred to as *parameter learning*. Structure learning can be performed in collaboration with domain experts. They identify the variables they find relevant in the description of the system, and how they are causally linked. For instance, clinical studies have demonstrated the correlation between the presence of bronchial dilatation on CT scans, and the diagnosis of obstructive airways disease. Therefore, radiologists can argue that in a Bayesian network emulating the diagnostic process of obstructive airways diseases, bronchial dilatation should be a variable represented by a node, and that it should be linked by a descending arc to the node representing the hypothesis that the patient suffers from obstructive airways disease, as illustrated in Figure (6.1).

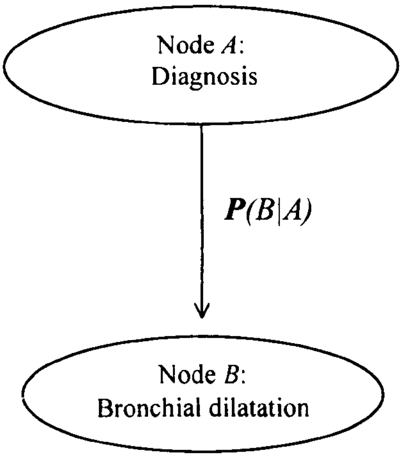


Figure 6.1. Example of a causal relationship represented in a probabilistic network. Expert domain knowledge indicates that bronchial dilatation is a sign of abnormal lung. The topology of the network reflects that knowledge: a node A representing diagnosis is linked to a node B representing bronchial dilatation. The causal relationship between the two nodes is expressed numerically in the probability matrix $\boldsymbol{P(B|A)}$ estimated with supervised parameter learning.

Parameter learning is achieved using a set of training images for which the observed variables are known, and the diagnostic conclusion was established by an expert. Assume that the nodes A and B are linked, and that their causal relationship is expressed by the conditional probability matrix $\boldsymbol{P(A|B)}$ denoted with the following integer ratios:

$$\begin{aligned}\mathbf{P}(B|A) &= \left(P(B_j|A_i)\right)_{(i,j)} \\ &= \left(\frac{b_{i,j}}{a_{i,j}}\right)_{(i,j)}\end{aligned} \qquad \textbf{(6.2)}$$

If we observe that node A is in state A_k and that node B is in state B_l, then we can update conditional probabilities by incrementing the corresponding entries in the matrix $\boldsymbol{P(A|B)}$ as follows:

$$P(B_j|A_i) = \frac{b_{i,j}+1}{a_{i,j}+1}, \quad \text{if } i = k \text{ and } j = l$$
$$= \frac{b_{i,j}}{a_{i,j}+1}, \quad \text{if } i = k \text{ and } j \neq l$$
$$= \frac{b_{i,j}}{a_{i,j}}, \quad \text{if } i \neq k \tag{6.3}$$

Theoretically, prior probabilities can also be learnt from the observed training set. We assume that the prior probabilities for A are given by vector ***P(A)***:

$$\mathbf{P}(A) = \left(P(A_i)\right)_i = \left(\frac{a_i}{n}\right)_i \tag{6.4}$$

When observing A_k, we can theoretically update the values $P(A_i)$ as follows:

$$P(A_i) = \frac{a_i+1}{n+1}, \quad \text{if } i = k$$
$$= \frac{a_i}{n+1}, \quad \text{if } i \neq k \tag{6.5}$$

Equation (6.5) relies on the assumption that the training set accurately reflects the prevalence of the hypothesis A, in the “real-world”. However, in the case of a medical decision-support system, the prevalence of the pathologies encountered is difficult to estimate. Such prior probabilities are dependent on the clinical context of the system. The prevalence of a rare lung disease in the population, for instance, is different from the prevalence of the disease amongst patients at a general hospital, which is again different from the prevalence of the disease at a specialised chest hospital where patients with unusual pathologies may be referred. Whereas it can be assumed that the training set models accurately the conditional probabilities linking variables within an image, it cannot be assumed that the training set models accurately prior probabilities. It is therefore preferable to set all prior probabilities to the same value, normalised to one.

All leaf nodes in visual recognition Bayesian networks represent instantiated or evidence variables. They contain the data derived from the image by the low-level feature extractors and represent the salient features extracted from the image, serving as a basis for the reasoning process. Probability propagation in the network is performed in a bottom-up fashion: from the instantiated leaf nodes towards the hypotheses root nodes. The result of propagation is the posterior probabilities of the root nodes, which represent the diagnostic conclusion of the network, given a set of evidence. Propagation is achieved in a bottom-up fashion, by computing at each leaf node a λ-message representing the evidence and by sending it to its parent. The algorithm for propagating probabilities, as established by Pearl [154], is as follows:

1. For all instantiated (leaf) nodes B^k, set $\lambda(B_i^k)$ using:

$$\begin{aligned} \lambda(B_i^k) &= 1 \quad \text{if state } i \text{ is instantiated} \\ &= 0 \quad \text{otherwise} \end{aligned} \tag{6.6}$$

2. From each leaf node B^k send a λ-message to its parent node A, using:

$$\lambda_{B^k}(A_i) = \sum_j P(B_j^k | A_i) \lambda(B_j^k) \tag{6.7}$$

3. For each parent node A, update its λ-value by multiplying the λ-messages received from its children *(B^k)*:

$$\lambda(A_i) = \prod_k \lambda_{B^k}(A_i) \tag{6.8}$$

4. Make the previous parent nodes A as the current leaf nodes. Repeat steps (2) and (3) until the root node is reached.
5. The posterior probability of the root A can then be computed as:

$$P(A_i | V) = \alpha \pi(A_i) \lambda(A_i) \tag{6.9}$$

where V denotes the set of evidence, $\pi(A_i)$ denotes prior probability $P(A_i)$, and α is a normalisation constant given by:

$$\alpha = \frac{1}{\sum_i \pi(A_i)\lambda(A_i)} \tag{6.10}$$

6.3 PROBABILISTIC NETWORK FOR THE DIAGNOSIS OF OBSTRUCTIVE LUNG DISEASES

6.3.1 Network topology

Based on domain expert knowledge described in Chapter 1, a probabilistic network for the diagnosis of obstructive lung disease can be designed. The leaf nodes represent the following variables:

1. the mean density difference δ between areas labelled as normal and areas labelled as hypo-attenuated, as described in Chapter 3,
2. a measure of bronchial dilatation, as described in Chapter 4,
3. a measure of bronchial wall thickening, as described in Chapter 4,
4. the extent of the parenchyma affected by each of the four textural patterns identified in Chapter 5.

The three first features do not reliably allow for differential diagnosis of centrilobular emphysema, panlobular emphysema and constrictive obliterative bronchiolitis. Therefore, a probabilistic network based on this data can only perform binary classification of normal *vs.* abnormal cases. A schematic representation of the network is given in Figure (6.2).

The topology of the network can be refined using the high-level clinical knowledge presented in Chapter 1. Once a case has been diagnosed as abnormal by the network illustrated in Figure (6.2), it is possible to distinguish between several types of obstructive lung diseases, using textural information. The network shown in Figure (6.3) allows for the differential diagnosis of normal lung, panlobular emphysema, centrilobular emphysema and constrictive obliterative bronchiolitis.

The leaf nodes of a probabilistic network must be conditionally independent given their parent [154]. Testing for independence can be achieved by measuring the correlation of pairs of features, as calculated with Pearson's correlation coefficient ρ [155], for instance. Although low correlation does not necessarily imply independence, it suggests that independence is a reasonable assumption [133]. Should the assumption of conditional independence be violated, then the topology of the network, as derived from expert knowledge, should be modified [133]. One of the leaf nodes may be eliminated, on the grounds that the information it provides is redundant with the information given by a correlated node. Alternatively, a new parameterisation can be found that combines the data contained in dependent nodes. Ultimately, the qualitative structure of the network may be redefined, in collaboration with domain experts. Conversely, if tests of correlation demonstrate the conditional independence of the leaf nodes, then the topology of the probabilistic network can be regarded as valid.

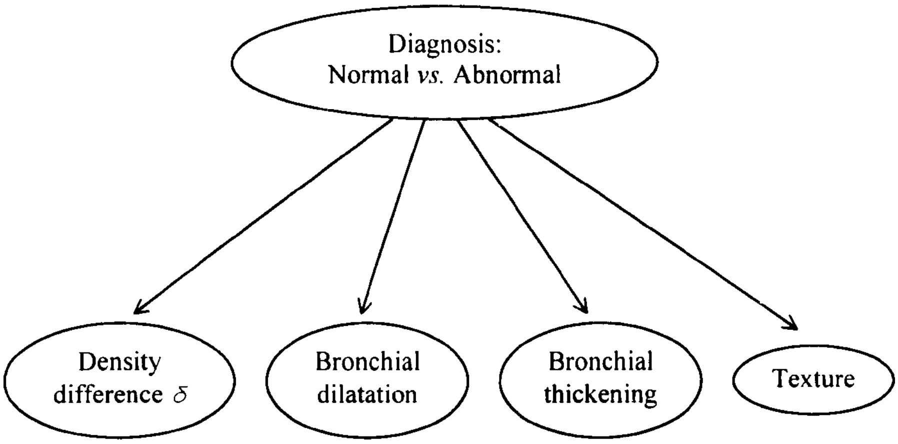


Figure 6.2. Probabilistic network for the binary classification of normal *vs.* abnormal lung. The leaf nodes represent the visual features extracted by the algorithms presented in Chapters 3, 4, and 5. The root node represents the diagnostic conclusion. The arrows represent the causal relationships, quantified by conditional probability matrices estimated with supervised learning.

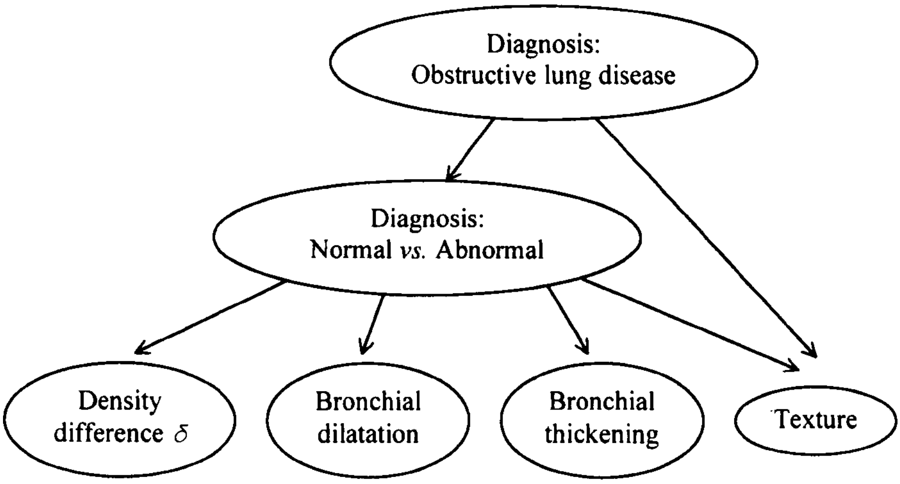


Figure 6.3. Probabilistic network for the differential diagnosis of obstructive lung diseases. All features are causally linked to a diagnosis of normal *vs.* abnormal lung, whereas texture only can allow for the differentiation amongst several types of obstructive lung diseases.

6.3.2 Variable discretisation

In order to express causal relationships through probability matrices, the variables instantiated at the leaf nodes need to be discrete. For each bronchus, dilatation and thickening are expressed on a discrete 4-point scale, as described in Chapter 4, and can be represented by 4-dimensional vectors ***d*** and ***t*** defined by:

$$\text{For } i = 0..3: \begin{cases} \mathbf{d}_i = 1, \text{if dilatation} = i \\ \mathbf{d}_i = 0, \text{otherwise} \end{cases} \tag{6.11}$$

$$\text{For } i = 0..3: \begin{cases} \mathbf{t}_i = 1, \text{if thickening} = i \\ \mathbf{t}_i = 0, \text{otherwise} \end{cases} \tag{6.12}$$

A discrete representation of bronchial morphology on an image can be obtained by summing and normalising all vectors ***d*** and ***t*** derived by ERS transform. A vector representing textural information is created by performing textural segmentation of the parenchyma. The extent of each of the four classes of patterns (normal, centrilobular emphysema, panlobular emphysema, constrictive obliterative bronchiolitis) is measured and stored in a normalised 4-dimensional vector. The density difference δ is represented on an N-point discrete scale. The minimum and maximum values of the scale were determined by the extrema values δ_{max} and δ_{min} of δ, as extracted from the training set of images, and the value of N was chosen equal to 4. The vector δ representing density differences is therefore defined by:

$$\text{For } i = 0..(N-1): \tag{6.13}$$

$$\begin{cases} \delta_i = 1, \text{if } \delta_{\text{min}} + \text{i} * \left(\frac{\delta_{\text{max}} - \delta_{\text{min}}}{N} \right) \leq \delta < \delta_{\text{min}} + (\text{i}+1) * \left(\frac{\delta_{\text{max}} - \delta_{\text{min}}}{N} \right) \\ \delta_i = 0, \text{otherwise} \end{cases}$$

6.4 MATERIAL AND METHODS

Supervised learning was achieved with a set of 88 HRCT scans acquired with the protocol described in Chapter 3, from 44 patients at two different levels (pulmonary venous confluence, and midway between pulmonary venous confluence and one centimetre above right hemi-diaphragm). For each image, an experienced radiologist established one predominant diagnostic conclusion: normal ($n = 24$) or abnormal ($n = 64$) (including centrilobular emphysema ($n = 16$), panlobular emphysema ($n = 17$), or constrictive obliterative bronchiolitis ($n = 31$)). Similarly, in order to evaluate the performance of the diagnostic network, 88 different HRCT images acquired at adjacent non-contiguous levels were labelled by a radiologist as normal ($n = 25$) or abnormal ($n = 63$) (including centrilobular emphysema ($n = 18$), panlobular emphysema ($n = 10$), or constrictive obliterative bronchiolitis ($n = 35$)). Binary classification (normal *vs.* abnormal) was achieved by employing the probabilistic network represented in Figure (6.2). Conditional independence of the leaf nodes was measured using Pearson's correlation coefficient ρ. To demonstrate the purpose of data fusion, classification was also performed on the basis of one single visual feature alone. For each of the feature extractors, supervised learning and classification were performed with a reduced probabilistic tree formed of one single leaf node and one root node. To reach a differential diagnosis, classification by the probabilistic network illustrated in Figure (6.3) was also performed.

6.5 RESULTS

6.5.1 Conditional independence

The value of Pearson's correlation coefficient ρ, for each pair of leaf nodes is given in Table (6.1).

Pair of features	ρ
Density difference δ and texture	0.188
Density difference δ and bronchial dilatation	-0.338
Density difference δ and bronchial wall thickening	0.089
Bronchial dilatation and bronchial wall thickening	0.216
Bronchial dilatation and texture	-0.102
Bronchial wall thickening and texture	0.064

Table 6.1. Pearson's correlation coefficient ρ, as measured for each pair of visual features. Low values of ρ suggest conditional independence.

Correlation was found to be low for all features, suggesting conditional independence of the leaf nodes. It is therefore reasonable to assume that the tree structures, given in Figure (6.2) and Figure (6.3), that encapsulate domain expert knowledge, define valid probabilistic networks.

6.5.2 Binary classification: normal *vs.* abnormal lung

The probabilistic network classified 84 out of 88 images correctly as normal or abnormal (accuracy 95.5 %). When employing one single visual feature extractor the accuracy was lesser: 48.9 % on the sole basis of bronchial wall thickening, 65.9 % on the sole basis of bronchial dilatation, 83.0 % on the sole basis of the density difference δ, and 84.1 % on the sole basis of textural information, as illustrated in Figure (6.4).

Combining through a probabilistic network all features extracted increases the accuracy of the classifier for the detection of normal *vs.* abnormal lung.

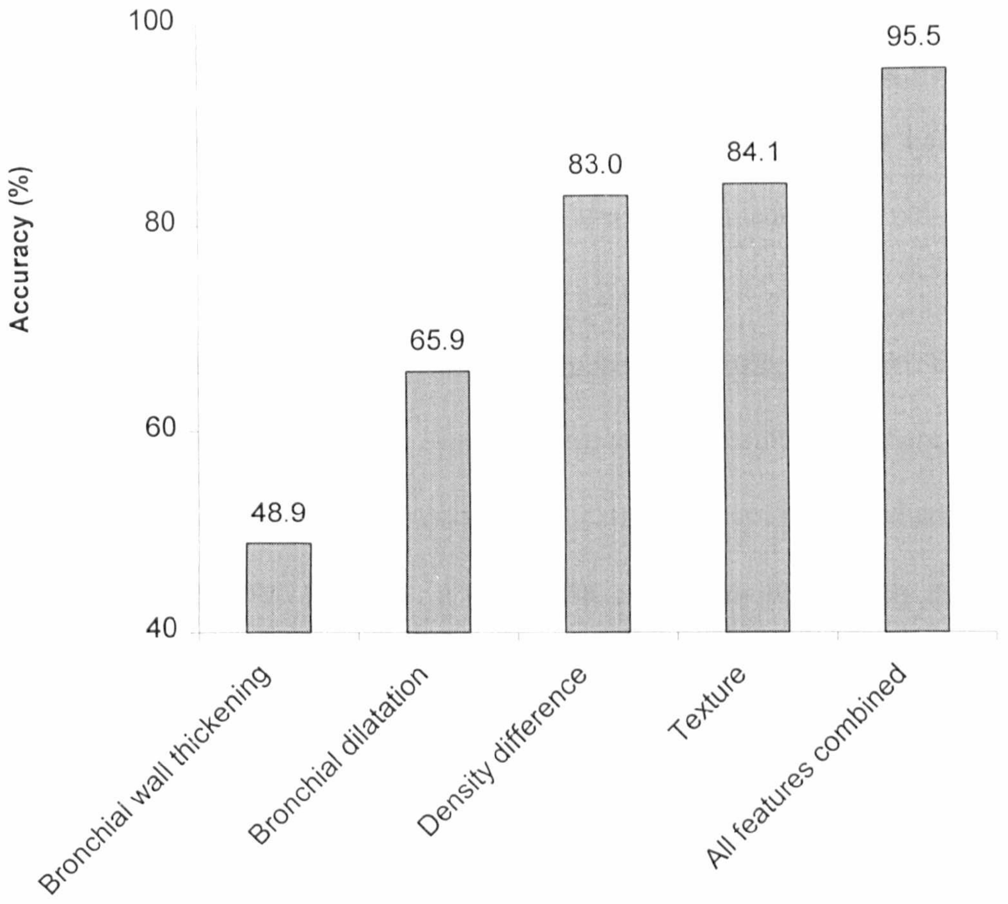


Figure 6.4. Accuracy of the binary probabilistic network, based on a single feature extractor, for each of the extractors employed. Multiplying and combining the inputs improves the performance of the classifier: an accuracy of 95.5 % in the detection of normal *vs.* abnormal lung is achieved by multiplying the inputs, a result superior to that of all feature extractors employed individually.

As an illustration of the diagnostic process, Figure (6.5) summarises the features extracted from the scan of a normal subject, shown in Figure (6.5a). Gradient-correction compensates for the density gradient due to gravity, as shown in Figure (6.5b). More than 90 % of the parenchyma classified by textural analysis was labelled as normal, as illustrated in Figure (6.5c). Figure (6.5d) reveals that some areas were classified as hypo-attenuated, but the density differences identified were small (δ = 34.3 H.U.) suggesting that they were not of pathological significance. Finally, the major bronchi identified by ERS transform, highlighted in Figure (6.5e), had normal morphology. The different feature extractors all indicate that no features of obstructive lung were present, and the probabilistic network concluded that the subject scanned was healthy. This diagnostic conclusion was in agreement with the assessment made by the radiologist.

The example of the scan of a patient diagnosed as suffering from constrictive obliterative bronchiolitis is given in Figure (6.6a), and the result of gradient-correction is shown in Figure (6.6b). Normal parenchymal texture was not identified anywhere, as shown in Figure (6.6c), and the predominant condition found was constrictive obliterative bronchiolitis. The segmentation of hypo-attenuated areas, illustrated in Figure (6.6d), revealed that 61.5 % of the lung was affected, and that the density differences were significant (δ = 75.0 H.U.). This finding was strengthened by the analysis of the morphology of the bronchi identified. The four first bronchi identified by ERS transform are circled in Figure (6.6e). Abnormal bronchial dilatation and bronchial wall thickening were present. All the findings support the diagnostic conclusion reached by the probabilistic network that the scan was one of diseased lungs.

Combining different features extractor helps resolve ambiguities and improve the overall accuracy of the diagnostic aid, as shown in Figure (6.7). The scan shown in Figure (6.7a) was regarded by the radiologist as showing the HRCT signs of constrictive obliterative bronchiolitis. However, due to the subtlety of the CT appearance of the condition and the relatively limited extent of diseased lung, textural segmentation only identified 20.6 % of the classified lung as characteristic of the disease. More than 70 % of the classified parenchyma was found to be normal. Nevertheless, the other feature extractors suggest that the scan was abnormal. The density differences shown in Figure (6.7d) were relatively large (66.0 HU) and the ERS transform detected abnormal bronchial morphology, as illustrated in Figure (6.7e). Therefore, the variety of inputs employed allowed the probabilistic network to reach a correct diagnostic conclusion of abnormal lungs.

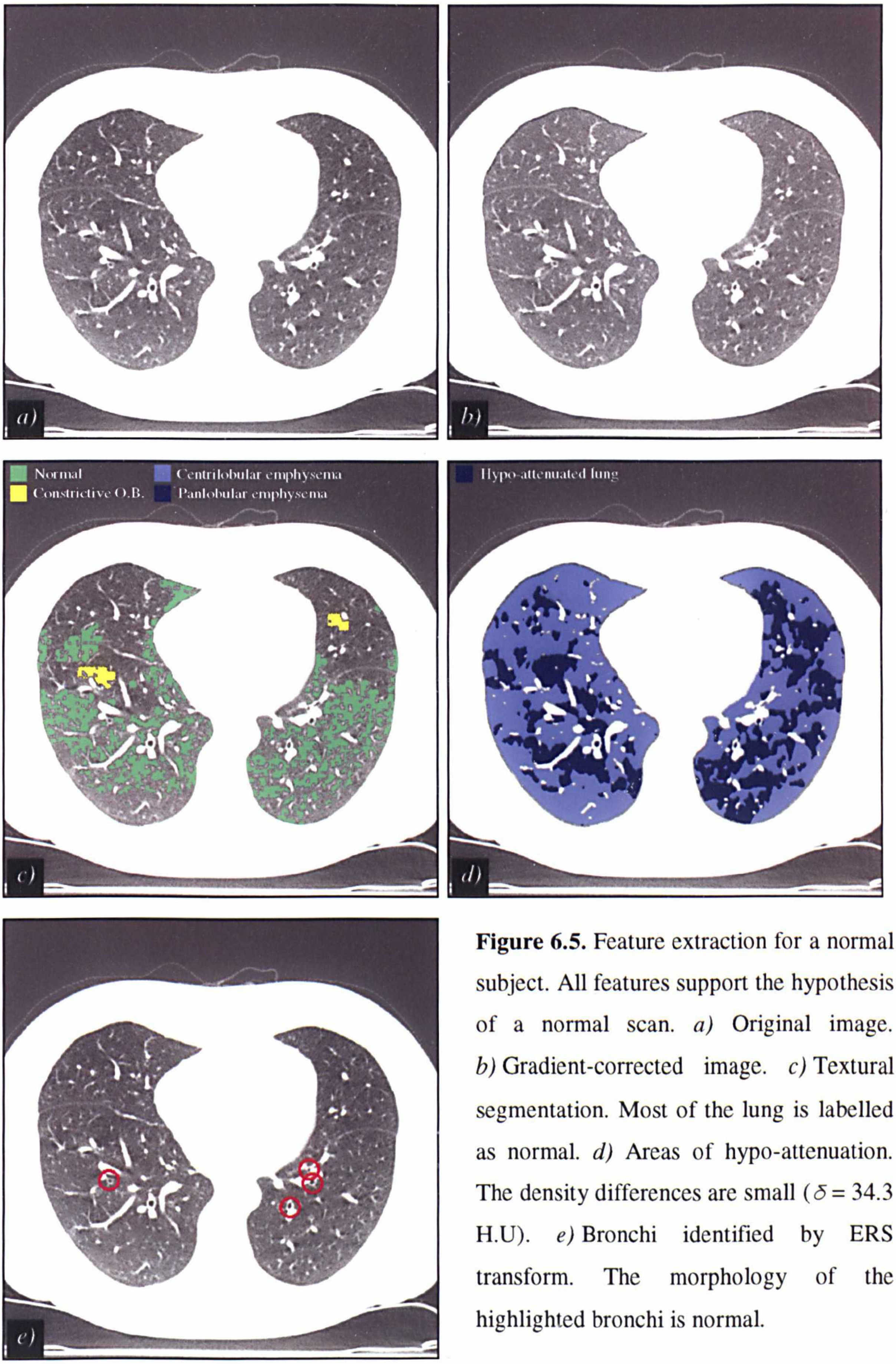


Figure 6.5. Feature extraction for a normal subject. All features support the hypothesis of a normal scan. *a)* Original image. *b)* Gradient-corrected image. *c)* Textural segmentation. Most of the lung is labelled as normal. *d)* Areas of hypo-attenuation. The density differences are small ($\delta = 34.3$ H.U). *e)* Bronchi identified by ERS transform. The morphology of the highlighted bronchi is normal.

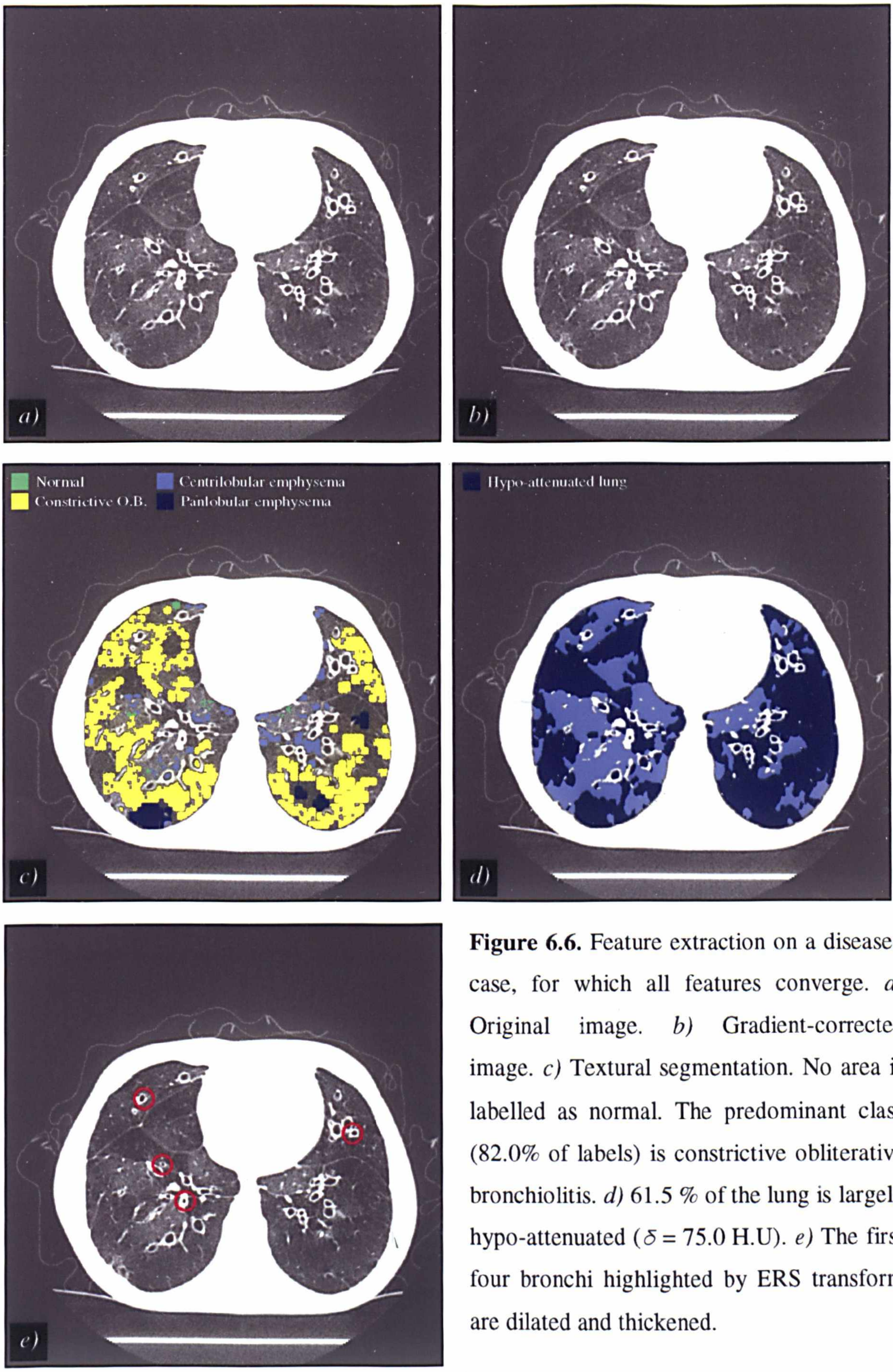


Figure 6.6. Feature extraction on a diseased case, for which all features converge. *a)* Original image. *b)* Gradient-corrected image. *c)* Textural segmentation. No area is labelled as normal. The predominant class (82.0% of labels) is constrictive obliterative bronchiolitis. *d)* 61.5 % of the lung is largely hypo-attenuated (δ = 75.0 H.U). *e)* The first four bronchi highlighted by ERS transform are dilated and thickened.

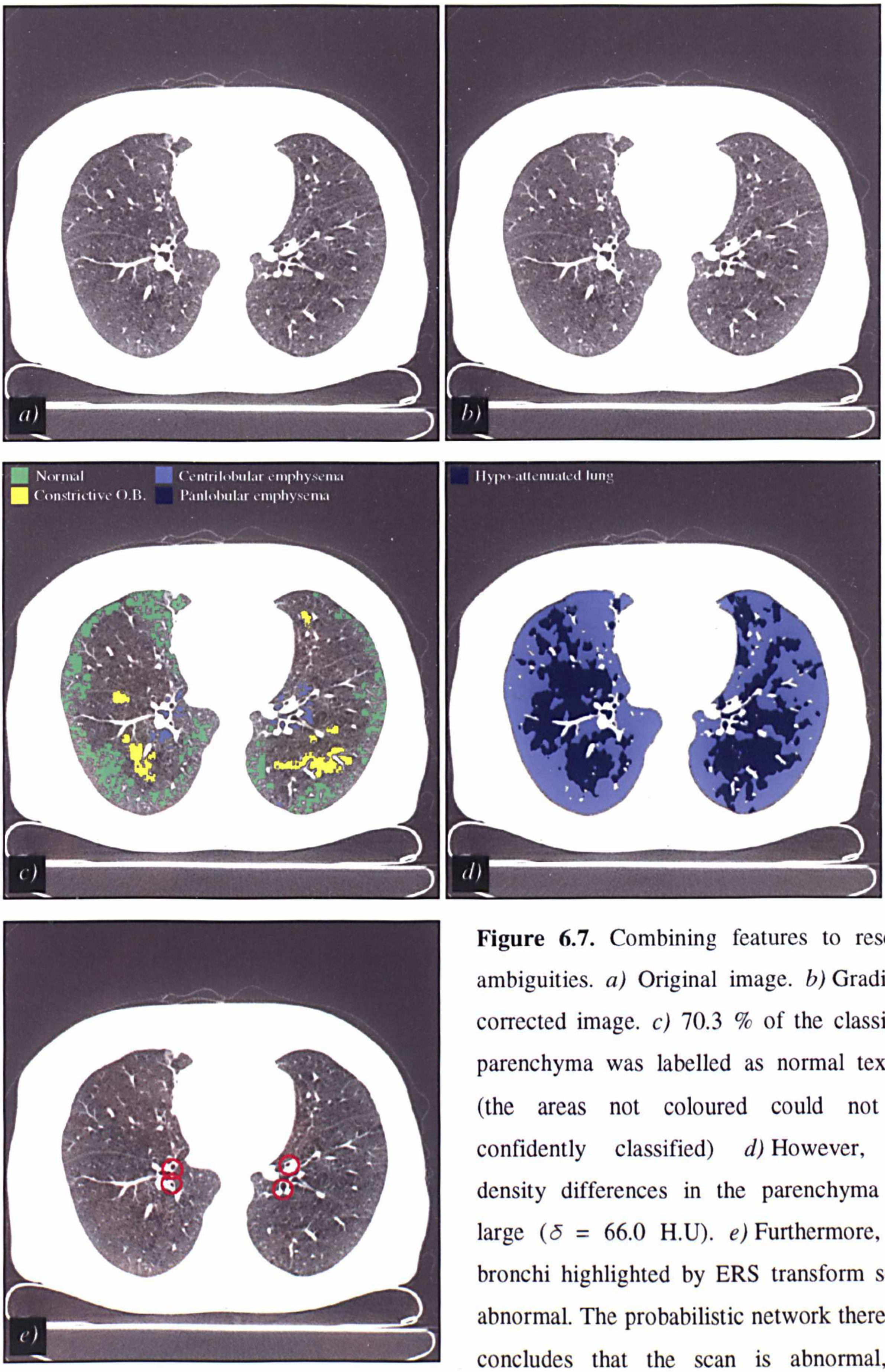


Figure 6.7. Combining features to resolve ambiguities. *a)* Original image. *b)* Gradient-corrected image. *c)* 70.3 % of the classified parenchyma was labelled as normal texture (the areas not coloured could not be confidently classified) *d)* However, the density differences in the parenchyma are large (δ = 66.0 H.U). *e)* Furthermore, the bronchi highlighted by ERS transform seem abnormal. The probabilistic network therefore concludes that the scan is abnormal, in accordance with an expert radiologist.

6.5.3 Differential diagnosis

Using the probabilistic network illustrated in Figure (6.3), a correct diagnosis of normal lung, panlobular emphysema, centrilobular emphysema or constrictive obliterative bronchiolitis was reached in 63 out of 88 instances, giving an accuracy of 71.6 %. Using textural information alone, the correct diagnosis was obtained in 57 out 88 cases, giving an accuracy of 64.8 %.

6.6 DISCUSSION

The low-level feature extractors presented in this thesis were incorporated in a probabilistic network encapsulating clinical knowledge through supervised learning. The accuracy of the method for the classification of normal *vs.* abnormal lung is high (95.5 %). A specific diagnosis of obstructive lung diseases can also be achieved with high accuracy (71.6 %).

The low correlation between each pair of visual features extracted, as measured by Pearson's correlation coefficient ρ, suggests that the leaf nodes of the probabilistic network may be regarded as conditionally independent. Low values of Pearson's correlation coefficient ρ do not guarantee strict conditional independence, but suggest that the assumption is reasonable. Also, a subjective threshold on ρ is required to determine which values can be regarded as acceptable. The maximum correlation found in the networks presented in this chapter is $\rho = 0.216$ for the nodes representing bronchial dilatation and bronchial wall thickening, as shown in Table (6.1). Higher values of Pearson's correlation coefficient ($\rho = 0.264$) have been regarded as appropriate for the definition of probabilistic networks that perform accurately [133]. Therefore, it can be assumed that the variables employed here are conditionally independent and that the topology of the probabilistic networks designed with domain expert knowledge is valid.

A limitation of the approach presented is revealed by cases presenting coexisting pathologies. It may happen that a single HRCT scan displays signs of both panlobular and centrilobular emphysema, for instance. The diagnostic conclusion of the probabilistic network will only mention the pathology that affects most of the parenchyma, on the basis of texture classification. It is important to acknowledge that the decision-support element provided by the probabilistic network only indicates a predominant condition, and that the careful assessment of the scans, and their textural segmentation, may reveal coexisting pathologies.

The diagnostic interpretation of the major bronchovascular bundles identified by ERS transform may prove problematic when the airways are near-bifurcation and/or close to the hila, as shown in Figure (6.7e). This limitation results from the fact that the automated detection of bronchi is based on geometric properties and does not include anatomical and functional knowledge, as discussed in Chapter 4. However, in most cases, the morphology of the bronchovascular bundles identified is unambiguous, as illustrated in Figure (6.6e), and therefore the ERS transform helps improve the accuracy of the classifier.

It was demonstrated that multiplying the inputs increases the reliability of the technique by resolving potential ambiguities. The probabilistic binary classifier for the identification of abnormal lung outperforms each of the single visual feature extractors. In the case of differential diagnosis, only textural information can help differentiate panlobular emphysema, centrilobular emphysema and constrictive obliterative bronchiolitis. Nevertheless, the other feature extractors, by supporting the hypothesis of normal or abnormal lung, help improve the reliability of differential diagnosis. On the consideration of parenchymal texture alone, the case illustrated in Figure (6.5) was wrongly regarded as normal, but correct diagnosis was obtained by incorporating the other findings. Employing several feature extractors also allows the probabilistic network to demonstrate clearly the inference mechanism used to reach a conclusion. This property can be of considerable value in the teaching of chest radiology.

CHAPTER 7

CONCLUSION

7.1 SUMMARY OF THE THESIS

Various conditions characterised by airflow obstruction may be demonstrated on HRCT. However, the signs of these obstructive lung diseases can be subtle, and subjective assessment is prone to observer variability, rendering the investigations of structure/function relationships difficult. This thesis has presented fully automated methods for the analysis of HRCT data of the lungs. Algorithms for the extraction of visual features used for the diagnosis of obstructive lung diseases have been described and evaluated. They provide reproducible measurements of specific CT features, highlight regions of interest, and serve as the basis for a diagnostic decision-support system.

A cardinal finding of obstructive lung disease on CT is the presence of areas of hypo-attenuation in the lung parenchyma. In order to perform automated and reproducible measurements of the extent of hypo-attenuated lung, an algorithm for the segmentation of the main anatomical structures was implemented. A method for the correction of density gradient due to gravity on lung CT was introduced. Based on histogram deconvolution and histogram energy maximisation, the algorithm was shown to perform gradient-correction with high accuracy on images acquired with a lung phantom (accuracy 99.3 %). The two-staged gradient-correction algorithm presented in this thesis is novel, general and non-parametric. It can be employed for the correction of non-uniformity associated with other image processing applications. Segmentation of the areas of hypo-attenuation with a model-based threshold optimisation technique was then performed. The automated measurement of the extent of hypo-attenuated lung on clinical images was compared with the assessment made by experienced radiologists. The accuracy of the technique was demonstrated to be high (average error: 8.2 % of the extent of hypo-attenuated lung).

Bronchial morphology, characterised by measurements of bronchial dilatation and bronchial wall thickening on major bronchi near-perpendicular to the plane of acquisition, is a feature useful in the interpretation of ambiguous scans, and is used by radiologists to refine the differential diagnosis in cases of obstructive lung disease. The ERS transform was introduced for the detection of elliptical rings, the characteristic appearance of major near-perpendicular bronchi on CT. Based on the statistical analysis of edge-distribution in local polar co-ordinates, the ERS transform uses geometric properties of ellipses. The ERS transform is a novel method for the recognition of elliptical patterns with a low computational cost. Measurements of bronchial morphology are obtained by comparing the size of bronchi found by ERS transform with adjacent pulmonary arteries identified with template matching. The robustness of the method was evaluated on synthetic images, and patient studies allowed for clinical validation. Agreement between the ERS transform and a radiologist, for the assessment of bronchial dilatation and bronchial wall thickening was remarkably similar to that between human observers (kappa statistics for the detection of bronchial dilatation and bronchial wall thickening: 0.63 and 0.55 respectively).

Parenchymal texture on HRCT provides essential information for the differentiation between the various causes of obstructive lung disease. Textural description was encapsulated using the statistical moments of CT densities, co-occurrence matrices and run-length parameters, in a set of representative regions of interest chosen by a radiologist. The textural features were computed on segmented images, after removal of the macroscopic pulmonary vessels. For each of the four classes of textural patterns, the distribution of the descriptors was modelled by a normal distribution, serving as a basis for a probabilistic Bayesian classifier. By setting a threshold on posterior probabilities, the textural classifier achieved high accuracy for the recognition of each class when tested on regions of interest from a new image set (sensitivity: 79.5 %, specificity: 93.2 %). Removal of the macroscopic anatomical structures prior to classification allowed for the accurate textural differentiation of four classes of patterns, as opposed to the binary classifiers previously developed for lung CT images. This result highlights the potential for texture analysis in medical imaging, when employed in conjunction with other feature extractors.

Using probabilistic networks, the outputs of the feature extractors presented above were combined, in order to provide a diagnostic aid. The topology of the networks employed was derived from expert knowledge. The validity of the topology was then validated by testing the conditional independence of the leaf nodes. The conditional probabilities that quantify the causal relationships between the leaf nodes (the visual features extracted automatically) and the root node (the diagnostic decision) were estimated with supervised learning. For each image of a training set of HRCT scans, a radiologist established which condition was dominant. Testing the system on another set of images demonstrated considerable accuracy for the detection of normal *vs.* abnormal lung. With a fully automated system, 95.5 % of tested images were correctly classified as normal or abnormal lung. Differential diagnosis of normal lung, centrilobular emphysema, panlobular emphysema and constrictive obliterative bronchiolitis was also performed with an accuracy of 71.6 %. The importance of combining classifiers was demonstrated: data fusion through a probabilistic network allowed for lifting diagnostic ambiguities. The performance of the diagnostic decision-support system is superior to that of all feature extractors employed individually.

7.2 FUTURE WORK

The image characteristics of obstructive lung diseases evaluated in the preceding investigations can be regarded as the major, but not exclusive, manifestations of this group of diseases. Decreased perfusion of the lung parenchyma is often associated with obstructive lung diseases and hypo-attenuation can be a manifestation of under-perfusion; in this context, the size and distribution of the macroscopic pulmonary vessels also provide important information. The automated segmentation of the macroscopic pulmonary vessels, as described in Chapter 3, provides a first estimate of the characteristics of perfusion on a lung image. However, the reliable interpretation of the morphology and distribution of the segmental and subsegmental pulmonary vessels requires an elaborate geometrical model incorporating high-level anatomical and functional knowledge. Indeed, the shape, branching pattern, and especially the calibre of vessels can only be interpreted in the light of their relative position in the vascular tree.

When combining classifiers in computer vision applications, contextual *spatial* information is important. It may be limiting to interpret the output of several image feature extractors with no reference to their spatial relationships. Relaxation labelling [156,157] is an iterative process for the propagation of spatial contextual information in vision systems. It may provide an appropriate framework for the extraction of diagnostic findings in imaging systems. For each feature extracted (bronchial abnormality, area of hypo-attenuation) a local spatial support should be defined, that characterises the spatial range for which the finding provides significant diagnostic information. Spatial contextual information should not be restricted to single images, but should make use of volumetric data. Reliable diagnosis from CT images requires the analysis of images acquired at different levels of the lungs. Also, the search for salient visual features by image processing should be directed by the partial conclusions of the decision-support systems [158].

Issues specific to the medical imaging diagnostic process need to be tackled, that have not been investigated in other fields of computer vision and Artificial Intelligence. Human visual pattern recognition is different from general reasoning and explicit domain knowledge representation often overlooks those factors that are subconsciously applied during visual recognition. Because of the *ad hoc* nature of grouping low level visual features, there has been no generic way of designing medical imaging decision-support systems. Each application is treated as a new problem, and requires considerable amount of interaction between clinical radiologists and computer scientists in order to identify intrinsic visual features that characterise the medical conditions concerned. The decision-support system presented in this thesis was made possible only through extensive interaction with domain experts, in a rare environment of combined clinical practice and computing research. These drawbacks call for the development of a new framework for a natural and systematic way of gathering knowledge from domain experts.

Tracking and analysing the eye movements of experienced radiologists would provide valuable insights into the mechanisms of selection and cross-examination of several findings to support or discard diagnostic hypotheses. Using dedicated eye movement tracking devices [159,160], it is possible to analyse how radiological visual search takes place [161]. At a conscious or subliminal level, radiologists focus on and compare particular regions of interest on different areas of a single projectional image (for example, a chest radiograph), or on the dozens of transaxial images that constitute a single CT examination. Even though the manner in which an expert radiologist assimilates information from a single image has been studied at a basic level [162], no research has so far studied the issue of visual features comparison and cross-examination of a set of numerous images. A coherent framework needs to be designed to investigate it. It would be helpful in refining the diagnostic models implemented in expert systems. It would have great potential for the teaching of the *modus operandi* of an experienced radiologist to trainees. The traditional training of radiologists is by a process of "directed search" initially for obvious localised abnormalities that can be detected by analysing all areas of the image systematically. Providing information on which image features are most thoroughly studied by experts, and what cross-checking is carried out, should improve the efficacy of training. By approaching the process of radiological diagnosis from observations of how expert visual search takes place, new possibilities for training and developing self-learning decision-support systems will emerge.

In an ideal environment, imaging expert systems should be seamlessly integrated in the clinical activity through PACS systems. Some early attempts to do so have been reported, as in digital mammography [131]. The expert systems could benefit from the clinical information associated with each case and provided by the clinicians. From the medical report of each patient, a set of facts derived from a knowledge base can be established and given as an additional input to the system. In return, the conclusion reached by the imaging expert system can be tagged to each case, as a decision-support element for the radiologist [127]. When faced with the difficulty of reporting ambiguous images, radiologists should be able to call, by the click of a button, a set of computerised tools to support their decision. Only with such an invisible front-end can expert systems be expected to become an integral part of the clinical activity.

Future applications will probably shift from pure diagnostic purpose to therapeutic management. In the iterative process of diagnosis and treatment, the response of a patient to a given therapy provides valuable information for supporting, discarding or refining diagnostic hypotheses. Such strategies have been successfully implemented in systems like T-HELPER for the management of patients with HIV disease [163]. They should become prevalent in expert systems for therapy planning and control. Imaging systems should provide decision-support for radiation therapy and surgery planning and assessment. Also, computerised decision-support systems can be expected to play a greater part in medical training in the future. As an example, the Quick Medical Reference (QMR) derived from the knowledge base established with the INTERNIST-1 system can be used as an "electronic textbook of medicine" [119]. Imaging expert systems, by automatically highlighting the areas of interest on an image, and outlining the findings extracted and their relationships to a possible diagnosis, should prove to be a useful and increasingly available tool for the teaching of radiology.

The generalisation of digital technology in medical imaging makes it possible to design computer-aided decision-support tools to help resolve the difficulties encountered by radiologists. It has already been demonstrated that image processing and automated feature extraction can help to improve diagnostic accuracy, and the work presented in this thesis provides a specific and successful example of such an application. Future challenges are the development of vision techniques tailored for a wider range of imaging modalities and pathologies and the design of systematic methods for the combination of several classifiers using contextual volumetric information. After evaluation of these tools, they should ideally be integrated in PACS systems, to provide clinicians with a documented second opinion and to highlight regions of interest. Then they can be used beneficially in the practice and teaching of radiology.

APPENDIX A

HISTOGRAM ENERGY MAXIMISATION OF A DISCRETE PIECE-WISE CONSTANT FUNCTION

Superimposing a linear intensity gradient to a piece-wise constant function decreases its histogram energy, under the assumption that the intensity distributions of the original constant pieces do not overlap under the effect of the gradient. This can be demonstrated by looking at a one-dimensional function. We will first consider the cases of a constant function, and a constant function with added linear term, before considering the case of a piece-wise constant function with added linear term. Piece-wise constant functions provide a good mathematical model for the theoretical CT image $I_u(x)$ of the lung parenchyma, consisting of homogeneous normal lung with a given CT density, and areas of homogeneous hypo-attenuated lung with another CT density.

A.1 HISTOGRAM ENERGY OF A DISCRETE CONSTANT FUNCTION

Let δy and δz be discretisation steps along the y and z axis.

Let $(y_i)_{0 \le i < N}$ be N consecutive sample points along the y-axis: $\forall i, y_i = i\delta y$

Let (z_j) be a sequence of sample points along the z-axis: $\forall j, z_j = j\delta z$

Let f be a constant function, sampled at $(y_i)_{1 \le i \le N}$, as illustrated in Figure (A.1a): $\forall i / 1 \le i \le N, f(y_i) = z_0 = j_0 \delta z$

Let H_f be the histogram of the function f, as illustrated in Figure (A.1b).

The energy of H_f is:

$$E(H_f) = N^2 \tag{A.1}$$

A.2 HISTOGRAM ENERGY OF A DISCRETE CONSTANT FUNCTION WITH ADDED LINEAR TERM

Let a linear gradient with a slope a be added to the function f:

$$\forall i, f'(y_i) = j_0\delta z + ay_i \tag{A.2}$$

The histogram $H_{f'}$ is given by the value of each bin b_j, b_j being the number of points y_i for which the sampled value of $f'(y_i)$ is equal to z_j. The sampled value of $f'(y_i)$ is equal to z_j if and only if :

$$\begin{aligned} & z_j \le f'(y_i) < z_{j+1} \\ \Leftrightarrow\ & j - j_0 \le ai\frac{\delta y}{\delta z} < j - j_0 + 1 \end{aligned} \tag{A.3}$$

Let I be the function that returns the integer part of any real number. Assuming that $\frac{\delta y}{\delta z} = 1$, the value of each bin b_j is then:

$$b_j = \sum_i \delta_{I(ai)-(j-j_0)} \text{ with } \delta_0 = 1 \text{ and } k \ne 0 \Rightarrow \delta_k = 0 \tag{A.4}$$

Therefore three cases arise, depending on the value of $|a|$:

- if $|a| \ge 1$, as illustrated in Figure (A.1c), the N sampled values of f' are distributed amongst N bins. N bins receive one point, the others receive none, as shown in Figure (A.1d):

  $$\forall j, 0 \le b_j \le 1 \tag{A.5}$$

  And the histogram energy is:

$$E(H_{f'}) = \sum_j b_j^2 = N \leq N^2 \qquad \textbf{(A.6)}$$

- if $\frac{1}{N} \leq |a| < 1$, as shown in Figure (A.1e), the sampled values of f' are distributed amongst $I(|a|N)$bins containing $I\left(\frac{1}{|a|}\right)$points each.. The histogram energy, as illustrated in Figure (A.1f), therefore verifies:

$$E(H_{f'}) = I(|a|N)\left[I\left(\frac{1}{|a|}\right)\right]^2 \leq \frac{N}{|a|} \leq N^2 \qquad \textbf{(A.7)}$$

- if $|a| < \frac{1}{N}$: sampled f' is equal to f.

$$b_0 = N \text{ and } \forall j \neq 0, b_j = 0 \qquad \textbf{(A.8)}$$

And the histogram energy is:

$$E(H_{f'}) = N^2 \qquad \textbf{(A.9)}$$

Therefore, histogram energy is maximum for the original constant function f.

A.3 HISTOGRAM ENERGY OF A DISCRETE PIECE-WISE CONSTANT FUNCTION WITH ADDED LINEAR TERM

The histogram of a piece-wise constant function exhibits several sharp peaks. After addition of the linear gradient, the peaks are broadened. If we assume that the distribution of values of the peaks do not overlap, then the histogram of a piece-wise constant function with added linear term is equivalent to the sum of the histograms of several constant functions with added linear terms. From section (A.2), we get that histogram energy is maximum for the original piece-wise constant function. We have therefore demonstrated that maximising intensity histogram energy allows for the correction of a linear gradient added to a piecewise constant function, under the assumption that the distributions of intensity classes do not overlap.

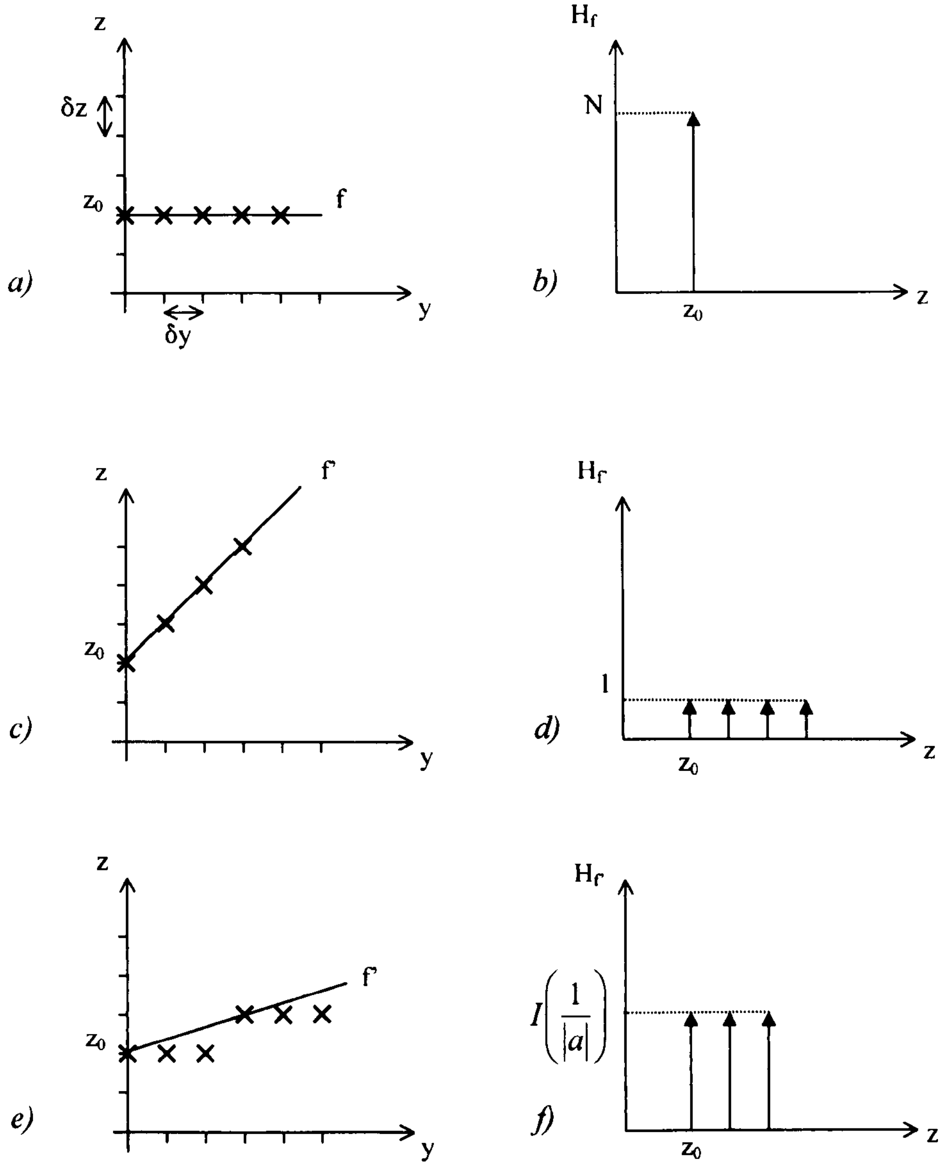


Figure A.1. Schematic representation of discrete functions and their histogram. *a)* Constant function. *b)* Corresponding histogram exhibiting one peak. *c)* Constant function with added linear term with a slope greater than one. *d)* Corresponding histogram. Each bin counts at most one point. *e)* Constant function with added linear term with a slope lesser than one. *f)* Corresponding histogram.

APPENDIX B

LUNG PHANTOM

For the validation of the gradient-correction and segmentation techniques presented in Chapter 3, a CT lung phantom was devised. Most of the phantoms developed for CT of the chest focus on mimicking the anatomy of the macroscopic high density structures. While the chest wall and ribs are accurately modelled, the lung parenchyma is usually represented by a homogeneous material (often air) [164]. Such models cannot be used to mimic the normal and abnormal intensity differences of the lung parenchyma [165,166]. Some authors have previously stressed the particular problem associated with CT lung densitometry [167], and the need for an accurate phantom of the low density structures. Because most clinical CT scanners are used for the examination of solid body parts (*e.g.* brain or abdomen) they are not often optimised by manufacturers to image the low density of the parenchyma (below -900 H.U.). Cavities within the lung tissues produce very inhomogeneous patterns with non-linear partial volume effect [14]. When trying to assess and quantify these issues, polyethylene foam (PE) was found to be a suitable material for the simulation of the lung parenchyma. PE foam is chemically pure and well-defined. Its "alveolar structure" produces a fine intricate pattern of air and higher-density material similar to the lung. PE foam is manufactured in various densities. The densities are relatively uniform. Nevertheless, the available density range is narrower than found in lungs [14,167,168]. To mimic the density differences between normal and hypo-attenuated lung, four different PE foams [167] (Jiffycel Polyethylene Foam, Jiffy Packaging Ltd, Winsford, UK) with the following densities were used: 25 kg/m^3, 35 kg/m^3, 65 kg/m^3, 150 kg/m^3. In each material, one large cylinder (radius 13.2 cm), and four small cylinders (radius 2.8 cm) were cut. The CT attenuation of each type of foam, as measured on a set of HRCT images, is given in Table (B.1).

PE Foam density (kg/m^3)	**Mean CT value (H.U.)**	**Standard Deviation (H.U.)**
25	-976.2	16.1
35	-970.3	14.3
65	-940.7	28.8
150	-864.0	60.2

Table B.1. Average and standard deviation of the CT attenuation, as measured on HRCT scans of PE foams of various densities.

The small cylinders were slotted into holes within the large cylinder to create patchy patterns of different densities analogous to heterogeneous lung disease. A larger water-filled container made of an inner and outer cylinders of Perspex surrounded the foam phantom to mimic the chest wall. The PE foam was put in position within the container prior to scanning. An illustration of the phantom ready for CT scanning is given in Figure (B.1).

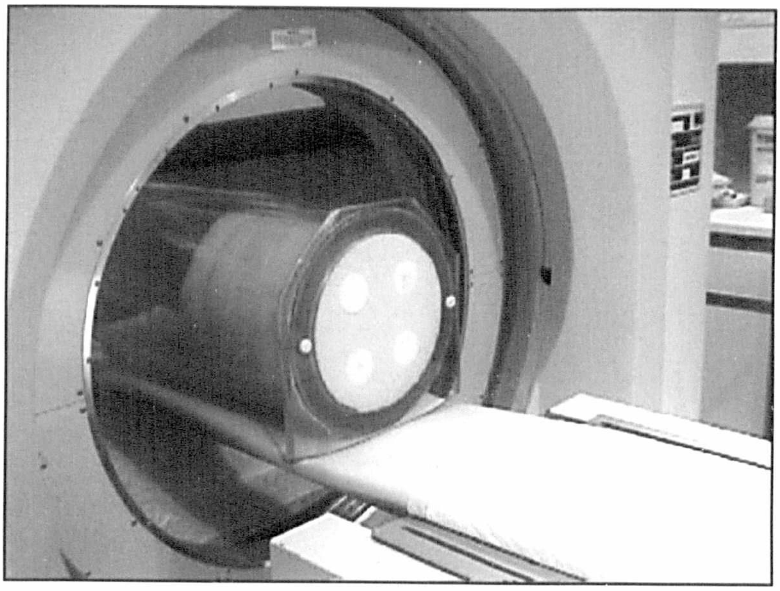


Figure B.1. Lung phantom placed in EBCT scanner. A Perspex container filled with water surrounds the polyethylene foams of various densities.

REFERENCES

[1] American Thoracic Society, Chronic bronchitis, asthma, and pulmonary emphysema: a statement by the committee on diagnostic standards for nontuberculous respiratory diseases *American Review of Respiratory Disease*, vol. 85, pp. 762-768, 1962.

[2] G.L. Snider, J. Kleinerman, and W.M. Thurlbeck, The definition of emphysema. Report of a National Heart, Lung and Blood Institute, division of lung diseases workshop *American Review of Respiratory Disease*, vol. 132, pp. 182-185, 1985.

[3] J.C. Hogg, P.T. Macklem, and W.M. Thurlbeck, Site and nature of airway obstruction in chronic obstructive lung disease *New England Journal of Medicine*, vol. 278, pp. 1355-1360, 1968.

[4] D.M. Hansell, M.B. Rubens, S.P.G. Padley, and A.U. Wells, Obliterative bronchiolitis: individual CT signs of small airways disease and functional correlation *Radiology*, vol. 203, pp. 721-726, 1997.

[5] D. Lamb. Chronic obstructive pulmonary disease: pathology. In: *Respiratory medicine*, eds. R.A.L. Brewis, B. Corrin, D.M. Geddes, and G.J. Gibson. London: Saunders, 1995. pp. 706-746.

[6] J.T. McFarlane. Pneumonia and other acute infections. In: *Respiratory medicine*, eds. R.A.L. Brewis, B. Corrin, D.M. Geddes, and G.J. Gibson. London: Saunders, 1995. pp. 706-746.

[7] W.M. Thurlbeck and J.L. Wright. *Thurlbeck's chronic airflow obstruction*, Oxford: Blackwell Science, 1999. pp. 9-16.

[8] W.M. Thurlbeck and G. Simon, Radiographic appearance of the chest in emphysema *AJR. American Journal of Roentgenology*, vol. 130, pp. 429-440, 1978.

[9] T.C. McLoud, G.R. Epler, T.V. Colby, E.A. Gaensler, and C.B. Carrington, Bronchiolitis obliterans *Radiology*, vol. 159, pp. 1-8, 1985.

[10] C.S. Ng, S.R. Desai, M.B. Rubens, S.P.G. Padley, A.U. Wells, and D.M. Hansell, Visual quantitation and observer variation of signs of small airways disease at inspiratory and expiratory CT *Journal of Thoracic Imaging*, vol. 14, pp. 279-285, 1999.

[11] P. Grenier, I. Mourey-Gerosa, M.W. Brauner, A.N. Leung, S. Lenoir, M.P. Cordeau, and B. Mazoyer, Abnormalities of the airways and lung parenchyma in asthmatics: CT observations in 50 patients and inter- and intraobserver variability *European Radiology*, vol. 6, pp. 199-206, 1996.

[12] T. Fotheringham, F. Chabat, D.M. Hansell, A.U. Wells, S.R. Desai, C. Gückel, S.P.G. Padley, M. Gibson, and G.Z. Yang, A comparison of methods for enhancing the detection of areas of decreased attenuation on CT caused by

airways disease *Journal of Computer Assisted Tomography*, vol. 23, pp. 385-389, 1999.

[13] D.P. Boyd and D.L. Parker. Basic principles of CT. In: *Computed Tomography of the body*, eds. A.A. Moss, G. Gamsu, and H.K. Genant. London: Saunders, 1983. pp. 1-21.

[14] G.J. Kemerink, R.J.S. Lamers, and G.R.P. Thelissen, The non-linear partial volume effect and CT densitometry of foam and lung *Medical Physics*, vol. 22, pp. 1445-1450, 1995.

[15] C.J. Bergin, N.L. Müller, and R.R. Miller, CT in the quantitative assessment of emphysema *Journal of Thoracic Imaging*, vol. 1, pp. 94-103, 1986.

[16] J.B. Cailes, R.M. Du Bois, and D.M. Hansell, Density gradient of the lung parenchyma on CT in patients with lone pulmonary hypertension and systemic sclerosis *Academic Radiology*, vol. 3, pp. 724-730, 1996.

[17] J.A. Verschakelen, L. Van Fraeyenhoven, G. Laureys, M. Demedts, and A.L. Baert, Differences in CT density between dependent and nondependent portions of the lung: influence of lung volume *AJR. American Journal of Roentgenology*, vol. 161, pp. 713-717, 1993.

[18] N. Tomiyama, N. Takeuchi, H. Imanaka, N. Matsuura, S. Morimoto, J. Ikezoe, T. Jokkoh, J. Arisawa, and T. Kozuka, Mechanism of gravity-dependent atelectasis. Analysis by nonradioactive xenon-enhanced dynamic computed tomography *Investigative Radiology*, vol. 28, pp. 633-638, 1993.

[19] W.R. Webb, E.J. Stern, N. Kanth, and G. Gamsu, Dynamic pulmonary CT: findings in healthy adult men *Radiology*, vol. 186, pp. 117-124, 1993.

[20] S. Morimoto, N. Takeuchi, H. Imanaka, M. Nishimura, J. Takezawa, N. Taenaka, N. Matsuura, K. Tomoda, J. Ikezoe, and J. Arisawa, Gravity-dependent atelectasis: radiologic, physiologic and pathologic correlation in rabbits on high-frequency oscillation ventilation *Investigative Radiology*, vol. 24, pp. 522-533, 1989.

[21] T. Hachenberg, H. Lundquist, L. Tokics, B. Brismar, and G. Hedenstierna, Analysis of lung density by computed tomography before and during general anesthesia *Acta Anaesthesiologica Scandinavia*, vol. 37, pp. 549-555, 1993.

[22] A. Strandberg, G. Hedenstierna, L. Tokics, H. Lundquist, and B. Brismar, Densities in dependent lung regions during anaesthesia: atelectasis or fluid accumulation? *Acta Anaesthesiologica Scandinavia*, vol. 30, pp. 256-259, 1986.

[23] S.A. Worthy, N.L. Müller, T.E. Hartman, S.J. Swensen, S.P.G. Padley, and D.M. Hansell, Mosaic attenuation pattern on thin-section CT scans of the lung: differentiation among infiltrative lung, airway, and vascular diseases as a cause *Radiology*, vol. 205, pp. 465-70, 1997.

[24] S.A. Worthy, C.S. Parks, J.S. Kim, and N.L. Müller, Bronchiolitis obliterans after lung transplantation: high-resolution CT findings in 15 patients *AJR. American Journal of Roentgenology*, vol. 169, pp. 673-677, 1997.

[25] D.M. Lau, M.J. Siegel, C.F. Hildebolt, and A.H. Cohen, Bronchiolitis obliterans syndrome: thin-section CT diagnosis of obstructive changes in infants and young children after lung transplantation *Radiology*, vol. 208, pp. 783-788, 1998.

[26] M. Remy-Jardin, J. Remy, S. Louvegny, D. Artaud, F. Deschildre, and A. Duhamel, Airway changes in chronic pulmonary embolism: CT findings in 33 patients *Radiology*, vol. 203, pp. 355-360, 1997.

[27] C. Gückel, A.U. Wells, D.A. Taylor, F. Chabat, and D.M. Hansell, Mechanism of mosaic attenuation of the lungs on computed tomography in induced bronchospasms *Journal of Applied Physiology*, vol. 86, pp. 701-708, 1999.

[28] G.Z. Yang, F. Chabat, and D.M. Hansell, Enhancement of subtle density differences of the lung parenchyma on CT *British Journal of Radiology*, vol. 71, pp. 686-690, 1998.

[29] C. Levi, J.E. Gray, E.C. McCullough, and R.R. Hatterry, The unreliability of CT numbers as absolute values *AJR. American Journal of Roentgenology*, vol. 139, pp. 443-447, 1982.

[30] F. Chabat, S.R. Desai, D.M. Hansell, and G.Z. Yang, Gradient correction and classification of CT lung images for the automated quantification of mosaic attenuation pattern *Journal of Computer Assisted Tomography* (In Press).

[31] F. Chabat, D.M. Hansell, and G.Z. Yang. ERS Transform for the detection of bronchi on CT of the lungs. In: *Image Computing and Computer-Assisted*

Intervention - MICCAI'99, eds. C. Taylor and A. Colchester. London: Springer, 1999. pp. 235-244.

[32] C.S. Parks, N.L. Müller, S.A. Worthy, J.S. Kim, N. Awadh, and M. Fitzgerald, Airway obstruction in asthmatic and healthy individuals: inspiratory and expiratory thin-section CT *Radiology*, vol. 203, pp. 361-367, 1997.

[33] B.R. Condon, J. Patterson, and D. Wyper, Image non-uniformity in magnetic resonance imaging: its magnitude and methods for its correction *British Journal of Radiology*, vol. 60, pp. 83-87, 1987.

[34] D.A.G. Wicks, G.J. Barker, and P.S. Tofts, Correction of intensity nonuniformity in MR images of any orientation *Magnetic Resonance Imaging*, vol. 11, pp. 183-196, 1993.

[35] J. Haselgrove and M. Prammer, An algorithm for compensation of surface-coil images for sensitivity of the surface coil *Magnetic Resonance Imaging*, vol. 4, pp. 469-472, 1986.

[36] B. Johnston, M.S. Atkins, B. Mackiewich, and M. Anderson, Segmentation of multiple sclerosis lesions in intensity corrected multispectral MRI *IEEE Transactions on Medical Imaging*, vol. 15, pp. 154-169, 1996.

[37] S.K. Lee and M.W. Vannier, Post-acquisition correction of MR inhomogeneities *Magnetic Resonance in Medicine* , vol. 36, pp. 275-286, 1996.

[38] B.M. Dawant, A.P. Zijdenbos, and R.A. Margolin, Correction of intensity variations in MR images for computer-aided tissue classification *IEEE Transactions on Medical Imaging*, vol. 12, pp. 770-781, 1993.

[39] C. Brechbuhler, G. Gerig, and G. Szekely, Compensation of spatial inhomogeneity in MRI based on parametric bias estimate *Proceedings of the Fourth International Conference on Visualization in Biomedical Computing*, pp. 141-146, 1996.

[40] W.M. Wells III, W.E.L. Grimson, R. Kikinis, and F.A. Jolesz, Adaptive segmentation of MRI data *IEEE Transactions on Medical Imaging*, vol. 15, pp. 429-442, 1996.

[41] S.P. Liou and R. Jain, An approach to three-dimensional image segmentation *Computer Vision, Graphics and Image Processing*, vol. 53, pp. 237-252, 1991.

[42] C.R. Meyer and P.H. Bland, Retrospective correction of intensity inhomogeneities in MRI *IEEE Transactions on Medical Imaging*, vol. 14, pp. 36-41, 1995.

[43] J.G. Sled, A.P. Zijdenbos, and A.C. Evans, A nonparametric method for automatic correction of intensity nonuniformity in MRI data *IEEE Transactions on Medical Imaging*, vol. 17, pp. 87-97, 1998.

[44] G.Z. Yang and D.M. Hansell, CT image enhancement with wavelet analysis for the detection of small airways disease *IEEE Transactions on Medical Imaging*, vol. 16, pp. 953-961, 1997.

[45] G.A. Gould, W. McNee, A. McLean, P.M. Warren, J.J. Best, D. Lamb, and D.C. Flenley, CT measurements of lung density in life can quantitate distal airspace enlargement - an essential defining feature of human emphysema *American Review of Respiratory Disease*, vol. 137, pp. 380-392, 1988.

[46] N.L. Müller, C.A. Staples, R.R. Miller, and R.T. Abboud, "Density Mask". An objective method to quantitate emphysema using computed tomography *Chest*, vol. 94, pp. 782-787, 1988.

[47] R.J. Knudson, J.R. Standen, W.T. Kaltenborn, D.E. Knudson, K. Rehm, M.P. Habib, and J.D. Newell, Expiratory computed tomography for assessment of suspected pulmonary emphysema *Chest*, vol. 99, pp. 1357-1366, 1991.

[48] D. Marr and E. Hildreth, Theory of edge detection *Proceedings of the Royal Society London*, vol. B207, pp. 187-217, 1980.

[49] D. Marr. *Vision : a computational investigation into the human representation and processing of visual information*, San Francisco: W.H. Freeman, 1982.

[50] M.H. Hueckel, A local visual edge operators which recognizes edges and lines *Journal of the ACM*, vol. 20, pp. 634-647, 1973.

[51] R.M. Haralick, The digital step edge from zero crossings of second directional derivatives *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 6, pp. 58-68, 1984.

[52] J.F. Canny, A computational approach to edge detection *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 8, pp. 679-698, 1986.

[53] A. Martelli, Edge detection using heuristic search methods *Computer Vision and Image Processing*, vol. 2, pp. 169-182, 1972.

[54] G.A. Hewer, C. Kenney, and B.S. Manjunath, Variational image segmentation using boundary functions *IEEE Transactions on Image Processing*, vol. 7, pp. 1269-1282, 1998.

[55] D.H. Ballard and C.M. Brown. *Computer Vision*, London: Prentice-Hall, 1982.

[56] M. Kass, A. Witkin, and D. Terzopoulos, Snakes: active contour models *Proceedings of the First International Conference on Computer Vision*, pp. 259-268, 1987.

[57] T.S. Denney, Estimation and detection of myocardial tags in MR images without user-defined myocardial contours *IEEE Transactions on Medical Imaging*, vol. 18, pp. 330-344, 1999.

[58] S. Ranganath, Contour extraction from cardiac MRI studies using snakes *IEEE Transactions on Medical Imaging*, vol. 14, pp. 328-338, 1995.

[59] S.A. Hojjatoleslami and J. Kittler, Region growing: a new approach *IEEE Transactions on Image Processing*, vol. 7, pp. 1079-1084, 1998.

[60] F. Moscheni, S. Bhattacharjee, and M. Kunt, Spatio-temporal segmentation based on region merging *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 20, pp. 897-915, 1998.

[61] P.C. Chen and T. Pavlidis, Image segmentation as an estimation problem *Computer Graphics and Image Processing*, vol. 12, pp. 172-182, 1980.

[62] M. Cheriet, J.N. Said, and C.Y. Suen, A recursive threhsolding technique for image segmentation *IEEE Transactions on Image Processing*, vol. 7, pp. 918-921, 1998.

[63] F.H.Y. Chan and F.K. Lam, Adaptive thresholding by variational method *IEEE Transactions on Image Processing* , vol. 7, pp. 468-473, 1998.

[64] L. Shafarenko, H. Petrou, and J. Kittler, Histogram-based segmentation in a perceptually uniform color space *IEEE Transactions on Image Processing*, vol. 7, pp. 1354-1358, 1998.

[65] J.S. Weszka, R.N. Nagel, and A. Rosenfeld, A threshold selection technique *IEEE Transactions on Computer*, vol. 23, pp. 1322-1326, 1974.

[66] D.L. Milgram and M. Herman, Clustering edge values for threshold selection *Computer Graphics and Image Processing*, vol. 10, pp. 272-280, 1979.

[67] D.P. Panda and A. Rosenfeld, Image segmentation by pixel classification in (gray level, edge level) space *IEEE Transactions on Computer*, vol. 27, pp. 875-879, 1978.

[68] S. Watanabe, An automated apparatus for cancer prescreening: CYBEST *Computer Graphics and Image Processing*, vol. 3, pp. 350-358, 1974.

[69] K. Haris, S.N. Efstratiadis, N. Maglaveras, and A.K. Katsaggelos, Hybrid image segmentation using watersheds and fast region merging *IEEE Transactions on Image Processing*, vol. 7, pp. 1684-1699, 1998.

[70] J. Seneterre, F. Paganin, J.M. Bruel, F.B. Michel, and J. Bousquet, Measurement of the internal size of bronchi using high-resolution computed tomography (HRCT) *European Respiratory Journal*, vol. 7, pp. 598-600, 1994.

[71] A.E. McNamara, N.L. Müller, M. Okazawa, J. Arntorp, B.R. Wiggs, and P.D. Pare, Airway narrowing in excised canine lungs measured by high-resolution computed tomography *Journal of Applied Physiology*, vol. 73, pp. 307-316, 1992.

[72] S.R. Desai, A.U. Wells, F.K. Cheah, P.J. Cole, and D.M. Hansell, The reproducibility of bronchial circumference measurements using computed tomography *British Journal of Radiology*, vol. 67, pp. 257-262, 1994.

[73] S.J. Schulte, R.L. Baron, S.A. Teefey, C.A. Rohrmann, P.C. Freeny, W.P. Shuman, and M.A. Foster, CT of the extraepathic bile ducts: wall thickness and contrast enhancement in normal and abnormal ducts *AJR. American Journal of Roentgenology*, vol. 154, pp. 79-85, 1990.

[74] D.B. Reiff, A.U. Wells, D.H. Carr, P.J. Cole, and D.M. Hansell, CT findings in bronchiectasis: limited value in distinguishing between idiopathic and specific types *AJR. American Journal of Roentgenology*, vol. 165, pp. 261-267, 1995.

[75] D.G. Altman. *Practical statistics for medical research*, London: Chapman & Hall, 1993.

[76] J. Cohen, Weighted kappa: nominal scale agreement with provision for scaled disagreement or partial credit. *Psychological Bulletin*, vol. 70, pp. 213-220, 1968.

[77] R.M. Haralick and L.G. Shapiro. *Computer and robot vision*, Wokingham: Addison-Wesley, 1992.

[78] A. Rosenfeld. *Picture processing by computers*, New York: Academic Press, 1969.

[79] H.K. Yuen, J. Illingworth, and J. Kittler, Detecting partially occluded ellipses using the Hough transform *Image and Vision Computing*, vol. 7, pp. 31-37, 1989.

[80] J. Illingworth and J. Kittler, A survey of the Hough transform *Computer Vision, Graphics and Image Processing*, vol. 44, pp. 87-116, 1988.

[81] C. Sun and W.G. Wee, Neighboring gray level dependence matrix for texture classification *Computer Vision, Graphics and Image Processing*, vol. 23, pp. 341-352, 1983.

[82] B. Julesz, Textons, the fundamental elements in preattentive vision and perception of textures *The Bell Systems Technical Journal*, vol. 62, pp. 1619-1645, 1983.

[83] R.M. Haralick, Statistical and structural approaches to texture *Proceedings IEEE*, vol. 67, pp. 786-804, 1979.

[84] R.M. Haralick and R. Bosley, Texture features for image classification *Proceedings of the Third ERTS Symposium, NASA SP-351*, pp. 1219-1228, 1973.

[85] M. Sonka, V. Hlavac, and R. Boyle. *Image processing, analysis and machine vision*, London: PWS Publishing, 1999.

[86] A. Pentland, Fractal-based description of natural scenes *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 6, pp. 661-674, 1984.

[87] F. Dastous and M.E. Jernigan, Texture discrimination based on detailed measure of the power spectrum *Proceeding of the Seventh International Conference on Pattern Recognition*, pp. 83-86, 1984.

[88] Y.P. Chien and K.S. Fu, Recognition of X-ray picture patterns *IEEE Transactions on Systems, Man, and Cybernetics*, vol. SMC-4, pp. 145-156, 1974.

[89] M. Buteau and S. Makram-Ebeid, A computer vision approach to emphysema detection and quantification *Proceedings of the 18th International Conference of the IEEE Engineering in Medicine and Biology Society*, pp. 1089-1090, 1996.

[90] R. Uppaluri, T. Mitsa, M. Sonka, E.A. Hoffman, and G. McLennan, Quantification of pulmonary emphysema from lung computed tomography images *American Journal of Respiratory and Critical Care in Medicine*, vol. 156, pp. 248-254, 1997.

[91] S. Katsuragawa, K. Doi, H. MacMahon, L. Monnier-Cholley, T. Ishida, and T. Kobayashi, Classification of normal and abnormal lungs with interstitial diseases by rule-based method and artificial neural networks *Journal of Digital Imaging*, vol. 10, pp. 108-114, 1997.

[92] K.R. Heitmann, H.U. Kauczor, P. Mildenberger, T. Uthmann, J. Perl, and M. Thelen, Automatic detection of ground glass opacities on lung HRCT using multiple neural networks *European Radiology*, vol. 7, pp. 1464-1472, 1997.

[93] K. Shimizu, T. Johkoh, J. Ikezoe, K. Ichikado, J. Arisawa, H. Nakamura, S. Tamura, and T. Nagareda, Fractal analysis for classification of ground-glass opacities on high-resolution CT: an in vitro study *Journal of Computer Assisted Tomography*, vol. 21, pp. 955-961, 1997.

[94] S. Delorme, M.A. Keller-Reichenbecher, I. Zuna, W. Schlegel, and G. Van Kaick, Quantitative assessment of high-resolution computed tomography findings by computer-assisted texture-based image analysis *Investigative Radiology*, vol. 32, pp. 566-574, 1997.

[95] A. Newell, J. Shaw, and H. Simon, Report on a general problem solving program *Proceedings of the International Conference on Information Processing*, pp. 256-264, 1959.

[96] E.H. Shortliffe. *Computer-based medical consultation: MYCIN*, New York: Elsevier, 1976.

[97] B.A. Draper, R.T. Collins, J. Brolio, A.R. Hanson, and E.M. Riseman. Issues in the development of a blackboard-based Schema system for image understanding. In: *Blackboard Systems*, eds. R. Engelmore and T. Morgan. Wokingham: Addison-Wesley, 1988.

[98] H. Li, R. Deklerck, and J. Cornelis, Integration of multiple knowledge sources in a system for brain CT-scan interpretation based on the blackboard model

Proceedings of the Tenth Conference on Artificial Intelligence for Applications, pp. 336-343, 1994.

[99] D. Michie. The state of the art in machine learning. In: *Introductory readings in expert systems*, ed. D. Michie. New York: Gordon and Breach, 1982. pp. 208-229.

[100] P.E. File and P.I. Dugard, Misunderstanding between knowledge engineers and domain experts: a re-evaluation of the use of induction for medical expert systems *International Journal of Medical Informatics*, vol. 46, pp. 113-118, 1997.

[101] J. Kingston, A. Tate, and N. Shadbolt, CommonKADS model for knowledge based planning *Proceedings of the AAAI 96*, vol. 1, pp. 477-482, 1996.

[102] J. Kingston, J. Doheny, and I. Filby, Evaluation of workbenches which support the CommonKADS methodology *Knowledge Engineering Review*, vol. 10, pp. 269-300, 1995.

[103] J. Fox, N. Johns, and A. Rahmanzadeh, Disseminating medical knowledge: the PROforma approach *Artificial Intelligence in Medicine*, vol. 14, pp. 157-181, 1998.

[104] A. Vollebregt, A. Ten Teije, F. Van Harmelen, J. Van der Lei, and M. Mosseveld, A study of PROforma, a development methodology for clinical procedures *Artificial Intelligence in Medicine* , vol. 17, pp. 195-221, 1999.

[105] P. Szolovits and S.G. Pauker, Categorical and probabilistic reasoning in medicine revisited *Artificial Intelligence*, vol. 59, pp. 167-180, 1993.

[106] R.S. Patil, P. Szolovits, and W.B. Schwartz, Causal understanding of patient illness in medical diagnosis *Proceedings IJCAI-81*, pp. 893-899, 1981.

[107] G. Shafer. *A mathematical theory of evidence*, Princeton: Princeton University Press, 1976.

[108] P.R. Cohen. *Heuristic reasoning about uncertainty: an artificial intelligence approach*, Boston: Pitman, 1985.

[109] P. Szolovits and S.G. Pauker, Categorical and probabilistic reasoning in medical diagnosis *Artificial Intelligence*, vol. 11, pp. 115-144, 1978.

[110] J. Pearl, Fusion, propagation, and structuring in belief networks *Artificial Intelligence*, vol. 29, pp. 241-288, 1986.

[111] D.E. Heckerman and B.N. Nathwani, Toward normative expert systems: Part II. Probability-based representations for efficient knowledge acquisition and inference *Methods of Information in Medicine*, vol. 31, pp. 106-116, 1992.

[112] F.J. Diez, J. Mira, E. Iturralde, and S. Zubillaga, DIAVAL, a Bayesian expert system for echocardiography *Artificial Intelligence in Medicine*, vol. 10, pp. 59-73, 1997.

[113] M. Shwe, B. Middleton, D. Heckerman, M. Henrion, E. Horvitz, and H. Lehmann, Probabilistic diagnosis using a reformulation of the INTERNIST-1/QMR knowledge base *Methods of Information in Medicine*, vol. 30, pp. 241-255, 1991.

[114] M. Korver and P.J. Lucas, Converting a rule-based expert system into a belief network *Medical Informatics*, vol. 18, pp. 219-41, 1993.

[115] D. Heckerman. An empirical comparison of three inference methods. In: *Uncertainty in artificial intelligence 4*, eds. T.S. Levitt, L.N. Kanal, and J.F. Lemmer. Amsterdam: Elsevier Science, 1990. pp. 283-302.

[116] S. Andreassen, A. Rosenfalck, B. Falck, K.G. Olesen, and S.K. Andersen, Evaluation of the diagnostic performance of the expert EMG assistant MUNIN *Electroencepahlography and Clinical Neurophysiology*, vol. 101, pp. 129-144, 1996.

[117] M.W. Frohlich, P.L. Miller, and J.S. Morrow, PATHMASTER: modelling differential diagnosis as "dynamic competition" between systematic analysis and disease-directed deduction *Computers and Biomedical Research*, vol. 23, pp. 499-513, 1990.

[118] J.G. Sotos, MYCIN and NEOMYCIN: two approaches to generating explanations in rule-based expert systems *Aviation Space and Environmental Medicine*, vol. 61, pp. 950-954, 1990.

[119] D.A. Wolfram, An appraisal of INTERNIST-1 *Artificial Intelligence in Medicine*, vol. 7, pp. 93-116, 1995.

[120] S.J. Rosenman, A.E. Korten, and C.T. Levings, Computerised diagnosis in acute psychiatry: validity of CIDI-Auto against routine clinical diagnosis *Journal of Psychiatric Research*, vol. 31, pp. 581-592, 1997.

[121] D. Piraino, B. Richmond, M. Schluchter, D. Rockey, and J. Schils, Radiology image interpretation system: modified observer performance study of an image interpretation expert system *Journal of Digital Imaging*, vol. 4, pp. 94-101, 1991.

[122] E.T. Keravnou, F. Dams, J. Washbrook, C.M. Hall, R.M. Dawood, and D. Shaw, Modelling diagnostic skills in the domain of skeletal dysplasias *Computer Methods and Programs in Biomedicine*, vol. 45, pp. 239-260, 1994.

[123] P. Taylor, Computer aids for decision-making in diagnostic radiology - a literature review *British Journal of Radiology*, vol. 68, pp. 945-957, 1995.

[124] A.R. Rao and R. Jain, Knowledge representation and control in computer vision systems *IEEE Expert*, vol. 3, pp. 64-79, 1988.

[125] N. Ayache, Medical image analysis: a challenge for computer vision research *14th International Conference on Pattern Recognition*, vol. 2, pp. 1255-1256, 1998.

[126] P. Kofakis and S. Orphanoudakis. Image indexing by context. In: *A second generation PACS concept*, eds. M. Osteaux, A.R. Bakker, D. Bell, R. Mattheus, D. Meyer-Ebrecht, S. Orphanoudakis, R. Van de Velde, and T. Wendler. London: Springler-Verlag, 1992. pp. 250-293.

[127] P. Taylor, J. Fox, and A. Todd-Pokropek, Computer aids for decision-making in radiology *IEE Colloquium on Artificial Intelligence in Medicine*, pp. 511-513, 1996.

[128] K. Doi, M.L. Giger, R.M. Nishikawa, K.R. Hoffmann, H. MacMahon, R.A. Schmidt, and K.G. Chua, Digital Radiography. A useful clinical tool for computer-aided diagnosis by quantitative analysis of radiographic images *Acta Radiologica*, vol. 34, pp. 426-439, 1993.

[129] D.P. Carmody, C.F. Nodine, and H.L. Kundel, An analysis of perceptual and cognitive factors in radiographic interpretation *Perception*, vol. 9, pp. 339-344, 1980.

[130] K.L. Junck, M. McEachern, S. Grandhi, and G. Lewey, Automated routing of DICOM CT, MR and CR images: solving the pitfalls of vendor-specific DICOM implementations *Journal of Digital Imaging*, vol. 11, pp. 131-133, 1998.

[131] K. Doi, M.L. Giger, R.M. Nishikawa, K.R. Hoffmann, H. MacMahon, and R.A. Schmidt, Potential usefulness of digital imaging in clinical diagnostic radiology: computer-aided diagnosis *Journal of Digital Imaging*, vol. 8, pp. 2-7, 1995.

[132] J. Kittler, M. Hatef, R.P.W. Duin, and J. Matas, On combining classifiers *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 20, pp. 226-239, 1998.

[133] L.E. Sucar and D.F. Gillies, Probabilistic reasoning in high-level vision *Image and Vision Computing*, vol. 12, pp. 42-60, 1994.

[134] J. Serra. *Image analysis and mathematical morphology*, London: Academic Press, 1982.

[135] D.P. Naidich, N.L. Müller, E.A. Zerhouni, G. McGuiness, W.R. Webb, G.A. Krinsky, and S.S. Siegelman. *Computed tomography and magnetic resonance of the thorax*, Philadelphia: Lippincott-Raven, 1999. pp. 9-23.

[136] L. Devroye. *Non-uniform random variate generation*, New York: Springer-Verlag, 1986.

[137] M.H. DeGroot. *Probability and statistics*, Wokingham: Addison-Wesley, 1986.

[138] R.L. Cannon, J.V. Dave, and J.C. Bezdek, Efficient implementation of the fuzzy c-means clustering algorithms *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 8, pp. 248-255, 1986.

[139] F. Chabat, D.M. Hansell, and G.Z. Yang, CT Lung image classification with correction for perfusion gradient *Proceedings of the 7th International conference on image processing and its applications IPA'99*, vol. 1, pp. 402-407, 1999.

[140] M.A. Sargent, R.A. Cairns, M.J. Murdoch, H.R. Nadel, D. Wensley, and K.R. Schultz, Obstructive lung disease in children after allogeneic bone marrow transplantation: evaluation with high-resolution CT *AJR. American Journal of Roentgenology*, vol. 164, pp. 693-696, 1995.

[141] W.H. Press, S.A. Teukolsky, W.T. Vetterling, and B.P. Flannery. *Numerical recipes in C: the art of scientific computing*, Cambridge: Cambridge University Press, 1993.

[142] N.L. Müller, C.A. Staples, and R.R. Miller, Bronchiolitis obliterans organizing pneumonia: CT features in 14 patients *AJR. American Journal of Roentgenology*, vol. 154, pp. 983-987, 1990.

[143] A.A. Bankier, D. Fleischmann, R. Mallek, A. Windisch, F.W. Winkelbauer, M. Kontrus, L. Havelec, C.J. Herold, and P. Hubsch, Bronchial wall thickness: appropriate window settings for thin-section CT and radiologic-anatomic correlation *Radiology*, vol. 199, pp. 831-836, 1996.

[144] P. Armstrong, A.G. Wilson, P. Dee, and D.M. Hansell. *Imaging of diseases of the chest*, London: Mosby, 2000.

[145] F. Chabat, D.M. Hansell, and G.Z. Yang, ERS transform for the automated assessment of bronchial morphology on CT of the lungs *Proceedings of Medical Image Understanding and Analysis 2000* (In Press).

[146] R.M. Haralick and L.G. Shapiro. *Computer and robot vision*, Wokingham: Addison-Wesley, 1992.

[147] J. Schurmann. *Pattern classification. A unified view of statistical and neural approaches*, Chichester: John Wiley & Sons, Inc., 1996.

[148] A.R. Gourlay. *Computational methods for matrix eigenproblems*, London: Wiley, 1973.

[149] R. Schalkoff. *Pattern recognition: statistical, structural and neural approaches*, New York: John Wiley & Sons, Inc., 1992.

[150] F. Chabat, D.M. Hansell, and G.Z. Yang, Texture classification for the distinction of obstructive lung diseases on CT images *Proceedings of Medical Image Understanding and Analysis 2000* (In Press).

[151] F. Chabat, D.M. Hansell, and G.Z. Yang, Computerised decision-support in medical imaging *IEEE Engineering in Medicine and Biology* (In Press).

[152] R.O. Duda, P.E. Hart, and N.J. Nilsson, Subjective Bayesian methods for rule-based inference systems *Proceedings AFIPS*, pp. 1075-1082, 1976.

[153] A. Tversky and D. Kahneman. Causal schemata in judgements under uncertainty. In: *Progress in social psychology*, ed. M. Fishbein. Hillsdale: Lawrence Erlbaum, 1977.

[154] J. Pearl. *Probabilistic reasoning in intelligent systems: networks of plausible inference*, San Mateo: Morgan Kaufmann, 1988.

[155] G. McPherson. *Statistics in scientific investigation*, New York: Springer-Verlag, 1990.

[156] R.A. Hummel and S.W. Zucker, On the foundations of relaxation labelling processes *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 5, pp. 267-286, 1983.

[157] M.W. Hansen and W.E. Higgins, Relaxation methods for supervised image segmentation *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 19, pp. 949-962, 1997.

[158] P. Taylor, J. Fox, and A.T. Pokropek, The development and evaluation of CADMIUM: a prototype system to assist in the interpretation of mammograms *Medical Image Analysis*, vol. 3, pp. 321-337, 1999.

[159] G.J. Zelinsky, R.N.P. Rao, M.M. Hayhoe, and D.H. Ballard, Eye movements reveal the spatiotemporal dynamics of visual search *Psychological Science*, vol. 8, pp. 448-453, 1997.

[160] G.J. Zelinsky and D.L. Sheinberg, Eye movements during parallel-serial visual search *Journal of Experimental Psychology*, vol. 23, pp. 244-262, 1997.

[161] J.R. Barrett, N.G. Johns, E.S. deParedes, T.E. Hutchinson, and S.J. Dwyer, Visual search patterns during diagnosis of repeatedly displayed mammograms *Proceedings of the 16th Annual Conference of the IEEE Engineering in Medicine and Biology Society*, vol. 1, pp. 588-589, 1994.

[162] E.A. Krupinski, C.F. Nodine, and H.L. Kundel, A perceptually based method for enhancing pulmonary nodules recognition *Investigative Radiology*, vol. 28, pp. 289-294, 1993.

[163] S. Tu, Y. Shanar, J. Dawes, J. Winkles, A. Puerta, and N. Musen, A problem-solving model for episodic skeletal-plan refinement *Knowledge Acquisition*, vol. 4, pp. 197-216, 1992.

[164] E.A. Zerhouni, M. Boukadoum, M.A. Siddiky, J.M. Newbold, D.C. Stone, M.P. Shirey, J.F. Spivey, C.W. Hesselman, F.P. Leo, F.P. Stitik, and S.S. Siegelman, A standard phantom for quantitative CT analysis of pulmonary nodules *Radiology*, vol. 149, pp. 767-773, 1983.

[165] D.V. Paranjpe and C.J. Bergin, Spiral CT of the lungs: optimal technique and resolution compared with conventional CT *AJR. American Journal of Roentgenology*, vol. 162, pp. 561-567, 1994.

[166] G. de Geer, G. Gamsu, C. Cann, and W.B. Webb, Evaluation of a chest phantom for CT nodule densitometry *AJR. American Journal of Roentgenology*, vol. 147, pp. 21-25, 1986.

[167] G.J. Kemerink, R.J.S. Lamers, G.R.P. Thelissen, and J.M.A. van Engelshoven, CT densitometry of the lungs: scanner performance *Journal of Computer Assisted Tomography*, vol. 20, pp. 24-33, 1996.

[168] G.J. Kemerink, R.J.S. Lamers, G.R.P. Thelissen, and J.M.A. van Engelshoven, Scanner conformity in CT densitometry of the lungs *Radiology*, vol. 197, pp. 749-752, 1995.