University of Alberta

Intelligent CAD System for Infectious TB Detection on Chest Radiographs

by

Tao Xu

A thesis submitted to the Faculty of Graduate Studies and Research
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

in

Biomedical Engineering

Department of Electrical and Computer Engineering

Abstract

Computer aided detection (CAD) or diagnosis (CADx) is rapidly entering the radiology mainstream due to the conversion from film-based to digital radiographic systems and the advances in computerized image analysis techniques over the past decades. However, little CAD work in chest radiology has been done beyond lung nodules. Our research focuses on developing an intelligent CAD system for automated detection of infectious tuberculosis (TB), which has typical radiographic features such as cavity and acinar shadows.

In this thesis, I first present a general conceptual framework of the CAD system consisting of several steps, such as image preprocessing, feature extraction and classification, and final decision analysis. I then propose an efficient technique for automatic lung field segmentation using edge-region force guided active shape model (ERF-ASM) which is an important preprocessing step in the CAD system. A coarse-to-fine dual scale (CFDS) feature classification technique is then proposed for TB cavity detection. In this technique, Gaussian-model-based template matching (GTM), local binary pattern (LBP) and histogram of oriented gradients (HOG) based features are applied at the coarse scale; while circularity, gradient inverse coefficient of variation (GICOV) and Kullback-Leibler divergence (KLD) measures are applied at the fine scale. Finally, a hybrid system using combined LBP, HOG and grey level co-occurrence matrix (GLCM) based features is proposed for acinar shadows detection. Experiments over 300 chest radiographs show promising results of the proposed techniques.

Acknowledgement

First and foremost, I would like to express my gratitude to my supervisor, Dr. Mrinal Mandal, for his continuous support in the Ph.D. program. He taught me how to create my own path, rather than defining one for me. He was a great mentor for me and will always serve as a role model in my life.

I would also like to thank Dr. Irene Cheng and Dr. Anup Basu for giving me the opportunity to be a member in their group. Special thanks to Dr. Irene Cheng for her encouragement and guidance during my study.

I am grateful to Dr. Richard Long and his colleagues including Dr. James Barrie, Dr. Christopher Winter, Angela Lau, and Abdel-Halim Elamy for their clinical support of this research work and for the help in collection of the TB image database.

I must thank the research students in Multimedia Computing and Communications Laboratory (MCCL) and Multimedia Research Centre (MRC): Gencheng, Guo, Cheng Lu, Yue Li, Tao Wang, and Rui Sheng. It was very enjoyable to study and work with them.

I would also like to thank Dr. Vicky Zhao, Dr. Roger Zemp and Dr. Ivan Bajić for being in my defense committee and giving helpful advice.

Finally, with great love and respect, I acknowledge the endless love and encouragement of my parents and my wife. This thesis is dedicated to my parents, my wife and my daughter.

Table of Contents

List of Figures

List of Tables

List of Abbreviations

ACS	Active Contour-based Segmentation
Adaboost	Adaptive Boosting
AS	Acinar Shadows
ASM	Active Shape Model
BVF	Boundary Vector Flow
CAD	Computer Aided Detection
CADx	Computer Aided Diagnosis
CFC	Contour-based Features Classification
CLAHE	Contrast-Limited Adaptive Histogram Equalization
CXR	Chest X-Ray image or Chest Radiograph
DBC-GVF	GVF snake with Dirichlet Boundary Conditions
DPF	Distance Potential Force
EM	Expectation Maximization
ERF-ASM	Edge-Region Force guided Active Shape Model
FD	Fractal Dimension
FDA	Food and Drug Administration
FS	Fourier Spectrum
FVF	Fluid Vector Flow
GAC	Geodesic Active Contours
GICOV	Gradient Inverse Coefficient of Variation
GLAM	Gray Level Appearance Model
GLCM	Grey Level Co-occurrence Matrix
GTM	Gaussian-model-based Template Matching
GVF	Gradient Vector Flow
HIE	Hessian-matrix-based image enhancement
HKG	Hybrid Knowledge-Guided

HOG	Histogram of Oriented Gradients
HS	Histogram-based Statistical features
IFVF	Improved Fluid Vector Flow
JSRT	Japanese Society of Radiological Technology
KLD	Kullback-Leibler Divergence
kNN	k-Nearest Neighbor
LBP	Local Binary Pattern
LHFC	LBP and HOG-based Feature Classification
LSSP	Level-Set with Shape Priors
MR	Missing Rate
MTB	Mycobacterium tuberculosis
PA	Postero-Anterior
PCA	Principal Component Analysis
PEE	Point Evolution Equation
PIG	Poisson Inverse Gradient
PPTB	Post-primary Pulmonary Tuberculosis
RBF	Radial Basis Function
ROC	Receiver Operating Characteristic
ROI	Region of Interest
SIFT	Scale Invariant Feature Transforms-based
SVM	Support Vector Machine
TB	Tuberculosis
TM	Tchebichef Moments
TMM	Tanimoto Measure
UAHCXR	CXR from the University of Alberta Hospital
ULZ	Upper Lung Zone
WHO	World Health Organization

List of Symbols

χ	Normalization operator.
∇	Nabla operator for calculating gradient.
Δ	Laplacian operator.
I	The image.
f	The gradient of the image.
B	The boundary map of the image.
$\hat{C}^{(i)}$	The shape model on the ith image, $i = 1,2, \ldots, M$.
$z_j^{(i)}$	The GLAM of the landmark point j in the ith image.
P_k	The eigenvectors of the shape model in typical ASM.
P_z	The eigenvectors of the GLAM in typical ASM.
$v(s)$	The parametric curve.
$E_{snake}(v(s))$	The energy of the parametric curve consisting of internal and external energy: $E_{int}(v(s))$ and $E_{ext}(v(s))$.
F_{int} and F_{ext}	The internal and external forces of the snake, usually $F = -\nabla E$.
F_{static} and $F_{dynamic}$	The static and dynamic forces in F_{ext}.
CM	The co-occurrence matrix measuring the spatial dependency of two grey levels.
$LBP_{P,R}$	The LBP operator based on a circularly symmetric neighbour set of P members on a circle of radius R.
$LBP_{P,R}^{riu2}$	The rotation invariant uniform LBP operator.
H_b	The block gradient orientation histogram.
H_c	The cell gradient orientation histogram.
Γ	A kernel function used in SVM.

D	The decision function in SVM.
sign	The sign function.
h	A decision stump used in SVM
div	The divergence operator.
$\mathbf{f_{ERF}}$	The edge-region force which is a linear combination of diffused edge force $\mathbf{f_E}$ and diffused region force field $\mathbf{f_R}$
H_{σ}	The Hessian matrix with a scale σ.
$h(S,O)$	The Hausdorff distance of two sets of points S and O.

Chapter 1

Introduction

Advances in computerized image analysis have a tremendous impact on the interpretation of medical images. Different types of computer-aided detection (CAD) or diagnosis (CADx) schemes have been developed for detection and classification of various lesions in many medical imaging modalities. The majority of CAD work in chest radiology has focused on lung nodules detection, although many other lung disease diagnostic procedures can benefit from CAD, such as Tuberculosis (TB). Worldwide, TB is a leading cause of morbidity and mortality, especially in developing countries. In Canada, since TB is uncommon in Canadian-born non-Aboriginals, it may not be suspected, even when radiographic features are typical. Such underreporting has been noted in some groups including Aboriginals and immigrants from TB endemic countries and could be prevented by a CAD system. Therefore, the primary goal of this thesis is to develop a comprehensive CAD system for automatic detection of typical radiographic features of infectious TB. In this chapter, the research motivation is presented along with the discussion of the background of TB and CAD schemes for TB detection. Thesis contributions are then summarized followed by the organization of the thesis.

1.1 Background and Research Motivation

1.1.1 Status of TB in the World and in Canada

Globally, TB is a common and potentially life threatening contagious disease caused by Mycobacterium tuberculosis (MTB). TB affects most commonly the lungs but can also affect the central nervous system, the lymphatic system, the circulatory system, the genitourinary system, bones, joints, and even the skin [1]. According to the World Health

1

Organization (WHO) [2], TB is the second leading cause of death from an infectious disease worldwide. Every year, 8-10 million new people contract TB and up to 2 million die from it. For example, in 2010, there were 8.8 million incident cases of TB worldwide, 1.1 million deaths from TB among HIV-negative people and an additional 0.35 million deaths from HIV-associated TB. In Canada, TB is often ignored by heath care professionals at emergency departments due to the fact that TB is a rare disease in Canadian-born non-Aboriginals. However, it remains prevalent in some groups including Aboriginals and immigrants from TB endemic countries [3]. According to Citizenship and Immigration Canada, more than 200,000 people immigrate to Canada, and more than 35 million people visit Canada each year. Most newcomers are from developing countries where TB is a common and potentially fatal disease. Therefore, it is important to develop some reliable and easy-to-implement diagnostic technique to prevent TB, especially the infectious TB.

1.1.2 Diagnosis of TB

The diagnosis of TB is difficult. For infectious pulmonary TB, although Sputum Culture is the gold standard in diagnosis, it is not easy to implement in many cases due to specimens collection, lab infrastructure and contamination rate of the culture [1]. Compared to other detection methods, chest radiography is increasingly important in the fight against pulmonary TB due to several reasons. First, chest radiography provides a noninvasive mean to reveal unsuspected pathologic alterations. Second, it has lower radiation dose and lower cost than CT. In addition, chest X-ray machines are cheaper and more accessible, especially in developing countries [4]. However, interpreting a chest radiograph (commonly known as chest X-ray image, henceforth referred to as CXR) is extremely challenging [5]. The difficulty primarily arises because the CXRs typically do not have sharp features. The pixel intensity values of the abnormalities are in many cases similar to other regions in CXRs, and the overlapping anatomical structures further obscure the detection of the abnormalities. Fig. 1.1 shows an example of CXR of normal case. Note that the accuracy for TB diagnosis

using chest radiography is low (sensitivity: 70%-80%, specificity: 60%-70%), and large interobserver and even intraobserver differences occur [1]. Double reading by a trained radiologist or a physician is a routine medical practice for reducing the error rate. But the shortage of radiologists and a huge amount of CXRs lead to a heavy workload for radiologists. Therefore, it will be of immense value to develop a CAD system to provide "second opinion" to assist radiologists' image readings [6].

Fig. 1.1. An example of normal CXR.

1.1.3 Existing CAD Systems in Chest Radiology

Since the first application of CAD in chest radiography emerged in 1963 [7], the CAD system in chest radiography has slowly but steadily made its way into the clinical arena. Recent CAD schemes have been focused on automating one of the many aspects involved in the evaluation of CXRs [9], such as detection of lung nodules, detection of interstitial diseases, size and shape measurements. Excellent reviews of the CAD technological development can be found in [8] [9]. Several commercial CAD software in chest radiology approved by the Food and Drug Administration (FDA) has been used in clinical routine for lung nodules detection, such as ClearRead +Detect™ (formerly OnGuard, RiverainMedical) [10], xLNA Enterprise (Philips) [11], IQQA-Chest (EDDA Technology) [12]. Existing CAD systems have already proved that CAD is best for two tasks: tedious tasks, such as looking

for a "needle in a haystack" (e.g., a very small lung nodule on CXR), and tasks that involve a complex combination of multiple image features (e.g., breast mass detection at mammography). The CAD is also helpful if there is high interobserver or intraobserver variation or a lack of trained observers. The physicians' performance is improved by the effective computer output which is obtained from the computerized analysis of CXRs based on a carefully selected combination of elaborate segmentation and pattern recognition (feature extraction and classification) techniques.

So far, CAD system for nodule detection has received the most attention. Little work has been done in CAD development beyond lung nodules [9]. Early work to detect interstitial changes using simple texture features such as moments of power spectrum [13], fractal dimension [14], histogram based features [15], and 1-D wavelet coefficients [16] can be helpful for TB CAD development. However, the interstitial pattern is not a reliable radiographic cue for infectious TB, and most of these CAD systems are not automatic. Regions of interest (ROIs) usually need to be selected manually by radiologists. More recently, Shen et al. [17] proposed a hybrid knowledge-guided framework (HKG) for TB cavity detection. Although this hybrid CAD scheme is semi-automated, the detection accuracy is still not satisfactory. It is sensitive to parameters selection and misses cavities with weak boundaries. Tan et al. [18] proposed a user-guided snake model for the lung field segmentation, and extracted the first order statistical texture features (histogram based features) from the lung fields to classify the TB and normal cases. Furthermore, Patil [19] investigated both first and second order statistical texture features from TB CXRs, and suggested the second order statistical texture features is more suitable for TB opacity detection. Although these CAD schemes using texture features achieve better accuracy, it is still semi-automated, and no other high level texture features are tested and compared for TB detection. Therefore, the techniques and methods used in existing CAD systems for TB detection are still in their infancy, and the performance achievable with the latest technology is far below the performance by radiologists.

1.1.4 Proposed Intelligent CAD System for Infectious TB Detection

The objective of our research is to develop an intelligent CAD system to aid radiologists to ease their workload, and to increase the detection accuracy of infectious TB. Based on our image database and the knowledge from the cooperant TB experts, we have developed a comprehensive CAD system for automatic extraction of major radiographic features of post-primary pulmonary tuberculosis (PPTB). The PPTB is a highly infectious TB, which is more likely to have transmission events or a public health consequence than other infectious TB. The PPTB is also named as typical TB by radiologists since it has typical radiographic features compared to other infectious TB named as atypical TB. The different CXR findings between typical and atypical TB [1] are listed in Table 1.1. While typical TB has radiographic features of position, volume loss, cavitation and acinar shadows (See Fig. 1.2 for a CXR example of typical TB), atypical TB usually has no cavitation and only lower lobes involvement. Atypical TB such as HIV-infected TB has hilar and mediastinal lymphadenopathy, which represent ill-defined opacification of swollen or enlarged lymph nodes (See Fig. 1.3 for such an example).

Table. 1.1. Radiographic features of typical and atypical TB

Typical TB	Atypical TB
(i) Position: apical-posterior segments of upper lobes or superior segment of lower lobes in 90%.	(i) Hilar and mediastinal lymphadenopathy, particularly in HIV-infected TB.
(ii) Volume Loss: a hallmark of TB as a result of its destructive and fibrotic nature.	(ii) Non-cavity infiltrates and lower lobes involvement.
(ii) Cavitation: defined as a parenchymal cyst greater than 1 cm diameter; usually seen at later stage and depends upon a vigorous immune response.	
(iv) Acinar shadows: opacities representing opacification of the individual pulmonary acinus. Multiple acinar shadows create the confluent, ill-defined opacity characteristic of consolidation.	
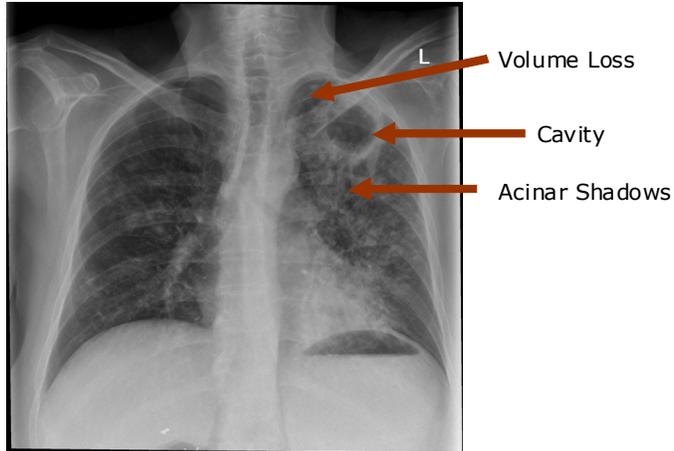
Fig. 1.2. An example of typical TB.
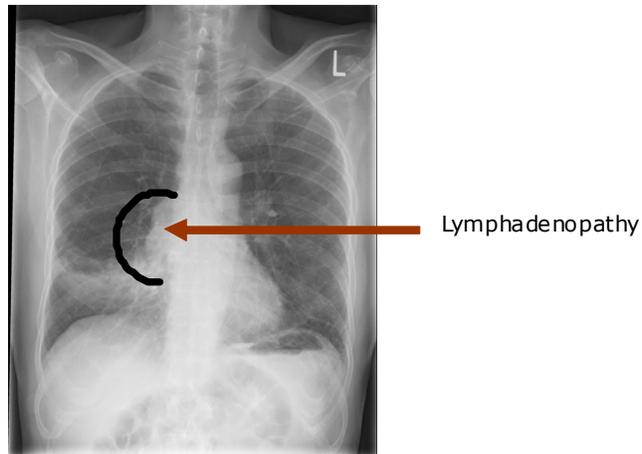
Fig. 1.3. An example of atypical TB.

To address the deficiencies of current TB-related CAD systems, we propose a conceptual framework of an automated and intelligent CAD system to detect infectious TB. The flow diagram of this framework is shown in Fig. 1.4. It consists of several steps, such as image preprocessing, identification of suspicious location, feature extraction and classification, further feature analysis and final decision analysis. It is observed that the CAD system can be divided into two major parts: CAD and CADx. The techniques discussed in this thesis are mainly focused on the CAD part. In particular, we propose an automatic lung field segmentation technique to overcome the problem of ROI's manual selection. We extract both low level (edge, regional and geometric) and high level (shape and textural) image features of those typical TB patterns (such as cavities and acinar

shadows). To overcome the variation of abnormalities, these features are fed into intelligent training based classifiers to complete the detection task. Our contribution will be explained in the next section. Note that, the scheme of our CAD and some techniques for automatic feature extraction are also expected to be applied into other CAD applications.
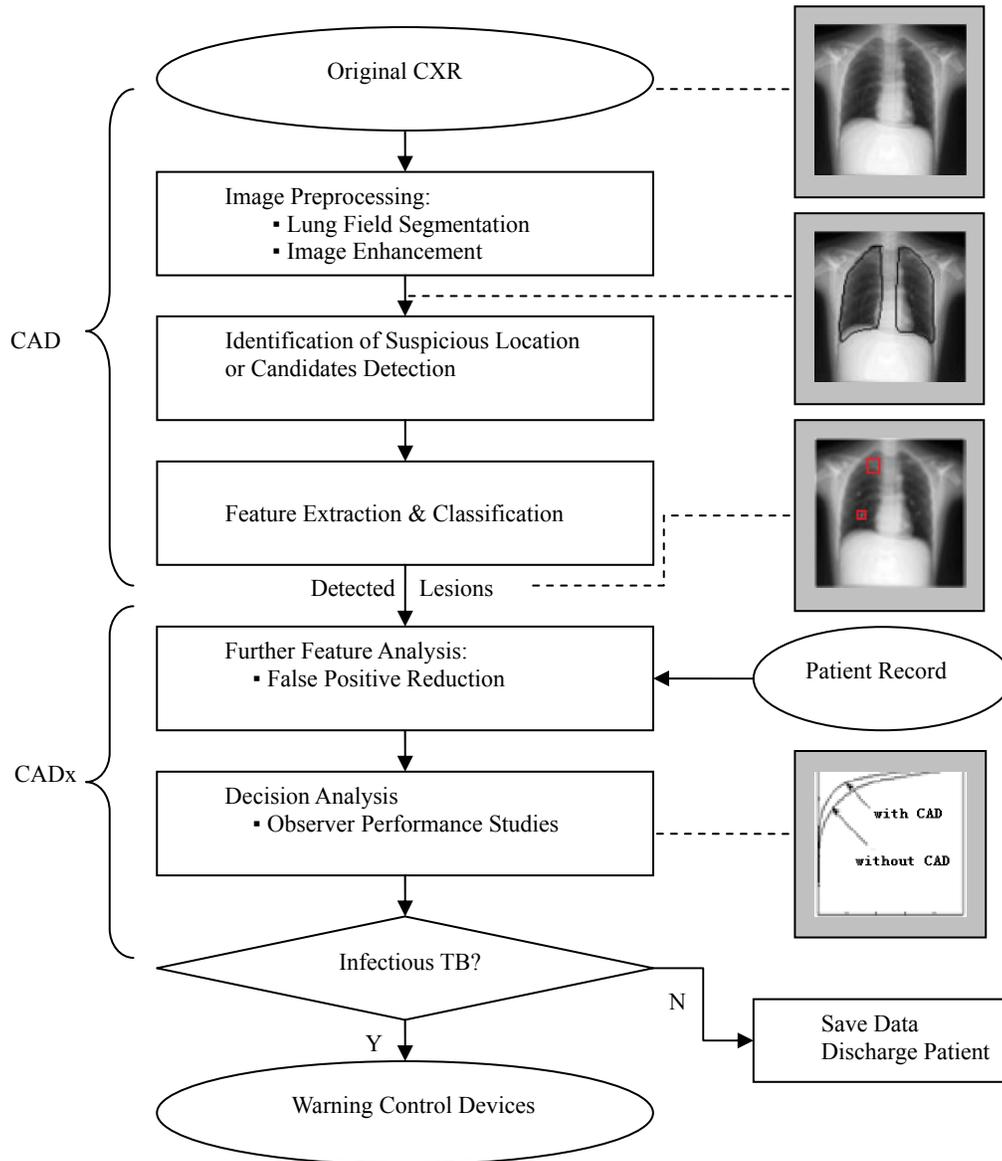
Fig. 1.4. Conceptual framework of the intelligent CAD system.

1.2 Contributions and Thesis Organization

The major contributions of this thesis are listed below:

Contribution I: An efficient technique using edge-region force guided active shape model (ERF-ASM) for automatic lung field segmentation has been proposed [20][21] (presented in Chap. 3).

Contribution II: A hybrid system using coarse-to-fine dual scale (CFDS) feature classification technique for TB cavity detection has been proposed [22][23][24] (presented in Chap. 4).

Contribution III: A hybrid system using multiple texture features based boosting technique for acinar shadows detection has been proposed [25] (presented in Chap. 5).

The rest of the thesis is organized as follows. Chap. 2 presents a review of related state-of-the-art techniques. Chap. 3 presents the proposed ERF-ASM lung field segmentation technique. The proposed hybrid systems for TB cavity and acinar shadows detection are presented in Chap. 4 and Chap.5. Chap. 6 concludes this thesis and mentions the potential future work. This is followed by the bibliography and appendix.

Chapter 2

Review of Related Works

Segmentation and pattern recognition are usually the two fundamental stages for automatic detection of abnormalities in most existing CAD systems [26][27]. Segmentation is the stage where anatomical structures (e.g. lung field) and abnormal objects (e.g. lung nodules and cavities) are delineated from background. Pattern recognition algorithms are then applied to the segmented structures to extract the essential medical features (such as shape, texture, and size), followed by the feature classification to determine normal and abnormal cases. In this chapter, we present a review of state-of-the-art segmentation and pattern recognition techniques used in CAD systems. Some of these techniques have been utilized in our proposed systems. Segmentation techniques are reviewed in Section 2.1, and the feature descriptors and classifiers are reviewed in Section 2.2.

2.1 Segmentation Techniques

A wide variety of segmentation techniques have been proposed for medical image processing, such as thresholding [56], edge tracing [57], expectation maximization [78] , active shape model (ASM) [28], and active contour model (ACM) [37]. Typically, a technique is chosen based on the application. In this section, we review two state-of-the-art segmentation techniques, ASM and ACM, which have been shown to provide excellent performance for medical images.

2.1.1 Active Shape Model (ASM)

The ASM is a statistical model of the object's shape, which iteratively deforms to fit to the object in a new image. Since shape prior knowledge is incorporated, ASM-based techniques

have been successfully applied in segmenting anatomical structures or organs in medical images, such as lung fields. The schematic of a typical ASM method [28] proposed by Cootes et al. is shown in Fig. 2.1. It contains three main stages: shape learning stage, initialization stage and segmentation stage.
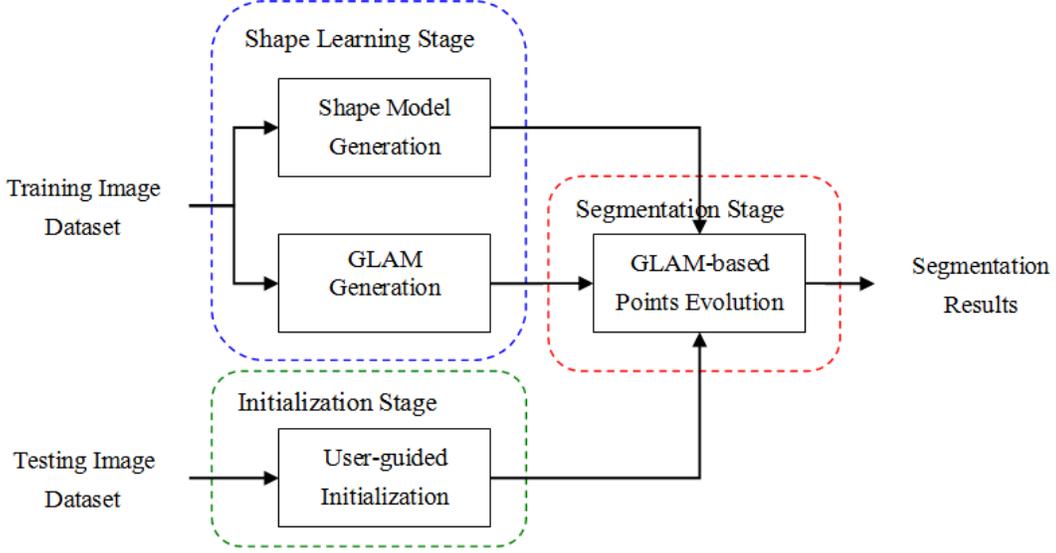
Fig. 2.1. Schematic of a typical ASM segmentation method.

2.1.1.1 Shape Learning Stage

In this stage, the shape prior and the gray level pattern of the objects are learnt by statistically analyzing the training image dataset. These two tasks are executed by shape model and gray level appearance model (GLAM) generation.

Shape Model Generation: A deformable shape model is computed from the manually annotated shapes on M training images. Assume that the object shape on the ith image is denoted by $\hat{C}^{(i)}$ and each shape has N points (henceforth referred to as the landmark points), the shape $\hat{C}^{(i)}$ can then be represented as:

$$\hat{C}^{(i)} = \left\{ (\hat{x}_1^{(i)}, \hat{y}_1^{(i)}), (\hat{x}_2^{(i)}, \hat{y}_2^{(i)}), \ldots (\hat{x}_N^{(i)}, \hat{y}_N^{(i)}) \right\}, \quad i = 1, 2, \ldots, M. \tag{2.1}$$

Before generating the shape model, all shapes must be aligned in the same way. An iterative method called generalized Procrustes analysis [29] is applied for the alignment. This method results in a set of aligned shapes $C^{(i)}$ with the least-square approach (See

details in Appendix A). Principal Component Analysis (PCA) is then performed to approximate any aligned shape in the training set as follows:

$$C^{(i)} \approx \overline{C} + P_k b_k^{(i)} \qquad (2.2)$$

where $\overline{C} = \dfrac{1}{M} \sum_{i=1}^{M} C^{(i)}$ is the mean shape of aligned shapes, $P_k=[p_1,p_2,...p_k]$ is the matrix corresponding to the first k eigenvectors, and $b_k^{(i)} =[b_1,b_2,...b_k]^T$ is a vector of weights that defines the shape parameters. Note that eigenvectors are calculated from the covariance matrix $S = \dfrac{1}{M} \sum_{i=1}^{M} (dC^{(i)})^T dC^{(i)}$, where $dC^{(i)} = C^{(i)} - \overline{C}$.

By varying shape parameters b, new shapes could be generated using Eq. (2.2). To make sure that a new shape is similar to one of those in the training set, shape parameters need to be constrained within suitable limits. Therefore, a shape model with a form of Eq. (2.2) is generated to model any allowable shapes.

GLAM Generation: Similar to the shape model generation, the PCA is applied to characterize the gray level pattern around each landmark point. In general, any region around a landmark point can be studied, but only a 1-D profile along the normal direction of the landmark point is generally considered in ASM. The GLAM of the landmark point j in the ith image can be approximated as:

$$z_j^{(i)} \approx \overline{z}_j + P_z b_z^{(i)} \qquad (2.3)$$

where $\overline{z}_j$ denotes the mean gradient profile along each point's normal direction, P_z is a matrix consisting of significant modes of gray level variations, and b_z is a vector of weights that defines the gray level parameters. Note that although being standard in typical ASM frameworks, the underlying assumption of a normal profile distribution often does not hold. Thus, optimal local features based on machine learning in the vicinity of landmark points were considered. For example, Ginneken et al. [30] used the local texture features learnt by the k-nearest neighbor (kNN) classifier. Shi et al. used the local Scale Invariant Feature

11

Transforms-based (SIFT) features [31] as the profile.

2.1.1.2 Initialization Stage

Given a test image, a close initialization is first needed to generate an approximate fit to the object shape. This is done by putting the mean shape $\overline{C} = \{(x_1, y_1), (x_2, y_2), ..., (x_N, y_N)\}$ in the center of test image with a user-guided transformation T, where T is a function of translation t, rotation θ, and scaling s of the mean shape as follows:

$$C^{(0)} = T(\overline{C}) = Q(s, \theta)\overline{C} + t = s\begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}\overline{C} + \begin{bmatrix} t_{x_j} \\ t_{y_j} \end{bmatrix} \quad (2.4)$$

where $Q(s, \theta) = s\begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$ and $t = \begin{bmatrix} t_{x_j} \\ t_{y_j} \end{bmatrix}$. The affine parameters θ, s, and t are manually chosen to ensure that the initial contour is reasonably close to the target shape. Several efforts have been made to automate the initialization [32][33][34][35][36].

2.1.1.3 Segmentation Stage

After an initial contour has been estimated, both shape model and gray level appearance model generated in the learning stage are used for segmentation using the following three iterative steps:

Step 1: Estimation of Desired Movements for Each Landmark Point Using GLAM

In this step, for each landmark point, a target point along the 1-D searching path is located by matching its gradient profile with GLAM of the corresponding landmark point. In Fig. 2.2 (a), landmark points of the initial contour are shown as (●), while target points on the object boundary are represented by (○). One of the landmark points' searching path is shown in Fig. 2.2 (b). For the landmark point (x_j, y_j), a straight line *CD* perpendicular to the straight line *AB* passing through the points (x_{j-1}, y_{j-1}) and (x_{j+1}, y_{j+1}) is treated as the searching path along the normal direction. Within this searching path, the target point is expected to have a gray level pattern that best matches the GLAM obtained from the

learning stage. Let (dx_j, dy_j) denote the movement of the jth landmark point to its target point. This procedure is repeated for all the landmark points ($1 \leq j \leq M$) to obtain the vector of desired movements dC, where $dC = \{(dx_1, dy_1), (dx_2, dy_2),\ldots, (dx_N, dy_N)\}$.
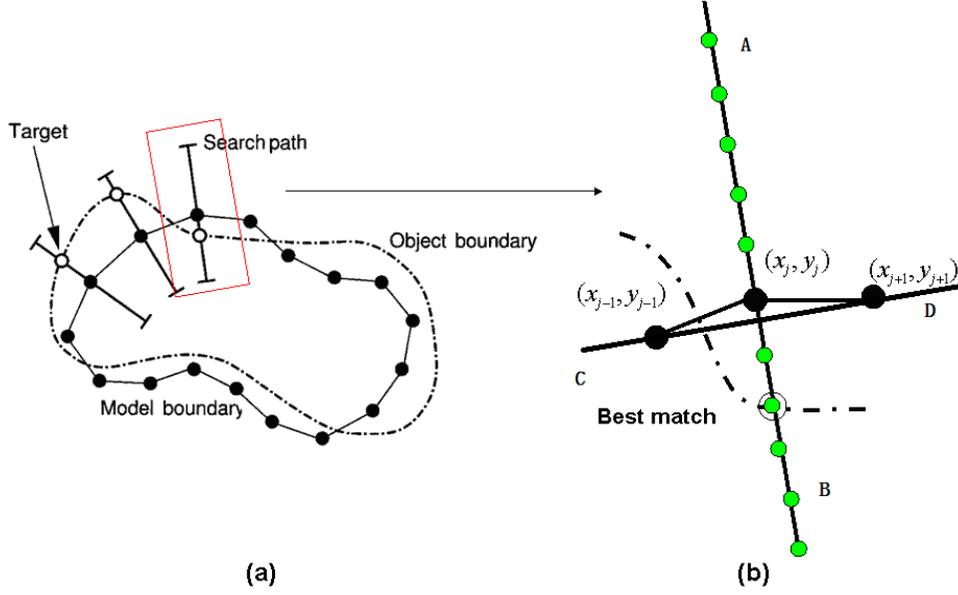
Fig. 2.2. Searching an approximate model fit for target points.

Step 2: Calculation of the Adjustments to the Affine and Shape Parameters

After obtaining the desired movements dC, the new shape $C_{new} = dC + C^{(n-1)}$ needs to be constrained within the shape model. In other words, the affine and shape parameters need to be adjusted in order to ensure the similarity between the new shape and training shapes. Let the adjustments of the affine and shape parameters be denoted as $(d\theta, ds, dt)$ and db_k, respectively. The adjustments at n^{th} iteration are then calculated as follows.

(1) $(d\theta, ds, dt)$ is obtained by aligning C_{new} to $C^{(n-1)}$ using the least-square approach explained in Appendix A.

(2) Calculate the affine adjusted shape, $C_{adj} = Q(s+ds, \theta+d\theta)C^{(n-1)}+t+dt$.

(3) Calculate the shape adjustment, $db_k = P_k^{T}(C_{new} - C_{adj})$.

(4) Calculate the final adjusted shape, $C^{(n)} = \bar{C} + P_k(b_k + db_k)$.

13

Step 3 – Iteration from Step 1 until Changes Become Negligible

Starting with the updated new contour $C^{(n)}$ after the previous affine and shape parameters adjustments, the procedure is iterated in the same way from step 1 to step 2, until movements of all landmark points are within a predefined limit. Note that the search strategy in the typical ASM is to find the target point by only inspecting the local edge gradients along the normal direction of landmark point, which can easily trap the search in a local optimum instead of reaching a global optimum.

2.1.2 Active Contour Model (ACM)

Active contours have been widely used in medical image segmentation, especially for segmenting abnormal objects with large shape variations or little shape prior such as tumors or cavities. In general, ACMs can be divided into two major types: parametric active contour (typically known as snake) and geometric active contour (typically known as level set). The major difference between the snake and the level set is in the representation of the contour. While the snake method represents the contour explicitly as a parameterized curve, the level set method implicitly represents the contour as the zero level set of a high dimensional function. We focus on the development of the snake-based methods for two reasons: 1) snakes are often faster than level sets in virtue of efficient numeric methods and lower complexity; 2) although level set handles topology changes better than snake, it may produce many false alarms in the presence of noise. Two examples of "false" objects segmented by level set are shown in Fig. 2.3.

| **(a)** | **(b)** | **(c)** | **(d)** |
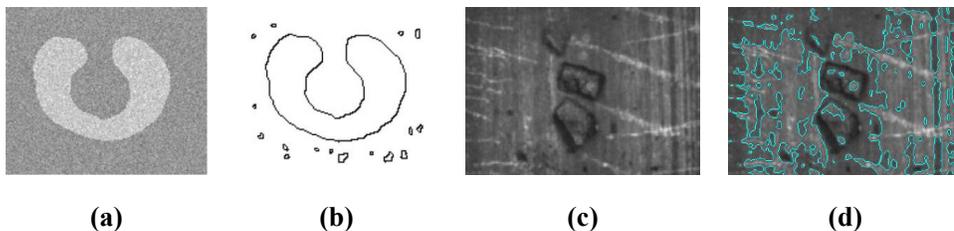
Fig. 2.3. Examples of "false" objects extracted by level set.

14

2.1.2.1 Traditional Snake and Its Development

The traditional snake model, first introduced by Kass et al. [37], is a parametric curve which deforms on the image domain and capture desired image features by minimizing the following energy functional:

$$E_{snake}(v(s)) = \int_0^1 \left[\frac{1}{2}(\alpha |v'(s)|^2 + \beta |v''(s)|^2) + E_{ext}(v(s)) \right] ds \qquad (2.5)$$

where $v(s)=(x(s),y(s))$, $s \in [0,1]$ represents a parametric curve, α and β are weighting parameters determining the strength of the model's tension (elasticity) and rigidity (stiffness), respectively, and $v'(s)$ and $v''(s)$ are the first and second derivatives of $v(s)$ with respect to s. E_{ext} denotes the external energy, the value of which is small at image features, such as edges. At the minima of Eq. (2.5), the contour must satisfy the Euler-Lagrange equation

$$\alpha v'' - \beta v'''' - \nabla E_{ext} = 0 \qquad (2.6)$$

which can be treated as a force balance equation

$$F_{int}(v) + F_{ext}(v) = 0 \qquad (2.7)$$

where $F_{int}(v) = \alpha v'' - \beta v''''$ is the internal force constraining contour's smoothness, and

$F_{ext}(v) = -\nabla E_{ext}(v)$ is the external force attracting the contour toward image features. Gradient descent and numerical approximation methods [37] are then applied to solve the Eq. (2.6). This parametric active contour model became popular quickly due to its natural ability of handling shape variations. However, the traditional external force relies directly on the image's gradient which inevitably limits the capture range.

Efforts made to improve the performance by designing different external forces can be generally classified as static forces and dynamic forces. The static forces calculated from the image remain unchanged while the dynamic forces dependent on the snake change as the snake deforms. Cohen [38] presented a static distance potential force (DPF) using negative gradient of a Euclidean distance map, which enlarges the capture range but has difficulties with concave shapes and leakage problem. Thus, Xu and Prince [39] proposed a
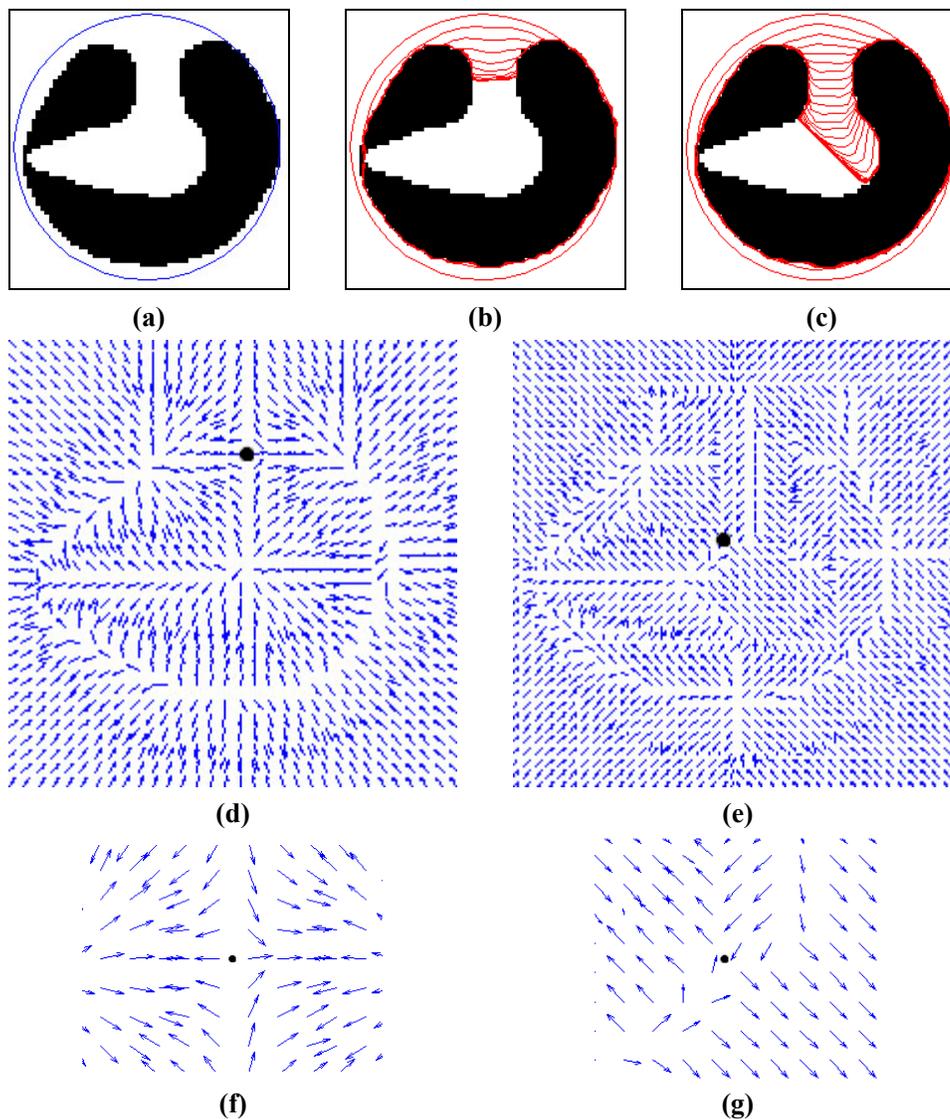
Fig. 2.4. Stationary or saddle point examples in GVF and BVF. (a) Acute concave shape with an initial contour as a blue circle; (b) Segmentation result using GVF; (c) Segmentation result using BVF;(d) Stationary point in GVF; (e) Saddle point in BVF; (f) GVF force field of the local area containing stationary point (the black point); (g) BVF force field of the local area containing saddle point.

new static external force model named as gradient vector flow (GVF) by diffusing the edge gradient. However, the capture range of GVF is still not the entire image and the diffusion is time-consuming. Sum and Cheung [40] presented a novel four-direction interpolation of the edge gradient called boundary vector flow (BVF) to extend the capture range with lower

Table. 2.1. Snakes using static external forces F_{ext}

Snakes	F_{ext}
Traditional [37]	$F_{ext} = \nabla \left\| \nabla (G_\sigma * I(x,y)) \right\|^2$ where G_σ is a 2-D Gaussian kernel with standard deviation σ.
DPF [38]	$F_{DPF} = \omega \delta \chi (\nabla d(x,y))$ where ω is a constant, $\delta = \pm 1$ controls the outward or inward direction, χ is a normalized operator, d is the Euclidean distance between a point (x,y) and the nearest point in the binary edge map.
GVF [39]	$F_{GVF} = (u,v)$, by solving $\begin{cases} g\nabla^2 u - (1-g)(u-f_x) = 0 \\ g\nabla^2 v - (1-g)(v-f_y) = 0 \end{cases}$ where $(f_x, f_y) = \nabla \left\| \nabla (G_\sigma * I(x,y)) \right\|^2$, $g = \exp(-(\frac{\|\nabla f\|}{k}))$, and k is a positive constant.
BVF[40]	$\begin{cases} F_{BVF1} = \Phi_1 = (\nabla\Psi_x, \nabla\Psi_y) \\ F_{BVF2} = \Phi_2 = (\frac{\sqrt{2}}{2}(\nabla\Psi_{xy} + \nabla\Psi_{yx}), \frac{\sqrt{2}}{2}(\nabla\Psi_{xy} - \nabla\Psi_{yx})) \end{cases}$ where Ψ_x, Ψ_y, Ψ_{xy}, and Ψ_{yx} are four potential functions computed using line-by-line interpolations in the horizontal, vertical, and two diagonal directions. The final F_{BVF} is obtained by applying two BVFs one by one.

computational cost. Note that the GVF and BVF force fields have stationary or saddle points where the sum of external forces is zero. At these points, the contour may fail to evolve to reach acute concavities. Fig. 2.4(a) shows an object with an acute concavity where the blue circle is the initial contour of the snake. Fig. 2.4(b) and (c) are the snake evolution results of the GVF and BVF, respectively. Fig. 2.4(d) and (e) are the external force fields of

the GVF and BVF, respectively. The corresponding local areas containing the stationary and saddle points in GVF and BVF are enlarged and shown in Fig. 2.4(f) and (g). It is observed that both GVF and BVF can not evolve further to reach the acute concavity due to the stationary and saddle points. Therefore, recent developments of snakes have been made by adding a dynamic force term into the static external force to overcome this problem, such as GVF snake with Dirichlet boundary conditions (DBC-GVF) [41] and fluid vector flow (FVF) [42]. The calculations of these static and dynamic external forces are listed in Table 2.1 and 2.2, respectively.

Table. 2.2. Snakes using dynamic external forces F_{ext}

Snakes	F_{ext}
DBC-GVF [41]	$F_{DBC-GVF} = (u,v)$, by solving $$\begin{cases} g\nabla^2 u - (1-g)(u-f_x) = 0 \ \text{in} \ (D \backslash C) \\ g\nabla^2 v - (1-g)(v-f_y) = 0 \ \text{in} \ (D \backslash C) \\ (u,v)_{\partial C} = \mathbf{n} \\ (u,v)_{\partial D} = \mathbf{0} \end{cases}$$ where D denotes the rectangular image domain, C is the region bounded by ∂C, $D \backslash C$ denotes the set difference of D and C, ∂D and ∂C are respectively the boundaries of D and C, and $\mathbf{n}$ is the unit outward normal to the boundary ∂C.
FVF [42]	$F_{FVF} = \chi(\chi(\nabla I(x,y)) + (1-\|f\|)\delta\chi(\nabla d(x,y)))$ where χ is a normalized operator, $\delta = \pm 1$ controls the outward or inward direction, $|f|$ is the binary boundary map, and d is the Euclidean distance between a point (x,y) (except the points on the object boundary) and the center point of the contour or the control point on the object boundary.

2.1.2.2 Fluid Vector Flow (FVF)

Among these snakes, the FVF not only has large capture range (i.e. the entire image), but also is able to extract the acute concave shapes due to its dynamic external force [42]. In FVF, the external force design is divided into two steps: vector flow initialization and FVF computation. The F_{ext} in vector flow initialization step is defined as follows:

$$F_{FVF} = \chi(f_x + \delta \cos\phi, f_y + \delta \sin\phi) \tag{2.8}$$

where χ is a normalized operator, $\delta = \pm 1$ controls the outward or inward direction, $f = (f_x, f_y)$ is the image gradient. Since ϕ is the angle between a point (x, y) (except the points on the object boundary) and the center point of the contour, $(\cos\phi, \sin\phi)$ is actually the normalized Euclidean distance between them.

To extract the curve to acute concavity, new dynamic F_{ext} is defined in FVF computation step.

$$F_{FVF} = \chi(f_x + \delta\chi(x - x_q), f_y + \delta\chi(y - y_q)) \tag{2.9}$$

where the normalized distance part is changed to the distance between a point (x, y) (except the points on the object boundary) and the point (x_q, y_q) picked up from the object boundary (named as control point). During curve evolution, the control point used for generating the additional distance force is moving along the object boundary which makes the external force dynamic and overcome stationary points problem. Since the F_{ext} on the object boundary is zero, the convergence will be achieved when the curve move close to these points.

Note that the two-step external forces in FVF could be consolidated by the following equation:

$$F_{FVF} = \chi(\chi(\nabla I(x, y)) + (1 - |f|)\delta\chi(\nabla d(x, y))) \tag{2.10}$$

where $|f|$ is the binary boundary map, and d is the Euclidean distance between a point (x,y) (except the points on the object boundary) and the center point of the contour or the control point on the object boundary.

From the Eq. (2.10), we found that the FVF is actually a combination of a static image gradient force and a dynamic DPF. It not only enlarges the capture range into the whole image, but also overcomes the convergence problem the static external forces such as GVF and BVF have by adding the dynamic term. However, it has two limitations. First, FVF suffers the same edge leakage problem as DPF. If the detected boundary encounters significant edge gaps, the FVF will fail to achieve accurate result. Fig. 2.5 shows an example, where the FVF leaks through the edge gaps. Secondly, control point selection for each deformation is critical. In vector flow initialization step, choice of the center point of the contour may not be robust. In the later FVF computation step, control point is sequentially selected along the object boundary, where all the boundary points are weighted equally. Ideally, when snaxels (points on the snake) get close to most part of the object boundary, the attraction force should be generated by the points located at the acute concavity, not at other points on the boundary. Otherwise, it may lead to inaccurate result. Fig. 2.6 shows an example of this effect. During the FVF computation step, if the control point (small black circle in Fig. 2.6) is selected from Fig. 2.6 (c) to (f), the segmentation of FVF will be inaccurate.

| (a) | (b) | (c) |
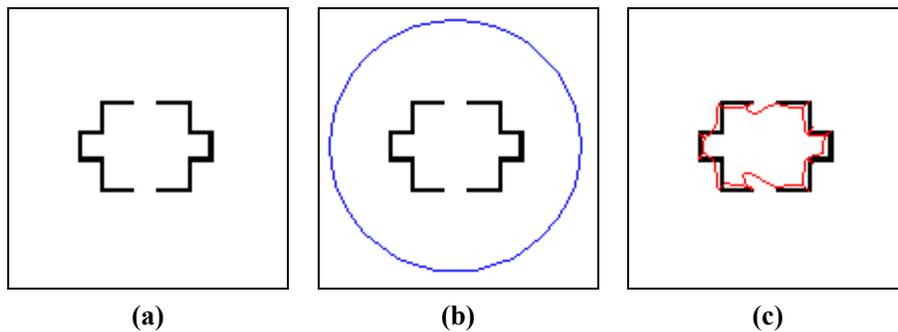
Fig. 2.5. Edge leakage problem in FVF. (a) Object with two edge gaps; (b) The initial contour; (c) Result of FVF.
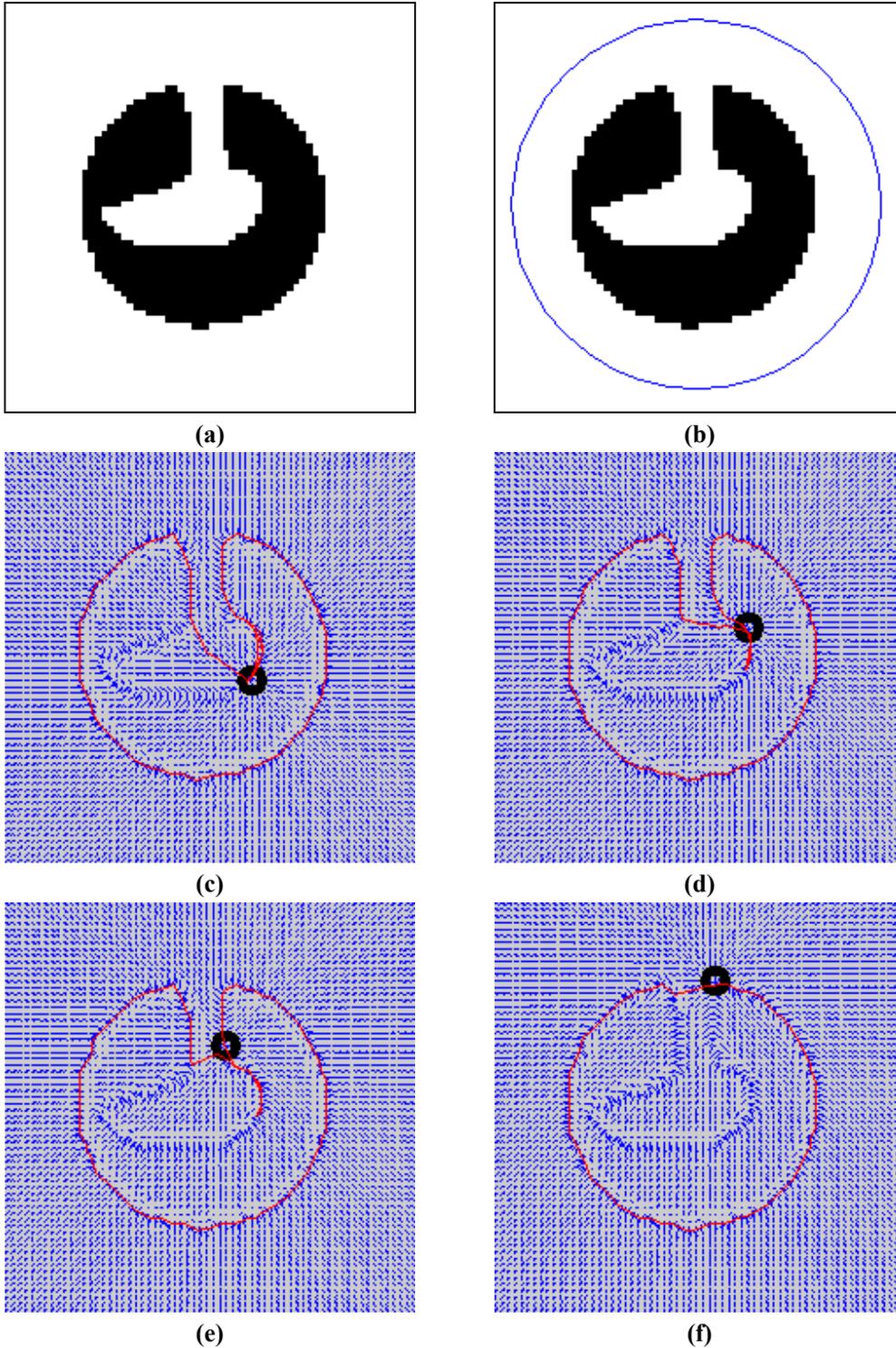
Fig. 2.6. Sensitivity of control point selection in FVF. (a) Object with acute concavity; (b) The initial contour; (c–f) Sequential results of FVF with its external force field plot; Small black circle is the control point's location and closed red contour is the evolved contour.

2.2 Feature Extraction and Classification Techniques

An image pattern can be distinguished from its neighbourhood by using an appropriate feature or a combination of features. These features are typically extracted based on different detection cues, such as intensity, shape, texture, and color [43]. Selected features are then fed to a classifier for making decisions about the presence of a specific pattern. Feature extraction is therefore a key step in pattern recognition. The classification accuracy critically depends on the quality of the feature descriptors. Various types of feature descriptors have been proposed and reviewed in the literature [43][44][45][46]. However, descriptors used in practice still depend on the intrinsic characteristics of the images and applications. Thus, in this section, we briefly provide an introduction to the feature extraction and classification methods which have a particularly significant impact on the TB detection task.

2.2.1 Feature Descriptors

2.2.1.1 Histogram based Statistical Features (HS)

Since image histogram carries rich information about image grey levels, histogram based statistical features (HS) are widely used in image texture analysis. They have also been used to discriminate the abnormal TB CXRs from normal ones in [15][18]. Given an image with L grey levels in the range [0, L-1], the normalized histogram of the image is a discrete probability density function $p(i) = \dfrac{\Psi(i)}{\Lambda}$, where $\Psi(i)$ is the number of pixels in the image with intensity i (i = 1,2,…, L-1), and Λ is the total number of pixels in the image. Instead of using histogram directly as a feature, statistical features based on histogram are usually applied in practice. Typically, six statistical features including average intensity (mean), average contrast (standard deviation), smoothness, skewness, uniformity and entropy are calculated as follows:

$$\text{mean}(\mu) = \sum_{i=0}^{L-1} ip(i)$$

$$\text{standard deviation}(\sigma) = \sqrt{\sum_{i=0}^{L-1} (i-\mu)^2 p(i)}$$

$$\text{smoothness} = 1 - \frac{1}{1+\sigma^2}$$

$$\text{skewness} = \sum_{i=0}^{L-1} (i-\mu)^3 p(i) \tag{2.11}$$

$$\text{uniformity} = \sum_{i=0}^{L-1} p(i)^2$$

$$\text{entropy} = -\sum_{i=0}^{L-1} p(i) \log_2(i)$$

2.2.1.2 Grey Level Co-occurrence Matrix (GLCM) based Features

Spatial GLCM based features are well-known and widely used texture features in computer vision. These features have also been used for texture analysis of TB CXRs [19]. Considering an image I with possible L grey levels, let the GLCM and its elements be denoted as CM and cm_{ij}, respectively ($0 \le i, j \le L-1$). Given a specific displacement (d, θ), $cm_{ij}(d, \theta)$ counts the number of pixel pairs (x_1, y_1) and (x_2, y_2) whose grey levels are i and j respectively, and $(x_2, y_2) = (x_1 + d\cos\theta, y_1 + d\sin\theta)$. It could be represented as:

$$cm_{ij}(d, \theta) = \left\| \{(x_1, y_1), (x_2, y_2) : I(x_1, y_1) = i, I(x_2, y_2) = j \} \right\| \tag{2.12}$$

where $\|\cdot\|$ is the cardinality of a set. Four angles $\theta = 0°, 45°, 90°, 135°$ are usually considered. Fig. 2.7 shows an example of generating $CM(1, 0°)$ for a given image I with $L = 3$. Before calculating the GLCM-based features, cm_{ij} usually needs to be normalized as

$cm'_{ij} = \dfrac{cm_{ij}}{\sum_{i=0}^{L-1} \sum_{j=0}^{L-1} cm_{ij}}$. Texture features, such as contrast, correlation, energy and homogeneity,

are then derived from the normalized co-occurrence matrix as follows:

$$\text{contrast} = \sum_{i=0}^{L-1} \sum_{j=0}^{L-1} cm_{ij}' (i-j)^2$$

$$\text{correlation} = \frac{\sum_{i=0}^{L-1} \sum_{j=0}^{L-1} (ij) cm_{ij}' - \mu_x \mu_y}{\sigma_x \sigma_y} \qquad (2.13)$$

$$\text{energy} = \sum_{i=0}^{L-1} \sum_{j=0}^{L-1} (cm_{ij}')^2$$

$$\text{homogeneity} = \sum_{i=0}^{L-1} \sum_{j=0}^{L-1} \frac{cm_{ij}'}{1+|i-j|}$$

where $\mu_x, \mu_y, \sigma_x,$ and σ_y are the means and standard deviations of cm_{ijx}' and cm_{ijy}',

respectively, and $cm_{ijx}' = \sum_{j=0}^{L-1} cm_{ij}'$, $cm_{ijy}' = \sum_{i=0}^{L-1} cm_{ij}'$.

0	0	0	1	2
1	1	2	1	1
2	2	1	0	0
1	1	0	2	0
0	0	1	0	1

Image *I*

	0	1	2
0	4	3	1
1	3	3	2
2	1	2	1

CM(1, 0°)
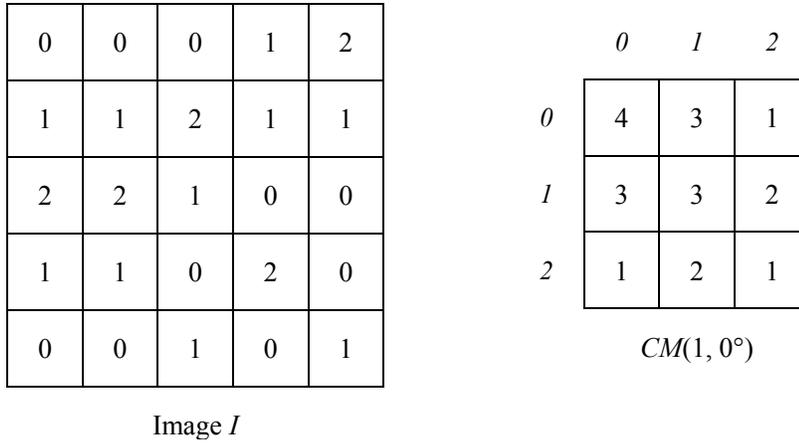
Fig. 2.7. An example of constructing a co-occurrence matrix *CM* of an image with *L* = 3 and (*d*, *θ*) = (1, 0°).

2.2.1.3 Local Binary Pattern (LBP) based Features

The LBP [47] is a hybrid texture feature widely used in image processing. It combines the traditionally divergent statistical and structural models of texture analysis. The LBP feature has some key advantages, such as its invariance to monotonic gray level changes and computational efficiency. Let the general LBP operator based on a circularly symmetric neighbor set of *P* members on a circle of radius *R* be denoted as $LBP_{P,R}$. It is obtained by

thresholding the gray levels of the neighborhood pixels with respect to the center pixel as follows

$$LBP_{P,R} = \sum_{p=0}^{P-1} 2^p s(g_p - g_c) \qquad \left[\text{Note: } s(x) = \begin{cases} 1 \text{ if } x \geq 0 \\ 0 \text{ otherwise} \end{cases} \right] \qquad (2.14)$$

where g_p, g_c are gray levels of the neighborhood pixels and center pixel, respectively. Different LBPs are generated using this LBP operator, for example, there are 256 possible binary patterns in a 3×3 neighborhood, i.e. $P = 8$, $R = 1$(Fig. 2.8 shows an example of one LBP 10010111 which is 151 in a clockwise 3×3 neighborhood). Statistical features based on the histogram of these LBP values can then be used as a texture descriptor, such as six HS features described in Eq. (2.11).

It was found that some LBPs are fundamental since they take up the vast majority of local texture patterns [47], and these fundamental patterns are called "uniform" patterns. Therefore, by introducing a uniformity measure U which corresponds to the number of spatial transitions (bitwise 0/1 changes) in the pattern, a new LBP operator, denoted as $LBP_{P,R}^{riu2}$, is calculated as follows:

$$LBP_{P,R}^{riu2} = \begin{cases} \sum_{p=0}^{P-1} s(g_p - g_c) & \text{if } U(LBP_{P,R}) \leq 2 \\ P+1 & \text{otherwise} \end{cases} \qquad (2.15)$$

The superscript ^{riu2} reflects the use of rotation invariant "uniform" patterns that have U values of at most 2. Eq. (2.15) assigns a unique label to each of the "uniform" patterns corresponding to the number "1" bits in the pattern, while the "nonuniform" patterns are grouped under the label $P+1$. Using this new LBP operator, the number of binary patterns will be greatly reduced. For example, 256 patterns are reduced to only 10 patterns in a 3×3 neighborhood. In Fig. 2.8, since the $U(LBP_{P,R}) = 4$, $LBP_{P,R}^{riu2} = 9$ belongs to the "nonuniform" patterns. Comparing to the $LBP_{P,R}$, the new rotation-invariant $LBP_{P,R}^{riu2}$ pays more attention to the "uniform" patterns and is more computationally efficient.
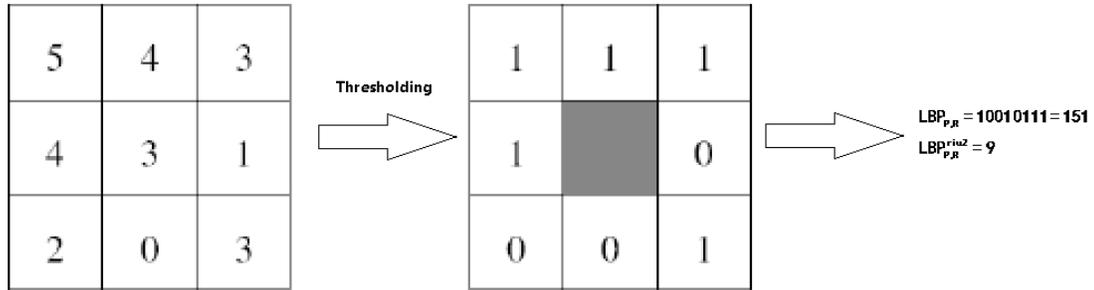
Fig. 2.8. An example of calculating LBP values in a 3×3 neighborhood.

2.2.1.4 Histogram of Oriented Gradients (HOG) based Features

The HOG feature [48], similar to Lowe's SIFT feature, is regarded as an excellent local feature descriptor to capture the edge and local shape information. It has a great advantage of being robust to changes in illumination or shadowing. The basic idea of HOG is that local object appearance and shape can often be characterized well by the distribution of local intensity gradients or edge directions. In practice, for a local image area, e.g. a rectangular image window with size of 64×64 pixels, its HOG feature is calculated as follows:

Step1. Gradient Computation: The gradient of each pixel in the window is calculated using two filter kernels: [-1, 0, 1] and [-1, 0, 1]^T. Let the magnitude and orientation of the gradient of the ith pixel ($1 \leq i \leq 4096$) be denoted by m_i and ϕ_i, respectively.

Step2. Orientation Histogram: Each window is first divided into non-overlapping small regions of equal dimension (called cells), e.g., a rectangular cell of 8×8. The orientation histogram is then generated by quantizing ϕ_i into one of the 9 major orientations: $\frac{(2k-1)\pi}{9} \pm \frac{\pi}{9}$, $1 \leq k \leq 9$. The vote of the pixel is weighted by its gradient magnitude m_i. Thus, a cell orientation histogram H_c is a vector with dimension of 1×9.

Step3. Block Normalization: For better invariance to illumination or shadowing, cells are grouped into larger spatial regions (called blocks) and the orientation histograms in a block are normalized to increase the local contrast. For example, we set the block size as 2 $\times$2 cells (i.e., 16×16 pixels), and the overlap between two neighboring blocks is 1/2 of the block size. Therefore, a whole window of 64×64 pixels contains 49 blocks. The block

divisions for an image window are shown in Fig. 2.9. The feature vector of one block H_b is concatenated by four cell histograms: $H_b = [H_{c1}\ H_{c2}\ H_{c3}\ H_{c4}]$. Note that, the orientation histogram of a block H_b is a vector with dimension of 1×36. The normalized HOG vector is then calculated as follows.

$$\hat{H}_b = \frac{H_b}{\|H_b\|} \tag{2.16}$$

where $\|.\|$ represents the L2-norm.
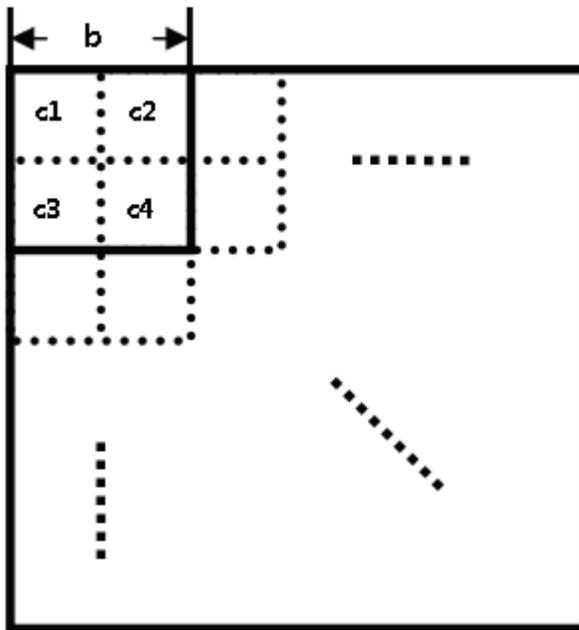
$H_b = [H_{c1}\ H_{c2}\ H_{c3}\ H_{c4}]$, where a cell orientation histogram H_{ci} is a vector with dimensions of 1×9, $i = 1, 2, 3, 4$.

Fig. 2.9. The block and cell divisions in a window image. Letters b and c stand for a block and a cell, respectively.

The final HOG feature vector of an image window (with 49 blocks) is a concatenated vector of all 49 normalized block orientation histogram ($\hat{H}_b$), and will have a dimension of 1×1764 in this example.

2.2.2 Feature Classifiers

To best utilize the prior knowledge of radiographic features in CXRs, supervised learning

based feature classification has been adopted in most CAD systems [49]. Although there are generative and discriminative approaches in supervised learning, discriminative approaches achieve more popularity in practice. That's because discriminative approaches generally outperform generative ones in classification tasks, and are not so sensitive to the presence of irrelevant features [50][51] . Therefore, in this section, we will review the basic idea of two state-of-the-art discriminative classifiers: Support Vector Machine (SVM) [52] and Adaptive Boosting (Adaboost) [53].

2.2.2.1 Support Vector Machine (SVM)

The SVM is one of the leading classifiers for its excellent performance and efficiency. In addition to performing linear classification, SVMs can efficiently perform non-linear classification using what is called the kernel trick, implicitly mapping their inputs into high-dimensional feature vector spaces [52]. For two-class classification, the optimal separating hyperplane in SVM to separate two sets of data in a feature vector space is defined by $\vec{w} \cdot \vec{x} + b = 0$, where $\vec{x}$ is the feature vector space, $\vec{w}$ is the normal vector to the hyperplane, and b is the offset of the hyperplane from the origin. Given M training feature vectors $\{\vec{x}_k, 1 \le k \le M\}$, and the corresponding ground truth classification result $\{y_k \in [1,-1], 1 \le k \le M\}$, the optimal hyperplane coefficients vector $\vec{w}$ is generated as follows:

$$\min \frac{1}{2}\|\vec{w}\|^2, \quad s.t. \ y_k\left[\Gamma(\vec{w},\vec{x}_k)+b\right] \ge 1, \ 1 \le k \le M \qquad (2.17)$$

where $\Gamma(\cdot)$ denotes a kernel function [52]. The SVM training builds a model that is able to distinguish the belonging class of any future data based on the Support Vectors obtained by the training dataset. Any new feature vector $\vec{x}_i$ is classified according to the output of the decision function:

$$D(\vec{x}_i) = \sum_{k=1}^{M} \alpha_k y_k \Gamma(\vec{x}_k, \vec{x}_i) + b \qquad (2.18)$$

where α_k is the Lagrange multiplier. If $D(\vec{x}_i) \geq 0$, it means $\vec{x}_i$ belongs to class $y = 1$,

and if $D(\vec{x}_i) < 0$, it means $\vec{x}_i$ belongs to class $y = -1$.

2.2.2.2 Adaptive Boosting (Adaboost)

Boosting is a general technique for improving performance of any given classifier. It can effectively combine a number of weak classifiers into a strong classifier which can achieve an arbitrarily low error rate given sufficient training data [53]. Compared to SVM, boosting techniques such as Adaboost, the most popular boosting method, have less parameters to tune, are more resistant to overfitting problem and do not require prior knowledge of the features [54]. The Adaboost algorithm forms a strong classifier by combining a set of weak learners linearly in an iterative manner. Although any classifier achieving accuracy larger than 50% can be used as the weaker classifier in Adaboost, a single level decision tree called decision stump is typically applied in practice to reduce the computational complexity [53]. Given N training examples $(x_1, x_2, ..., x_N)$ and corresponding labels $(y_1, y_2, ..., y_N)$ with $y_i \in \{-1, 1\}$, the pseudo-code of the Adaboost combing M decision stumps is described as follows:

   1. Initialize the observation weights, $w_i^+ = \dfrac{1}{2N^+}, w_i^- = \dfrac{1}{2N^-}$, where $i = 1, 2, ..., N$,

   $N^+ + N^- = N$, '+' and '-' represent positive and negative samples, respectively.

   2. For $m = 1$ to M

   (a) Fit a decision stump $h_m(x)$ to the training data using weights $w_i^{(m)}$, where

   $h_m(x_i) = sign(x_i - t_m)$, t_m is some feature value chosen as a threshold for the

   decision stump.

(b) Compute $err_m = \dfrac{\sum\limits_{i=1}^{N} w_i^{(m)} I(y_i \neq h_m(x_i))}{\sum\limits_{i=1}^{N} w_i^{(m)}}$, where I is an indicator function.

(c) Compute $\alpha_m = \dfrac{1}{2} \ln(\dfrac{1-err_m}{err_m})$.

(d) Update $w_i^{(m+1)} = w_i^{(m)} \exp(\alpha_m I(y_i \neq h_m(x)))$.

3. Combine weak learners into a strong classifier $D(x) = sign(\sum\limits_{m=1}^{M} \alpha_m h_m(x))$.

Chapter 3

Automatic Lung Field Segmentation

Automatic and accurate lung field segmentation is an essential step for developing an automated computer-aided diagnosis system for CXRs. Although ASM has been useful in many medical imaging applications, lung field segmentation remains a challenge due to the superimposed anatomical structures. In this chapter, we propose an efficient technique using edge-region force guided active shape model (ERF-ASM) to address the inadequacy of ASM in lung field extraction. Experimental results using both normal and abnormal CXRs show that the proposed technique provides better performance and can achieve 3-6 % improvement on accuracy, sensitivity and specificity compared to traditional ASM techniques.

3.1 Introduction

A CAD procedure in chest radiography involves various steps including lung field segmentation, feature extraction and analysis. Since only the information inside the lung is required, automatic lung field segmentation becomes a mandatory pre-processing step for computerized analysis of CXRs. Accurate segmentation results will provide useful information for the later feature extraction and analysis steps. A robust lung field segmentation technique, accurately isolating the target region from the background, will reduce the high false positive rates.

Previous efforts on lung field segmentation approaches can be classified into two categories: low level and high level methods. Low level methods focus on pixels and edges, such as thresholding, edge detection and linking, and pixel-based classification or clustering, to guide segmentation. For example, Armato et al.[56] used a combination of gray-level thresholding and contour smoothing for lung field segmentation. Duryea et al. [57]

proposed a heuristic edge tracing approach to extract lung regions. McNitt-Gray et al. [58] developed a method using feature-based classification of pixels for segmenting the lung. Shi et al. [59] proposed a Gaussian kernel-based fuzzy clustering algorithm with spatial features for automatic lung field segmentation. These techniques without prior knowledge are usually fast and automatic but often not reliable due to the variability that exists in an image, *e.g.*, abnormal anatomy and poor image quality.

High level methods, on the other hand, try to utilize the prior knowledge of generic thoracic images to overcome the shortcomings of the low level methods. Two types of high level methods have been applied on lung field segmentation. The first type is knowledge-based model matching techniques. For example, Brown et al. [60] developed a knowledge-based system which matches image edges to a lung boundary model. This knowledge-based method was later refined by Park et al. [61] and Luo et al. [62]. However, the anatomical models used in these knowledge-based systems are fixed and sensitive to the lung shape variation or patient's positioning.

Another type of high level methods is active contours incorporated with shape priors. An example of active contour techniques is level-set with shape priors (LSSP). Three types of LSSP have been successfully applied into different medical imaging modalities: edge-based LSSP, region-based LSSP, and hybrid (both edge and region-based) LSSP. For example, Leventon et al. [63] first incorporated shape priors into the edge-guided geodesic active contours (GAC) using level-set. These techniques are sensitive to image noise and weak edges. Tsai et al. [64] proposed to combine the shape constraint into the Chan and Vese (CV) model [65] to overcome the drawbacks of edge-based LSSP. However, the CV model's assumption (foreground and background regions are statistically homogenous) causes these methods to fail when segmenting images with intensity inhomogeneity. Appropriate computation of foreground and background in local regions need to be considered [66]. Bresson et al. [67] finally consolidated the GAC and CV models with shape priors under the level set framework. Similar hybrid LSSP [68] is proposed for lung

segmentation by considering specific image features in CXRs, in which the region related energy term assumes the foreground and background regions follow the Gaussian distribution, and the edge related energy term combines the low level canny edge map and castrophenic angle corner feature. The hybrid LSSP techniques and the edge-based and region-based LSSP compensate each other, but increases the complexity, such as higher computational cost and more parameters to adjust (e.g., weights of edge, region and shape related energy terms). More discussion on LSSP can be found in the Appendix B.

Active contours of snake and ASM have achieved more popularity for lung field segmentation in recent years. Iglesias et al. [69] first introduced the snake with shape constraint for detecting lung contours, and studied the influence of the different parameters of the snake. Yu et al. [70] derived the nonlinear shape statistics which were used for shape regularization in snake. Nevertheless, snake suffers from sensitivity to parameters selection and limitation of one object segmentation. ASM was then proposed by Cootes et al. [28] to overcome these problems. Ginneken et al. [30] applied an ASM with optimal texture features for lung field segmentation based on machine learning. Shi et al. [31] combined the local SIFT feature to the ASM for detecting lung field. Iakovidis et al. [71] applied ASM supported by selective thresholding for detecting lung boundaries in portable CXRs. Although ASM-based techniques have achieved satisfactory segmentation results, these techniques focus on exploiting local features and thus can fail to locate global features. Without global regulation and given the superimposed anatomical structures on CXRs, shape restrictions plus local features alone are unable to achieve optimal segmentation [72]. Hence, satisfactory convergence relies very much on how close their initialization is to the actual lung field contour. To overcome this initialization problem, a user-guided manual initialization is used in the typical ASM [28], making this method supervised and not automatic. Cootes et al. [32] proposed a multi-resolution approach, which could extend the search to include points farther away but still within a certain distance limit. Brejl et al. [33] used an exhaustive search for those situations in which the object might be located

anywhere in the image. Specific features such as intensity feature introduced by Li et al. [34] and color feature introduced by Mahoor et al. [35] are used for automatic initialization, but these techniques are task-specific. Cosio [36] proposed a more complex automatic ASM which combines pixel classification with a multi-population genetic algorithm. Similar to the exhaustive search, the computational cost is very expensive.

In this chapter, we propose an efficient automatic segmentation technique for lung boundary detection in CXRs. The proposed method applies the global edge and region force (ERF) field to guide the ASM framework, which makes the automatic initialization and segmentation stages more robust. Experimental results show that our method significantly improves the typical ASM [28] and performs better than the hybrid LSSP [68] segmentation methods in terms of accuracy, sensitivity and specificity, as well as time efficiency.

3.2 Proposed ERF-ASM for Lung Field Segmentation

As pointed out in Chap. 2, due to its local search scheme the performance of ASM relies on an initialization sufficiently close to the target. In this section, we present our global ERF-ASM for lung field segmentation to address the limitations of typical ASM. The proposed method also applies PCA analysis to learn the lung fields' shape prior, which is then applied to regularize the later ERF-based segmentation. Our contributions compared to the typical ASM [28] include: (1) Generating the initial shape automatically based on the global edge and region information.; (2) Applying a new point evaluation technique when locating the target contour of the lung field; (3) Removing the constraint of placing initial landmark points sufficiently close to the target contour. The proposed procedure (Fig. 3.1) contains three stages: shape learning stage, automatic initialization stage, and segmentation stage. The novelty lies in the last two stages, which are described below.
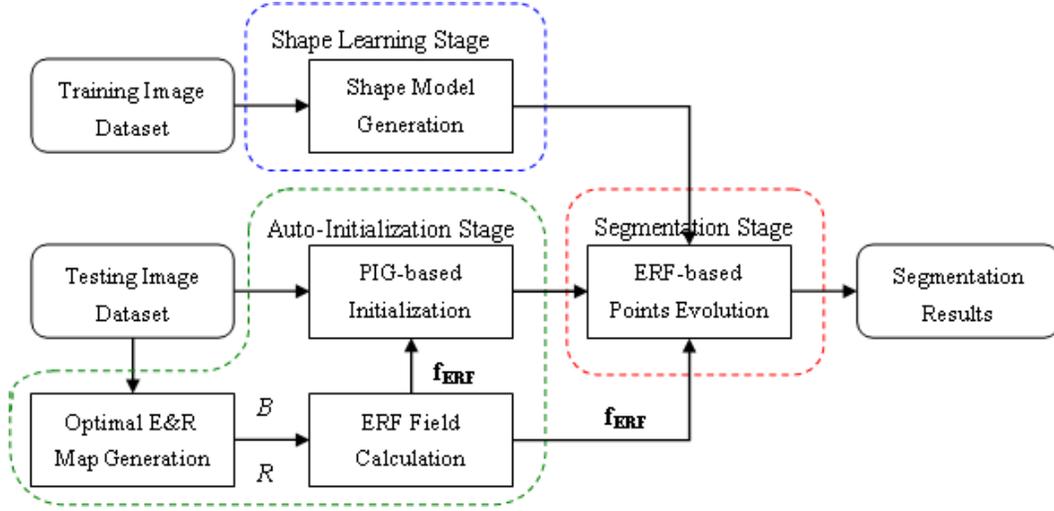
Fig. 3.1. Schematic of our proposed ERF-ASM for automatic lung field segmentation.

3.2.1 Automatic Initialization Stage

In order to obtain accurate segmentation, a preliminary but crucial step is to estimate the affine parameters for automatic initialization. Different automatic initialization approaches for ASM and their weaknesses have been discussed in the Introduction. Given the low contrast in CXRs, to achieve automatic initialization is a major challenge. In order to overcome the limitations of previous methods, we introduce global edge and region features analysis into the Poisson Inverse Gradient (PIG) initialization process.

The PIG initialization method estimates the energy field from the force field that corresponds to the object shape in the image, and uses this estimated energy field to determine the most likely initial contour for the segmentation [73]. This automatic initialization is achieved in two steps.

Step 1 – Estimating the Energy from the Given Force by Solving Poisson's Equation

The PIG approach estimates the energy field E such that the negative gradient of E is the closest vector field to the force field $\mathbf{F}$ in the L2-norm sense. Then, the problem becomes:

$$E = \arg\min_{E} \iint \left| -\nabla E(x,y) - \mathbf{F}(x,y) \right|^2 dxdy \tag{3.1}$$

It can be shown that E is the unique solution of the following Poisson's equation:

$$\Delta E = -div\, \mathbf{F}(x, y) \qquad\qquad\qquad (3.2)$$

where $\Delta = \dfrac{\partial^2}{(\partial x)^2} + \dfrac{\partial^2}{(\partial y)^2}$ is the Laplacian operator, and $div\, \mathbf{F} = \dfrac{\partial u}{\partial x} + \dfrac{\partial v}{\partial y}$ is the

divergence of $\mathbf{F} = (u, v)$. In discrete domain (e.g., images), numerical methods are often applied to calculate E.

Step 2 – Selecting the Isomodel of the Estimated Energy for Initialization

After calculating the estimated E from $\mathbf{F}$, different isovalues of E with contours of isolines can be achieved. Those isolines are candidate isomodels. Since the final result in Snake's evolution is usually expected to correspond with image edges, an ideal initial contour should be close to those edges. Thus given an edge force, the isomodel with the lowest energy is chosen as the optimal initial contour. However, using edge force alone can lead to an inaccurate initialization as explained below. To address this issue, we propose a three-step automatic initialization in our procedure.

Step 1 – Optimal Edge and Region Map Generation

In this step, an edge map B and a region map R are generated for the lung field. Since the CXRs are noisy, an edge-preserving and image smoothing scheme is necessary. We apply a speckle reducing anisotropic diffusion technique [74] for smoothing the image, preserving the strong edges, and suppressing speckle noise. The edge map B is then extracted using an edge detector, such as the Canny edge detector [75]. Fig. 3.2 (b) shows an example of the optimized edge map of the right lung image in Fig. 3.2 (a).

The region map R is generated by using a two-region segmentation technique in which the lung region is treated as the foreground and the rest of the image is treated as the background. Several region segmentation methods [76][77][78] were evaluated in our experiments, and it was found that the expectation maximization (EM) method proposed in [78] provides the best performance. Hence, in this work, we use EM technique to achieve optimal region map R. Fig. 3.2 (c) shows the region map corresponding to the image shown

in Fig. 3.2 (a). Note that the broken contour in the edge map (Fig. 3.2 (b)) is enclosed after applying region information.

Step 2 – ERF Field Calculation

The global edge and region field vector $\mathbf{f_{ERF}}$ is calculated by a linear combination of diffused edge force field $\mathbf{f_E}$ and diffused region force field $\mathbf{f_R}$ as follows:

$$\mathbf{f_{ERF}} = \lambda\mathbf{f_E} + (1-\lambda)\mathbf{f_R} \tag{3.3}$$

where the parameter $\lambda \in (0,1)$ controls the relative weight of $\mathbf{f_E}$ and $\mathbf{f_R}$ in $\mathbf{f_{ERF}}$.

The edge force is the gradient of the image's edge map B. We adopt the gradient vector flow (GVF) technique proposed by Xu et al. [39] to globally diffuse the edge force field. The GVF field $\mathbf{g} = (u(x,y), v(x,y))$ is defined by minimizing the energy functional:

$$E_{ext} = \arg\min_{\mathbf{g}} \iint \eta(u_x^2 + u_y^2 + v_x^2 + v_y^2) + |\nabla B|^2 \cdot |\mathbf{g} - \nabla B|^2 \, dxdy \tag{3.4}$$

where η is the smoothing factor, and ∇B is the gradient of the edge map B. Using variational calculus, the GVF field can be found by solving the Euler equations of Eq. (3.4). Finite difference approach is then applied to achieve its numerical solution on a discrete grid. Fig. 3.2 (d) shows the diffused edge force field $\mathbf{f_E}$ of the CXR given in Fig. 3.2 (a).

The diffused region force field is generated similarly with an assumption that the gradient of the optimal region map R is the region force field. Fig. 3.2 (e) shows the diffused region force field $\mathbf{f_R}$ of the example CXR, and Fig. 3.2 (f) is our final ERF vector field after combining the global edge and region force fields.

Step 3 – Initialization using global ERF

PIG-based method was reported in [73] to be more efficient than other automatic initialization techniques. However, using edge force alone in the PIG method can lead to an inaccurate initialization result. For example, the isoline passing through the gap between broken edges will result in an inaccurate initialization (Fig. 3.2 (h) & (j)). Therefore, we introduce region force and mean shape alignment to obtain better results. We perform two

37

processes as follows:

First, we replace the edge force with our ERF field. Thus, the Poisson Eq. (3.2) becomes:

$$\Delta E = -div\, \mathbf{f}_{\mathbf{ERF}}(x, y) \tag{3.5}$$

For example, given the ERF field in Fig. 3.2 (f), the estimated energy field by solving Eq. 3.5 is shown in Fig. 3.2 (g). Next, a mean shape alignment is applied to achieve the final optimal initialization. The mean shape, which we derived from the shape learning stage, is placed in the middle of the image, and each landmark point on the mean shape seeks its corresponding point on the isoline obtained from the previous step. The optimal initialization is achieved by minimizing the distance between the mean shape and the isoline. All the corresponding points will gradually form an approximate shape, and the initialization is eventually finished by aligning the mean shape with this approximate shape. Fig. 3.2 (i) & (k) shows our automatic initialization results of the example images, and Fig. 3.2 (h) & (j) are the results of the original PIG approach. Note that our auto-initialization technique preserves a smoother and more accurate lung field contour, and the initial contour could be inside, outside or overlapping the exact lung field.
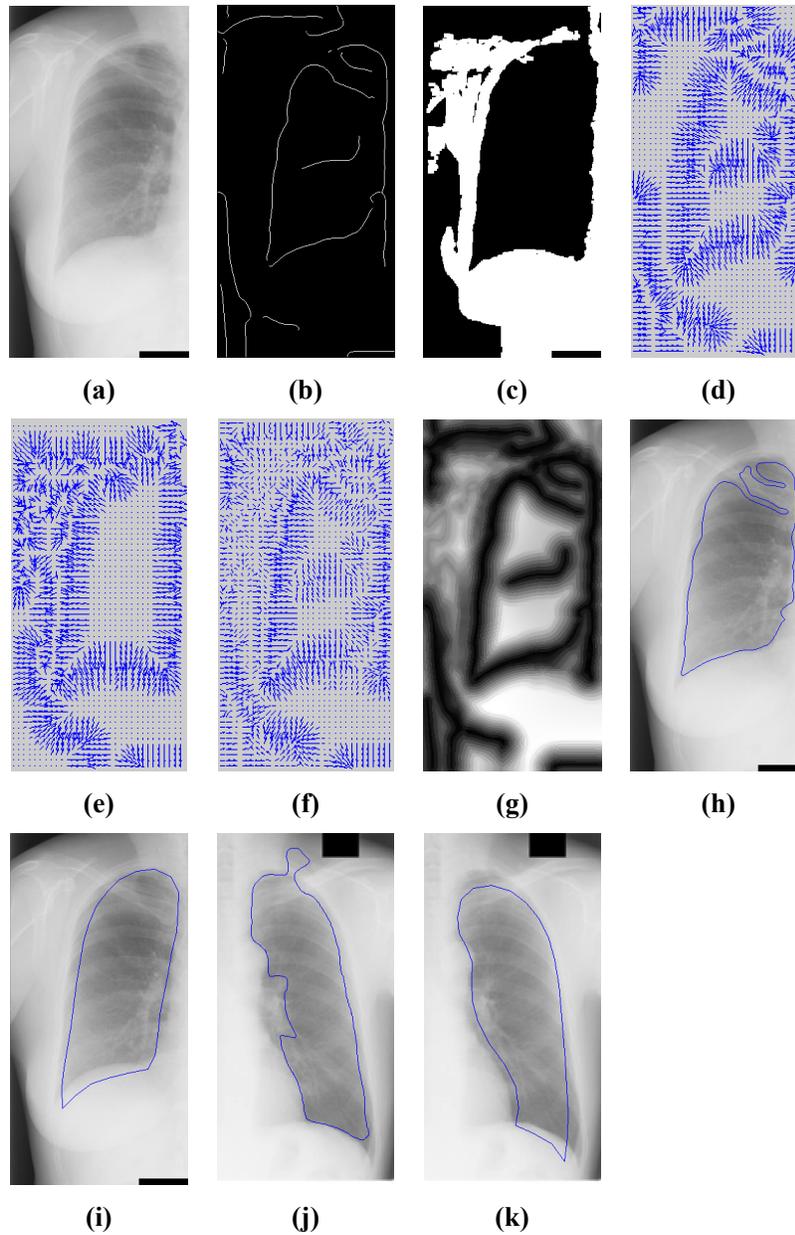
**Fig. 3.2. An example of our automatic initialization using edge and region force fields:
(a) The original right half lung image; (b) Optimized edge map B; (c) Optimized
region map R; (d) Diffused edge force field f_E; (e) Diffused region force field f_R; (f)
Total ERF field f_{ERF}; (g) Estimated energy field of the force f_{ERF}; (h) PIG-based
initialization; (i) Our initialization. Another comparison is shown in (j) PIG-based
initialization and (k) our initialization.**

3.2.2 Segmentation Stage

We introduce a new point evolution equation (PEE), which applies the global ERF field generated in the initialization stage, to lead the landmark points out of the local optima. Xu et al. [79] introduced the GVF field **g** into the ASM's points evolution. Their PEE, including a given step size w and an annealing factor k, is given as follows:

$$C = C' + w \cdot k \cdot \mathbf{g}(C') \tag{3.6}$$

where C and C' denote the current and previous contours respectively, and the annealing factor $k = \dfrac{p}{t^q}$ (constants $p, q \geq 1$) decreases as the iteration time t increases.

Steered by both the direction and magnitude of **g**, this GVF-ASM improves the robustness and accuracy of the segmentation. However, Eq. (3.6) has two limitations: first, the three parameters w, p and q need to be chosen before searching for the target points and may need to be adjusted for different test images; second, the lung contour in a CXR is usually blurred, which means that the magnitudes of the gradient vectors close to the contour edges change only slightly. But in Eq. (3.6) the power function $1/t^q$ decreases too fast compared to the corresponding changes in the gradient vectors, resulting in an early convergence. In order to reduce the complexity and improve the search accuracy, we propose a new PEE below, based on our preliminary study published in [20], which can provide more stability and accuracy:

$$C = C' + w \cdot sign(\mathbf{g}(C')) \cdot e^{-\left|\mathbf{g}(C')\right|} \tag{3.7}$$

where sign function $sign$ keeps the GVF vector's direction and the function $exp(-|\mathbf{g}(C')|)$ works as a smooth monotonically decreasing function (or a speed function), which attracts the points to edges along the lung contour. Since the boundary of the lung field does not always correspond with continuous edges, we improve the performance by incorporating both edge and region information of the lung fields. This is done by substituting the GVF vector **g** with our ERF vector $\mathbf{f_{ERF}}$. The final PEE is computed as:

$$C = C' + w \cdot sign(\mathbf{f}_{\mathbf{ERF}}(C')) \cdot e^{-\left|\mathbf{f}_{\mathbf{ERF}}(C')\right|}. \qquad (3.8)$$

We illustrate the difference between PEE (3.6) and (3.7) by both 1-D and 2-D force field examples. Fig. 3.3 shows point evolution guided by 1-D GVF field $\mathbf{g} = \{0.01, 0.02, ..., 1, -1, -0.99, ..., -0.01\}$. Note that $\mathbf{g}$ is a GVF field corresponding to 200 points position ($1 \le x \le 200$), and the strong edges lie at $x = 100$ (when $\mathbf{g} = 1$) and $x = 101$ (when $\mathbf{g} = -1$). The horizontal axes represent the iteration time t, and the vertical axes represent the landmark point's new coordinate x at different t. The graphs in Fig. 3.3 show the evolution of landmark point x over time t. Observe that in Fig. 3.3 (a) evolutions corresponding to both PEE (3.6) and (3.7) converge to the expected strong edges ($x = 100$ and $x = 101$) with parameters $w = p = q = 1$ for PEE (3.6) and $w = 1$ for PEE (3.7). However, in Fig. 3.3 (b), the evolution of landmark point using PEE (3.6) leads to an early false edge convergence with $w = q = 2$, $p = 1$, while the result of using our proposed PEE (3.7) is still accurate with $w = 2$. In other words, PEE (3.6) is sensitive to parameters selection.

Fig. 3.4 shows another example with a different 1-D GVF whose value changes more rapidly compared to the $\mathbf{g}$ considered in Fig. 3.3, $\mathbf{g} = \{0.01, 0.06, ..., 0.96, -0.96, -0.91, ..., -0.01\}$, where the strong edges lie at $x = 20$ (when $\mathbf{g} = 0.96$) and $x = 21$ (when $\mathbf{g} = -0.96$). It can be observed that the point evolution using PEE (3.6) has strong ringing effect, and jumps too far away from the strong edges position.

In the analysis depicted in Figs. 3.3 and 3.4, the proposed PEE is more accurate and stable; this can also be proved in the 2-D domain. Fig. 3.5 shows some evolution results of our point evolution process compared to Eq. (3.6).

After all the landmark points converge, the shape model's parameters are updated accordingly. The stopping criteria can be defined by the number of iterations or a threshold specifying the Euclidean distance between two consecutive iterations. In the experiments, we used a fixed number of iterations.
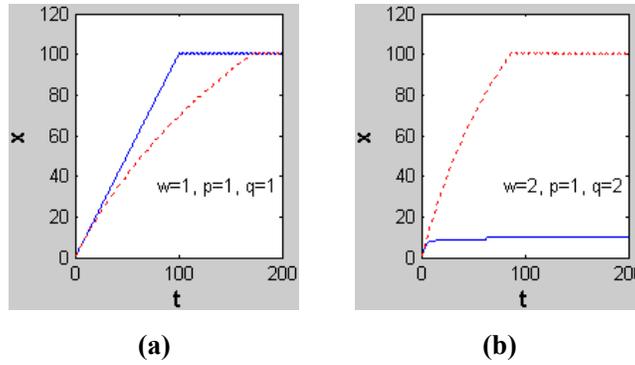
Fig. 3.3. Example of 1-D point evolution using PEE (3.6) (solid line) and PEE (3.7) (dotted line). While both converging to the strong edges in graph (a), the point evolution by using PEE (3.6) fails in graph (b) showing that it is sensitive to even a small change in the parameter values.
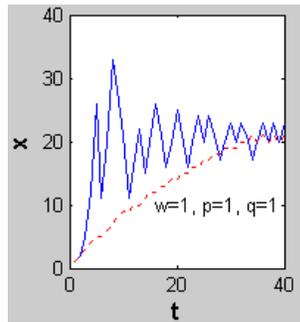
Fig. 3.4. Another example of 1-D point evolution using PEE (3.6) (solid line) and PEE (3.7) (dotted line). When changes are more rapid in a 1-D GVF field, point evolution using PEE (3.6) has strong ring effect while the result of PEE (3.7) appears more stable.
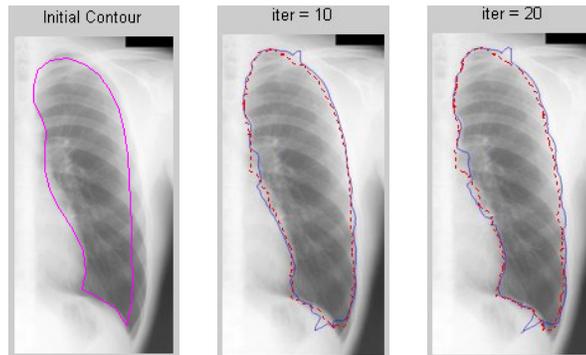
Fig. 3.5. 2-D points evolution using PEE (3.6) with w=1, p=20, q=1 (solid line) and PEE (3.7) with w=1 (dotted line). The points evolution using PEE (3.6) jumps to some false edges at iteration time t = 10 and 20 comparing to the more accurate points evolution result of PEE (3.7).

3.3 Performance Evaluation

In this section, we compare the performance of our proposed method with the typical ASM [28] and LSSP [68].

3.3.1 Database Used and Lung Field Definition

Standard postero-anterior (PA) CXRs from two different databases (normal and abnormal) are used in our experiments: (1) a publicly available Japanese Society of Radiological Technology (JSRT) database [80]; (2) a CXR database from the University of Alberta Hospital (UAHCXR). The JSRT database contains 93 normal PA-view images with resolution of 2048×2048, and the UAHCXR image database has 50 pulmonary tuberculosis PA-view images with different resolutions.

We downsample the images in the JSRT database to 512×512 pixels and images in our UAHCXR database to approximately 512×512 due to the following reasons: First, in our experiments 512×512 is a suitable resolution agreed to by radiologists to be adequate for them to define ground truth and perform visual assessments. Second, for computational efficiency it is common to segment the lung field in lower resolution, as an initial screening step and then analyze higher resolutions if necessary. For example, Refs. [30][31][60][68] use 256×256, and Refs. [62][71] use approximately 512×512.

We define the lung field as the part of the CXR which contains lungs not obscured by diaphragm, mediastinum and heart [81]. A professional radiologist helps to manually draw lung field contours for both test and training images using a mouse cursor on an interactive interface.

3.3.2 Parameters Configuration

Parameters in our proposed method are set as follows: in the shape learning stage, for

images from the JSRT database, we use an aligned training set of M = 23 left or right lung field contours, annotated with N = 30 landmark points. Twelve principal component vectors, which represent 90% of the total variance of the training set, are used to model the shape variation of the aligned lung field training set. We apply a similar process for our UAHCXR database with fewer CXRs, M = 20, N = 30. In the automatic initialization stage, the smoothness factor η is set to 0.2 empirically following previous research [39], and the optimal weight parameter λ of 0.6 in Eq. (3.3) is found to generate good results by a simulation experiment which will be explained in detail in Section 3.3.3. In the segmentation stage, the evolution step size w in Eq. (3.8) is fixed at 1, and the number of iterations is set to 100. Fig. 3.6 shows an example of the segmentation process.

<div align="center">

(a)   **(b)**   **(c)**   **(d)**

</div>
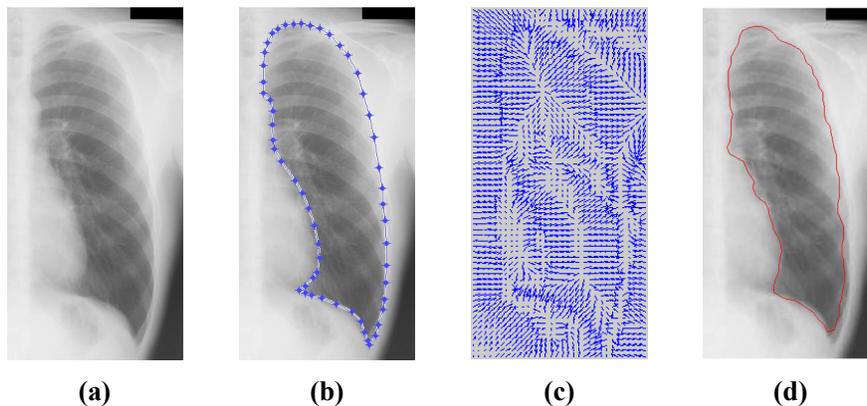
Fig. 3.6. Performance of the proposed method for lung field segmentation: (a) left lung image; (b) ground truth; (c) generated ERF field; and (d) automatic segmentation result of proposed method.

As for the parameters in the comparative techniques of typical ASM [28] and hybrid LSSP [68], same parameters are followed in both training and initialization stages. In the segmentation stage of typical ASM, the length of searching profile along normal direction is selected as 19 pixels including the landmark point. In the segmentation stage of hybrid LSSP, after carefully tuning, the weight of shape related energy term β is set to 0.005, the smoothing parameter (weight of length term) λ is set to 0.1, and the balancing parameter of low level features w is set to 0.4. (β, λ and w are the same symbols used in [68])

3.3.3 Experiments and Analysis

The automatic segmentation results of the test images are compared with ground truth data, which was manually outlined by a professional radiologist. By considering the problem as segmentation between lung and background, the performance of the proposed method is evaluated based on three indices: accuracy, sensitivity and specificity [81], which are calculated as follows:

$$\text{Accuracy}=\frac{N_{TP}+N_{TN}}{N_{TP}+N_{TN}+N_{FP}+N_{FN}}$$

$$\text{Sensitivity}=\frac{N_{TP}}{N_{TP}+N_{FN}} \tag{3.9}$$

$$\text{Specificity}=\frac{N_{TN}}{N_{TN}+N_{FP}}$$

where N_{TP} is the true positive fraction (part of the image correctly classified as lung), N_{TN} is the true negative fraction (part of the image correctly classified as background), N_{FP} is the false positive fraction (part of the image incorrectly classified as lung), and N_{FN} is the false negative fraction (part of the image incorrectly classified as background). Another evaluation parameter for comparison between methods is the time-cost factor. It is defined as the curve evolution time based on the same initial contour to the same number of iterations. Our experimental platform is Matlab 2007b on an Intel Pentium 4 CPU 2.8G Hz with 2G RAM computer.

Following the above evaluation criteria, we designed two categories of experiments: (1) simulation experiments to determine the optimal weight parameter in ERF generation; (2) comparison experiments to evaluate segmentation results between hybrid LSSP technique [68], typical ASM technique [28] and our proposed ERF-ASM method.

The simulation experiment is performed by applying our ERF technique with variable λ on 20 CXRs which are randomly selected from our image database. The segmentation results of both left and right lung images are illustrated in Fig. 3.7 (a) and (b). We select $\lambda = 0.6$ to optimize accuracy, sensitivity and specificity.
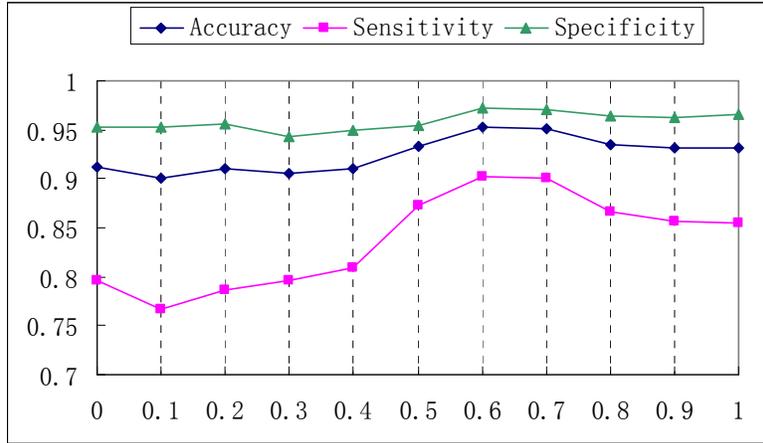
(a)
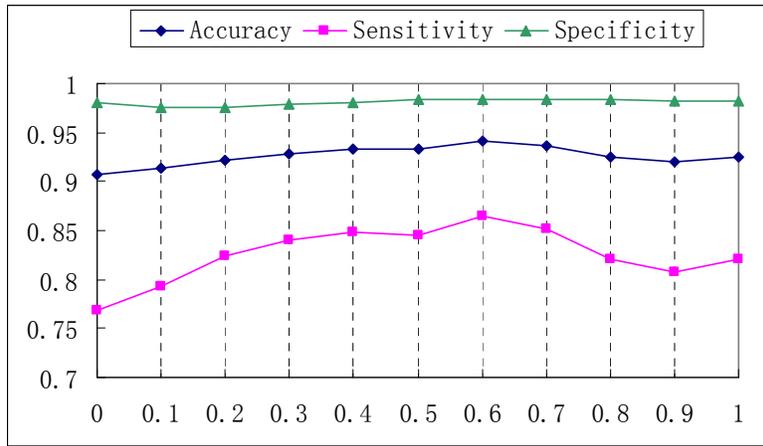
(b)

Fig. 3.7. Simulation experiment with tuning weighted parameter λ in ERF generation where x axis stands for λ value changing from 0 to 1, and y axis is the average accuracy, sensitivity and specificity values over 20 randomly selected CXRs. (a) Results of 20 left lung images; (b) Results of 20 right lung images.

The comparison experiments are applied to all the normal and abnormal images in our databases. The evaluation results are listed in Table 3.1. Accuracy, sensitivity and specificity of different methods on both left and right lung CXRs are compared. The values of these three indices are represented by the average value ± standard deviation. It is found that the proposed ERF-ASM method outperforms the typical ASM and LSSP techniques in average accuracy, sensitivity and specificity. For example, while the average accuracy of the left lung field segmentation in the JSRT image database is 91.7% using the typical ASM

technique, the accuracy of the proposed method is 95.2%. The sensitivity and specificity also increase from 81.3% to 89.7% and from 95.9% to 97.2% respectively, which indicate that the proposed method has higher true positive rate and lower false positive rate. Considering both normal and abnormal left and right lung images, the proposed method provides around 4%, 6% and 3% overall improvement on the segmentation's accuracy, sensitivity and specificity, respectively. Similar improvement can be found by comparing the proposed method to hybrid LSSP technique. In addition, the standard deviations of the three evaluation criteria drop significantly when using the proposed method, suggesting our technique is more robust. As for the time-cost, LSSP is more than 800 times and typical ASM is more than 5 times than the proposed method, indicating our method is more efficient for real-time operation. Fig. 3.8 shows some segmentation results for comparison using the same initial contour for LSSP, ASM and our proposed method on both normal and abnormal left and right lung images. It can be observed that our proposed method gives more accurate results with respect to the ground truth.

Table. 3.1. Evaluation of different segmentation methods using the same automatic initialization.

Database	Segmentation Method	Accuracy	Sensitivity	Specificity	Time Cost(s)
JSRT (Left Lung)	LSSP[68]	0.909 ± 0.020	0.857 ± 0.047	0.931 ± 0.028	320.18
	ASM[28]	0.917 ± 0.039	0.813 ± 0.083	0.959 ± 0.032	2.21
	Proposed	**0.952 ± 0.013**	**0.897 ± 0.038**	**0.972 ± 0.014**	**0.38**
JSRT (Right Lung)	LSSP[68]	0.913 ± 0.045	0.859 ± 0.061	0.944 ± 0.040	321.95
	ASM[28]	0.924 ± 0.034	0.847 ± 0.071	0.959 ± 0.035	2.38
	Proposed	**0.955 ± 0.014**	**0.912 ± 0.030**	**0.976 ± 0.018**	**0.35**
UAHCXR (Left Lung)	LSSP[68]	0.910 ± 0.013	0.835 ± 0.090	0.943 ± 0.039	321.45
	ASM[28]	0.903 ± 0.034	0.825 ± 0.087	0.933 ± 0.038	2.37
	Proposed	**0.946 ± 0.015**	**0.883 ± 0.050**	**0.969 ± 0.019**	**0.39**
UAHCXR (Right Lung)	LSSP[68]	0.924 ± 0.045	0.885 ± 0.086	0.947 ±0.036	320.22
	ASM[28]	0.899 ± 0.040	0.858 ± 0.085	0.912 ± 0.042	2.31
	Proposed	**0.953 ± 0.017**	**0.900 ± 0.044**	**0.977 ± 0.020**	**0.33**
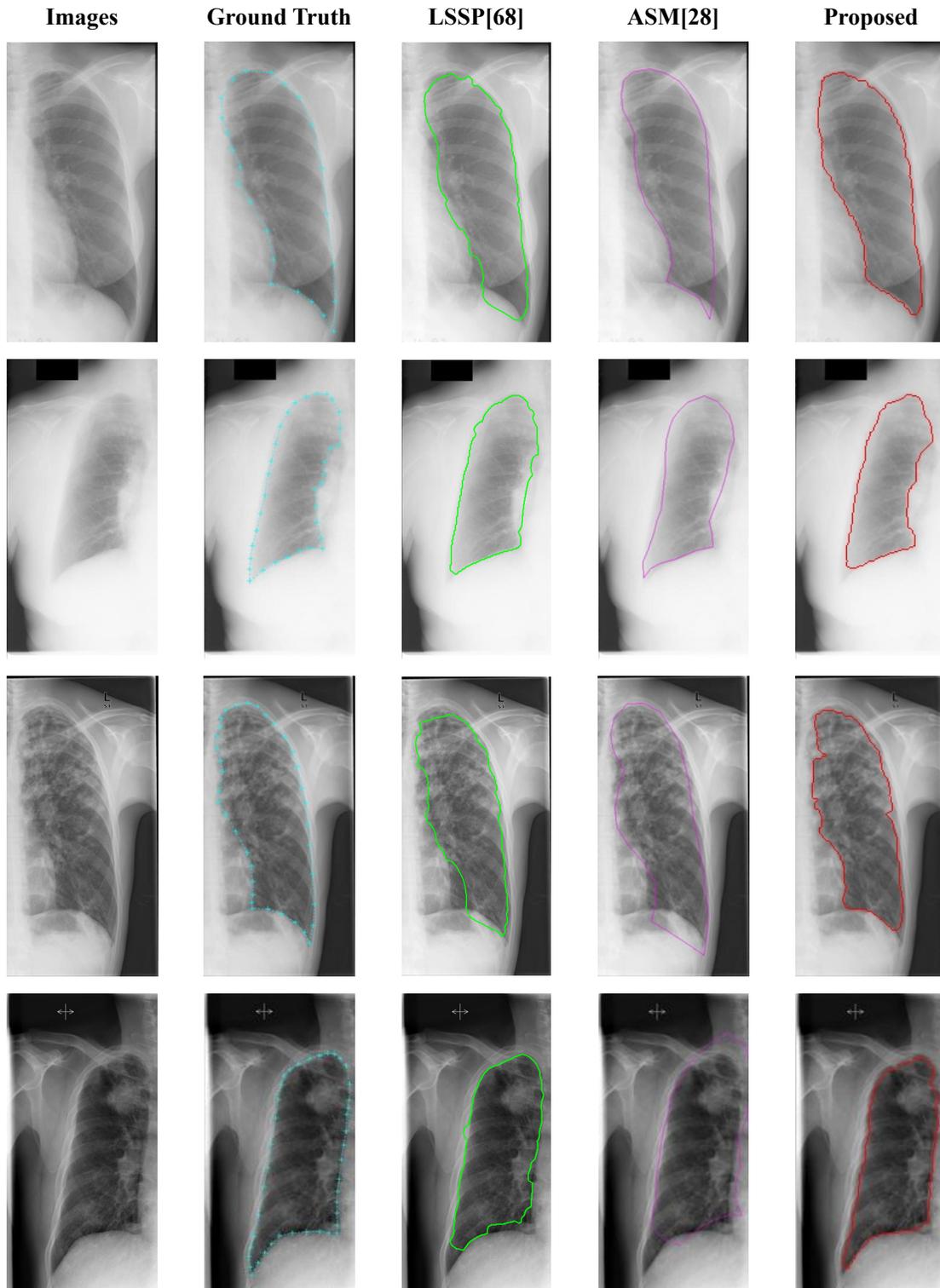
Fig. 3.8. Comparisons of segmentation results, using the same initialization and ground truth, between LSSP, typical ASM and our proposed on normal and abnormal images (both left and right lungs). Images from top row to bottom row show normal left lung, normal right lung, abnormal left lung and abnormal right lung.

The complexity involved in segmenting medical images varies depending on the anatomical structures. Lung field segmentation is challenging mainly because of the superimposed features, which blur the lung field region as well as create discontinuous edges along the contour. Although typical ASM techniques have been proved to be successful in other types of segmentation, they have their limitations when applied to lung fields. Hybrid LSSP incorporating specific edge and region features improves the accuracy of lung field segmentation in UAHCXR database. However, due to its higher computational cost and more parameters needing to be tuned, our proposed technique is much more preferable than LSSP in real-time application.

3.4 Summary

In this chapter we proposed an automatic global edge and region force (ERF) field guided method with non-linear exponential point evolution for lung field segmentation. Experimental results demonstrated that the proposed method is time efficient and improves the accuracy, sensitivity, specificity and robustness of the segmentation results, compared to the typical ASM and hybrid LSSP. Our automatic initialization also has better performance than the original PIG based initialization. Although we run our experiments on CXRs, the proposed ERF technique can easily be adapted to other image segmentation applications. Future developments include analyzing more complex ERF fields and other non-linear statistical shapes to improve the current model.

Chapter 4

Efficient TB Cavity Detection

For typical TB (PPTB), cavitation in the upper lung zone (ULZ) is a strong indicator that the disease has developed into a highly infectious state. Automatic and accurate lung field segmentation proposed in Chap. 3 has already paved the way for finding the location of ULZ. Our next task is to detect the presence of cavities in this area. In order to address the deficiency of existing computer-aided TB cavity detection methods, in this chapter, we propose an efficient coarse-to-fine dual scale (CFDS) technique for cavity detection in CXRs. Gaussian-model-based template matching, LBP and HOG-based features are applied at the coarse scale; while circularity, gradient inverse coefficient of variation (GICOV) and Kullback-Leibler divergence (KLD) measures are applied at the fine scale. Experimental results demonstrate that the proposed technique outperforms other existing techniques with respect to true cavity detection rate and segmentation accuracy.

4.1 Introduction

Cavitation in the ULZ is a typical radiographic feature of PPTB [1][82]. In chest radiography, a cavity is typically defined as a parenchymal cyst greater than one cm in diameter, containing either air or fluid or both [83]. Since the cavities are created by tissue necrosis within nodules or masses, their radiographic patterns are usually demonstrated as annular rings with variable wall thickness. Fig. 4.1(a) shows a CXR with a typical cavity (inside the rectangle region), which manifests as a focal lucent area on the image and appears as a "hole" in the patient's left ULZ. However, these holes might be blurred due to the overlapping projection of anatomical structures or some other abnormalities in the neighborhood, which makes the identification of cavities a difficult task for radiologists. Fig. 4.1(b) is another example of a TB cavity obscured by the left clavicle. Fig. 4.1(c) shows an

example where the cavity is overlapped with airspace opacities.
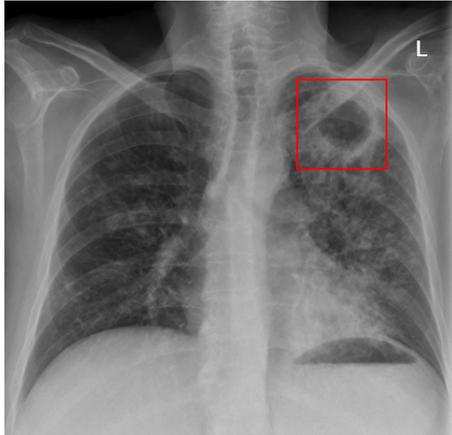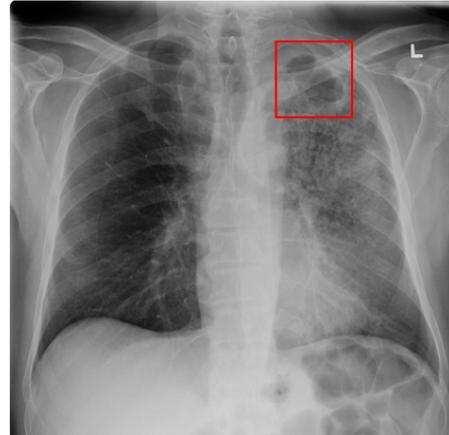
(a)   (b)
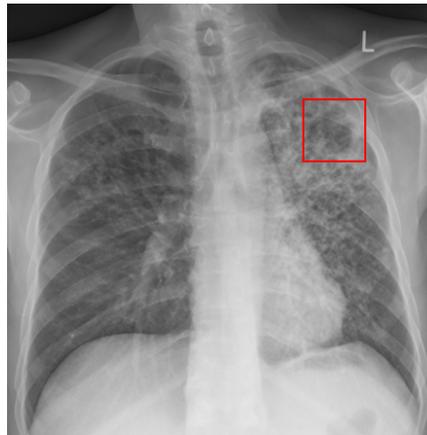
(c)

Fig. 4.1. Examples of cavities in CXRs (in the red rectangle).

So far, insufficient research has been done for efficient detection of TB cavities. Shen et al. [17] recently proposed a hybrid knowledge-guided framework (HKG) for TB cavity detection, which contains three major steps. In Step 1, the cavity candidates are detected using adaptive thresholding on the mean-shift clustered CXRs. In Step 2, a segmentation technique is applied to the candidates to generate contours of important objects present in the CXR. In Step 3, the contour-based circularity and GICOV features are extracted for the final cavity classification using a Bayesian classifier. Although, this technique provides a good performance, it has several limitations. First, due to cavity size variation and the occlusion from neighboring superimposed anatomical structures, the mean shift cluster

result is sensitive to the parameter values used. Secondly, the adaptive threshold, which is a quadratic polynomial of GICOV score, does not perform well when the cavity boundary is weak. These two limitations lead to a high missing rate of true cavities. To overcome these problems, we propose a dual scale feature classification strategy for TB cavity detection in CXRs. First, a *coarse* feature classification step is performed to detect the cavity candidates by capturing the geometric, textural and gradient features in the lung field. Second, a Hessian matrix based technique is applied to enhance the cavity candidates, which leads to a more accurate contour segmentation. Finally, *fine* features based on the shape, edge and region are extracted from the segmented contours for the final cavity classification. Experimental results show that the performance of the proposed candidates detection, segmentation and cavity classification modules are superior compared to the results obtained using other related CAD systems.

4.2 Proposed CFDS for TB Cavity Detection

Computer-aided pattern recognition in CXRs is comparatively more challenging than pattern recognition in medical images of other body parts because of the rib cage and other superimposed anatomical structures in the lung field as illustrated in Fig. 4.1. After examining the geometric, textural and photometric characteristics of TB cavities, we propose a coarse-to-fine feature classification technique for cavity detection. Fig. 4.2 shows a schematic of the proposed technique. It is observed that there are three major steps: (i) coarse feature classification, (ii) contour segmentation and (iii) fine feature classification. A CXR is first divided into patches. In the first step, a coarse feature classification is performed on each image patch to identify candidates which are suspected to contain cavities. Two modules are used to capture the coarse features: **G**aussian-model-based **T**emplate **M**atching (GTM), and **L**BP and **H**OG-based **F**eature **C**lassification (LHFC). In the second step, contours of the chosen candidates are segmented using two modules: **H**essian-matrix-based **I**mage **E**nhancement (HIE) and **A**ctive **C**ontour-based **S**egmentation

52

(ACS). The HIE is used to boost the cavity edges. The Edge-based ACS is then applied to segment the enhanced images. In the third step, a **C**ontour-based **F**eature **C**lassification (CFC) module is applied. Fine features including shape, edge and region are extracted from the contours. Cavity classification is then performed based on these features. A detailed description of these five modules is presented in the following sections.
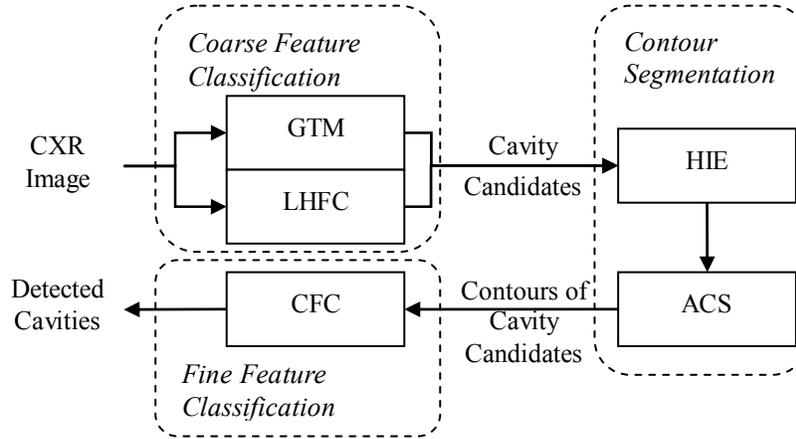
Fig. 4.2. Schematic of the proposed CAD framework.

4.2.1 Gaussian-model-based Template Matching (GTM)

The template matching is a widely used technique in pattern recognition, where the presence of a pattern in an image is detected by comparing different parts of an image with a reference pattern known as template. In many TM techniques, instead of comparing a given template directly, a transformation of the template is matched with similar transformation of a candidate region using a similarity measure. Normalized cross correlation is often used to measure similarity because of its fast implementation using the fast Fourier transform. Since traditional template matching is sensitive to rotation and scale, rotation and scale invariant transform such as Fourier-Mellin transform [84], or ring-projection transform [85] can be incorporated into TM. However, these transforms provide good results only when a cavity shape/size deviates very little from the template shape/size. To avoid missing true cavities, a solution is to use a large set of templates

covering different cavity sizes and rotation angles.

Using a large set of templates can be computationally expensive but still cannot guarantee to detect all cavities. Therefore, the proposed technique makes use of prior knowledge given by TB experts to generate a customized template database specific for TB cavities. Observe that in the "hole" like cavity shown in Fig. 4.1(a), line-cut intensity profiles in various directions of the cavity region appear similar. Fig. 4.3(a) shows the magnified region of a cavity, and Fig. 4.3(b)-(e) show plots of the four intensity lines passing through the image center at 0°,45°,90°,135°. Each line's intensity profile appears as a bi-modal Gaussian function. Based on the similarity of these intensity profiles, it is reasonable to mimic the cavity pattern using rotationally symmetric pattern such as 2D circular or elliptical Gaussian ring distribution (as shown in Fig. 4.3(f)). Note that if a line-cut intensity profile of Fig. 4.3(f) is calculated, a bimodal Gaussian distribution is obtained where the two major peaks correspond to the two sides of the ring.
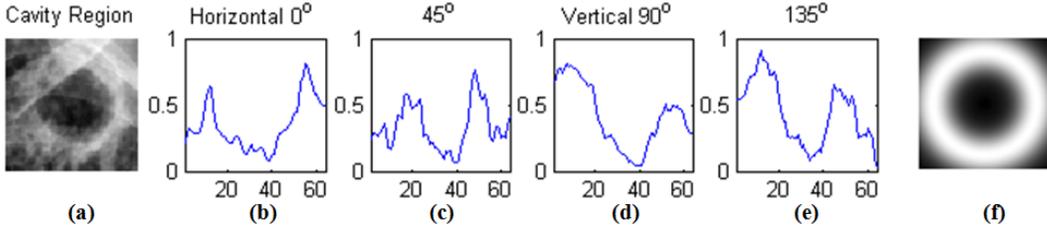
Fig. 4.3. Line-cut intensity profile analysis of 'hole' like cavity region. (a) a cavity region; (b)-(d) Line-cut intensity profile in four directions; (f) customized template for mimicking the cavity pattern.

A generic 2D Gaussian ring is defined as follows:

$$I(x,y) = e^{-\frac{(1-w)^2(x^2+y^2)}{2\xi^2}}$$ (4.1)

where $w = \dfrac{ab}{\sqrt{a^2y^2 + b^2x^2}}$, a and b are the two radii (distance between the origin and the peaks on x and y axes), $I(x,y)$ is the image intensity function in the 2D domain, and ξ is the standard deviation of the Gaussian distribution which determines the wall thickness of the ring. Noting that when $a = b = r$, Eq. (4.1) represents a 2D circular Gaussian ring, where

r is the inner radius. Rotated patterns can be generated by incorporating a rotation angle θ into the following coordinate transformation:

$$\begin{cases} x = x'\cos\theta + y'\sin\theta \\ y = y'\cos\theta - x'\sin\theta \end{cases} \tag{4.2}$$

where *x'*, *y'* are the pixel's location before rotation. Using Eqs. (4.1) and (4.2), the template database can be built with various sizes, wall thicknesses and rotation angles by changing the value of parameters *a*, *b*, ξ and θ. For example, given a 512×512 CXR with a pixel spacing [0.8 mm, 0.8 mm], the physical size represented by the image is 40.96 cm×40.96 cm. Since the diameter of the largest cavity is usually less than 6 cm, we define the template size as 75×75. While the wall thickness is within the range of [4mm, 16mm], parameter ξ is varied from 5 to 20 pixels. Fig. 4.4 shows a set of templates, with various radii, rotation angle and wall thickness.
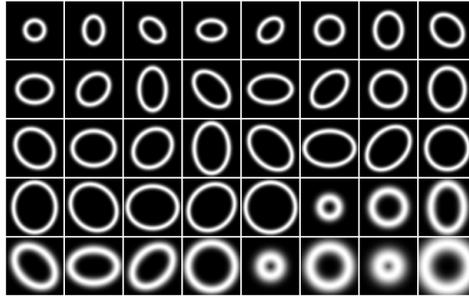
Fig. 4.4. An example of cavity templates. a/b<1.6, wall thickness within [5,20], and θ **= 0°,45°,90°,135°.**

4.2.2 LBP and HOG-based Feature Classification (LHFC)

Although the proposed GTM module works well for cavities of typical shape and intensity, it is difficult to detect cavities obscured by anatomical structures or some other abnormalities in the lung field. To address this issue, we combine LBP and HOG features, which have been shown to be useful in human detection in handling partial occlusion [86]. As explained in Chap. 2, the LBP [47] is a hybrid texture feature widely used in image processing. It combines the traditionally divergent statistical and structural models of

texture analysis. The LBP feature has some key advantages, such as its invariance to monotonic gray level changes and computational efficiency. The HOG feature [48] is regarded as an excellent descriptor to capture the edge or local shape information. It has a great advantage of being robust to changes in illumination or shadowing. These two features are expected to complement well the GTM technique, especially in blurred regions containing cavities, to detect TB cavity candidates,

In the LHFC module, a feature vector, which combines the LBP and HOG features, is calculated for each candidate window. The feature vector is then fed to a classifier, which is trained offline using ground truth (cavity and non-cavity) training data. The classifier will assess the windows as cavity candidates (positive samples) or not (negative samples). The candidate windows are generated using a sliding-window paradigm where an image is scanned from the top left to the bottom right with overlapping rectangular sliding windows. The windows are scanned row-wise. The window size is consistent with the template size in GTM, i.e., each window has a size of 75x75. The overlap between two consecutive windows is 2/3 of the window size. The computation of the combined LBP and HOG feature vector and the classification using SVM are now explained below.

Step 1 - Combined LBP and HOG based Feature Extraction

Given an image window, the LBP values are first calculated as explained in Section 2.2.1.3, six statistical features (mean, standard deviation, smoothness, skewness, uniformity and entropy) based on the LBP histogram are then extracted. Fig. 4.5(b) shows the 6 LBP features calculated from the image window shown in Fig. 4.5(a). To compute the HOG-based feature vector, for computational convenience, we first resize each 75×75 image window into a 64×64 window using bicubic interpolation, the feature vector is then extracted using the 3 steps explained in Section 2.2.1.4. With the cell size of 8×8 pixels, and the block size of 2×2 cells, the final HOG-based feature vector of a window (with 49 blocks) is a concatenated vector of all 49 normalized block orientation histogram, and will have a dimension of 1×1764. The final combined LBP and HOG features, a feature vector

of size 1×1770 is obtained for each image window. Fig. 4.5(c) shows the plot of the HOG feature vector of the image window shown in Fig. 4.5(a).

Average	33.33585
Std	19.10606
Smoothness	0.06248
Skewness	-1.40448
Uniformity	0.035923
Entropy	5.326193

(a)   **(b)**

(c)
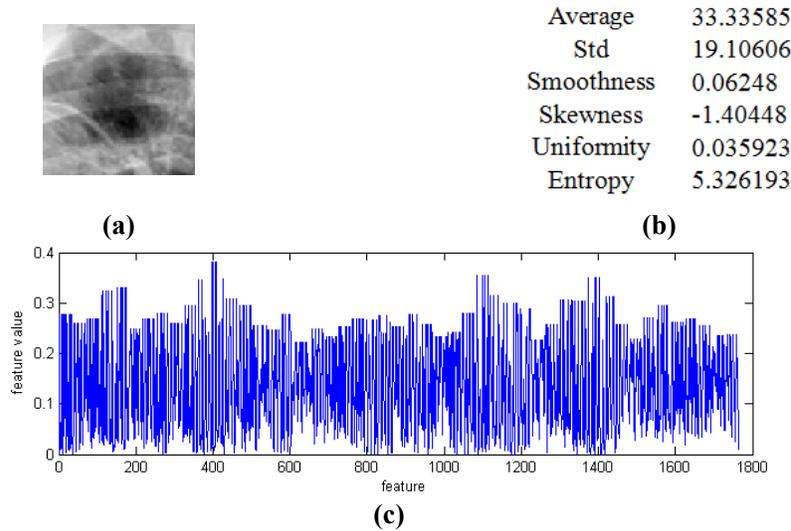
Fig. 4.5. An example of the LBP and HOG features. (a) An image window containing a cavity; (b) Six LBP features corresponding to (a); (c) The HOG feature vector (1x1764) corresponding to (a).

Step 2 – Feature Classification using SVM

These features vectors are fed to the SVM classifier which we explained in Section 2.2.2.1. To train the SVM classifiers, we apply the 'leave-one-out' method since the size of samples with cavities is small [87]. The training set contains the combined LBP and HOG feature vectors extracted from windows with and without cavities (positive and negative samples) in CXRs. Note that the negative samples for training were selected from the contralateral position of the positive samples based on the approximate symmetry of the lung field. We use the radial basis function (RBF) kernel function in SVM which performs better than other kernels such as linear, polynomial, and sigmoid kernels in our tasks. Based on the trained SVM model, candidate windows are detected by SVM as the positive ones.

An example of cavity candidate detection using GTM+LHFC is shown in Fig. 4.6. Fig. 4.6(a) shows the original CXR, and Fig. 4.6(b) shows three detected TB cavity candidates C1, C2, C3. The magnified images of these candidates are also shown in Fig. 4.6(c). To eliminate the false positive candidates (C1 and C3), further contour segmentation and fine

feature classification are necessary.
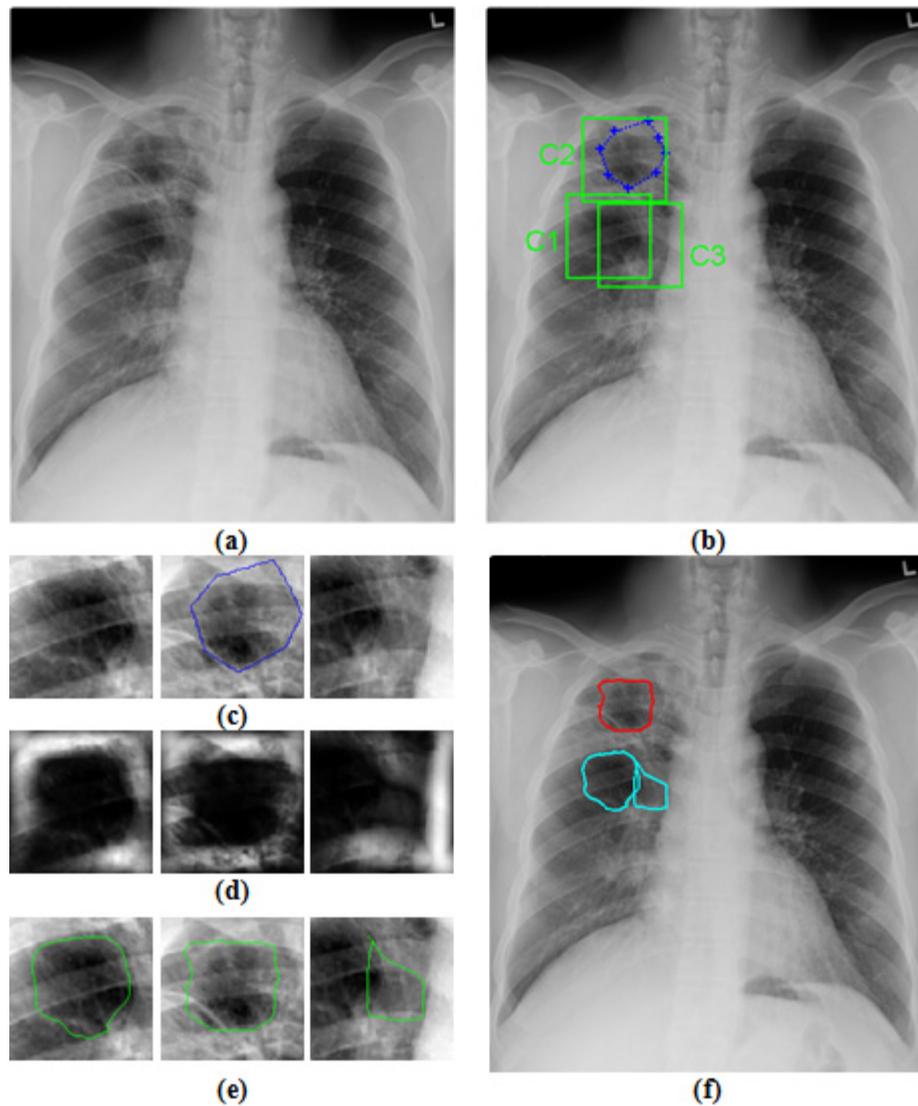
Fig. 4.6. An example of cavity candidates detection using the proposed technique. (a) Original CXR; (b) Candidate detection results in ULZ obtained using GTM+LHFC where the green rectangular windows (C1, C2, C3) represent the candidates, and the blue dotted contour is the true cavity annotated by radiologists; (c) Magnified candidate windows: C1-C3 (left to right); (d) HIE results of C1-C3; (e) IFVF results of C1-C3 with the help of HIE; (f) Final cavity detection results using fine feature classification. Red contour is the detected cavity, while the cyan ones are classified as non-cavity contours.

4.2.3 Hessian-matrix-based Image Enhancement (HIE)

As shown in Fig. 4.6 (b), the GTM+LHFC detects a large number of cavity candidates some of which may be false positives (e.g., C1 and C3 shown in Fig. 4.6(b)). In this section, we present a technique to enhance the cavity feature in a candidate, which will help in reducing the number of false positives. In order to reduce the effect of noise and irrelevant anatomical structures or abnormalities, we apply the HIE to enhance the candidates. Note that the Hessian matrix has been applied in the literature to enhance local patterns such as plate-like, line-like or blob-like structures [88]. The proposed HIE has three steps, which are described in the following:

Step 1 - Laplacian of Gaussian Smoothed Image

In this step, three Laplacians (in three directions) of a Gaussian smoothed image, at scale σ, are obtained by convolving a cavity candidate with the 2nd derivative of Gaussians as follows:

$$
\begin{cases}
L_{xx}(x,y,\sigma) = \sigma^2 I(x,y) * G_{xx}(x,y,\sigma) \\
L_{xy}(x,y,\sigma) = \sigma^2 I(x,y) * G_{xy}(x,y,\sigma) \\
L_{yy}(x,y,\sigma) = \sigma^2 I(x,y) * G_{yy}(x,y,\sigma)
\end{cases}
\tag{4.3}
$$

where $I(x,y)$ is the candidate and G is the Gaussian kernel. Note that for a candidate of size 75×75, each of the three L matrices in Eq. (4.3) will have a size of 75×75. Fig. 4.7 shows the 2nd derivative of a 1D Gaussian kernel. The intrinsic characteristic of this analysis is that the 2nd derivative of the Gaussian kernel at scale σ generates a probe kernel that measures the contrast between the regions inside and outside the range $(-\sigma, \sigma)$ in the direction of the derivative.
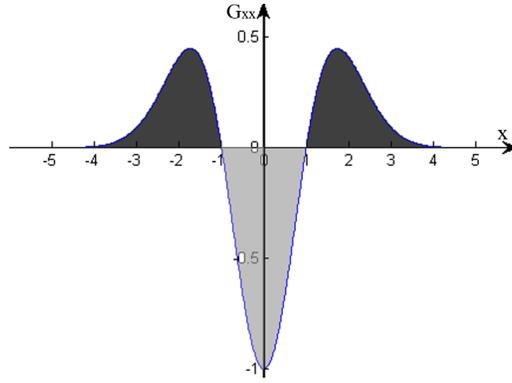
Fig. 4.7. The 2nd derivative of a 1D Gaussian kernel probes inside/outside contrast of the range ($-\sigma$, σ). In this example, $G_{xx}(x) = (\dfrac{x^2}{\sigma^4} - \dfrac{1}{\sigma^2})e^{-\frac{x^2}{2\sigma^2}}$ **,** σ **=1.**

Step 2 - Hessian Matrix Calculation

For a given σ value, the Hessian matrix corresponding to pixel (x_i, y_i) in the candidate is calculated as follows:

$$H_\sigma(x_i, y_i) = \begin{bmatrix} L_{xx}(x_i, y_i, \sigma) & L_{xy}(x_i, y_i, \sigma) \\ L_{yx}(x_i, y_i, \sigma) & L_{yy}(x_i, y_i, \sigma) \end{bmatrix}. \tag{4.4}$$

where $L_{xy}(x_i, y_i, \sigma) = L_{yx}(x_i, y_i, \sigma)$. A known problem of multi-scale analysis using Hessian matrix is that over-blurring can occur during the multi-scale smoothing, which may increase false detections [89]. Therefore, we set the σ value equal to the object scale calculated using the method described in Ref. [90]. The object scale at every pixel is defined as the radius of the largest hyperball centered at the pixel such that all pixels within the ball satisfied a predefined image intensity homogeneity criterion. Object scale represents the geometric information (size) of the local structure. Object scale at the center of a blob-like structure is approximately equal to the radius of the blob in pixel size.

Step 3 - Image Enhancement Using Eigenvalues of Hessian Matrix

The pixel (x_i, y_i) in the candidate with intensity $I(x_i, y_i)$, is enhanced using the following equation:

$$I_E(x_i, y_i) = |\lambda_1| I(x_i, y_i) \qquad\qquad (4.5)$$

where λ_1 and λ_2 are eigenvalues of $H_\sigma(x_i, y_i)$, and $|\lambda_1| \geq |\lambda_2|$. The intuition in Eq. (4.5) of using only the largest eigenvalue for cavity enhancement is based on the fact that the Hessian matrix has a strong edge effect (for those strong edge points, $|\lambda_1| >> |\lambda_2| \approx 0$) [91]. Although cavities are usually embedded in noisy surroundings due to the neighboring necrosis caused by cavitation, the inside of a cavity (filled with air or fluid or both) still has lower intensity than the background. Thus the strong edge between the inside and outside of a cavity gives a good clue to indentify the contour of cavity. Different techniques of edge enhancement were evaluated in our study, such as contrast-limited adaptive histogram equalization [92], fuzzy C means [93] and speckle reducing anisotropic diffusion [74] technique, and the proposed HIE technique achieves the best performance.

The enhanced window candidates C1-C3 are shown in Fig. 4.6(d). It is observed that the annular ring-like structure is greatly enhanced.

4.2.4 Active Contour-based Segmentation (ACS)

In this module, we propose a snake-based technique known as improved fluid vector flow (IFVF) [22] to segment the cavity. As explained in Section 2.1.2.1, a snake contour represented by v evolves through the candidate window to reach a force balance equation $F_{int}(v) + F_{ext}(v) = 0$, where $F_{int}(v)$ is the internal force constraining contour's smoothness, and $F_{ext}(v)$ is the external force attracting the contour toward image features. The development of snakes mainly focus on design of new static and/or dynamic terms in the external force, where $F_{ext}(v) = F_{static}(v) + F_{dynamic}(v)$.

From the discussion about FVF in Section 2.1.2.2, we improve the two limitations of FVF as follows:

First, the F_{static} in FVF is replaced by a static external force which overcomes the edge

leakage problem, such as GVF [39] or BVF [40]. In our task, we use BVF as the F_{static} due to its larger capture range and higher computational efficiency than GVF.

Next, the new proposed F_{dynamic} is achieved in three steps:

Step 1 - Given a HIE enhanced candidate image, a binary edge map B is generated using a Canny edge detector [75].

Step 2 - By comparing the edge map points to the current snake contour points (snaxels), a new control point (x_c, y_c) is selected by considering the point which contributes more to the distance between snake contour and object boundary [22]. We use the Hausdorff distance to find such a point. Assuming two sets of points S and O, the Hausdorff distance is then defined as $h(S, O) = \max_{o \in O} \left\{ \min_{s \in S} \left\{ d(s, o) \right\} \right\}$ where $d(s,o)$ is the Euclidean distance between a snaxel s and a object boundary point o. So the control point is chosen as the point on the object boundary which has the Hausdorff distance value. For example, in Fig. 4.8, point o3 will be selected as the control point for the current snake evolution.
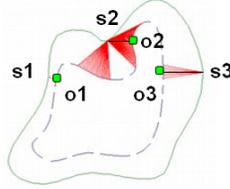
Fig. 4.8. An example of control point selection. s1-s3, o1-o3 are points on the snake contour and object boundary, respectively.

Step 3 - For any snaxel (x,y) on the contour v, its $F_{\text{dynamic}}(x,y)$ is then calculated as follows:

$$F_{\text{dynamic}}(x, y) = (1 - B)\delta \frac{\nabla d'(x, y)}{\|\nabla d'(x, y)\|} \tag{4.6}$$

where $\delta = \pm 1$ controls the outward or inward direction, and $d'(x,y)$ is the Euclidean distance between points (x,y) and (x_c, y_c). Note that the term $(1 - B)$ makes the F_{dynamic} zero for those snaxels which already reach edges. As a special case, when the snaxel (x,y) reaches the control point ($\|\nabla d'(x, y)\| = 0$), the F_{dynamic} is zero since the control point is

selected from the edge map ((1-*B*)=0). Based on the edge map generated from the enhanced

candidates images using HIE, the IFVF segmentation result of these candidates C1-C3 are

shown in Fig. 4.6(e). The stopping criterion of the evolution is determined by computing the

difference in locations of the corresponding snaxels between two consecutive iterations. If it

is less than a convergence threshold *t*, the active contour evolution will be stopped.

4.2.5 Contour-based Features Classification (CFC)

The last module in our proposed technique is the CFC, which performs the fine scale feature

classification. Three types of contour-based features: shape, edge and region, are extracted

for the final cavity detection. These features include circularity measure [94], GICOV [95],

and KLD [96] between the pixel intensity distributions inside and outside the contour. A

SVM classifier is then applied using these contour-based features to identify cavity contours

in a similar way as in LHFC. The computation of three features is explained below:

(1) *Circularity*: Assuming a contour has one centroid, *L* points are selected from the

contour in *L* cardinal directions. The circularity of the contour is then calculated as scaled

variance as follows:

$$Circularity = \frac{\text{var}(d(x_i, y_i))}{\text{max}(d(x_i, y_i))}, \quad i = 1, 2, ..., L \quad (4.7)$$

where $d(x_i, y_i)$ is the distance from the centroid to the contour point (x_i, y_i) in the ith direction.

In this study, we use $L = 16$. The circularity feature is a feature which could effectively

reduce the false positives.

(2) GICOV: Based on the observation that the inner boundary of a cavity often has

dark-to-bright transition, GICOV value of *L* points on the contour is calculated as follows:

a) For the contour point (x_i, y_i) in the ith direction, its gradient in normal direction

$g_n(x_i, y_i)$ is calculated as $g_n(x_i, y_i) = \nabla I(x_i, y_i) \cdot \vec{n}(x_i, y_i)$, where $\vec{n}(x_i, y_i)$ is the

unit outward normal vector at this point.

b) The mean and standard deviation of g_n, denoted by m and s, are then calculated as

$$m = \frac{1}{L}\sum_{i=1}^{L} g_n(x_i, y_i) \quad \text{and} \quad s^2 = \frac{1}{L-1}\sum_{i=1}^{L}(g_n(x_i, y_i) - m)^2 \ .$$

c) The GICOV value of the contour is finally achieved using following equation:

$$GICOV = \frac{m}{s/\sqrt{L}} \tag{4.8}$$

(3) *KLD*: Given the probability distributions, P and Q, of the pixel intensity values inside and outside the cavity in the candidate window respectively, the KLD is calculated as follows:

$$KLD = \sum_{i=1}^{B} P(i) \ln \frac{P(i)}{Q(i)} \tag{4.9}$$

where B is the number of bins in the histogram span by P and Q. The KLD compares the difference in gray level distribution between the pixels inside and outside the contour.

Table 4.1 shows the above feature values corresponding to three contours shown in Fig. 4.6(e). As in the coarse feature classification step, we select the SVM as the fine feature classifier in this step. Based on the feature values (in Table 4.1), the trained SVM classifier identifies the Contour-2 as a positive and Contour-1 and Contour-3 as negatives. The final detected cavity (corresponding to Contour-2) in the CXR is shown in Fig. 4.6(f) as the red contour. The result matches with the ground truth.

Table. 4.1. Fine feature values of three contours in Fig. 4.6(e)

	Circularity	GICOV	KLD
Contour-1	0.11	15.33	1.49
Contour-2	**0.15**	**13.68**	**2.28**
Contour-3	0.69	15.26	0.28

4.3 Performance Evaluation

In this section, we evaluate our proposed CFDS technique with respect to three aspects: the

effectiveness of candidate selection; the accuracy of contour segmentation; and the accuracy of final cavity detection.

4.3.1 Database Used and Parameters Configuration

A cavity dataset of 35 CXRs containing 50 cavities is obtained from the UAHCXR. All the CXRs were independently read by three experienced chest radiologists who are specialized in TB diagnosis. The presence of TB cavities was confirmed by the agreement of at least two radiologists. The sample histograms of cavity properties such as diameter, circularity and wall thickness are shown in Fig. 4.9. From the histograms, it can be seen that the cavities vary in diameters while their circularities range mainly from 0.15 to 0.2 and most of them have intermediate thickness. For computational efficiency, the original CXRs are resized as 512×512 (or close to this size) with a fixed pixel spacing [0.8 mm, 0.8 mm]. The proposed cavity detection technique is implemented in Matlab 2007b on an Intel Pentium 4 CPU 2.8G Hz with 2G RAM computer. All the parameters in the proposed technique are listed in Table 4.2. The SVM classifiers in both coarse and fine feature classification are built using LIBSVM software [97].

Table. 4.2. Parameters configuration in the proposed technique

Modules	Parameters Names	Parameters Values
GTM	Template size	75×75 pixels
	Wall thickness σ	[5, 20]
	Aspect ratio a/b	[1, 1.6]
	Rotation angle θ	{0°,45°,90°,135°}
LHFC	Window size	75×75 pixels
	Cell size	8×8 pixels
	Block size	2×2 cells
	Block overlap	2/3
	SVM parameters	Default values in LIBSVM software
ACS	Snake evolution direction δ	1
	Convergence threshold t	0.05
CFC	SVM parameters	Default values in LIBSVM software
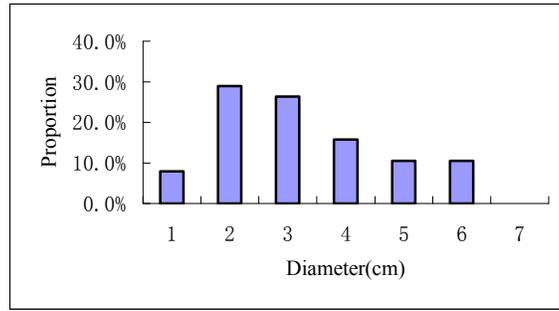
(a)
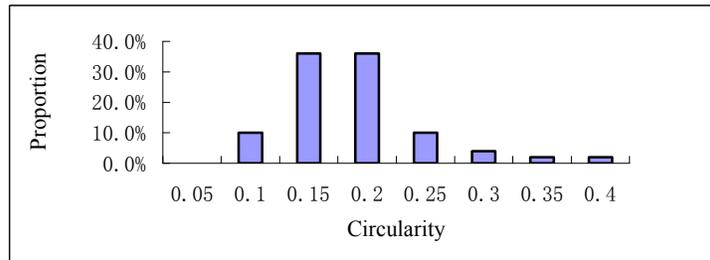
(b)
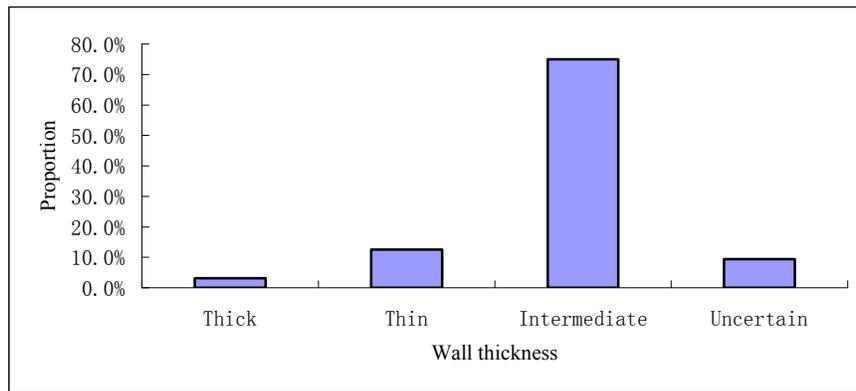
(c)

Fig. 4.9. Sample histograms of cavity properties. (a) histogram of diameter; (b) histogram of circularity; (c) histogram of wall thickness of four categories: "Thick" (≥16mm), "Intermediate"(4-15 mm), "Thin"(< 4mm), and "Uncertain" (wall not discernible).

4.3.2 Experiments and Analysis

Experiment 1 - Effectiveness of Candidate Selection

The proposed coarse feature classification technique for candidate detection is evaluated by

the missing rate (MR), which is calculated as follows:

$$MR = \frac{\text{\# of Cavities Excluded from Candidates}}{\text{Total \# of True Cavities}} \times 100\% \qquad (4.10)$$

A preliminary experiment using only GTM for candidate detection has already been reported in [23]. We anticipate that by integrating with other novel techniques, a better result can be obtained. Thus we used different combinations of LBP and/or HOG features together with GTM, and checked whether the missing rate could be reduced. Table 4.3 shows our test results.

Table. 4.3. Candidates detection results

	HKG [17]	GTM [23]	GTM+LBP	GTM+HOG	GTM+LBP+HOG
# of Cavities	50	50	50	50	50
# of Candidates	170	124	315	229	160
# of Missing Cavities	22	18	10	17	**7**
MR	44%	36%	20%	34%	**14%**

From the results, we observe that the HKG framework for TB cavity detection [17] missed more cavities than our proposed approach. HKG is based on an adaptive thresholding on the mean-shifted clustered image for candidate detection. Its high missing rate is due to two reasons. First, the mean-shift clustering approximates nearest neighbors intensities and space information but neglects the texture. Second, the adaptive threshold, which is a quadratic polynomial of the GICOV feature, is not suitable for modeling all shapes, especially when the boundary of a cavity is weak. Fig. 4.10 compares the candidates detection results of HKG and our technique. The green regions and boxes represent cavity candidates. In Fig. 4.10(a), HKG cannot identify both cavities due to the failure of mean-shift clustering in the noisy ULZ, while our technique is able to cover the two cavities (Fig. 4.10(b)). Fig. 4.10(c) is yet another example showing the adaptive threshold value used in HKG unable to identify the cavity. However, our technique is able to detect the cavity correctly (Fig. 4.10(d)).

(a) **(b)**

(c) **(d)**
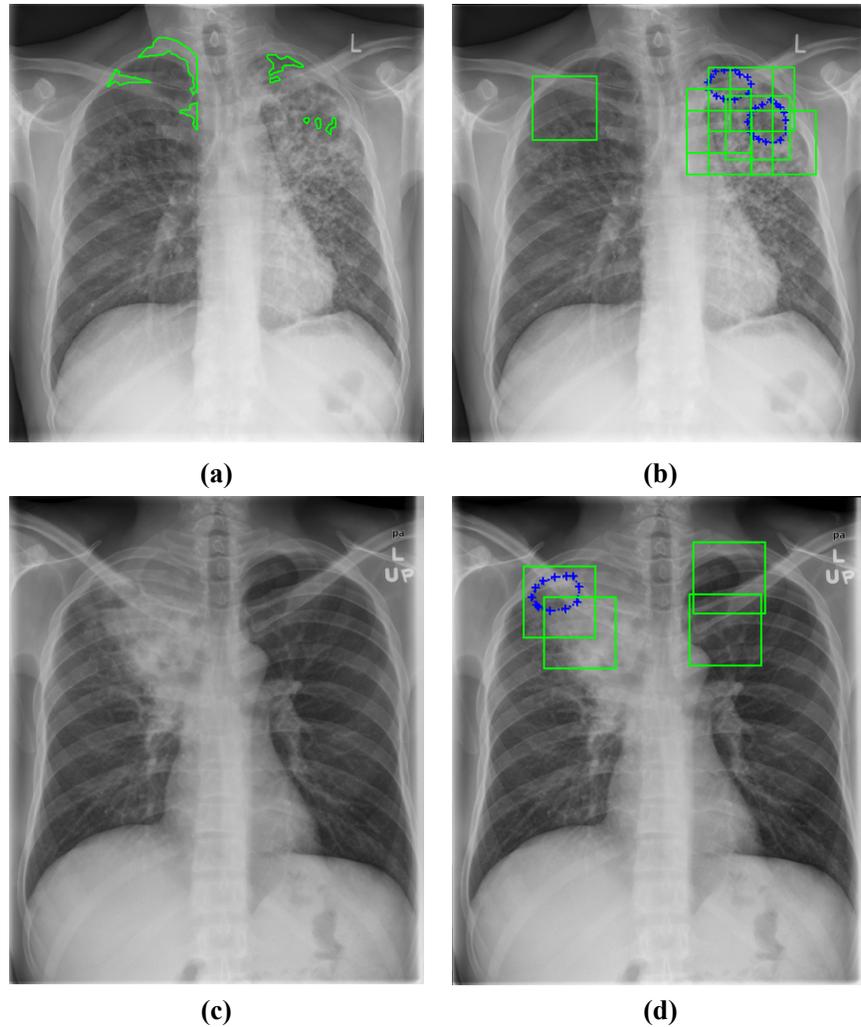
Fig. 4.10. Comparison of candidates detection between HKG [17] and the proposed technique. (a), (c) are the results of HKG, while (b), (d) are generated from the proposed technique. Green regions in the images are cavity candidates regions reported by different techniques and blue dotted contours are the true cavities annotated by radiologists.

Using the same parameter values for LBP and HOG as in the literature, we found that a combination of LBP and HOG together with GTM achieved better performance. Our finding is consistent with the results in human detection using LBP and HOG features [86]. HOG performs poorly when the background is cluttered with noises. LBP is able to alleviate this deficiency. It can filter out noises following the uniform pattern estimation. However, if LBP is used alone without HOG, the entire ULZ will be extracted if some other abnormalities are also present in the area. In that case, the HOG helps to reduce the false

positives based on the available edge information. Fig. 4.11 illustrates the complementary effect of LBP and HOG. The window reported by the classifier should contain a complete cavity in order to be qualified as a positive candidate. Note that in the 1st row 2nd column when using only HOG, no reported window contains a complete cavity. The HOG performs poorly when the background is cluttered with noises, and the edge information is no longer reliable. Similarly, in the 2nd row 1st column, when using only the LBP, the small cavity is missing because no reported window contains the complete small cavity, and only the larger cavity is fully contained in a reported window.

GTM+LBP	GTM+HOG	GTM+LBP+HOG

(a)   (b)   (c)

(d)   (e)   (f)
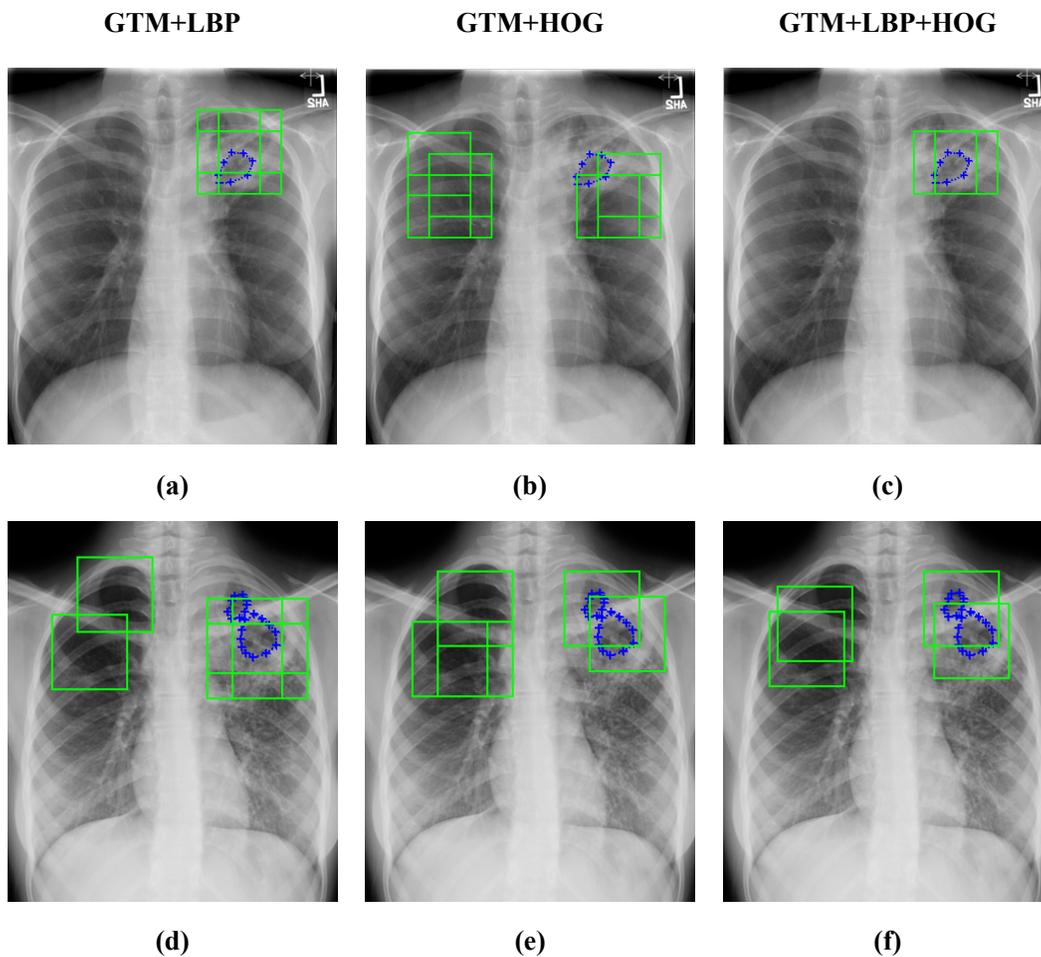
Fig. 4.11. Comparison of candidate detection in the coarse feature classification step using (a,d) GTM+LBP, (b,e) GTM+HOG, (c,f) GTM+LBP+HOG. Note that in the first row HOG misses the cavity but LBP is able to detect it. In the second row, LBP misses the small cavity but HOG can detect it. In both rows our technique is able to detect all the cavities.

The above test results show that combining the LBP and HOG features for capturing the texture and gradient information around the cavity region, and using the GTM for shape recognition, contributes to the low missing rate of the proposed coarse feature classification technique.

Experiment 2 – Accuracy of Contour Segmentation

To make the comparison consistent, we evaluate the segmentation accuracy using the same Tanimoto measure (TMM) in [17]:

$$TMM = \frac{\left\| R_c \cap R_g \right\|}{\left\| R_c \cup R_g \right\|}$$
(4.11)

where R_c denotes the region enclosed by the contour generated by the segmentation techniques, such as DBC-GVF [17] and our IFVF [22]; R_g denotes the region of a TB cavity that is enclosed by the ground truth contour manually drawn by radiologists; and $\left\| \cdot \right\|$ denotes the cardinality (# of pixels). *TMM* = 0 indicates that the segmented contour has no intersection with the ground truth, while *TMM* = 1 indicates that the segmented contour is identical to the exact cavity. To improve the segmentation accuracy, we apply the HIE on the candidates before segmentation.

The performance of the DBC-GVF and the IFVF techniques with and without the HIE is shown in Table 4.4. Note that around 10% accuracy improvement is achieved for both DBC-GVF and IFVF when HIE is incorporated. The results are also more robust as demonstrated by the lower standard deviations of the TMM. Fig. 4.12 presents subjective comparison of different segmentation techniques. With the HIE, the segmented contours are closer to the ground truth compared to the same techniques without the HIE.

Table. 4.4. Segmentation accuracy evaluation

	DBC-GVF Without HIE	DBC-GVF With HIE	IFVF Without HIE	IFVF With HIE
Average of TMM	55.1%	64.6%	56.8%	**67.1%**
Std. of TMM	15.8%	12.6%	12.2%	**9.3%**
Mean of TMM	58.2%	64.9%	59.3%	**66.1%**

Ground Truth	HIE Result	DBF-GVF without HIE	DBF-GVF with HIE	IFVF without HIE	IFVF with HIE
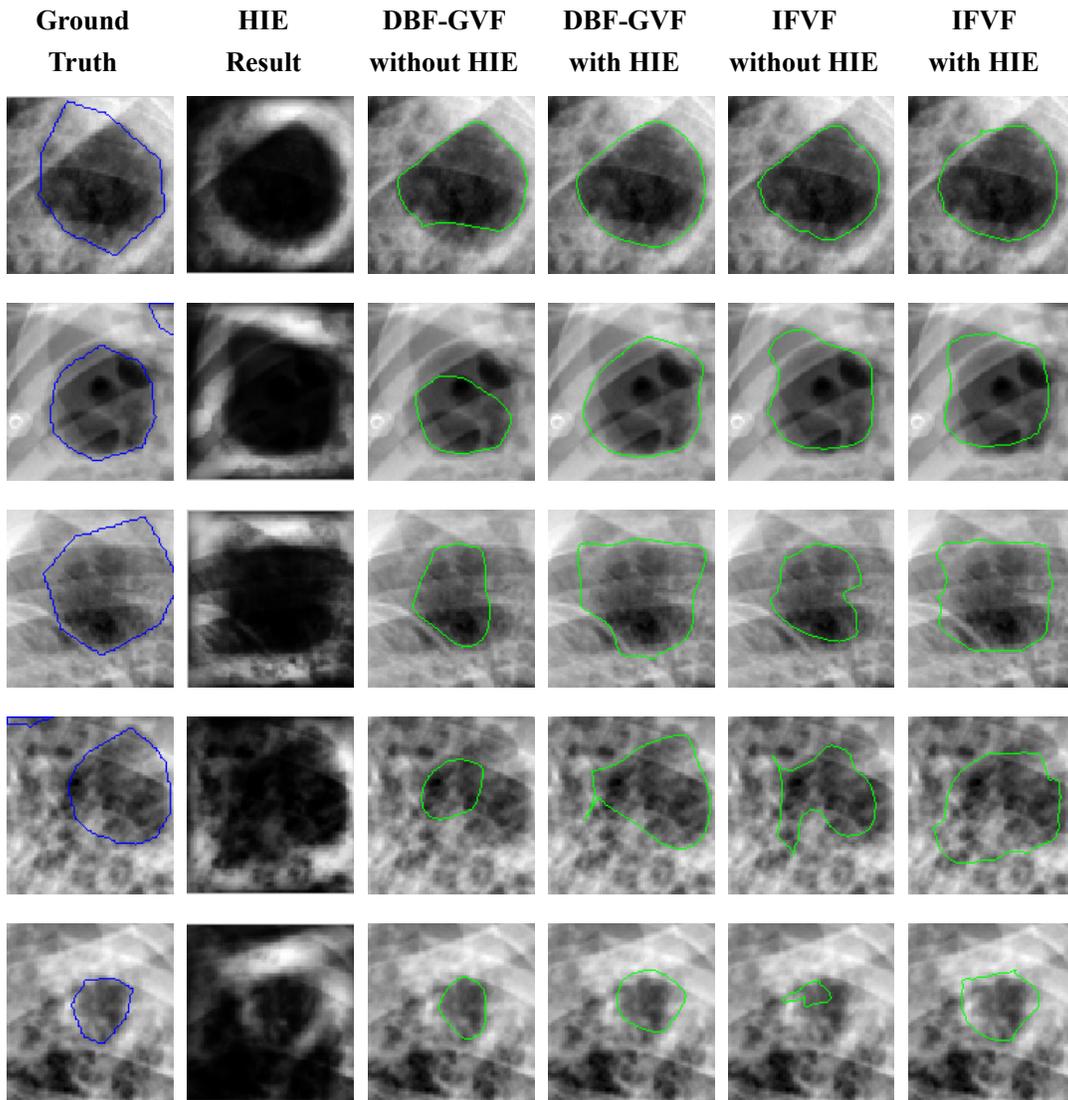
Fig. 4.12. Cavity segmentation result comparison using different edge-based snakes with and without HIE. From top to bottom, the cavity is more and more difficult to identify. Blue contours are the true cavities annotated by radiologists. Green contours are the computer segmentation results.

Note that image patterns, even without cavities, may generate close to ring-like shape after the HIE step. Fig. 4.13 shows some of these cases. For example, the image in the bottom row contains a pattern similar to a cavity. To eliminate this type of candidates, the fine scale feature classification step in our approach is necessary. The accuracy of our final cavity detection is evaluated in the next section.

Candidates without Cavity	HIE Result	IFVF with HIE
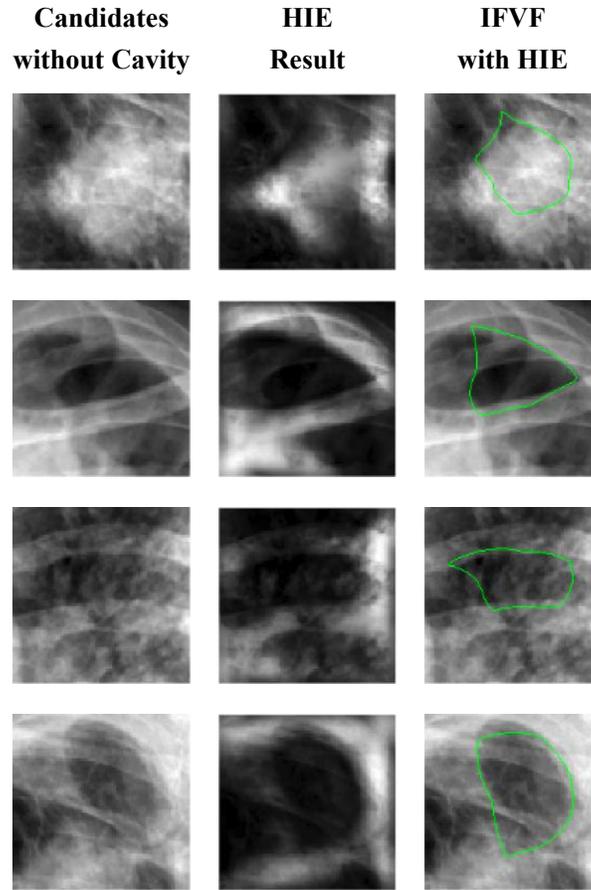
Fig. 4.13. Segmentation results of candidates without cavity.

Experiment 3 – Accuracy of Final Cavity Detection

Before performing the final cavity detection, 160 candidate contours are divided into cavity and non-cavity contours. Candidate region reported after the LHFC (as highlighted by the green windows in Fig.4.11) may not contain true cavities. Also, even if a reported window contains the entire cavity, its segmented contour may not be the same as the ground truth. To evaluate the accuracy of the final contour classification, we need to impose a value TMM>0.7 (based on the segmentation accuracy of 67.1% reported in Table 4.4), in order to qualify a candidate to be a true cavity; otherwise it is considered as non-cavity. Three contour-based features (Circularity, GICOV and KLD) are extracted from the candidate contours for the final cavity classification. To evaluate the performance of classification, sensitivity, specificity and accuracy are calculated as follows:

$$\text{Accuracy} = \frac{\text{\# of Correctly-Detected Contours}}{\text{Total \# of Candidates Contours}} \times 100\%$$

$$\text{Sensitivity} = \frac{\text{\# of Correctly-Detected Cavity Contours}}{\text{Total \# of Cavity Contours}} \times 100\% \qquad (4.12)$$

$$\text{Specificity} = \frac{\text{\# of Correctly-Detected Non-Cavity Contours}}{\text{Total \# of Non-Cavity Contours}} \times 100\%$$

The classification result using SVM for the 160 candidate contours is shown in Table 4.5. It can be observed that the detection accuracy is increased by more than 8% in our approach after adding KLD feature. Fig. 4.14 shows cavity detection results of HKG [17] and the proposed technique, which demonstrate that our technique can detect more true cavities and detect fewer false cavities. As illustrated in Fig. 4.14, the proposed cavity detection system identifies all cavities annotated by the radiologists and there is only one false alarm. The presence of cavities in the upper half of the lungs, especially when there are multiple or bilateral cavities, should raise suspicion of TB in the appropriate epidemiologic and/or clinical context. Unfortunately, in practice, a lot of these findings are not mentioned in the radiologist's report, because the epidemiologic or clinical information, necessary to raise suspicion, is not provided by the ordering physician on the requisition. This is often the case in geographic regions where TB rate is low. Based on the clinician's perspective, a relatively higher false positive rate is better than false negatives because the latter can cause an infectious TB to spread. Even with false positives, clinicians find automatic cavity detection system helpful in reducing a large number of true negatives and radiograph examinations. This is beneficial given the limited radiologists available particularly in remote communities and developing countries.

Table. 4.5. Cavity detection evaluation

	Sensitivity	Specificity	Accuracy
Circularity+GICOV [17]	62%	46%	54%
Circularity+GICOV+KLD	**70%**	**60%**	**65%**
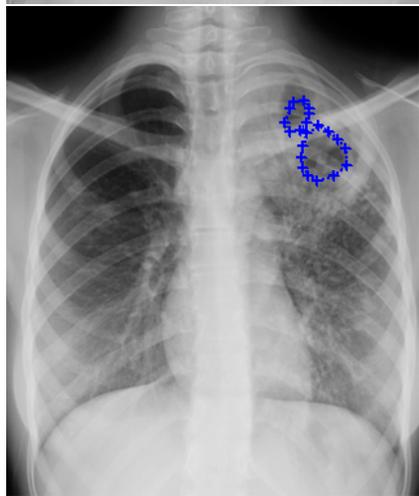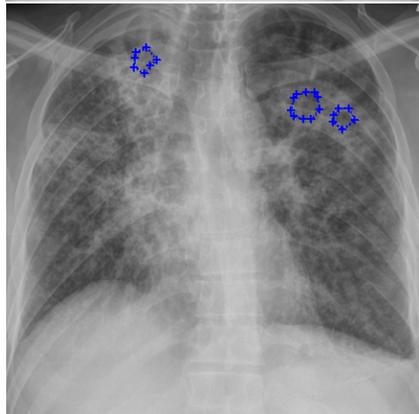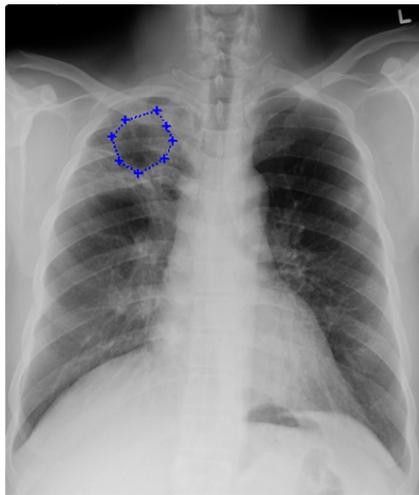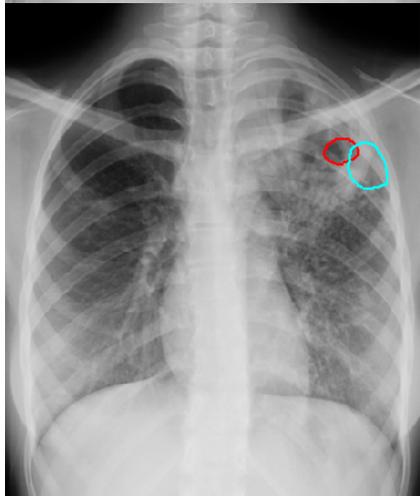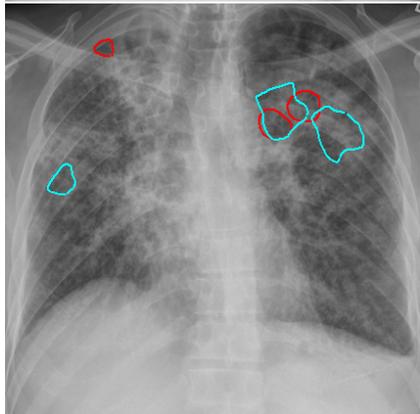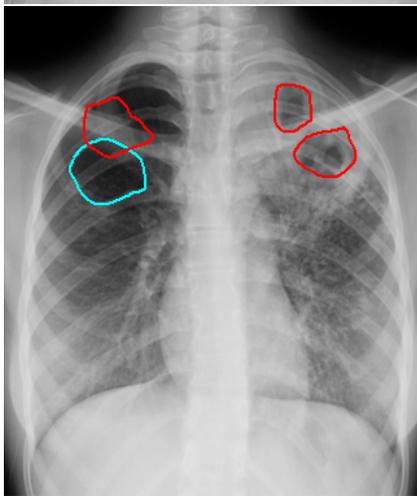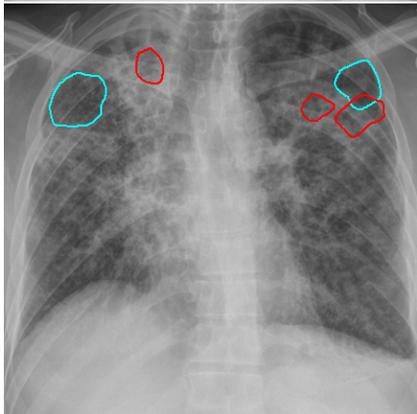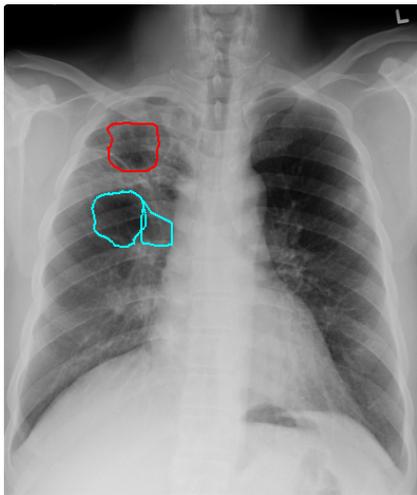
Fig. 4.14. Cavity detection comparison between HKG [17] and the proposed technique. Blue dotted contours are the true cavities annotated by radiologists. Red contours are the detected cavities, while the cyan ones are the non-cavity contours.

The radiologists also classified the true cavity contours into two categories: E-Group and D-Group, containing cavities which are 'easy' or 'difficult' to identify, respectively. The D-Group contains cavities even radiologists found them difficult to identify without other demographic or additional information. False cavity contours were then combined with each of these two groups. The SVM classification results of these groups are shown in Table 4.6 and 4.7. Observe that on average the classification accuracy in each group is higher than the result reported in Table 4.5. The performance of the E-Group is significantly improved by adding the KLD feature. In the D-Group, although the intensity variation inside and outside a cavity changes only slightly making it very difficult to identify the contour even for radiologists, there is still improvement in the detection result. This shows that the classifier can perform better if trained using more specific knowledge.

Table. 4.6. Cavity detection evaluation of E-Group

	Sensitivity	Specificity	Accuracy
Circularity+GICOV [17]	65%	78.2%	71.6%
Circularity+GICOV+KLD	**78.8%**	**86.8%**	**82.8%**

Table. 4.7. Cavity detection evaluation of D-Group

	Sensitivity	Specificity	Accuracy
Circularity+GICOV [17]	57.6%	88%	72.8%
Circularity+GICOV+KLD	**69.4%**	**81.6%**	**75.5%**

4.4 Summary

In this chapter, we proposed an efficient CFDS feature classification technique for TB cavity detection in CXRs. Experimental results demonstrate that the proposed technique outperforms existing methods in three aspects. First, a lower missing rate is achieved because in the proposed method local cavity region-related coarse features, such as geometric, textural and gradient features, are taken into consideration. Second, edge-based

segmentation becomes more accurate by incorporating HIE to enhance the contours. Third, the final cavity detection accuracy is greatly increased by introducing the fine scale feature classification using three types of contour-related features, which includes shape, edge and region. Our work contributes in the development of CAD systems for infectious TB diagnosis, because of the higher detection rate and lower missing rate compared to other techniques.

Chapter 5

Efficient TB Acinar Shadows Detection

For typical TB (PPTB), the co-occurrence of other abnormalities, such as AS in the vicinity of a cavity is a very useful and complementary radiographic cue. Sometimes, multiple AS will raise suspicion of infectious TB to radiologists, even if cavitation is undetectable. Therefore, in this chapter, we propose an effective CAD system aimed for AS regions detection in CXRs. This system exploits textural and photometric features extraction techniques which include LBP, GLCM and HOG to analyze target regions in CXRs. Classification of AS using Adaboost is then deployed to verify the performance of a combination of these techniques. Comparative study in different image databases shows that the proposed CAD system delivers consistent high accuracy in detecting AS.

5.1 Introduction

Typical radiographic patterns of PPTB as mentioned in many radiology handbooks include cavities, volume loss, AS and so on [1] [83]. In the diagnosis of pulmonary TB on CXR, AS reflect the presence of endobronchial spread of disease, the spread of tuberculous 'caseous' material within the bronchial tree. Such spread is a universal feature of pulmonary TB [1]. The AS are either within the vicinity of the major focus of disease (for example a cavitated area), immediately dependent from it or occasionally in the contralateral lung – indicating position or posture-related drainage from the major focus. The presence of AS in the vicinity of an upper lung zone infiltrate, especially if cavitary, further adds to the probability that the infiltrate reflects the presence of infectious PPTB. Our previous work in Chap. 4 focused on the extraction of TB cavities on CXRs. Thus, in current work, we focus on the automatic detection of AS on CXRs, which is defined as "round or ovoid poorly defined pulmonary opacities approximately 5-8 mm in diameter, presumed to represent an anatomic

acinus rendered opaque by consolidation" [83]. An example of multiple AS could be found in Fig 5.1.
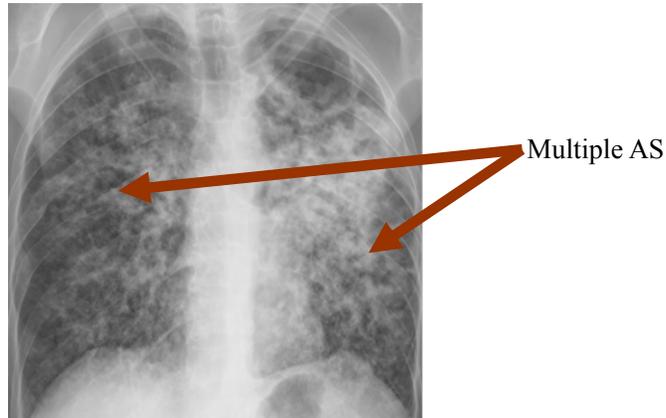
Fig. 5.1. An example of CXR with multiple acinar shadows.

So far, to the best of our knowledge, no automatic algorithm has been developed that can detect TB AS from CXRs accurately. As we mentioned in Section 1.1.3, there are several CAD schemes using texture features such as moments of power spectrum [13], fractal dimension [14], histogram based features [15], and 1-D wavelet coefficients [16] to identify interstitial changes. However, not only these CAD systems are not fully automatic, but also the interstitial opacity is not a reliable radiographic cue for infectious TB and is pathologically different with airspace opacity such as AS. More recently, Tan et al. [18] proposed a semi-automatic CAD system to detect airspace opacity of TB. They used a user-interactive ASM technique for the lung field segmentation and applied the first order statistical texture features (histogram based features) to the Adaboost classifier to identify TB from normal cases. Further more, Patil [19] investigated both first and second order statistical texture features from TB CXRs, and suggested the second order statistical texture features (GLCM-based features) are more suitable for TB opacity detection. Although these new CAD schemes achieve high accuracy, they are still not automatic and extracted only individual texture features without considering other high level features or combination of different types of features. Besides, the database they use only contains TB and normal cases, ignoring other lesions need to be distinguished from airspace opacity such as discrete

nodules. Therefore, in this chapter, we conduct a comprehensive study applying different feature descriptors and investigate their influence on classification using the popular SVM or Adaboost classifiers for AS detection. Based on the comparative experiments in different image datasets including typical TB with AS, atypical TB without AS, normal cases without AS, and nodule cases without AS, we finally propose a hybrid CAD system for AS detection using combined LBP, HOG and GLCM based features with Adaboost classifier. To further reduce false positive rate, we introduce an efficient morphological operation to achieve better performance.
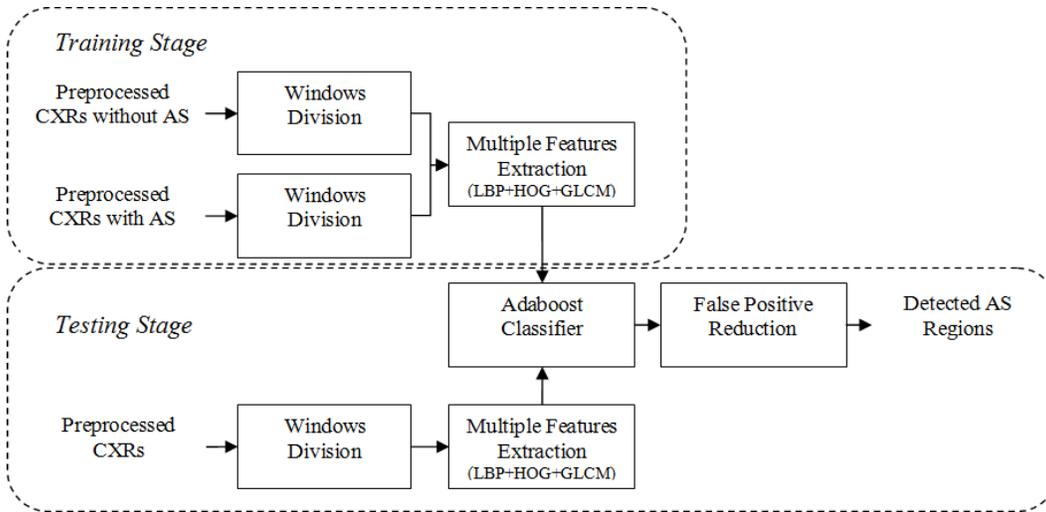
Fig. 5.2. Schematic of the proposed CAD system for AS detection.

5.2 Proposed Hybrid CAD System for AS Detection

The proposed computer-aided AS detection system follows the state-of-the-art sliding window paradigm in object detection, which is similar as applied in Section 4.2.2. Each CXR is first preprocessed as a contrast-enhanced subimage containing only lung fields. The preprocessed subimage is then divided into non-overlapping windows, whose size is fixed as 16×16 in our experiments. Multiple features are calculated from each window to form a multi-dimensional feature vector. Based on these feature vectors, a classifier is trained offline to distinguish windows containing instances of AS from other windows. The test

image scanned with sliding windows is then analyzed and classified using the model generated from the training set. We conducted comparative experiments using different types of features on different classifiers. Finally, combined LBP, HOG and GLCM-based features are extracted to the Adaboost classifier, followed by a morphological smoothing technique to reduce the false positive windows. Fig. 5.2 shows the schematic of the proposed AS detection technique.

5.2.1 Preprocessing

Before a CXR is sent to our proposed hybrid CAD system for AS regions detection, it needs to be preprocessed in two steps: First, the CXR will be cropped as a rectangular subimage which contains the border of lung fields using our ERF-ASM technique (See Chap. 3). Next, this subimage is locally contrast-enhanced using contrast-limited adaptive histogram equalization (CLAHE) technique [92]. Fig. 5.3 shows subimages from different image datasets and their improved qualities using the contrast enhancement step.

| Typical TB with AS | Atypical TB without AS | Normal Case | Nodule Case without AS |
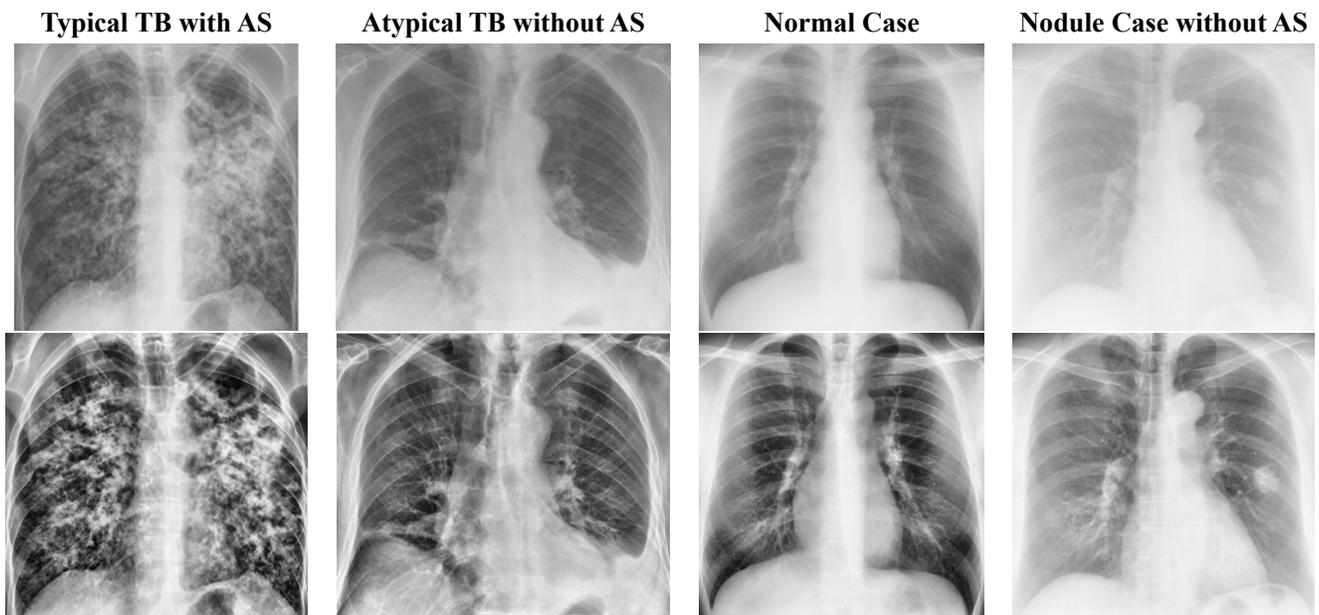
Fig. 5.3. Preprocessed CXRs from four different image databases. CXRs in the 1st row: cropped subimages with lung fields limitation; CXRs in the 2nd row: contrast-enhanced subimages.

5.2.2 Multiple-Feature Extraction

To recognize AS pattern, texture features will be the most important cue [98]. Although thousand of texture features have been developed, the selection of proper feature descriptors is often application dependent and largely based on the intrinsic characteristics of the images. Generally, texture features can be divided into four categories: statistical features, structural features, signal processing based features, and model based features [99]. Therefore, a variety of features from different categories are selected and compared in our study, including texture features applied in the existing CAD schemes in chest radiology such as histogram based statistical features (HS) [18], Fourier spectrum (FS) based features [13], GLCM based features [19], and fractal dimension (FD) [14]. We also test some advanced feature descriptors which receive more and more popularity in many pattern recognition tasks in medical images, including LBP based features [47], HOG based features [48], and Tchebichef moments based features(TM) [100]. Finally, based on the results of comparative experiments (See Section 5.3.2), we combined LBP, GLCM and HOG to achieve the best performance.

5.2.3 Feature Classification

Multiple features are evaluated on two state-of-the-art classifiers SVM and Adaboost. Since the details of these two classifiers have already been discussed in Section 2.2.2, no more will be repeated here. We finally apply the Adaboost classification technique into our hybrid CAD system due to its better performance than SVM (See Section 5.3.2).

5.2.4 False Positive Reduction

Notice that most of the false positive windows are discrete, thus, we introduce a simple smoothing technique based on morphological operation to reduce the false positives so that

radiologist diagnosis can be more effective. Considering a block of 3×3 windows, for each center window, if five or more of its 8-surrounding windows have different classification labels from the center window, the center window is then reassigned the same label as the majority of its neighbourhood. Otherwise, it keeps its label.

5.3 Performance Evaluation

In this section, we investigate multiple features including HS, FS, GLCM, FD, LBP, HOG, and TM on both SVM and Adaboost classifiers for the AS detection. Based on the comparative experiments, we combine LBP, HOG and GLCM-based features with the Adaboost classifier to detect AS regions. Performance with and without the proposed false positive reduction step is also evaluated.

5.3.1 Database Used and Parameters Configuration

Standard PA view CXRs from two image databases are used in this study. The first database obtained from the UAHCXR consists of 58 cases of sputum smear positive pulmonary TB. 37 cases in this database are typical infectious TB with AS and 21 cases are atypical TB without AS. Ground truth information was determined by a panel of three independent expert chest radiologists. One of the experts helped to draw the region of AS. The second database obtained from the JSRT [80] contains 93 CXRs of normal cases and 154 CXRs of abnormal cases with solitary nodule which is defined as a discrete, well-marginated, rounded opacity less than or equal to 3 cm in diameter [83]. Images in this database were also evaluated with the consensus of three chest radiologists. Finally, CXRs in these two image databases are grouped into four datasets: D1 − typical infectious TB with AS (37 cases); D2 − atypical TB without AS (21 cases); D3 − Normal cases without AS (93 cases); D4 − Nodule cases without AS (154 cases).

For computational efficiency, the original CXRs from D1-D4 are resized as 512×512

(or close to this size) with a fixed pixel spacing [0.8 mm, 0.8 mm]. All the experiments are implemented in Matlab 2007b on an Intel Pentium 4 CPU 2.8G Hz with 2G RAM computer. Parameters used in Preprocessing (ERF-ASM and CLAHE), SVM and Adaboost classification are set as default values in these algorithms. The selected multiple features used for comparison include 6 features of HS, 5 features based on FS, 4 features based on GLCM, one feature of FD value, 6 features using the LBP histogram, 36 features based on HOG and 6 features based on TM. Details of these features are listed in Table 5.1.

Table. 5.1. Multiple features used for comparison

Multiple Features from	No. of Features	Used Features
HS	6	Mean, standard deviation, smoothness, skewness, uniformity, and entropy calculated from the intensity histogram.
FS	5	Energy, variance, entropy, low frequency energy ratio, low/high frequency energy ratio calculated from FS.
GLCM	4	Contrast, correlation, energy, homogeneity calculated from the co-occurrence matrix in horizontal, vertical and diagonal directions.
FD	1	Fractal dimension
LBP	6	Mean, standard deviation, smoothness, skewness, uniformity, and entropy calculated from the histogram of LBP values.
HOG	36	4 cells of 9 major orientation of the gradient
TM	6	Moments of T_{00}, T_{01}, T_{10}, T_{11}, T_{12}, T_{21}

5.3.2 Experiments and Analysis

Before doing the comparative experiments, we need to figure out the ground truth vagueness problem. Although radiologist has helped to extract the ground truth region containing AS, some sliding windows are vague to be defined as positive windows with AS due to the incomplete coverage of the target. Fig. 5.4 shows an example where the red block is one of the sliding windows and white closed contour is the ground truth drawn by the radiologist. To resolve such ambiguity, we define an area ratio R = (area of AS in the window) / (area of the window) to divide the positive and negative windows. Given a

threshold value t, a positive window should satisfy $R \geq t$, and vice versa. For example in Fig. 5.4, if $t = 3/4$, the red block will not be defined as the positive windows with AS.
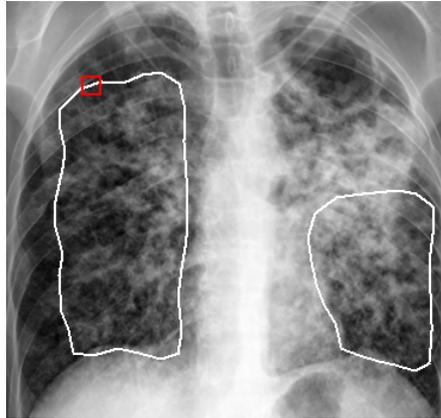
Fig. 5.4. An example of a window (red) containing part of AS (white contour).

We use sensitivity, specificity and precision, which are widely used in medical domain, to evaluate the classification performances. These parameters are defined as follows:

$$\text{Sensitivity} = \frac{\# \text{ of true positives}}{\# \text{ of true positives} + \# \text{ of false negatives}}$$

$$\text{Specificity} = \frac{\# \text{ of true negatives}}{\# \text{ of true negatives} + \# \text{ of false positives}}. \qquad (5.1)$$

$$\text{Precision} = \frac{\# \text{ of true positives}}{\# \text{ of true positives} + \# \text{ of false positives}}$$

Note that sensitivity measures the proportion of the correctly classified AS windows over the entire ground truth AS windows, specificity measures the proportion of the correctly classified non-AS windows over the entire ground truth non-AS windows, and precision measures the proportion of the correctly classified AS windows over all detected AS windows.

Comparative Experiment 1 – Individual Feature Descriptors Evaluation

Since the region of AS only occupies a small part of a CXR, to make the sample size between normal and abnormal cases relatively balanced in training, we randomly select 25 preprocessed CXRs in D1 to obtain positive windows and 6 preprocessed CXRs from D2 to D4 to get negative windows. For testing, we select the rest from D1 and randomly select 10

images from D2 to D4. The classification results using only one type of features with SVM are listed in Table 5.2. The corresponding receiver operating characteristic (ROC) curves by tuning threshold t is shown in Fig. 5.5. The Area Under the Curve (AUC) is also calculated and listed in Table 5.2. In comparison, the classification results of using individual features with Adaboost and the corresponding ROC curves are illustrated in Table 5.3 and Fig. 5.6.

Table. 5.2. SVM classification results using individual features.

Results	t	1/8	2/8	3/8	4/8	5/8	6/8	7/8	1	Avg	AUC
Only HS	Sen	94.5%	94.1%	93.2%	92.4%	90.4%	89.6%	89.2%	90.1%	91.7%	
	Spe	34.7%	37.1%	41.7%	42.4%	48.3%	50.6%	52.7%	55.6%	45.4%	0.729
	Pre	37.5%	36.6%	36.7%	35.5%	36.1%	35.0%	33.9%	33.4%	35.6%	
Only FS	Sen	97.9%	98.3%	98.2%	98.1%	98.2%	52.8%	6.3%	0.6%	68.8%	
	Spe	26.1%	27.7%	29.0%	30.3%	31.8%	93.2%	99.4%	100.0%	54.7%	0.796
	Pre	35.8%	34.8%	33.7%	32.9%	32.0%	70.0%	75.4%	80.0%	49.3%	
Only GLCM	Sen	93.6%	93.7%	93.1%	92.5%	92.7%	91.9%	92.0%	91.7%	92.6%	
	Spe	80.9%	82.3%	84.0%	85.2%	85.8%	87.5%	88.3%	90.0%	85.5%	0.913
	Pre	67.4%	67.5%	68.1%	68.5%	68.1%	68.9%	68.5%	69.7%	68.3%	
Only FD	Sen	74.6%	71.1%	69.2%	66.7%	64.1%	62.5%	60.0%	56.5%	65.6%	
	Spe	79.1%	82.7%	84.6%	86.7%	88.8%	91.1%	91.8%	93.1%	87.2%	0.804
	Pre	56.0%	57.7%	58.4%	59.7%	61.2%	64.1%	63.2%	63.6%	60.5%	
Only LBP	Sen	97.4%	97.8%	97.9%	98.3%	98.6%	98.8%	97.9%	98.7%	98.2%	
	Spe	89.3%	89.1%	89.1%	90.3%	90.6%	90.7%	91.4%	91.7%	90.3%	**0.946**
	Pre	77.9%	76.4%	76.6%	77.0%	76.7%	75.6%	75.8%	75.0%	76.4%	
Only HOG	Sen	95.3%	95.7%	96.3%	96.2%	96.5%	96.5%	95.5%	95.9%	96.0%	
	Spe	72.5%	73.7%	75.7%	77.3%	78.0%	79.2%	80.8%	82.0%	77.4%	0.885
	Pre	55.3%	54.7%	55.3%	55.5%	54.8%	54.1%	53.8%	53.1%	54.6%	
Only TM	Sen	97.7%	90.0%	67.7%	51.0%	42.2%	15.9%	1.0%	0.3%	45.7%	
	Spe	34.3%	47.7%	63.6%	76.8%	84.5%	98.1%	99.9%	100.0%	75.6%	0.745
	Pre	34.7%	36.4%	36.7%	39.3%	42.9%	68.3%	77.8%	100.0%	54.5%	
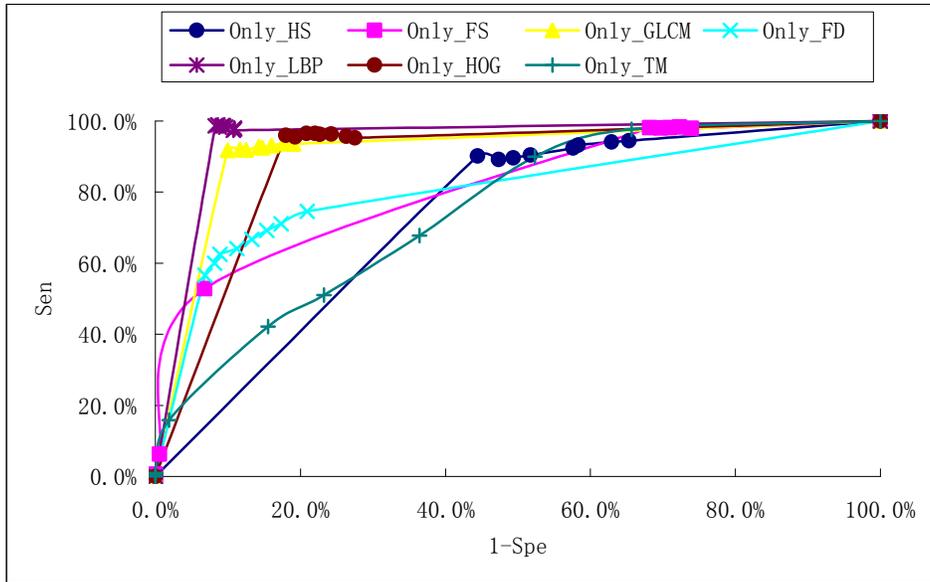
Fig. 5.5. ROC curves of different features using SVM.

Table. 5.3. Adaboost classification results using individual features.

Results	t	1/8	2/8	3/8	4/8	5/8	6/8	7/8	1	Avg	AUC
Only HS	Sen	86.9%	84.0%	85.4%	79.9%	79.7%	73.4%	70.4%	68.8%	78.6%	0.737
	Spe	49.2%	53.5%	53.4%	60.1%	60.7%	66.9%	70.3%	72.4%	60.8%	
	Pre	56.7%	54.1%	52.9%	49.8%	48.2%	44.1%	41.0%	38.1%	48.1%	
Only FS	Sen	91.6%	86.7%	86.0%	83.6%	81.9%	82.2%	77.7%	76.7%	83.3%	0.858
	Spe	71.8%	77.3%	79.1%	80.4%	80.3%	81.6%	83.5%	85.0%	79.9%	
	Pre	72.0%	69.4%	67.8%	66.0%	64.1%	62.3%	58.9%	56.2%	64.6%	
Only GLCM	Sen	93.4%	93.6%	92.7%	92.4%	92.3%	92.0%	92.4%	92.6%	92.7%	0.917
	Spe	83.6%	85.3%	84.7%	87.3%	88.0%	89.6%	90.1%	90.8%	87.4%	
	Pre	81.0%	79.9%	78.7%	77.7%	76.6%	75.0%	73.5%	71.5%	76.7%	
Only FD	Sen	81.6%	82.3%	78.7%	73.6%	74.2%	74.5%	72.9%	66.6%	75.5%	0.822
	Spe	73.8%	73.8%	77.4%	81.2%	81.3%	81.7%	83.4%	89.6%	80.3%	
	Pre	73.7%	72.4%	70.2%	67.6%	66.4%	64.6%	62.1%	57.6%	66.8%	
Only LBP	Sen	96.2%	96.3%	96.3%	96.4%	96.7%	96.6%	96.6%	96.4%	96.4%	**0.959**
	Spe	94.1%	94.4%	94.5%	95.4%	95.7%	96.2%	96.4%	95.8%	95.3%	
	Pre	91.4%	90.9%	90.1%	89.8%	89.3%	88.4%	87.5%	85.5%	89.1%	
Only HOG	Sen	89.7%	91.1%	90.4%	90.8%	92.1%	90.9%	87.7%	90.5%	90.4%	0.885
	Spe	81.0%	81.4%	82.9%	83.7%	84.9%	86.2%	87.7%	87.7%	84.4%	
	Pre	72.3%	71.2%	69.7%	68.6%	67.5%	65.4%	62.6%	61.1%	67.3%	
Only TM	Sen	80.7%	78.1%	76.1%	75.7%	72.9%	71.2%	67.4%	64.4%	73.3%	0.785
	Spe	67.9%	70.2%	72.8%	73.6%	76.2%	77.7%	80.2%	85.0%	75.5%	
	Pre	65.7%	63.4%	61.3%	59.8%	57.4%	54.8%	51.3%	47.7%	57.7%	
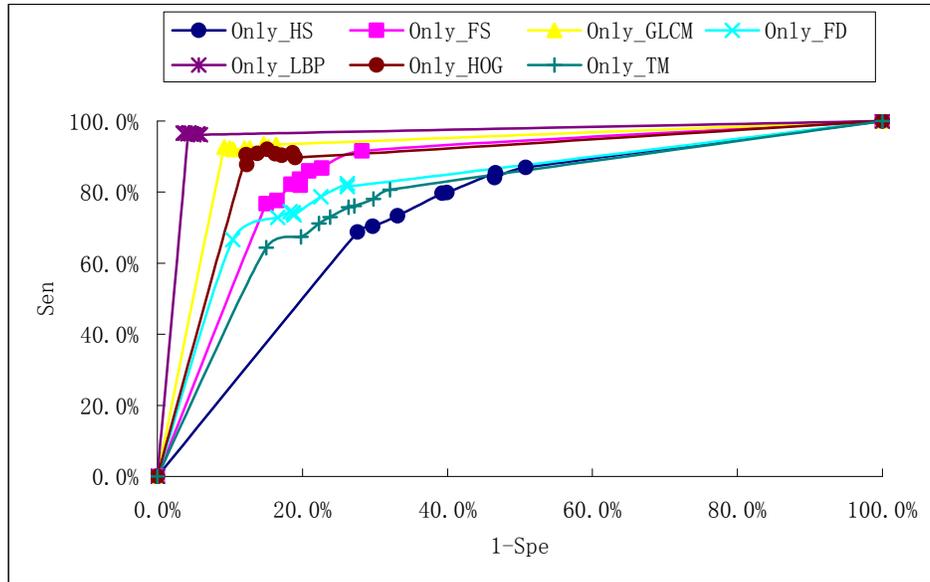
Fig. 5.6. ROC curves of different features using Adaboost.

From the above experimental results, it is observed that LBP outperforms the other features with largest AUC. GLCM and HOG also achieve good performance. While keeping the similar sensitivity, classifier Adaboost greatly increase the precision and specificity outperforming SVM. Note that some ROC curves are not monotonically increasing due to the ambiguity of defining true positives using the threshold value *t*.

Comparative Experiment 2 – Different Combination Strategies of LBP, HOG and GLCM

Our hypothesis is to use multiple strong features to deliver better outcome. Thus, we perform the second comparison experiments using different combination strategies of LBP, HOG and GLCM based features. Adaboost classification results using LBP combined with other features are listed in Table 5.4 and 5.5, respectively. The corresponding ROC curves with different *t* are shown in Figs. 5.7 and 5.8. Since it is difficult to discriminate different curves, the rectangular regions in Figs. 5.7 and 5.8 are enlarged and shown in the middle of the figures. It could be observed that the combination of LBP, GLCM and HOG achieves the overall best performance in both SVM and Adaboost classification. For the performance comparison between SVM and Adaboost classifiers, Adaboost greatly outperforms SVM in specificity and precision while keeping similar high sensitivity. The Adaboost classifier

using LBP, GLCM and HOG based features provides around 5% and 12% improvement in average specificity and precision, which means it not only reduces the false positives but also increases the accuracy of the total detected positives significantly.

Table. 5.4. SVM classification results using combined features of LBP, GLCM and HOG.

Results	t	1/8	2/8	3/8	4/8	5/8	6/8	7/8	1	Avg	AUC
LBP +HOG	Sen	97.0%	97.7%	98.0%	97.6%	98.4%	98.9%	98.6%	98.8%	98.1%	0.946
	Spe	87.8%	88.1%	89.3%	90.1%	90.8%	90.9%	91.7%	92.1%	90.1%	
	Pre	74.0%	73.1%	74.0%	74.3%	74.7%	73.6%	73.7%	72.8%	73.8%	
LBP +GLCM	Sen	97.5%	97.9%	98.2%	98.1%	98.3%	98.2%	98.2%	98.1%	98.1%	0.950
	Spe	88.5%	88.7%	89.7%	90.4%	90.8%	91.6%	92.1%	92.5%	90.5%	
	Pre	78.3%	77.5%	77.9%	78.2%	78.0%	78.1%	77.7%	77.7%	77.9%	
LBP +GLCM +HOG	Sen	98.5%	98.6%	98.7%	98.6%	99.2%	99.4%	99.5%	99.4%	99.0%	**0.956**
	Spe	88.5%	88.9%	89.4%	90.4%	90.7%	91.5%	92.0%	92.6%	90.5%	
	Pre	78.4%	77.7%	77.4%	78.1%	77.7%	77.8%	77.5%	77.0%	77.7%	

Table. 5.5. Adaboost classification results using different combination strategies of LBP, HOG and GLCM based features.

Results	t	1/8	2/8	3/8	4/8	5/8	6/8	7/8	1	Avg	AUC
LBP +HOG	Sen	95.9%	96.2%	96.2%	97.0%	96.7%	96.5%	96.7%	96.2%	96.4%	0.962
	Spe	94.4%	95.1%	95.3%	95.7%	96.0%	96.7%	96.4%	96.7%	95.8%	
	Pre	91.2%	90.7%	90.1%	89.7%	89.1%	88.2%	87.3%	86.1%	89.1%	
LBP +GLCM	Sen	96.1%	96.3%	96.8%	97.0%	97.6%	97.3%	97.4%	96.6%	96.9%	0.963
	Spe	93.4%	93.6%	94.0%	94.7%	95.0%	95.9%	96.3%	96.5%	94.9%	
	Pre	92.4%	91.9%	91.5%	91.0%	90.5%	89.7%	88.9%	87.9%	90.5%	
LBP +GLCM +HOG	Sen	97.1%	97.5%	97.1%	97.4%	97.8%	98.4%	97.4%	97.7%	97.5%	**0.968**
	Spe	94.2%	94.7%	94.8%	95.0%	95.4%	95.6%	96.5%	96.5%	95.3%	
	Pre	92.0%	91.4%	91.0%	90.5%	90.0%	89.3%	88.4%	87.4%	90.0%	
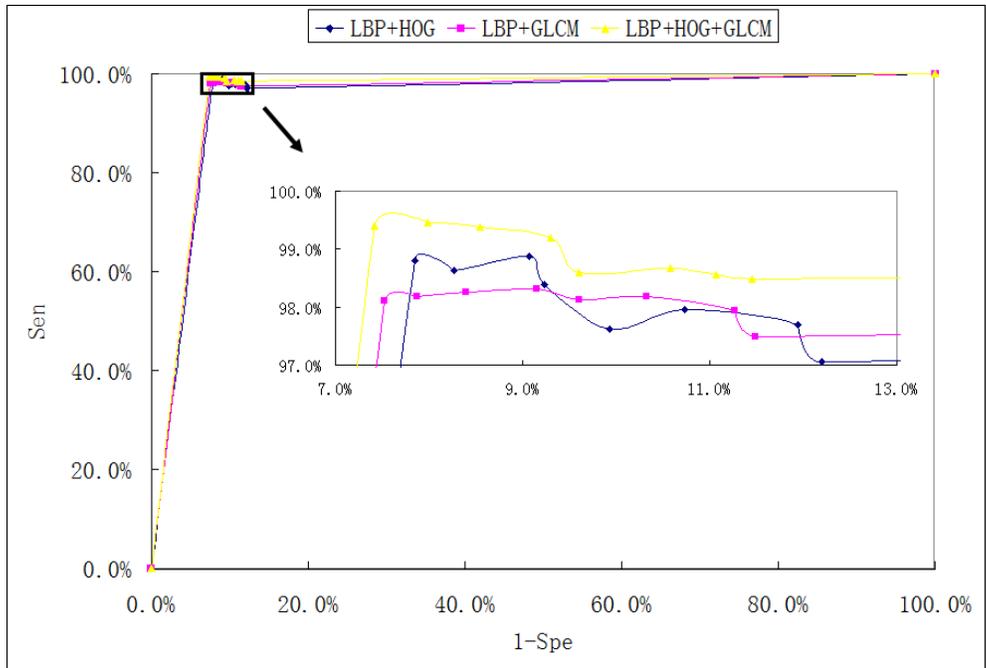
Fig. 5.7. ROC curves using different combination of LBP, GLCM and HOG features with SVM. To better discriminate different curves, the rectangle area has been enlarged and shown in the middle.
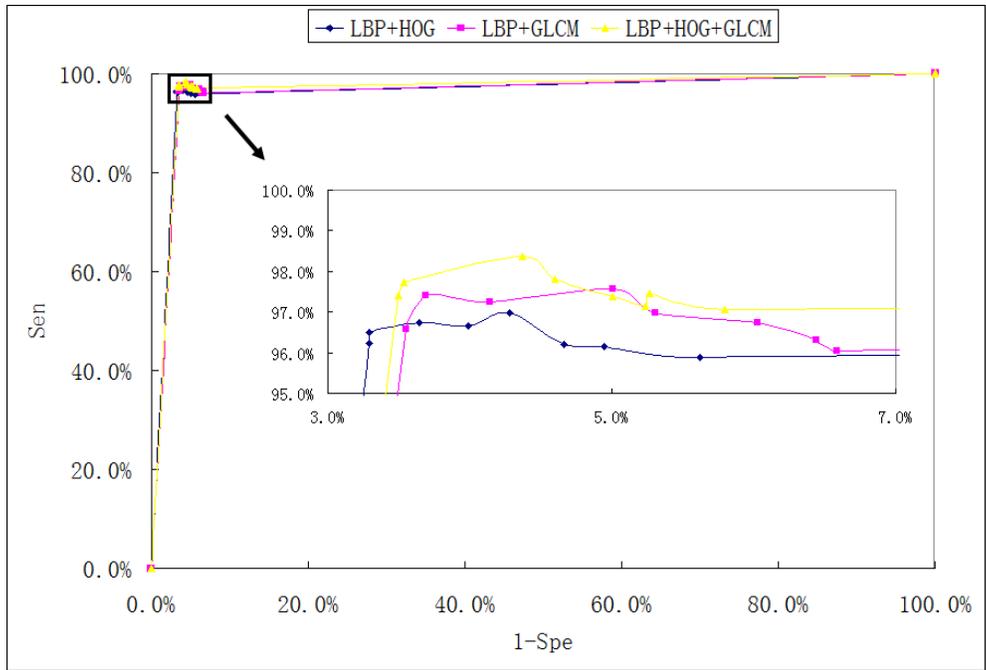
Fig. 5.8. ROC curves using different combination of LBP, GLCM and HOG features with Adaboost. To better discriminate different curves, the rectangle area has been enlarged and shown in the middle.

Comparative Experiment 3 – Final AS Detection with and without False Positive Reduction

Based on the outcome of our comparison analysis, we apply LBP+GLCM+HOG features and Adaboost classifier in the final AS detection system. Considering different characteristics among D2, D3 and D4 (atypical TB without AS, normal cases without AS and nodule cases without AS, respectively), we conduct tests for the datasets D1 with D2 (D1D2), D1 with D3 (D1D3), and D1 with D4 (D1D4), respectively. Notice that in Table 5.5, the best performance of using LBP+GLCM+HOG with Adaboost is achieved when $t = 3/4$, thus the threshold for the whole datasets tests is chosen as $t = 3/4$. The performance of the final AS detection without further false positive reduction is shown in Table 5.6.

Table. 5.6. Final AS detection results without false positive reduction

Results	D1D2	D1D3	D1D4
Sen	98.37%	98.12%	98.25%
Spe	93.50%	99.99%	99.98%
Pre	63.34%	99.49%	98.13%

Quantitative analysis shows that the proposed CAD system achieves both high sensitivity and specificity. Examples of AS detection results of CXRs from D1 are shown in Fig. 5.9. It could be observed that the detected positive windows are quite consistent with the ground truth drawn by radiologist. However, the specificity and precision in the test of D1D2 is relatively lower. It is because lots of false positive windows are detected in the images of D2 comparing to D3 and D4. This higher false positive rate (lower specificity) in D1D2 test could be explained as more lesions similar to AS caused by atypical TB are observed in CXRs in D2. Although the specificity in tests of D1D3 and D1D4 are close to 100%, there are still several images in D3 and D4 containing false positive windows. Examples of those false positive windows are shown in Fig. 5.10.
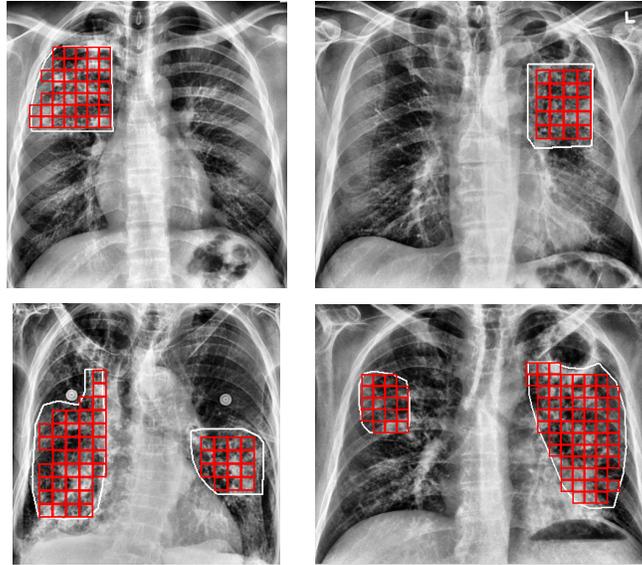
Fig. 5.9. True positive windows detected in D1 comparing with the ground truth (white contour).

| CXR1 from D2 | CXR2 from D2 | CXR3 from D3 | CXR4 from D4 |
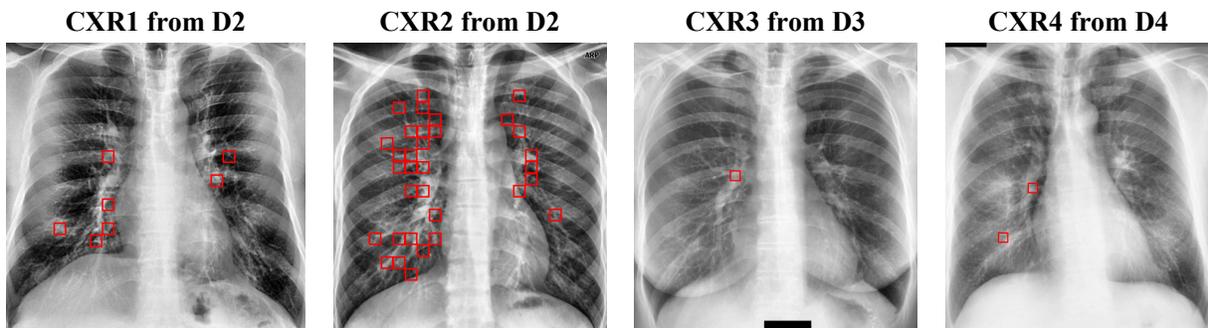
Fig. 5.10. False positive windows detected in D2, D3 and D4.

After applying the smoothing technique, the final AS detection results are greatly refined for the specificity and precision, while keeping the high sensitivity. See Table 5.7 for the refined results. Examples of false positive windows in CXR1, CXR3 and CXR4 in Fig. 5.10 are all removed except a few left in CXR2 (See Fig. 5.11 for the rest of false positive windows).

Table. 5.7. Final AS detection results with false positive reduction.

Results	D1D2	D1D3	D1D4
Sen	92.37%	91.74%	92.49%
Spe	97.43%	100%	100%
Pre	80.39%	100%	100%

CXR2 from D2
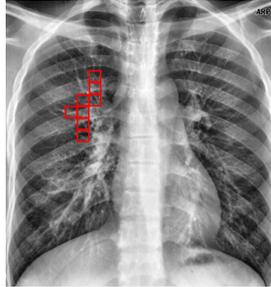
Fig. 5.11. False positive windows detected after the smoothing technique.

5.4 Summary

In this chapter, a hybrid intelligent system is proposed to detect acinar shadow regions in CXRs. This novel CAD system takes advantages of integrating multiple features of LBP, GLCM and HOG into the Adaboost classifier. False positives are further reduced by introducing a morphological smoothing technique. The proposed CAD system shows an outstanding performance with more than 92% sensitivity, 97% specificity and 80% precision, which make it an effective tool to improve diagnostic performance. Our CAD system is efficient to eliminate a large number of irrelevant cases so that the radiologists can focus on a smaller set of significant cases.

Chapter 6

Conclusions and Future Work

Despite the technological advances in medical diagnosis, accurate detection of infectious TB still poses challenges due to complex image features and thus infectious TB continues to be a public health problem of global proportions. Currently, the detection of TB is mainly conducted visually by radiologists examining CXRs. To reduce the backlog of CXR examination and provide more precise quantitative assessment, CAD systems for potential lung lesions have been increasingly adopted and commercialized for clinical practice. However, little work has been done beyond lung nodules. Thus, our research focuses on developing an intelligent CAD system to automatically detect typical radiographic patterns of infectious TB in CXRs. This CAD system is expected to work as a supporting tool to alert radiologists on suspected infectious TB features that could have been neglected. In this chapter, our research work is concluded and new directions for future research are discussed.

6.1 Conclusions

Followed by the general framework of our proposed CAD system stated in Section 1.1.4, the research work can be summarized in three parts:

Part I – Preprocessing using ERF-ASM for Lung Field Segmentation

Automatic and accurate lung field segmentation has become a mandatory preprocessing step in most CAD systems in chest radiology. Considering the deficiency of existing lung field segmentation techniques, we propose an efficient ERF-ASM algorithm by introducing the global ERF into PIG-based automatic initialization stage and a new non-linear exponential point evolution based segmentation stage. Experimental results on both normal and TB CXRs shows the superior performance (higher accuracy and lower computational

cost) of our ERF-ASM over other existing techniques. The segmented lung fields using ERF-ASM will help to restrict the detection area of TB radiographic patterns.

Part II – TB Cavity Detection

Since cavitation in the ULZ is a strong indicator of typical TB, our next research effort focuses on efficient TB cavity detection. To compare with the existing cavity detection HKG system, we propose an efficient CFDS technique which first detects cavity candidates at coarse feature level using GTM and LHFC and then identifies cavity contours at fine feature level using CFC. To accurately extract the cavity contour, we propose the HIE for cavity edge enhancement and develop a snake model using new dynamic external force. Experimental results show that our proposed CFDS outperforms HKG at all aspects including accuracy of candidates selection; accuracy of contour segmentation; and accuracy of final cavity contour classification.

Part III – TB AS Detection

Since the presence of AS in the vicinity of a cavity increases the probability of typical TB, we use a sliding window paradigm in MFB to detect AS candidates. We conduct a comprehensive study applying different feature descriptors (HS, FS, GLCM, FD, LBP, HOG, and TM) and investigate their influence on classification using the well-known SVM and Adaboost classifiers. Based on the comparative experimental results, the final hybrid system for AS detection utilizes the combined LBP, HOG and GLCM features with Adaboost classification. We also apply a simple smoothing technique to remove the false positive windows efficiently. Note that *Part II* and *Part III* could work in parallel for real practice.

6.2 Future Work

This research work can be extended in the following topics to make the CAD system more beneficial for clinical practice:

Topic I – Other Radiographic Patterns Detection

Volume loss is a hallmark of TB as a result of its destructive and fibrotic nature. Since the lung fields could be extracted using our ERF-ASM, we would like to study the variation of lung volumes to recognize the volume loss. Besides, hilar and mediastinal lymphadenopathy could be a good indicator for HIV-related TB. The detection of lymphadenopathy will also help to reduce the false positive rate of typical TB.

Topic II – Incorporation of Patient Record

An intelligent CAD system is not only to accomplish tasks as detecting suspicious TB radiographic patterns from images, but also needs to learn and judge from other patient record information such as epidemiologic risk, pulmonary or constitutional symptoms, leukocyte count number, and medical history. Therefore, more efficient classifiers which could handle with different data structure need to be studied and proposed.

Topic III – Cooperation with Radiologists for the Evaluation of the CAD system

All in all, the objective of the CAD system is to provide ''second opinion'' to assist radiologists' image readings. The proposed CAD system needs to be tested with both experienced TB experts and inexperienced radiologists to evaluate its impact on the readers' diagnostic judgment. With their feedback, further feature analysis might be necessary.

Bibliography

[1] R. Long and E. Ellis, *Canadian tuberculosis standards (6th ed.)*, Public Health Agency of Canada, 2007.

[2] "Global tuberculosis control," *World Health Organization Report*, 2011.

[3] "Tuberculosis in Canada," http://www.phac-aspc.gc.ca/tbpc-latb/index-eng.php

[4] H.P. McAdams, E. Samei, J. Dobbins III, et al, "Recent advances in chest radiography," *Radiology*, 2006, 241(3):663-83.

[5] C.J. Vyborny, "The AAPM/RSNA physics tutorial for residents: image quality and the clinical radiographic examination," *Radiographics*, 1997, 17:479-498.

[6] K. Doi, "Overview on research and development of computer-aided diagnostic schemes," *Seminar in Ultrasound CT & MR*, 2004, 25: 404-410.

[7] G.S. Lodwick, T.E. Keats, and J.P. Dorst, "The coding of Roentgen images for computer analysis as applied to lung cancer," *Radiology*, 1963, 81(2):185–200.

[8] B.V. Ginneken, B.M.T.H. Romeny, and M. Viergever, "Computer-aided diagnosis in chest radiography: a survey," *IEEE Trans. on Medical Imaging*, 2001, 20(12):1228-1241.

[9] B.V. Ginneken, L. Hogeweg, and M. Prokop, "Computer-aided diagnosis in chest radiography: beyond nodules," *European Journal of Radiology*, 2009, 72(2):226-30.

[10] http://www.riveraintech.com/ClearRead-detect-product-overview.html

[11] http://www.healthcare.philips.com/cn_zh/products/xray/products/radiography/cad_chest

[12] http://edda-tech.com/Products/IQQAChest.html

[13] S. Katsuragawa, K. Doi, H. MacMahon, et al, "Quantitative computer-aided analysis of lung texture in chest radiographs," *Radiographics*, 1990, 10(2):257-69.

[14] S. Kido, J. Ikezoe, H. Naito, et al, "Fractal analysis of interstitial lung abnormalities in chest radiography," *Radiographics*, 1995, 15(6):1457-64.

[15] B.V. Ginneken, S. Katsuragawa, B.M.T.H. Romeny, et al, "Automatic detection of abnormalities in chest radiographs using local texture analysis," *IEEE Trans. on Medical Imaging*, 2002, 21(2):139-149.

[16] N.M. Noor, O.M. Rijal, H. Shaban, et al, "Discrimination between two lung diseases using chest radiographs," in *Proc. of the IEEE EMBC'05, Shanghai,* 2005, 3320-3323.

[17] R. Shen, I. Cheng, and A. Basu, "A hybrid knowledge-guided detection technique for screening of infectious pulmonary tuberculosis from chest radiographs," *IEEE Trans. on Biomedical Engineering*, 2010, 57(11):2646-2656.

[18] J. Tan, U.R. Acharya, C. Tan, et al, "Computer-assisted diagnosis of tuberculosis: a first order statistical approach to chest radiograph," *Journal of Medical Systems*, 2012, 2751-2759.

[19] S.A. Patil, "Texture analysis of TB X-ray images using image processing techniques," *Journal of Biomedical and Bioengineering*, 2012, 3(1):53-56.

[20] T. Xu, M. Mandal, R. Long, et al, "Gradient vector flow based active shape model for lung field segmentation in chest radiographs," in *Proc. of the IEEE EMBC'09, Minnesota*, 2009, 3561-3564.

[21] T. Xu, M. Mandal, R. Long, et al, "An edge-region force guided active shape approach for automatic lung field detection in chest radiographs," *Computerized Medical Imaging and Graphics*, 2012, 36(6):452-463.

[22] T. Xu, I. Cheng, and M. Mandal, "An improved fluid vector flow for cavity segmentation in chest radiographs," in *Proc. of the IEEE ICPR'10, Istanbul*, 2010, 3376-3379.

[23] T. Xu, I. Cheng, and M. Mandal, "Automated cavity detection of infectious pulmonary tuberculosis in chest radiographs," in *Proc. of the IEEE EMBC'11, Boston*, 2011, 5178-5181.

[24] T. Xu, I. Cheng, R. Long, et al, "Novel coarse-to-fine dual scale technique for tuberculosis cavity detection in chest radiographs," *EURASIP Journal on Image and Video Processing*, 2013, 2013:3.

[25] T. Xu, I. Cheng, R. Long, et al, "Computer-aided detection of acinar shadows in chest radiographs," submitted.

[26] I. N. Bankman, *Handbook of medical imaging: process and analysis*, Academic Press, 2000.

[27] K. Doi, "Computer-aided diagnosis in medical imaging: historical review, current status and future potential," *Computerized Medical Imaging and Graphics*, 2007, 31:198-211.

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

[29] J.C. Gower, "Generalized Procrustes analysis," *Psychometrika*, 1975, 40:33-51.

[30] B.V. Ginneken, A.F. Frangi, J.J. Staal, et al, "Active shape model segmentation with optimal features," *IEEE Trans. on Medical Imaging*, 2002, 21(8):924-933.

[31] Y. Shi, F. Qi, Z. Xue, et al, "Segmenting lung fields in serial chest radiographs using both population-based and patient-specific shape statistics," *IEEE Trans. on Medical Imaging*, 2008, 27(4):481-494.

[32] T.F. Cootes, C.J. Taylor, and A. Lanitis, "Multi-resolution search with active shape models," in *Proc. of the IEEE ICPR'94*, 1994, A:610-612.

[33] M. Brejl and M. Sonka, "Object localization and border detection criteria design in edge-based image segmentation: automated learning from examples," *IEEE Trans. on Medical Imaging*, 2000, 19:973-985.

[34] H. Li and O. Chutatape, "Boundary detection of optic disk by a modified ASM method," *Pattern Recognition*, 2003, 36:2093-2104.

[35] M.H. Mahoor and M. Abdel-Mottaleb, "Facial features extraction in color images using enhanced active shape model," *FGR*, 2006, 144-148.

[36] F.A. Cosío, "Automatic initialization of an active shape model of the prostate," *Medical Image Analysis*, 2008, 12(4):469-83.

[37] M. Kass, A.Witkin, and D. Terzopoulus, "Snakes: active contour model," *International Journal of Computer Vision*, 1988, 1(4):321-331.

[38] L. Cohen and I. Cohen, "Finite-element methods for active contour models and balloons for 2-d and 3-d images," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 1993, 15(11):1131-1147

[39] C. Xu and J. Prince, "Snakes, shapes, and gradient vector flow," *IEEE Trans. on Image Processing*, 1998, 7(3):359-369.

[40] K. W. Sum and P. Y. S. Cheung, "Boundary vector field for parametric active contours," *Pattern Recognition*, 2007, 40(6):1635-1645.

[41] N. Ray, S. T. Acton, T. Altes, et al, "Merging parametric active contours within homogeneous image regions for MRI-based lung segmentation," *IEEE Trans. On Medical Imaging*, 2003, 22(2):189-199.

[42] T. Wang, I. Cheng, and A. Basu, "Fluid vector flow and applications in brain tumor segmentation", *IEEE Trans. on Biomedical Engineering*, 2009, 56(3):781-789.

[43] T. Tuytelaars and K. Mikolajczyk, "Local invariant feature detectors: a survey", *Foundations and Trends in Computer Graphics and Vision*, 2007, 3(3):177-280.

[44] K. Mikolajczyk and C. Schmid, "A performance evaluation of local descriptors," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 2005, 27(10):1615-1630.

[45] M. Yang, K. Kpalma, and J. Ronsin. "A survey of shape feature extraction techniques," *Pattern Recognition Techniques, Technology and Applications*, P.Y. Yin, Ed. Vukovar,

Croatia: In-Teh, 2008, 43-90.

[46] M. Mellor, B. Hong, and M. Brady, "Locally rotation, contrast, and scale invariant descriptors for texture analysis," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 2008, 30(1):52-61.

[47] T. Ojala, M. Pietikainen, and T. Maenpaa, "Multiresolution gray-scale and rotation invariant texture classification with local binary pattern," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 2002, 24(7):971-987.

[48] N. Dalal and B. Triggs, "Histograms of oriented gradients for human detection", in *Proc. of IEEE CVPR*, San Diego, 2005, 1:886-893.

[49] U. Bagci, M. Bray, J. Caban, et al, "Computer-assisted detection of infectious lung diseases: A review," *Computerized Medical Imaging and Graphics*, 2012, 36(1):72-84.

[50] A.Y. Ng and M.I. Jordan, "On discriminative vs. generative classifiers: a comparison of logistic regression and naive Bayes," in *Proc. of NIPS'01*, Vancouver 2001, 841-848.

[51] I. Ulusoy and C.M. Bishop, "Comparison of generative and discriminative techniques for object detection and classification," in *Proc. of Toward Category-Level Object Recognition*, Taormina, 2006, 173-195.

[52] C. Cortes and V. Vapnik, "Support-vector networks," *Machine Learning*, 1995, 20(3):273-297.

[53] R.E. Schapire, "A brief introduction to boosting," in *Proc. of IJCAI'99*, Stockholm, 1999, 1401-1406.

[54] H. Zhang and C. Gu, "Support vector machines versus boosting," *SVM Report*, 2008, Electrical Engineering, UC Berkeley.

[55] A. McCarthy, "Computer-aided detection chest X-ray helps identify lung lesions," *Medscape Today News Article*, 2009 (ARRS 2009 Annual Meeting: Abstract 103).

[56] S.G. Armato III, M.L. Giger, K. Ashizawa, et al, "Automated lung segmentation in digitized posteroanterior chest radiographs," *Academic Radiology*, 1998, 5:245-255.

[57] J. Duryea, and J.M. Boone, "A fully automated algorithm for the segmentation of lung fields on digital chest radiographic images," *Medical Physics*, 1995, 22 (2):183-191.

[58] M.F. McNitt-Gray, H.K. Huang, J.W. Sayre, "Feature selection in the pattern classification problem of digital chest radiograph segmentation," *IEEE Trans. on Medical Imaging*, 1995, 14(3):537-547.

[59] Z. Shi, P. Zhou, L. He, et al. "Lung segmentation in chest radiographs by means of Gaussian kernel-based FCM with spatial constraints," in *Proc. of Fuzzy Systems and Knowledge Discovery'09*, Tianjin, 2009, 428-432.

[60]M.S. Brown, L.S. Wilson, B.D. Doust, et al, "Knowledge-based method for segmentation and analysis of lung boundaries in chest X-ray images," *Computerized Medical Imaging and Graphics*, 1998, 22:463-477.

[61]M. Park, L.S. Wilson, J.S. Jin, "Automatic extraction of lung boundaries by a knowledge-based method," *Visual Information Processing*, 2001, 2:14-19.

[62]H. Luo, R. Gaborski, R. Acharya, "Automatic segmentation of lung regions in chest radiographs: a model guided approach," in *Proc. of IEEE ICIP*, Vancouver, 2000, 2:483-486.

[63]M. Leventon, W. Grimson, and O. Faugeras, "Statistical shape influence in geodesic active contours," in *Proc. of IEEE CVPR*, Hilton Head Island, 2000, 316-323.

[64]A. Tsai, A. Yezzi, W. Wells, et al. "Model-based curve evolution techniques for image segmentation," in *Proc. of IEEE CVPR*, Kauai, 2001, 463-468.

[65]L.A. Vese and T.F. Chan, "A multiphase level set framework for image segmentation using the Mumford and Shah model," *International Journal of Computer Vision*, 2002; 50(3):271-293.

[66]Y. Wang, L. Wang, S. Xiang, et al, "Level set evolution with locally linear classification for image segmentation," in *Proc. of IEEE ICIP*, Brussels, 2011, 3361-3364.

[67]B. Xavier, V. Pierre and T. Jean-Philippe, "A variational model for object segmentation using boundary information and shape prior driven by the Mumford-Shah functional," *International Journal of Computer Vision*, 2006; 28(2):145-162.

[68]P. Annangi, S. Thiruvenkadam, A. Raja, et al, "A region based active contour method for x-ray lung segmentation using prior shape and low level features," in *Proc. of IEEE ISBR*, Rotterdam, 2010, 892-895.

[69]I. Iglesias, P.G. Tahoces, M. Souto, et al, "Lung segmentation on postero-anterior digital chest radiographs using active contours," *LNCS*, 2004, 3138:538-546-.

[70]T. Yu, J. Luo, A. Singhal, et al, "Shape regularized active contour based on dynamic programming for anatomical structure segmentation," in *Proc. of SPIE Medical Imaging 2005: Image Processing*, 5747:419-430.

[71]D.K. Iakovidis, M.A. Savelonas, and G. Papamichalis, "Robust model-based detection of the lung field boundaries in portable chest radiographs supported by selective thresholding," *Measurement Science and Technology*, 2009, 20:104019.

[72]G. Behiels, F. Maes, D. Vandermeulen, et al, "Evaluation of image features and search strategies for segmentation of bone structures in radiographs using active shape models," *Medical Image Analysis*, 2002, 6:47-62.

[73]B. Li and S.T. Acton, "Automatic active model initialization via Poisson inverse

gradient," *IEEE Trans. on Image Processing*, 2008, 17(8):1406-1420.

[74] Y. Yu and S.T. Acton, "Speckle reducing anisotropic diffusion," *IEEE Trans. on Image Processing*, 2002; 11(11):1260-1270.

[75] J. Canny, "A computational approach to edge detection," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 1986; 8(6):679-714.

[76] D. Pham and J. Prince, "Adaptive fuzzy segmentation of magnetic resonance images," *IEEE Trans. on Medical Imaging*, 1999, 18:737-752.

[77] D. Comaniciu, P. Meer, "Mean shift: a robust approach toward feature space analysis," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 2002, 24(5):603-619.

[78] K.V. Leemput, F. Maes, D. Vandermeulen, et al, "Automated Model-Based Tissue Classification of MR Images of the Brain," *IEEE Trans. on Medical Imaging*, 1999, 18(10):897-908.

[79] Y. Xu, B. Giritharan, and J. Oh, "Gradient vector flow driven active shape for image segmentation," in *Proc. of IEEE ICME*, Beijing, 2007, 2058-2061.

[80] Japanese Society of Radiological Technology. http://www.jsrt.or.jp/english.html.

[81] B.V. Ginneken, "Computer-aided diagnosis in chest radiographs," Ph.D. dissertation, Utrecht Univ., Utrecht, The Nederlands. 2001.

[82] A. Lau, J. Barrie R. Long, et al, "The public health consequences of smear positive pulmonary tuberculosis in patients with typical and atypical chest radiographs," TB Report, 15th Annual International Union Against Tuberculosis and Lung Disease – North American Region Conference, Vancouver, February 24-26, 2011.

[83] C. L. Daley, M. B. Gotway, R. M. Jasmer, *Radiographic Manifestations of Tuberculosis: A Primer for Clinicians*, Francis J. Curry National Tuberculosis Center, 2009.

[84] B.S. Reddy and B.N. Chatterji, "An FFT-based technique for translation, rotation, and scale-invariant image registration," *IEEE Trans. on Image Processing*, 1996, 5(8):1266-1271.

[85] Y.H. Lin and C.H. Chen, "Template matching using the parametric template vector with translation, rotation and scale invariance," *Pattern Recognition*, 2008, 41(7):2413-2421.

[86] X. Wang, T.X. Han and S. Yan, "An HOG-LBP human detector with partial occlusion handling," in *Proc. of IEEE ICCV*, Kyoto 2009, 32-39.

[87] T. Hastie, R. Tibshirani and J. Friedman, *The Elements of Statistical Learning: Data Mining, Inference, and Prediction (2nd ed.)*, Springer, 2009.

[88] Y. Sato, S. Nakajima, N. Shiraga, et al, "3D multiscale line filter for segmentation and visualization of curvilinear structure in medical images," *Medical Image Analysis*, 1998,

2(2):143-168.

[89] J. Liu, J.M. White and R.M. Summers, "Automated detection of blob structures by Hessian analysis and object scale," in *Proc. of IEEE ICIP*, HongKong, 2010, 841-844.

[90] P.K. Saha, and J.K. Udupa, "Scale-based image filtering preserving boundary sharpness and fine structure," *IEEE Trans. on Medical Imaging*, 2001, 20(11):1140-1155.

[91] H. Zhang, M. Wan and Z. Bian, "Complementary tensor-driven image coherence diffusion for oriented structure enhancement," *EURASIP Journal of Advanced Signal Processing*, 2011, 70-70.

[92] K. Zuiderveld, "Contrast limited adaptive histogram equalization," Chapter VIII.5, *Graphics Gems IV*, P.S. Heckbert (Eds.), Cambridge, MA, Academic Press, 1994, pp. 474-485.

[93] M. Gil, E.G. Sarabia, et al, "Fuzzy c-means clustering for noise reduction, enhancement and reconstruction of 3D ultrasonic images," in *Proc. of ETFA'99*, Barcelona, 1999, 465-472.

[94] C. Di Ruberto and A. Dempster, "Circularity measures based on mathematical morphology," *Electronic Letters*, 2000, 36(20):1691-1693.

[95] G. Dong, N. Ray, and S. Acton, "Intravital leukocyte detection using the gradient inverse coefficient of variation," *IEEE Trans. On Medical Imaging*, 2005, 24(7):910-924.

[96] C. Bishop, *Pattern Recognition and Machine Learning*, Springer, 2006, ch. 1.

[97] C. C. Chang and C. J. Lin, "LIBSVM: a library for support vector machines," 2011, software available at: http://www.csie.ntu.edu.tw/~cjlin/libsvm/

[98] G.D. Tourassi GD, "Journey toward computer-aided diagnosis: role of image texture analysis," *Radiology*, 1999, 213(2):317-320.

[99] M. Mirmehdi, X. Xie, and J. Suri, *Handbook of Texture Analysis*, Imperial College Press, London, 2008.

[100] R. Mukundan, S.H. Ong, and P.A. Lee, "Image analysis by Tchebichef moments", *IEEE Trans. On Image Processing*, 2001, 10(9):1357-1364.

Appendices

Appendix A: Procrustes Alignment Method

The flow diagram of the generalized Procrustes analysis [29] is shown in Fig. A.1. It could be observed that the alignment works in an iterative way and the critical part is to align two shapes.
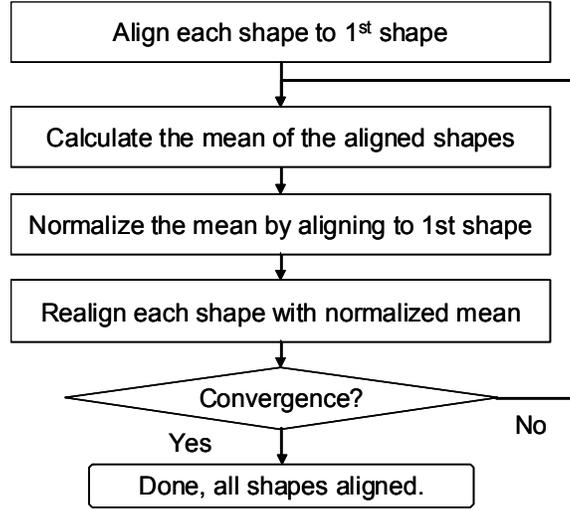
```
┌─────────────────────────────────────────┐
│      Align each shape to 1st shape       │
└─────────────────────────────────────────┘
                    │
┌─────────────────────────────────────────┐
│   Calculate the mean of the aligned shapes │
└─────────────────────────────────────────┘
                    │
┌─────────────────────────────────────────┐
│  Normalize the mean by aligning to 1st shape │
└─────────────────────────────────────────┘
                    │
┌─────────────────────────────────────────┐
│   Realign each shape with normalized mean │
└─────────────────────────────────────────┘
                    │
              ◇ Convergence? ◇ ── No
                    │ Yes
┌─────────────────────────────────────────┐
│        Done, all shapes aligned.         │
└─────────────────────────────────────────┘
```

Fig. A.1 The flow diagram of the generalized Procrustes analysis.

Considering a pair of shapes $C^{(1)}$ and $C^{(2)}$, the alignment is achieved using the following least-square approach:

$$\arg Min(E) = \arg Min\left\{[T(C^{(2)}) - C^{(1)}]^T [T(C^{(2)}) - C^{(1)}]\right\} \tag{A.1}$$

where $T(C^{(2)}) = s \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} x_i^{(2)} \\ y_i^{(2)} \end{bmatrix} + \begin{bmatrix} t_x \\ t_y \end{bmatrix}$ is an affine transform, and s, θ, $\begin{bmatrix} t_x \\ t_y \end{bmatrix}$

are the scaling, rotation and translation factors. If we denote $a = s\cos\theta, b = s\sin\theta$, and

let $\begin{bmatrix} x_i \\ y_i \end{bmatrix} = \begin{bmatrix} x_i^{(2)} \\ y_i^{(2)} \end{bmatrix}, \begin{bmatrix} x_i' \\ y_i' \end{bmatrix} = \begin{bmatrix} x_i^{(1)} \\ y_i^{(1)} \end{bmatrix}$, Eq. (A.1) then becomes:

$$\arg \min\left[\sum_{i=1}^{N} (ax_i - by_i + t_x - x_i')^2 + (bx_i + ay_i + t_y - y_i')^2 \right] \tag{A.2}$$

Differentiating with respect to each of the variables a, b, t_x, t_y, the solution to Eq. (A.2) leads to four linear Eqs.:

$$
\begin{cases}
\dfrac{\partial E}{\partial a} = 0 \\[2mm]
\dfrac{\partial E}{\partial b} = 0 \\[2mm]
\dfrac{\partial E}{\partial t_x} = 0 \\[2mm]
\dfrac{\partial E}{\partial t_y} = 0
\end{cases}
\Rightarrow
\begin{cases}
a(S_{xx} + S_{yy}) + t_x S_x + t_y S_y = S_{xx'} + S_{yy'} \\
b(S_{xx} + S_{yy}) - t_x S_x + t_y S_y = S_{xy'} - S_{yx'} \\
aS_x - bS_y + t_x = S_{x'} \\
bS_x + aS_y + t_y = S_{y'}
\end{cases}
\tag{A.3}
$$

where $S_x = \dfrac{1}{N}\sum\limits_{i=1}^{N} x_i, S_{xx} = \dfrac{1}{N}\sum\limits_{i=1}^{N} x_i^2, S_{xx'} = \dfrac{1}{N}\sum\limits_{i=1}^{N} x_i x_i'$, and similar to other S. Linear

Eqs. (A.3) could be rewritten as:

$$
\begin{bmatrix}
S_{xx} + S_{yy} & 0 & S_x & S_y \\
0 & S_{xx} + S_{yy} & -S_x & S_y \\
S_x & -S_y & 1 & 0 \\
S_y & S_x & 0 & 1
\end{bmatrix}
\begin{bmatrix} a \\ b \\ t_x \\ t_y \end{bmatrix}
=
\begin{bmatrix}
S_{xx'} + S_{yy'} \\
S_{xy'} - S_{yx'} \\
S_{x'} \\
S_{y'}
\end{bmatrix}
\tag{A.4}
$$

Here I just prove the 1st Eq. in (A.3) as follows:

$$
\frac{\partial E}{\partial a} = \sum_{i=1}^{N} 2x_i(ax_i - by_i + t_x - x_i') + 2y_i(bx_i + ay_i + t_y - y_i')
$$

$$
\Rightarrow \sum_{i=1}^{N} a(x_i^2 + y_i^2) + t_x x_i + t_y y_i = \sum_{i=1}^{N} x_i x_i' + y_i y_i'
\tag{A.5}
$$

Assuming S is centered, i.e. $S_x = S_y = 0$, then we could get:

$$
\begin{cases}
a = \dfrac{S_{xx'} + S_{yy'}}{(S_{xx} + S_{yy})} \\[4mm]
b = \dfrac{S_{xy'} - S_{yx'}}{(S_{xx} + S_{yy})} \\[4mm]
t_x = S_{x'} \\[1mm]
t_y = S_{y'}
\end{cases}
\tag{A.6}
$$

Appendix B: LSSP for Lung Field Segmentation

To compare the ASM-based technique to the level-set with shape priors (LSSP), we applied a hybrid LSSP [68] which is more specific for lung field segmentation. In addition, we carefully selected the parameters in this implemented LSSP detailed below.

Before discussing the parameters' selection, we first explain the difference between LSSP in [67] and [68], and why LSSP [68] is more specific for lung field segmentation.

The hybrid LSSP framework is based on the minimization of the following energy functional:

$$E = w_r E_{region} + w_b E_{boundary} + w_s E_{shape} \tag{B.1}$$

where w_r, w_b, w_s are the weights of three energy terms. The major differences between LSSP in [67] and [68] are the region and boundary energy terms. The comparison is listed in Table B.1.

Table B.1. Comparison of Energy Terms between LSSP in [67] and [68]

	LSSP [67]	LSSP [68]
E_{region}	$\int_{\Omega_{in}} H(\phi)(\lvert I - u_{in}\rvert^2 + \mu\lvert\nabla u_{in}\rvert^2)d\Omega +$ $\int_{\Omega_{out}} (1 - H(\phi))(\lvert I - u_{out}\rvert^2 + \mu\lvert\nabla u_{out}\rvert^2)d\Omega$	$\int_{\Omega_{in}} H(\phi)((\frac{I - c_1}{\sigma_1})^2 + \ln(\sigma_1))d\Omega +$ $\int_{\Omega_{out}} (1 - H(\phi))((\frac{I - c_2}{\sigma_2})^2 + \ln(\sigma_2))d\Omega$
$E_{boundary}$	$\int_{\Omega} \delta(\phi)\lvert\nabla(\phi)\rvert g(\lvert\nabla I\rvert)d\Omega$	$\int_{\Omega} \delta(\phi)\lvert\nabla(\phi)\rvert(d_E^2 + \lambda)d\Omega$
where u_{in} and u_{out} are smooth approximations of the original image I in Ω_{in} and Ω_{out}. (c_1, σ_1) and (c_2, σ_2) are the parameters of the distributions inside and outside. H and δ are the Heaviside and Dirac functions, respectively. g is an edge detecting function. $d_E = wd_{E1} + d_{E2}$ denotes the sum of distance transforms of the Canny edge map d_{E1} and the castrophenic angle corner feature d_{E2}. μ and w are balancing weights. λ is the weight of length term.		

It can be seen that LSSP [68] is much more specific for lung field segmentation due to:

(1) the incorporation of Gaussian distributions assumption of the inside and outside the

contour in the region term, which is more precise than the homogeneity assumption in the CV model; (2) low level edge features including castrophenic angle corner feature are considered, which is more specific than the general image gradient.
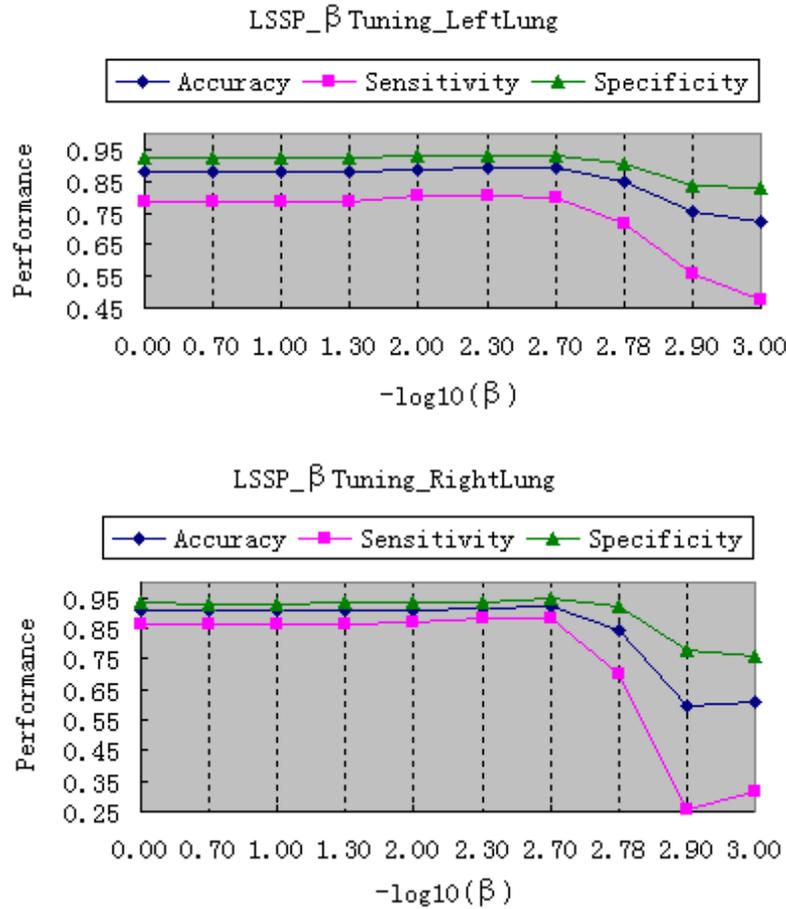

Fig. B.1. Segmentation performance in 20 randomly selected CXRs with changing β and fixed λ and w.

Paper [68] simplifies the weights between edge, region and shape terms as $w_r = w_b = 1$, and only uses one parameter β to control the weight of shape term. Thus, three explicit parameters in LSSP [68] need to be decided: weight of shape term β, weight of length term (smoothing parameter) λ and low level edge feature balancing weight w. We first fix λ and w with appropriate values, tune β in the test of randomly selected

20 CXRs (See the following Fig. B.1 illustrating the results of tuning β). Based on the segmentation performance we find the optimal β. Then using this β, we fix w to find the optimal λ. Finally optimal w could also be decided. The final segmentation results shown in the revised paper are based on these optimal parameters. Note that the similarity coefficient of 0.86 in our LSSP implementation is consistent with the result (0.88) reported in [68]. This shows that the parameter value selection is not biased.