Syracuse University

SURFACE

2011

Noise-Enhanced and Human Visual System-Driven Image Processing: Algorithms and Performance Limits

Renbin Peng
Syracuse University

Follow this and additional works at: https://surface.syr.edu/eecs_etd

Part of the Electrical and Computer Engineering Commons

ABSTRACT

This dissertation investigates the problem of image processing based on stochastic resonance (SR) noise and human visual system (HVS) properties, where several novel frameworks and algorithms for object detection in images, image enhancement and image segmentation as well as the method to estimate the performance limit of image segmentation algorithms are developed.

Object detection in images is a fundamental problem whose goal is to make a decision if the object of interest is present or absent in a given image. We develop a framework and algorithm to enhance the detection performance of suboptimal[1] detectors using SR noise, where we add a suitable dose of noise into the original image data and obtain the performance improvement. Micro-calcification detection is employed in this dissertation as an illustrative example. The comparative experiments with a large number of images verify the efficiency of the presented approach.

Image enhancement plays an important role and is widely used in various vision tasks. We develop two image enhancement approaches. One is based on SR noise, HVS-driven image quality evaluation metrics and the constrained multi-objective optimization (MOO) technique, which aims at refining the existing suboptimal image enhancement methods. Another is based on the selective enhancement framework, under which we develop several image enhancement algorithms. The two approaches are applied to many low quality images, and they outperform many existing enhancement algorithms.

Image segmentation is critical to image analysis. We present two segmentation algorithms driven by HVS properties, where we incorporate the human visual perception

---

[1]Suboptimality may be due to inaccurate statistical models, model mismatch and system limitation, such as fixed decision threshold.

factors into the segmentation procedure and encode the prior expectation on the segmentation results into the objective functions through Markov random fields (MRF). Our experimental results show that the presented algorithms achieve higher segmentation accuracy than many representative segmentation and clustering algorithms available in the literature.

Performance limit, or performance bound, is very useful to evaluate different image segmentation algorithms and to analyze the segmentability of the given image content. We formulate image segmentation as a parameter estimation problem and derive a lower bound on the segmentation error, i.e., the mean square error (MSE) of the pixel labels considered in our work, using a modified Cramér–Rao bound (CRB). The derivation is based on the biased estimator assumption, whose reasonability is verified in this dissertation. Experimental results demonstrate the validity of the derived bound.

Noise-Enhanced and Human Visual System-Driven Image Processing: Algorithms and Performance Limits

A Thesis
Presented to
The Academic Faculty
by

Renbin Peng

In Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy

Department of Electrical Engineering and Computer Science
Syracuse University
August 2011

To my grandparents,

Yuan, Shulan and Ren,Wenxuan

and also to my parents,

Ren, Huiling and Peng, Shiyuan

and my wife He, Jing

ACKNOWLEDGEMENTS

First and foremost I would like to gratefully and sincerely thank my advisor Dr. Pramod K. Varshney. His patience, support, encouragement, mentoring and guidance helped me finish this dissertation. It was his valuable advice on both academic matters and personal affairs which helped me overcome many difficulties in these years. It has been an honor to be his Ph.D. student. I would also like to give special thanks to Dr. Lixin Shen and Dr. Ruixin Niu for their comments, suggestions and discussions during my research work. I am also heartily thankful to Dr. Chilukuri Mohan, Dr. Hao Chen and Dr. Ioana Coman whose suggestions and encouragement from the preliminary to the concluding levels broadened my horizon and benefited my research greatly. Special thanks are also due to all my defense committee members for their time in reading my dissertation and making helpful comments to improve its quality.

The members of Dr. Varshney's research group have contributed immensely to my personal and professional time at Syracuse. The group has been a source of friendships as well as good advice and collaboration. In particular, I would like to thank Dr. Min Xu, Dr. Long Zuo, Dr. Priyadip Ray, Arun Subramanian, Satish Iyengar and Dr. Engin Masazade.

Finally, I would like to thank my family for all their love and encouragement. For my grandparents and parents who raised me and supported me in all my pursuits. And for my loving, supportive, encouraging and patient wife, Jing He, whose faithful support in these years is so appreciated. Thank you.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ACRONYMS AND ABBREVIATIONS

2SR-GBAD	two-peak stochastic resonance noise-enhanced Gaussian background assumption-based detection
2SR-GBAD-M	two-peak stochastic resonance noise-enhanced Gaussian background assumption-based detection with memory
ANN	artificial neural network
APD	amplitude probability distribution
BEBS	boundary element-based segmentation
BiHE	bi-histogram equalization
CAD	computer-aided diagnosis
CLAHE	contrast limited adaptive histogram equalization
CM	classification-maximization
CRB	Cramér–Rao bound
CSF	cerebrospinal fluid
CSF	contrast sensitivity function
Det-En	iterative detection-based enhancement method
DDSM	digital database for screening mammography
DWT	discrete wavelet transform
EM	expectation-maximization
FNM	finite normal mixture

FPI	false positives per image
GaSR	Gaussian stochastic resonance noise
GBAD	Gaussian background assumption-based detector
GGD	general Gaussian detector
GGD_ID	GGD-based iterative detector
GM	gray matter
GMM	Gaussian mixture modeling
HDSS-1	human visual system-driven segmentation scheme 1
HDSS-2	human visual system-driven segmentation scheme 2
HOSLW	high order statistics method based on local maxima detection and adaptive wavelet transform
HVS	human visual system
ICM	iterative conditional modes
i.i.d.	independent and identically-distributed
It-En	iterative enhancement
JND	just-noticeable difference
Joint-En	joint enhancement
*k***-NN**	k-nearest neighbor
LSEWRI	level set evolution-based method
MAP	maximum *a posteriori* probability
MCMC	Markov chain monte carlo
MDL	minimum description length

MI	mutual information
MIAS	mammographic image analysis society
MLE	maximum likelihood estimation
MNCut	multi-scale normalized cuts-based segmentation
MOO	multi-objective optimization
MRF	Markov random fields
MSE	mean square error
PDE	partial differential equations
pdf	probability density function
P_F	probability of false alarm
P_D	probability of detection
pmf	probability mass function
RACM	region-based active contour model
ROI	region of interest
RVM	relevance vector machine
SαS	symmetric α stable
SNR	signal-to-noise ratio
SR	stochastic resonance
SSIM	structural similarity
TP	true-positive fraction
TrSR	triangle SR noise
TwSR	two peak SR noise

UnSR   uniformly distributed SR noise

VCM   varying coefficient model

WM   white matter

CHAPTER I

INTRODUCTION

We are living in a world where we are immersed in a variety of visual information, which includes different forms and shapes, colors and textures, motion and tranquility. As the carrier of this information, still images, graphic and videos play more and more important role in our lives. Human perception is capable of acquiring, integrating and interpreting the visual information around us, while it is very challenging if we expect to impart such capability to a machine [1]. Often images, graphics and videos are of varying quality due to the acquisition procedure [2], network/communication conditions [3][4][5], image/video compression [6], etc. These factors further add to the difficulties for a machine to extract the useful information from different visual scenarios. It is, therefore, very important to understand and develop superior techniques to process images, graphics and videos, such as detection, enhancement/restoration, segmentation, registration and fusion.

The goal of this dissertation is to develop novel image processing algorithms which improve the capability when a machine is employed to interpret the information embedded in images. In this chapter, we briefly review the basic problems in several image processing areas that are addressed in this dissertation. They include object detection in images, image enhancement and image segmentation.

1.1 Object Detection in Images

The goal of object detection in images is to determine whether a specified object is present or not in a given image, and, if present, to determine its location, size or other parameters. This task could range from identifying a location to identifying and registering components of a particular object class at various levels of detail [7]. We could be required to outline precisely the object in the image, to detect a certain number of well-defined landmarks on the object, or to determining a deformation from a prototype of the object in the image. There are various types of deformations, e.g., a simple 2D Affine map or a more detailed non-linear map. The object itself may have different degrees of variability. It may be a rigid 2D object, such as a fixed desk surface, or a 2D view of a 3D object, or it may be a highly deformable object such as a breast. All these issues are included while considering object detection problems. Detection implies identifying some aspects of the particular way the object is present in the image, i.e., some partial description of the object instantiation. The image processing pipeline for object detection shown in Fig. 1.1, which includes the feature extraction step followed by a detection algorithm.
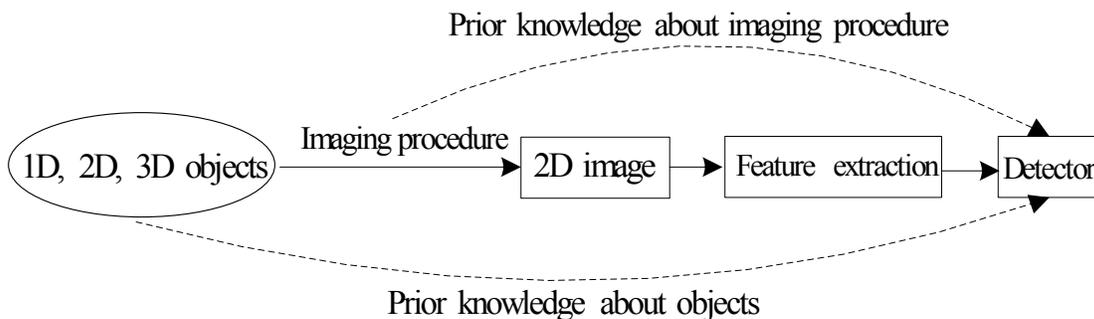
Figure 1.1: Processing pipeline for object detection in images**.**

The general problem of object detection in static images is challenging, because the object detection system is required to distinguish a particular class of objects from all others [8]. This requires the algorithm to possess a model of the object class that has high inter-class and low intra-class variability. Another difficulty is that a robust object detection algorithm should be able to detect objects in uneven illumination, objects which are rotated, and objects that are partially occluded or whose parts blend in with the background. The outline of an object is, under all of the above conditions, usually altered and its entire form may not be discernible. The problem becomes even more challenging if recognition, a step that usually follows detection, is required. Recognition refers to the classification among objects or subclasses of a general class of objects, which is present in a particularly isolated region of the image [7].

The object detection systems can be classified into three major categories [8]. The first category consists of systems that are model-based, in which a model is defined for the object of interest and the system attempts to find a match between this model and different parts of the image [9]. The second category includes the image invariance methods where matching is based on a set of image pattern relationships, e.g., brightness levels. In the ideal case, the pattern relationships uniquely determine the objects being searched for [10]. The third set of object detection systems refer to the example-based learning algorithms [11]. These systems learn the features of a class from sets of labeled positive and negative examples, i.e., training data.

From the machine learning perspective, we can also divide all object detection methods as generative methods and discriminative methods [12]. In generative methods, we attempt to learn significant features of an object, and then combine these features in a

suitable way to synthesize a new image. In a classification scenario, these models can be reversed and indicate the probability that this particular pixel was generated from this model. In the discriminative methods, we learn a decision rule, i.e., classifier, and assign features representations of images to different classes.

In this dissertation, we focus on the problem of detecting objects in images based on statistical decision theory. We model the probability density functions (*pdf*s), i.e., features, of the object and the background (non-object) under two hypotheses, i.e., presence or absence of the object in a particular location of an image. The detection is carried out under the Neyman-Pearson criterion [13], because we assume that the prior probabilities of occurrence of the object and the background are not available. In real-world applications, it is usually not easy to have the full knowledge of the probability distributions of the object and background, so we learn and update the *pdf*s under the two hypotheses in an iterative manner with the detection proceeding. Thus, our method falls in the categories of both model-based and example-based learning algorithms, and is a combination of the generative method and discriminative method.

More specifically, we consider a binary statistical decision problem, where we wish to choose between the two hypotheses

$$\begin{cases} H_0 : p_{\bar{y}}(\bar{y}; H_0) = p_0(\bar{y}) \\ H_1 : p_{\bar{y}}(\bar{y}; H_1) = p_1(\bar{y}) \end{cases} \tag{1.1}$$

where $\bar{y}$ is an N-dimensional data vector, i.e., $\bar{y} \in R^N$. $p_0(\bar{y})$ (or $p(\bar{y}; H_0)$) and $p_1(\bar{y})$ (or $p(\bar{y}; H_1)$) are the *pdf*s of $\bar{y}$ under H_0 (background) and H_1 (object) hypotheses, respectively. During the decision process, a test is necessary to choose between the two

hypotheses, which can be completely characterized by a critical function, or decision function, $\phi(\bar{y})$, $0 \leq \phi(\bar{y}) \leq 1$, and

$$\phi(\bar{y}) = \begin{cases} 1 : T(\bar{y}) > \gamma \\ \beta : T(\bar{y}) = \gamma \\ 0 : T(\bar{y}) < \gamma \end{cases} \tag{1.2}$$

where T is the test statistic which is a function of $\bar{y}$. γ is the threshold, and $0 \leq \beta \leq 1$ is a suitable number used for randomization.

The detection performance of this test can be evaluated in terms of probability of detection (P_D) and probability of false alarm (P_F),

$$P_D^{\bar{y}} = \int_{R^N} \phi(\bar{y}) p_1(\bar{y}) d\bar{y} \tag{1.3}$$

$$P_F^{\bar{y}} = \int_{R^N} \phi(\bar{y}) p_0(\bar{y}) d\bar{y} \tag{1.4}$$

where $P_D^{\bar{y}}$ and $P_F^{\bar{y}}$ represent the P_D and P_F of the detector based on the input $\bar{y}$, respectively.

It is well known that under the Neyman-Pearson criterion, the optimal detector is a likelihood ratio test given as

$$L(\bar{y}) = \frac{p_1(\bar{y})}{p_0(\bar{y})} \underset{\underset{decide\ H_0}{<}}{\overset{\overset{decide\ H_1}{>}}{}} \gamma \tag{1.5}$$

The threshold γ is found from

$$P_F^{\bar{y}} = \int_{\{\bar{y}:L(\bar{y})>\gamma\}} p_0(\bar{y}) d\bar{y} = \alpha \tag{1.6}$$

where α is the desired value of the P_F.

5

1.2 Image Enhancement

Image enhancement aims at producing images with improved brightness/contrast and detail, so as to better represent the visual information. It is widely used in many areas, such as vision, remote sensing, dynamic scene analysis, autonomous navigation and biomedical image analysis. Fig. 1.2 shows an example of enhancing an MRI image, where some details in the image become more obvious after enhancement (Fig. 1.2 (b)).

(a)   (b)
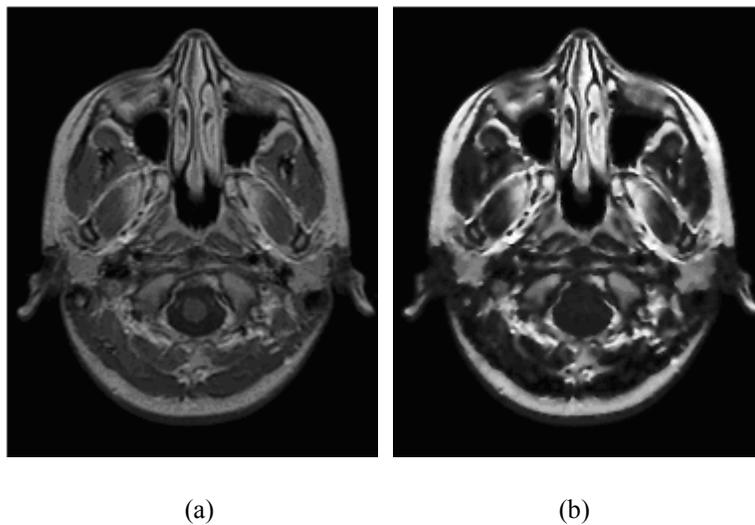
Figure 1.2: An image enhancement example. (a) Original image with low contrast; (b) enhanced image with higher contrast.

A comprehensive survey of image enhancement techniques can be found in [14]. Based on the methodology involved, image enhancement methods fall into four main categories: point operations, spatial operations, transform operations and pseudo-coloring [2].

Point operations are zero memory operations in which a given gray level is mapped into another gray level according to a transformation. Point operations include contrast

stretching, window slicing, noise clipping and histogram modeling. Histogram equalization is one of the well-known point operation methods. Standard histogram equalization [2] processes images globally, thus often causes intensity saturation and the image is far from a natural one. Its improved versions, such as bi-histogram equalization (BiHE) [15] and contrast limited adaptive histogram equalization (CLAHE) [16], attempted to achieve the natural contrast enhancement by preserving the brightness of an original image and processing images locally.

Spatial operations perform enhancement on local neighborhoods of the input pixels, where the image is convolved with a finite impulse response filter. Typical methods include noise smoothing, median filtering and unsharp masking [17]. A standard spatial operation is median filtering. It performs well in some cases, such as suppressing the pepper-and-salt noise. Another popular spatial operator is adaptive Wiener filter [18]. It can carry out the filtering procedure locally and adaptively. For a digital mammography application, J. Dengler *et al.* [19] used a morphological filter to improve the perspicacity of micro-calcifications by enhancing the small square.

Transform operations are carried out on a transformed image and followed by the inverse transformation for image reconstruction. Typical transform operations include the classical Fourier transform, fuzzy logic transform and wavelet decomposition. Fuzzy image enhancement can achieve superior performance in mammogram enhancement [20], compared with several popularly used methods.

Pseudo-coloring methods map a set of images into a color image, whose disadvantage is that extensive interactive trials are required to determine a satisfactory mapping.

Since humans are often the ultimate evaluators of image quality, it is desirable to include human visual system (HVS) aspects in image enhancement. Recently, several methods have used HVS models during enhancement [21][22]. Their goal is to simulate the function of HVS for discriminating between useful and useless data [23] to improve the enhancement procedure. For example, in [22], HVS-based segmentation was carried out in spatial domain to generate three image regions prior to image enhancement. HVS considerations could be incorporated in any of the above four classes of enhancement methods, thereby improving enhancement performance from human visualization point of view.

1.3 Image Segmentation and Performance Limit of Segmentation Algorithms

Image segmentation plays a critical role in image analysis. It subdivides an image into its constituent parts in order to extract information regarding objects of interest, and has an impact on all the subsequent image analysis tasks, such as object classification and scene interpretation [24]. Fig. 1.3 shows an example of image segmentation.

Image segmentation is a challenging problem in computer vision, and a wide variety of solutions have been presented. These include thresholding techniques [25], Markov random fields (MRF)-based approaches [26][27], multi-resolution algorithms [28] and partial differential equations (PDE)-based methods [29]. Surveys of image segmentation techniques can be found in [24][30][31]. Based on the image information being employed for the segmentation task, image segmentation algorithms can be classified into three

categories: region-based segmentation, boundary- or edge-based segmentation and the methods combining both region and boundary (edge) information.

Figure 1.3: An image segmentation example (from MATLAB Central). (a) Original image; (b) segmented image.

Region-based segmentation methods aim at exploiting the image contextual information, such as spatial dependency or spatial distribution. The segmented images are expected to consist of regions within which the image content is homogeneous, while the contrast between neighboring regions is high. Typical methods falling into this category include region growing, watershed, some MRF-based methods [26], mean-shift [32] and the recently presented lossy data compression-based approach [33]. Segmentation methods based on the boundary or edge information are designed to exploit the discontinuity of the image features, such as difference in texture or pixel intensity, on the two sides of the boundary. Typical methods in this group include gradient-based methods, such as the Canny edge detector [34], line detection methods, such as the Hough transform [35], those taking into account the interaction between boundaries (or edges)

[36][37][38], and the methods derived from physics models [39][40]. There also exist algorithms that combine region-based and boundary-based segmentations in order to benefit from fusing these two complementary approaches. There are two types of algorithms that belong to this category. The first type of algorithms carry out region and boundary segmentations sequentially [41][42][43], where one segmentation method is employed as the preprocessing or initialization step of another. The second type performs segmentation by considering region and boundary information simultaneously [44][45].

While development of efficient segmentation algorithms is highly desirable, the assessment of their performance is also very important. There are basically three groups of methods for segmentation evaluation [24]. These include analysis methods, empirical goodness methods and empirical discrepancy methods [24]. The analysis methods treat the algorithms for segmentation directly, such as the evaluation of the convergence rate, the computation speed and the reasonability of the objective function design. Empirical goodness methods judge the segmented image so as to indirectly assess the performance of algorithms using quantities such as intra-region uniformity, inter-region contrast and region shape. Empirical discrepancy methods compare the segmented image with the reference image and use their difference to evaluate the performance of algorithms. For instance, position and number of mis-segmented pixels and feature values of segmented objects are all performance indicators falling into this class. Surveys of the evaluation techniques for image segmentation can be found in [24][43][44].

Much progress has been made recently in evaluating the segmentation results, but performance of such methods tends to vary as widely as the techniques themselves. As a result, the performance of the evaluation methods is far from being satisfactory. In [24],

the authors listed some of the factors which limit the advancement of evaluation methods and, in turn, the performance improvement of segmentation algorithms. These factors include a lack of common mathematical models or general strategy for evaluation, the challenges in defining wide-ranging performance metrics and statistics, the difficulties in defining the ground truth, large costs in performing comprehensive evaluations and the fact that the testing data are not representative enough for actual applications.

We note that given a specific image, among all the factors possibly affecting the performance assessment of segmentation algorithms, the most important factor is the image content. Therefore, an investigation of the performance limit or performance bound, which is only associated with the available image data and is independent of the segmentation algorithms, will be very helpful to evaluate the efficiency of image segmentation techniques. A tight performance bound can tell us what the best achievable performance of any image segmentation algorithm is for the specific image content. Thus, performance bounds can also be used to study how the image content or image preprocessing operations affect segmentation performance. The gap between the actual segmentation error of an approach and a tight bound can provide us with the efficiency of that segmentation approach and available room for improvement.

Several contributions in the literature have developed the bounds from a statistical perspective, e.g., the work in [46] based on the finite normal mixture (FNM) model assumption, that presented in [47] for studying the performance of multi-spectral image segmentation based on Rissanen's minimum description length (MDL) criterion, and the one presented in [48] based on MRF-based assumption. All these algorithms attempt to determine the performance bounds which can serve as benchmarks for the image dataset

and segmentation algorithms.

1.4 Image Processing Based on Stochastic Resonance Noise

Basically, there exist two approaches to improve the performance of an image processing system. One involves the redesign of the existing systems or algorithms, while another involves the preprocessing of the image data before sending them to the original processing system. In our work, we propose a novel preprocessing approach to improve the system performance via stochastic resonance (SR) noise, where system performance is improved by adding some suitable noise to the input image signal but the original image processing system is kept unchanged.

Traditionally considered as a nuisance, noise can sometimes play a constructive role in signal processing. SR is one such nonlinear physical phenomenon where the output signals of some nonlinear systems can be amplified by adding noise to the input. Fig. 1.4 shows an image segmentation result from our experiments, where we add noise to the original image and better segmentation is observed (Figs. 1.4 (c) and (d)) while the original image segmentation algorithm remains fixed.

First proposed as an explanation for the ice ages in 1981 [49], the SR effect has been observed and applied in numerous nonlinear systems ranging from a bistable system to a crayfish [50]. Classic SR signatures include a signal-to-noise ratio (SNR) gain and a mutual information (MI) gain. In signal detection theory, SR also plays a very important role in improving signal detectability. In [51] and [52], improvement of detection performance of a weak sinusoid signal is reported. In [53], an optimal SR based procedure for a general binary hypotheses detection problem was addressed, the

underlying mechanism of the SR phenomenon was explored, the improvability conditions via SR were established subject to a false alarm constraint and the optimum SR noise form was determined. The work was later extended to the more general case where the detector itself can also be modified [54]. In this dissertation, we employ SR noise to refine several image enhancement algorithms and also to improve some existing lesion detection algorithms for breast cancer diagnosis.

<div style="text-align:center">(a)   (b)   (c)   (d)</div>
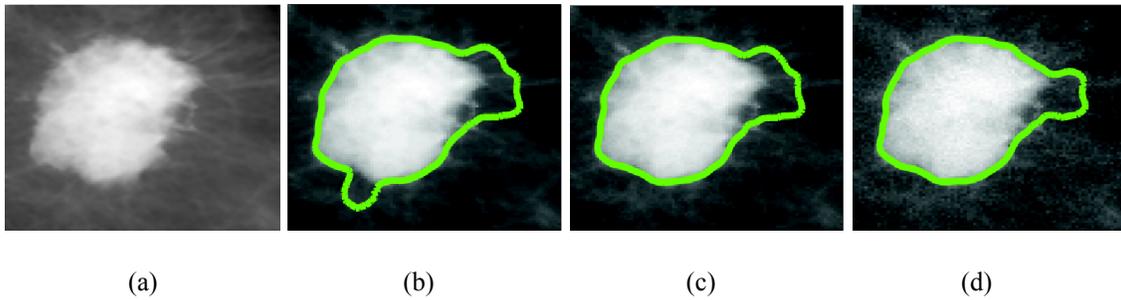
Figure 1.4: Stochastic resonance noise-enhanced image segmentation. (a) Original image; (b) segmentation result of the level set evolution-based method (LSEWRI) [121] without SR noise; (c) segmentation results of LSEWRI with Gaussian SR noise (zero mean and variance of 60); (d) segmentation results of LSEWRI with uniform SR noise (zero mean and variance of 25).

1.5 Human Visual System-Driven Image Processing

HVS-driven image processing is motivated by the fact that, in most circumstances, humans are the ultimate judge of the quality of the image processing results. So, an image processing algorithm is likely to yield more satisfactory results if HVS factors are taken into account during the processing procedure. For example, images can be more accurately segmented, enhancement can be better carried out, and detection can be more efficiently fulfilled, to mention a few major advantages.

In many different image processing applications, the limitations of the HVS can be exploited to improve the performance from a visual quality point of view, based on which many image processing algorithms have been developed. For example, the image enhancement algorithms [21][22][23], as mentioned in Section 1.2, and the image segmentation algorithms [55][56][57] are all driven by HVS, and outperform many "classical" pixel-based algorithms.

Even if the specific requirements for each of these image processing applications are different, the common element of building a computational model of the HVS is always essential. These computational models are closely dependent on the visual properties of the HVS which are characterized by both the psychophysical experiments and the understanding of the physiological evidence. Please refer to [58] for a detailed review of the vision physiology.

HVS models account for a number of psychophysical effects [59], like the luminance and color, multi-channel decomposition, contrast and adaptation, contrast sensitivity and masking [60]. In our work, we incorporate HVS models into the objective functions for image enhancement and image segmentation, and encode the preference of HVS to a good image processing result in a mathematical form, which may make the image processing output more favored by humans.

1.6 *Main Contributions and Dissertation Organization*

The design of efficient object detection, image enhancement and image segmentation algorithms is very important for a variety of applications. In this dissertation, several aspects of these problems are discussed. We present several schemes to improve object

detection, image enhancement and image segmentation based on SR noise and HVS properties. We also analyze the performance limit of image segmentation algorithms and present quality metrics for evaluate the image enhancement results. Main contributions and organization of the dissertation are described as follows.

This chapter has introduced the background of object detection, image segmentation and image enhancement. The significance of investigating the performance limit of image segmentation algorithms was also discussed. The concept and motivation of image processing based on SR noise and HVS properties have also been presented.

In Chapter 2, we show the feasibility of SR noise-enhanced image processing through several examples, where we improve the performance of image processing algorithms by adding noise to the images before the processing procedure while the original image processing algorithms are kept unchanged. We investigate the effect of SR noise on image enhancement, dithering, edge detection and image segmentation.

In Chapter 3, we consider the problem of improving the detection of micro-calcifications in mammograms using SR noise. We develop a SR noise-based detection algorithm and a general detection enhancement framework to improve the performance of the suboptimal detectors. We attempt to reduce the dependence of the determination of the optimum SR noise on the knowledge of the *pdf*s of the object (lesion) and background (normal tissues) by employing iterative learning procedures. We also develop an iterative SR noise-based detection enhancement scheme with memory to improve the efficiency and robustness of the SR noise-based detection systems. Moreover, a more general SR noise-based detection enhancement framework is presented. Our algorithms and the framework are tested on a set of 75 representative abnormal mammograms. They yield

superior performance when compared with several classification and detection approaches developed in our work as well as those available in the literature.

In Chapter 4, we present two image enhancement approaches. One is based on SR noise and HVS properties, and another is based on selective enhancement technique. In the first algorithm, we develop an SR noise-based image enhancement framework and apply it to improve a number of existing suboptimal image enhancement methods. The enhanced image and the optimum parameters of the SR noise are obtained by solving a constrained multi-objective optimization (MOO) problem, where a weighting method is used to mimic different HVS preferences and to reduce the size of the non-dominant solution set as discussed later. The principle of the SR noise-refined image enhancement is explored and the corresponding image enhancement system is presented. Additionally, a novel image quality evaluation metric based on HVS is developed which is used as one of the objective functions.

In the second enhancement approach, we present a selective enhancement framework based on image segmentation techniques, which starts with region of interest (ROI) selection, and is followed by ROI enhancement and background suppression. Several point operation based algorithms under this framework are presented. Compared with some popular enhancement algorithms, our methods enjoy the advantages of flexibility, robustness and low computational burden, which suggest that the presented methods are suitable for real-world CAD applications. We also present a quality metric to evaluate the enhancement results by fusing sub-quality metrics.

In Chapter 5, we present two image segmentation algorithms. In the first algorithm, we aim at designing an algorithm based on HVS properties, with the segmentation

performance robust to the variations in the parameter values of the algorithm. More specifically, we integrate region label estimation for each pixel with boundary localization for each region, according to the quality metrics for region-based and boundary-based segmentation evaluations. These metrics attempt to mimic the preferences of human vision to good segmentation and thus make the segmentation HVS-driven. Under a Bayesian framework, the HVS-driven quality metrics are encoded in the MRF as the priors of the *a posteriori* distribution, which is the objective function for segmentation. Segmentation is carried out by optimizing the objective function which reflects the desired properties of segmentation from both global and local perspectives. Three variations of the algorithm are developed. The first one integrates the region and boundary information simultaneously during segmentation. The second one carries out region-based segmentation and boundary-based segmentation iteratively. The third one takes advantage of only the boundary information for segmentation.

In the second algorithm, we present an image segmentation framework, which is based on a "soft" objective function and considers the effect of the segmentation result for a single pixel on the segmentation performance in local regions. A specific performance measure, the probability of successful detection, is used to show the efficiency and utility of this framework. Moreover, a contrast sensitivity function (*CSF*), as an object feature enhancer, is employed for further improving the segmentation performance, which also makes the segmentation procedure HVS-driven.

In Chapter 6, we formulate image segmentation as a statistical parameter estimation problem and derive Cramér–Rao bounds (CRB) on the performance measure, namely on the mean square error (MSE) of the resulting pixel labels, based on the biased estimator

assumption and Affine bias model. In addition, an approximation is made when computing the expectation of the inverse Fisher information matrix to reduce the computational burden. Bootstrapping technique and empirical approximation to the second-order statistics are employed to overcome the difficulty that the probability distribution of the images is unknown. Our final goal is to derive a tight performance bound for the image segmentation problem and compare the bound with the performance of various segmentation algorithms when applied to different image datasets.

Finally, in Chapter 7, we present the main conclusion of this dissertation and provide an outline of our future research plan.

CHAPTER II

NOISE-ENHANCED IMAGE PROCESSING

In this chapter, we show the efficiency of stochastic resonance (SR) noise-enhanced image processing through several illustrative examples, where the performance of some image processing algorithms are improved through the introduction of SR noise into the original images while the image processing algorithms themselves remain unchanged. The image processing tasks we investigate include image enhancement, dithering, edge detection and image segmentation.

In this chapter, we adjust the parameters of the SR noise manually and the quality of the image processing results is evaluated by observation. In Chapters 3 and 4 we will employ SR noise to improve object detection and image enhancement, where we formulate these two tasks as optimization problems and develop systematic schemes to find the optimum SR noise parameters automatically.

2.1 Types of Stochastic Resonance Noise

Three types of SR noises are investigated in this chapter. They are Gaussian SR noise (GaSR), uniformly distributed SR noise (UnSR), and two peak SR noise (TwSR) [53]. In Section 4.2.4 we will also investigate triangle SR noise (TrSR). The adjustable parameters of the GaSR and UnSR are means (μ) and variances (σ^2), and those of TrSR are amplitude range and the point where its mass function has maximum value. Probability mass function (*pmf*) of TwSR has the form

$$p_{n-Twopeak}(n) = \alpha\delta(n - n_1) + (1 - \alpha)\delta(n - n_2) \tag{2.1}$$

where α and $1-\alpha$ are the occurrence probabilities of the suitable constants n_1 and n_2, $0 \leq \alpha \leq 1$. Its adjustable parameters are α, n_1 and n_2. The SR noise is introduced in the degraded image by point-wise addition.

2.2 Illustrative Examples

2.2.1 Standard Histogram Equalization-Based Image Enhancement

Fig. 2.1 shows the results of enhancing a mammogram with speculated lesions using standard histogram equalization [2] and SR noise. The result shown in Fig.2.1 (b) is based on the standard histogram equalization method. This method processes images globally, and we can see that it does not take into account the features of the lesions which are characterized by the local image contents. The background normal tissues surrounding the lesions are also enhanced, such that the contrast between the lesions and the normal tissues are not increased. In contrast, the SR noise-enhanced histogram equalization method (Fig. 2.1 (c)), where we add Gaussian SR noise, increases the contrast between lesion and background, and also makes the enhanced image more natural than Fig. 2.1 (b) from visualization perspective.

2.2.2 Contrast limited adaptive histogram equalization-Based Image Enhancement

Contrast limited adaptive histogram equalization (CLAHE) [16] was originally designed for enhancing medical images, which processes images locally and is an improved version of the standard histogram equalization method. The experiments are carried out using mammograms with two types of lesions, mass and spiculated lesions, as well as the

microscope cell image.

<div align="center">(a)   (b)   (c)</div>
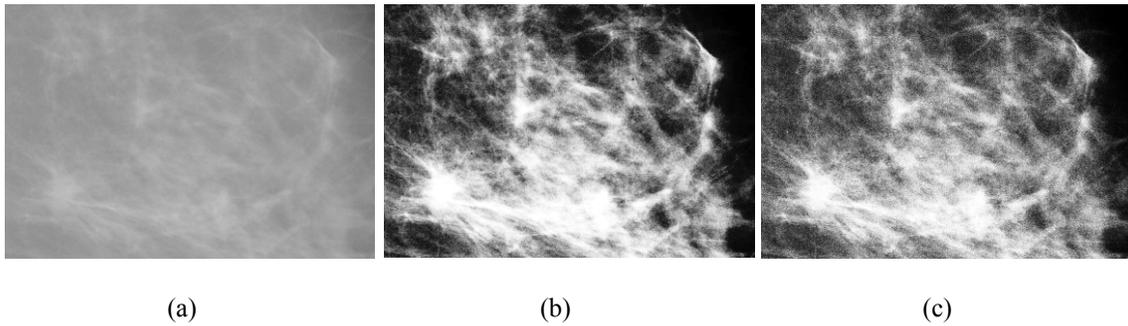
Figure 2.1: Enhancement results of the mammogram with spiculated lesions. (a) Original mammogram; (b) enhanced by standard histogram equalization; (c) enhanced by standard histogram equalization with GaSR ($\mu = 0$ and $\sigma^2 = 220$).

Figs. 2.2 (b), 2.3 (b), 2.4 (b) and 2.5 (b) show the results of image enhancement using CLAHE only. We can see that CLAHE performs better than the standard equalization method and achieves contrast improvement. However, some background normal tissues in the mammograms are also enhanced (Figs. 2.2 (b), 2.3 (b) and 2.4 (b)). As a result, the lesions are still not easily recognizable. For the cell image (Fig. 2.5 (b)), CLAHE only stands out the cores of the cells, but does not make the cell body more visible. On the other hand, SR noise-enhanced CLAHE increases the intensity of the lesions and at the same time decreases that of the normal tissues around lesions, such that the lesions are more obvious. Besides, the SR noise-enhanced CLAHE increases the visibility of the entire cell bodies.

(a)   (b)

(c)   (d)   (e)
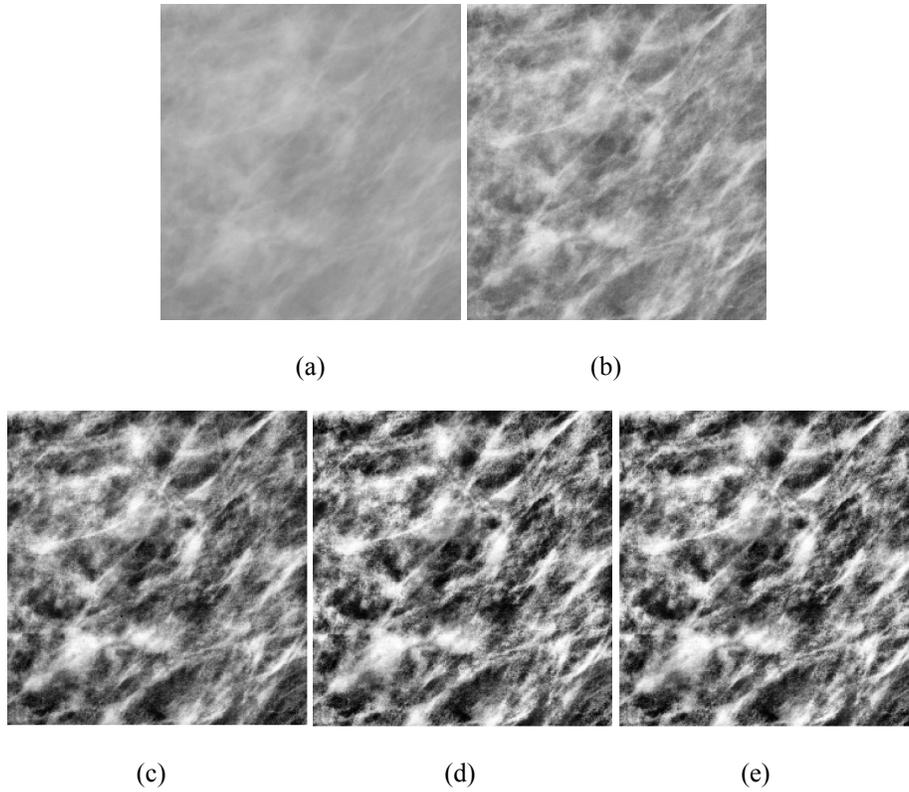
Figure 2.2: Enhancement results of the mammogram with masses. (a) Original mammogram; (b) enhanced by CLAHE; (c) enhanced by CLAHE with TwSR ($\alpha = 0.5$, $n_1 = 2.5$ and $n_2 = -2.5$); (d) enhanced by CLAHE with GaSR ($\mu = 0$ and $\sigma^2 = 15$); (e) enhanced by CLAHE with UnSR ($\mu = 0$ and $\sigma^2 = 25$).
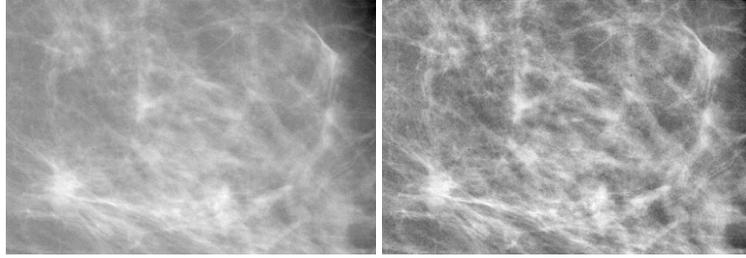
(a)   (b)
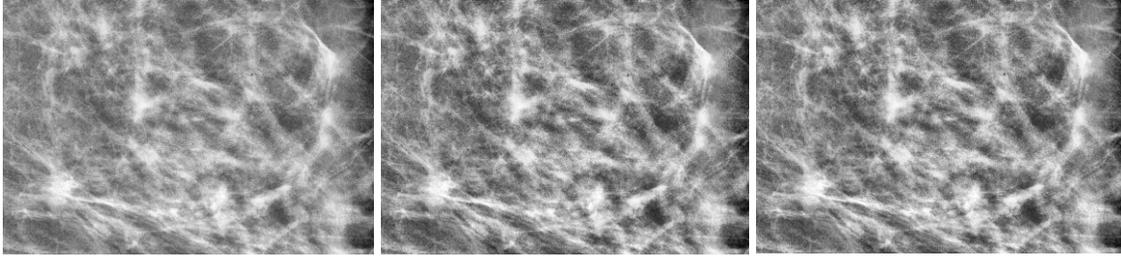
(c)   (d)   (e)

Figure 2.3: Enhancement results of the mammogram with spiculated lesions. (a) Original mammogram; (b) enhanced by CLAHE; (c) enhanced by CLAHE with TwSR ($\alpha = 0.5$, $n_1 = 5$ and $n_2 = -5$); (d) enhanced by CLAHE with GaSR ($\mu = 0$ and $\sigma^2 = 60$); (e) enhanced by CLAHE with UnSR ($\mu = 0$ and $\sigma^2 = 65$).

(a)   (b)

(c)   (d)
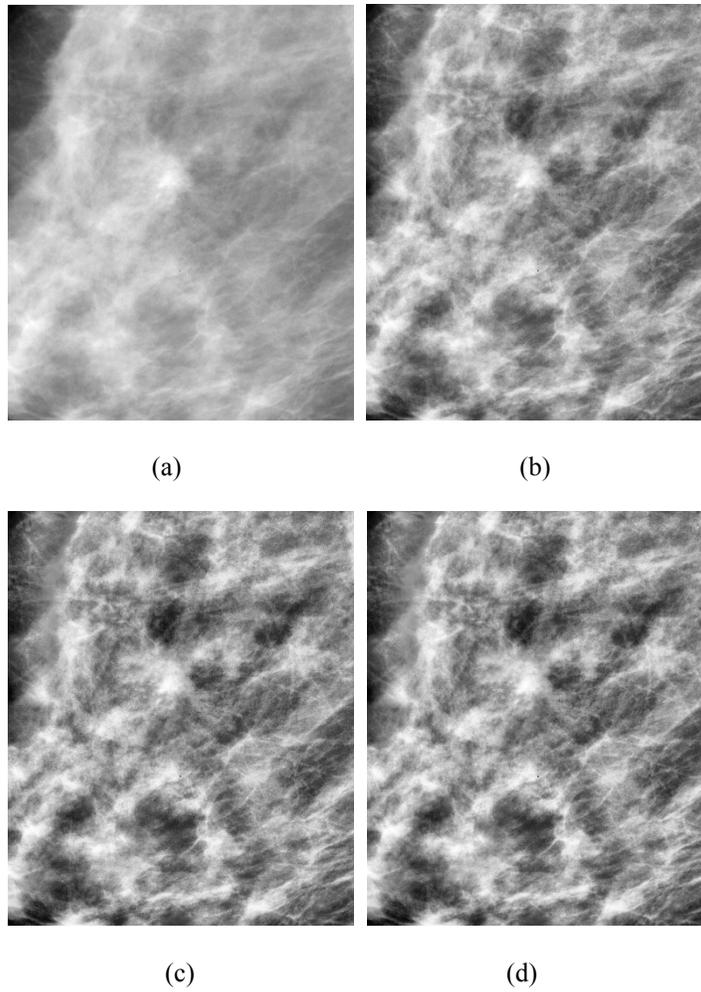
Figure 2.4: Enhancement results of the mammogram with spiculated lesions. (a) Original mammogram; (b) enhanced by CLAHE; (c) enhanced by CLAHE with GaSR ($\mu = 0$ and $\sigma^2 = 25$); (d) enhanced by CLAHE with UnSR ($\mu = 0$ and $\sigma^2 = 25$).
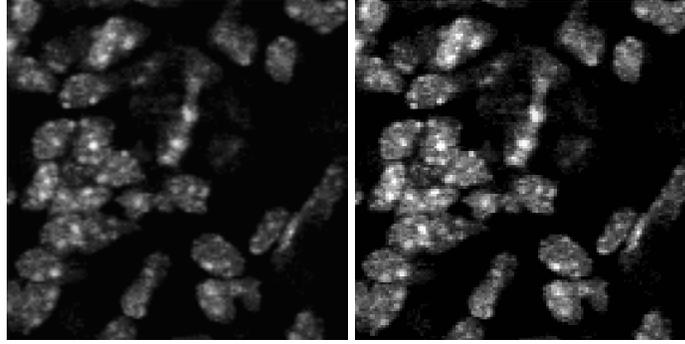
(a)   (b)
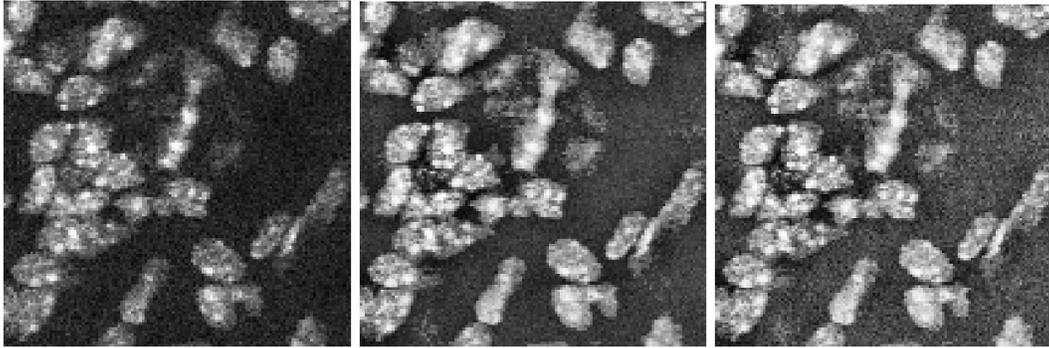
(c)   (d)   (e)

Figure 2.5: Enhancement results of the cell image. (a) Original cell image; (b) enhanced by CLAHE; (c) enhanced by CLAHE with TwSR ($\alpha = 0.5$, $n_1 = 30$ and $n_2 = -30$); (d) enhanced by CLAHE with GaSR ($\mu = 0$ and $\sigma^2 = 85$); (e) enhanced by CLAHE with UnSR ($\mu = 0$ and $\sigma^2 = 55$).

2.2.3 Dithering

In this section, we investigate using SR noise to improve the image dithering algorithm [18][61]. Fig. 2.6 (b) shows the dithering result without SR noise, and we can see many granularities in the image and the image is far from being a natural one. On the contrary, the SR noise-enhanced dithering yields a better result (Fig. 2.6 (c)) and the image becomes much smoother and more natural.

(a)   (b)   (c)
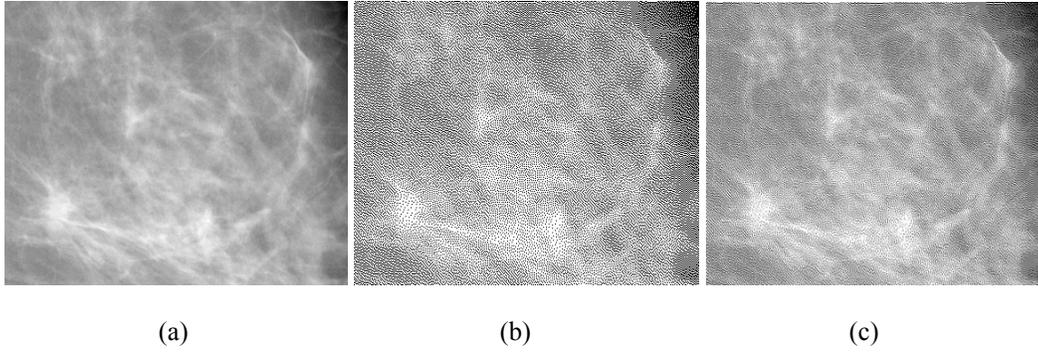
Figure 2.6: Dithering results of the mammogram with spiculated lesions. (a) Original mammogram with spiculated lesions; (b) dithering without SR noise; (c) dithering with UnSR ($\mu = 0$ and $\sigma^2 = 25$).

2.2.4 Edge Detection and Image Segmentation

Fig. 2.7 shows the results of SR noise-enhanced edge detection using Sobel detector [18][62]. We can see from Fig. 2.7 (b) that the original Sobel detector finds all the edges of the lesions but also generates many false positives. The SR noise-enhanced Sobel detectors (Figs. 2.7 (c) and (d)) keep the lesion edges while dramatically reduce the number of false detections.

(a)   (b)   (c)   (d)
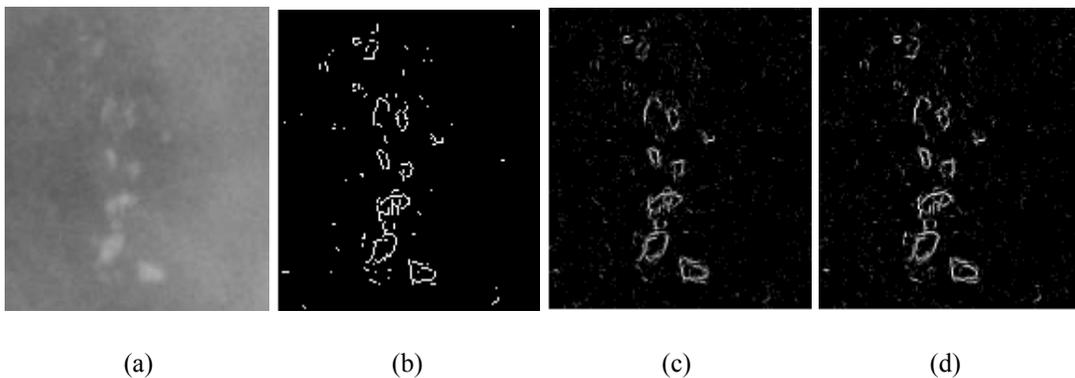
Figure 2.7: Edge detection results of the mammogram with micro-calcifications. (a) Original mammogram; (b) detected positives (detected "edges") by Sobel; (c) detected positives by Sobel with GaSR ($\mu = 0$ and $\sigma^2 = 45$); (d) detected positives by Sobel with UnSR ($\mu = 0$ and $\sigma^2 = 45$).

Fig. 2.8 shows the results of sequential image enhancement and edge detection, where we first employ CLAHE to enhance the original cell image and then use Sobel detector to locate the edges of the cells. Fig. 2.8 (b) shows the edge detection result using Sobel detector only, and Fig. 2.8 (c) shows the result of using CLAHE followed by Sobel. We can see from the two figures that the edge detection result is not improved. This is because the cores of the cells have much higher intensities than other part of the cell body, which drives the result of the intensity-based edge detector, such as Sobel. However, as shown in Fig. 2.5 (b), CLAHE only enhance the visibility of the cores rather than the cell body, so the detection results are still dominated by the cores, which yields the detected edges with incomplete cell contour and even false cell boundary. After adding SR noise to the enhancement and detection processes respectively, edge detection is improved a lot, and the cell location and shape can be recognized much easier (Figs. 2.8 (d) and (e)).

An example of SR noise-enhanced image segmentation using a level set-based algorithm can be found in Fig. 1.4 in Section 1.4.

2.3 Discussion

The goal of this chapter was to examine the feasibility of improving performance of image processing algorithms when noise is added prior to processing. By means of several examples, we observed that performance did improve when images are seen qualitatively. In the following chapters, we develop more systematic procedure for performance enhancement and evaluation.
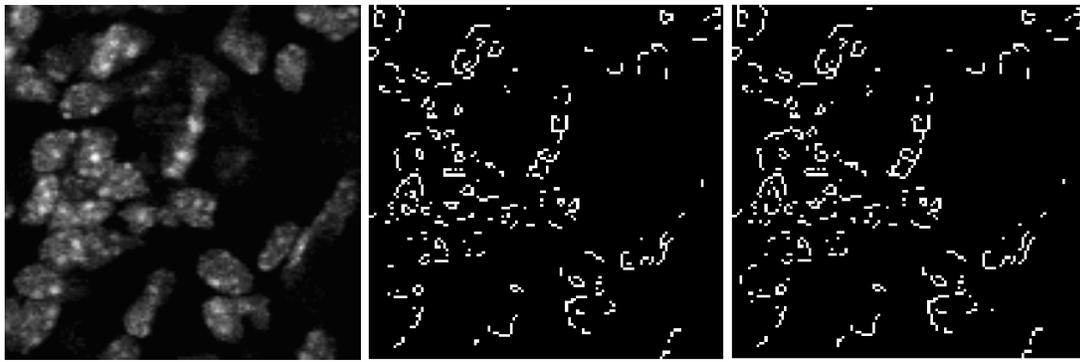
(a)   (b)   (c)
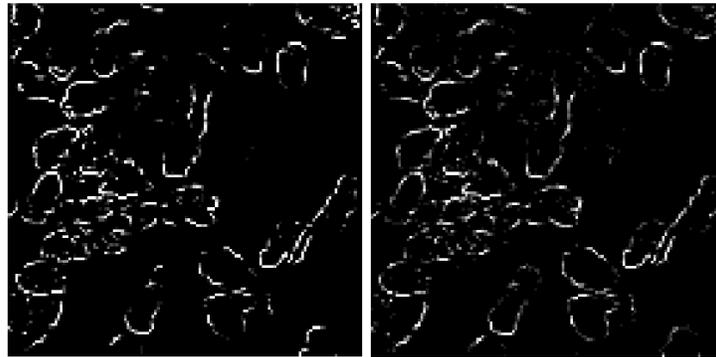
(d)   (e)

Figure 2.8: Cell image enhancement and edge detection. (a) Original cell image; (b) detected positives (detected "edges") by Sobel edge detection; (c) detected positives (detected "edges") by CLAHE enhancement and Sobel edge detection; (d) detected positives by CLAHE enhancement (GaSR: $\mu = 0$ and $\sigma^2 = 10$) and Sobel edge detection (GaSR: $\mu = 0$ and $\sigma^2 = 100$); (e) detected positives by CLAHE enhancement (GaSR: $\mu = 0$ and $\sigma^2 = 10$) and Sobel edge detection (UnSR: $\mu = 0$ and $\sigma^2 = 100$);

CHAPTER III

NOISE-ENHANCED DETECTION OF MICRO-CALCIFICATIONS IN DIGITAL MAMMOGRAMS

In this chapter, we investigate novel algorithms for the detection of micro-calcifications using stochastic resonance (SR) noise. In these algorithms, a suitable dose of noise is added to the abnormal mammograms such that the performance of a suboptimal lesion detector is improved without altering the detector's parameters. Unlike what has been discussed in Chapter 2, in this chapter we formulate the detection problem as an optimization problem and develop a systematic scheme to find the optimum parameters of SR noise and carry out the detection automatically.

3.1 Introduction

Breast cancer is a serious disease with high occurrence rate in women [63]. There is clear documented evidence which shows that early diagnosis and treatment of breast cancer can significantly increase the chance of survival for patients [64]. One of the important early symptoms of breast cancer in the mammograms is the appearance of micro-calcification clusters. An accurate detection of micro-calcifications is highly desirable to ensure early diagnosis of breast cancer.

Computer-aided diagnosis (CAD) improves the diagnostic performance of radiologists [65][66] and is an effective method for early diagnosis thereby increasing survival time for women with breast cancer. While advances have been made in the area of CAD for

digital mammograms, the main challenge of accurately identifying breast cancer in digital mammograms still remains, which is due to the small sizes and subtle contrast of the lesions compared with the surrounding normal breast tissues.

Much effort has been made for detecting micro-calcifications by using CAD techniques. Some methods tried to detect micro-calcifications through a modeling procedure. For example, Bazzani *et al.* [67] and Gurcan *et al.* [68] detected the micro-calcifications by using Gaussianity tests in the difference and filtered mammograms, respectively. Karssemeijer [69] modeled the mammograms using Markov random fields. Nakayama *et al.* [70] used a Gaussian probability density function (*pdf*) to model the abnormal regions in the subband mammograms generated by a filter bank. Regentova *et al.* [71] characterized the *pdf*s of the magnitudes of the wavelet coefficients, which are assumed to correspond to two hidden Markov states, as zero mean Gaussian distributions with different variances. Deepa and Tessamma [72] used a deterministic fractal model to characterize breast background tissues. The challenge for these model-based methods is that an accurate model is generally not easy to obtain and model mismatch is hard to avoid, so the detection results are deteriorated. There are also some methods that attempt to avoid the necessity of modeling during the detection process. For example, in [73], a relevance vector machine (RVM) was employed as a micro-calcification classifier, and its parameters were determined through a supervised learning procedure. Catanzariti *et al.* [74] trained a three-layer feed-forward artificial neural network (ANN) to detect micro-calcifications using the features extracted by a bank of Gabor filters. Strickland *et al.* [75], Lemaur *et al.* [76] and Li and Dong [77] proposed wavelet domain thresholding techniques to obtain the information of interest for the detection of micro-calcifications.

These methods partially bypassed the modeling problem, but determination of the optimum parameters, such as the threshold, is still a very challenging task, and the detection performance was often affected by the suboptimum parameters.

Basically, lesion detection can be considered as an anomaly detection problem [78]. Performance of the detectors is heavily dependent on the accuracy of the mathematical models and the detector parameters. However, as discussed before, appropriate models and optimum parameter values are generally very difficult to obtain in practical applications, which often results in unsatisfactory detection performance in terms of high probability of false alarm (P_F) and low probability of detection (P_D).

In this chapter, we investigate the use of SR noise to enhance the detection of micro-calcifications in mammograms. We first develop three lesion detectors based on the Gaussian assumption. We will see that they are all suboptimal detectors, suffering from model mismatch. Then, we present the main results of some prior work on SR noise-enhanced signal detection under the Neyman-Pearson criterion, where the optimum form of the SR noise is determined. After that, we develop a SR noise-based detection algorithm for lesion detection that attempts to improve the suboptimal detectors. An iterative detection scheme involving the use of SR noise with memory is also presented. A more general SR noise-based detection enhancement framework based on the iterative detection scheme is then presented. Finally, experimental results and the performance evaluation of several detection and classification algorithms are presented in Section 3.3.

3.2 Noise-Enhanced Detection of Micro-Calcifications in Digital Mammograms

3.2.1 Problem Statement and Gaussian Assumption-Based Lesion Detection

In this section, we first introduce the lesion detection problem from a statistical hypothesis testing point of view, and then present three Gaussian assumption-based detectors for the lesion detection task. We will show via experiments that the performance of these detectors is not satisfactory due to the large number of false alarms. This is due to model mismatch, and it results in suboptimality of the detectors.

Higher pixel intensity than the surrounding normal tissues distinguishes lesions from the normal structures in mammograms, which is one of the most important features of abnormal mammograms. The algorithms developed in this chapter perform the detection by exploring the pixel intensity information. We deal with this anomaly detection problem using statistical hypothesis testing methods. Formally, we want to choose one of the two hypotheses corresponding to the absence and presence of micro-calcifications on a pixel-by-pixel basis,

$$
\begin{aligned}
H_0 &: y[m] = w[m] \\
H_1 &: y[m] = s[m] + w[m]
\end{aligned}
\tag{3.1}
$$

where m is the pixel index corresponding to the pixel observation under consideration, $y[m]$ is the observed pixel intensity, larger than or equal to zero, $s[m]$ is the lesion signal, and $w[m]$ is the background noise that is assumed to obey Gaussian distribution with mean μ_b and variance σ_b^2. The noise is assumed to be additive, independently distributed and independent of the noise-free mammogram data. A processing window is employed

with the pixel under consideration being at the center of the window. This window is employed to estimate the parameters of the detector by using the pixels included in it.

3.2.1.1 *Gaussian Background Assumption-Based Detector (GBAD)*

The Gaussian background assumption leads to linear and tractable solutions [13]. The micro-calcifications, the signals of interest here, are brighter spots than the surrounding normal background tissues. So the micro-calcification is modeled as a signal with constant amplitude, and the lesion detection problem is to detect a constant signal in Gaussian noise, which we refer to as the Gaussian background assumption-based detector (GBAD). The lesion signal $s[m]$ in (3.1) is, therefore, a constant intensity.

As mentioned in Section 1.1, for this anomaly detection problem, the *a priori* probabilities of the background and lesion pixels are unavailable, so we employ the Neyman-Pearson criterion [13] for the detection task.

Under the Gaussian background and constant signal assumptions, $p(\bar{y}; H_1)$ and $p(\bar{y}; H_0)$ all obey Gaussian distribution with the same variance σ_b^2, and the optimal test given in (3.1) can be expressed in terms of the GBAD test statistic $T_{GBAD}(y)$ as follows

$$T_{GBAD}(y) = y - \mu_b \underset{\underset{decide \ H_0}{<}}{\overset{\overset{decide \ H_1}{>}}{}} \gamma_1 \qquad (3.2)$$

where y is the intensity of the pixel under consideration, and the threshold γ_1 is determined from the desired P_F and the statistical parameters, i.e., mean and variance, of the pixels in the processing window.

To estimate the detector's parameters, an initial detection is first carried out in the

processing window to perform a coarse detection, and the resulting detected negatives (H_0) and positives (H_1) are employed to estimate the parameters. Many methods, such as a local maxima filter [79] or adaptive thresholding techniques [80], can perform the initial detection. Local maxima filter is employed in this work because the lesion pixels generally have a higher intensity than the surrounding normal background tissues.

For cancer diagnosis, the most serious mistake is to miss any lesions. To reduce the probability of miss, we use a "safer" initial detection and attempt to exclude all the lesion pixels from the background. It can be realized by using a local maxima filter with appropriate window size and local threshold, permitting more pixels having relatively higher intensities in the local regions to be classified into the lesion part.

The detection results for GBAD are shown in Figs. 3.2 (b), 3.3 (b) and 3.4 (b) in Section 3.3, corresponding to three types of mammograms with micro-calcifications. From the figures, we can see that the micro-calcifications are completely detected but with a higher P_F than the desired value, 0.01, used in our work. At this point, it suffices to say that the performance of the detector is not satisfactory. A detailed discussion of the experimental results is postponed to Section 3.3.

3.2.1.2   *General Gaussian Detector (GGD)*

Micro-calcifications, especially micro-calcification clusters, have a small size but generally do not have a constant intensity, so a Gaussian model as opposed to a constant signal model is proposed in this section to be a more reasonable model to represent the signal part. Thus, the problem can be considered to be the problem of detection of Gaussian signals in Gaussian noise. We refer to this detector as the general Gaussian

detector (GGD) [13]. The lesion signal $s[m]$ in (3.1) under this assumption obeys Gaussian distribution, i.e., $s[m] \sim N(\mu_s, \sigma_s^2)$.

The detected positive pixels (corresponding to lesions) and negative pixels (corresponding to the background) in the initial detection are employed to coarsely estimate the means and variances of the lesion and background pixel intensity *pdf*s. Under the GGD assumption, $p(\bar{y}; H_1)$ and $p(\bar{y}; H_0)$ obey Gaussian distribution but with different variances, and the optimal test is still the likelihood ratio test given in (1.5). The optimum test can be expressed in terms of the GGD test statistic $T_{GGD}(y)$ as follows

$$T_{GGD}(y) = \frac{\sigma_s^2}{\sigma_b^2}(y - \mu_b)^2 + 2\mu_s(y - \mu_b) \underset{\substack{< \\ decide\ H_0}}{\overset{\substack{decide\ H_1 \\ >}}{}} \gamma_2 \tag{3.3}$$

where the threshold γ_2 is determined from the desired P_F. The statistical parameters, namely the means and the variances, of the initially detected positive and negative pixels can be estimated using the processing window with the pixel under consideration at the center of the window.

We can see that when $\sigma_s^2 \to 0$, the first term on the right side of (3.3) tends to zero, and (3.3) reduces to a form similar to (3.2), which corresponds to $s[m]$ being a constant signal. We also notice that (3.3) is a detector with two thresholds because the test statistic is quadratic. Due to the nature of the abnormal mammograms, i.e., lesion pixels have intensities that are generally higher than the surrounding normal background tissues and the probability of the intensities of lesion pixels falling below the lower threshold is extremely small, thus, only the higher threshold is employed to classify the mammogram pixels into background and lesions. Therefore, the higher threshold of the test in (1.5) is used in our work for the detection task.

The detection results for GGD are shown in Figs. 3.2 (c), 3.3 (c) and 3.4 (c) in Section 3.3, where all the micro-calcifications are discovered by GGD, with less false positives compared with the GBAD.

3.2.1.3   GGD-Based Iterative Detector (GGD_ID)

Encouraged by the improvement achieved by the GGD over GBAD, we propose an iterative method to further improve the performance of GGD by increasingly improving the estimation of statistical parameters in an iterative manner.

At each step of the iteration, the GGD is designed with the parameters, μ_s, μ_b, σ_s^2 and σ_b^2, corresponding to the background and micro-calcifications, estimated from the detection result in the preceding iteration as opposed to keeping them fixed during all iterations, which results in different thresholds at each iteration.

The procedure of the iterative detection algorithm is described as follows:

Initialization: Initial detection using the coarse detector described in Section 3.2.1.1.

Step 1: Means and variances of the detected positive (lesion) and negative (background) pixels are calculated.

Step 2: Detection is performed using the GGD (3.3) with the desired P_F and the updated parameters, μ_s, μ_b, σ_s^2 and σ_b^2, calculated in Step 1. If there are no differences in the detected positives and negatives between two successive detections, terminate the algorithm, else go to Step 1.

The presented GGD_ID algorithm is similar in spirit to Gaussian assumption-based dynamic clustering (GADC), in which both background and lesions are assumed to obey Gaussian distributions, and the detection (or clustering) and parameter updating are

performed in an iterative manner [81]. The difference is that the method presented here incorporates an additional constraint in terms of the desired value of P_F. The reason we include P_F in the algorithm is that at each step of the iteration, some detected negative pixels have intensities much larger than the mean of the detected background pixels and are close to that of the detected lesion pixels. In other words, some pixels have a non-negligible and, in fact, fairly high probability to belong to the lesion part. Since we do not want to miss any lesions, these pixels are classified into the lesion part by the desired P_F value, such as the value 0.01 used in this work. We will observe in Section 3.3 that the iterative detection method presented here performs better than the GADC.

The detection results for this detector are shown in Figs. 3.2 (d), 3.3 (d) and 3.4 (d) in Section 3.3, where the GGD is employed iteratively four times on the mammograms. Experiments show that the method generally converges within 5 iterations.

3.2.1.4   Model Mismatch Analysis

From the experimental results, we can see that the detection performance has improved with the melioration of the detection schemes, but the final results are still not satisfactory as seen via inexact lesion contours and large number of false positives. The resulting diagnosis may result in additional testing and biopsies for spots on mammograms that finally turn out to be harmless, which is a weakness of many CAD systems exhibit currently [82].

One major reason for the unsatisfactory detection is that the Gaussian assumption does not accurately model the background distribution and the resulting test including the detection threshold is not optimal. A more accurate model for the background, heavy-

tailed symmetric α stable ($S\alpha S$) distribution, was proposed in [83]. For verification, we draw the amplitude probability distribution (APD) [84] plots of real-world mammogram background data of a mammogram from the MIAS Mini-mammographic Database, simulated Gaussian distribution and heavy-tailed $S\alpha S$ distribution data on a log-log scale (see Fig. 3.1). Plotting APD is a commonly used method to test impulsive noise. It is defined as the probability that the noise amplitude is above some threshold. We can see from Fig. 3.1 that for small amplitudes, the simulated heavy-tailed distribution and Gaussian distribution provide good fits to the mammogram data. At larger amplitudes (i.e., at the tails), the simulated heavy-tailed $S\alpha S$ distribution is shown to be a better fit than the Gaussian one. In addition, the plots of the mammogram data and the simulated heavy-tailed $S\alpha S$ data decay linearly with a constant slope compared with that of the Gaussian data. These two observations indicate that the heavy-tailed $S\alpha S$ distribution [83][84] is a better model than the Gaussian model for the background pixel intensities of a digital mammogram. Hence, there exists empirical support for the existence of the $S\alpha S$ noise distribution in mammogram background (as opposed to the Gaussian distribution). Theoretical analysis and more detailed discussion on this can be found in [83].
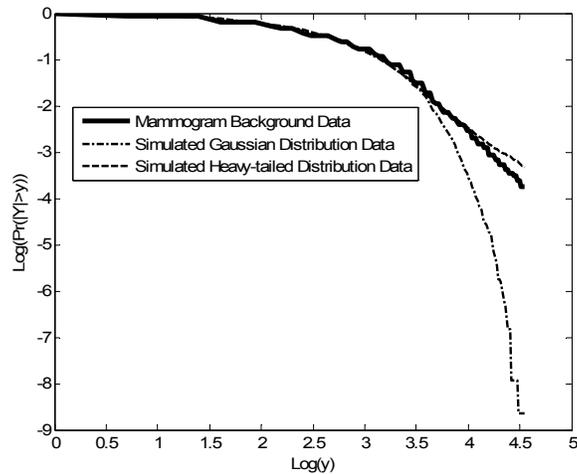
Figure 3.1: The APD plots of the real-world mammogram background data, simulated Gaussian distribution and heavy tailed $S\alpha S$ distribution data on a log-log scale. It shows that the mammogram pixel intensities obey heavy- tailed $S\alpha S$ distribution more closely

One approach to the design of the optimal lesion detector is to derive the optimal test under the Neyman-Pearson formulation when the background is modeled as the $S\alpha S$ distribution. However, the difficulties in learning the parameters of the $S\alpha S$ distribution from the real-world data as well as the off-line integration when calculating the detection threshold constrains the practical application of the optimal $S\alpha S$ -based detectors. In the following sections, we will investigate an alternate approach, namely the application of SR noise, to the lesion detection problem. We will continue to use the suboptimal detectors designed based on the Gaussian noise background assumption. Admittedly there is a model mismatch, and we will attempt to overcome the deterioration in the detector performance by adding SR noise at the input to the detector. We will see that the SR noise-based detector yields significant performance enhancement and is easy to implement.

3.2.2 Optimum SR Noise-Enhanced Signal Detection

One of the main goals of this chapter is to develop SR noise-enhanced detection methods for lesion detection in mammograms. We first discuss in this section how to find the optimum SR noise.

Based on the binary statistical decision formulation (1.1)~(1.6) in Section 1.1, we popose to add an appropriate noise $\bar{n}$ to the original data $\bar{y}$, which yields a new data vector $\bar{z}$

$$\bar{z} = \bar{y} + \bar{n} \tag{3.4}$$

where $\bar{n}$ is either a random vector with *pdf* $p_{\bar{n}}(.)$ or a nonrandom signal.

The binary hypotheses testing problem for this new observed data can be expressed as

$$\begin{cases} H_0 : p_{\bar{z}}(\bar{z};H_0) = p_0(\bar{z}) = \int\limits_{R^N} p_0(\bar{y}) p_{\bar{n}}(\bar{z} - \bar{y}) d\bar{y} \\ H_1 : p_{\bar{z}}(\bar{z};H_1) = p_1(\bar{z}) = \int\limits_{R^N} p_1(\bar{y}) p_{\bar{n}}(\bar{z} - \bar{y}) d\bar{y} \end{cases} \tag{3.5}$$

In our work, we consider the SR noise-enhanced fixed detectors whose parameters, such as the thresholds, are unchanged before and after adding the SR noise, so the critical function ϕ of $\bar{z}$ is the same as that of $\bar{y}$. Therefore,

$$\begin{aligned} P_D^{\bar{z}} &= \int\limits_{R^N} \phi(\bar{z}) p_1(\bar{z}) d\bar{z} = \int\limits_{R^N} p_{\bar{n}}(\bar{y}) \left(\int\limits_{R^N} \phi(\bar{z}) p_1(\bar{z} - \bar{y}) d\bar{z} \right) d\bar{y} \\ &= \int\limits_{R^N} F_1(\bar{y}) p_{\bar{n}}(\bar{y}) d\bar{y} \end{aligned} \tag{3.6}$$

And similarly we have

$$P_F^{\bar{z}} = \int\limits_{R^N} \phi(\bar{z}) p_0(\bar{z}) d\bar{z} = \int\limits_{R^N} F_0(\bar{y}) p_{\bar{n}}(\bar{y}) d\bar{y} \tag{3.7}$$

where $F_i(\bar{y}) = \int\limits_{R^N} \phi(\bar{z}) p_i(\bar{z} - \bar{y}) d\bar{y}$, corresponds to hypothesis H_i.

Thus, our goal is to find the optimum *pdf* of the SR noise which is the solution to the

following optimization problem.

$$\text{Maximize } P_D^{\bar{z}}$$

subject to (3.8)

$$P_F^{\bar{z}} \leq P_F^{\bar{y}} \quad p_{\bar{n}}(\bar{y}) \geq 0 \text{ and } \int_{R^N} p_{\bar{n}}(\bar{y})d\bar{y} = 1$$

In this section, we discuss two methods to find the optimum solution to (3.8).

3.2.2.1 Linear Programming

The first method is based on linear programming.

From (3.8) we observe that the objective function $P_D^{\bar{z}}$ and the constraints are all linear functions of the *pdf* $p_{\bar{n}}(.)$. Therefore, we could use linear programming methods to find the optimum $p_{\bar{n}}(.)$. Simplex algorithm [85] has been employed in our experiments for the optimization task and we achieved very satisfactory results.

3.2.2.2 Convex Hull Theorem

The second optimization method is based on the Convex Hull Theorem, which is discussed in detail in [53]. We present the fundamental results on SR noise-enhanced signal detection using the Convex Hull Theorem [53] in this section.

The sufficient condition for improvability of detection via SR noise is given in Theorem 1 [53].

Theorem 1: If $J(P_F^{\bar{y}}) > P_D^{\bar{y}}$ or $J''(P_F^{\bar{y}}) > 0$ when $J(t)$ is second-order continuously differentiable around $P_F^{\bar{y}}$, then there exists at least one noise process $\bar{n}$ with *pdf* $p_{\bar{n}}(.)$ that can improve the detection performance, where $J(t)$ is defined as the maximum value of

f_1 given f_0, i.e., $J(t) = \sup(f_1 : f_0 = t)$. f_0 and f_1 are the given values of F_0 and F_1, respectively.

Theorem 2 [53] determines the form of the optimum SR noise when the detector is improvable.

Theorem 2: To maximize $P_D^{\bar{z}}$, under the constraint that $P_F^{\bar{z}} \leq P_F^{\bar{y}}$, the optimum noise can be expressed as $p_{\bar{n}}^{opt}(\bar{n}) = \lambda\delta(\bar{n} - \bar{n}_1) + (1-\lambda)\delta(\bar{n} - \bar{n}_2)$, where λ and $1-\lambda$ are the occurrence probabilities of the suitable *N*-dimensional vectors $\bar{n}_1$ and $\bar{n}_2$, $0 \leq \lambda \leq 1$.

The approach to determine λ, $\bar{n}_1$ and $\bar{n}_2$ is discussed in detail in [53]. They can be determined in practice using numerical methods. Since the optimum SR noise is a randomization of two deterministic vectors, we call it the "Two-peak SR noise" in this chapter.

The advantage of a SR noise-enhanced fixed detector is that the parameters, such as the threshold, of the original detector do not need to be changed, yet better detection performance is expected. In other words, model mismatch can be handled fairly easily by using this approach. However, to obtain the optimum SR noise, full knowledge of the *pdf*s under the two hypotheses is required, which in real-world applications is generally not available. In the next two sections, we will discuss how to find the suitable SR noise for enhancing a suboptimal lesion detector when the knowledge of the *pdf*s is incomplete.

3.2.3 SR Noise-Enhanced Gaussian Assumption-Based Detection

In this section, the SR noise-enhanced detection approach is employed for finding lesions and enhancing the previously discussed suboptimal detectors based on the Gaussian assumption. We perform pixel-by-pixel detection. The suboptimal detectors to be

improved result from the model mismatch and the lack of information about the mammogram statistics. These detectors are excellent candidates for the application of the SR noise-enhanced detection scheme.

The basic idea of the SR noise-enhanced detection is to obtain the optimum additive SR noise based on the knowledge of the *pdf*s of the lesion and the background signals. Since these *pdf*s are not known, they need to be estimated from the given mammogram itself. The mammogram is modified with the optimum additive SR noise determined using the estimated *pdf,* and then the original suboptimal detector performs the detection. Two SR noise-based schemes are presented here for improving lesion detection.

3.2.3.1   *Two-Peak SR Noise-Enhanced Gaussian Background Assumption-Based Detection (2SR-GBAD)*

In this algorithm, we try to reduce the dependence of the SR noise determination on the knowledge of the true *pdf*s and increasingly enhance the suboptimal detectors through an iterative procedure.

We first use the SR noise to enhance the GBAD discussed in Section 3.2.1.1.To achieve this goal, we perform the coarse detection of the lesion and background using the local maxima filter mentioned in Section 3.2.1.1. The detection threshold is calculated for the GBAD, which is suboptimum due to model mismatch. Then, the probability densities under H_1 and H_0 are obtained using the kernel density estimation method [86] based on the detected positives and negatives. The parameters of the SR noise are calculated from the suboptimum threshold and the estimated densities. The SR noise is added to the original mammogram. Detection is performed on the SR noise-modified data using the

original detector. This procedure is repeated in an iterative manner until the difference between two successive detection results is very small[2]. The procedure of the 2SR-GBAD detection algorithm is described as follows:

Initialization: Initial detection using the coarse detector described in Section 3.2.1.1.

Step 1: Mean μ_b and variance σ_b^2 of the background are estimated based on the detected negative pixels. The detection threshold is updated based on the desired P_F as well as μ_b and σ_b^2 using (1.6), where we assume that the background obeys Gaussian distribution (see GBAD in Section 3.2.1.1).

Step 2: The pixels are detected with the updated threshold found in Step 1. The resulting detected positive and negative pixels are employed for estimating probability densities under the two hypotheses using the kernel density estimation method.

Step 3: The updated threshold in Step 1 and the newly estimated probability densities in Step 2 are used to determine the SR noise with the method mentioned in Section 3.2.2.

Step 4: The mammogram data is modified by adding to the original pixel intensities the SR noise determined in Step 3.

Step 5: Detection is performed with the detector updated in Step 1 using the modified data from Step 4. If the difference between two successive detection results is very small, terminate the algorithm else go to Step 1.

According to the experiments, a good initialization can be generated by schemes such as a maxima filter with an appropriate window size and threshold, such that satisfactory detection can still be obtained even when the threshold update procedure in Step 1 is not

---

[2] There could be many methods to define and evaluate the difference. In our work, the difference is defined as the ratio of the   number of differently labeled pixels in two successive detections to the total number of pixels in the mammogram. The labeled pixel here means a pixel classified as a positive (lesion) pixel or a negative (background) pixel. The iterative procedure is terminated when the ratio is smaller than a desired value.

performed during the iterations. In other words, the critical function can remain fixed throughout the iterations if the initial detection is good enough. We can also perform threshold updating every several iterations to improve the computation speed.

In a similar manner, we can use the above procedure to design the SR noise-enhanced GGD test, i.e., 2SR-GGD, where the means, μ_s and μ_b, and variances, σ_s^2 and σ_b^2, of the detected positives and negatives as well as the desired P_F are used together to update the threshold in Step 1. The rest of the four steps of 2SR-GGD are the same as those of 2SR-GBAD. Since GGD is a more accurate model for abnormal mammograms, which can be seen in the comparison between the detection results of GBAD and GGD, 2SR-GGD yields better performance than 2SR-GBAD, according to the experiments. The 2SR-GGD method also shows improvement over GGD detection. Moreover, the presented algorithm generally needs fewer iterations than GGD_ID discussed in Section 3.2.1.3 to reach similar detection results. Also, the final results of the presented algorithm are better than GGD_ID.

3.2.3.2 *Two-Peak SR Noise-Enhanced Gaussian Assumption-Based Detection with Memory (2SR-GBAD-M)*

The experiments show the improved performance of 2SR-GBAD. In this section, we further improve its efficiency and robustness by introducing memory in the detection enhancement scheme.

As we know, to find the optimum SR noise, the exact knowledge of the probability distribution under the two hypotheses and the determination of the solution for a set of equations are required. However, in real-world applications, due to incomplete

information about the distribution, limitations on the accuracy when solving the equations and various contents of mammograms, high efficiency and robustness of the SR noise-enhanced detection system may not be achievable using the SR noise-based enhancement procedure only once. Multiple applications of the procedure may yield further enhancement of detection performance. We, therefore, apply suitably arranged multiple two-peak SR noises multiple times to increase the efficiency and robustness of the detection system, which we refer to as 2SR-GBAD-M.

Formally, for the SR noise-based scheme with memory, we have

$$\bar{z} = \bar{y} + \bar{n}^* \tag{3.9}$$

where $\bar{n}^*$ represents multiple-peak SR noises instead of a single two-peak SR noise added to the original mammogram data in Step 4 of the algorithm presented in Section 3.2.3.1, and

$$p_{\bar{n}^*} = \sum_{k=1}^{r} w_k p_{\bar{n}_k} \tag{3.10}$$

where w_k is the weight or probability of occurrence of the kth two-peak SR noise, $0 \le w_k \le 1$ and $\sum_{k=1}^{r} w_k = 1$. r is the number of two-peak SR noises which in our current work equals the number of iterations already run plus 1 (i.e., the SR noise determined from the estimated probability mass function (*pmf*) and the updated threshold at current iteration is also included, where *pmf* is used as the specific form of the probability distribution for discrete digital mammogram data), and

$$p_{\bar{n}_k}(\bar{n}) = \lambda_k \delta(\bar{n} - \bar{n}_{1k}) + (1 - \lambda_k)\delta(\bar{n} - \bar{n}_{2k}) \tag{3.11}$$

Of course, we can change the memory size by using different values of r, but in any case the latest r two-peak SR noises should be employed. When $r=1$, a single two-peak noise

is used, and the scheme reduces to the scheme without memory.

At each step of iteration, a larger weight, i.e., higher probability is allocated to the SR noise calculated from the currently estimated *pmf*s, and the weights for the rest of the SR noises are inversely proportional to the distances between their corresponding *pmf*s and the currently estimated ones. The distance D between the *pmf*s obtained during the lth iteration and the latest estimated *pmf*s is defined as

$$D_l = \sum_{i=0}^{B_n} [\,|\, PMF_{lH_0}(i) - PMF_{eH_0}(i)\,| + |\, PMF_{lH_1}(i) - PMF_{eH_1}(i)\,|\,] \tag{3.12}$$

where PMF_{lH_j} denotes the *pmf* under hypothesis H_j obtained during the lth iteration, and $j=0,1$. B_n is related to the resolution of the image data. For example, $B_n = 255$ if an 8-bit image is used. PMF_{eH_j} is the estimated *PMF* under hypothesis H_j obtained at the current iteration. The summation is over all possible image intensity values. This approach to incorporate memory has resulted in encouraging results as will be seen in the Section 3.3. The detection results of the two-peak SR noise enhanced GBAD tests with memory are shown in Figs. 3.2 (f), 3.3 (f) and 3.4 (f), from which we can see the Gaussian assumption-based detection suffering from model mismatch is improved through the addition of SR noise. Experiments also show that 2SR-GGD-M yields better performance than 2SR-GBAD-M.

3.2.4 SR Noise-Based Detection Enhancement Framework

We have presented a SR noise-based detection enhancement method in Section 3.2.3 to reduce the model mismatch resulting from the Gaussian assumption. When models other than Gaussian models are used to fit data, there may still exist model mismatches,

resulting in detector performance degradation, and SR noise may enhance the detector performance. In this section, we extend the SR noise-based detection scheme and present a more general SR noise-based detection enhancement framework. This framework provides much more flexibility and higher efficiency. In this chapter, we only consider the detectors (or classifiers) which we have control of, e.g., we can change their thresholds.

The framework is developed by modifying the first two steps of the detection procedure presented in Section 3.2.3 and is shown as follows.

Initialization: Initial detection.

Step 1: Probability density estimates are obtained under the two hypotheses using the detected positive (lesion) and negative (background) pixels. The detection threshold (or the classifier) is updated according to the estimated probability density information.

Step 2: The pixels are classified (or detected) with the updated threshold or the classifier in Step 1. The resulting detected positive and negative pixels are employed for estimating probability densities under the two hypotheses.

Step 3: The updated threshold or classifier in Step 1 and the newly estimated probability densities in Step 2 are used to determine the SR noise with the method mentioned in Section 3.2.2.

Step 4: The mammogram data is modified by adding SR noise to the original pixel intensities.

Step 5: Detection is performed with the detector or classifier updated in Step 1 using the modified data from Step 4. If the difference between two successive detection results is very small, terminate the algorithm else go to Step 1.

To improve the efficiency and robustness of the detection framework, the two-peak SR noise scheme with memory, which yields multi-peak SR noise, can also be used in Step 4.

We note that no specific constraints are put on the initialization, threshold or classifier updating and *pdf* estimation methods used in this framework. Any reasonable approaches could be employed. In the current work, we illustrate the ability of our framework by considering different algorithms for threshold or classifier updating and *pdf* estimation. For initialization, we still use the maxima filter discussed in Section 3.2.1.1. A. For threshold or classifier updating, one may use the methods that can converge when there is no SR noise added, such as GADC and iterative mode separation (IMS) algorithms [81]. IMS is an unsupervised learning pattern classification approach, which employs kernel density estimation technique to determine the *pdf* and performs clustering in an iterative manner. For *pdf* estimation, one may use non-parametric methods, such as kernel density estimation, *k*-nearest neighbor density estimation [86] and Bootstrap methods [87][88], etc., because we want to reduce the model mismatch during the *pdf* estimation as well as the dependence of the framework on modeling, and to make the framework more generally usable. In this chapter, for performance comparison, we employ the kernel density estimation approach and threshold update using (3.13) [89], same as those used in IMS. We will observe in Section 3.3 that the SR noise-based method can further improve the performance of IMS. The threshold updating is carried out by using

$$P_0 p_0(y^*) = P_1 p_1(y^*) \tag{3.13}$$

where y^* is the updated detection threshold during the current iteration. P_0 and P_1 are the *a priori* probabilities of the detected negatives and positives, which can be estimated by

$\hat{P}_i = n_i / n$, where n_i is the number of negatively detected pixels when $i=0$ and positively detected pixels when $i=1$, and n is the total number of pixels in the mammogram. This generates a suboptimal detector because the threshold is determined from the coarsely estimated *a priori* probabilities and *pdf*s by using the plug-in rule [81].

Experimental results show that the SR noise-based algorithm presented in this chapter generally needs fewer number of iterations than IMS to reach similar detection results. Also, the final results of the SR noise-based algorithm are better than IMS, where the final results are attained when the difference between two successive detection results is very small. In addition, given a good initialization, satisfactory detection can still be obtained even when the threshold or classifier update procedure in Step 1 is not performed during the iterations.

It can be seen that the above iterative procedure includes a scheme for *pdf* estimation, but in our current detection (or clustering) application, the estimated *pdf*s are not of interest as an end in themselves. Instead, we are more interested in the detection results which, of course, depend on the estimate. At the same time, an accurate *pdf* estimate can also be obtained from an accurate detection. So, the detection results are used in this chapter as an alternative way to evaluate the performance of the *pdf* estimation algorithm. For comparison, a Gaussian mixture modeling (GMM) [90]-based clustering method which performs the detection based on the GMM-fitted *pdf* is employed with the detection results shown in Section 3.3.

3.3   *Experimental Results*

3.3.1   Experimental Data

The majority of the mammograms used in the experiments are from Digital Database for Screening Mammography (DDSM) [91], and a few of them are from the Mammographic Image Analysis Society (MIAS) Mini-mammographic Database [92], so the experimental parameters, such as the processing window size, are determined mainly based on DDSM. DDSM has 2620 cases available in 43 volumes. A case consists of between 6 and 10 mammograms, where the grey levels are quantized to 16 bits, and resolution of the images is 50 microns/pixel. The MIAS Database includes 330 mammograms with the resolution being 200 microns/pixel. The location and types (malignant or benign) of the mammogram lesions are identified by expert radiologists and used as the ground truth in our work. In this chapter, our emphasis is on location detection based on the ground truth.

We choose three types of representative abnormal mammograms with micro-calcifications (clusters) including one having homogeneous background with a small number of isolated micro-calcifications, one having homogeneous background with a large number of micro-calcifications (clusters) and one having inhomogeneous background with a moderate number of micro-calcifications (clusters), respectively. These three types of mammograms cover a broad spectrum of mammogram micro-calcification (cluster) cases. Seventy five images selected from the three types of mammograms, 25 for each type, are employed to test the algorithms.

Micro-calcifications are very small, their sizes are in the range of 0.05–1.00mm [93], and the average is 0.3mm. Those smaller than 0.1 mm cannot be easily distinguished in the film-screen mammography from the high-frequency noise [94]. The width of the

majority of the micro-calcifications in our study is in the range between 0.25 and 0.5 mm. A micro-calcification cluster is considered to be a group of 3 to 5 or more micro-calcifications, 5 mm apart [95]. We choose the processing window size of 49 by 49 which is based on experiments that we conducted as well as the characteristics and the size of the lesions. Our experiments also indicated that the detection results were not very sensitive to the choice of window size provided that the window size was in the range between 31 and 61 when processing the data.

Since we carry out pixel by pixel detection, any isolated detected positive should not be considered to be a lesion due to the micro-calcification size mentioned above as well as the fact that the high-frequency noise may have serious influence on an individual pixel. Therefore, a micro-calcification (cluster) is declared to be detected only if at least 4 by 4 positively detected pixels are in a clump.

3.3.2 Performance Comparison and Analysis

In this section, performance of several lesion detection algorithms is compared and analyzed. These algorithms include GBAD, GGD and GGD_ID discussed in Section 3.2.1, GADC, 2SR-GBAD-M, IMS, GMM-based clustering method, high order statistics method based on local maxima detection and adaptive wavelet transform (HOSLW) [96] and the SR noise-based detection enhancement framework using a procedure similar to IMS, i.e., SR_IMS.

The first four algorithms are based on the Gaussian distribution assumption and are parametric approaches. GMM is a semi-parametric technique for *pdf* estimation, in which the superposition of a number of parametric densities, e.g., Gaussian distribution, are

used to approximate the underlying *pdf*. It offers a useful compromise between the non-parametric methods mentioned in Section 3.2.4, and the parametric estimation methods, such as those mentioned above. For the clustering application, we first fit the GMM given in (3.14) by using the Expectation-maximization algorithm [90]

$$f(y) = \sum_{i=1}^{g} P_i f_i(y) \tag{3.14}$$

where $f(y)$ is the density of the observation y, and $f_i(y)$ are the component densities of the mixture. g is the number of components, which can be preset or automatically determined according to the data statistics. In this chapter, we set g=2 to facilitate two-class clustering. P_i are the mixing proportions or weights, $0 \leq P_i \leq 1$ (i=1,…,g) and $\sum_{i=1}^{g} P_i = 1$.

Clustering is performed by using the plug-in rule given in (3.15) based on the Bayes rule [90]

$$R(y) = i \text{ if } \hat{P}_i \hat{f}_i(y) \geq \hat{P}_j \hat{f}_j(y) \tag{3.15}$$

for j=1,…,g, where $R(y) = i$ denotes that the allocation rule $R(y)$ assigns the observation y to the ith component of the mixture model. $\hat{P}_i$ and $\hat{f}_i(y)$ are the fitted values of P_i and $f_i(y)$, respectively.

The HOSLW algorithm is proved to have superior performance compared with other existing methods [97] in terms of efficiency and reliability. In this method, local maxima of the mammogram are determined as the lesion candidates, and the adaptive wavelet transform is employed to generate subbands which permit the rank of these maxima in the subband mammogram using a higher order statistical test for lesion detection.

For fairness, we use the same initial detection for the algorithms compared in the experiments. In 2SR-GBAD-M and SR_IMS, the weights of the two-peak SR noise

calculated from the currently estimated pdfs are set to 0.5. We have carried out the experiments using 75 images and present the results of five detection/classification algorithms in Table 3.1, in terms of true-positive fraction (TP) and false positives per image (FPI) [98], where TP is defined as the ratio of the number of the true positive marks to the number of lesions and FPI is defined as the average number of false positives per image. In our work, if a detected positive area has more than 50% overlap with the ground truth area, we consider the detected area to be a TP lesion. Otherwise, we consider it to be a false positive. This is the same definition as used in [99].

We first present the qualitative evaluation of these algorithms. Figs. 3.2, 3.3 and 3.4 show the experimental results for the three ROIs cut from three representative mammograms, where the detected positive pixels are labeled with small dots.

In the experiment shown in Fig. 3.2, a fixed threshold is employed in the 2SR-GBAD-M and SR_IMS algorithms throughout the iterations. The complexity of the mammogram used in these experiments is the lowest compared with the other two to be discussed next. From the figures, we can see that the GBAD and GGD methods find all the lesions, but at the same time generate many false alarms (see Figs. 3.2 (b) and 3.2 (c)). GGD_ID (Fig. 3.2 (d)) is a more robust method. It improves the detection of GGD and performs better than the GADC and IMS methods shown in Figs. 3.2 (e) and 3.2 (g), but it still fails to reduce the false positives satisfactorily. The advantage of the GADC is that it converges quickly, generally in no more than 8 iterations in our experiments, while IMS may converge to local extrema. HOSLW method (Fig. 3.2 (h)) can find the lesions efficiently, but it fails to determine lesion shape which plays a very important role in discriminating the benign tumors from the malignant ones. Moreover, its detection performance depends

on how accurately we can estimate the number of lesion pixels, which is generally not available or known in real-world cases. These detectors suffer from model mismatch and parameter suboptimality resulting in suboptimum detection threshold, and their performances are degraded. The GMM-based detector finds all the lesions but still does not avoid the high P_F (see Fig. 3.2 (i)), which is due to the inaccuracy when GMM is used to fit the mammogram data. In contrast, the presented 2SR-GBAD-M and SR_IMS algorithms yield good detection results in terms of lesion localization, lesion contour exploration and P_F reduction (see Figs. 3.2 (f) and (j)), which demonstrates the capability of the SR noise-based method for enhancing the detectors with model mismatch and parameter suboptimality. Comparing Figs. 3.2 (f) and (j), we can see that SR_IMS performs a little better than 2SR-GBAD-M in reducing false alarms and determining lesion boundaries.

| (a) | (b) | (c) | (d) | (e) |

| (f) | (g) | (h) | (i) | (j) |
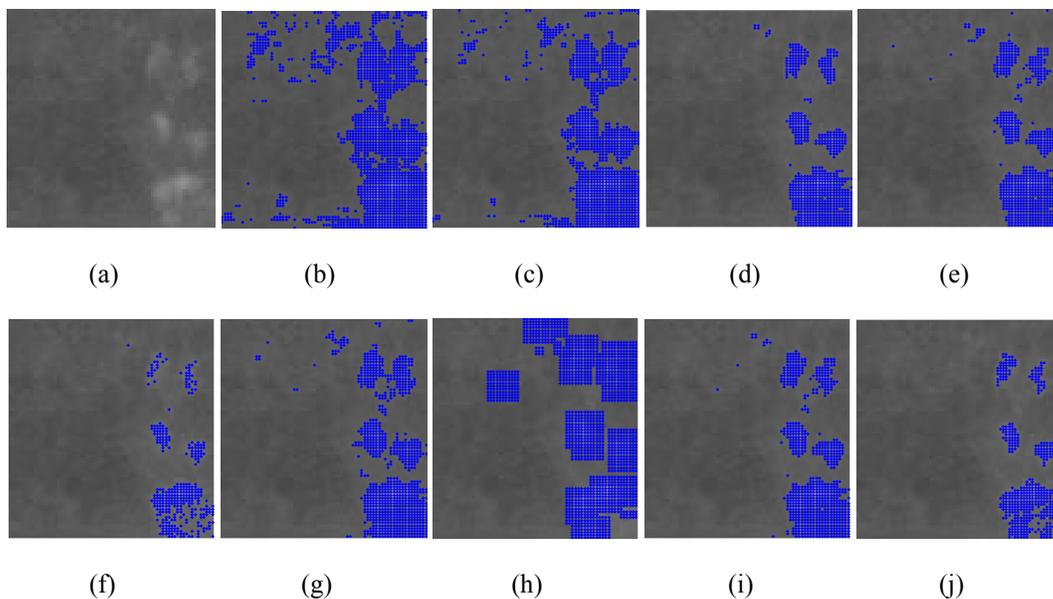
Figure 3.2: Original abnormal mammogram and the detection results *(Abnormal mammogram type 1: homogeneous background with small number of isolated micro-calcifications).* (a) Original mammogram with micro-calcifications; (b) GBAD; (c) GGD; (d) GGD_ID; (e) GADC; (f) 2SR-GBAD-M; (g) IMS; (h) HOSLW; (i) GMM-based detection; (j) SR_IMS. The detected positive pixels are labeled with dots.

Fig. 3.3 shows a more complex case, where both isolated micro-calcifications and crowded clusters exist and the number of lesions is large. We can see that still the 2SR-GBAD-M and SR_IMS algorithms yield better detections with clearer lesion contours and less false positives (see Figs. 3.3 (f) and 3.3 (j)). Compared with GBAD and GGD in Figs. 3.3 (b) and 3.3 (c), GGD_ID and GADC method shown in Figs. 3.3 (d) and 3.3 (e) perform better but still with high P_F's. IMS fails to find some lesions (see Fig. 3.3 (g)). HOSLW (Fig. 3.3 (h)) does not give us much useful information about the lesion positions in this crowded micro-calcifications (clusters) case. This is because its detection operation is performed in subband images which have a quarter of the size of the original mammogram, so the area of the detected positives will be four times of those in the subband images when the detection result is shown in the original mammogram. When the micro- calcifications (clusters) are close to each other, their boundaries and locations are hard to determine. GMM performs better than the rest of the methods (except for Figs. 3.3 (f) and 3.3 (j)), but still generates many false alarms.

(a)   (b)   (c)   (d)   (e)

(f)   (g)   (h)   (i)   (j)
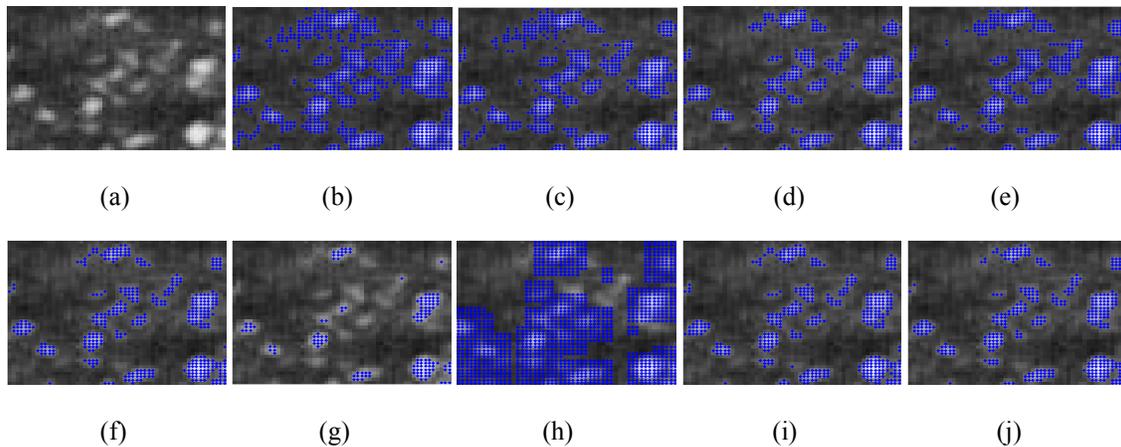
Figure 3.3: Original abnormal mammogram and the detection results *(Abnormal mammogram type 2: homogeneous background with large number of micro-calcifications (clusters)).* (a) Original mammogram with micro-calcifications; (b) GBAD; (c) GGD; (d) GGD_ID; (e) GADC; (f) 2SR-GBAD-M; (g) IMS; (h) HOSLW; (i) GMM-based detection; (j) SR_IMS. The detected positive pixels are labeled with dots.

Fig. 3.4 is the most complex case, where the background distribution is inhomogeneous and some background pixels have their intensities approaching the lesion pixels. It is hard to model the background using just a univariate probability distribution. Finite mixture models [90] may be a choice, but to determine the model type and parameters is also a challenging task. Also, their performance could be deteriorated by the non-stationary nature of the images. Therefore, model mismatch in this type of images is more serious and unavoidable. In our experiment, we still use univariate Gaussian distribution to model the pixel intensity distributions of the background and lesion, respectively, through which the model mismatch is simulated. From Fig. 3.4, we can see that the performance of all the detectors degrades to some extent with higher P_F and lower P_D values as well as more imprecise lesion contours compared with the previous two cases. But the presented 2SR-GBAD-M and SR_IMS algorithms (see Figs. 3.4 (f) and 3.4 (j)) still stand out with better detection results, which again demonstrate their efficiency in reducing the negative influences of model mismatch and suboptimum parameters.

Next, we present the results of quantitative performance evaluation. We select three methods to compare with 2SR-GBAD-M and SR_IMS, and present the results in Table 3.1. The reason we choose GADC and IMS is that they are all classical pattern classification methods and also based on iterative procedures, like 2SR-GBAD-M and SR_IMS. GADC may suffer from model mismatch due to the Gaussian assumption and IMS may have suboptimum threshold value due to the inaccuracy of the *pdf* estimation when processing mammogram data. Additionally, HOLSW is said to be superior to several micro-calcification detectors [97].

(a)   (b)   (c)   (d)   (e)

(f)   (g)   (h)   (i)   (j)
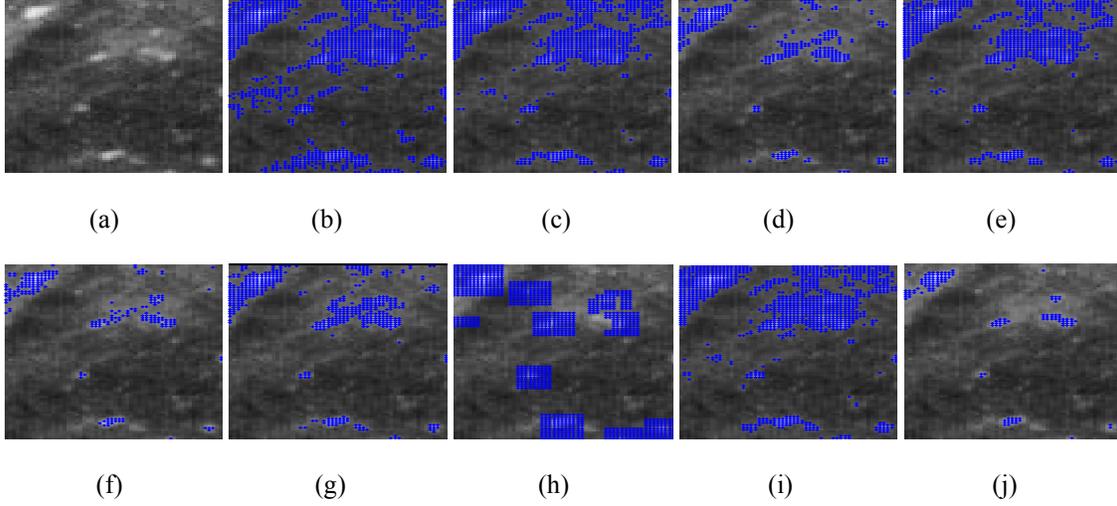
Figure 3.4: Original abnormal mammogram and the detection results (*Abnormal mammogram type 3: inhomogeneous background with moderate number of micro-calcifications (clusters))*. (a) Original mammogram with micro-calcifications; (b) GBAD; (c) GGD; (d) GGD_ID; (e) GADC; (f) 2SR-GBAD-M; (g) IMS; (h) HOSLW; (i) GMM-based detection; (j) SR_IMS. The detected positive pixels are labeled with dots.

TABLE 3.1
DETECTION PERFORMANCE OF FIVE ALGORITHMS

METHODS		GADC	IMS	2SR-G	SR_I	HO
RESULTS						
TP	Range	[0.61,1]	[0.58,1]	[0.80,1]	[0.81,1]	[0.81,1]
	Mean	0.89	0.90	0.93	0.94	0.94
	Standard deviation	0.25	0.28	0.12	0.11	0.11
FPI	Range	[0, 20]	[0, 17]	[0, 9]	[0, 7]	[0, 14]
	Mean	8.16	7.89	4.91	3.12	5.22
	Standard deviation	6.18	7.08	3.94	2.95	4.82

2SR-G: 2SR-GBAD-M; SR_I: SR_IMS; HO: HOLSW.

TP and FPI are employed as the metrics. The means of TP and FPI represent the average performance of each method, and their standard deviations are used as a measure of the robustness of each method when applied to different types of images. A better method is identified to be one with higher mean TP value but lower mean FPI value as well as lower TP and FPI standard deviations.

Since HOLSW requires the knowledge of the number of lesions, which is generally not available in real-world applications, we adjust the lesion number manually, such that the TPs of HOLSW and SR_IMS for each image are the same, and then FPI is employed as a criterion for their performance comparison.

From Table 3.1, we can see that 2SR-GBAD-M and SR_IMS achieve superior performance than the classical methods, GADC and IMS, both in true positive detection and in false positive reduction. HOLSW can attain a similar true positive detection performance as 2SR-GBAD-M and SR_IMS, but it is worse than the two SR noise-enhanced detectors in terms of FPI reduction. 2SR-GBAD-M and SR_IMS have similar detection results, (actually SR_IMS performs a little better) but SR_IMS yields more satisfactory results in terms of FPI reduction. This is because 2SR-GBAD-M updates the threshold based on the Gaussian assumption, and is, therefore, affected by the model mismatch.

It should be emphasized that the detection performance of our detectors may be further improved if image enhancement techniques are employed before detection [100] and post-processing methods, such as pattern classifiers embedded with other lesion features, are used after the detection procedure.

3.4   Summary

Automatic detection techniques for micro-calcifications are very important for breast cancer diagnosis and treatment. Therefore, it is imperative that the detection techniques be developed that detect micro-calcifications accurately. This chapter first developed a lesion detection approach based on SR noise for enhancing the Gaussian assumption-

based detectors which suffer from model mismatch, and furthermore presented a more general SR noise-based detection enhancement framework. Comparative performance evaluation was carried out via experiments between the presented SR noise-based detection enhancement schemes and several detection and classification techniques with three types of representative abnormal mammograms. The results show that the presented algorithm and the framework resulted in highly encouraging performance in terms of flexibility, detection efficiency and system robustness, which demonstrates SR noise's capability of enhancing the suboptimal detectors and supports its real-world CAD application.

CHAPTER IV

IMAGE ENHANCEMENT BASED ON STOCHASTIC RESONANCE NOISE AND SELECTIVE ENHANCEMENT FRAMEWORK

In this chapter, we present two image enhancement approaches. In the first approach, we add a suitable dose of noise to the lower quality images such that the performance of a suboptimal image enhancer is improved without altering its parameters. In the second approach, we present a framework for image enhancement that is based on the selective enhancement technique. Several enhancement algorithms under this framework are developed. Mammogram enhancement is used as an example to illustrate the efficiency of the framework.

4.1 Introduction

As described in Section 1.2, there are mainly four types of image enhancement algorithms. Many popular enhancement algorithms more or less have some disadvantages, which restrain their real-world applications. For example, the standard histogram equalization method [2] processes images globally, thus often causes intensity saturation and the enhanced image is far from a natural one. Its improved versions, such as bi-histogram equalization (BiHE) [15] and contrast limited adaptive histogram equalization (CLAHE) [16], cannot achieve the naturalness of an original image. Median filtering is a standard spatial operation. It often blurs the images and yields poor results when the noise is Gaussian or the number of noise pixels in the processing window is large [18].

Another popular spatial operator, adaptive Wiener filter [18], often results in over-smoothness on the edge. According to [101], a comparative analysis of transform based image enhancement shows that the existing transform-based enhancement techniques have some commonly occurring problems, such as the introduction of artifacts and partial enhancement. In addition, enhancement algorithms attempting to incorporate human visual system (HVS) information are well-motivated, but the unsuitable usage of the HVS properties or the simplified HVS computational models often degrade their performance.

In this chapter, we investigate two enhancement approaches to improve image quality. In the first approach, we attempt to improve image enhancement via the use of stochastic resonance (SR) noise. In recent years, image enhancement methods based on SR noise, such as [102][103] and some work from our group, such as those presented in the previous chapters in this dissertation and the one presented in [104], have appeared in the literature. These methods improve image quality by introducing appropriate noise into the image. But they either have complicated procedures to determine suitable SR noises or are only suitable for enhancing specific image contents. In Chapter 3, we have developed a novel approach to improve the performance of suboptimal breast cancer detection systems by adding suitable noise to the mammograms. In these SR noise-based image processing work presented in [102]~[104] and Chapter 3, however, HVS characteristics during the processing procedures were not considered. In this Chapter, we present a new practical approach for image enhancement based on SR noise, where a systematic methodology is developed to determine SR noise parameters and promising results are shown via a number of illustrative examples. We first formulate the image

enhancement problem as a constrained multi-objective optimization (MOO) problem. Then, we investigate the principle of the SR noise-refined image enhancement scheme, and present an image enhancement system. We also present the method to determine the optimum parameters of the SR noise based on MOO, where a weighting scheme is employed to mimic HVS preferences while selecting solutions from the dominant solution set. In our work, two objective functions are discussed to illustrate the capability of the enhancement scheme, and the experimental results together with performance evaluation are provided in Section 4.2.4.

In the second approach, we present a selective enhancement framework based on image segmentation techniques, which starts with region of interest (ROI) selection, and is followed by ROI enhancement and background suppression. We first introduce the enhancement framework. Then, we introduce thresholding-based enhancement methods, including two basic enhancement techniques, i.e. weighted mean gray value- and fuzzy cross-over point-based thresholding, and two improved methods, a joint enhancer and an iterative enhancement method. After that, we discuss statistical detection-based enhancement schemes. Finally, we show the experimental results and the performance evaluation of the presented approach in Section 4.3.

4.2   Noise-Refined Image Enhancement Using Multi-Objective Optimization

4.2.1   Problem Formulation

The problem of improving the quality of images from degraded observations is an ill-

posed problem. Formally, we have the degraded image model [105],

$$Y = \phi(H(X))\Theta\eta \tag{4.1}$$

where X is the original image, and Y is the observed and degraded image; H corresponds to a shift-invariant point spread function (PSF). The blurred image $H(X)$ is produced and recorded by a sensor, which is often accompanied by a nonlinear transformation of $H(X)$, i.e., $\phi(\cdot)$. Also, noise η can be introduced by the sensor. Θ denotes the process by which measurement noise gets introduced, e.g., additive or multiplicative noise.

In this section, we first formulate the image enhancement problem as a MOO problem, where two objective functions $f(\hat{X})$ and $g(\hat{X},Y)$ are optimized at the same time. Formally, we want to

1. maximize $f(\hat{X})$

2. maximize $g(\hat{X},Y)$ \hfill (4.2)

where $f(\hat{X})$ denotes the desired enhancement characteristic of the enhanced image $\hat{X}$, and $g(\hat{X},Y)$ is the measure of similarity between the enhanced image and the degraded one, to avoid over-enhancement. The specific objective functions used in this work will be discussed in Section 4.2.4 where several illustrative examples are also presented. $\hat{X} = D(Y)$, and D is an existing image enhancement algorithm. Suppose $Y \in \overline{Y}$ and $\hat{X} \in \overline{\hat{X}}$, where $\overline{Y}$ is a subset of $R^{K \times L}$, and $\overline{\hat{X}}$ is a subset of $R^{K' \times L'}$. R is the set of real numbers. (Here, we assume that the pixel intensities of an image can only take certain values, e.g., the integers between 0 and 255, which is a subset of R.) $K \times L$ and $K' \times L'$ are the image sizes of the original image and enhanced image, respectively. In this work, without loss of generality, we assume that $K = K'$ and $L = L'$. Obviously, $\overline{\hat{X}}$ depends

on $\overline{Y}$ for a given algorithm D such that $\overline{\hat{X}} = \left\{ \hat{X} \middle| \hat{X} = D(Y), Y \in \overline{Y} \right\}$. The two objective functions f and g can be considered as quality metrics which evaluate the performance of the enhancer D. To regularize the enhancement result and improve the performance, we further introduce a constraint in the MOO problem, i.e., $g(\hat{X}, Y) \geq b$, that is, we require the processed image to have at least certain similarity to the observed image, which also determines the feasible set. Thus, we have the constrained MOO problem

$$1. \text{ maximize } f(\hat{X})$$

$$2. \text{ maximize } g(\hat{X}, Y)$$

$$\text{subject to } g(\hat{X}, Y) \geq b. \tag{4.3}$$

where the selection of the constraint b will be discussed in Section 4.2.4. The feasible set of the constrained MOO problem is denoted by $\overline{\hat{X}}^F$, and

$$\overline{\hat{X}}^F = \{ \hat{X} \mid \hat{X} = D(Y), Y \in \overline{Y}, g(\hat{X}, Y) \geq b \}.$$

In the current image enhancement research, lots of efforts have been made to model the imaging process (4.1) and design the algorithm D to improve the image quality. These are very challenging tasks, because the mismatches between the assumed model and the actual underlying data features and the imaging process are very difficult to avoid, which often result in poor quality of the processed images.

Rather than developing new models and algorithms to obtain better enhancement, in this work we investigate how to improve an existing image enhancement algorithm D while still keeping the structure including the parameters of D unchanged. From the formulation of the constrained MOO problem (4.3), we can see that for fixed f, g and D,

one important way to improve the enhancement is to alter the feasible set of solutions $\bar{\hat{X}}^F$, such that the resulting set possibly contains better solutions. This could be achieved by introducing suitable SR noise into the observation Y. The performance improvement will be achieved by finding the solution of problems posed in (4.9) and (4.11) from the altered feasible set. Details of this process will be discussed in the next two sections.

4.2.2   Stochastic Resonance Noise-Refined Image Enhancement

In this section, we present the principle of SR noise-refined image enhancement, and describe the scheme and system for improving the image enhancement performance.

4.2.2.1   Principle of SR Noise-Refined Image Enhancement

Given a degraded image $Y=y$, we obtain the unique enhanced version $\hat{X} = D(Y)$ employing the existing enhancer D. Suppose $\hat{X}$ satisfies $g(\hat{X}, Y) \geq b$, then the feasible set $\bar{\hat{X}}^F$ includes only one element $\hat{X} = \hat{x} = D(y)$, and $\bar{\hat{X}}^F = \bar{\hat{X}}$. Our goal is to enlarge $\bar{\hat{X}}^F$ via the introduction of SR noise. This enlarged $\bar{\hat{X}}^F$ may potentially contain solutions that are better than $D(Y)$. In this chapter, we only consider the additive SR noise case. Other ways of introducing SR noise, such as multiplicative SR noise, can be analyzed in a similar manner.

Suppose there is a SR noise n with the probability density function (*pdf*) $P_n(n)$, and after adding the SR noise into Y in a pixel-wise manner, we obtain the modified image data Z,

$$Z = Y + N \tag{4.4}$$

The enhancer D is applied to the modified data Z to obtain the modified enhancement result

$$\hat{X}' = D(Z) = D\big(Z(N)\big) \qquad\qquad (4.5)$$

where N is a SR noise matrix with the same size as Y, and the notation $Z(N)$ explicitly indicates that Z is a function of N. Each element (pixel) of the $K \times L$ matrix N of the SR noise n is generated using the *pdf* $P_n(n)$ in an independent manner. In our work, we assume that the added SR noises are independent for different pixel positions, so that the *pdf* of N can be written as

$$P_N(N) = \prod_{i=1}^{K \times L} P_n(n_i) \qquad\qquad (4.6)$$

where i is the pixel index of an image.

Let $\Gamma = \big\{ D(Z(N)) \mid N \in R^{K \times L} \big\}$ be the set of all enhancement results of the original algorithm D using the SR noise-modified images with all the possible $P_n(n)$ that are candidates for use as SR noise. Γ includes the original single element set $\overline{X} = \big\{ \hat{X} \mid \hat{X} = D(Y), Y = y \big\}$ when no SR noise is added. Thus, for a given $Y = y$, we extend the set $\overline{X}$ to a larger set $\overline{X}' = \Gamma$ by adding SR noise. So the original constrained MOO problem will have a potentially [3] larger feasible set $\overline{X}^{F}{}' \in \overline{X}'$, where $\overline{X}^{F}{}' = \{ \hat{X}' \mid \hat{X}' = D(Z(N)), Y = y, g(\hat{X}', Y) \geq b \}$, from which better solutions might be generated.

---

[3] It is possible that some elements of Γ generated by adding SR noise may not satisfy $g(\hat{X}, Y) \geq b$. But if there exists at least one enhanced image which satisfies the condition, we obtain a larger feasible set and one or more elements of this set might yield better solutions than the original one without SR noise, in terms of the values of the two objective functions.

Here the term "better solutions" corresponds to a "more dominant" Pareto-optimal front[4] for the constrained MOO problem (4.9) to be discussed later. We call this scheme the single SR noise refinement scheme as a single realization of N is employed.

Since the larger feasible set provides more chances to find a better image enhancement result, we may further improve the SR noise refinement scheme by creating an even larger feasible set which includes $\bar{\hat{X}}^{F}{'}$ and may yield an "even more dominant" Pareto-optimal front. A convex combination using $P_N(N)$ is proposed in this chapter for the task. Formally, we have

$$D'(Z) = \int_N P_N(N) D(Z(N)) dN \tag{4.7}$$

Let $\Xi = \left\{ D'(Z) \mid D'(Z) = \int_N P_N(N) D(Z(N)) dN, N \in R^{K \times L} \right\}$ be the set of all $D'(Z)$, which includes Γ and may be viewed as a "convexified" version of $D(Z(N))$ on $R^{K \times L}$. Thus, through randomization we further extend the set $\bar{\hat{X}}{'}$ to a larger set $\bar{\hat{X}}{''} = \Xi$ and obtain potentially an even larger feasible set $\bar{\hat{X}}^{F}{''}$ of solutions, i.e., enhanced images, where $\bar{\hat{X}}^{F}{''} \in \bar{\hat{X}}{''}$ and $\bar{\hat{X}}^{F}{''} = \{\hat{X}{''} \mid \hat{X}{''} = D'(Z) = \int_N P_N(N) D(Z(N)) dN, Y = y, g(\hat{X}{''}, Y) \geq b\}$. We call this scheme the multiple SR noise refinement scheme, since from (4.7) we can see that multiple SR noise matrices N need to be generated to produce the combination coefficients $P_N(N)$. We will employ this scheme in the chapter to refine the enhancement performance. The multiple SR noise matrices are generated by repeating the procedure for yielding single realization of matrix N in the single SR noise refinement scheme

---

[4] The Pareto-optimal front is a collection of all the solutions which are not dominated by any other solutions. A solution X_1 is said to dominate another solution X_2, if and only if $f(X_1) \geq f(X_2)$ and $g(X_1, Y) > g(X_2, Y)$, or vice versa .

multiple times, i.e., multiple realizations of the SR noise matrix with the *pdf* $P_N(N)$.

In real-world applications, it is not easy to find the exact result of (4.7) since the integration of D over $P_N(N)$ may be intractable. Therefore, we use the empirical average to approximate the expectation over $P_N(N)$, which is shown in (4.8).

$$D_{Emp}'(Z) = \frac{1}{M} \sum_{i=1}^{M} D(Z(N_i)) \qquad (4.8)$$

where M is the number of the realizations of the SR noise matrix, $N_1, N_2, ..., N_M$, with *pdf* $P_N(N)$. Given enough number of SR noise realizations, we may obtain sufficiently accurate $D'(Z)$ [106]. Another advantage of using empirical average is that the output of the averaged enhancement result has the same expected output as that using only one SR noise realization, but with a smaller variance [104]. This scheme is employed for system realization in the next section.

4.2.2.2   *SR Noise-Refined Image Enhancement Scheme and System*

So far, we have discussed the principle of SR noise-refined enhancement and the underlying mechanism. However, to determine the conditions for improvability of an algorithm and to find the *pdf* of the optimum SR noise for an improvable algorithm are tough tasks, since they depend on the properties of D, f, and g and on the observed image. So, instead of finding the form of optimum SR noise to achieve the maximum performance improvement for any specific properties as mentioned above, in this chapter, we pre-decide the form of the SR noise and then determine its optimum parameters which yield maximum performance improvement. Of course, this procedure may allow

us to only find the sub-optimum SR noise because we fix the form of the SR noise. But it makes the procedure more practical, since the knowledge of the required properties in real-world applications may not be available or very difficult to determine. Once more information, like the monotonicity, continuity and convexity properties of D, f and g and the observation, is available, more analytical results may be obtained and the performance could be further improved by determining the form of the optimum SR noise *pdf*.

The SR noise-refined image enhancement scheme can be summarized as follows. We choose the form of a suitable SR noise from a SR noise pool which contains several SR noise candidates with controllable parameters, e.g., Gaussian noise with adjustable mean and variance. The Pareto-optimal front, consisting of the SR noise-modified non-dominant solutions, is determined by using a MOO algorithm for the constrained MOO problem (4.9). The final solution of the enhanced image and, therefore, the optimum parameters of the SR noise are selected by optimizing a cost function shown in (4.11), which will be discussed in the next section.

$$1. \text{ maximize } f\big(D'(Z)\big)$$

$$2. \text{ maximize } g\big(D'(Z),Y\big)$$

$$\text{subject to } g\big(D'(Z),Y\big) \geq b. \tag{4.9}$$

where the optimization is with respect to the parameter θ_n of the SR noise distribution $P_n(n)$.

The SR noise-refined image enhancement system is shown in Fig. 4.1, where the image quality is evaluated in terms of the two objective functions and the constraint. It

should be noted that the SR noise can be introduced either in spatial domain or in transform domain. The original enhancers can be pixel, spatial or transform operators, making the SR noise-refined enhancement system applicable to improving a wide variety of algorithms. If the SR noise is introduced in spatial domain, i.e., introduced into the original image directly, there is no "Transform" block in Fig. 4.1, and $Y' = Y$. If the SR noise is introduced in the transform domain, Y in (4.4) is replaced with Y', and $\hat{X} = D(Y)$ mentioned above is changed to $\hat{X} = D(Y')$, where Y' is the transformed image without SR noise. The image enhancers in Fig. 4.1 carry out image enhancement in the transform domain and then transform back the processed image to spatial domain once the enhancement is done. In Fig. 4.1, satisfaction with the result in the decision block means that there exists θ_n^* which is the solution of the problems posed in (4.9) and (4.11). As mentioned before, M realizations of the SR noise matrix, $N_1, N_2,..., N_M$, with pdf $P_N(N)$, are employed, which yield M modified images, $Z_1, Z_2,..., Z_M$, where $Z_i = Y + N_i$, $i = 1,2,...,M$. Each image is enhanced by the original enhancer, D, to yield M enhanced images, $D(Z(N_1)), D(Z(N_2)),..., D(Z(N_M))$. The empirical expectation as discussed earlier is found by the averaging operation.
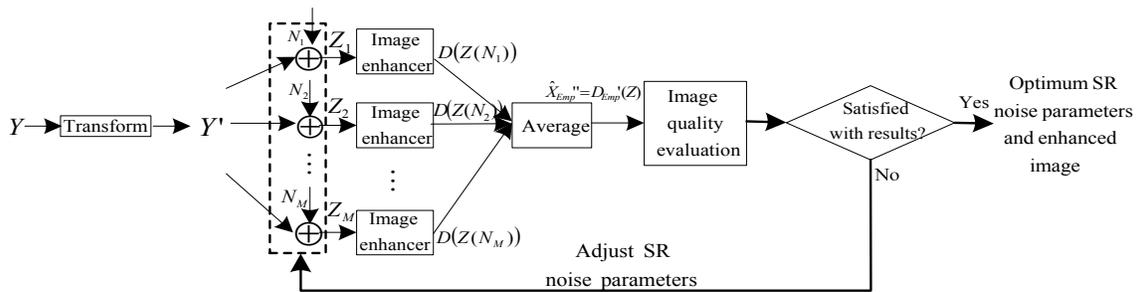
Figure 4.1: SR noise-refined image enhancement system.

4.2.3 Selection of Solution Based on Human Visual System Preferences

In general, the analytical solution to the multiple-step image processing problem shown in Fig. 4.1 is not easy to obtain, which is highly dependent on the properties of D, g, f as well as on the content of a specific image. Therefore, we employ a genetic algorithm-based MOO method, non-dominated sorting genetic algorithm–II (NSGA-II) [107], to find the Pareto-optimal front for the problem posed in (4.9). For illustration, Fig. 4.2 shows a typical Pareto-optimal front of the constrained MOO problem (4.9) after introducing SR noise into the image, where the solid segment of the front corresponds to the solutions satisfying the constraint in (4.9) and the dashed line corresponds to the solutions violating the constraint. Points A and B correspond to two types of image contents, Y_1 and Y_2, respectively. They represent two original solutions without SR noise, where point A is below the front and point B is on the front. The original enhancement results corresponding to the two solutions could be improved if (4.11) is satisfied, even if there is a solution lying on the Pareto optimal front. This is because we will use a non-dominance criterion to select the solution from the Pareto-optimal front, which incorporates HVS information and will be discussed next.

We note that the point corresponding to the original enhancement will not lie above the front. This is because the Pareto-optimal front consists of all the dominant solutions. If the solution corresponding to the original enhancement result (without SR noise) dominates the other solutions and, therefore, lies on the Pareto-optimal front such as the point B in Fig. 4.2, we can set the SR noise equal to zero.

$f(D'(Z))$

Pareto-optimal front

$B : \big(f(D(Y_2)), g(D(Y_2), Y_2)\big)$

$A : \big(f(D(Y_1)), g(D(Y_1), Y_1)\big)$

b

$g(D'(Z), Y = y)$
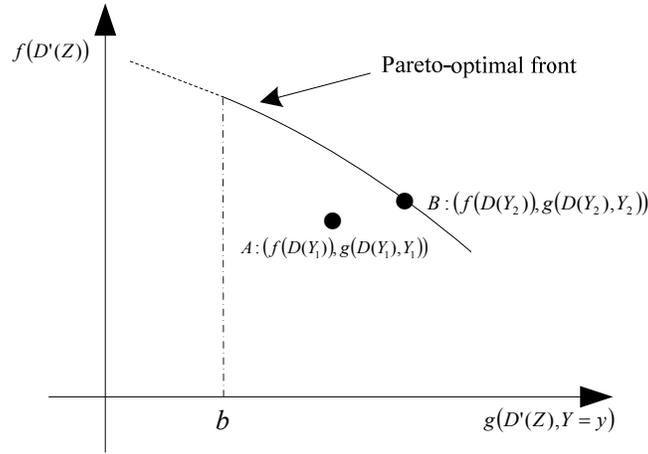
Figure 4.2: A typical Pareto-optimal front of the constrained MOO problem after introducing SR noise.

Admittedly, it cannot be guaranteed that all the solutions from the Pareto-optimal front generated by adding SR noise can yield better enhancement results than the original enhancer in terms of human visualization. Therefore, we need a criterion and a strategy to select the solutions from the front which correspond to better quality of the enhanced image. In our work, we employ an HVS-weighting scheme for the selection of solutions, which is described as follows.

According to the previous discussion, we observe that the two objective functions represent different image characteristics which are optimized to enhance visualization by humans, so a combination of the objective functions may take into account these characteristics at the same time. We also note that different people have different visual preference for the same object. So, to include human preferences while selecting the solution from the Pareto-optimal front, we use a solution selection function, in terms of a linear combination of the two objective functions, as the criterion to evaluate the quality

of an image and to select the solutions from the front. More specifically, we weigh the two objective functions with non-negative real numbers, that is, g is multiplied by a weight w_1 and f is multiplied by w_2, where $0 \leq w_1, w_2 \leq 1$ and $w_1 + w_2 = 1$. Thus, we have

$$L(D'(Z), w_1, w_2) = w_1 g(D'(Z), Y) + w_2 f(D'(Z)) \tag{4.10}$$

where w_1 and w_2 denote the tradeoff between the consistency and the desired characteristics of the enhanced image. A larger value of w_2 indicates that HVS prefers the desired image characteristic, while a larger value of w_1 represents the HVS's preference for consistency. A larger value of L denotes better quality of the enhanced image for a specific HVS preference. Therefore, a solution from the Pareto-optimal front, and correspondingly the parameters of the SR noise, is considered to represent an improved performance over the original enhancer for a specified HVS preference, i.e., a specified (w_1, w_2) pair, if (4.11) is satisfied.

$$L(D'(Z), w_1, w_2) - L(D(Y), w_1, w_2) > 0 \tag{4.11}$$

where $L(D(Y), w_1, w_2) = w_1 g(D(Y), Y) + w_2 f(D(Y))$, corresponding to the enhancement result when no SR noise is introduced.

We can see that this selection scheme is different from the one based on the dominance criterion in which we require $f(D'(Z)) \geq f(D(Y))$ and $g(D'(Z), Y) > g(D(Y), Y)$ (or, $f(D'(Z)) > f(D(Y))$ and $g(D'(Z), Y) \geq g(D(Y), Y)$) at the same time. The advantage of this scheme is its flexibility in terms of its ability to select solutions which emphasize the desirable goals, where the linear combination is employed to consider the different characteristics at the same time. Moreover, if the original enhanced image already has some acceptable characteristic which does not need to be refined further, this scheme

permits us to improve other desired image attributes while still maintaining the image characteristic within an acceptable level. As we will observe in the experiment involving SR noise-refined CLAHE in Section 3.2.4, the original CLAHE performs satisfactorily in maintaining the similarity between the enhanced image and the degraded one, which can be seen in Fig. 4.4 (b). Therefore, in this case, the SR noise is mainly employed to improve other characteristics of the enhanced image, such as the contrast information represented by f, while the similarity is still maintained to an acceptable extent. Another advantage of the selection scheme is that it can reduce the solution size dramatically and take into account different preferences of the HVS. If additional information about the HVS preferences becomes available, in terms of new weights, it is not necessary to rerun the algorithm. We can choose an alternative solution from the non-dominant solution set obtained from the MOO which is best suited for the new weights.

4.2.4 Illustrative Examples of Noise-Refined Image Enhancement

For illustration, in this section we employ a specific objective function pair to illustrate the efficiency of the proposed scheme. As mentioned in Section 2.1, we investigate four types of SR noises in this chapter. They are Gaussian SR noise (GaSR), uniformly distributed SR noise (UnSR), triangle SR noise (TrSR) and two peak SR noise (TwSR) [53].

4.2.4.1   Objective Function Design

One objective function is employed to maintain the consistency between the observation and the enhanced image, and to reduce over-enhancement, which is evaluated by the

structural similarity (*SSIM*) index [108] in this work. *SSIM* measures the similarity between the processed image and the original one using image structure information,

$$SSIM\ (\hat{X},Y) = \frac{(2\mu_{\hat{X}}\mu_Y + C_1)(2\sigma_{\hat{X}Y} + C_2)}{(\mu_{\hat{X}}^2 + \mu_Y^2 + C_1)(\sigma_{\hat{X}}^2 + \sigma_Y^2 + C_2)} \tag{4.12}$$

where $\mu_{\hat{X}}, \mu_Y$ and $\sigma_{\hat{X}}, \sigma_Y$ as well as $\sigma_{\hat{X}Y}$ denote mean intensity and contrast as well as the correlation coefficient of images $\hat{X}$ and Y, respectively; C_1 and C_2 are constants used to avoid instabilities for very small μ or σ. The value of $SSIM(\hat{X},Y)$ is between 0 and 1. A higher value means more similarity between two images.

The contrast sensitivity information of an image provides an important indication of image quality. We, therefore, expect that the enhanced image contains high contrast sensitivity information. We represent the contrast sensitivity information by a novel metric, $ConSen(\hat{X})$, developed as follows.

In [109], a model of the contrast sensitivity function (*CSF*) for luminance images has been considered, which describes human's sensitivity to spatial frequencies f_r and is shown in (4.13),

$$CSF(f_r) = 2.6(0.192 + 0.114 f_r)e^{-(0.114 f_r)^{1.1}} \tag{4.13}$$

When a 4-level discrete wavelet transform (DWT) of an image is carried out, the relation between *CSF* and the spatial frequency ranges is shown in Fig. 4.3 [110]. We notice from the figure that the sensitivity information is mainly distributed in the 3rd and 4th level DWT. Therefore, we weigh the DWT coefficients of these two levels with the *CSF* and add the absolute values of these weighted coefficients together. The summation is then divided by the total number of pixels in the image to reduce the effect of the image size, as shown in (4.14). This is employed as a quality evaluation metric to

measure the average contrast sensitivity information included in an image. A higher value indicates that one of the image characteristics, contrast in this chapter, is favored more by HVS, which represents a better image quality. This metric is denoted by *ConSen*[5]. The wavelet we use is Daubechies-8.

$$ConSen(\hat{X}) = \frac{\{\sum_{u=1}^{[K/4]} \sum_{v=1}^{[L/4]} | Coef(u,v) \cdot CSF(f_r) |\}}{KL} \tag{4.14}$$

where *K* and *L* provide the image size as mentioned before. *Coef(u,v)* are the decomposition coefficients, and *(u,v)* (in cycles/degree) are the directions in decomposition domain, with $f_r = \sqrt{u^2 + v^2}$. [*a*] denotes the smallest integer larger than *a*.
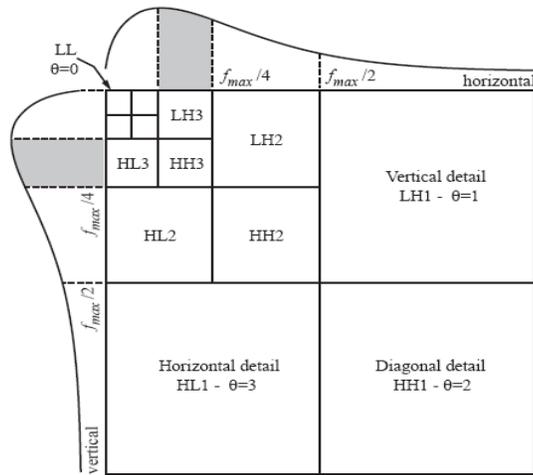
Figure 4.3: Luminance *CSF* along horizontal and vertical directions of four-level wavelet decomposition [110].

Thus, we have the pair of objective functions

$$f(\hat{X}) = ConSen(\hat{X}) \tag{4.15}$$

$$g(\hat{X}, Y) = SSIM(\hat{X}, Y) \tag{4.16}$$

---

[5] This metric is based on joint work with Mr. Vijay Chintham Reddy.

The constraint is specified in terms of g, i.e., $g(\hat{X}, Y) \geq b$, where b is a pre-chosen constant, $0 \leq b \leq 1$. It represents the least consistency the enhanced image is required to satisfy, which defines the feasible set of solutions for the optimization problem.

4.2.4.2 Experimental Results

We examine the SR noise-refined CLAHE and fuzzy [111] enhancers in enhancing medical images. We also investigate the improvement of image de-noising using median and Wiener filters.

The population size used in the MOO is 100 and the number of generations is 50. The number of noise realizations, M, is 35. In the weighting scheme, we choose uniformly spaced weight vectors in the interval [0, 1], i.e., w_1 changes from 0 to 1 with the increment of 0.1, that is, there are 11 simulated HVS preferences for each SR noise system. b is set equal to 0.7.

4.2.4.2.1 Medical Image Enhancement

Medical images typically suffer from impairments such as low resolution, high level of noise, low contrast, geometric deformations and the presence of imaging artifacts [112]. Image enhancement is an important part of the computer-aided diagnostic (CAD) technique. In this section, contrast enhancement of mammograms is considered, where the location and lesion types are identified by the radiologists as the ground-truth used in our work. The mammograms are from Digital Database for Screening Mammography (DDSM) [91].

4.2.4.2.1.1   CLAHE

We examine the application of CLAHE to enhance the X-ray mammogram by adding SR noise in spatial domain. The enhancement results of mammogram with masses are provided in Fig. 4.4. We show the results corresponding to the weight vectors {0.3, 0.7} and {0.6, 0.4} for the four types of SR noise-refined systems, representing two HVS preferences.

From Fig. 4.4, we can visually see that the SR noise system can improve the enhancement corresponding to different simulated HVS preferences. Next, we provide quantitative evaluation using a non-reference image quality metric. Since the increase of contrast is an important sign of improvement of the medical image quality, we choose to use the metric presented in [113], which is designed to measure the sharpness of an image. We call it "sharpness index". The higher the value of the index, the better the contrast. (Sharpness provides important contrast information, and $ConSen$ is closely related to sharpness. Besides the sharpness, we also want to keep the consistency between the enhanced image and the original one to some extent, so both $ConSen$ and *SSIM* are employed to guide the enhancement procedure.) Table 4.1 shows the comparison of the enhancement results in Fig. 4.4 with different SR noises.
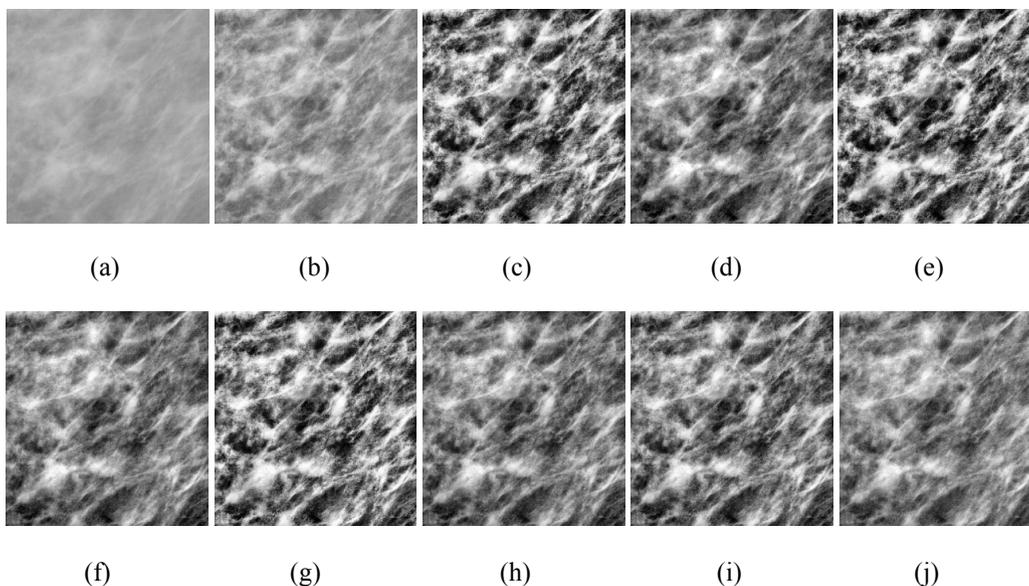
Figure 4.4: Enhancement results of the mammogram with masses. (a) Original mammogram; (b) enhanced by CLAHE; (c) enhanced by CLAHE with GaSR ({0.3, 0.7}); (d) enhanced by CLAHE with GaSR ({0.6, 0.4}); (e) enhanced by CLAHE with UnSR ({0.3, 0.7}); (f) enhanced by CLAHE with UnSR ({0.6, 0.4});(g) enhanced by CLAHE with TrSR ({0.3, 0.7}); (h) enhanced by CLAHE with TrSR ({0.6, 0.4}); (i) enhanced by CLAHE with TwSR ({0.3, 0.7}); (j) enhanced by CLAHE with TwSR ({0.6, 0.4}).

From Table 4.1, we can see that in the mammogram mass enhancement, for the HVS preference simulated by {0.3, 0.7} and {0.6,0.4}, GaSR yields best results and TwSR performs the worst, but all of the four SR noises improve CLAHE. In the rest of the experiments, we will see that the performance ranks of the SR noise change with different enhancers, weight vectors and image contents, which indicates that the efficiency of the SR noise is problem-dependent and is also closely related to HVS preferences.

TABLE 4.1
SHARPNESS INDEX OF THE ENHANCED IMAGES

				Sharpness index
Mammogram with masses: CLAHE	No SR noise	Original degraded image		6.3317
		Image enhanced by CLAHE only		6.5372
	With SR noise	GaSR	{0.3,0.7}	7.4873
			{0.6,0.4}	6.8093
		UnSR	{0.3,0.7}	7.1189
			{0.6,0.4}	6.7213
		TrSR	{0.3,0.7}	7.0793
			{0.6,0.4}	6.7094
		TwSR	{0.3,0.7}	7.0498
			{0.6,0.4}	6.6424
Cell image: CLAHE	No SR noise	Original degraded image		6.5365
		Image enhanced by CLAHE only		6.7090
	With SR noise	GaSR	{0.2,0.8}	7.3389
			{0.7,0.3}	6.7490
		UnSR	{0.4,0.6}	7.3749
			{0.8,0.2}	6.7208
		TrSR	{0.4,0.6}	7.1233
			{0.7,0.3}	6.7112
		TwSR	{0.2,0.8}	7.0689
			{0.8,0.2}	6.7225
Mammogram with micro-calcifications: Fuzzy logic enhancer	No SR noise	Original degraded image		6.6518
		Image enhanced by Fuzzy logic enhancer only		6.8225
	With SR noise	GaSR	{0.2,0.8}	7.4139
			{0.7,0.3}	7.2419
		UnSR	{0.4,0.6}	7.1926
			{0.8,0.2}	7.0817
		TwSR	{0.2,0.8}	7.4624
			{0.8,0.2}	7.3954

Fig. 4.5 shows the enhancement results of the electroscope cell image using the presented SR noise-refined enhancement system, where different weight vectors, simulating different HVS preferences, are employed. We can visually see that the SR

noise system improves the enhancement for different simulated HVS preferences. The quantitative evaluation results using the sharpness index are shown in Table 4.1.

| (a) | (b) | (c) | (d) | (e) |

| (f) | (g) | (h) | (i) | (j) |
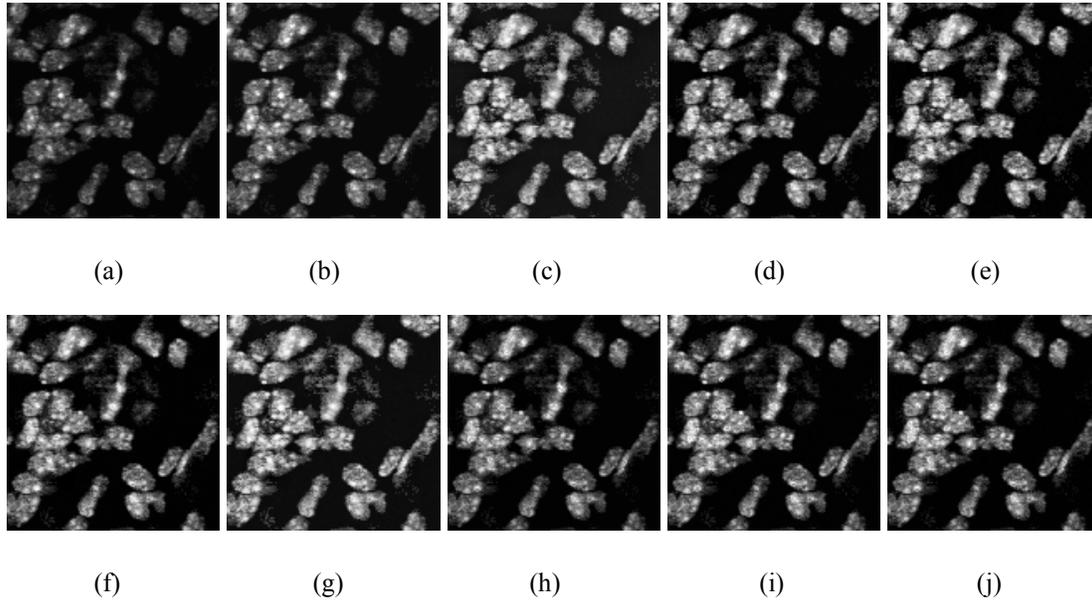
Figure 4.5: Enhancement results of the cell image. (a) Original cell image; (b) enhanced by CLAHE; (c) enhanced by CLAHE with GaSR ({0.2, 0.8}); (d) enhanced by CLAHE with GaSR ({0.7, 0.3}); (e) enhanced by CLAHE with UnSR ({0.4, 0.6}); (f) enhanced by CLAHE with UnSR ({0.8, 0.2}); (g) enhanced by CLAHE with TrSR ({0.4, 0.6}); (h) enhanced by CLAHE with TrSR ({0.7, 0.3}); (i) enhanced by CLAHE with TwSR ({0.2, 0.8}); (j) enhanced by CLAHE with TwSR ({0.8, 0.2}).

From these results, we can see that uniformly distributed SR noise (UnSR) produces the best result for the given weight vectors.

4.2.4.2.1.2   Fuzzy Logic Histogram Hyperbolization

Fuzzy logic enhancers are designed for modeling vague and ambiguous information, two characteristics of medical and biological images [114]. We now use SR noise to improve the fuzzy logic histogram hyperbolization method [115], where the SR noise is added

after the original image is transformed to the fuzzy domain.

From the experiments, we notice that there is no solution satisfying (4.9) and (4.11) at the same time when triangle noise is used, which indicates that triangle noise cannot enhance the image quality for the given weights, D, f, g and the degraded image. The enhancement results for the other three types of SR noises are shown in Fig. 4.6, and the quantitative evaluation results are also shown in Table 4.1. From the figures, we can see that all three of the SR noise-refined enhancers increase the contrast in the images, and also reduce the over-enhancement phenomenon of the original fuzzy logic enhancer. For the weight vectors used here, TwSR yields the best enhancement.

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

(e)   (f)   (g)   (h)
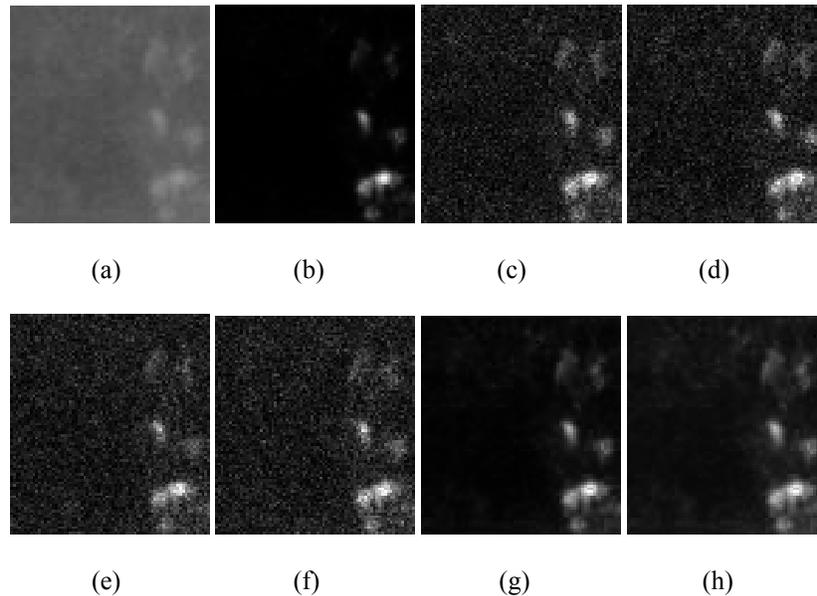
Figure 4.6: Enhancement results of the mammogram with micro-calcifications. (a) Original mammogram image; (b) enhanced by fuzzy enhancer; (c) enhanced by fuzzy enhancer with GaSR ({0.4, 0.6}); (d) enhanced by fuzzy enhancer with GaSR ({0.7, 0.3}); (e) enhanced by fuzzy enhancer with UnSR ({0.3, 0.7}); (f) enhanced by fuzzy enhancer with UnSR ({0.7, 0.3}; (g) enhanced by fuzzy enhancer with TwSR ({0.2, 0.8}); (h) enhanced by fuzzy enhancer with TwSR ({0.8, 0.2}).

4.2.4.2.2   Image De-noising

In this section, we consider the application of the SR noise-refined system to improve the quality of images disturbed by Gaussian mixture noise, and present the results representing a single HVS preference. We investigate two commonly used filters, i.e., median filter for "Lena" image with signal-to-noise ratio (SNR) of 16.61dB and Wiener filter for "Cameraman" image with SNR of 12.32 dB. The de-noising results are shown in Figs. 4.7 and 4.8. The weight vector for the images is {0.5, 0.5}. Note that the Wiener filter is a linear operation, but it involves the estimation of the local second-order statistics of the image signal and noise in the observation, so the effect of the additive noise will not be washed off by the averaging operation. Here, we determine the statistics locally because image contents are non-stationary.

A full-reference quality metric, mean square error (MSE), between the de-noised image and the ground-truth image, is employed for quantitative performance evaluation. The MSE is defined as

$$MSE(\hat{X}, X_0) = \frac{\sum_{r=1}^{K}\sum_{s=1}^{L}[\hat{X}(r,s) - X_0(r,s)]^2}{KL} \tag{4.17}$$

where $\hat{X}$ and X_0 denote the de-noised and the ground-truth noise-free images, respectively, and (r,s) are the pixel locations in an image. For each type of SR noise, the noise-refined filter is investigated to de-noise the image using 11 HVS preferences. Figs. 4.9 (a) and (b) show the evaluation results based on the MSE for median and Wiener filters, respectively. For each HVS preference, the enhancement improvement is identified by the MSEs of the SR noise-refined enhancement results which are lower than that of the enhancement without SR noise. Lower MSE denotes more enhancement

improvement. The mean value of each curve shows the average performance of one type of SR noise, and its variance measures the robustness of the system using this SR noise. For clarity, the enhancement results with very large MSE are not shown. The HVS preference indices correspond to the weight w_1 changing from 0 to 1 with the increment of 0.1.

(a)   (b)   (c)

(d)   (e)   (f)   (g)
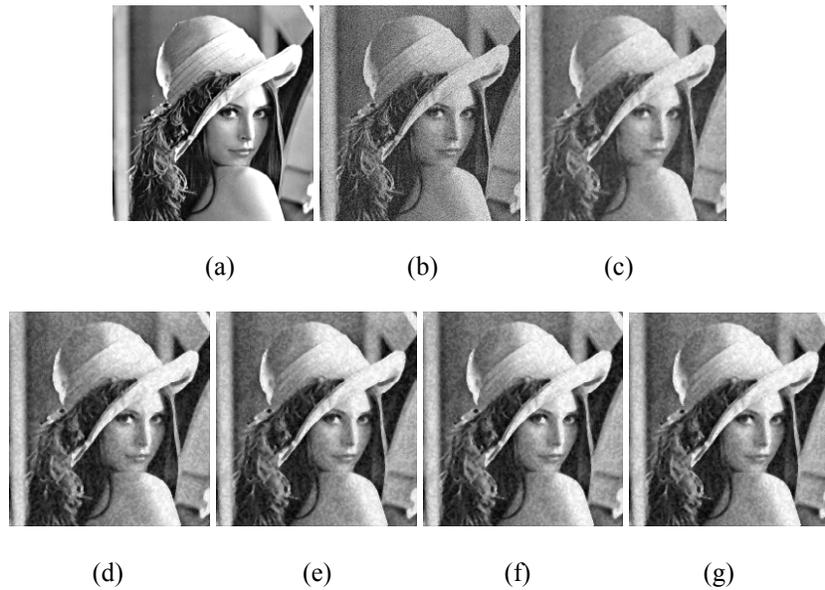
Figure 4.7: Median filter for de-noising "Lena" image. (a) Noise-free image; (b) image disturbed by Gaussian mixture noise with the SNR of 16.61dB; (c) de-noised by median filter; (d) de-noised by median filter with GaSR; (e) de-noised by median filter with UnSR; (f) de noised by median filter with TrSR; (g) de-noised by median filter with TwSR. The weight vector used is{0.5,0.5}.

(a)　　　　　　(b)　　　　　　(c)

(d)　　　　(e)　　　　(f)　　　　(g)
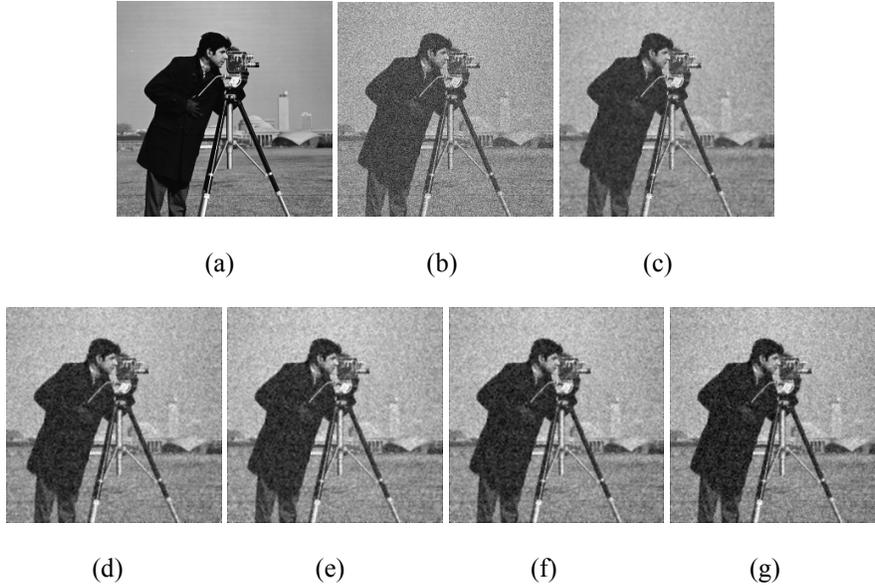
Figure 4.8: Wiener filter for de-noising "Cameraman" image. (a) Noise-free image; (b) image disturbed by Gaussian mixture noise with the SNR of 12.32dB; (c) de-noised by Wiener filter; (d) de-noised by Wiener filter with GaSR; (e) de-noised by Wiener filter with UnSR; (f) de-noised by Wiener filter withTrSR; (g) de-noised by Wiener filter with TwSR. The weight vector used is{0.5,0.5}.
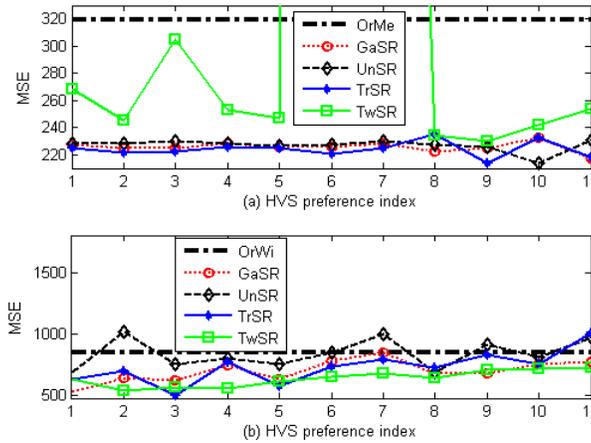
Figure 4.9: MSE of the de-noising results using median and Wiener filters. (a) Median filter-based de-noising; (b) Wiener filter-based de-noising. OrMe and OrWi mean the de-noising using original median and Wiener filter, respectively, without SR noise.

From Figs. 4.7 and 4.8 and the experiments, we observe that the TrSR noise performs the best for the "Lena" image, TwSR performs the best for the "Cameraman" image. Also, in an informal evaluation, the quality of the enhanced images is deemed better than the enhanced image without SR noise by human judges for all the weights. However, using MSE as the criterion, we can see from Fig. 4.9 that for some weights, SR noise worsens the enhancement. This illustrates that the objective function pair, as a quality metric, is more consistent with human evaluation results, while MSE has some disadvantages when used as metric for image quality evaluation [116].

Tables 4.2 shows the quantitative evaluation results of image de-noising based on MSE statistics, for median and Wiener filters. The MSE statistics include the range, mean and variance of the MSE, as well as the number of weight vectors corresponding to the enhancement results which have lower MSE than the SR noise-free case. The mean value shows the average system performance and the other three statistics illustrate the robustness of the system.

TABLE 4.2

MSE STATISTICS OF MEDIAN AND WIENER FILTERS DE-NOISING

				MSE statistics
De-noising using median filter ("Lena")	No SR noise	MSE value of noisy image		354.0
		MSE value of the de-nosed image by median filter only		319.7
	With SR noise	GaSR	Range	[216.8,232.4]
			LN*	11
			Mean	225.7
			Variance	15.4
		UnSR	Range	[214.9, 230.6]
			LN*	11
			Mean	227.0
			Variance	18.8
		TrSR	Range	[213.6,235.5]
			LN*	11
			Mean	224.0
			Variance	36.6
		TwSR	Range	[230.2,3385.0]
			LN*	9
			Mean	627.9
			Variance	924870
De-noising using Wiener filter ("Camera-man")	No SR noise	MSE value of noisy image		1466.4
		MSE value of the de-noised image by Wiener filter only		849.8
	With SR noise	GaSR	Range	[534.3, 850.9]
			LN*	10
			Mean	700.1
			Variance	8073.4
		UnSR	Range	[688.5,1019.0]
			LN*	6
			Mean	842.2
			Variance	1390.3
		TrSR	Range	[505.2,1011.0]
			LN*	10
			Mean	731.5
			Variance	1817
		TwSR	Range	[536.6,726.0]
			LN*	11
			Mean	640.7
			Variance	4224.5

*LN means the number of the weight vectors corresponding to the enhancement results which have lower MSE than the SR noise-free case. In this paper there are 11 weight vectors for each SR noise, i.e., w_1 changes from 0 to 1 with the increment of 0.1.

4.2.4.2.3   Image De-blurring

We also employ the SR noise-refined system to improve two image de-blurring

algorithms, Lucy-Richardson algorithm and regularized filter, which are applied when images are suffering from convolution and motion effects, respectively. The convolution effect is simulated by using a low-pass Gaussian filter to smooth a ground-truth image. The size of the filter is 13 by 13 and its variance is 36. The motion effect is simulated by convoluting an image with a filter which models a linear motion of camera with 20 pixels and with an angle of 10 degrees in a counterclockwise direction. The images shown use the weight vector {0.4,0.6}.

The experimental results are shown in Figs. 4.10, 4.11 and 4.12. We can see that triangle noise works best when Lucy-Richardson algorithm is used for de-blurring, but works worst in the regularized filter case although the resulting enhanced image quality is not much visually worse than the SR noise-free case. Uniformly distributed SR noise generates the most significant improvement in regularized filter de-blurring.

Table 4.3 shows the quantitative evaluation results for image de-blurring based on MSE statistics for Lucy-Richardson algorithm and regularized filter.

(a)   (b)   (c)

(d)   (e)   (f)   (g)
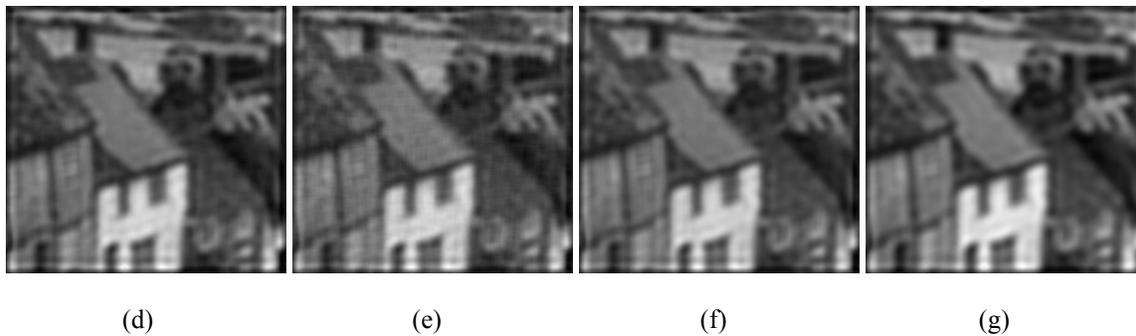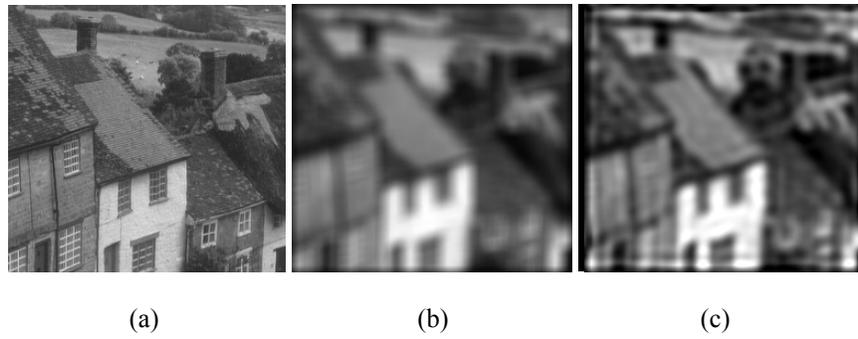
Figure 4.10: Lucy-Richardson algorithm for de-blurring "House" image. (a) Ground-truth image; (b) blurred image; (c) de-blurred by Lucy-Richardson; (d) de-blurred by Lucy-Richardson with GaSR; (e) de-blurred by Lucy-Richardson with UnSR; (f) de-blurred by Lucy-Richardson with TrSR; (g) de-blurred by Lucy-Richardson with TwSR. {0.4,0.6}.
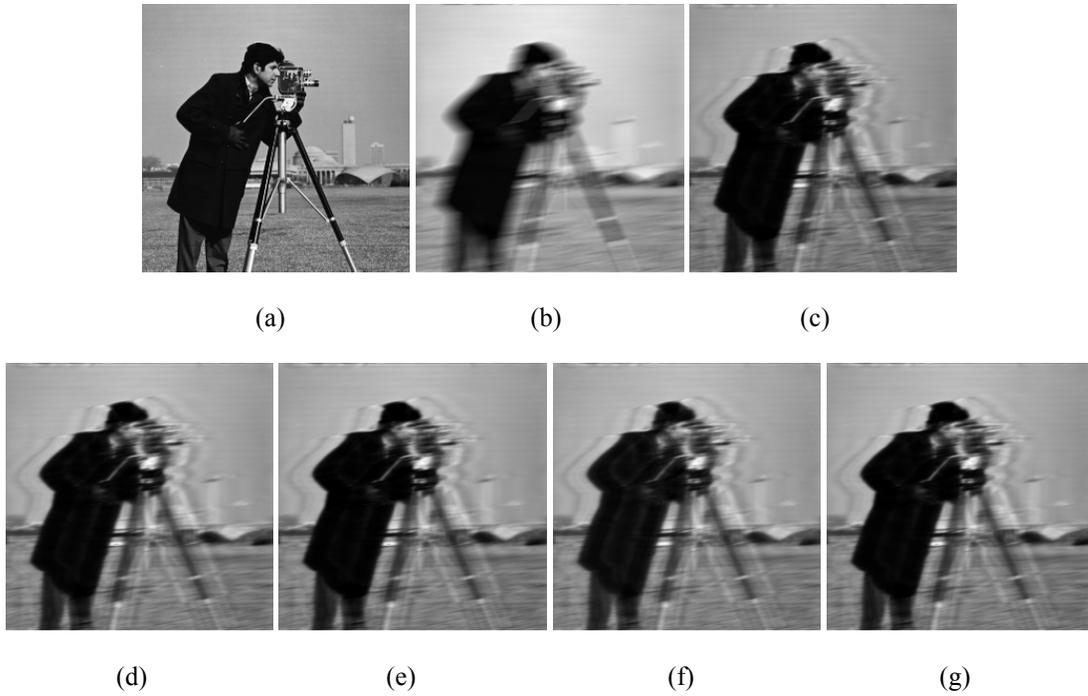
Figure 4.11: Regularized filter for de-blurring "Cameraman" image. (a) Ground-truth image; (b) blurred image; (c) de-blurred by regularized filter; (d) de-blurred by regularized filter with GaSR; (e) de-blurred by regularized filter with UnSR; (f) de-blurred by regularized filter with TrSR; (g) de-blurred by regularized filter with TwSR. {0.4,0.6}.
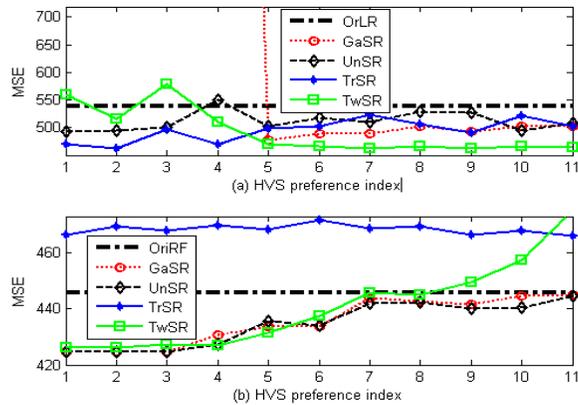
Figure 4.12: MSE of the de-blurring results using Lucy-Richardson algorithm and regularized filter. (a) Lucy-Richardson algorithm de-blurring; (b) regularized filter-based de-blurring. OrLR and OriRF mean the de-blurring using Lucy-Richards algorithm and regularized filter, respectively, without SR noise.

TABLE 4.3

MSE Statistics of The Lucy-Richardson Algorithm And Regularized Filter De-Blurring

				MSE statistics
De-blurring using Lucy-Richardson algorithm ("House")	No SR noise	MSE value of blurred image		561.6
		MSE value of the de-blurred image by Lucy-Richardson algorithm only		538.9
	With SR noise	GaSR	Range	[477.7, 4036.1]
			LN	7
			Mean	1679.3
			Variance	2775.1
		UnSR	Range	[493.3, 549.4]
			LN	10
			Mean	511.6
			Variance	312.3
		TrSR	Range	[462.4,524.2]
			LN	11
			Mean	494.9
			Variance	416.3
		TwSR	Range	[462.3,579.5]
			LN	9
			Mean	492.9
			Variance	1816.0
De-blurring using regularized filter ("Camera-man")	No SR noise	MSE value of blurred image		1719.3
		MSE value of the de-blurred image by regularized filter only		445.7
	With SR noise	GaSR	Range	[424.6, 445.1]
			LN	11
			Mean	435.5
			Variance	71.7
		UnSR	Range	[424.6, 444.4]
			LN	11
			Mean	434.5
			Variance	62.9
		TrSR	Range	[465.9,471.6]
			LN	0
			Mean	468.3
			Variance	2.9
		TwSR	Range	[426.2,475.2]
			LN	7
			Mean	440.8
			Variance	247.5

We can see from Table 4.3 as well as Figs. 4.10, 4.11 and 4.12 that the TrSR performs the best in Lucy-Richardson algorithm-based de-blurring, and performs the worst in the regularized filter case although the resulting enhanced image is not much worse visually than the SR noise-free case (Fig. 4.11 (f)). UnSR yields the most significant improvement in regularized filter de-blurring.

From these experimental results, we see that the SR noise-refined system can improve the performance of several types of image enhancement algorithms when dealing with different distortion situations.

4.3 Image Enhancement Based on Selective Enhancement Framework

4.3.1 The Selective Enhancement Framework

We first introduce the "selective enhancement" framework, as shown in Fig. 4.13. In the framework, the ROI, e.g., the region containing lesions in a mammogram, is determined, and then some image enhancement and noise suppression techniques, such as the gray level stretching method, are used to increase the contrast in the ROI and suppress the background noise. More generally, we can divide an image into several ROIs according to their relative importance and other characteristics, and employ suitable enhancement and suppression algorithms for different ROIs. This chapter only deals with the two-class case, one ROI and the background. More general cases can be considered in a similar manner.
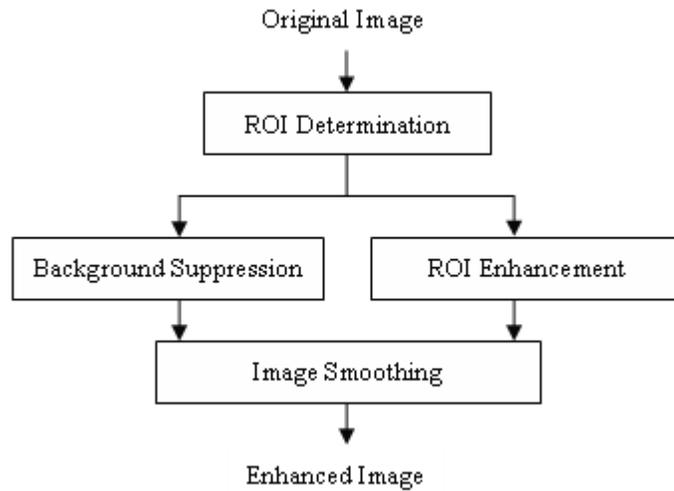
Figure 4.13: "Selective enhancement" framework (two-class case).

From the diagram, we can see that ROI determination is the first and fundamental step of this framework, which leads to the next steps. Based on ROI determination, different operations, i.e. enhancement and suppression, are carried out in different areas of the image. The main steps of the algorithm, i.e. ROI determination and enhancement as well as background suppression, are described as follows. Gray level stretching technique is employed in this chapter to carry out the enhancement and suppression operations. Other techniques can also be employed under our framework. This type of methodology where different operations are used in different parts of the image has been employed in the past (e.g. [66][117]). To the best of our knowledge, the framework proposed here has not been presented elsewhere. Our main emphasis in this section is to propose a number of methods for ROI determination that result in superior enhancement performance.

In this section, we will use a specific application of this framework, i.e., mammogram enhancement, to illustrate the efficiency of the presented approach.

First, the gray values of the original mammogram are rescaled to the range from zero to the full pixel intensity, e.g., 255 in this section. A method is used to find the ROI (determination of the ROI will be discussed later).

Second, a gray level stretching method is employed on the ROI, according to (4.18),

$$g(x,y) = \frac{N-n}{M-m}[f(x,y)-m]+n \qquad (4.18)$$

where $f(x, y)$ and $g(x, y)$ are the gray values of the pixel at location (x, y) in the rescaled and stretched ROI respectively. $[m,M]$ and $[n,N]$ are the gray value ranges of the rescaled and stretched ROI, where M=255 due to the rescaling operation and m is the lowest gray value in the rescaled ROI. Here, N is set to 255, the highest gray value, to make the enhanced ROI involve the brightest part in the mammogram.

Because ROI may contain both actual lesions and some mis-labeled background pixels with lower gray values, increase of contrast in the ROI is necessary for a better visualization. Therefore, we need $N-n>M-m$. We can make n=0, such that all the non-ROI parts have zero intensity. However, the non-ROI part may also include valuable information for diagnosis, so maintenance of the non-ROI information to some moderate extent is necessary. Our experiments show that setting n equal to the mean gray value of the background of the rescaled mammogram yields satisfactory results, both in having lesions stand out and in suppressing the background. To reduce the influence of a small number of pixels with extreme values, the mean values in this chapter are all calculated by weighting each gray value with the ratio of the number of pixels having the gray value to the total number of pixels in the mammogram, i.e. weighted mean is computed and used.

Thirdly, the gray value range of the stretched ROI is from the mean gray value of the background of the rescaled mammogram to 255. So we need to change the background gray values to some extent, such that background suppression, similarity maintenance and image smoothing are achieved at the same time. A reasonable choice is to make the background pixel gray values range from zero to the minimum gray value of the stretched ROI, using a method similar to (4.18).

Finally, an adaptive filter, such as an adaptive Wiener filter, is used for de-noising and further smoothing. For abnormal mammograms that include micro-calcifications, a matched filter with a Gaussian appearance may be used instead of the adaptive filter to further enhance the lesions [67].

As can be seen from the above procedure, ROI determination plays a fundamentally important role in enhancement. There are many methods dealing with this problem, such as thresholding methods [118][119] and segmentation methods [120][121]. According to the pixel intensity properties of the mammograms, we have developed several algorithms for selecting the ROI, which will be discussed in the next few sections.

4.3.2   Thresholding-Based Mammogram Enhancement Algorithms

In this section, several thresholding methods are presented for ROI determination. These are based on the pixel intensities of the mammogram under consideration.

Higher pixel gray values than the surrounding normal tissues distinguish lesions from the normal structures in mammograms, which is one of the most important features of abnormal mammograms and widely used in lesion detection and mammogram enhancement. The presented algorithm tries to increase the contrast of the lesions against

the normal background by exploring the pixel gray value information.

For the micro-calcification cases, we set the so-called cross-over point as the threshold, which corresponds to the intensity of micro-calcifications and can be determined automatically [99]. However, sometimes, the cross-over point value is a little higher than the optimal value and some lesions are excluded from the ROI, such that these lesions are suppressed during enhancement.

The second method is to set the mean gray value of the whole mammogram as the threshold, because the lesion intensity generally is above the average gray level of the whole mammogram. (Of course, a lower threshold can be used to make sure that all the lesions are considered with a higher confidence.) With this threshold, lesions are included in the ROI together with more background pixels than the first thresholding scheme, because it is a little lower than the optimal value. As a result, contrast between the lesions and background is not explored enough through the enhancement process.

The enhancement results corresponding to the two thresholding methods are shown in Figs. 4.14 (h) and (i) in Section 4.3.4.

The two enhancement methods can increase the contrast between the lesions and background, benefiting visualization, but a technique to set an appropriate threshold is desired. In general, it is not easy to find the optimal threshold by using analytical methods, but the tradeoff between the previous two threshold determination schemes might yield satisfactory results. Here, we use a simple image fusion technique to find the tradeoff, i.e., to average the enhanced mammograms resulting from the two threshold schemes respectively. This is named joint enhancer, and the corresponding enhancement

result is shown in Fig. 4.14 (j) in Section 4.3.4.

As mentioned before, the lesion pixel intensity generally is above the average gray value of the whole mammogram, so the thresholding scheme 2 may include all the lesions in the ROI. The lesions become more visible through the enhancement process, and the gray values of most background pixels will be suppressed in the range from zero to the minimum gray value of the stretched ROI through one enhancement operation. Accordingly, larger contrast may be expected if the enhancement is iteratively used on the mammogram. This is because the threshold includes some background pixels in the ROI with relatively lower gray values compared with lesion pixels. And due to the gray level stretching applied to ROI, the intensities of some background pixels will be decreased below the threshold. Therefore, if we set another appropriate threshold on the enhanced mammogram, some background pixels in the previously stretched ROI may be moved to the new non-ROI region and therefore be suppressed through the gray level stretching process applied to the new non-ROI region. Thus, more and more background pixels will be suppressed with the iteration process, which may further increase the contrast. By changing thresholds of the mean value thresholding scheme in each iteration, we can determine the ROIs in the iterative enhancement procedure. Of course, other thresholds, e.g. those larger than a quarter of the full gray level and smaller than the mean value threshold, may also yield good results. A more flexible scheme is to use variable thresholds, e.g. thresholds decreasing with the number of iterations. We can keep iterating until the enhancement result is satisfactory. The stopping criterion can be formed simply as follows: at iteration $n+1$, if the lesions that stood out at iteration n disappear or are weakened, then we choose to stop at iteration n. At each step of iteration,

some pixels of the old ROI in the prior enhanced mammogram will be excluded. So if the iteration number is too large, over-enhancement will arise where some lesion pixels are suppressed due to enhancement. Therefore, an appropriate number of iterations is required. Through experiments, we find that 3 iterations is a good choice when we use the mean value threshold. The results of the iterative enhancement method applied to micro-calcifications, mass and spiculated lesions are shown in Fig. 4.14 (k) and Fig. 4.15 (h) and Fig. 4.16 (h) in Section 4.3.4.

4.3.3 Statistical Detection-Based Mammogram Enhancement

In this section, decision theory based statistical detection methods are employed for ROI determination, where the detected positives are classified as ROI. Two models will be considered for the micro-calcifications.

A low probability of false alarm (P_F) is very important to our enhancement problem. This is because if the actual P_F is too high, more background pixels will be included in the ROI, which worsens the enhancement result. We will show that an acceptable P_F is achieved with the refinement of the detection schemes.

4.3.3.1   Gaussian Background Assumption-Based Detector and Enhancement

The micro-calcification detection problem is actually an anomaly detection problem [78], and we assume the asymptotic distribution, when the number of reference samples approaches infinity, of the background, i.e. normal tissues, to be Gaussian. This leads to

linear and tractable solutions [13] and is named the Gaussian Background Assumption-Based Detector (GBAD).

The micro-calcifications, which are considered as the signal of interest here, are brighter spots than the surrounding normal tissues. So the signal with constant amplitude larger than the average pixel gray value is a gross model for the micro-calcifications. Thus, the lesion detection problem is converted to detecting a constant signal embedded in Gaussian noise. Recalling (3.1) in Chapter 3, we wish to choose between the two hypotheses

$$H_0 : y[m] = w[m]$$
$$H_1 : y[m] = A + w[m]$$

$$(4.19)$$

where $y[m]$ is the pixel gray value, $w[m]$ is the Gaussian background noise with mean μ_b and variance σ_b^2, and A is a positive constant, larger than μ_b. m is the index of the pixel to be processed in each detection, and $m = 0,1,\cdots,M-1$. In the section, we extend the statistical test to include more pixels in each detection, instead of the signal pixel detection discussed in Chapter 3.

Based on the Neyman-Pearson criterion [13], we get the GBAD statistic $T(y)$ and the test as follows

$$T(y) = \frac{1}{N} \sum_{m=0}^{M-1} (y[m] - \mu_b) \underset{\substack{< \\ decide \; H_0}}{\overset{\substack{decide \; H_1 \\ >}}{}} \gamma_1 \qquad (4.20)$$

where the threshold γ_1 is determined from the given P_F and the statistical parameters, mean and variance, of the pixels surrounding the pixel to be detected.

The parameters are estimated as follows,

$$\mu_b = \frac{1}{M_b} \sum_m y_b[m] \tag{4.21}$$

$$\sigma_b^2 = \frac{1}{M_b} \sum_m (y_b[m] - \mu_b)^2 \tag{4.22}$$

where the subscript b means that the estimation is carried out for background pixels; M_b denotes the number of the background pixels in each processing window, which are used to estimate these parameters; m is the pixel index.

If we assume a single pixel target [78], that is, in each detection, only one pixel is taken as input, then $M=1$. The statistical test will reduce to the case discussed in Chapter 3. The detection result is shown in Fig. 4.17 (h) in Section 4.3.4, where the micro-calcifications are completely detected but with a higher P_F than the preset value, 0.01. An important reason is that the lesion and background distributions are not well modeled. And, therefore, the contrast in the mammogram (see Fig. 4.17 (i)) is not explored enough.

4.3.3.2   *General Gaussian Detector and Enhancement*

Micro-calcifications, especially micro-calcification clusters, are small but not with a constant intensity, so a Gaussian model as opposed to a constant signal model may be more reasonable to represent them, which gives rise to a General Gaussian Detector (GGD) test.

To design the GGD, we refer to the GBAD. The detected positives of the GBAD, which include all the micro-calcifications and some background pixels, are employed to roughly estimate the mean, μ_s and variance, σ_s^2, of the micro-calcifications for the Gaussian model, using an approach similar to that used in (4.21) and (4.22).

In the general Gaussian case, recalling (3.1), we wish to choose between the two hypotheses

$$
\begin{aligned}
H_0 &: y[m] = w[m] \\
H_1 &: y[m] = s[m] + w[m]
\end{aligned}
\tag{3.1}
$$

where $s[m]$ denotes the micro-calcification signal obeying Gaussian distribution, $N(\mu_s, \sigma_s^2)$. We therefore obtain the statistical test

$$
T(y) = \sum_{m=0}^{M-1} (y[m] - \mu_b + \frac{\sigma_b^2}{\sigma_s^2} \mu_s)^2 \underset{\underset{decide\ H_0}{<}}{\overset{\overset{decide\ H_1}{>}}{}} \gamma_2
\tag{4.23}
$$

where the threshold γ_2 is determined from the given P_F and the statistical parameters, mean and variance, of the pixels surrounding the pixel to be detected.

The detection result is shown in Fig. 4.17 (j) in Section 4.3.4, where all the micro-calcifications are discovered by GGD but with less false positives than the GBAD, and therefore the enhancement result (Fig. 4.17 (k)) shows improvement.

4.3.3.3   Iterative Detection and Enhancement Method

Although the previous two detection-based enhancement algorithms yield good results, the detection and accordingly the enhancement may be further improved. Here we employ the GGD-based iterative detector (GGD_ID) discussed in Section 3.2.1.3 to further improve the detection and therefore the enhancement performance. The detection and enhancement results are shown in Figs. 4.17 (l) and (m).

4.3.4   Performance Evaluation of Selective Enhancement Framework

The performances of the enhancement algorithms are evaluated both subjectively (qualitatively) and objectively (quantitatively). The subjective evaluation is carried out through human eye judgment, in terms of the comparison with the original mammograms and several representative enhancement algorithms.

For comparison, we implemented six representative enhancement algorithms. These algorithms include standard histogram equalization, CLAHE, unsharp masking, fuzzy histogram hyperbolization with *S*-function and *norma*-function [111][114], and fuzzy logic possibility distribution method [122][123] techniques. The qualities of some fuzzy contrast enhancement algorithms were evaluated in [113], where fuzzy histogram hyperbolization and fuzzy logic possibility distribution method stand out because of their good performances in providing global contrast improvement and enhancing the highest density regions of a mammogram, respectively.

4.3.4.1   Qualitative Evaluation

Fig. 4.14 compares several representative enhancement algorithms with the presented algorithms for enhancing mammograms with micro-calcifications.

From the figures, we see that although the two threshold determination schemes can enhance the lesions to some extent, threshold 1 (h) suppresses some lesions while threshold 2 (i) enhances too many background pixels such that some lesions are not obvious enough. The joint enhancer (j) emphasizes all the lesions, and there are less

enhanced background pixels compared with the other algorithms. The iterative enhancement method (k) also yields good result.

(a)

(b)   (c)   (d)   (e)   (f)

(g)   (h)   (i)   (j)   (k)
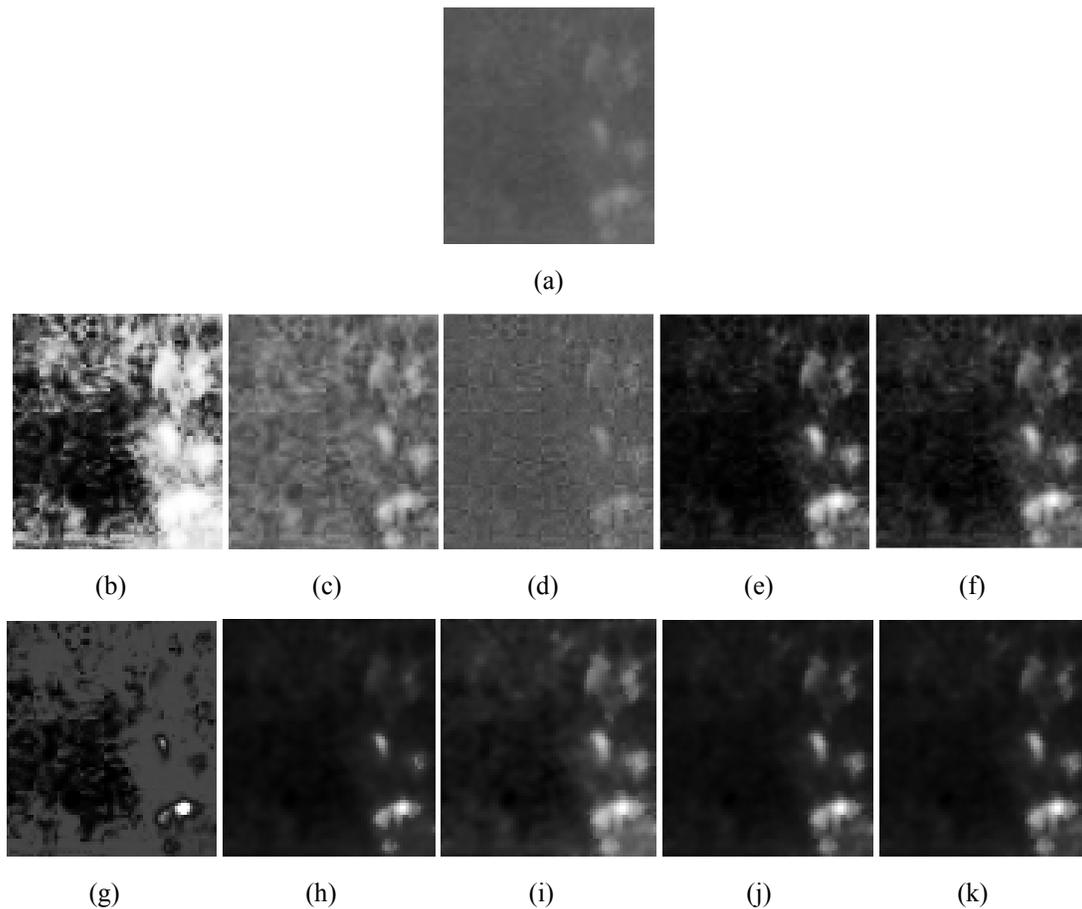
Figure 4.14: Original abnormal mammogram with micro-calcifications and its enhanced versions. (a) Original mammogram with micro-calcifications; (b) enhanced mammogram with standard histogram equalization; (c) enhanced mammogram with CLAHE; (d) enhanced mammogram with unsharp masking; (e) enhanced mammogram with fuzzy histogram hyperbolization (*S*-functionand); (f) enhanced mammogram with fuzzy histogram hyperbolization (*norma*-function); (g) enhanced mammogram with fuzzy logic possibility distribution method; (h) enhanced mammogram with threshold 1; (i) enhanced mammogram with threshold 2; (j) enhanced mammogram using the joint enhancer; (k) enhanced mammogram using iterative enhancement method.

Figs. 4.15 and 4.16 compare the enhancement results of mammograms with mass and spiculated lesions, respectively, using the presented iterative enhancement method and several other enhancement algorithms.

(a)	(b)	(c)	(d)
(e)	(f)	(g)	(h)
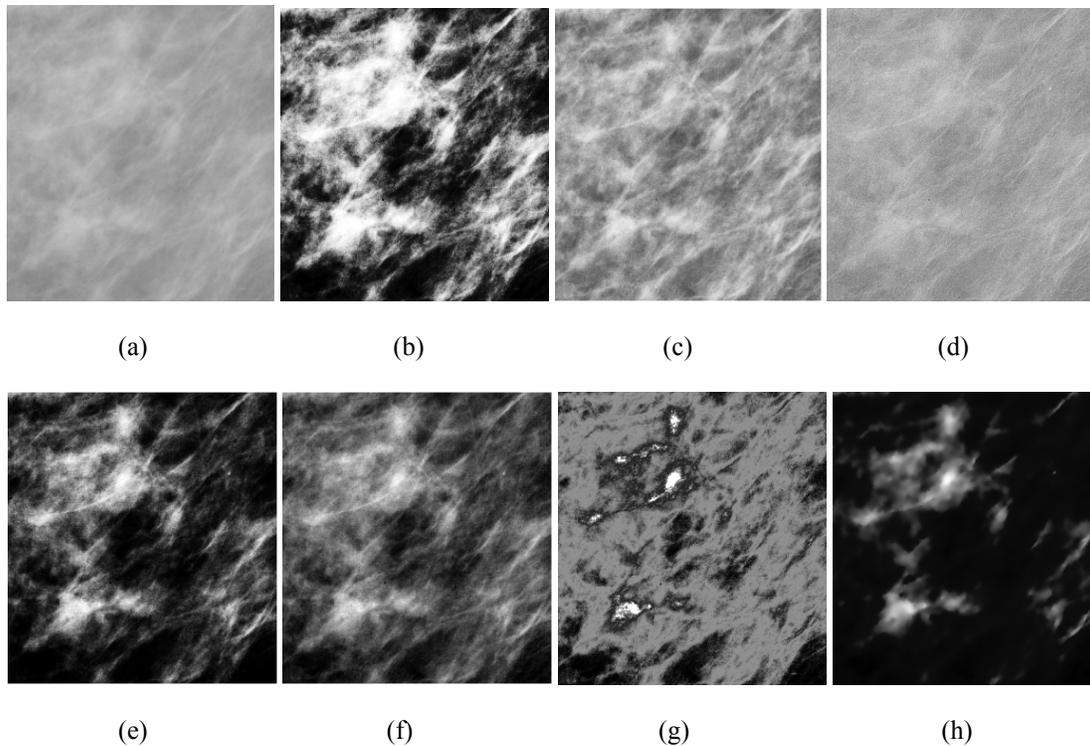
Figure 4.15: Mammogram with mass and its enhanced versions. (a) abnormal mammogram with mass; (b) enhanced mass with standard histogram equalization; (c) enhanced mass with CLAHE; (d) enhanced mass with unsharp masking; (e) enhanced mass with fuzzy histogram hyperbolization (S-function); (f) enhanced mass with fuzzy histogram hyperbolization (norma-function); (g) enhanced mass with fuzzy logic possibility distribution method; (h) enhanced mass with iterative enhancement.

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

(e)   (f)   (g)   (h)
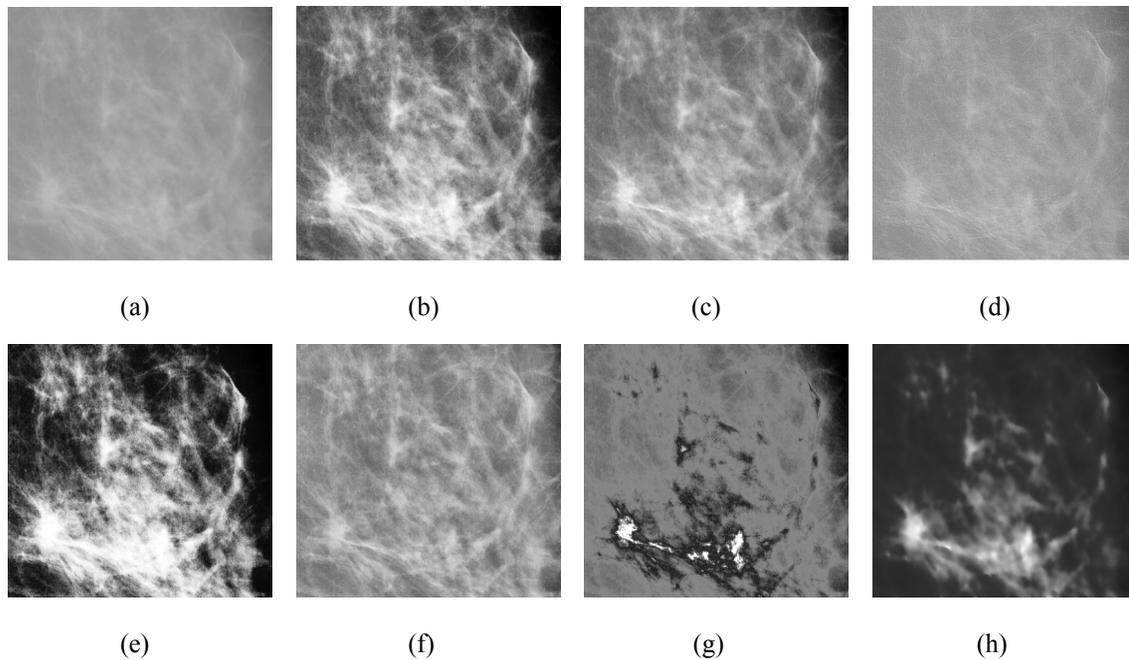
Figure 4.16: Mammogram with spiculated lesions and its enhanced versions. (a) abnormal mammogram with spiculated lesions; (b) enhanced mass with standard histogram equalization; (c) enhanced mass with CLAHE; (d) enhanced mass with unsharp masking; (e) enhanced mass with fuzzy histogram hyperbolization (S-function); (f) enhanced mass with fuzzy histogram hyperbolization (norma-function); (g) enhanced mass with fuzzy logic possibility distribution method; (h) enhanced mass with iterative enhancement.

From Figs. 4.15 and 4.16, we can see that the iterative scheme keeps all the lesions and dramatically reduces the false positives compared with other algorithms. The corresponding enhanced mammograms (Figs. 4.15 (h) and 4.16 (h)) show very good results.

The performance of the presented statistical detection-based enhancement is compared with other enhancement algorithms in Fig. 4.17. For the ease of comparison, in Fig. 4.17 we show the enhancement results of several other algorithms provided in Fig. 4.14 again.

(a)　　　　(b)　　　　(c)　　　　(d)　　　　(e)

(f)　　　　(g)　　　　(h)　　　　(i)

(j)　　　　(k)　　　　(l)　　　　(m)
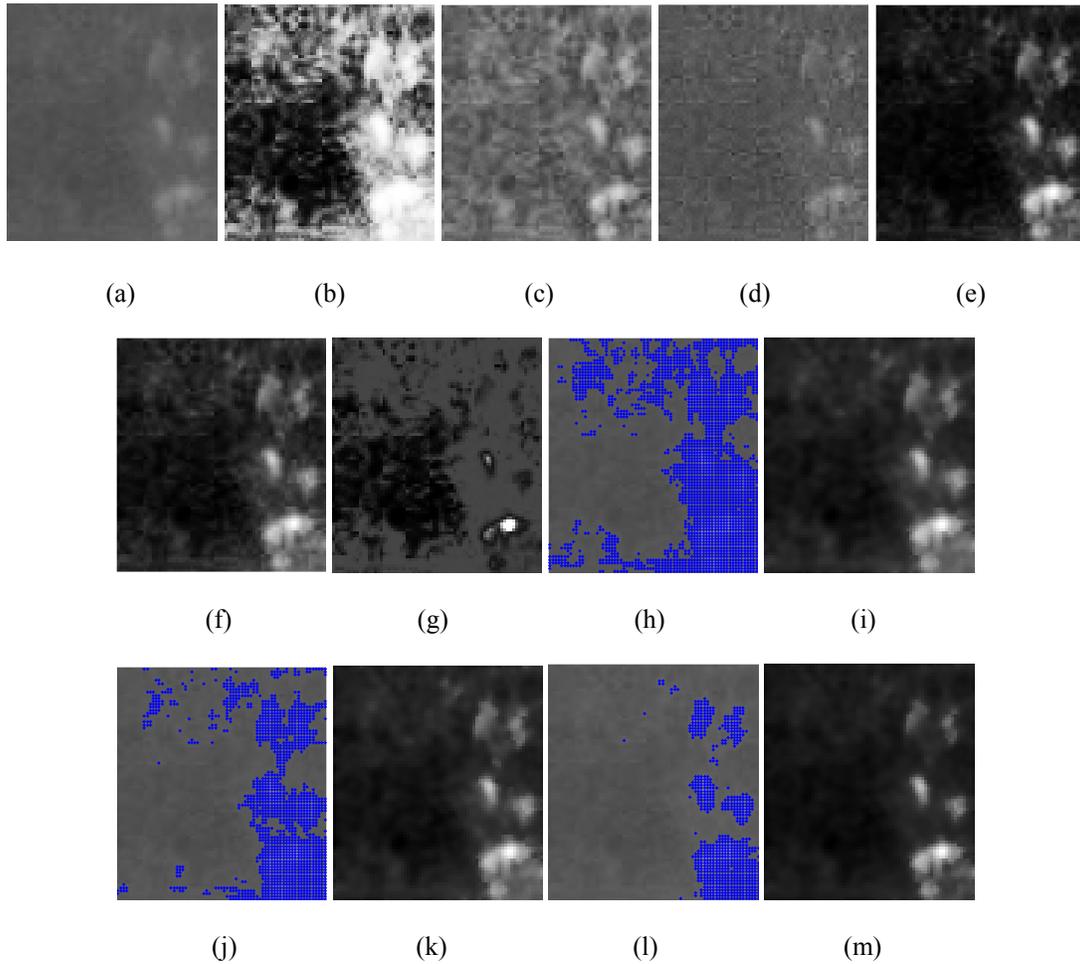
Figure 4.17: Detected positives and enhanced mammograms. (a) Original mammogram with micro-calcifications; (b) enhanced mammogram with standard histogram equalization; (c) enhanced mammogram with CLAHE; (d) enhanced mammogram with unsharp masking; (e) enhanced mammogram with fuzzy histogram hyperbolization (S-functionand); (f) enhanced mammogram with fuzzy histogram hyperbolization (norma-function); (g) enhanced mammogram with fuzzy logic possibility distribution method; (h) detected positives with GBAD; (i) enhanced mammogram based on GBAD; (j) detected positives with GGD; (k) enhanced mammogram based on GGD; (l) detected positives with iterative detection; (m) enhanced mammogram based on the iterative detection. Detected positives are marked using (blue) dots. P_F is equal to 0.01.

From Fig. 4.17, we can see that with the refinement of the detectors ((h), (j), (l)), all the micro-calcifications and located, the number of false positives are dramatically reduced, and better enhancements are achieved ((i), (k), (m)).

4.3.4.2   Quantitative Evaluation

The objective evaluation is carried out with quantitative metrics. Several existing quality metrics were employed but their evaluation results were inconsistent with the human subjective judgment.

For visualization and detection, a good enhancement method should be equipped with the following three properties:

(i) The contrast between the objects and background should be increased as much as possible, which can be represented by the increased difference between the mean gray value of the objects and the maximum value of the background.

(ii) The ratio of the dynamic gray value ranges between the objects and background should be increased as much as possible, which stands the object out and can be represented by the increased ratio of gray value variances between the objects and background.

(iii) The number of suspicious pixels, i.e. the pixels in the object and background having similar gray values, should be reduced as much as possible.

(iv) The average distance between the suspicious pixels in the object and background should be increased as much as possible.

(v) To avoid over-enhancement, the similarity between the enhanced image and the original one should be maintained to some extent, that is, the enhanced image should be similar as the original one as much as possible, which can be measured with the similarity metric [108].

Based on these criteria, a novel quality evaluation scheme is presented. We first calculate 5 quantities for a given image, corresponding to the 5 criteria or metrics mentioned above, and then normalize the 5 quantities into the range of [-1, 1], respectively. Finally, the summation of the 5 normalized quantities is employed as the grade of image quality. The higher the grade is, and the better the image quality will be.

Table 4.4 shows the evaluation results based on the combination of the 5 criteria. The enhancement algorithms include joint enhancer, iterative enhancer, iterative detection-based enhancer, standard histogram equalization, CLAHE, unsharp masking and fuzzy logic enhancement methods.

TABLE 4.4
ENHANCEMENT EVALUATION

Method	Joint-En	It-En	Det-En	Hist-Eq	CLAHE	Unsharp	FuzH-S	FuzH-N	Fuz-PD
Grade	3.0043	2.5025	4.3282	1.7538	1.9624	1.3907	2.4348	2.3452	2.1025

Joint-En: joint enhancer; It-En: iterative enhancer; Det-En: iterative detection-based enhancer; Hist-Eq: standard histogram equalization method; FuzH-S: fuzzy histogram hyperbolization with S-function; FuzH-N: fuzzy histogram hyperbolization with *norma*-function; Fuz-PD: fuzzy logic possibility distribution method.

From Table 4.4 we can see that the iterative detection-based enhancement method (Det-En) achieves the best performance, and the joint enhancement (Joint-En) and

iterative enhancement (It-En) methods also perform well. Fuzzy histogram hyperbolizeation methods yield good results, but their performances are highly parameter-dependent, which restricts their CAD applications.

4.4 Summary

Image enhancement plays a fundamentally important role in nearly all of the vision and image processing systems. In this chapter, we presented two novel image enhancement approaches. In the first approach, we developed an image enhancement system based on SR noise under the constrained MOO framework, for improving the suboptimal image enhancers which suffer from model mismatch and yield unsatisfactory enhancement results. The principle of SR noise-refined image enhancement was investigated, and a genetic algorithm-based MOO method was used to find the SR noise solution, in which the HVS preference was included via a weighting scheme for reducing the size of the solution set. Four types of SR noises were employed in the system and a number of enhancement algorithms were investigated in this chapter. Experimental results show that the presented system has highly encouraging performance in terms of simplicity, flexibility, efficiency and robustness, which demonstrates SR noise's capability of improving the suboptimal enhancers and supports its real-world CAD application.

In the second approach, we employ mammogram enhancement as an example to illustrate the efficiency of the presented selective enhancement framework. We have presented several schemes to automatically determine the ROI. ROI enhancement and background suppression were carried out by using the adaptive gray level stretching

technique. Experiments were based on real-world mammograms containing different types of lesions, and the comparison with several representative methods showed that the presented algorithms can achieve superior performance in terms of both subjective and objective evaluations.

CHAPTER V

HUMAN VISUAL SYSTEM-DRIVEN IMAGE SEGMENTATION

In this chapter, we present two human visual system (HVS)- driven image segmentation approaches. In the first approach, the quality metrics for evaluating the segmentation result based on human visual perception properties, from both region-based and boundary-based perspectives, are integrated into an objective function. The objective function encodes the HVS properties into a Markov random fields (MRF) framework, where the just-noticeable difference (JND) model is employed when calculating the difference of the image contents. In the second approach, we consider image segmentation as a detection problem and present a framework for image segmentation. In this framework, a "soft" segmentation objective function, in terms of the detection performance measured in local regions, is employed to guide the segmentation procedure. The human visual system information is incorporated into the segmentation procedure to improve the efficiency of the framework through the introduction of a contrast sensitivity function (*CSF*)-filtering operation in the wavelet domain.

5.1   *Introduction*

As briefly introduced in Section 1.3, image segmentation is a challenging problem, and lots of algorithms have been presented to deal with this task. Design of a suitable objective function is crucial to the performance of image segmentation approaches. Good segmentation algorithms require an efficient scheme for parameter adjustment and an

appropriate description of the desired properties of the segmentation result, which, of course, are all very challenging tasks. In real-world applications, the performance of some segmentation algorithms is influenced by their dependence on the parameters of these algorithms. But the optimum parameters and, therefore, satisfactory segmentation results are not easy to obtain. Some segmentation algorithms only partially incorporate the feature information from region and boundary perspectives, and fail to fully take advantage of fusing the two types of information. For example, Markov Chain Monte Carlo (MCMC) has been employed [124] to solve the maximum *a posteriori* probability (MAP)-MRF estimation problem for generative image segmentation. Due to many constraints involved in this generative approach, the selection of suitable parameters for satisfactory segmentation becomes difficult. It is also not easy for the objective function in [36] to yield a satisfactory balance between connecting the boundary and labeling the pixels, since there are many parameters which need to be chosen carefully. In [38], the proposed objective function does not exploit fully the connectivity property of the neighboring edge components. The normalized cut methods [125][126] can capture salient parts of an image. However, due to the ad hoc approximations introduced when relaxing this NP-hard computational problem, these methods do not exploit well the image content information which is useful for segmentation. As a result, the algorithms often perform unsatisfactorily.

Among all the existing image segmentation models, MRF models are very popular ones and have been used to represent contextual information in many pixel-based segmentation problems, because they can be employed to characterize the spatial dependency or spatial distribution. A statistical method, namely the MAP approach, is

often used during MRF-based image segmentation, which has been investigated comprehensively. The MAP-MRF method maximizes an objective function consisting of the *a priori* density in terms of the Gibbs distribution and the conditional probability density function (*pdf* for continuous data, and probability mass function, *pmf*, for discrete data) of the observed image data given the distribution of the segmented region, in which some image features are often embedded [127][128]. However, some strong assumptions and inaccurate estimates of the conditional *pdf* corresponding to intensity values of single pixels limit its performance and application.

Another weakness of many existing segmentation algorithms is that they are developed based on the information provided only by the image data and neglect the fact that the human is the best and usually the ultimate evaluator of the segmentation result. That is, these algorithms do not consider the impact of the HVS on object interpretation and information extraction. As a result, many algorithms are inconsistent with the preferences of human vision. There do exist efforts to incorporate HVS information into image segmentation, e.g., [126][127][128], but their performances were constrained by the simplistic computational models as well as an insufficient consideration of the HVS properties when designing the objective functions.

Our first effort is to develop a segmentation algorithm which takes into account HVS preference during the segmentation procedure and is also robust to the parameter configuration. We first introduce the HVS-driven image segmentation model under the MAP-MRF framework. Second, the criteria for evaluating region-based segmentation and the resulting energy function are discussed. Thirdly, the boundary-based evaluation criteria are discussed and encoded into the energy function via the development of a

novel concept, called boundary element in our work, which describes the interaction between pixel labels, boundary configuration and the image content. Then, we integrate the objective function that includes both region and boundary information, where the optimization method and the three variations of the HVS-driven segmentation algorithm are discussed. Experimental results and performance comparisons between the presented algorithms and other representative segmentation and clustering algorithms are presented in Section 5.2.5.

In our second segmentation approach described in Section 5.3, we attempt to address two problems mentioned above, i.e., the constraints of the conventional MRF-based algorithms and the lack of HVS information during the segmentation procedure. We consider image segmentation to be a detection problem, and present a novel image segmentation framework. We employ the probability of successful detection as a metric when designing the objective function to show the efficiency of this framework. A *CSF*, which takes into account HVS preference to the image content, is used as an object feature enhancer to further improve the segmentation performance. Experimental results are shown in Section 5.3.4.

5.2 A Human Visual System-Driven Image Segmentation Algorithm

5.2.1 MAP-MRF Framework

In this section, we develop our segmentation model under the MAP-MRF framework, which incorporates the information from both region-based and boundary-based

segmentation perspectives.

Under the MAP framework, image segmentation can be obtained by solving the following optimization problem

$$\left(\hat{L}, \hat{B}\right) = \arg\max_{L \in \Omega_L, B \in \Omega_B} P(L, B \mid Y) = \arg\max_{L \in \Omega_L, B \in \Omega_B} P(Y \mid L, B) P(L, B) \qquad (5.1)$$

where $P(L, B \mid Y)$ is the *a posteriori* distribution of the label field, L, and the boundary field, B, given the observed image, Y. L and B are assumed to have the MRF property, and they consist of pixel labels and boundary elements, respectively. The boundary elements will be defined in Section 5.2.3. Ω_L and Ω_B are the configuration spaces of L and B. $\Omega_L = \{l_1, l_2, \cdots, l_M\}$, where l_i is the label of the pixel with the index i. $l_i = \{0, 1, \cdots, A-1\}$, and A is the number of possible region types. For example, for binary segmentation, $A=2$. $\Omega_B = \{b_1, b_2, \cdots, b_M\}$, where b_i is the boundary element of the pixel with the index i. M is the total number of pixels in an image. A segmented image region is composed of the pixels with the same label. In this section, label-based segmentation is equivalent to region-based segmentation, and the two terms will be used interchangeably.

Thus, we obtain our segmentation model under the MAP-MRF framework with the region label MRF L and the boundary MRF B. In our work, the label field L and the boundary field B are defined as functions of the image data Y, that is, $L = L(Y)$ and $B = B(Y)$. $L(Y)$ and $B(Y)$ will be precisely defined in the next two subsections. Therefore, the likelihood term in (1.1) has the form

$$P(Y \mid L, B) = P(Y \mid L(Y), B(Y)) = 1 \qquad (5.2)$$

So (5.1) is reduced to

$$\left(\hat{L}, \hat{B}\right) = \underset{L \in \Omega_L, B \in \Omega_B}{\arg\max} \; P(L(Y), B(Y)) \tag{5.3}$$

Since both L and B have been assumed to exhibit MRF properties, according to the Hammersley-Clifford theorem [127], they can be represented in terms of the Gibbs distribution and the optimization problem of (5.3) can be written as

$$\left(\hat{L}, \hat{B}\right) = \underset{L \in \Omega_L, B \in \Omega_B}{\arg\max} \; \frac{1}{Z} \exp\{-U(L(Y), B(Y))\} \tag{5.4}$$

where $U(L(Y), B(Y))$ is the energy function, denoting the interaction between label and boundary configurations as well as the observation. Z is included for normalization and is a function of the MRF parameters. For given MRF parameters, (5.4) is equivalent to

$$\left(\hat{L}, \hat{B}\right) = \underset{L \in \Omega_L, B \in \Omega_B}{\arg\min} \; U(L(Y), B(Y)) \tag{5.5}$$

The energy function $U(L(Y), B(Y))$ consists of two factors, corresponding to region-based and boundary-based segmentations. Here, as in some prior work [24][44], we express the energy function in the following additive form

$$U(L(Y), B(Y)) = U_L(L(Y)) + U_B(B(Y)) \tag{5.6}$$

where the energy functions $U_L(L(Y))$ and $U_B(B(Y))$ can be considered as the quality metrics corresponding to region- and boundary-based evaluations, respectively. We want to emphasize that L and B here are two different aspects of the same segmentation result, and the corresponding metrics are the complementary evaluations for the same segmentation result from region and boundary perspectives, respectively.

5.2.2 Energy Function for Region-Based Image Segmentation

In this section, the energy function $U_L(L(Y))$ in (5.6) corresponding to region-based segmentation is developed. The approach is based on human preference for good segmentation, from a region-based segmentation perspective.

We note that a human often evaluates the segmentation result in both global and local manners, that is, the fitness of a segmentation result to the entire image content and the local image region are considered simultaneously. Therefore, both region-based and boundary-based segmentation evaluations should be taken into account. In this subsection, we only consider region-based evaluation, and postpone the consideration of boundary-based evaluation to the next subsection. We summarize the desirable properties for good segmentation in terms of region-based evaluation as follows.

 (i) The contrast of pixel intensities between two neighboring regions, i.e., inter-region contrast, should be large;

 (ii) The contrast of pixel intensities within a region, i.e., intra-region contrast, should be small;

 (iii) The pixel labels should correspond to homogeneous regions, that is, neighboring pixels prefer having the same label.

Criteria (i) and (ii) represent the global properties of a good segmentation, and criterion (iii) is a local property which indicates that the segmentation should yield large-sized regions.

Thus, the region-based segmentation evaluation metric $U_L(L(Y))$ should consist of two types of measures, namely, global inter- and intra-region contrast measures and a local

label homogeneity measure. We express the composite measure also using an additive form as follows

$$U_L\big(L(Y)\big) = U_{Global}\big(L(Y)\big) + U_{Local}\big(L(Y)\big) \tag{5.7}$$

We can see from the three desirable properties that the contrasts between the neighboring regions and those between the neighboring pixels need to be calculated in order to obtain a quantitative metric for evaluating segmentation quality. In this work, we incorporate the HVS properties into the contrast measure via the JND model [129][130][131]. HVS is capable of only perceiving pixel intensity changes above a certain visibility threshold, which, in turn, is determined by the underlying physiological and psychophysical mechanisms. JND refers to the minimum visibility threshold above which visual contents can be distinguished. The JND model plays an important role in perceptual image and video processing, and has been successfully used in measuring the difference or distortion of the image contents [132][133]. In this chapter, we use the spatial JND model, i.e., pixel-wise JND, presented in [134], which is defined as a nonlinear additive model,

$$\text{JND}_p(i) = T^L(i) + T^t(i) - C^{L,t}(i) \cdot \min\big\{T^L(i), T^t(i)\big\} \tag{5.8}$$

where $\text{JND}_p(i)$ is the JND threshold of the pixel indexed by i. $T^L(i)$ and $T^t(i)$ are the visibility thresholds due to luminance adaptation and texture masking, respectively, and $C^{L,t}(i)$ represents the overlapping effect in masking where $0 < C^{L,t}(i) \leq 1$. Details on the definition of JND and its computation are available in [129][134].

5.2.2.1 Energy Function for Global Contrast

In order to incorporate the desired characteristic of global inter- and intra-region contrasts into the energy function, we define a global neighborhood system, 61 pixels by 61 pixels in this chapter, with pixel of interest at the center. The segmentation quality metric based on criteria (i) and (ii) can be expressed as

$$Q_{Global} = a_1 \cdot C_{Inter} - a_2 \cdot C_{Intra} \tag{5.9}$$

where C_{Inter} and C_{Intra} denote the inter- and intra-region contrasts of the pixel intensity based on JND, respectively, and a_1 and a_2 are two non-negative weights that control the contributions of the two types of contrasts to the energy function. A higher value of Q_{Global} means a better segmentation quality.

Due to the Markov property assumed in our work, the quality metric value corresponding to a single pixel, s, is independent on other pixels given the segmentation result in the global neighborhood system of s. Therefore, we have the energy function for the global contrast as

$$U_{Global}(L(Y)) = -\sum_{s=1}^{M} Q_{Global-s|\Gamma^s}(s) = -\sum_{s=1}^{M} \left[a_1 \cdot C_{Inter-s|\Gamma^s}(s) - a_2 \cdot C_{Intra-s|\Gamma^s}(s) \right] \tag{5.10}$$

where $s \mid \Gamma^s$ means that the contrast is a function of the label of the pixel s and is calculated given that the labels of the rest of the pixels in the global neighborhood system of s, Γ^s, are fixed. To speed up the computation, instead of calculating the contrasts in a pixel-wise manner, the following region-wise measure of the inter-region contrast is employed,

$$C_{Inter-s|\Gamma^s}(s) = \varphi \left\{ \Xi_{s|\Gamma^s}(\mu_1, \mu_0) \right\} \tag{5.11}$$

where the calculation is carried out on $\Gamma^s \cup \{s\}$, the set consisting of the pixels included

in Γ^s plus pixel s, and " $\cup$ " means "union". Here,

$\Xi_{s|\Gamma^s}(\mu_1, \mu_0) = |\mu_1 - \mu_0| / \min[\mathrm{JND}_R(1), \mathrm{JND}_R(0)]$. $\mu_u = \sum_{i=1, \text{ and } l_i=u}^{N_u} y_i / N_u(s)$, which is the mean

value of the pixel intensity in the uth type region, $u \in \{0,1\}$ for binary segmentation,

and $N_u(s)$ is the number of pixels with the label u. y_i is the ith pixel with the label u.

$\mathrm{JND}_R(u) = \sum_{i=1, \text{ and } l_i=u}^{N_u} \mathrm{JND}_p(i) / N_u(s)$, which is the average JND value of the regions with

the label u. Thus, $\Xi_{s|\Gamma^s}(\mu_1, \mu_0)$ is the measure of the average intensity difference of the two

types of regions weighted by the minimum of the average JND values of the two types of

regions. $\varphi(D)$ is a robust function that reduces the impact of outliers and is defined in [33]

as

$$\varphi(D) = \frac{D^4}{G + D^4} \tag{5.12}$$

where G is a small positive constant.

For the intra-region contrast, we define and employ the metric,

$$C_{Intra-s|\Gamma^s}(s) = \frac{1}{N_1(s) + N_2(s)} \left\{ \Phi_{s|\Gamma^s}(0) + \Phi_{s|\Gamma^s}(1) \right\} \tag{5.13}$$

Still, the calculation is carried out on $\Gamma^s \cup \{s\}$, where $\Phi_{s|\Gamma^s}(0) = \sum_{i=1 \text{ and } l_i=0}^{N_0} \varphi\left(|y_i - \mu_0| / \mathrm{JND}_p(i)\right)$

and $\Phi_{s|\Gamma^s}(1) = \sum_{i=1 \text{ and } l_i=1}^{N_1} \varphi\left(|y_i - \mu_1| / \mathrm{JND}_p(i)\right)$. Here, $|y_i - \mu_j| / \mathrm{JND}_p(i)$, $j=0$ and 1 for binary

segmentation, is the weighted difference between the intensity of the ith pixel and the

average intensity of the region to which the pixel belongs. So $\Phi_{s|\Gamma^s}(0)$ and $\Phi_{s|\Gamma^s}(1)$ in (5.13)

measure the "variation", or the inhomogeneity, of the two types of regions.

We can see that (5.11) and (5.13) measure how the segmentation result of a single pixel s, i.e., the label of s, affects the segmentation in a global manner, and take into account the inter- and intra-region contrasts at the same time. Therefore, the energy function defined in (5.10) reduces the risk of being too biased when segmenting an image [125]. For example, an algorithm may become very greedy in finding small but tight clusters in the image data if only intra-region contrast is considered [125]. Additionally, (5.11) and (5.13) do not require modeling of the probability distributions of the noise and the image data, which may improve the robustness of the presented algorithm.

5.2.2.2   *Energy Function for Local Homogeneity*

In this subsection, we discuss the energy function based on criterion (iii), incorporating local homogeneity for good segmentation. A second-order neighborhood system, that includes 8 nearest neighboring pixels of the pixel of interest s, is employed for describing local homogeneity. The pixels, excluding s, included in the solid (yellow) rectangular shown in Fig. 5.1 (b) form a configuration of the second-order neighborhood system of s. Besides the label homogeneity, we also incorporate information about the dynamics of the pixel intensity in order to make the metric adaptive to non-stationary image contents. The energy function is defined as

$$U_{Local}\left(L(Y)\right) = -\sum_{s=1}^{M}\left\{\sum_{\eta_s \in NB_2(s)} \psi(l_s, l_{\eta_s}) \exp\left[-\varphi\left(\Delta_{s,\eta_s}(y)\right)\right]\right\} \tag{5.14}$$

where $\Delta_{s,\eta_s}(y) = |y_s - y_{\eta_s}| / \min[\mathrm{JND}_P(s), \mathrm{JND}_P(\eta_s)]$ represents the JND-weighted contrast between pixel s and its second-order neighbor η_s. $NB_2(s)$ denotes the set of all of the

second-order neighbors of s. The cost function for the label configuration of the neighboring pixel pair is denoted by $\psi(l_s, l_{\eta_s})$ and defined in (5.15) for binary segmentation.

$$\psi(l_s, l_{\eta_s}) = \begin{cases} \beta, & \text{if } l_s = l_{\eta_s}, \\ -\beta, & \text{otherwise.} \end{cases} \tag{5.15}$$

where β is a non-negative real number and is the cost used to define the label homogeneity measure of the neighboring pixel pair. Eq. (5.14) has a form similar to the generalized Potts model [135], except for the additional robust estimation shown in (5.12) and the JND-weighting operation.

5.2.3 Energy Function for Boundary-Based Image Segmentation

In this subsection, we develop the energy function for boundary-based segmentation, which is also derived from the desirable properties for good segmentation in terms of a boundary-based evaluation. These properties are listed below,

(i) Region boundary should be smooth and of as small a length as possible. In other words, the boundary should avoid containing too many sharp angles or turns;

(ii) The intensity contrast of a neighboring pixel pair on the two sides of the boundary should be large, while the contrast within a region enclosed by the boundary curve should be small;

(iii) The pixels lying on the boundary curve should be connected.

Criterion (i) represents a property which is characterized by the image contents in both global and local manners. Criteria (ii) and (iii) are properties of a good segmentation in small regions, and can be measured locally.

When designing the energy function $U_B(B(Y))$, due to the huge computational burden for the global boundary-based feature measure, we only employ the local properties described in the boundary-based segmentation evaluation. That is,

$$U_B(B(Y)) = U_{Local}(B(Y)) \qquad (5.16)$$

Before deriving the energy function, we first discuss the neighborhood system and define a novel concept called the boundary element in the next section.

5.2.3.1   Element and Neighborhood System

The energy function of boundary-based segmentation is calculated based on a novel concept, the boundary element, defined for each boundary pixel. A boundary pixel is a pixel for which at least one of its second-order neighbors has a label different from it. A boundary element consists of an angle together with its two directed edges. Suppose s is a boundary pixel, the angle of the boundary element of s originates from s, and two edges of the angle point to the two neighboring boundary pixels of s which have the same labels as s. The two edges separate the pixels into different regions according to whether their labels are different from or same as that of s. Some examples of the angles are shown in Fig. 5.2. The value of the angle is constrained to be in the range $[0, \pi]$, and is related to the smoothness of the boundary curve. Small angles correspond to sharp turns and therefore to a wiggly boundary, while large angles correspond to a smooth region

boundary. Naturally, we prefer large angles. The edges of the boundary element are related to the connectedness of the boundary pixels and thus the continuation of the boundary contour. The edges also play an important role in determining the cross-boundary pixel pair, which will be discussed in Section 5.2.3.2.1. The boundary elements for two special segmentation configurations, isolated segmentation and interior pixel segmentation, will be discussed in Section 5.2.3.2.2, where the pixel of interest has either a different label from or the same label as its second-order neighbors.

A boundary element is determined from the interaction between the pixel of interest, say, s, and its 8 second-order neighbors. The total energy of a neighborhood system centered at s is dependent on the boundary elements of s and its 8 neighbors. A change in the label of s may affect the boundary elements of its 8 neighbors. Therefore, we define the neighborhood system for boundary-based segmentation as one consisting of the 8 nearest neighbors in the second-order neighborhood system of s plus all the second-order neighbors of these 8 pixels, excluding s. This is, in fact, a third-order neighborhood system of s, $NB_3(s)$, i.e., $NB_3(s) = \left\{ \cup NB_2(\eta_s) \right\} \cup \left\{ NB_2(s) \right\} \backslash s$, where " $\backslash s$ " means "excluding s". $NB_3(s)$ is shown in Fig. 5.1, in which the pixels, excluding s, included in the rectangular bold dot-dash (light blue) line constitute the third-order neighborhood system of s for boundary-based segmentation. In Fig. 5.1, s lies at the center of the neighborhood system. The pixels represented by the solid points belong to one type of region with label zero (0 region). The small circles represent pixels belonging to the region with label one (1 region). The thin solid (green) line and the bold solid (black) line denote the boundary curves of the 0 and 1 regions, respectively. The solid (red) directed edges, paired together with the angle between the edge pair, form the boundary elements

of the pixels in the 0 region. The dashed (black) directed edge pairs and the corresponding angles form the boundary elements of the pixels in the 1 region. Now, let us consider the situation when the label of a pixel changes. The impact of the label change of *s* on the boundary elements of its 8 neighbors is shown in Fig. 5.1 (b), where *s* changes its label from 0 (Fig. 5.1 (a)) to 1 (Fig. 5.1 (b)) and results in the change of the boundary elements of its second-order neighbors.

(a)   (b)
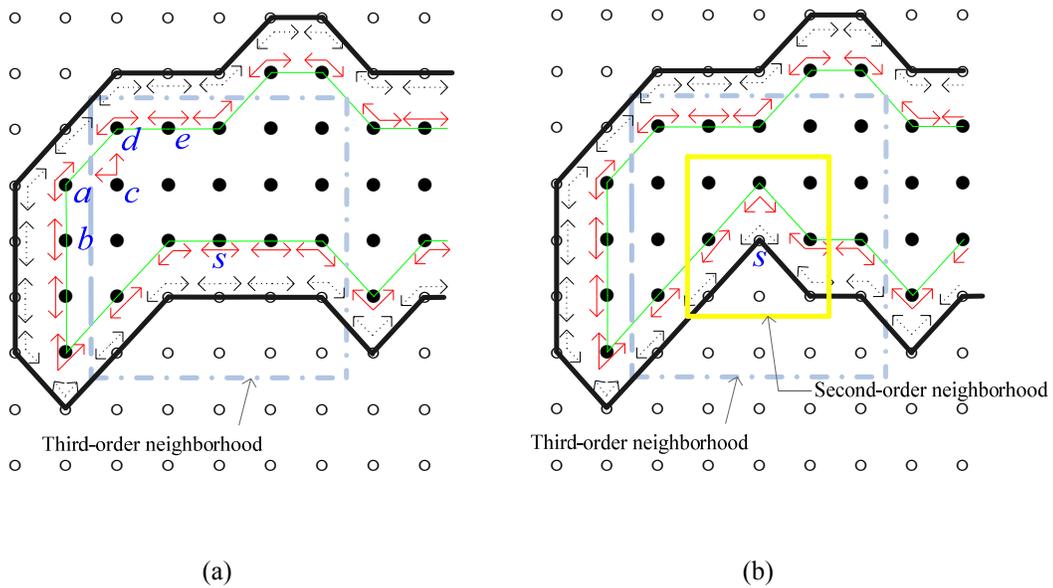
Figure 5.1: Boundary elements, second-order and third-order neighborhood systems. (a) Typical boundary elements and third-order neighborhood system; (b) impact of label change of *s* on the boundary elements of its second-order neighbors.

From Fig. 5.1, we can see that the boundary curve of a region is determined by the boundary pixels together with their edge pairs. However, not all the boundary pixels contribute directly to the boundary curve, and it is possible that the impact of some

boundary pixels is "hidden" by the neighboring boundary pixels when forming the curve. For example, pixel c in Fig. 5.1 (a) is a boundary pixel, but the boundary curve, the thin solid (green) line, does not pass through c and its two edges, $\vec{ca}$ and $\vec{cd}$. This is because the boundary elements a and d together with their edges, $\vec{ab}$, $\vec{ad}$ and $\vec{da}$, $\vec{de}$, hide the contribution of c to the curve. However, the change in the label of c will affect the boundary elements of a and d, as mentioned above. Thus, we still need to consider the boundary element and the energy corresponding to pixel c when updating the pixel labels, the parameters of the MRF and therefore the shape of the boundary curve. In other words, boundary pixel c impacts the boundary curve in an indirect or implicit manner. The energy function corresponding to boundary elements will be discussed in the next section.

From Fig. 5.1, we notice that the angles of the boundary elements reflect the variation in the boundary shape, and the directed edges represent the interaction and the relative locations of the neighboring boundary pixels.

5.2.3.2   *Energy Function for Boundary-Based Segmentation*

As indicated earlier, we prefer smooth boundaries, i.e., gentle "turns". This corresponds to large angles, and the energy function is, thus, designed as a monotonically decreasing function of the angle value. At the same time, a "reasonable" turn, which results from the significant contrast of pixel intensities across the boundary, should also be maintained. Furthermore, the intensity contrast of pixels on the same side of the boundary and belonging to the same region should be small. In this way, we may make the smoothness measure of the boundary curve dynamic and adaptive to non-stationary image content.

Thus, the energy function of a boundary element is composed of three terms: the first one is related to the angle between the two edges, the second one is related to the intensity contrast across the boundary, and the third one is related to the contrast on the same side of the boundary which is called intra-pie slice contrast. We define a "pie slice" next.

5.2.3.2.1   Cross-Boundary Contrast and Intra-Pie Slice Contrast

Consider a boundary pixel, s. A pie-slice of s, by definition, consists of s, as the origin, and some other pixels in the second-order neighborhood system of s. These pixels must have the same labels as s, and there are no pixels with different labels from s in the pie slice. The cross-boundary contrast is calculated as the intensity difference between pixels in a pie slice of s and the pixels in the second-order neighborhood system of s but with labels different from s. All possible configurations of the single pie slice in a second-order neighborhood system are shown in Fig. 5.2. The configuration in each figure is valid when rotated by $\pi / 2$. The configurations of the multi-pie slices, where a second-order neighborhood system contains more than one pie slice, can be determined in a similar manner and an example of it is shown in Fig. 5.3.

In Fig. 5.2, s is the pixel of interest. Crosses "×" represent the pixels with the same label as s (including s), and form a pie slice of s. The remaining pixels have labels different from s, and are represented by "+", "-" and "*". The cross-boundary contrast of a pie slice is determined by averaging the contrast of the so-called cross-boundary pixel pairs, which are determined by the two edges of the boundary element. More specifically, in the figure the cross-boundary pixel pairs corresponding to the pixels represented by "-" and "×" are determined by the dashed (red) edge, those corresponding to the pixels

represented by "+" and "×" are determined by the solid (blue) edge, and those corresponding to the pixels represented by "*" and "×" are determined by both of the two types of edges mentioned above. The cross-boundary pixel pair configuration in each figure is valid when rotated by $\pi/2$.

$$
\begin{array}{ccc}
+ \quad + \quad + & \times \quad + \quad + & \times_t \quad \times \quad + \\
\times \leftarrow \times \quad * & \times \quad \times \quad + & \times \quad \times \quad + \\
- \quad - \quad - & - \quad - \quad * & \overline{P_{2\text{-}2}} \quad \overline{P_{s\text{-}3}} \quad * \\
\end{array}
$$

Figure symbols — (a) (b) (c) (d) (e); (f) (g) (h) (i) (j); (k) (l) (m) (n)
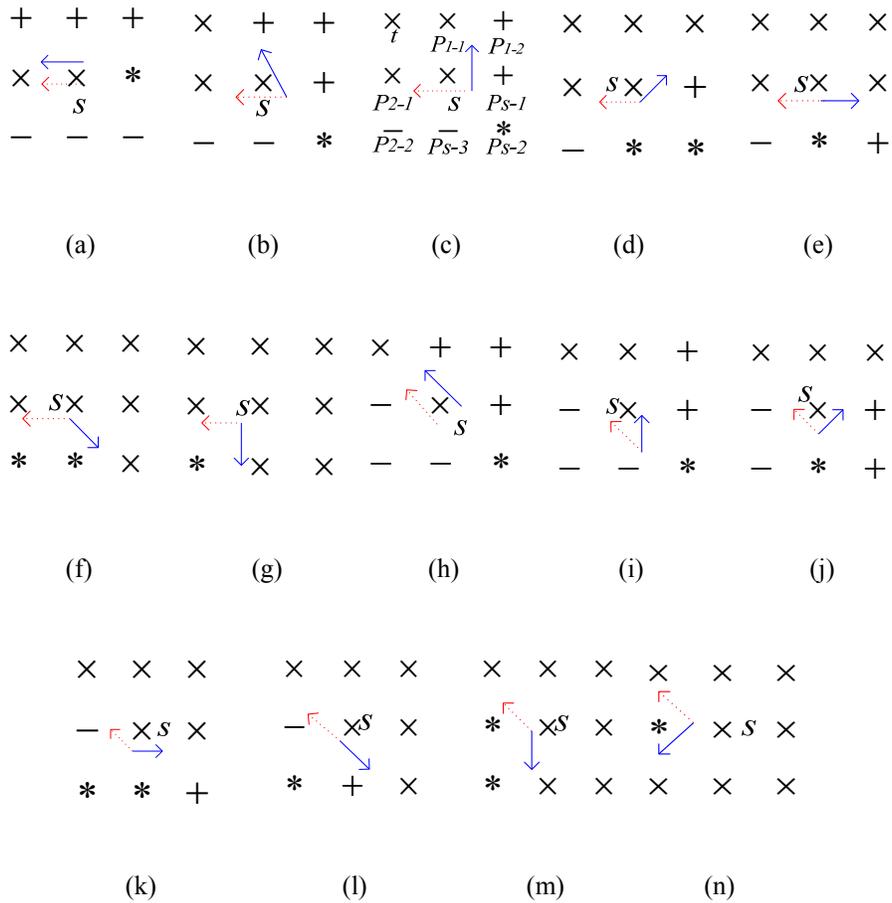
Figure 5.2: Possible configurations of the single pie slice in a second-order neighborhood system. The configuration in each figure is valid when rotated by $\pi/2$.

129

Suppose there are $N_{PS}(s)$ pie slices included in a second-order neighborhood system with the origin at s. The cross-boundary contrast of the ith pie slice consists of two terms, one including s and another without s, as shown in (5.17)

$$Contr\ (s)_{CB-i} = \frac{1}{N_r + N_j}\left[\sum_{r=1}^{N_r}\varphi\left(\frac{\left|y_{P_{r-1}} - y_{P_{r-2}}\right|}{\min[\ \text{JND}\ _P(P_{r-1}),\ \text{JND}\ _P(P_{r-2})]}\right) + \sum_{j=1}^{N_j}\varphi\left(\frac{\left|y_s - y_{P_{s-j}}\right|}{\min[\ \text{JND}\ _P(s),\ \text{JND}\ _P(P_{s-j})]}\right)\right] \quad (5.17)$$

where the first summation is related to the cross-boundary contrast calculated from the cross-boundary pixel pairs corresponding to the pixels in the ith pie slice of s (excluding s), and the second summation is related to that corresponding to s. Here, y_k denotes the intensity value of pixel k. P_{r-1} and P_{r-2} denote the boundary pixels inside and outside the ith pie slice of s (excluding s). They form the rth cross-boundary pixel pair, in which P_{r-2} is the closest neighboring boundary pixel of P_{r-1} and has a label different from P_{r-1}, and N_r denotes the number of such pairs. Here, $i = 1,...,N_{PS}(s)$ with $N_{PS}(s) \le 4$, where $N_{PS}(s)$ denotes the number of pie slices in the second-order neighborhood system of s and 4 is the maximum number of pie slices in a second-order neighborhood system. P_{s-j} are the boundary pixels outside the ith pie slice and have the labels different from s. s and P_{s-j} form the jth cross-boundary pixel pair, and N_j denotes the number of such pairs. By annotating the pixel in Fig. 5.2 (c), we show an example for calculating the cross-boundary contrast using (5.17), where the four "×" pixels form a pie slice and the cross-boundary pixel pairs are determined by the solid and dashed edges. Here, pixel t is not involved in the calculation since it is not a boundary pixel when we calculate the cross-boundary contrast of s. As mentioned before, in this case there is only one pie slice in the second-order neighborhood system.

130

The intra-pie slice contrast of the ith pie slice is defined as

$$Contr(s)_{PS-i} = \frac{1}{N_{PS-i}(s)} \sum_{j=1}^{N_{PS-i}(s)} \varphi\left(\frac{\left|y_{i-j} - \mu_{PS-i}\right|}{JND_P(j)}\right) \qquad (5.18)$$

where $N_{PS-i}(s)$ is the number of pixels included in the ith pie slice in the second-order neighborhood system of s. μ_{PS-i} is the average intensity value of the pixels in the ith pie slice, and y_{i-j} is the observed pixel intensity of the jth pixel in the ith pie slice.

5.2.3.2.2   Energy Function for the Second-Order Neighborhood System

In this subsection, we develop the energy function for the second-order neighborhood system of the pixel of interest.

As mentioned before, the energy function is designed to encourage a large turn angle and large cross-boundary contrast. At the same time, we would like to have the contrast of the pixel intensity within each pie slice to be as small as possible. We first find all the pie slices in each second-order neighborhood system of the pixel of interest s, and then the contrasts across boundary and inside the pie slice are determined by (5.17) and (5.18) for each pie slice. The energy corresponding to a second-order neighborhood system can then be calculated as the summation of the energy of each pie slice,

$$\phi(s) = \sum_{i=1}^{N_{PS}(s)} \exp\left[-a_3 \cdot \alpha_i(s) - a_4 \cdot Contr(s)_{CB-i} + a_5 \cdot Contr(s)_{PS-i}\right] \qquad (5.19)$$

where $\alpha_i(s)$, $Contr(s)_{CB-i}$ and $Contr(s)_{PS-i}$ are the angle value, cross-boundary contrast and intra-pie slice contrast of the ith pie slice, respectively. The weights a_3, a_4 and a_5 are non-negative real numbers.

Fig. 5.3 shows an example for calculating the energy of a multi-pie slice segmentation, where there exist two pie slices Δasb and $\Delta csed$. The corresponding energy is

$$\phi(s) = \exp\{-a_3 \cdot \alpha_{\Delta asb} - a_4 \cdot Contr(s)_{CB-\Delta asb} + a_5 \cdot Contr_{PS-\Delta asb}(s)\}$$
$$+ \exp\{-a_3 \cdot \alpha_{\Delta csed} - a_4 \cdot Contr(s)_{CB-\Delta csed} + a_5 \cdot Contr_{PS-\Delta csed}(s)\} \qquad (5.20)$$
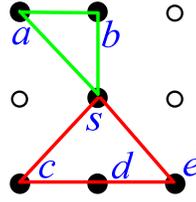
Figure 5.3: A typical multi-pie slice configuration, where there exist two pie slices Δasb and $\Delta csed$.

We can see that the multi-pie slice configurations, which correspond to complicated and less preferred segmentation, have more terms than the single pie slice case. The larger the $N_{PS}(s)$, the more "messy" the segmentation is. We show in APPENDIX A that the energy function (5.19) assigns higher energy and therefore lower probability to the multi-pie slice configurations, which is consistent with our expectation.

In many practical situations, there exist two special segmentation configurations, isolated segmentation and interior pixel segmentation, as mentioned before. Isolated segmentation corresponds to a special boundary element, where the angle of the isolated segmentation is -2π and the cross-boundary contrast is determined by averaging the JND-weighted intensity differences between s and its 8 second-order neighbors (with the robust function considered). The intensity contrast within the pie slice for the isolated

segmentation is zero since the pie slice includes only one pixel, s itself. Admittedly, the interior pixel is not a boundary pixel and does not have the boundary element as defined. However, for updating the pixel label we still assign a special boundary element to this type of configuration, where the angle of the interior pixel is set to be 2π and the corresponding cross-boundary contrast is zero. The intensity contrast within the pie slice can be calculated using (5.18), where s and its 8 neighboring pixels included in the second-order neighborhood system construct a pie slice.

5.2.3.2.3 Energy Function for Boundary-Based Segmentation

As mentioned before, a change in the label of s might impact the boundary elements and therefore the energy function values of the 8 second-order neighborhood systems centered at the 8 second-order neighbors of s, plus the second-order neighborhood system of s. Let $NB_{2+s}(s)$ denote the set of pixels in the second-order neighborhood system of s plus s itself, so $NB_{2+s}(s) = NB_2(s) \cup \{s\}$, where $NB_2(s)$ represents the 8 second-order neighbors of s, as defined in Section 5.2.2.2. Thus, the energy function values of the second-order neighborhood systems of the pixels in $NB_{2+s}(s)$ are necessary to calculate the energy corresponding to the label configuration of s. Therefore, energy function (5.16) can be expressed as

$$U_B\big(B(Y)\big) = \sum_{s=1}^{M} \left[\frac{1}{N_H(s)} \sum_{h=1}^{N_H(s)} \phi(h) \right] \tag{5.21}$$

where $\sum_{h=1}^{N_H(s)} \phi(h) / N_H(s)$ is the average energy of the second-order neighborhood systems included in the third-order neighborhood system of s. This energy function takes into

account the impact of the change in the label of s on the local region. The pixels in $NB_{2+s}(s)$ are indexed by h, and $N_H(s)$ denotes the number of pixels in $NB_{2+s}(s)$. We can see that $N_H(s) = 9$ if s is an internal pixel of an image.

To speed up the computation, only the energy function of the second-order neighborhood system of s is calculated in this chapter, which is sufficient to produce satisfactory results and is therefore implemented in the experiments. Thus, we have a simplified energy function for boundary-based segmentation,

$$U_B(B(Y)) = \sum_{s=1}^{M} \phi(s) \tag{5.22}$$

5.2.4 Overall Energy Function and Optimization

Having defined the energy functions for region-based and boundary-based segmentations, we formulate the overall optimization problem. According to the previous discussion, the image segmentation problem can be transformed to an optimization problem with respect to the pixel label and region boundary configurations,

$$
\begin{aligned}
(\hat{L}, \hat{B}) &= \underset{L \in \Omega_L, B \in \Omega_B}{\arg\min} U(L(Y), B(Y)) = \underset{L \in \Omega_L, B \in \Omega_B}{\arg\min} \{U_L(L(Y)) + U_B(B(Y))\} = \underset{L \in \Omega_L, B \in \Omega_B}{\arg\min} \{U_{Global}(L(Y)) + U_{Local}(L(Y)) + U_{Local}(B(Y))\} \\
&= \underset{L \in \Omega_L, B \in \Omega_B}{\arg\min} \left\{
\begin{array}{l}
-\sum_{s=1}^{M}\left(a_1 \cdot \varphi\left[\Xi_{s|\Gamma^s}(\mu_1, \mu_0)\right] - a_2 \cdot \dfrac{1}{N_1(s) + N_2(s)}\left[\Phi_{s|\Gamma^s}(0) + \Phi_{s|\Gamma^s}(1)\right]\right) \\
-\gamma_1 \sum_{s=1}^{M}\left[\sum_{\eta_s \in NB_2(s)} \psi(l_s, l_{\eta_s}) \exp\left[-\varphi\left(\Delta_{s,\eta_s}(y)\right)\right]\right] \\
+\gamma_2 \sum_{s=1}^{M}\left[\sum_{i=1}^{N_{PS}(s)} \exp\left(-a_3 \cdot \alpha_i(s) - a_4 \cdot Contr(s)_{CB-i} + a_5 \cdot Contr(s)_{PS-i}\right)\right]
\end{array}
\right\}
\end{aligned}
\tag{5.23}
$$

where γ_1 and γ_2 are two hyper-parameters controlling the contributions of the label and the boundary energy functions to the total energy. Finding the optimal configuration of L and B includes the minimization of (5.23) with respect to L and B followed by the

maximum likelihood estimation (MLE) of β if γ_1 and γ_2 are given. The procedure proceeds in an iterative manner. The JND weights only need to be calculated once before segmentation. We call this scheme HVS-driven segmentation scheme 1 (HDSS-1).

The hyper-parameters γ_1 and γ_2 can be determined either by trial-and-error or by using the method presented in [45]. In our work, we use the trial-and-error method to choose these parameters and the experimental results show that the performance of the algorithm is not sensitive to these parameters. To further reduce the effort of choosing the hyper-parameters, the optimization of the energy function is carried out in two steps. In the first step, only the energy terms related to region-based segmentation are minimized. In the second step, the boundary energy function is minimized which is based on the segmentation result from the first step. Steps 1 and 2 are iterated until we achieve a satisfactory result. (In our work, the number of iterations is determined by trial-and-error, but the experiments show that the performance of the algorithm is not sensitive to the number of iterations if the number is larger than a certain value.) We see that in this scheme only the hyper-parameter γ_1 needs to be chosen. We call this scheme the HVS-driven segmentation scheme 2 (HDSS-2). The two-step procedure is given by

$$
\text{Step 1:} \quad \hat{L} = \underset{L \in \Omega_L}{\arg\min} \left\{ \begin{array}{l} -\sum_{s=1}^{M} \left(a_1 \cdot \varphi\!\left[\Xi_{s|\Gamma^s}(\mu_1,\mu_0)\right] - a_2 \cdot \frac{1}{N_1(s)+N_2(s)}\left[\Phi_{s|\Gamma^s}(0)+\Phi_{s|\Gamma^s}(1)\right] \right) \\ -\gamma_1 \sum_{s=1}^{M}\left[\sum_{\eta_s \in NB_2(s)} \psi(l_s,l_{\eta_s})\exp\!\left[-\varphi\!\left(\Delta_{s,\eta_s}(y)\right)\right]\right] \end{array} \right\}
$$

$$
\text{Step 2:} \quad \hat{B} = \underset{B \in \Omega_B}{\arg\min}\left\{ \sum_{s=1}^{M}\left[\sum_{i=1}^{N_{PS}(s)} \exp\!\left(-a_3 \cdot \alpha_i(s) - a_4 \cdot Contr(s)_{CB-i} + a_5 \cdot Contr(s)_{PS-i}\right)\right]\right\} \quad (5.24)
$$

In our experiments, besides the implementation of the above mentioned segmentation schemes, we will also present the results when only the boundary element is considered

135

during the optimization, that is, when only the energy function in Step 2 of (5.24) is optimized. We call this segmentation method the boundary element-based segmentation (BEBS).

To the best of our knowledge, HDSS-1, HDSS-2 and BEBS are distinctive from existing segmentation algorithms in terms of the design of the objective functions from region- and boundary-based perspectives.

5.2.5 Experimental Results

In this subsection, comparative results of the segmentation of two types of medical datasets, mammogram and MRI brain images, and one natural image are shown. The comparisons are carried out between the presented algorithms and several representative segmentation and clustering algorithms.

5.2.5.1 Experiment Configuration

The mammogram used in the experiment is from the Digital Database for Screening Mammography (DDSM) [91]. The MRI brain image is T1-weighted and is from SUNY Upstate Medical University. We use Cameraman image as an example of the natural image.

In mammogram segmentation, our goal is to find the pixels which represent lesions. The segmented positive pixels by the presented algorithms are marked in blue. In segmenting the MRI brain image, we wish to segment the white matter (WM) from the gray matter (GM) and cerebrospinal fluid (CSF). The segmented non-WM tissues are

shown using purple and black colors. For the Cameraman image, we want to segment out the man as the foreground from the background, i.e., the building and the meadow. We also treat the camera as a background object and do not attempt to segment it out. The segmentation result is represented by a binary image, where the dark part is the foreground.

5.2.5.2   *Segmentation Algorithms*

The algorithms used for comparison purposes are briefly described as follows.

The Gaussian assumption-based dynamic clustering algorithm (GADC) [81], as described in Section 2.2.1.3, assumes that both background and lesions obey Gaussian distributions, where the detection (or clustering) and parameter update are performed in an iterative manner.

IMS [81], as described in Section 2.2.4, is an unsupervised learning pattern classification approach, which employs kernel density estimation technique to determine the probability distribution and performs clustering iteratively.

The HOSLW algorithm [96], as described in Section 2.3.2, has been shown to have superior performance compared with other existing methods for breast cancer detection in digital mammograms [97], in terms of efficiency and reliability.

 The conventional MRF-based algorithm [26] is also employed for comparison purpose, since the presented algorithms also include the MRF assumptions on the pixel label field and boundary element field. In our implementation, the likelihood term of the conventional MRF-based algorithm is determined under the Gaussian assumption and the

prior term is derived from the Gibbs distribution. An adaptive window size is employed to improve its performance, and the parameters of the Gibbs distribution are updated as the segmentation process proceeds.

Otsu thresholding [127][128] is a classical and effective method for image segmentation. It is widely used in software for MRI brain image segmentation, such as 3Dslice. It searches for the threshold by minimizing the intra-class variance. In our experiments, the original bi-level thresholding method [136] is employed for mammogram and Cameraman segmentations, where our goals are to find the lesion pixels and the Cameraman, respectively, while the multi-level version [137] is employed for the MRI brain image to segment out the WM.

The level set evolution-based method, LSEWRI, presented in [121] is a recent variational formulation. It forces the level set function to be close to a signed distance function, and does not need the costly re-initialization procedure.

The region-based active contour model (RACM) algorithm [138] is also based on level set evolution, which aims at overcoming the difficulties of segmentation due to inhomogeneous intensity. The authors employ a region-based active contour model which draws upon the intensity information in local regions at a controllable scale. A contour and two fitting functions that locally approximate the image intensities on the two sides of the contour are defined as the data fitting energy. A variational level set formulation incorporates the energy with a level set regularization term, and then energy minimization is carried out for the derived curve evolution equation.

The multi-scale normalized cuts-based segmentation (MNCut) algorithm [139] uses the normalized cut graph partitioning framework of image segmentation, where a graph

that encodes pair-wise pixel affinity is constructed and partitioned for image segmentation. The algorithm works simultaneously across the graph scales, with an inter-scale constraint to ensure communication and consistency between the segmentations at each scale, such that both coarse and fine level details are captured.

We test HDSS-1 and BEBS for mammogram data, HDSS-1, HDSS-1 and BEBS for MRI brain image data, and HDSS-1for Cameraman image. To be fair, we use the same initialization for all the algorithms implemented in the experiments. We set a_1, a_2, a_3 and a_5 to be 1, and a_4 to be 15. In HDSS-1, γ_2 is chosen to be 1. γ_1's in HDSS-1 and HDSS-2 are set to be 0.1. From the experimental results, we notice that the segmentation performance is not sensitive to the choice of these parameters. We use the iterative conditional modes (ICM) algorithm [140] as the optimization method for all the three algorithms.

5.2.5.3   *Experimental Results*

Fig. 5.4 shows the segmentation results of the mammogram with lesions, where the blue points denote the segmented positives. The regions enclosed by the light green curves in Figs. 5.4 (f) and (g) correspond to the segmented lesion regions by LSEWRI and RACM, respectively.
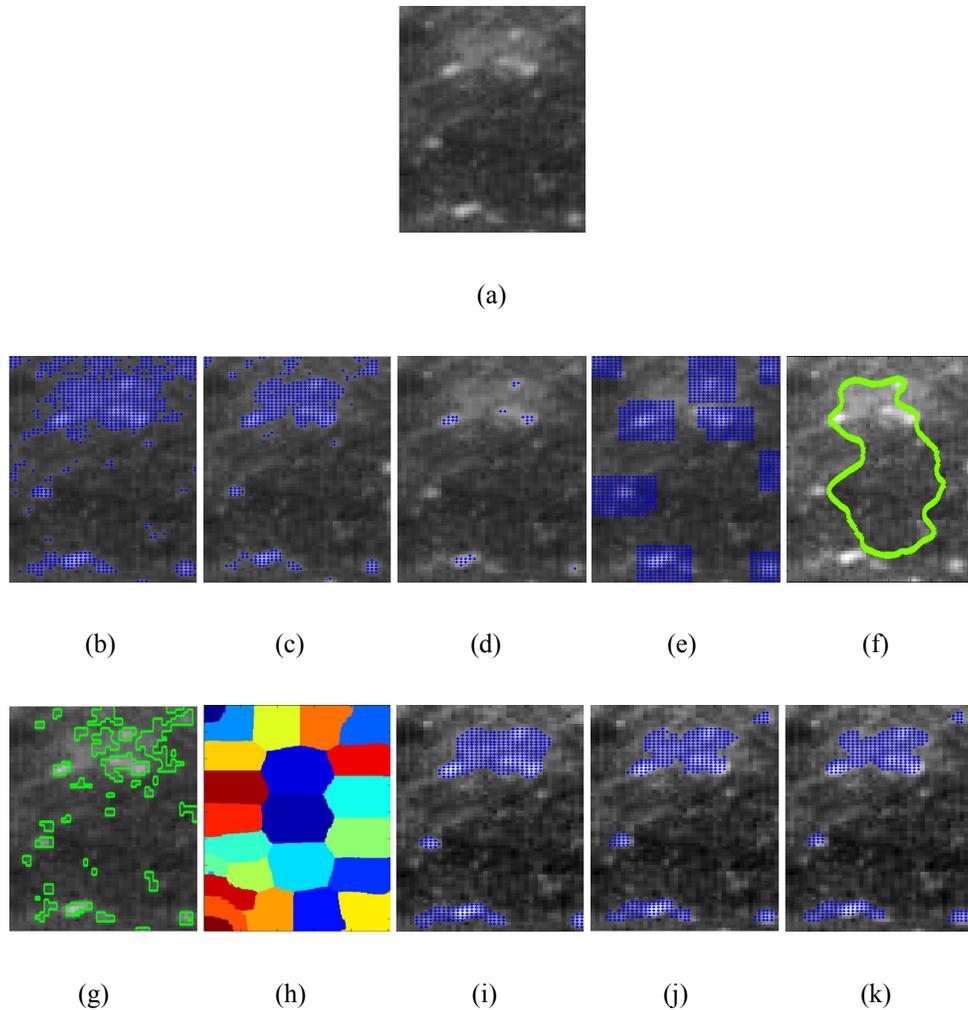
Figure 5.4: Original mammogram and the segmentation results. (a) Original mammogram with lesions; (b) segmentation by Otsu thresholding; (c) segmentation by GADC; (d) segmentation by IMS; (e) segmentation by HOSLW; (f) segmentation by LSEWRI; (g) segmentation by RACM; (h) segmentation by MNCut; (i) segmentation by conventional MRF; (j) segmentation by BEBS; (k) segmentation by HDSS-1.

From the figures, we can see that BEBS and HDSS-1, shown in Figs. 5.4 (j) and (k), yield better results than the other methods. Otsu thresholding, (b), yields too many false alarms. Obviously, for an image in which the intensity contrast is not very high, like the mammogram, intra-class variance measure is insufficient for yielding good segmentation.

The advantage of GADC (c) is that it converges quickly, but it yields many false positives. IMS (d) may converge to local extrema and misses many lesions. HOSLW method (e) can find the lesions efficiently, but it still generates false alarms and fails to determine the shape of lesion which, however, plays a very important role in discriminating the benign tumors from the malignant ones. Moreover, the segmentation performance of HOSLW depends on how accurately we can estimate the number of lesion pixels, which is usually not available in real-world applications. The LSEWRI method, shown in (f), also performs poorly and yields many mis-segmentations. RACM (g) finds all the lesions but with many false alarms. Besides, it also fails to determine the lesion shapes. The number of iterations of the two level set-based algorithms is set to be 500, which is sufficient for them to converge. For the MNCut method, we tried several numbers of segments, but did not observe any satisfactory results. A typical segmentation is shown in (h). Conventional MRF (i) does not find all the lesions and it also fails to determine lesion shape. This is because the conventional MRF only emphasizes intra-region homogeneity and label smoothness, such that it is too conservative and works poorly when the image contents are complex. Besides, all the methods used for comparison purpose fail to mimic the adaptation of HVS to the complexity and non-stationarity of the image contents.

As shown in Fig. 5.4 (j), BEBS performs satisfactorily, but the lesion contour is not smooth. This is because the boundary energy function does not emphasize the label homogeneity of the neighboring pixels as well as the global contrast of the image. On the contrary, HDSS-1 (k) integrates the boundary information, global contrast and pixel label homogeneity, and therefore yields a better result.

Fig. 5.5 shows the results for MRI brain image segmentation, where the non-brain background is first removed from the image and segmentation is carried out only on the brain pixels. Similar to the mammogram case, our presented algorithms yield better segmentations than the representative ones, in terms of smaller mis-segmented area and higher accuracy in object boundary determination. In Figs. 5.5 (b), (c), (g), (h) and (i), white regions denote the segmented WM by different algorithms. The regions enclosed by the light green curves in Figs. 5.5 (d) and (e) correspond to the segmented WM by the two level set-based methods with the number of iterations being 2000, which is sufficient for them to converge.

| | | | | |
| (a) | (b) | (c) | (d) | (e) |

| | | | |
| (f) | (g) | (h) | (i) |
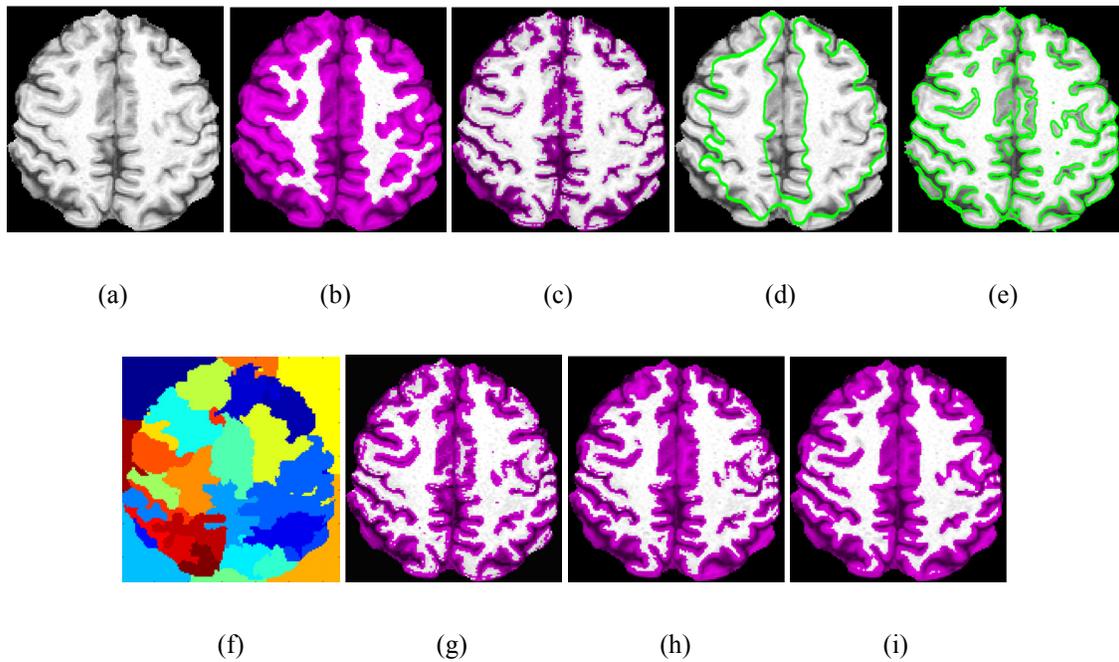
Figure 5.5: Original MRI brain image and the segmentation results. (a) Original MRI image; (b) segmentation by conventional MRF; (c) segmentation by Otsu thresholding; (d) segmentation by LSEWRI; (e) segmentation by RACM; (f) segmentation by MNCut; (g) segmentation by BEBS; (h) segmentation by HDSS-1; (i) segmentation by HDSS-2.

From the figures, we can see that the conventional MRF-based method (b) is conservative, as mentioned before. It puts more weight on pixel label homogeneity so misses many fine structures of the WM. Otsu (c) yields an unsuitable threshold such that some GM and CSF regions are segmented into the WM part. LSEWRI (d) produces a large number of mis-segmentations. RACM (e) includes some GM into WM and also yields segments with a very small size in the right half part of the brain image. MNCut (f) fails to characterize the fine structures of the WM and GM. The result from BEBS (g) is satisfactory, but it also yields some isolated pixels as well as some false positives in the bottom and boundary of the image. In contrast, HDSS-1 (h) and HDSS-2 (i) produce better segmentations in terms of reduced number of isolated pixels and the strengthened homogeneity of the neighboring pixel labels.

Fig. 5.6 shows the results of segmenting the Cameraman image. For visualization purposes, the results of the Otsu and HDSS-1 algorithms are shown as a binary image. From the figures, we can see that Otsu (b) yields a good segmentation of the human body, but generates many mis-segmentations in the background building and meadow parts. Most of them are segments with small sizes. LSEWRI (c) generates no mis-segmentation in the meadow but misses part of the human legs and also segments out a large area of background between the man and camera. RACM (d) yields good segmentation of the man. It even finds some fine structures of the image. But this method generates a large number of mis-segmentations in the meadow. We adjusted the number of segments of the MNCut method (e), and observed that it generates many homogeneous patches. Obviously, some post-processing algorithms, like region merging, need to be applied to

find the human from these patches. (f) shows the result of the presented HDSS-1 method. We can see that the man is segmented out with fewer mis-segmentations, especially in the meadow part, when compared to other methods. Also, a smaller portion of the camera is segmented out by this method.

<div align="center">(a)   (b)   (c)</div>

<div align="center">(d)   (e)   (f)</div>
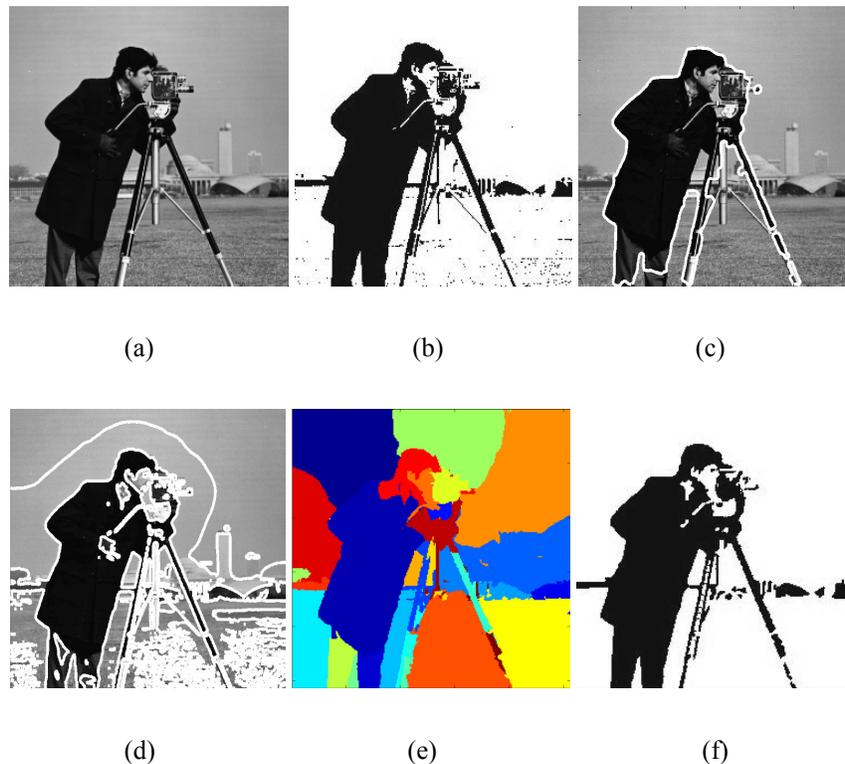
Figure 5.6: Original cameraman image and the segmentation results. (a) Original image; (b) segmentation by Otsu thresholding; (c) segmentation by LSEWRI; (d) segmentation by RACM; (e) segmentation by MNCut; (f) segmentation by HDSS-1.

5.3 HVS-Driven Image Segmentation Framework Using Local Segmentation Performance Measure

5.3.1 Image Segmentation Based on MAP-MRF

Context is important in image segmentation because contiguous pixels are likely to

belong to the same region and form homogeneous areas. MRF is an appropriate prior contextual model because it can identify the local properties of image regions by introducing context or dependence among neighboring pixels. MAP is a frequently used approach to obtain the solutions to the MRF-based image segmentation problems. Formally, in this approach, we have

$$x^* = \arg\max_{x} P(Y = y \mid X = x, \theta) P(X = x \mid \theta) \tag{5.25}$$

where x denotes the segmented image region label, $x=(x_1,\ldots, x_s,\ldots,x_M)$; x_s is the label of the pixel at location s in an image, $x_s=0,1,\ldots, L\text{-}1$; L is the number of region types, and M is the total number of pixels in an image; y is the observed image data, which is often modeled as the noise-free image corrupted by additive noise. θ is the MRF model parameter, relevant to the type of the cliques. A clique is a set of pixels that are neighbors of each other. $P(Y=y|X=x,\theta)$ is the conditional density of the observed image given the distribution of regions. $P(X=x|\theta)$ is the *a priori* density of the region process given θ, which according to the Hammersley-Clifford theorem [127], can be described by a Gibbs density of the form

$$P(X = x \mid \theta) = \exp\{-\sum_{C} V_c(X = x \mid \theta)\} / Z \tag{5.26}$$

Here Z is a constant used for normalization, and the summation is carried out over all cliques C. $V_c(X=x|\theta)$ is called the potential function associated with clique c.

Due to the Markov property and the assumptions that the additive noise is independent of the image signal and independent and identically-distributed (*i.i.d.*), e.g., *i.i.d.* Gaussian noise [140], (5.25) is often approximated by

$$x^* = \arg\max_x \prod_s P(Y_s = y_s \mid X_s = x_s, \theta) P(X_s = x_s \mid X_{\eta|s} = x_{\eta|s}, \theta) \tag{5.27}$$

where $\eta \mid s$ is the neighborhood of the pixel s. The terms of $P(Y_s=y_s|X_s=x_s,\theta)P(X_s=x_s|X_{\eta|s}=x_{\eta|s},\theta)$ can be considered as objective functions for representing the performance of single pixel segmentation.

5.3.2 Image Segmentation Based on Local Segmentation Performance Measure

In this section, we avoid of the noise modeling, and present a "soft" objective function, in which we consider image segmentation as a detection problem, and employ a local segmentation error to measure the impact of individual pixel segmentation on the local region and further on the objective function value and the segmentation result corresponding to the entire image, such that the interaction between the neighboring pixel segmentation as well as the uncertainty of the individual pixel segmentation can be taken into account.

Image segmentation can be considered as a detection problem. For example, in the binary class segmentation case, a pixel can be considered as a detected positive if the pixel is segmented into the object region, or a detected negative if segmented into the background region. This equivalence is also applicable to multi-class segmentation. Therefore, many metrics relevant to detection performance, such as the Ali-Silvey distance measures [141], could be appropriate choices for the segmentation objective function assignment. In this section, the probability of successful detection (P_{Suc}) is employed as an objective function to show the efficiency of this scheme.

The novel objective function for the algorithm is

$$C_{P_{Suc}} = \sum_s P_{Suc_s}(X_s = x_s)$$

(5.28)

where $\sum_s$ means the summation over the entire image. $P_{Suc_s}(X_s = x_s)$ is a measure for successful detection, reflecting the probability of successful segmentation of the pixel s when it is labeled with x_s. It aims to reflect the impact of the decision of a pixel on the segmentation performance of the entire image, and is defined as

$$P_{Suc_s}(X_s = x_s) = \gamma_{x_s}(1 - P_{e|x_s}(e \mid X_s = x_s))$$

(5.29)

where γ_{x_s} is the *a priori* probability of the pixel s labeled with x_s, and can be obtained from the Gibbs distribution, as shown in (5.26). $P_{e|x_s}(e \mid X_s = x_s)$ is the conditional probability of error given the label x_s. We define the optimal segmentation to be the one that maximizes (5.28).

For simplicity, we first consider the number of the types of image regions to be two and define $P_{e|x}(e|X_s=x_s)$ as

$$\begin{aligned} P_{e|x_s}(e \mid X_s = x_s) &= \pi_{0|x_s}\int_{-\infty}^{\infty}\phi_{x_s}(y)P_{0|x_s}(y)dy \\ &+ \pi_{1|x_s}\int_{-\infty}^{\infty}(1 - \phi_{x_s}(y))P_{1|x_s}(y)dy \end{aligned}$$

(5.30)

Here $\pi_{0|x_s}$ and $\pi_{1|x_s}$ are the *a priori* probabilities of occurrence of regions 0 and 1, with $\pi_{1|x_s} = 1 - \pi_{0|x_s}$. They can be estimated from the available segmentation results, i.e., $\pi_{i|x_s} = m_i / M$, where m_i is the number of pixels segmented into region i, and as mentioned before M is the total number of pixels in an image. We will see that the estimation of the probabilities of occurrence is not necessary in the binary class segmentation case. $P_{0|x_s}(y)$ and $P_{1|x_s}(y)$ are the *pmf*s of the discrete digital image data y for

the two types of regions, respectively. $\phi_{x_s}(y)$ is the critical region indicator function which gives the probability of assigning label 1 to x_s based on the observations.

Due to the instability of image contents and the fact that there may exist correlation between neighboring pixels, a "hard" threshold, in terms of a "hard" critical function, may not be suitable for segmentation. So we define a "soft" critical function, namely a "soft" decision rule, according to the data distribution, which is given as follows,

$$\phi_{x_s}(y) = P_{1|x_s}(y) / (P_{1|x_s}(y) + P_{0|x_s}(y)) \tag{5.31}$$

Substituting (5.30) and (5.31) into (5.29), we get

$$P_{Suc_s} = \gamma_{x_s} \cdot (1 - \int_{-\infty}^{\infty} \frac{P_{0|x_s}(y)P_{1|x_z}(y)}{P_{0|x_z}(y) + P_{1|x_s}(y)} dy) \tag{5.32}$$

Thus, the estimation of $\pi_{0|x_s}$ and $\pi_{1|x_s}$ is not necessary.

For the segmentation with multiple region types, we have

$$P_{Suc_s} = \gamma_{x_s} \cdot (1 - P_{e|x_s}) = \gamma_{x_s} \cdot (1 - \sum_{i=0}^{L-1} \sum_{j \neq i}^{L-1} \pi_{j|x_s} \int_{R_i} P_{j|x_s}(y)dy)$$
$$= \gamma_{x_s} \cdot (1 - \sum_{i=0}^{L-1} \pi_{i|x_s} \int_{-\infty}^{\infty} \phi_{i|x_s}(y)P_{i|x_s}(y)dy) \tag{5.33}$$

where R_i denotes the region indexed by i; $\pi_{i|x_s}$ are the *a priori* probabilities of occurrence of the type i region, with $\sum_{i=0}^{L-1} \pi_{i|x_s} = 1$, and can be estimated using the method mentioned before; $\phi_{i|x_s}(y)$ are soft decision rules for region i, defined as

$$\phi_{i|x_s}(y) = \sum_{j \neq i}^{L-1} P_{j|x_s}(y) / ((L-1)\sum_{j=0}^{L-1} P_{j|x_s}(y)) \tag{5.34}$$

and $\sum_{i=0}^{L-1} \phi_{i|x_s}(y) = 1$. $P_{j|x_s}(y)$ are estimated by the pixels labeled with j given that pixel s is classified as type x_s.

Now, we can see that the objective function is a soft one. It evaluates the result of single pixel segmentation where the statistical information of surrounding pixels is taken into account. From (5.32) or (5.33) we notice that γ_{x_s}, $P_{j|x_s}(y)$ and therefore P_{Suc_s} of different pixels are coupled, such that the label configuration of x_s's, which aims at achieving the global maximum value of (5.28), are very difficult to obtain. To simplify the optimization process, we assume that the label configuration obeys the MRF model, and define a local P_{Suc}, denoted as LP_{Suc}. Thus, LP_{Suc_s}'s for different pixel locations are independent of each other given the neighboring pixels around s. The calculation of LP_{Suc_s}'s and the update of x_s's can be carried out pixel by pixel in an independent manner. Formally, we have

$$
\begin{aligned}
C_{LP_{Suc}} &= \sum_s P_{Suc|x_s, x_{\eta|s}} (X_s = x_s \mid X_{\eta|s} = x_{\eta|s}) \\
&= \sum_s \gamma_{x_s|x_{\eta|s}} \cdot (1 - P_{e|x_s, x_{\eta|s}} (e \mid X_s = x_s, X_{\eta|s} = x_{\eta|s}))
\end{aligned}
\tag{5.35}
$$

Thus, the objective function is a measure of the segmentation performance in local regions. Each term of (5.35) is calculated from the pixel s and those around s. The optimal segmentation based on this criterion is achieved when (5.35) is maximized. The parameters of the Gibbs field are updated by fixing the pixel label X_s and finding the parameters which maximize (5.35). Therefore, $C_{LP_{Suc}}$ in our work is maximized by updating the pixel label and Gibbs field parameters in an alternative manner.

As mentioned before, other metrics relevant to detection performance could also be employed instead of probability of successful detection. Thus, we have developed a framework of image segmentation incorporating a soft objective function and local segmentation measure. It can be seen that only the pixel intensity information is used in

this chapter for segmentation. Further work considering other image features beyond intensities may yield better results.

5.3.3   HVS-Driven Image Segmentation

Many image segmentation algorithms devote a great deal of effort in extracting image features from the image itself, but overlook the properties of the HVS for characterizing the image content. In this section, we present a segmentation scheme that takes care of the HVS properties during segmentation. In the scheme, we first transform the original image into the wavelet domain and obtain the "wavelet image", and then the wavelet coefficients are weighted by the frequency response of the HVS given by a specified *CSF* in the wavelet domain. The *Invariant Single Factor Weighting* scheme discussed in [142] is employed here for the weighting task. Finally, the weighed "wavelet image" is transformed back to the spatial domain, which yields the *CSF*-enhanced image y_{CSF}. Segmentation is then carried out as before on y_{CSF} with HVS information embedded.

A number of different *CSF*s have been proposed in the literature. In this chapter, we choose the one discussed in [109] and used in (4.13) of Section 4.2.4.1.

$$A(f_r) = 2.6(0.192 + 0.114 f_r)e^{-(0.114 f_r)^{1.1}} \tag{4.13}$$

where f_r is the radian spatial frequency in cycles/degree. The wavelet we use is Daubechies-8 and the number of decomposition levels is 4.

5.3.4 Experimental Results

In this section, experiments are carried out on real-world mammogram data from the Digital Database for Screening Mammography (DDSM) [91], where the location of the mammogram lesions are identified by expert radiologists and used as the ground truth in our work.

Three segmentation algorithms are compared, namely the conventional MAP-MRF, LP_{Suc} and CSF-enhanced LP_{Suc}. Simulated annealing (SA) is employed for optimizing (5.35). An annealing schedule of $c_0/\log(1+k)$ [143] is used for the SA procedure, in which k is the discrete time variable and c_0 is specified by trial and error. In our experiments, a second-order neighborhood system, containing 8 pixels with the pixel of interest at the center, is employed for calculating $\gamma_{x_s|x_{\eta|s}}$.

The *pmf* in all the three methods is estimated by the kernel density technique [86], which uses the data in a processing window with the pixel to be labeled at the center of the window. The choice of the processing window size is critical. It should be small enough such that the effect of the single pixel segmentation result can influence LP_{Suc} and CSF-enhanced LP_{Suc}, and the local image information can be included, but not too small because an accurate *pmf* estimation is difficult to achieve given a very small data set. To find the tradeoff, we use a processing window shrinking scheme, similar to that used in [144]. In the scheme, window size is reduced with the iterations proceeding. The reason is that in the early stages of the iterations, a large window is necessary for robust estimation of the *pmf*s. As the algorithm progresses, the segmentation is improved and smaller windows provide more local information as well as more reliable *pmf* values. The window shrinking stops when a minimum window size is reached which is set to be 5×5

in our experiment.

For a fair comparison, the initial segmentations for the three algorithms are assumed to be the same, obtained using the local maxima filter [96]. Fig. 5.7 shows the original mammogram containing micro-calcifications, as well as the segmentation results of the conventional MAP-MRF, LP_{Suc} and the CSF-enhanced LP_{Suc} algorithms, where the detected positives are marked with (blue) dots, the false positives and missed lesions are pointed out by dashed circles and solid circles, respectively.

| (a) | (b) | (c) | (d) |
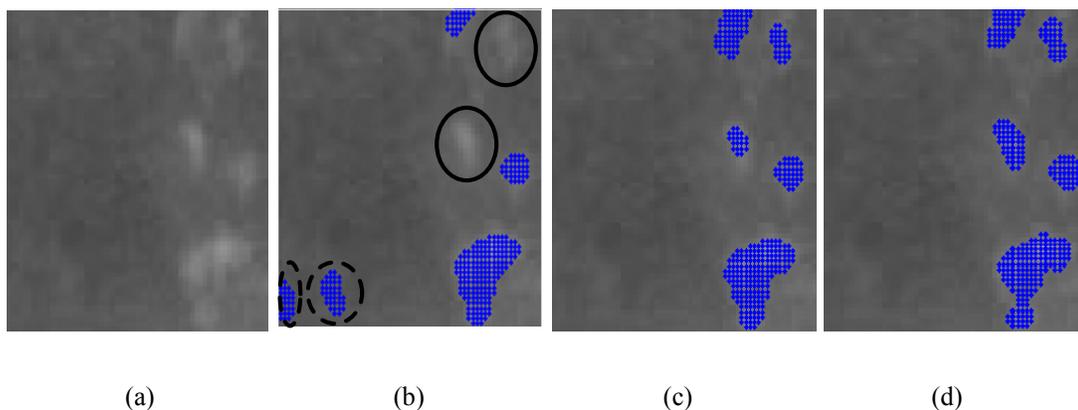
Figure 5.7: Original mammogram and segmentation results. (a) Original mammogram with micro-calcifications; (b) segmentation result with MAP-MRF; (c) segmentation result with LPSuc; (d) segmentation result with CSF-enhanced LPSuc. The detected positives are marked with dots, the false positives and missed lesions are pointed out by dashed circles and solid circles, respectively.

From the results, we can see that the conventional MAP-MRF (Fig. 5.7 (b)) yields two false positives at the left-bottom of the mammogram, and misses two true positives at the right-top and right-middle of the mammogram. LP_{Suc} (Fig. 5.7 (c)) finds all the lesions and does not generate any false positives. CSF-enhanced LP_{Suc} (Fig. 5.7 (d)) also detects

all the lesions but with more accurate lesion location and clearer lesion boundary than LP_{Suc}. The proposed algorithms provide better segmentation than the conventional MAP-MRF in terms of more accurate lesion localization and shape determination.

5.4   Summary

Image segmentation is a very important technique but a challenging problem for computer vision and image analysis. We presented two HVS-driven image segmentation approaches in this chapter. In the first approach, the objective function for segmentation was designed by considering the preference of HVS to good segmentation, from both region-based and boundary-based perspectives. The metrics were encoded into the MRF and the JND model was used to calculate the contrast of the image contents. Comparative performance evaluation was carried out via the experiments between the three variations of the presented algorithm and several representative segmentation and clustering algorithms available in the literature. The results show that the presented algorithms resulted in highly encouraging performance in terms of segmentation efficiency, robustness and convergence speed.

In the second approach, we considered image segmentation as a detection problem, and developed a novel image segmentation framework by introducing a local "soft" objective function to steer the segmentation. Moreover, the segmentation was further improved when the HVS information is embedded into images through a *CSF*-filtering procedure. Real-world mammogram data were used in the experimental comparison, which shows that the presented framework has highly encouraging performance.

CHAPTER VI

PERFORMANCE LIMIT OF IMAGE SEGMENTATION ALGORITHMS

In this chapter, image segmentation is formulated as a statistical parameter estimation problem based on varying coefficient model (VCM), and a modified Cramér–Rao bound (CRB) combined with the Affine bias model is employed to determine the performance limit of image segmentation algorithms. A fuzzy segmentation formulation is considered, of which hard segmentation is a special case. The effect of the factors, such as the contrast of the image pixel intensity, on the segmentation result is investigated via the bound, which gives us insights into the achievable accuracy of a segmentation algorithm in segmenting a specific image. Experimental results are obtained where we compare the performance of several representative image segmentation algorithms with the derived bound on both synthetic and real-world image data.

6.1 Introduction

As discussed in Section 1.3, the performance bound of image segmentation algorithms, which is dependent only on the given image contents, is significant both to the development of segmentation algorithms and to the evaluation of the quality of the segmentation results.

As mentioned in Section 1.3, there do exist efforts on bounding the segmentation performance from a statistical perspective. The work in [46] is based on the finite normal mixture (FNM) model assumption, where the model parameters, means and variances,

are estimated using Expectation-Maximization (EM) and Classification-Maximization (CM) algorithms. CRB on the variances of these estimates are derived. However, the FNM model is not universally applicable to all the images, and also, the unbiased estimator assumption made in [46] does not hold for many real-world segmentation algorithms, which will be seen in our experimental results. While studying multi-spectral image segmentation [47], the performance of the Markov random fields (MRF)-based segmentation algorithms was predicted using false alarm rate which was based on Rissanen's minimum description length (MDL) criterion. The analysis in [48] covered many detailed scenarios of segmentation, but the computational complexity, the MRF-based assumption and the use of multi-spectral image data constrained its application. In [48], the true segmentation label and two performance level parameters (sensitivity and specificity) were estimated using the EM algorithm. This scheme did not decouple the performance bound, i.e., the best achievable segmentation result for the given image data, from the specific segmentation algorithm, i.e., the EM algorithm used in [48]. In addition, the EM algorithm only guarantees to yield a locally optimal solution, which may not be appropriately used as a performance benchmark or bound, a global concept.

In this chapter, we derive the bound, or limit, on the performance measure of the segmentation result, i.e., the mean square error (MSE) of the pixel labels, based on the CRB, biased estimator assumption and Affine bias model. We first show that image segmentation problem fits the VCM [138] and image segmentation can be formulated as a parameter estimation problem. Second, in order to derive the biased bound, the CRB based on the unbiased estimator assumption is discussed as a necessary intermediate step. Thirdly, the biased bound and the optimum parameters for the Affine bias model are

determined, where the methods used to calculate the bound are also discussed. In the experiment section, we compare the derived biased bound with several representative image segmentation algorithms using synthetic and real-world image data. We also show the results of the unbiased bound, and demonstrate the unsuitability of the unbiasedness assumption.

6.2   *Performance Limit of Image Segmentation Algorithms*

6.2.1   Problem Formulation

Image segmentation is a very challenging problem, and many segmentation algorithms have been proposed. However, there is a fundamental question to be asked as to whether there exists a theoretical limit to image segmentation performance and, more importantly, how much room we have to improve the existing algorithms. In this section, as a first step to attempt to answer this question, we model the image segmentation problem as a linear estimation problem using a VCM, where the parameters of interest, i.e., the pixel labels indicating which region a pixel belongs to, are considered to be the coefficients of the VCM.

6.2.1.1   *Varying-Coefficient Model* [145]

In this section, we briefly introduce the VCM. Consider a random variable s whose distribution is dependent on a parameter η. In the VCM, η can be expressed as

$$\eta = F_0 + h_1 F_1(\chi_1) + \cdots + h_M F_M(\chi_M) \tag{6.1}$$

where $h_1, h_2, \cdots, h_M$ and $\chi_1, \chi_2, \cdots, \chi_M$ are known as the predictors for η, and $F_1, F_2, \cdots, F_M$ are functions that enable the representation of η. F_0 is the intercept term. Thus, the model is linear in the regressors, while their coefficients are allowed to change smoothly with the value of other variables which we call "effect modifiers". η is called the linear predictor, which is related to the mean $\Lambda = E\{s\}$ via the link function $\eta = \kappa(\Lambda)$. In the simplest case of the Gaussian model, $\kappa(\Lambda) = \Lambda$ and the data s is normally distributed with mean η, and model (6.1) has the form

$$s = h_1 F_1(\chi_1) + \cdots + h_M F_M(\chi_M) + \varepsilon \tag{6.2}$$

where $E\{\varepsilon\} = 0$, $\text{var}(\varepsilon) = \sigma_\varepsilon^2$. Other commonly used models are log-linear models, for which $\eta = \log(\Lambda)$ and s has a Poisson distribution, and the linear logistic model with $\kappa\{\Lambda\} = \log\{\Lambda/(1-\Lambda)\}$ and s is a binomial variable. A special case occurs when χ_k's are the same variable, such as time, age or pixel coordinates as used in our work.

There are many ways to model the functions $F_k(\chi_k)$. For example, we could use flexible parametric representations, such as Fourier series, piecewise polynomials, or otherwise and more generally nonparametric functions. In our work, the *B*-spline function (tensor product *B*-splines) is employed.

6.2.1.2   *Image Segmentation Model*

In this section, we model the image segmentation problem using VCM. Suppose we have an image with *N* pixels whose observed intensity values are $y(x)$, where x are pixel indices and ordered through zig-scanning, starting from the top-left to bottom-right in an

image, and $x = 1, 2, \cdots, N$. The image segmentation problem can be formulated, based on Gaussian model (6.2), as

$$
\begin{aligned}
y(x) &= s(x) + w'(x) \\
&= \left[h_1(x)F_1(x) + \cdots + h_M(x)F_M(x) + \varepsilon \right] + w'(x) \\
&= h_1(x)F_1(x) + \cdots + h_M(x)F_M(x) + w(x)
\end{aligned}
\tag{6.3}
$$

where $s(x)$ are the noise-free intensity values of the pixel x. This model has the signal effect modifying variable x, where M is the number of segmented regions, and $M \leq N$. (Note that the pixels which have the same features or characteristics should be classified into the same class, but these pixels classified into the same class need not be connected to each other, that is, they may be located in separate regions. The method used to calculate the bound in this chapter is based on regions, not on classes, so we will consider regions one by one, no matter whether they belong to the same class or not.) $h_k(x)$ is the pixel label of x, which can be considered as the membership function, representing the degree to which the pixel x belongs to the kth region, $0 \leq h_k(x) \leq 1$ and $\sum_{k=1}^{M} h_k(x) = 1$ for every x. In the rest of the chapter, the terms "label value" and "membership function value" will be used interchangeably. This definition enables the model to represent a general image segmentation scenario, i.e., fuzzy segmentation [146] where each pixel can belong to different regions at the same time. As a special case of fuzzy segmentation, a pixel in hard or crisp segmentation has the membership function $h_k(x) \in \{0,1\}$. In addition to providing a more general formulation, another important reason to study fuzzy segmentation is that the CRB fails to limit the MSE if the space of a parameter becomes finite [147], i.e., the hard segmentation case.

In (6.3), the noise term $w(x)$ consists of two parts, the image noise $w'(x)$ and the smoothing error ε. We assume that we have a very powerful smoother and the smoothing error is very small compared with the additive noise, so the image noise dominates the noise term, i.e., $w(x) = \varepsilon + w'(x) \approx w'(x)$. In this work, the noise is considered to be independent and identically distributed (*i.i.d.*) Gaussian random variable with zero mean and variance σ^2. Also, in our work, $F_k(x)$ is modeled using the 2D *B*-spline function with the coefficient vector β_k. Let $F_k(x) = \phi(x;\beta_k)$ represent the intensity of the pixel x in the kth region, and $\phi(x;\beta_k) = \sum_{l=1}^{m} \beta_{kl} b_l(x)$, where $b_l(x)$ are *B*-spline basis functions and m is the number of knots in an image. l is the index of the knots which are ordered through zig-scanning starting from the top-left to bottom-right in an image. For simplicity, the knots are uniformly deployed on the entire image plane.

Thus, (6.3) can be written in a matrix form as

$$
\begin{aligned}
y(x) &= h(x)^T \cdot \phi(x;\beta) + w(x) \\
&= h(x)^T \cdot \beta \cdot b(x) + w(x)
\end{aligned}
\tag{6.4}
$$

where T denotes the matrix transpose, $h(x) = [h_1(x), h_2(x), \cdots, h_M(x)]^T$ and $\phi(x;\beta) = [\phi(x;\beta_1), \phi(x;\beta_2), \cdots, \phi(x;\beta_k), \cdots, \phi(x;\beta_M)]^T$. Here, $\phi(x;\beta_k) = \beta_k^T \cdot b(x)$, where $\beta_k = [\beta_{k1}, \beta_{k2}, \cdots, \beta_{km}]^T$ and $b(x) = [b_1(x), b_2(x), \cdots, b_m(x)]^T$. So $\phi(x;\beta) = [\beta_1^T, \beta_2^T, \cdots, \beta_M^T]^T \cdot b(x) = \beta \cdot b(x)$, where $\beta = [\beta_1^T, \beta_2^T, \cdots, \beta_M^T]^T$.

We note that a similar formulation has been used in [148][149] for developing image segmentation algorithms. In [148][149], $h_i(x)$ is considered to be equal to or very close to 0 or 1, that is, hard segmentation, while in our formulation we consider a more general segmentation configuration, i.e., fuzzy segmentation where $h_i(x)$ lies in [0,1]. In addition,

in [148][149], it was argued that the pixel label, with the given Gibbsian distribution as the prior, is independent of the image content represented by β. In contrast, we do not make any assumptions on the dependence or the prior distribution.

There are several advantages to represent the image using the smoothing coefficients β, instead of the original pixel intensity information: (i) we can denote regions with various shapes and sizes, i.e., different number of pixels, using a "uniform" representation, i.e., the basis $b(x)$ and the smoothing coefficients β_k with known or controllable dimensions. Thus, the segmentation problem can be conveniently represented by some linear models, like VCM, and the analysis can be simplified; (ii) smoothing can reduce the impact of a small number of pixels with large difference in intensity from their neighboring pixels, i.e., outliers, so as to enhance the homogeneity of the image regions. It is also helpful in reducing the possibility of yielding regions with very small size, i.e., region with very few pixels; (iii) spatially varying intensity and interactions between the neighboring image areas can be taken into consideration by the smoothing representation to some extent; (iv) the smoothing procedure can represent the image content using much smaller number of coefficients compared with the number of original image pixels, and, therefore, simplifies the computation.

From (6.4), we can see that there are two sets of parameters $h(x)$ and β in the model, but we are only interested in the estimation of $h(x)$. We pack $h(x)$ into a large vector H and obtain $H = [h_1(1), h_2(1), \cdots, h_M(1), h_1(2), h_2(2), \cdots, h_M(2), \cdots, h_1(N), h_2(N), \cdots h_M(N)]^T$. In this chapter, we assume that the segmentation algorithms are biased estimators, that is, the output, $\hat{h}(x)$, of a segmentation algorithm is a biased estimator of the true pixel

label $h(x)$. More details about this assumption as well as its justification can be found in Section 6.2.3 and Appendix C. Before deriving the MSE bound under the biased estimator assumption, we first discuss the Fisher information matrix and the bound based on the unbiased estimator assumption in the next section, where the segmentation algorithm is assumed to yield an unbiased estimate of the true pixel label. We will see that the bound under the unbiasedness assumption is very useful in finding the bound under the biasedness assumption and is also helpful in the experimental part to verify the validity of the biased estimator assumption.

6.2.2 Fisher Information and Cramér–Rao Bound for Unbiased Estimator

In this section, we derive the Fisher information matrix and the Cramér–Rao bound based on the unbiased estimator assumption.

For an estimation problem with two unknown parameters, like H and β in our work, one parameter, say, H, can be considered to be the wanted parameter and the other one, β, can be considered as the unwanted one. Both of them are assumed to be random. Based on this formulation, the performance of four variations of the Bayesian bound for estimating the wanted parameter was compared in [150][151]. However, determination of all of the bounds requires either the computation of derivatives and expectation over the joint probability distributions of the observation Y and the wanted parameter or the observation and the whole parameter set, i.e., $P(Y,H)$ or $P(Y,H,\beta)$, which is a very challenging task given the variety of image contents. Here $Y = [y(1), \cdots, y(N)]^T$.

In our work, we assume H and β to be random so as to find a bound with reasonable complexity. We first determine the conditional CRB given H and β, and then find the expectation of the conditional bound with respect to H and β to obtain the global one. We will see that during the computation of expectation it is not necessary to determine the joint probability $P(H,\beta)$ and to even consider the potential dependence between H and β.

6.2.2.1   Fisher Information Matrix

In this section, we derive the Fisher information matrix conditioned on H and β, and propose a scheme to deal with the singularity of the matrix which may exist in the single image segmentation scenario. Assume that the noise $w(x)$ is *i.i.d.* Gaussian random variable with zero mean and variance σ^2, and the observed pixel intensity is also *i.i.d.* given the membership function H and the smoothing coefficient β. Then the conditional *pdf* of the observation is

$$P(Y;H,\beta) = \left(\frac{1}{\sqrt{2\pi\sigma^2}}\right)^N \exp\left(\frac{-\sum_{x=1}^{N}[y(x)-h(x)^T \cdot \beta \cdot b(x)]^2}{2\sigma^2}\right) \qquad (6.5)$$

So the log likelihood function is given by

$$L = \ln[P(Y;H,\beta)] = -\frac{N}{2}\ln 2\pi\sigma^2 - \frac{1}{2\sigma^2}\sum_{x=1}^{N}[y(x)-h(x)^T \cdot \beta \cdot b(x)]^2 \qquad (6.6)$$

We are only interested in estimating H and assume that the information about β is available, which can be estimated from the image contents and the ground-truth

segmentation results. This assumption on the availability of β is helpful in simplifying the determination of the bound and also in eliminating the ambiguity in model (6.4) due to the multiplication of H and β. So we focus on the Fisher information matrix corresponding to H and obtain

$$[J_F(H)]_{ij} = E_{Y|H,\beta}\left\{\left[\frac{\partial L}{\partial h_i}\right]\left[\frac{\partial L}{\partial h_j}\right]^T\right\} \tag{6.7}$$

The detailed derivation and the resulting Fisher information matrix are provided in Appendix B.

We notice from (B.6) that $J_F(H)$ is singular, which can be verified by multiplying the first row of $J_F(H)$ by $\beta_2^T b(1)$ and the second row by $\beta_1^T b(1)$. This is because the dimension of H is usually higher than the available observation Y, especially for the case of single image segmentation, which can be seen more clearly from (6.4). For multi-spectral image segmentation, there may not exist such a problem, since we have more observed image data, and the resulting Fisher information matrix for this case is shown in Appendix C. In this chapter, we focus on the derivation of the bound for the segmentation of single images, and the bound for multi-spectral image segmentation can be derived in a similar manner.

To overcome the singularity problem, we transform the multi-region segmentation problem, where $M>2$, to a binary-region segmentation problem, i.e., $M=2$, by maintaining the information regarding the region of interest, say, the ith region, and by considering the remaining regions as a single "super" region. That is, the membership functions and the smoothing coefficients corresponding to the pixels in the ith region remain fixed, and the rest of the regions are merged to form a "super" region whose membership functions and

the smoothing coefficients are recalculated based on the image contents of the "super" region. Thus, the segmentation model (6.4) can be written as

$$
\begin{aligned}
y(x) &= h(x)^T \cdot \beta \cdot b(x) + w(x) \\
&= h_i(x) \cdot \beta_i^T \cdot b(x) + \left[\sum_{j=1, j\neq i}^{M} \left(h_j(x) \cdot \beta_j^T \right) \right] \cdot b(x) + w(x) \\
&= h_i(x) \cdot \beta_i^T \cdot b(x) + h_{i^s}(x) \cdot \beta_{i^s}^T \cdot b(x) + w(x) \\
&= h_i(x) \cdot (\beta_i^T - \beta_{i^s}^T) \cdot b(x) + \beta_{i^s}^T \cdot b(x) + w(x)
\end{aligned}
\tag{6.8}
$$

where $h_i(x)$ and β_i are the original parameters of the ith region, and $h_{i^s}(x)$ and

$\beta_{i^s}^T$ correspond to the "super" region. $h_{i^s}(x) \cdot \beta_{i^s}^T \cdot b(x) = \left[\sum_{j=1, j\neq i}^{M} \left(h_j(x) \cdot \beta_j^T \right) \right] \cdot b(x)$, with $h_{i^s}(x) \geq 0$,

and $h_i(x) + h_{i^s}(x) = 1$, $i = 1, 2, \cdots, M$.

Based on (6.8), the Fisher information matrix of $H_i = [h_i(1), \cdots, h_i(N)]$, corresponding to the ith region, can be calculated as (6.9), by following a similar procedure as in Appendix A but with the "super" region considered.

$$
J_F(H_i) = \frac{1}{\sigma^2}
\begin{bmatrix}
\left(\beta_i^T b(1) - \beta_{i^s}^T b(1) \right)^2 & 0 & \cdots & 0 & 0 \\
0 & \left(\beta_i^T b(2) - \beta_{i^s}^T b(2) \right)^2 & \cdots & 0 & 0 \\
\vdots & \vdots & \ddots & \vdots & \vdots \\
0 & 0 & \cdots & \left(\beta_i^T b(N-1) - \beta_{i^s}^T b(N-1) \right)^2 & 0 \\
0 & 0 & \cdots & 0 & \left(\beta_i^T b(N) - \beta_{i^s}^T b(N) \right)^2
\end{bmatrix}_{N \times N}
\tag{6.9}
$$

which is not singular if $\beta_i^T b(x) - \beta_{i^s}^T b(x) \neq 0$. Since the resulting bound also requires the determination of the expectation of $\beta_i^T b(x) - \beta_{i^s}^T b(x)$ with respect to β, which will be seen in (6.14), we discuss the invertibility of the Fisher information matrix in the next section.

Thus, for $\beta_i^T b(x) - \beta_{i^s}^T b(x) \neq 0$, we have

$$J_F^{-1}(H_i) = \sigma^2 \begin{bmatrix} \dfrac{1}{\left(\beta_i{}^T b(1) - \beta_{i^s}{}^T b(1)\right)^2} & 0 & \cdots & 0 & 0 \\ 0 & \dfrac{1}{\left(\beta_i{}^T b(2) - \beta_{i^s}{}^T b(2)\right)^2} & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & \dfrac{1}{\left(\beta_i{}^T b(N-1) - \beta_{i^s}{}^T b(N-1)\right)^2} & 0 \\ 0 & 0 & \cdots & 0 & \dfrac{1}{\left(\beta_i{}^T b(N) - \beta_{i^s}{}^T b(N)\right)^2} \end{bmatrix}_{N \times N} \quad (6.10)$$

The same result can be obtained using the constrained CRB [152] with the "super" region scheme where the constraint is $h_i(x) + h_{i^s}(x) = 1$.

6.2.2.2   Cramér–Rao Bound for Unbiased Estimator

In this section, we derive the Cramér–Rao bound under the unbiased estimator assumption, and employ Jensen's inequality for matrix measures [153] to simplify the expectation determination procedure. We assume that the segmentation algorithms yield unbiased estimates of the pixel labels. Based on the formulation in the last section, the unbiased bound of multi-region segmentation can be calculated in a region by region manner. For the ith region, we calculate the Fisher information matrix $J_F(H_i)$ and its inverse $J_F^{-1}(H_i)$ which corresponds to the conditional bound of the covariance matrix of $\hat{H}_i$. We find the expectation of $J_F^{-1}(H_i)$ with respect to H and β, and obtain the global bound for $\hat{H}_i$, which is different from the bounds discussed in [150][151] as mentioned at the beginning of Section 6.2.2. Repeating the procedure for all the regions and averaging the resulting bounds, we obtain the average unbiased bound for the entire image. In this way, we decompose the estimation problem with the dimensionality equal to MN into M sub-problems, each of which has the dimensionality N, the same size as the

number of observations (the total number of pixels in an image), and therefore overcome the ambiguity due to insufficient number of observations.

Now, we study the bound on the covariance of the estimate $\hat{H}$ under the unbiasedness assumption. The conditional covariance matrix of $\hat{H}_i$, i.e., $Cov(\hat{H}_i \mid H, \beta)$, for the unbiased estimator can be written as

$$Cov(\hat{H}_i \mid H, \beta) = E_{Y \mid H, \beta}\{(\hat{H}_i - \hat{\mu}_{\hat{H}_i \mid H, \beta})(\hat{H}_i - \hat{\mu}_{\hat{H}_i \mid H, \beta})^T\} \geq J_F^{-1}(H_i) \qquad (6.11)$$

where $\hat{\mu}_{\hat{H}_i \mid H, \beta} = E(\hat{H} \mid H, \beta)$, and the corresponding conditional bound $CRB_{Unbiased}(\hat{H}_i \mid H, \beta)$ is

$$CRB_{Unbiased}(\hat{H}_i \mid H, \beta) = \text{Tr}[J_F^{-1}(H_i)] = \sigma^2 \sum_{x=1}^{N} \frac{1}{(\beta_i^T b(x) - \beta_{i^s}^T b(x))^2} \qquad (6.12)$$

where $\text{Tr}(U)$ denotes the trace of the matrix U.

The global bound for H_i is determined by finding the expectation of $CRB_{Unbiased}(\hat{H}_i \mid H, \beta)$ with respect to H and β, i.e., $E_{H,\beta}\{CRB_{Unbiased}(\hat{H}_i \mid H, \beta)\}$.

The average bound for the unbiased estimator for an individual region can be found by averaging the global bounds of all the regions, that is,

$$
\begin{aligned}
CRB_{Unbiased-Ave}(\hat{H}) &= \frac{1}{M} \sum_{i=1}^{M} E_{H,\beta}\{CRB_{Unbiased}(\hat{H}_i \mid H, \beta)\} \\
&= \frac{1}{M} \sum_{i=1}^{M} E_{H,\beta}\{\text{Tr}[J_F^{-1}(H_i)]\} \\
&= \frac{1}{M} \sum_{i=1}^{M} \text{Tr}\{E_\beta[J_F^{-1}(H_i)]\}
\end{aligned} \qquad (6.13)
$$

where the last equality holds since $\text{Tr}[J_F^{-1}(H_i)]$ is not a function of H.

In our work, we further average $CRB_{Unbiased-Ave}(\hat{H})$ over all the pixels in an image and the average pixel-level bound serves as the bound on the performance of image segmentation. Since H_i's have the same dimensions, i.e., the number of pixels included in an image, we obtain the average pixel-level bound by dividing $CRB_{Unbiased-Ave}(\hat{H})$ with the total number of pixels, N, in an image, which is shown in (6.14).

$$CRB_{P-Unbiased-Ave}(\hat{H}) = \frac{1}{N} CRB_{Unbiased-Ave}(\hat{H})$$
$$= \frac{1}{MN} \sum_{i=1}^{M} \text{Tr}\{E_\beta[J_F^{-1}(H_i)]\} \tag{6.14}$$

We notice from (6.10) that it is not easy to find the expectation of $J_F^{-1}(H_i)$ over β, so we employ an approximation when calculating the bound, by performing the expectation operation on $J_F(H_i)$ first and then finding its inverse, i.e., $(E_\beta[J_F(H_i)])^{-1}$. According to **Theorem 4.2** (Jensen's inequality for matrix measures) and the **Tracial Jensen inequalities** in [153], we have

$$E_\beta[J_F^{-1}(H_i)] \geq (E_\beta[J_F(H_i)])^{-1} \text{ and } \text{Tr}\{E_\beta[J_F^{-1}(H_i)]\} \geq \text{Tr}\{(E_\beta[J_F(H_i)])^{-1}\} \tag{6.15}$$

where

$$(E_\beta[J_F(H_i)])^{-1} = \sigma^2 \begin{bmatrix} 1/E_\beta\{(\beta_i^T b(1) - \beta_{i^*}^T b(1))^2\} & 0 & \cdots & 0 & 0 \\ 0 & 1/E_\beta\{(\beta_i^T b(2) - \beta_{i^*}^T b(2))^2\} & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1/E_\beta\{(\beta_i^T b(N-1) - \beta_{i^*}^T b(N-1))^2\} & 0 \\ 0 & 0 & \cdots & 0 & 1/E_\beta\{(\beta_i^T b(N) - \beta_{i^*}^T b(N))^2\} \end{bmatrix}_{M \times N} \tag{6.16}$$

Thus, a looser bound is found to ease computation, which is called the modified CRB in this chapter and is indicated by the superscript *Mod*. Therefore, from (6.14) we have

$$CRB^{Mod}_{P-Unbiased-Ave}\left(\hat{H}\right)=\frac{1}{MN}\sum_{i=1}^{M}\text{Tr}\left\{\left(E_{\beta}[J_{F}(H_{i})]\right)^{-1}\right\}$$ (6.17)

We now discuss a special situation, where $E_{\beta}\left\{\left(\beta_{i}{}^{T}b(x)-\beta_{i^{s}}{}^{T}b(x)\right)^{2}\right\}$ in (6.16) has very small values such that its inverse is very large. In this case, the resulting average CRB value might be large. We note that the very small values of $E_{\beta}\left\{\left(\beta_{i}{}^{T}b(x)-\beta_{i^{s}}{}^{T}b(x)\right)^{2}\right\}$ correspond to an extreme situation where two image regions are not distinguishable at x. Because $E_{\beta}\left\{\left(\beta_{i}{}^{T}b(x)-\beta_{i^{s}}{}^{T}b(x)\right)^{2}\right\}$ evaluates the average intensity difference between the two regions with the center at x (due to the expectation operation with respect to β), it reduces the effect when the two different regions have similar pixel intensities at x, by making use of the intensity information of a group of pixels. Therefore, there are very few components of $E_{\beta}\left\{\left(\beta_{i}{}^{T}b(x)-\beta_{i^{s}}{}^{T}b(x)\right)^{2}\right\}$ in (6.16) with very small values, given that the two image regions are reasonably separable, which has also been verified by our experiments. Thus, in our work we simply ignore the contribution of the components to the bound when they have very small values. This operation yields a reasonable tight bound. However, if we do not incorporate the expectation operation when calculating the bound, the performance of the resulting bound might be deteriorated when different regions have similar pixel intensities at x, which can be seen in the experimental results shown in Figs. 6.1 (c), 6.2 (c) and 6.3 (c).

From (6.16), we can see that $E_{\beta}\left\{\left(\beta_{i}{}^{T}b(x)-\beta_{i^{s}}{}^{T}b(x)\right)^{2}\right\}$ actually measures the square of the difference between the intensities at pixel x contributed by the region of interest and the "super" region. It indicates the interaction between different regions at x. A smaller difference means a higher similarity between the two image regions. This result corresponds to the image content which is more difficult to segment apart, and the

variance of the segmentation label is larger. Here, the intensity difference evaluation is carried out by using the spline coefficients and the expectation operation, and, thus, the effect of the contribution of the neighboring pixels to the intensity at *x*, i.e., the correlation between neighboring pixels, is also taken into account. It is also interesting to notice that the separability of the two regions, which is reflected by the segmentation variance, is independent of the membership values and only related to the contrast between the intensities of the neighboring regions overlapping at a pixel. Additionally, a larger noise energy, i.e., larger σ^2, has a larger negative influence on the segmentation result, which corresponds to a higher value of the bound. We can see that the bound of (6.17) is consistent with these intuitive expectations.

The bound (6.17) has been obtained under the unbiasedness assumption but as we will see in the next section that a biased estimator is a more reasonable assumption for real-world image segmentation algorithms. Therefore, the result obtained in this section is not applicable in practice. However, it will be very useful in deriving the bound for the biased estimator case.

6.2.3   Cramér–Rao Bound for Biased Estimator

In this section, we assume the estimator of *H* to be biased, and derive the bound on the MSE of the segmentation results. We continue to consider the transformed binary segmentation problem in this section.

6.2.3.1 Cramér–Rao Bound for Biased Estimator

From both theoretical and practical points of views, unbiased estimators do not always exist. Moreover, biased estimators often have the advantage of lower MSE over unbiased ones if they exist [154]. MSE actually includes the tradeoff between bias and covariance. In addition, unbiased estimators tend to yield very large variance, especially for some ill-posed problems, such as image segmentation. Regularization is widely used to solve ill-posed problems and the resulting estimators are often biased [155]. Many state-of-the-art image segmentation algorithms are designed under a regularization framework, in which an objective function consisting of both a fidelity term and a penalty term is optimized, resulting in biased estimators.

Following the same steps as when deriving the average bound for the unbiased estimator in the last section, we first write the expression of the conditional MSE in terms of bias and covariance, as shown in (6.18)

$$E\left\{\left\|\hat{H}_i - H_i\right\|^2 \mid H, \beta\right\} = \left\|g(H_i)\right\|^2 + \mathrm{Tr}\left\{Cov(\hat{H}_i \mid H, \beta)\right\} \tag{6.18}$$

where $g(H_i) = E\{\hat{H}_i\} - H_i$ is the bias vector of $\hat{H}_i$.

Under suitable regularity conditions on $P(Y \mid H, \beta)$, the covariance of a biased estimator of $\hat{H}$ is bounded by the CRB [150]

$$Cov(\hat{H}_i \mid H, \beta) \geq A J_F^{-1}(H_i) A^T \tag{6.19}$$

where

$$A = I + \frac{\partial g}{\partial H} \tag{6.20}$$

and I is the identity matrix.

In our work, we assume that the behavior of the bias model can be approximated by an Affine function. The Affine model has been justified and employed to study the MSE bound for estimation problems in [156]. The details of the justification of the Affine bias assumption in image segmentation can be found in Appendix D. Formally, we have

$$g(H_i) = K_i H_i + u_i \tag{6.21}$$

where K_i and u_i are Affine parameters for the ith region. So, following the same steps as in the last section and considering (6.18)-(6.21), we have the conditional MSE bound of a biased estimator for $\hat{H}_i$ as follows

$$E\left\{\left\|\hat{H}_i - H_i\right\|^2 \mid H, \beta\right\} \geq (K_i H_i + u_i)^T (K_i H_i + u_i) + \mathrm{Tr}\left((I + K_i)J_F^{-1}(H_i)(I + K_i)^T\right) \tag{6.22}$$

Therefore, the global MSE bound for $\hat{H}_i$, i.e., $CRB_{Biased}(\hat{H}_i)$, is given by

$$E\left\{\left\|\hat{H}_i - H_i\right\|^2\right\} \geq CRB_{Biased}(\hat{H}_i) = \int \left\{(K_i H_i + u_i)^T (K_i H_i + u_i) + \mathrm{Tr}\left((I + K_i)J_F^{-1}(H_i)(I + K_i)^T\right)\right\} P(H, \beta)dHd\beta \tag{6.23}$$

The average MSE bound, i.e., $CRB_{Biased-Ave}(\hat{H})$, can be found by averaging the global bound for each region, and we, therefore, obtain

$$CRB_{Biased-Ave}(\hat{H}) = \frac{1}{M}\sum_{i=1}^{M} \int \left\{(K_i H_i + u_i)^T (K_i H_i + u_i) + \mathrm{Tr}\left((I + K_i)J_F^{-1}(H_i)(I + K_i)^T\right)\right\} P(H, \beta)dHd\beta \tag{6.24}$$

6.2.3.2   Optimum Affine Bias Model

In this section, we determine the optimum $\{K_i^*, u_i^*\}$ of the Affine bias model which yield the minimum value of the bound in (6.23), that is,

$$\{K_i^*, u_i^*\} = \arg\min_{K_i, u_i}\left\{\int \left[(K_i H_i + u_i)^T (K_i H_i + u_i) + \mathrm{Tr}\left((I + K_i)J_F^{-1}(H_i)(I + K_i)^T\right)\right] P(H, \beta)dHd\beta\right\} \tag{6.25}$$

There are basically two schemes to find the solution of the optimization problem posed in (6.25). The first one is to assume that K_i and u_i are functions of H or β and $\{K_i^*, u_i^*\}$ are found as the solution to the following optimization problem, as discussed in [156],

$$\{K_i^*, u_i^*\} = \arg\min_{K_i, u_i} \{MSEB(K_i, u_i, H_i) - MSEB(0, 0, H_i)\} \qquad (6.26)$$

where $MSEB(K_i, u_i, H_i) = (K_i H_i + u_i)^T (K_i H_i + u_i) + \text{Tr}\left((I + K_i) J_F^{-1}(H_i)(I + K_i)^T\right)$, and $MSEB(0, 0, H_i)$ corresponds to the unbiased estimator case. As derived in [156], the resulting optimum Affine bias parameters are $K_i^* = -\text{Tr}\{J_F^{-1}(H_i)\}/\left(\text{Tr}\{J_F^{-1}(H_i)\} + c_i\right)I$ and $u_i^* = -\text{Tr}\{J_F^{-1}(H_i)\}/\left(\text{Tr}\{J_F^{-1}(H_i)\} + c_i\right)v_i$, where $\|H_i - v_i\|^2 \le c_i$ for some vector v_i and scalar $c_i > 0$. The calculation of the bound requires the expectation of the function in (6.23) over $P(H, \beta)$, which is usually not tractable.

We, therefore, use the second scheme, in which we assume that M_i and u_i are not functions of H and β. As a further simplification, by using the result of (6.15) and also observing that $(I + K_i)\left\{E_\beta\left[J_F^{-1}(H_i)\right] - \left(E_\beta\left[J_F(H_i)\right]\right)^{-1}\right\}(I + K_i)^T$ are positive semi-definite, we obtain a modified bound $CRB_{Biased}^{Mod}(\hat{H}_i)$ for the biased estimator, which is looser than $CRB_{Biased}(\hat{H}_i)$ shown in (6.23). Thus, we have

$$CRB_{Biased}(\hat{H}_i) \ge CRB_{Biased}^{Mod}(\hat{H}_i) = \int\left\{\text{Tr}\left((I + K_i)\left(E_\beta\left[J_F(H_i)\right]\right)^{-1}(I + K_i)^T\right)\right\}P(H)dH + \int(K_i H_i + u_i)^T(K_i H_i + u_i)P(H, \beta)dHd\beta \qquad (6.27)$$
$$= \text{Tr}\left((I + K_i)\left(E_\beta\left[J_F(H_i)\right]\right)^{-1}(I + K_i)^T\right) + \int(K_i H_i + u_i)^T(K_i H_i + u_i)P(H)dH$$

The last equality in (6.27) holds because $\text{Tr}\left((I + K_i)\left(E_\beta\left[J_F(H_i)\right]\right)^{-1}(I + K_i)^T\right)$ is not a function of H and $(K_i H_i + u_i)^T(K_i H_i + u_i)$ is not a function of β.

Inspired by [157], the optimum Affine model parameters can be obtained by setting the derivative of $CRB_{Biased}^{Mod}(\hat{H}_i)$ with respect to the two parameters to zero, i.e.,

$$\frac{\partial CRB_{Biased}^{Mod}(\hat{H}_i)}{\partial K_i} = 0 \text{ and } \frac{\partial CRB_{Biased}^{Mod}(\hat{H}_i)}{\partial u_i} = 0 \tag{6.28}$$

Thus, we obtain the optimum parameter pair

$$K_i^* = -\left(E_\beta\left[J_F(H_i)\right]\right)^{-1}\left\{\left(E_\beta\left[J_F(H_i)\right]\right)^{-1} + Cov(H_i)\right\}^{-1} \tag{6.29}$$

and

$$u_i^* = \left(E_\beta\left[J_F(H_i)\right]\right)^{-1}\left\{\left(E_\beta\left[J_F(H_i)\right]\right)^{-1} + Cov(H_i)\right\}^{-1}E_{H_i}(H_i) \tag{6.30}$$

Substituting K_i^* and u_i^* into $CRB_{Biased}^{Mod}(\hat{H}_i)$, we obtain the modified bound for the ith region

$$CRB_{Biased}^{Mod}(\hat{H}_i)^* = \mathrm{Tr}\left\{\left(E_\beta\left[J_F(H_i)\right]\right)^{-1} - \left(E_\beta\left[J_F(H_i)\right]\right)^{-1}\left\{\left(E_\beta\left[J_F(H_i)\right]\right)^{-1} + Cov(H_i)\right\}^{-1}\left(E_\beta\left[J_F(H_i)\right]\right)^{-1}\right\} \tag{6.31}$$

The details of the above derivation for the parameters and the bound can be found in Appendix E.

So the average MSE bound is

$$\begin{aligned} CRB_{Biased-Ave}^{Mod}(\hat{H}) &= \frac{1}{M}\sum_{i=1}^{M}CRB_{Biased}^{Mod}(\hat{H}_i)^* \\ &= \frac{1}{M}\sum_{i=1}^{M}\mathrm{Tr}\left\{\left(E_\beta\left[J_F(H_i)\right]\right)^{-1} - \left(E_\beta\left[J_F(H_i)\right]\right)^{-1}\left\{\left(E_\beta\left[J_F(H_i)\right]\right)^{-1} + Cov(H_i)\right\}^{-1}\left(E_\beta\left[J_F(H_i)\right]\right)^{-1}\right\} \end{aligned} \tag{6.32}$$

As before, we obtain the average pixel-level MSE bound by averaging $CRB_{Biased-Ave}^{Mod}(\hat{H})$ with respect to the total number of pixels, N, in an image, and we have

$$CRB^{Mod}_{P-Biased-Ave}(\hat{H}) = \frac{1}{N} CRB^{Mod}_{Biased-Ave}(\hat{H})$$

$$= \frac{1}{MN} \sum_{i=1}^{M} \mathrm{Tr}\left\{ \left(E_\beta[J_F(H_i)]\right)^{-1} - \left(E_\beta[J_F(H_i)]\right)^{-1}\left\{\left(E_\beta[J_F(H_i)]\right)^{-1} + Cov(H_i)\right\}^{-1}\left(E_\beta[J_F(H_i)]\right)^{-1}\right\} \tag{6.33}$$

We notice from (6.33) that the decomposition of the terms containing H and β makes the solution easily computable and no explicit expression of the joint probability $P(H, \beta)$ is required. It also avoids the study of the dependence between H and β.

6.2.3.3   *Calculation of the MSE Bound*

Computation of (6.33) requires the determination of $E_\beta[J_F(H_i)]$ and $Cov(H_i)$ for the ith segmented region. In this section, we discuss the schemes to calculate these quantities.

6.2.3.3.1   Calculation of $E_\beta[J_F(H_i)]$

We notice that calculation of $E_\beta[J_F(H_i)]$ is not straightforward even if we are able to find the distribution of β, which, of course, is also a challenging task given various image contents. So we propose to use an empirical approximation to find the expectation value.

$E_\beta[J_F(H_i)]$ is a diagonal matrix, with the diagonal elements $E_\beta\left\{\left(\beta_i^T b(x) - \beta_{i^s}^T b(x)\right)^2\right\}$. Therefore, without loss of generality, we only investigate this term.

$$E_\beta\left\{\left(\beta_i^T b(x) - \beta_{i^s}^T b(x)\right)^2\right\} = \sum_{q=1}^{m} b_q(x)^2 \left\{E_{\beta_i}\left(\beta_{iq}^2\right) + E_{\beta_{i^s}}\left(\beta_{i^s q}^2\right)\right\} - 2\sum_{q=1}^{m}\sum_{t=1}^{m} E_\beta\left(\beta_{iq}\beta_{i^s t}\right)b_q(x)b_t(x)$$

$$+ 2\sum_{q=1}^{m}\sum_{t=q+1}^{m} b_q(x)b_t(x)\left\{E_{\beta_i}\left(\beta_{iq}\beta_{it}\right) + E_{\beta_{i^s}}\left(\beta_{i^s q}\beta_{i^s t}\right)\right\} \tag{6.34}$$

Thus, we may use empirical estimation to approximate the second-order statistics in (6.34) and thereafter find the overall expectation of $\left(\beta_i^{\ T}b(x)-\beta_{i^s}^{\ T}b(x)\right)^2$, which avoids the step of finding the probability distribution of β. More specifically, given the noise-free image with the segmentation label H, we determine H_i and H_{i^s} as well as the pixels belonging to i^{th} and i^{sth} regions. In this way, we separate an image into two layers, one corresponding to the ith region, called the ith layer, and another corresponding to the i^{sth} region, called the i^{sth} layer. In particular, if $h_i(x)=1$, the pixel intensity at x of the ith layer is set equal to $s(x)$, which is the pixel intensity of the original noise-free image at x; if $h_i(x)=0$, the intensity at x of the ith layer is set equal to zero. Otherwise, for a fuzzy pixel $s(x)=h_i(x)s_i^{Hard}(x)+h_{i^s}(x)s_{i^s}^{Hard}(x)$ with $h_i(x)\in(0,1)$, the pixel intensity at x of the ith layer is set equal to $s_i^{Hard}(x)$. Here $s_i^{Hard}(x)$ is the "original" hard component from the ith region, which contributes to the fuzzy pixel. This is motivated by the result of (6.16) in Section 6.2.2.2 that the segmentation error is only dependent on the intensity difference between the original regions, irrespective of the membership function values. The same procedure is carried out for the i^{sth} region. We then use the tensor B-splines to find the smoothing coefficients β_i and β_{i^s} for the two layers, respectively.

As we know that the empirical statistics will be closer to the true ones if more samples from the same distribution are used. To obtain enough valid samples of β_i and β_{i^s}, we use a "non-local" technique. That is, for β_{iq}, we search the coefficients with statistics similar to β_{iq} in the ith layer. β_{iq} together with the other similar coefficients are collected to form an ensemble, and they are considered to be various realizations of the same

random variable. Then the terms including the second-order statistic of β_{iq} in (6.34) are calculated empirically using the collected coefficients in the ensemble. For example, $E_{\beta_i}\left(\beta_{iq}{}^2\right)$ is approximated by $\left(\sum_{d=1}^{D}\beta_{iq(d)}{}^2\right)/D$, where $\beta_{iq(d)}$ is the dth collected coefficient in the ensemble of β_{iq}, and D is the number of these coefficients, i.e., the size of the ensemble. The same procedure is carried out for β_{is}. The second-order statistics including β_i and β_{is} are calculated using the collected coefficients from both the ith and i^{sth} layers.

Ignoring the approximation error, the smoothing coefficients and the pixels are two ways to represent the same image content, so we use pixel level features to search for similar coefficients because usually the number of pixels is much larger than the coefficients and, therefore, the statistics of the pixel level are more reliable. For example, when we search for coefficients similar to β_{iq}, we divide the image into patches centered at each knot with a fixed size, that is, the location of each smoothing coefficient is at the center of the patch. A suitable metric is employed to find the patches with a similar structure to the patch centered at β_{iq}, and the corresponding smoothing coefficients will be put in the ensemble of β_{iq}. As will be seen in Section 6.3, we will use a metric called structural similarity (*SSIM*) index [108]. In other words, we use the similarity of the patches to represent the similarity of the smoothing coefficients.

6.2.3.3.2   Calculation of $Cov(H_i)$

The analytical solution to estimate $Cov(H_i)$ requires the knowledge of the distribution of

H_i which is unknown and also not easy to find. Therefore, we use the bootstrapping technique [158]. Bootstrapping is an approach for statistical inference, and used to estimate the properties of an estimator ($Cov(H_i)$ in our work) by measuring those properties when sampling from an approximating distribution. It generates the empirical distribution of the observed data by constructing a number of resamples of the observed dataset, i.e., H_i in our work, with the same size as the observed dataset. These resamples are obtained by random sampling with replacement from the original dataset. Bootstrapping procedure is independent of the distribution, and provides an indirect method to assess the properties of the distribution which determine the sample and the parameters of interest [159]. Besides, bootstrapping is robust with respect to possibly small number of samples.

In our work, random sampling with replacement is carried out L times on H_i, and we obtain L bootstrap samples, from which the covariance matrix is calculated. This procedure is repeated R times, and the resulting R calculated covariance matrices represent an empirical bootstrap distribution of $Co\hat{v}(H_i)$ obtained from the available dataset. We accept the average of the estimated covariance matrices as the estimate of $Co\hat{v}(H_i)$. From this empirical bootstrap distribution, we can derive a bootstrap confidence interval which is also the confidence interval of the estimate of the bound and can be considered as the variance of the bootstrap estimates.

Formally, we have

$$Co\hat{v}(H_i)^r = \frac{1}{L-1}\sum_{l=1}^{L}\left(H_i^{rl} - \hat{\mu}_{H_i^{lr}}\right)\left(H_i^{rl} - \hat{\mu}_{H_i^{lr}}\right)^T \qquad (6.35)$$

and

$$Co\hat{v}(H_i) = \frac{1}{R}\sum_{r=1}^{R} Co\hat{v}(H_i)^r \qquad (6.36)$$

where H_i^{rl} is the lth bootstrap sample of the same size as H_i when generating the rth covariance matrix from the empirical bootstrap distribution, and $\hat{\mu}_{H_i^{rl}}$ is the mean vector of L bootstrap samples H_i^{rl}.

Repeating the above procedure of estimating $E_\beta\left[J(H_i)\right]$ and $Cov(H_i)$ for all the M regions, and plugging these results into (6.33), we obtain the average pixel-level MSE bound of image segmentation for the whole image. By substituting the estimated $E_\beta\left[J(H_i)\right]$ into (6.17), we can also obtain the average pixel-level unbiased bound, which will be used in the next section for comparison purposes.

6.3   Experiments and Analysis

In this section, we verify the efficiency of the presented MSE bound by comparing it with the segmentation results of several representative image segmentation algorithms using both synthetic and real-world image data.

6.3.1   Experiment Configuration

The two synthetic images considered here include one image with hard labels and one with hybrid labels. Hybrid here means that some pixels have hard labels and others have fuzzy labels. The real-world image is a cut of a mammogram, containing micro-

calcifications, which is from the Digital Database for Screening Mammography (DDSM) [91]. The micro-calcifications are identified by the radiologists, which are used as the ground-truth in our work.

When calculating the empirical second-order statistics, we employ the *SSIM* index [108] to find similar image patches, as mentioned before. *SSIM* measures the similarity between two images using structure information, which was shown in (4.12) and recapped in (6.37)

$$SSIM\ (Y_1, Y_2) = \frac{(2\mu_{Y_1}\mu_{Y_2} + C_1)(2\sigma_{Y_1Y_2} + C_2)}{(\mu_{Y_1}^2 + \mu_{Y_2}^2 + C_1)(\sigma_{Y_1}^2 + \sigma_{Y_2}^2 + C_2)} \tag{6.37}$$

where μ_{Y_1}, μ_{Y_2} and σ_{Y_1}, σ_{Y_2} as well as $\sigma_{Y_1Y_2}$ denote mean intensity and contrast as well as the correlation coefficient of images Y_1 and Y_2, respectively; C_1 and C_2 are constants used to avoid instabilities for very small μ or σ. The value of $SSIM(Y_1, Y_2)$ is between 0 and 1. A higher value means more similarity between two images. In this section, Y_1 and Y_2 are two image patches under comparison, instead of the entire images used in (4.12).

Admittedly, the patch size, the number of similar patches found for one coefficient, the spline type and even the distance between two neighboring knots have an impact on the resulting bound. We have carried out the experiments by varying these parameters over reasonable ranges and found that the following configuration yields robust and efficient bounds. The patch size is 13 by 13 pixels, the knots are deployed every 4 pixels in both horizontal and vertical directions, and the spline function is cubic *B*-spline. There are two constraints to determine the number of patches: (i) the patches with the *SSIM* index larger than 0.7 are considered as patches similar to the underlying patch; (ii) the first 20 patches with the largest index values are considered as similar patches if the

number of patches selected by (i) exceeds 20.

As a further verification of the biased estimator assumption and Affine bias model, the unbiased bound discussed in Section 6.2.2 is also calculated for comparison purposes.

6.3.2   Segmentation Algorithms

The algorithms for hard image segmentation include the MRF-based algorithm [1275], Otsu thresholding [136][137], Gaussian assumption-based dynamic clustering (GADC) algorithm [80], the region-based active contour model (RACM) [138], and the multi-scale normalized cuts-based segmentation (MNCut) [139], where RACM and MNCut are more recent and can be considered as the state-of-the-art segmentation algorithms. Those for fuzzy image segmentation include fuzzy *C*-means [160], fuzzy *k*-nearest neighbor (fuzzy *k*-NN) [161], and the Gath-Geva algorithm [162].

MRF-based algorithm, Otsu thresholding, GADC, RACM and MNCut algorithms have been introduced in Section 5.2.5.2 and Section 3.2.1.3.

The fuzzy *C*-means clustering algorithm is based on the minimization of the *C-means functional* which is used as the objective function. The minimization of the *C-means functional* is a nonlinear optimization problem that can be solved by using a variety of available methods. The most popular one is a Picard iteration through the first-order conditions for the stationary points of the *C-means functional*. The algorithm yields the weighted mean of the data items that belong to a cluster, where the weights are the membership values.

Fuzzy *k*-NN is a fuzzy version of the crisp *k*-NN algorithm, in which fuzzy sets are introduced into the algorithm. The basic step of the fuzzy *k*-NN algorithm is to assign

membership of a vector as a function of the vector's distance from its *k*-nearest neighbors and those neighbors' memberships in the possible classes.

The Gath-Geva algorithm uses a distance norm based on the fuzzy maximum likelihood estimates. This distance norm involves an exponential term and thus decreases faster than the inner-product norm. The membership degrees are interpreted as the posterior probabilities of selecting the ith cluster given a data point. Gath and Geva [162] reported that the fuzzy maximum likelihood estimates clustering algorithm is able to detect clusters of varying shapes, sizes and densities.

6.3.3   Experimental Results

Fig. 6.1 (a) shows a synthetic hard image with three intensity values, where the square in the upper-left corner has the intensity 90, the central arc has intensity 88, and the rest has intensity 80. White Gaussian noise is added into the image with zero mean and variance σ^2. Fig. 6.1 (b) shows the MSE curves of the segmentation results using the above five hard image segmentation algorithms as well as the bound calculated using (6.33) based on the biased estimator assumption and Affine bias model. Fig. 6.1 (c) shows the variance curves of these segmentation algorithms and the bound calculated using (6.17) where we assume that the segmentation algorithms are unbiased estimators. The bounds, MSEs and variances are calculated for the particular image of Fig. 6.1 (a) under different noise strengths, i.e., different SNRs. At each SNR, the MSE and variance of each segmentation algorithm are the averages of 100 segmentation results. This procedure is used for all the experiments in this chapter.
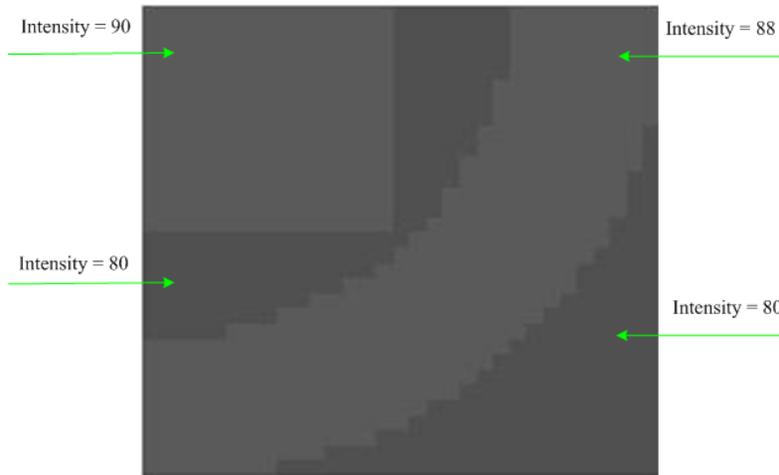
(a)
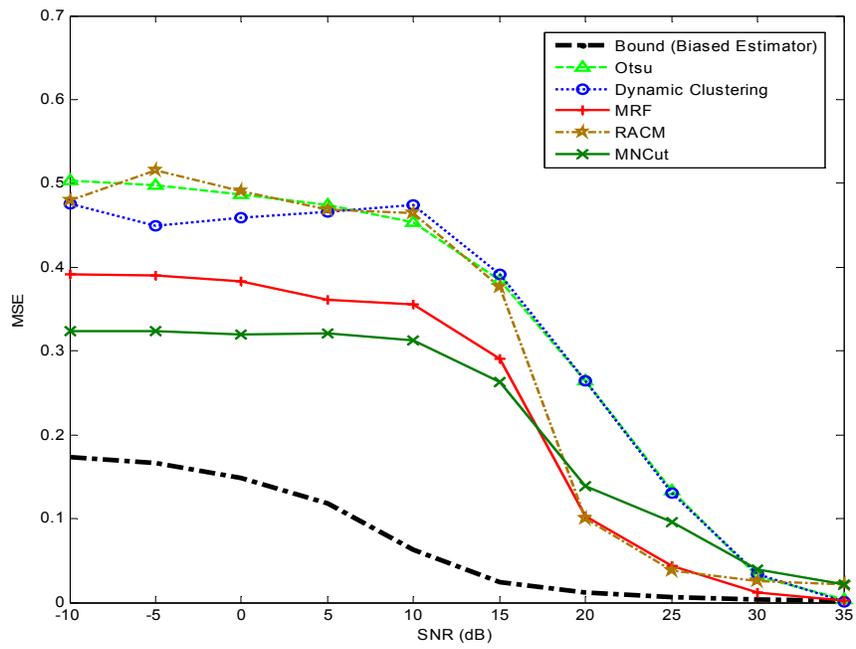
(b)
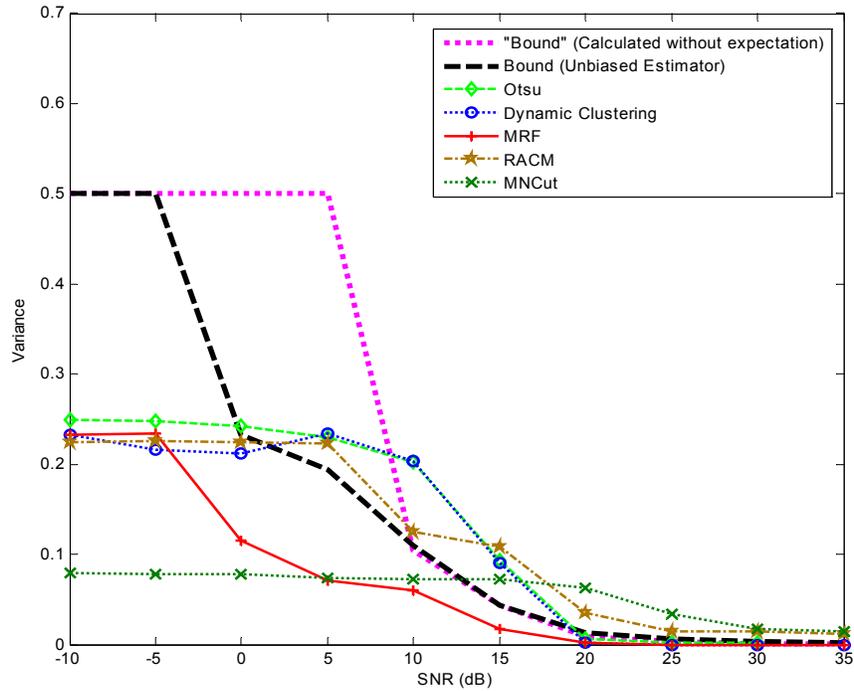
(c)

Figure 6.1: Bounds for hard image segmentation (synthetic image). (a) Synthetic hard image; (b) MSEs and bound under the biased estimator assumption; (c) variances and bound under the unbiased estimator assumption.

From Fig. 6.1 (b) we can see that the MSE bound (the bold dashed-dot line in the lower part of the figure) derived under the biased estimator assumption bounds the MSEs of these algorithms from below. With the increase of SNR, the bound and the MSEs decrease. When the SNR is very high, the MSEs converge to the bound. These expected results show that the bound in (6.33) provides a valid performance prediction of the segmentation algorithms and a benchmark of the segmentation results. In comparison, the bound in Fig. 6.1 (c) based on the unbiased estimator assumption, the bold dashed line, fails to bound the variance of these algorithms, which again verifies the reasonability of

the biased estimator and the Affine model assumptions. In Fig. 6.1 (c), we use the bound values of 0.5 to represent the invalid cases where the variances calculated from the unbiased estimator assumption are very large. However, the variance should have a small value, given that the value of the pixel membership function lies in a small range of [0, 1].

From Fig. 6.1 (b), we can see that the MRF-based segmentation algorithm exploits the correlation between neighboring pixels and yields a better result, in terms of smaller MSE, than the methods which consider pixels to be independent when carrying out segmentation, such as dynamic clustering. This also shows the reasonability of our representation of the image using smoothing coefficients and the expectation operation with respect to β when calculating the bound, which take into account the correlation information contained in an image. As a further verification, in Fig. 6.1 (c) we draw the "bound" curve, the dotted line at the right hand side of the unbiased bound, which is based on the unbiased estimator assumption but calculated by using $\left(\beta_i^T b(x) - \beta_{i^s}^T b(x)\right)^2$ directly from the pixel intensity and without the expectation operation with respect to β. We can see that not taking correlation into account yields an even worse result. Similar results can also be seen in Fig. 6.2 (c) and Fig. 6.3 (c).

Fig. 6.2 shows the results when calculating the bounds and MSEs using the real-world mammogram data. We can see that the presented biased estimator-based bound performs satisfactorily in predicting the performance limit of the algorithms, while the one based on the unbiased assumption fails.

(a)
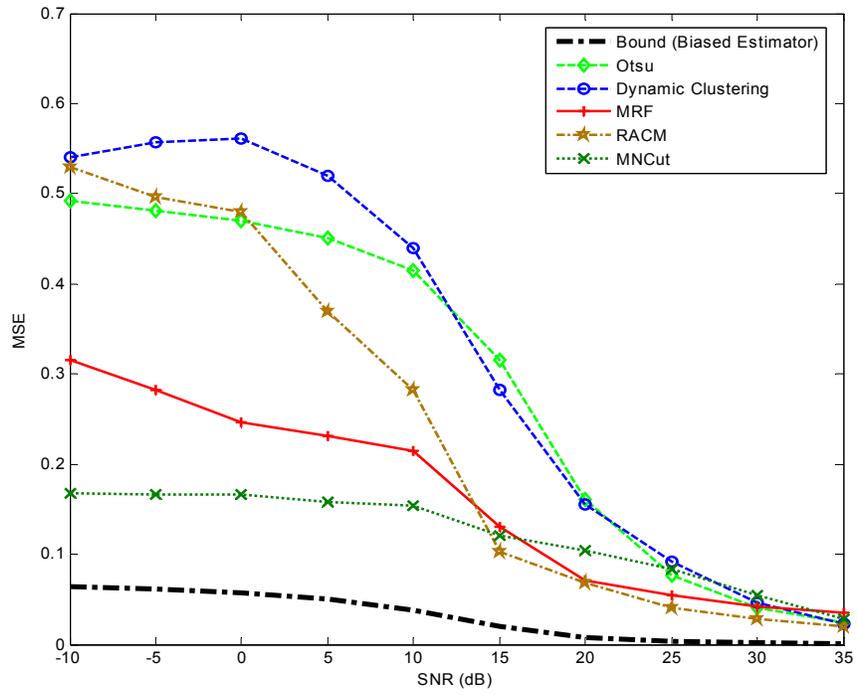
(b)
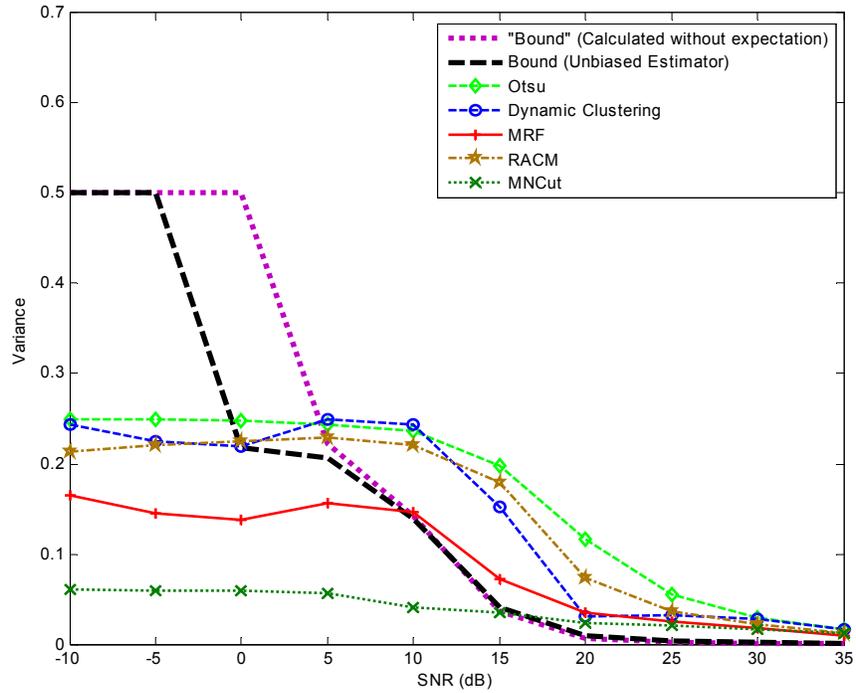
(c)

Figure 6.2: Bounds for hard image segmentation (real-world image). (a) Mammogram with micro-calcifications; (b) MSEs and bound for biased estimator assumption; (c) variances and bound for unbiased estimator assumption.

Fig. 6.3 deals with hybrid image segmentation for the synthetic image shown in Fig. 6.3 (a). There are four basic image regions, corresponding to the intensity values of 120, 90, 60 and 20, respectively. The three arc regions at the left side of the diagonal curves are fuzzy regions, denoted as Regions A, B, and C, and have membership values of [0.5, 0.2, 0.2, 0.1], [0, 0.6, 0.3, 0.1] and [0, 0, 0.8, 0.2], respectively. The rest of the four regions are hard ones with the intensity values mentioned above. Once again white Gaussian noise is added into the image with zero mean and variance σ^2. Fig. 6.3 (b) shows the MSE curves of the segmentation results using the three fuzzy image

segmentation algorithms when the biased estimator assumption and Affine bias model are employed. Fig. 6.3 (c) shows the variance curves of the segmentation algorithms, the bound calculated using (6.17) for the unbiased estimator assumption and the "bound" determined by ignoring the expectation operation. We can see from the figures that the bound based on the biased estimator assumption is valid but those based on the unbiased estimator assumption fail again.
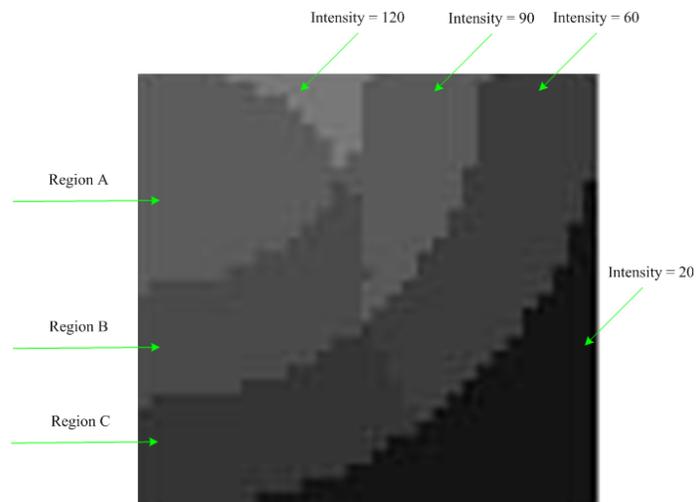
(a)
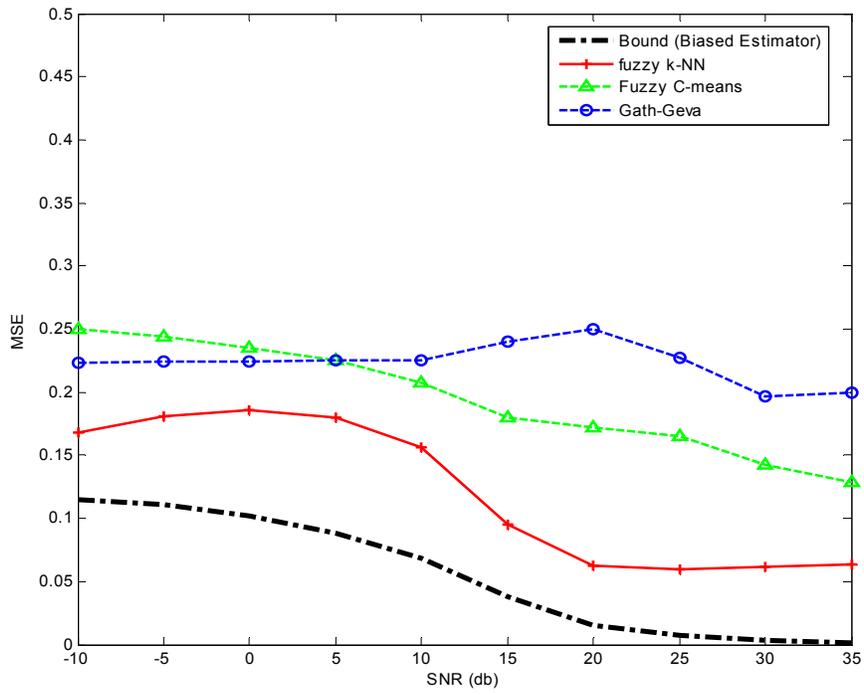
(b)
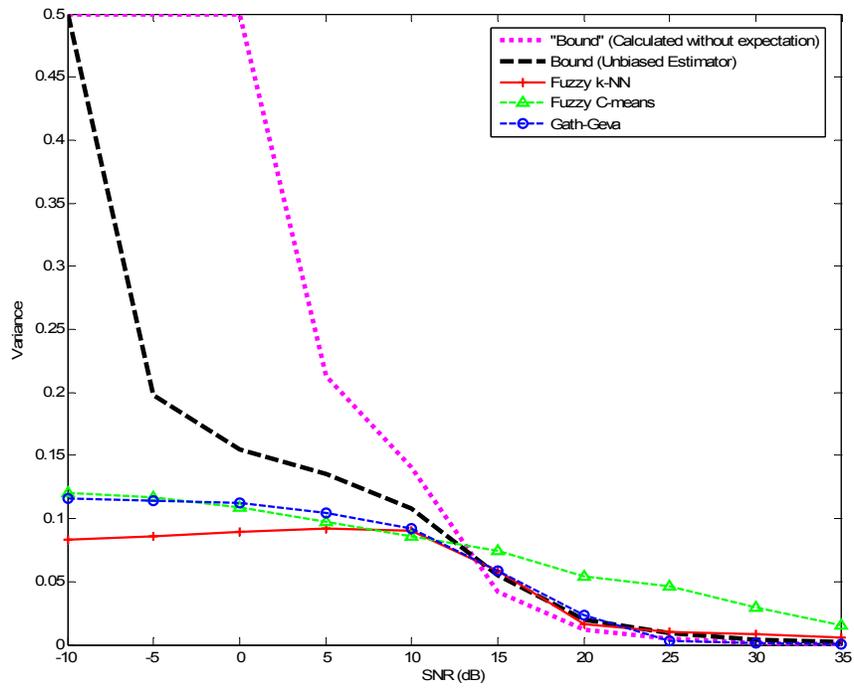
(c)

Figure 6.3: Bounds for hybrid image segmentation (synthetic image). (a) Synthetic hybrid image; (b) MSEs and bound for biased estimator assumption; (c) variances and bound for unbiased estimator assumption.

6.4 Summary

Image segmentation is very important but also very challenging for computer vision and image analysis. However, performance limit of segmentation algorithms are seldom studied from a statistical perspective, which plays a fundamental critical role in developing segmentation algorithms and evaluating segmentation results. This chapter developed a systematic method to determine a lower bound on the MSE of image segmentation algorithms under a statistical estimation framework. The bound was based on the biased estimator assumption and Affine bias model, where an approximation was employed to simplify the computation when determining the expectation on the inverse of the Fisher information matrix. Additionally, non-local searching and boostrapping techniques were used to approximate the unknown second-order statistics during the computation of the bound. The theoretical analysis and experimental results show that the presented bound is efficient and robust in bounding the performance of the segmentation algorithms and providing a benchmark for the segmentation problem.

CHAPTER VII

CONCLUSIONS AND SUGGESTIONS FOR FUTURE WORK

7.1   Conclusions

Object detection in images, image enhancement and image segmentation are all critical but challenging problems in image processing. In this dissertation, we have investigated these three problems based on stochastic resonance (SR) noise and human visual system (HVS) properties. Several frameworks and algorithms have been presented to improve the performance of object detection, image enhancement and image segmentation. In addition, a statistical performance bound has been derived for evaluating and analyzing image segmentation algorithms.

To detect objects of interest in an image is a difficult problem, especially when the image quality is low, which often happens in medical images. As a result, many detection algorithms suffer from suboptimality and yield unsatisfactory results. In this dissertation, we developed a framework and algorithm to improve the performance of suboptimal detectors based on SR noise. We have also developed several schemes to improve the SR noise-based detection improvement system, such as probability density function (*pdf*) learning and SR noise with memory (multi-peak SR noise). The experimental results with large number of image data show that the presented framework and algorithm are flexible, efficient and robust.

Image enhancement is a widely used technique but also a challenging task. Many image enhancement algorithms fail to increase the quality of the image due to either the complexity of the image content and the image degradation process or the unsuitable design of the enhancement procedure. We have presented two image enhancement approaches in this dissertation. The first one was based on SR noise, where we developed a SR noise-refined image enhancement system, and employed HVS-driven objective functions and constrained multi-objective optimization (MOO) techniques to find the optimum parameters of the SR noise distribution. The second approach was based on the selective enhancement framework, where we enhanced the extracted region of interest (ROI) and suppressed the background. Several enhancement algorithms under this framework have been developed. The experimental results with various types of degraded images show that the two presented approaches can achieve superior performance in terms of both subjective and objective evaluations compared with many representative image enhancement algorithms.

Image segmentation plays a fundamentally important role in image analysis. Its challenge lies in the complexity of the image contents and the difficulty in defining appropriate segmentation criteria. As a result, many segmentation algorithms fail either in the objective function design or in the parameter setting. We have developed two HVS-driven image segmentation approaches to improve the segmentation performance. One of the approaches took into account the preference of HVS to good segmentation from both region-based and boundary-based perspectives. Markov random fields (MRF) and the just-noticeable difference (JND) model have been employed to encode the HVS preference into the objective function for image segmentation. In our second algorithm,

we formulated image segmentation as a detection problem, and developed an image segmentation framework by introducing a local "soft" objective function for segmentation. We further employed contrast sensitivity function (*CSF*) as a filter to preprocess the image, which embedded HVS information into the segmentation procedure. Experiments with real-world image data show that the presented approaches outperformed many representative segmentation and clustering algorithms.

Finally, we investigated a very important but seldom studied problem in image segmentation, i.e., the statistical bound or performance limit of segmentation algorithms. We have developed a systematic method to determine a lower bound on the mean square error (MSE) of segmentation algorithms under a statistical estimation framework, based on the biased estimator assumption and Affine bias model. We compared the experimental performance of several representative segmentation algorithms with respect to this performance bound and a bound derived from the unbiased estimator assumption. The efficiency of the bound and the biasedness assumption are verified therein. We also analyzed the impact of image contents on the bound, and explained the factors that lead to the gap between the bound and the actual MSE performance. In particular, we presented analysis and experimental evidence which suggest that the consideration of pixel correlation benefits image segmentation. Furthermore, we showed that studying the performance bounds provides much insight into the image segmentation problem. We expect that this type of analysis would offer guidance to the practitioner for choosing and evaluating segmentation algorithms for a given image.

7.2   Suggestions for Future Work

In this section, we suggest the following problems that are worth studying in the future.

- The first part of the future work is about the SR noise-enhanced micro-calcification detection approach presented in Chapter 3. The investigation on optimizing the SR noise-based technique with memory, by determining the optimum weights for two-peak SR noises, will be very useful to further improve the efficiency and robustness of the SR noise-based detection enhancement scheme. Extension of the SR noise-based technique to enhancing fixed multiple threshold detectors is also an important research issue. The performance of SR enhanced variable detectors [54] has been shown to be superior to the fixed ones, where both the SR noise and the critical function can be jointly designed to enhance detection. So SR noise-based detectors incorporating variable critical function are likely to be promising. In our current work, we only considered the case where signal and background noise are all independently distributed. Future research on the correlated signal and noise case may further improve the detection performance. In addition, the application of the detection schemes developed in this dissertation to other two types of mammogram lesions, i.e., mass and spiculated lesions, and even other medical images, will be of great interest. Finally, the SR noise-enhanced scheme may also be useful in color images, which could be an excellent extension of our work to more real world applications.

- The second part of the future work is about the SR noise-refined image enhancement scheme presented in Chapter 4. In our experiments, we investigated the effect of SR noise on different distortion situations, such as image sharping, noise reduction and

image de-convolution, and showed some of the results in this dissertation. The very encouraging results indicate that we may achieve significant performance improvement when applying the presented SR noise refinement scheme to other enhancement and restoration methods, such as super-resolution. Moreover, the extension to video enhancement may also be very promising. In this dissertation, we introduced independent SR noise in the image, but the performance may be further improved if correlated SR noise is employed. An investigation on other types of SR noises will be a very interesting topic to improve the quality of the enhanced image. Future research on ameliorating the weighting scheme may also improve the enhancement performance. Finally, the SR noise-refined image enhancement scheme may also be useful in color images, which could be an excellent extension of our work to more real world applications.

- The third part of the future work is about the HVS-driven image segmentation algorithm presented in Chapter 5. An investigation on the segmentation with multiple region types or multiple pixel labels, instead of the binary labels, 0 and 1, as discussed in this dissertation, will be an interesting extension to the algorithm. This would, of course, require the presented objective function to be adjusted accordingly. In our current work, we mainly discussed the problem of hard or crisp segmentation, that is, a pixel belongs to either region 0 or region 1. Future research on fuzzy segmentation, based on the objective function presented in Chapter 5, will be another interesting research topic. Perhaps, the scheme of designing the objective function for fuzzy MRF, as discussed in [146], will be very helpful. Finally, in Chapter 5 we employed iterative conditional modes (ICM) for the optimization task,

and the research on approaches which can further improve the computational speed and segmentation accuracy will be both theoretically and practically useful. Graph-cut based methods [27][135] may be promising options.

- The fourth part of the future work is about the method to determine the performance limit of image segmentation algorithms, which was discussed in Chapter 6. An investigation on the probability distribution estimation techniques may be helpful to improve the computation of the expectation involved in the bound, where statistical learning methods may be helpful. In our current work, we mainly discussed the problem of segmenting a single image, and only mentioned multi-spectral image segmentation in Appendix C and did not consider the 3D scenario. Future research on the extension of the developed bound to the multi-spectral and 3D images will be an interesting research topic. When developing the bound, the ground truth information about the noise-free image and the membership value of each pixel label is required. Research on approaches which can reduce the dependence of the bound on such information will be both theoretically and practically useful. Perhaps image denoising and linear regression techniques will be helpful in handling it. Finally, the presented bound may also be useful in color images, which could be an excellent extension of our work to more real-world applications.

APPENDIX A

REASONABILITY OF ENERGY FUNCTION (5.19)

In this appendix, we show that the multi-pie slice configuration, with $N_{PS}(s)$ pie slices, may have a large angle value, which can be obtained by adding the angles of the $N_{PS}(s)$ pie slices together, but will unnecessarily result in a smooth boundary and thus a lower energy due to the exponential function in (5.19).

We assume that the cross-boundary contrast and the interior contrast of each pie slice are the same as each other. Then (5.19) can be written as

$$\phi(s) = \sum_{i=1}^{N_{PS}(s)} \exp[-a_3 \cdot \alpha_i(s) - a_4 \cdot Contr(s)_{CB-i} + a_5 \cdot Contr(s)_{PS-i}] = \exp[Contrast(s)] \sum_{i=1}^{N_{PS}(s)} \exp(-a_3 \cdot \alpha_i(s)) \quad \text{(A.1)}$$

where $Contrast(s) = -a_4 \cdot Contr(s)_{CB-i} + a_5 \cdot Contr(s)_{PS-i}$. Suppose we have a favorite segmentation with a single pie slice in the second-order neighborhood system of s. Its contrast is equal to $Contrast(s)$, and its angle value, $\alpha_{favorite}(s)$, equals the summation of the angle values in (A.1), i.e., $\alpha_{favorite}(s) = \sum_{i=1}^{N_{PS}(s)} \alpha_i(s)$. Then, it is not difficult to prove that

$$\exp(-a_3 \cdot \alpha_{favorite}(s)) = \exp\left(-a_3 \cdot \sum_{i=1}^{N_{PS}(s)} \alpha_i(s)\right) \le \sum_{i=1}^{N_{PS}(s)} \exp(-a_3 \cdot \alpha_i(s)) \quad \text{(A.2)}$$

given that $\alpha_i(s) \ge 0$. Therefore, we have

$$\phi(s)_{favorite} \le \phi(s) \quad \text{(A.3)}$$

Thus, the favorite segmentation has lower energy than the multi-pie slice configuration and therefore has a higher probability to survive.

APPENDIX B

CALCULATING FISHER INFORMATION MATRIX (FOR SINGLE IMAGE)

Assume that the noise $w(x)$ is *i.i.d.* Gaussian random variable with zero mean and variance σ^2, and the observed pixel intensity is also *i.i.d.* given the membership H and the coefficient β. Then the conditional *pdf* of the observation is

$$P(Y;H,\beta) = \left(\frac{1}{\sqrt{2\pi\sigma^2}}\right)^N \exp\left(\frac{-\sum_{x=1}^{N}[y(x)-h(x)^T \cdot \beta \cdot b(x)]^2}{2\sigma^2}\right) \tag{B.1}$$

So

$$
\begin{aligned}
L &= \ln[P(Y;H,\beta)] \\
&= -\frac{N}{2}\ln 2\pi\sigma^2 - \frac{1}{2\sigma^2}\sum_{x=1}^{N}[y(x)-h(x)^T \cdot \beta \cdot b(x)]^2 \\
&= -\frac{N}{2}\ln 2\pi\sigma^2 - \frac{1}{2\sigma^2}\sum_{x=1}^{N}\left\{ \begin{aligned} &y(x)-[h_1(x)(\beta_{11}b_1(x)+\beta_{12}b_2(x)+\cdots+\beta_{1m}b_m(x))+h_2(x)(\beta_{21}b_1(x)+\beta_{22}b_2(x)+\cdots+\beta_{2m}b_m(x)) \\ &+\cdots+h_M(x)(\beta_{M1}b_1(x)+\beta_{M2}b_2(x)+\cdots+\beta_{Mm}b_m(x))] \end{aligned} \right\}^2
\end{aligned}
\tag{B.2}
$$

Therefore,

$$\frac{\partial L}{\partial h_k(x)} = \frac{1}{\sigma^2}\beta_k^{\ T} \cdot b(x) \cdot w(x) \tag{B.3}$$

where $k = 1,2,\cdots,M$.

$$
\begin{aligned}
E\left\{\left[\frac{\partial L}{\partial h_k(x)}\right]\left[\frac{\partial L}{\partial h_{k'}(x')}\right]^T\right\} &= E\left\{\frac{1}{\sigma^4}\beta_k^{\ T} \cdot b(x) \cdot w(x) \cdot w(x') \cdot b(x')^T \cdot \beta_{k'}\right\} \\
&= \begin{cases} \frac{1}{\sigma^2}\beta_k^{\ T} \cdot b(x) \cdot b(x)^T \cdot \beta_{k'}, & if \quad x = x' \\ 0, & if \quad x \neq x' \end{cases}
\end{aligned}
\tag{B.4}
$$

Fisher information matrix is determined as follows,

$$[J_F(H)]_{ij} = E_{Y|H,\beta}\left\{\left[\frac{\partial L}{\partial h_i}\right]\left[\frac{\partial L}{\partial h_j}\right]^T\right\} \tag{B.5}$$

and

$$J_F(H) = E_{Y|H,\beta}\left\{\left[\frac{\partial L}{\partial h_k(x)}\right]\left[\frac{\partial L}{\partial h_{k'}(x')}\right]^T\right\}_{MN \times MN} =$$

$$\frac{1}{\sigma^2}\begin{bmatrix}
\beta_1^T b(1)b(1)^T \beta_1 & \cdots & \beta_1^T b(1)b(1)^T \beta_M & 0 & \cdots & 0 & \cdots & 0 & \cdots & 0 \\
\beta_2^T b(1)b(1)^T \beta_1 & \cdots & \beta_2^T b(1)b(1)^T \beta_M & 0 & \cdots & 0 & \cdots & 0 & \cdots & 0 \\
& \vdots & & & \vdots & & \vdots & & \vdots & \\
\beta_M^T b(1)b(1)^T \beta_1 & \cdots & \beta_M^T b(1)b(1)^T \beta_M & 0 & \cdots & 0 & \cdots & 0 & \cdots & 0 \\
0 & \cdots & 0 & \beta_1^T b(2)b(2)^T \beta_1 & \cdots & \beta_1^T b(2)b(2)^T \beta_M & \cdots & 0 & \cdots & 0 \\
0 & \cdots & 0 & \beta_2^T b(2)b(2)^T \beta & \cdots & \beta_2^T b(2)b(2)^T \beta_M & \cdots & 0 & \cdots & 0 \\
& \vdots & & & \vdots & & \cdots & & \vdots & \\
0 & \cdots & 0 & \beta_M^T b(2)b(2)^T \beta_1 & \cdots & \beta_M^T b(2)b(2)^T \beta_M & \cdots & 0 & \cdots & 0 \\
& \vdots & & & \vdots & & \cdots & & \vdots & \\
0 & \cdots & 0 & 0 & \cdots & 0 & \cdots & \beta_1^T b(N)b(N)^T \beta_1 & \cdots & \beta_1^T b(N)b(N)^T \beta_M \\
0 & \cdots & 0 & 0 & \cdots & 0 & \cdots & \beta_2^T b(N)b(N)^T \beta_1 & \cdots & \beta_2^T b(N)b(N)^T \beta_M \\
& \vdots & & & \vdots & & \cdots & & \vdots & \\
0 & \cdots & 0 & 0 & \cdots & 0 & \cdots & \beta_M^T b(N)b(N)^T \beta & \cdots & \beta_M^T b(N)b(N)^T \beta_M
\end{bmatrix}$$

$$\tag{B.6}$$

APPENDIX C

CALCULATING FISHER IFNROMATION MATRIX (FOR MULTI-SPECTRAL IMAGES)

For a multi-spectral image set including P images, H is the same for all of them, b can be different if the smoothing configuration, such as the number, position and size of the spacing of knots, are different from one image to another, but β usually are different for different images. Therefore, we have the segmentation model

$$y^i(x) = h(x)^T \phi^i(x;\beta) + w^i(x) = h(x)^T \cdot \beta^i \cdot b^i(x) + w^i(x) \tag{C.1}$$

where $i = 1,2,\cdots,P$, $x = 1,2,\cdots,N$, $b^i(x) = [b_1^i(x), b_2^i(x),\cdots,b_{m_i}^i(x)]^T$, $\beta^i = [\beta^{i\,T}_1, \beta^{i\,T}_2,\cdots,\beta^{i\,T}_M]^T$, and

$\beta^i_k = [\beta^i_{k_1}, \beta^i_{k_2},\cdots,\beta^i_{km_i}]^T$. The noise may be different, so we assume $w^i(x)$ are *i.i.d.* Gaussian noise with zero mean and variance σ_i^2 . For simplicity, we use the same knot configuration for every image. Then the model is simplified to

$$y^i(x) = h(x)^T \cdot \beta^i \cdot b(x) + w^i(x) \tag{C.2}$$

We still assume that the observed pixel intensities are *i.i.d.* given the membership H and the coefficient β , so the conditional *pdf* of the observation is

$$P(Y;H,\beta^1,\beta^2,\cdots,\beta^P) = \left(\prod_{i=1}^{P} \frac{1}{\sqrt{2\pi\sigma_i^2}}\right)^N \exp\left(-\sum_{x=1}^{N}\sum_{i=1}^{P} \frac{[y^i(x) - h(x)^T \cdot \beta^i \cdot b(x)]^2}{2\sigma_i^2}\right)$$

$$= \prod_{i=1}^{P}\left(\frac{1}{\sqrt{2\pi\sigma_i^2}}\right)^N \exp\left(-\sum_{i=1}^{P} \frac{\sum_{x=1}^{N}[y^i(x) - h(x)^T \cdot \beta^i \cdot b(x)]^2}{2\sigma_i^2}\right) \tag{C.3}$$

So

$$L = \ln[P(Y; H, \beta^1, \beta^2, \cdots, \beta^P)]$$

$$= -\frac{N}{2}\sum_{i=1}^{P}\ln 2\pi\sigma_i^2 - \sum_{i=1}^{P}\frac{1}{2\sigma_i^2}\sum_{x=1}^{N}[y^i(x) - h(x)^T \cdot \beta^i \cdot b(x)]^2 \tag{C.4}$$

$$= -\sum_{i=1}^{P}\frac{1}{2\sigma_i^2}\sum_{x=1}^{N}\{y^i(x) - [h_1(x)(\beta_{11}{}^i b_1(x) + \beta_{12}{}^i b_2(x) + \cdots + \beta_{1m}{}^i b_m(x)) + \cdots + h_M(x)(\beta_{M1}{}^i b_1(x) + \beta_{M2}{}^i b_2(x) + \cdots + \beta_{Mm}{}^i b_m(x))]\}^2$$

$$\quad -\frac{N}{2}\sum_{i=1}^{P}\ln 2\pi\sigma_i^2$$

Therefore,

$$\frac{\partial L}{\partial h_k(x)} = \sum_{i=1}^{P}\frac{1}{\sigma_i^2}\beta^i{}_k{}^T \cdot b(x) \cdot w^i(x) \tag{C.5}$$

$$E\left\{\left[\frac{\partial L}{\partial h_k(x)}\right]\left[\frac{\partial L}{\partial h_{k'}(x')}\right]^T\right\} = E\left\{\left[\sum_{i=1}^{P}\frac{1}{\sigma_i^2}\beta^i{}_k{}^T \cdot b(x) \cdot w^i(x)\right]\left[\sum_{i=1}^{P}\frac{1}{\sigma_i^2}\beta^i{}_k{}^T \cdot b(x') \cdot w^i(x')\right]^T\right\}$$

$$= E\left\{\sum_{j=1}^{P}\sum_{i=1}^{P}\frac{1}{\sigma_i^2}\beta^i{}_k{}^T \cdot b(x) \cdot w^i(x)\frac{1}{\sigma_j^2}b(x')^T \cdot \beta^j{}_{k'} \cdot w^j(x')\right\} \tag{C.6}$$

$$= E\left\{\sum_{j=1}^{P}\sum_{i=1}^{P}\frac{1}{\sigma_i^2\sigma_j^2}\beta^i{}_k{}^T \cdot b(x) \cdot w^i(x) \cdot w^j(x') \cdot b(x')^T \cdot \beta^j{}_{k'}\right\}$$

$$= \begin{cases} \sum_{i=1}^{P}\frac{1}{\sigma_i^2}\beta^i{}_k{}^T \cdot b(x) \cdot b(x)^T \cdot \beta^i{}_{k'}, & \text{if } x = x' \text{ and } i = j \\ 0, & \text{if } x = x' \end{cases}$$

So we have

$$J_F(H) = E_{Y|H\beta}\left\{\left[\frac{\partial L}{\partial h_k(x)}\right]\left[\frac{\partial L}{\partial h_{k'}(x')}\right]^T\right\}_{MN \times MN} =$$

$$\tag{C.7}$$

APPENDIX D

JUSTIFICATION OF THE BIASED ESTIMATOR ASSUMPTION AND AFFINE BIAS MODEL

The estimation problem in linear models was analyzed in [163]~[165]. The linear model is

$$Y = Q\theta + n \tag{D.1}$$

where Y is the observation, θ is a parameter vector, Q is a model matrix, and n is zero-mean random vector. The estimator of θ is assumed to be linear, i.e., $\hat{\theta} = GY$, which estimates θ by performing a weighted average operation over the observation. Linear estimators are quite frequently used for least squire estimation problems, whose forms have been established by solving optimization problems, with the constraints put on Q, θ and even n. These constraints can be considered as the prior information on these parameters and the penalties under the regularization framework.

Similarly, image segmentation can also be modeled as a linear estimation problem, as shown in (6.4)

$$\begin{aligned} y(x) &= h(x)^T \cdot \phi(x; \beta) + w(x) \\ &= h(x)^T \cdot \beta \cdot b(x) + w(x) \end{aligned} \tag{6.4}$$

where $\phi(x; \beta)^T$ can be considered as the model matrix and $h(x)$ is the label parameter vector to be estimated. During the segmentation procedure, some prior information about $\phi(x; \beta)$, $h(x)$ and $w(x)$ is usually employed as the penalty terms of the objective functions for segmentation, to reduce the solution space under regularization framework. For example, the smoothness assumption is often made on the labels of the neighboring pixels, like that

used in the MRF-based algorithms, which equivalently brings the constraint on $h(x)$.

Moreover, local information is often used during the estimation procedure, that is, $h(x)$ is

often estimated by using the observation Y around the coordinate x. Thus, it is reasonable

to assume that many image segmentation algorithms, especially the state-of-the-art ones,

perform the label estimation using linear estimators $\hat{H} = GY$.

Here, we consider the penalty or prior information resulting from the label smoothness

assumption, and assume that H forms Gaussian MRF

$$
\begin{aligned}
h(x_j) &= \sum_{x_{j_l} \in \eta_{x_j}} a_{j_l} h\big(x_{j_l}\big) + \omega(x_j) \\
&= \sum_{x_{j_l} \in \eta_{x_j}, x_{j_l} \neq x} a_{j_l} h\big(x_{j_l}\big) + \omega(x_j) + a_x h(x)
\end{aligned}
\tag{D.2}
$$

where x_{j_l} denotes the indices of the lth neighbor of the pixel x_j in the neighborhood

system η_{x_j} of x_j, and $\omega(x_j)$ is zero mean Gaussian noise vector. Pixel x also belongs to

η_{x_j}. a_{j_l} and a_x are the model parameters. In this dissertation, two pixels are called

neighbors if they are close to each other spatially and their observations have an impact

on the estimation of the pixel labels of each other. So it is not compulsory for two

neighboring pixels to be deployed in a way that one is followed immediately by another

spatially.

With the neighboring information incorporated in the segmentation procedure, the

linear estimator finds the weighted average over the observation in a local window. We

can also consider that the weighted average is carried out over the whole set of

observations in an image, but the weights decrease with the increase of the distance

between the coordinates of the observations and the pixel of interest. Here, we only

consider the observations which are neighbors of the pixel of interest. We have

$$\hat{h}(x) = G_x Y_x \tag{D.3}$$

where G_x and Y_x are the weighting matrix and observation vector corresponding to a neighborhood system of the pixel at x. More specifically, $G_x = [g_1, g_2, \cdots, g_M]^T$ and $g_i = [g_{i1}, g_{i2}, \cdots, g_{iC(x)}]^T$, where $i = 1, 2, \cdots, M$, and $C(x)$ is the total number of neighboring pixels of pixel x. $C(x)$ is equal to the size of η_x, and may be different from pixel to pixel . $Y_x = [y(x_1), y(x_2), \cdots, y(x_{C(x)})]^T$, which is the vector consisting of the neighboring pixels of x.

We claim that if pixel x_j is the neighbor of pixel x, then pixel x is the neighbor of pixel x_j. Thus, substituting (6.4) and (D.2) into (D.3), we have

$$
\begin{aligned}
\hat{h}(x) &= G_x Y_x \\
&= \sum_j g_j y(x_j) \\
&= \sum_j g_j \left(\phi(x_j; \beta)^T \cdot h(x_j) + w(x_j) \right) \\
&= \sum_j g_j \left\{ \phi(x_j; \beta)^T \cdot \left(\sum_{x_{j_l} \in \eta_{x_j}, x_{j_l} \neq x} a_{j_l} h(x_{j_l}) + \omega(x_j) + a_x h(x) \right) + w(x_j) \right\} \\
&= \sum_j g_j \left\{ \phi(x_j; \beta)^T \cdot a_x h(x) + \phi(x_j; \beta)^T \cdot \left(\sum_{x_{j_l} \in \eta_{x_j}, x_{j_l} \neq x} a_{j_l} h(x_{j_l}) + \omega(x_j) \right) + w(x_j) \right\} \\
&= \sum_j g_j \left[\phi(x_j; \beta)^T \cdot a_x \right] h(x) + \sum_j g_j \left[\phi(x_j; \beta)^T \cdot \left(\sum_{x_{j_l} \in \eta_{x_j}, x_{j_l} \neq x} a_{j_l} h(x_{j_l}) + \omega(x_j) \right) + w(x_j) \right]
\end{aligned}
\tag{D.4}
$$

where g_j is the jth column of the matrix G_x and $j = 1, 2, \cdots, C(x)$. $g_j \left[\phi(x_j; \beta)^T \cdot a_x \right]$ is a $M \times M$ matrix, and $\phi(x_j; \beta)^T \cdot \left(\sum_{x_{j_l} \in \eta_{x_j}, x_{j_l} \neq x} a_{j_l} h(x_{j_l}) + \omega(x_j) \right)$ is a scalar. The expected value of this linear estimator, given the true value of $h(x)$, is

$$E\{\hat{h}(x)\mid h(x)\}=E\left\{\sum_{j}g_{j}\left[\phi(x_{j};\beta)^{T}\cdot a_{x}\right]h(x)+\sum_{j}g_{j}\left[\phi(x_{j};\beta)^{T}\cdot\left(\sum_{x_{j_{l}}\in\eta_{x_{j}},x_{j_{l}}\neq x}a_{j_{l}}h(x_{j_{l}})+\omega(x_{j})\right)+w(x_{j})\right]\right\}$$

$$=E\left\{\sum_{j}g_{j}\left[\phi(x_{j};\beta)^{T}\cdot a_{x}\right]h(x)\right\}+E\left\{\sum_{j}g_{j}\left[\phi(x_{j};\beta)^{T}\cdot\left(\sum_{x_{j_{l}}\in\eta_{x_{j}},x_{j_{l}}\neq x}a_{j_{l}}h(x_{j_{l}})+\omega(x_{j})\right)+w(x_{j})\right]\right\}$$

$$=E\left\{\sum_{j}g_{j}\left[\phi(x_{j};\beta)^{T}\cdot a_{x}\right]\right\}\cdot h(x)+E\left\{\sum_{j}g_{j}\left[\phi(x_{j};\beta)^{T}\cdot\left(\sum_{x_{j_{l}}\in\eta_{x_{j}},x_{j_{l}}\neq x}a_{j_{l}}h(x_{j_{l}})+\omega(x_{j})\right)+w(x_{j})\right]\right\}$$

(D.5)

So the bias vector of the linear estimator is

$$g(h(x))=E\{\hat{h}(x)\mid h(x)\}-h(x)$$

$$=E\left\{\sum_{j}g_{j}\left[\phi(x_{j};\beta)^{T}\cdot a_{x}\right]\right\}\cdot h(x)+E\left\{\sum_{j}g_{j}\left[\phi(x_{j};\beta)^{T}\cdot\left(\sum_{x_{j_{l}}\in\eta_{x_{j}},x_{j_{l}}\neq x}a_{j_{l}}h(x_{j_{l}})+\omega(x_{j})\right)+w(x_{j})\right]\right\}-h(x)$$ (D.6)

$$=\left(E\left\{\sum_{j}g_{j}\left[\phi(x_{j};\beta)^{T}\cdot a_{x}\right]\right\}-I\right)\cdot h(x)+E\left\{\sum_{j}g_{j}\left[\phi(x_{j};\beta)^{T}\cdot\left(\sum_{x_{j_{l}}\in\eta_{x_{j}},x_{j_{l}}\neq x}a_{j_{l}}h(x_{j_{l}})+\omega(x_{j})\right)+w(x_{j})\right]\right\}$$

$$=K_{x}\cdot h(x)+u_{x}$$

where $K_{x}=E\left\{\sum_{j}g_{j}\left[\phi(x_{j};\beta)^{T}\cdot a_{x}\right]\right\}-I$ and $u_{x}=E\left\{\sum_{j}g_{j}\left[\phi(x_{j};\beta)^{T}\cdot\left(\sum_{x_{j_{l}}\in\eta_{x_{j}},x_{j_{l}}\neq x}a_{j_{l}}h(x_{j_{l}})+\omega(x_{j})\right)+w(x_{j})\right]\right\}$.

The subscript "x" of K_{x} and u_{x} means that these two quantities are relevant to pixel x. K_{x} and u_{x} can be further decomposed for each region type. That is, $K_{x}=[K_{x:1},K_{x:2},\cdots,K_{x:M}]^{T}$ and $u_{x}=[u_{x:1},u_{x:2},\cdots,u_{x:M}]^{T}$. Here, for the ith region type, $K_{x:i}=[K_{x:i,1},K_{x:i,2},\cdots,K_{x:i,M}]^{T}$, a $M\times1$ vector, and $u_{x:i}$ is a scalar, $i=1,2,\cdots,M$.

In the "super" region scheme employed in our work, we have two regions, i.e., i^{th} and i^{sth} regions, when we consider the segmentation performance for the ith region. So, $M=2$, $K_{x}=[K_{x:i},K_{x:i^{s}}]^{T}$, $u_{x}=[u_{x:i},u_{x:i^{s}}]^{T}$ and $K_{x:i}=[K_{x:i,1},K_{x:i,2}]^{T}$. From (D.6) we have

$$
\begin{aligned}
g(h_i(x)) &= E\{\hat{h}_i(x) \mid h(x)\} - h_i(x) \\
&= E\{\hat{h}_i(x) \mid h_i(x), h_{i^S}(x)\} - h_i(x) \\
&= \left(K_{x:i,1} \cdot h_i(x) + K_{x:i,2} \cdot h_{i^S}(x)\right) + u_{x:i} \\
&= \left(K_{x:i,1} - K_{x:i,2}\right) \cdot h_i(x) + K_{x:i,2} + u_{x:i} \\
&= K'_{x:i} \cdot h_i(x) + u'_{x:i}
\end{aligned}
\tag{D.7}
$$

where $K'_{x:i} = K_{x:i,1} - K_{x:i,2}$ and $u'_{x:i} = K_{x:i,2} + u_{x:i}$ and we have employed the relation of

$h_i(x) + h_{i^S}(x) = 1$ in the derivation. Therefore, we have

$$
\begin{aligned}
g(H_i) &= E\{\hat{H}_i \mid H_i\} - H_i \\
&= E\{\hat{H}_i \mid H_i, H_{i^S}\} - H_i \\
&= \left(\begin{bmatrix} K_{1:i,1}h_i(1) \\ K_{2:i,1}h_i(2) \\ \vdots \\ K_{x:i,1}h_i(x) \\ \vdots \\ K_{N:i,1}h_i(N) \end{bmatrix} + \begin{bmatrix} K_{1:i,2}h_{i^S}(1) \\ K_{2:i,2}h_{i^S}(2) \\ \vdots \\ K_{x:i,2}h_{i^S}(x) \\ \vdots \\ K_{N:i,2}h_{i^S}(N) \end{bmatrix}\right) + \begin{bmatrix} u_{1:i} \\ u_{2:i} \\ \vdots \\ u_{x:i} \\ \vdots \\ u_{N:i} \end{bmatrix} \\
&= \begin{bmatrix} \left(K_{1:i,1} - K_{1:i,2}\right)h_i(1) \\ \left(K_{2:i,1} - K_{2:i,2}\right)h_i(2) \\ \vdots \\ \left(K_{x:i,1} - K_{x:i,2}\right)h_i(x) \\ \vdots \\ \left(K_{N:i,1} - K_{N:i,2}\right)h_i(N) \end{bmatrix} + \begin{bmatrix} K_{1:i,2} + u_{1:i} \\ K_{2:i,2} + u_{2:i} \\ \vdots \\ K_{x:i,2} + u_{x:i} \\ \vdots \\ K_{N:i,2} + u_{N:i} \end{bmatrix} \\
&= K_i \cdot H_i + u_i
\end{aligned}
\tag{D.8}
$$

where
$$
K_i = \begin{bmatrix}
\left(K_{1:i,1} - K_{1:i,2}\right) & 0 & \cdots & 0 & \cdots & 0 \\
0 & \left(K_{2:i,1} - K_{2:i,2}\right) & \cdots & 0 & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\
0 & \cdots & \cdots & \left(K_{x:i,1} - K_{x:i,2}\right) & \cdots & 0 \\
\vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\
0 & 0 & \cdots & \cdots & \cdots & \left(K_{N:i,1} - K_{N:i,2}\right)
\end{bmatrix}_{N \times N}, \quad \text{a} \quad N \times N
$$

diagonal matrix, and $u_i = \left[K_{1:i,2} + u_{1:i}, \quad K_{2:i,2} + u_{2:i}, \quad \cdots, \quad K_{x:i,2} + u_{x:i}, \quad \cdots, \quad K_{N:i,2} + u_{N:i}\right]^T$, a

$N \times 1$ vector .

From the above analysis, we can see that in many segmentation problems, the bias of the segmentation label is an affine function of the true label.

APPENDIX E

DETERMINATION OF THE OPTIMUM PARAMETERS FOR THE MODIFIED CRAMER-RAO BOUND

We first find the optimum values of K_i and u_i for the modified Cramér–Rao bound (6.27) by setting the derivative of (6.27) with respect to K_i and u_i to zero, respectively. Then the modified Cramér–Rao bound is obtained through submitting the resulting K_i^* and u_i^* into (6.27).

$$\frac{\partial CRB_{Biased}^{Mod}(\hat{H}_i)}{\partial u_i} = \frac{\partial}{\partial u_i} \left\{ \text{Tr}\left((I+K_i)\left(E_\beta[J_F(H_i)]\right)^{-1}(I+K_i)^T\right) + \int (K_iH_i+u_i)^T(K_iH_i+u_i)P(H)dH \right\} = 0$$

$$\Rightarrow \int \left\{ \frac{\partial}{\partial u_i}(K_iH_i+u_i)^T(K_iH_i+u_i) \right\} P(H)dH = 0$$

$$\Rightarrow \int \{2(K_iH_i+u_i)\}P(H)dH = 0$$

$$\Rightarrow u_i = -K_i \int H_i p(H)dH = -K_i E_{H_i}(H_i)$$

(E.1)

Using (E.1), we have

$$\frac{\partial CRB_{Biased}^{Mod}(\hat{H}_i)}{\partial K_i} = \frac{\partial}{\partial K_i} \left\{ \text{Tr}\left((I+K_i)\left(E_\beta[J_F(H_i)]\right)^{-1}(I+K_i)^T\right) + \int (K_iH_i+u_i)^T(K_iH_i+u_i)P(H)dH \right\} = 0$$

$$\Rightarrow 2(I+K_i)\left(E_\beta[J_F(H_i)]\right)^{-1} + \int 2(K_iH_i+u_i)H_i^T P(H)dH = 0$$

$$\Rightarrow K_i\left\{\left(E_\beta[J_F(H_i)]\right)^{-1} + \int H_iH_i^T P(H)dH\right\} = -\left\{\left(E_\beta[J_F(H_i)]\right)^{-1} + \int u_iH_i^T P(H)dH\right\}$$

$$\Rightarrow K_i\left\{\left(E_\beta[J_F(H_i)]\right)^{-1} + \int H_iH_i^T P(H)dH\right\} = -\left(\left(E_\beta[J_F(H_i)]\right)^{-1} - \int_H K_iE_{H_i}(H_i)H^T p(H)dH\right)$$

$$\Rightarrow K_i\left\{\left(E_\beta[J_F(H_i)]\right)^{-1} + E\left(H_iH_i^T\right)\right\} = -\left\{\left(E_\beta[J_F(H_i)]\right)^{-1} - K_iE_{H_i}(H_i)E_{H_i}(H_i)^T\right\}$$

$$\Rightarrow K_i\left\{\left(E_\beta[J_F(H_i)]\right)^{-1} + E\left(H_iH_i^T\right) - E_{H_i}(H_i)E_{H_i}(H_i)^T\right\} = -\left(E_\beta[J_F(H_i)]\right)^{-1}$$

$$\Rightarrow K_i\left\{\left(E_\beta[J_F(H_i)]\right)^{-1} + Cov(H_i)\right\} = -\left(E_\beta[J_F(H_i)]\right)^{-1}$$

$$\Rightarrow K_i^* = -\left(E_\beta[J_F(H_i)]\right)^{-1}\left\{\left(E_\beta[J_F(H_i)]\right)^{-1} + Cov(H_i)\right\}^{-1}$$

(E.2)

So

$$u_i^* = -K_i E_{H_i}(H_i)$$

$$= \left(E_\beta[J_F(H_i)]\right)^{-1}\left\{\left(E_\beta[J_F(H_i)]\right)^{-1} + Cov(H_i)\right\}^{-1} E_{H_i}(H_i)$$

(E.3)

Substituting K_i^* and u_i^* into $CRB_{Biased}^{Mod}(\hat{H}_i)$, we obtain the modified bound for the ith region as follows

$$CRB_{Biased}^{Mod}(\hat{H}_i)^* = \mathrm{Tr}\left((I + K_i^*)(E_\beta[J_F(H_i)])^{-1}(I + K_i^*)^T\right) + \int (K_i^* H_i + u_i^*)^T (K_i^* H_i + u_i^*) P(H)dH$$

$$= \mathrm{Tr}\left((I + K_i^*)(E_\beta[J_F(H_i)])^{-1}(I + K_i^*)^T\right) + \int (K_i^* H_i - K_i E_{H_i}(H_i))^T (K_i^* H_i - K_i E_{H_i}(H_i)) P(H)dH$$

$$= \mathrm{Tr}\left((I + K_i^*)(E_\beta[J_F(H_i)])^{-1}(I + K_i^*)^T\right) + \int (H_i - E_{H_i}(H_i))^T K_i^{*T} K_i^* (H_i - E_{H_i}(H_i)) P(H)dH$$

$$= \mathrm{Tr}\left((I + K_i^*)(E_\beta[J_F(H_i)])^{-1}(I + K_i^*)^T\right) + E_{H_i}\left\{(H_i - E_{H_i}(H_i))^T K_i^{*T} K_i^* (H_i - E_{H_i}(H_i))\right\}$$

$$= \mathrm{Tr}\left((I + K_i^*)(E_\beta[J_F(H_i)])^{-1}(I + K_i^*)^T\right) + E_{H_i}\left\{\mathrm{Tr}\left[K_i^*(H_i - E_{H_i}(H_i))(H_i - E_{H_i}(H_i))K_i^{*T}\right]\right\}$$

$$= \mathrm{Tr}\left((I + K_i^*)(E_\beta[J_F(H_i)])^{-1}(I + K_i^*)^T\right) + \mathrm{Tr}\left\{K_i^* E_{H_i}\left[(H_i - E_{H_i}(H_i))(H_i - E_{H_i}(H_i))\right]K_i^{*T}\right\}$$

$$= \mathrm{Tr}\left\{(I + K_i^*)(E_\beta[J_F(H_i)])^{-1}(I + K_i^*)^T + K_i^* E_{H_i}\left[(H_i - E_{H_i}(H_i))(H_i - E_{H_i}(H_i))\right]K_i^{*T}\right\}$$

$$= \mathrm{Tr}\left\{(I + K_i^*)(E_\beta[J_F(H_i)])^{-1}(I + K_i^*)^T + K_i^* Cov_{H_i}(H_i)K_i^{*T}\right\}$$

$$= \mathrm{Tr}\left\{(E_\beta[J_F(H_i)])^{-1} + 2K_i^*(E_\beta[J_F(H_i)])^{-1} + K_i^*(E_\beta[J_F(H_i)])^{-1}K_i^{*T} + K_i^* Cov_{H_i}(H_i)K_i^{*T}\right\}$$

$$= \mathrm{Tr}\left\{\begin{array}{l}(E_\beta[J_F(H_i)])^{-1} - 2(E_\beta[J_F(H_i)])^{-1}\left\{(E_\beta[J_F(H_i)])^{-1} + Cov(H_i)\right\}^{-1}(E_\beta[J_F(H_i)])^{-1} \\ + (E_\beta[J_F(H_i)])^{-1}\left\{(E_\beta[J_F(H_i)])^{-1} + Cov(H_i)\right\}^{-1}\left\{(E_\beta[J_F(H_i)])^{-1} + Cov_{H_i}(H_i)\right\} \\ \left\{(E_\beta[J_F(H_i)])^{-1}\left\{(E_\beta[J_F(H_i)])^{-1} + Cov(H_i)\right\}^{-1}\right\}^T\end{array}\right\}$$

(E.4)

$$= \mathrm{Tr}\left\{\begin{array}{l}(E_\beta[J_F(H_i)])^{-1} - 2(E_\beta[J_F(H_i)])^{-1}\left\{(E_\beta[J_F(H_i)])^{-1} + Cov(H_i)\right\}^{-1}(E_\beta[J_F(H_i)])^{-1} \\ + (E_\beta[J_F(H_i)])^{-1}\left\{(E_\beta[J_F(H_i)])^{-1} + Cov(H_i)\right\}^{-1}(E_\beta[J_F(H_i)])^{-1}\end{array}\right\}$$

$$= \mathrm{Tr}\left\{(E_\beta[J_F(H_i)])^{-1} - (E_\beta[J_F(H_i)])^{-1}\left\{(E_\beta[J_F(H_i)])^{-1} + Cov(H_i)\right\}^{-1}(E_\beta[J_F(H_i)])^{-1}\right\}$$

REFERENCES

[1]   T. Acharya and A. K. Ray, *Image Processing-Principles and Applications*, Wiley InterScience, 2005.

[2]   A. K. Jain, *Fundamentals of Digital Image Processing,* Prentice Hall, Upper Saddle River, NJ, 1988.

[3]   J. Proakis, *Digital Communications (4 edition)*, McGraw-Hill Science/Engineering/Math, Aug. 15, 2000.

[4]   B. Niu, O. Simeone, O. Somekh and A. M. Haimovich, "Ergodic and outage performance of fading broadcast channels with 1-bits feedback," *IEEE Trans. Vehicular Technology,* vol. 59, no. 3, pp. 1282-1293, Mar. 2010.

[5]   B. Niu and A. M. Haimovich, "Interference Subspace Tracking for Network Interference Alignment in Cellular Systems" in *Proc. IEEE Global Communications Conference (Globecom)*, pp. 1-5, Honolulu, Nov. 2009.

[6]   P. Symes, *Digital Video Compression*, McGraw-Hill/TAB Electronics, Oct. 2, 2003.

[7]   Y. Amit, *2D Object Detection and Recognition: Models, Algorithms and Networks*, The MIT Press, Nov. 1, 2002.

[8]   A. Mohan, C. Papageorgiou and T. Poggio, "Example-based object detection in images by components*," IEEE Trans. Pattern Analysis and Machine Intelligence,* vol. 23, no. 4, pp. 349-361, 2001.

[9]   A. Yuille, "Deformable Templates for face recognition," *Journal of. Cognitive Neuroscience,* vol. 3, no. 1, pp. 59-70, 1991.

[10]   P. Sinha, "Object recognition via image invariants: a case study," *Investigative Ophthalmology and Visual Science,* vol. 35, pp. 1735-1740, 1994.

[11]   C. Papageorgiou and T. Poggio, "A Trainable System for Object Detection," *International Journal of. Computer Vision,* vol. 38, no. 1, pp. 15-33, 2000.

[12] N.V. Balaji and M. Punithavalli, "Machine learning approach for object detection-a survey approach," *International Journal of Computer Science and Information Security,* vol. 8, no. 7, 2010.

[13] S. M. Kay, *Fundamentals of Statistical Signal Processing-Detection Theory*, Prentice Hall PTR, 1998.

[14] A. C. Bovik, *Handbook of Image and Video Processing, (second ed.),* Academic Press, Burlington, Jun., 2005.

[15] Y. T. Kim, et al., "Contrast enhancement using brightness preserving bi-histogram equalization," *IEEE Trans. Consumer Electronics,* vol. 43, pp. 1-8, 1997.

[16] K. Zuiderveld, "Contrast limited adaptive histogram equalization," in: P. S. Hechbert, (Ed.), *Graphics Gems IV,* AP Professional, Boston, pp. 474-485,1994.

[17] G. Ramponi, N. Strobel, S. K. Mitra and T.-H. Yu, "Nonlinear unsharp masking methods for image contrast enhancement," *Journal of Electronic Imaging,* vol. 5, pp. 353-366, 1996.

[18] J. S Lim, *Two-Dimensional Signal and Image Processing,* Englewood Cliffs, Prentice Hall, NJ, 1990.

[19] J. Dengler, S. Behrens and J.F. Desaga, "Segmentation of micro-calcifications in mammograms," *IEEE Trans. Medical Imaging,* vol. 12, pp. 634-642, 1993.

[20] S. Singh and K. Bovis, "An evaluation of contrast enhancement techniques for mammographic breast masses," *IEEE Trans. Information Technology in Biomedicine,* vol. 9, pp. 109-119, 2005.

[21] J. Dušek and K. Roubík, "Testing of new models of the human visual system for image quality evaluation," in *Proc. IEEE 7th International Symposium on Signal Processing and its Application,* vol. 2, pp. 621–622, 2003.

[22] K. A. Panetta, E. J. Wharton and S. S. Agaian, "Human visual system-based image enhancement and logarithmic contrast measure*," IEEE Trans. Systems, Man, and Cybernetics—Part B: Cybernetics,* vol. 38, pp. 174-88, 2008.

[23]   R. A. Nobakht and S. A. Rajala, "An image coding technique using a human visual system model and image analysis criteria," in *Proc. IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP),* vol. 12, pp. 1358–1361, 1987.

[24]   Y. Zhang (Editor), *Advances in Image And Video Segmentation*, IRM Press, USA May 2, 2006.

[25]   P. K. Sahoo, S. Soltani, A. K. C.Wong, and Y. C. Chen, "A survey of thresholding techniques," *Computer Vision, Graphics, and Image Processing archive*, vol. 41, no. 2, Feb., 1988.

[26]   S. Z. Li*, Markov Random Field Modeling in Computer Vision*. New York: Springer-Verlag, 2001.

[27]   V. Kolmogorov and R. Zabih, "What energy functions can be minimized via graph cuts?" *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 26, no. 2, pp. 147–159, Feb. 2004.

[28]   M. Unser, "Texture classification and segmentation using wavelet frames," *IEEE Trans. Image Process.*, vol. 11, no. 4, pp. 1549–1560, Apr. 1995.

[29]   D. Mumford and J. Shah, "Optimal approximation by piecewise smooth functionals and associated variational problems," *Commun. Pure Appl. Math.*, vol. 42, pp. 577–685, 1989.

[30]   D. Pham, C Xu, and J. Prince, "A survey of current methods in medical image segmentation," *Annual Review of Biomedical Engineering*, vol. 2, pp. 315-337, 2000.

[31]   N. R. Pal and S. K. Pal, "A review on image segmentation techniques," *Pattern Recognition*, vol. 26, no. 9, pp. 1277-1294, 1993.

[32]   D. Comaniciu and P. Meer, "Mean shift: A robust approach toward feature space analysis," *IEEE Trans. Pattern Analysis and Machine Intelligence,* vol. 24, no. 5, pp. 603–619, May 2002.

[33]   Y. Ma, H. Derksen, W. Hong and J. Wright, "Segmentation of multivariate mixed data via lossy coding and compression," *IEEE Trans. on Pattern Analysis and Machine Intelligence,* vol. 29, no. 9, pp. 1546-1562, 2007.

[34]   J. Canny, "A computational approach to edge detection", *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 8, no. 6, pp. 679–698, 1986.

[35]   L. Shapiro and G. Stockman, *Computer Vision*, Prentice-Hall, Inc. 2001.

[36]   D. Geman, "Stochastic model for boundary detection," *Image and Vision Computing*, vol. pp. 61-65, 1987.

[37]   D. Geman, S. Geman, Chr. Graffigne, and P. Dong, "Boundary detection by constrained optimization," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 12, no. 7, pp. 609-628, 1990.

[38]   D. Geman and S. Geman, "Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 6, pp. 721-741, 1984.

[39]   M. Kass, A. Witkin and D. Terzopoulos, "Snakes: Active contour models," *International Journal of Computer Vision*, vol. 1, no. 4, pp. 321–331, 1988.

[40]   D. Cremers, M. Rousson and R. Deriche, "A review of statistical approaches to level set segmentation: Integrating color, texture, motion and shape," *International Journal of Computer Vision*, vol. 72, no. 2, pp. 195–215, 2007.

[41]   J. Freixenet, X. Munoz, D. Raba, J. Marti and X. Cuffi, "Yet another survey on image segmentation: region and boundary information integration," in *Proc. the 7th European Conference on Computer Vision-Part III*, pp. 408 – 422, 2002.

[42]   X. M. Pujol, "Image segmentation integrating color, texture and boundary information," *Ph.D. Thesis*, University of Girona, Spain, 2003.

[43]   J. Fan, D.K.Y. Yau, A.K. Elmagarmid and W.G. Aref, "Automatic image segmentation by integrating color-edge extraction and seeded region growing," *IEEE Trans. Image Processing*, vol. 10, no. 10, pp. 1454 – 1466. 2001.

[44]   J. Wu, and A. C. S. Chung, "A segmentation model using compound Markov random fields based on a boundary model," *IEEE Trans. Image Processing*, vol. 16, no. 1, Jan. 2007.

[45]   M. S. Alili and D. Ziou, "An approach for dynamic combination of region and boundary information in segmentation," in *Proc. 19th International Conference on Pattern Recognition*, 2008.

[46]   T. Lei and J. K. Udupa, *"Performance evaluation of finite normal mixture model-based image segmentation techniques,"* *IEEE Trans. on Image Processing*, vol. 12, no. 10, pp. 1153-1169, Oct. 2003.

[47]   I. B. Kerfoot and Y. Bresler, "Theoretical analysis of multi-spectral image segmentation criteria," *IEEE Trans. on Image Processing*, vol. 8, no. 6, pp. 798-820, Jun. 1999.

[48]   S. K. Warfield, K. H. Zou and W. M. Wells, "Simultaneous truth and performance level estimation (STAPLE): an algorithm for the validation of image segmentation," *IEEE Trans. on Medical Imaging*, vol. 23, no. 7, pp. 903-921, Jul. 2004.

[49]   R. Benzi, A. Sutera and A. Vulpiani, "The mechanism of stochastic resonance," *Journal of Physics A: Mathematical and General*, vol. 14, pp. L453 – L457, 1981.

[50]   L. Gammaitoni, P. H¨a nggi, P. Jung and F. Marchesoni, "Stochastic resonance," *Rev. Mod. Phys.*, vol. 70, no. 1, pp. 223–287, Jan. 1998.

[51]   A. Asdi and A. Tewfik, "Detection of weak signals using adaptive stochastic resonance," in *International Conference on Acoustics, Speech, and Signal Processing, (ICASSP' 95)*, vol. 2, May 1995, pp. 1332–1335.

[52]   S. Zozor and P.-O. Amblard, "On the use of stochastic resonance in sine detection," *Signal Processing*, vol. 7, pp. 353–367, March 2002.

[53]   H. Chen, P. K.Varshney, J. H. Michels and S. M. Kay, "Theory of the stochastic resonance effect in signal detection: Part 1 - fixed detectors," *IEEE Transactions on Signal Processing*, vol. 55, no. 7, pp. 3172–3184, July 2007.

[54]   H. Chen and P. K.Varshney, "Theory of the stochastic resonance effect in signal detection: Part 2 - variable detectors," *IEEE Transactions on Signal Processing*, vol. 56, no. 10, pp. 5031–5041, 2008.

[55]   B. Kim and D. Park, *"Novel non-contrast-based edge descriptor for image segmentation,"* *IEEE Trans. Circuits and Systems for Video Technology*, vol. 16, no. 9, Sep. 2006.

[56]   J. Chen*, T. N. Pappas*, A. Mojsilovic´*, and B. E. Rogowitz*, "Adaptive perceptual color-texture image segmentation," *IEEE Trans. Image Processing*, vol. 14, no. 10, Oct. 2005.

[57]   J. Zhang and D. Wang, "Image segmentation by multi-grid Markov random field optimization and perceptual considerations," *Journal of Electronic Imaging*, vol. 7, 52, 1998.

[58]   B. Wandell, *Foundations of Vision*, Sinauer Associates, May, 1995.

[59]   S. Winkler, "Issues in vision modeling for perceptual video quality assessment," *Signal Processing,* vol. 78, no. 2, pp. 231–252, 1999.

[60]   M. J. Nadenau, S. Winkler, D. Alleysson and M. Kunt, "Human vision models for perceptually optimized image processing – a review," *Proceedings of the IEEE*, 2000.

[61]   R. W. Floyd, and L. Steinberg, "An Adaptive Algorithm for Spatial Gray Scale," *International Symposium Digest of Technical Papers*, Society for Information Displays, p. 36, 1975.

[62]   J. R. Parker, *Algorithms for Image Processing and Computer Vision*, New York, John Wiley & Sons, Inc., 1997.

[63]   E. J. Feurer and L. M. Wun, *Devcan: probability of developing or dying of cancer software (Version 4.0)*, National Cancer Institute, Bethesda, MD, 1999.

[64]   R. A. Smith, "Epidemiology of breast cancer categorical course in physics," *Tech. Aspects Breast Imaging, Radiol. Soc. N. Amer.*, pp. 21–33, 1993.

[65]   H. P. Chan *et al.*, "Improvement in radiologist's detection of clustered micro-calcifications on mammograms," *Investigative Radiol.*, vol. 25, pp. 1102–110, 1990.

[66]   S. Astley *et al*, "Automation in mammography: Computer vision and human perception," *International Journal of Pattern Recognition Artificial Intelligence*, vol. 7, no. 6, pp. 1313–1338, 1993.

[67]   A. Bazzani *et al*, "Automatic detection of clustered micro-calcifications in digital mammograms using an SVM classifier," in *Proc. of European Symposium on Artificial Neural Networks Plastics*, Bruges, 26-28, April, 2000..

[68]   M. Nafi Gurcan, Y. Yardimci and A. E. Getin, "Influence function based gaussianity tests for detection of micro-calcifications in mammogram images," in *Proc. International Conference on Image Processing* ,vol.3, pp. 407-411, 1999.

[69]   N. Karssemeijer, "Adaptive noise equalization and recognition of micro-calcification clusters in mammograms," *International Journal of Pattern Recognition Artificial Intelligence,* vol. 7, no. 6, pp. 1357–1376, 1993.

[70]   R. Nakayaman *et al*, "Computer-aided diagnosis scheme using a filter bank for detection of micro-calcification clusters in mammograms," *IEEE Trans. Biomedical Engineering*, vol. 53, no. 2, pp. 273-283, Feb. 2006.

[71]   E. Regentova *et al*, "Detecting micro-calcifications in digital mammograms using wavelet domain hidden markov tree model," in *Proc. 28th Annual International Conference of the IEEE Engineering in Medicine and Biology Society 2006 (EMBS '06)*, pp. 1972-1975, 30,Aug. –3, Sep., 2006.

[72]   S. Deepa and T. Tessamma, "Fractal modeling of mammograms based on mean and variance for the detection of micro-calcifications," in *Proc. International Conference on Computational Intelligence and Multimedia Applications*, vol. 2, pp. 334-348, 13-15, Dec., 2007.

[73]   L.Wei *et al*, "Relevance vector machine for automatic detection of clustered micro-calcifications," *IEEE Trans. Medical Imaging,* vol. 24, no. 10, pp. 1278-1285, Oct. 2005.

[74]   E. Catanzariti *et al*, "A CAD system for the detection of mammographyc micro-calcifications based on Gabor Transform," in *Proc. Nuclear Science Symposium Conference Record*, vol. 6, pp. 3599- 3603, 16-22 Oct. 2004

[75]   R. N. Strickland, "Wavelet transform methods for objects detection and recovery," *IEEE Trans. Image Processing*, vol. 6, pp. 724–735, May, 1997.

[76]   G. Lemaur, K. Drouiche and J. DeConinck, "Highly regular wavelets for the detection of clustered micro-calcifications in mammograms," *IEEE Trans. Medical Imaging*, vol. 22, no. 3, Mar., 2003.

[77] K. Li and Z. Dong, "A novel method of detecting calcifications from mammogram images based on wavelet and sobel detector," in *Proc. 2006 IEEE International Conference on Mechatronics and Automation*, pp.1503-1508, Jun. 2006.

[78] D. W. J. Stein *et al*, "Anomaly detection from hyperspectral imagery," *IEEE Signal Processing Magazine*, vol.19, pp. 58-69, Jan., 2002.

[79] M. N. Giircan, Y. Yardimci and A. E. cetin, "2-D adaptive prediction based Gaussianity tests in micro-calcification detection," *Proceedings of SPIE, Visual Communications and Image Processing*, vol. 3309, 1997. *, and Signal Processing (ICASSP'02)*, vol.4, pp. IV-3856- IV-3859, 2002.

[80] M. Sezgin and B. Sankur, "Survey over image thresholding techniques and quantitative performance evaluation," *Journal of Electronic Imaging*, 13(1), pp. 146–165, Jan. 2004.

[81] P. A. Devijver, and J. Kittler, *Pattern Recognition: A Statistical Approach*, Prentice Hall, Englewood Cliffs, London, 1982.

[82] G. Kolata, "Study questions computerized breast cancer detection," *The New York Times National*, Thur., Apr. 5, 2007.

[83] A. Banerje and R. Chellappa, "Tumor detection in digital mammograms," in *Proc. International Conference on Image Processing (ICIP'00)*, vol.3, pp. 432-435, Vancouver, BC, Canada, 2000.

[84] C. L. Nikias and M. Shao, *Signal Processing with Alpha-Stable Distributions and Applications,* John Wiley and Sons,Inc.,1995.

[85] T. H. Cormen, C.E. Leiserson, R. L. Rivest and C.Stein, *Introduction to Algorithms (3rd edition)*, The MIT Press, 2009.

[86] K. Fukunaga, *Introduction to Statistical Pattern Recognition (second edition)*, Academic Press, Sep., 1990.

[87] B. Efron and R. J. Tibshirani, *An Introduction to the Bootstrap*, Chapman & Hall, 1993.

[88]   J. I. D. Rosa and G. A. Fleury, "Bootstrap Methods for a Measurement Estimation Problem," *IEEE Trans. Instrumentation and Measurement*, vol. 55, no. 3, pp. 820-827, Jun., 2006.

[89]   J. Kittler and J. Lllingworth, "Minimum error thresholding," *Pattern Recognition*, vol.19, no. 1, pp. 41-47, 1986.

[90]   G. Mclachlan and D. Peel, *Finite Mixture Models*, A Wiley-Interscience Publication, 2, Oct., 2000.

[91]   [online]. Available: http://marathon.csee.usf.edu/Mammography/Database.html.

[92]   [online]. Available: http://peipa.essex.ac.uk/info/mias.html.

[93]   M. Lanyi, *Diagnosis and Differential Diagnosis of Breast Calcifications*. Berlin, Germany: Springer-Verlag, 1988.

[94]   C. García-Orellana, R Gallardo-Caballero, M Macías-Macias and H González-Velasco, "SVM and Neural Networks comparison in mammographic CAD", in *Proc. of the 29th Annual International Conference of the IEEE EMBS*, Lyon, France, August 23-26, 2007.

[95]   C. J. D'Orsi and D. B. Kopans, "Mammographic feature analysis," *Semin. Roentgenol. XXVIII*, pp. 204-230, 1993.

[96]   A. M. Bagct, Y. Yardimci and A. E. Cetin, "Detection of micro-calcification clusters in mammogram images using local maxima and adaptive wavelet transform analysis," in *Proc. Of International Conference on Acoustics, Speech, and Signal Processing (ICASSP'02)*, vol.4, pp. IV-3856- IV-3859, 2002.

[97]   W. M. Diyana, J. Larcher and R. Besar, "A comparison of clustered micro-calcifications automated detection methods in digital mammogram," in *Proc. International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, vol. 2, Page(s): II - 385-8 6-10, Apr., 2003.

[98]   M. Sampat, M. Markey and A. Bovik, "Computer-aided detection and diagnosis in mammography", *Handbook of Image and Video Processing*, Academic Press, 2005.

[99]   H. Cheng, Y. M. Lui and R. I. Freimanis, "A novel approach to micro-calcification detection using fuzzy logic technique," *IEEE Trans. on Medical Imaging*, vol. 17, no. 3, Jun. 1998.

[100] J. Suckling *et al*, "The mammographic image analysis society mammogram database," in *Proc. 2nd International Workshop on Digital Mammography*, York, U.K., pp. 375–378, 1994.

[101] D. Wang, A. H. Vagnucci, "Digital image enhancement," *Computer Vision, Graphics, and Image Processing,* vol. 24, pp. 363–381, 1981.

[102] R. J. Marks *et al,* "Stochastic resonance of a threshold detector: image visualization and explanation," in Proc. *IEEE International Symposium on Circuits and Systems (ISCAS),* vol. 4, pp. IV521-523, pp. 2002.

[103] Q. Ye, H. Huang and C. Zhang, "Image enhancement using stochastic resonance," in *Proc. International Conference in Image Processing (ICIP),* vol. 1, pp. I 263-266, 2004.

[104] H. Chen, P. K. Varshney and J. H. Michels, "De-noising image with noise," in *Proc. 41st Asilomar Conference on Signals, Systems and Computers,* pp. 513-517, 2007.

[105] M. Rivera and J. L. Marroquin, "Adaptive rest condition potentials: first and second order edge-preserving regularization," *Computer Vision and Image Understanding,* vol. 88. pp. 76-93, 2002.

[106] V. N. Vapnik, *Statistical Learning Theory,* Wiley-Interscience, 1998.

[107] K. Deb, A. Pratap, S. Agarwal and T. Meyarivan, "A fast elitist multi-objective genetic algorithm: NSGA-II," *IEEE Transaction on Evolutionary Computation,* vol. 6, no. 2, 182-197, 2002.

[108] Z. Wang, A. C. Bovik, H. R. Sheikh and E. P. Simoncelli, "Image quality assessment: from error visibility to structural similarity," *IEEE Trans. Image Processing,* vol. 13, pp. 600-612, 2004.

[109] J. Mannos and D. Sakrison, "The effects of a visual fidelity criterion the encoding of images," *IEEE Trans. Information Theory,* vol. 20, pp. 525–536, 1974.

[110] J. N. Ellinas, D. E. Manolakis, "A robust watermarking scheme based on edge detection and contrast sensitivity function," *VISAPP Proc. International Conference on Computer Vision Theory and Applications,* Barcelona, 2007.

[111] G. Kell *et al,* "Enhancement and associative restoration of electronic portal images in radiotherapy," *International Journal of Medical Informatics,* vol. 49, pp. 157-171, 1998.

[112] S. Angenent, E. Pichon and A. Tannenbaum, *Mathematical Methods in Medical Image Processing,* Bulletin of the American Mathematical Society, 2006.

[113] J. Caviedes, and S. Gurbuz, "No-reference sharpness metric based on local edge kurtosis," in *Proc. International Conference in Image Processing (ICIP),* vol. 3, pp. III-53-56, 2002.

[114] M. A. Wirth and D. Nikitenko, "Quality evaluation of fuzzy contrast enhancement algorithms," in *Proc. Annual Meeting of the North American Fuzzy Information Processing Society (NAFIPS),* 2005.

[115] H. R. Tizhoosh, G. Krell and B. Michaelis, "On fuzzy enhancement of megavoltage images in radiation therapy," *presented at IEEE International Conference on Fuzzy Systems,* Barcelona, Spain, 1997.

[116] Z. Wang and A. C. Bovik, "Mean squared error: love it or leave it?" I*EEE Signal Processing Magazine,* vol. 26, pp. 98-117, 2009.

[117] M. Penedo *et al,* "Region-based wavelet coding methods for digital mammography," *IEEE Trans. Medical Imaging,* vol. 22, no. 10, 2003.

[118] C. Chang, K Chen, J Wang and M. L.G. Althouse, "A Relative entropy-based approach to image thresholding," Pattern Recognition, vol. 27, no. 9, pp. 1275-1289, 1994.

[119] N.R.Pal and S.k.Pal, "Entropy thrsholding," Signal Processing, vol. 16, 97-108, 1989.

[120] M. L. Comer, S. Liu and E. J. Delp, "Statistical segmentation of mammograms," in *Proc. the 3rd International Workshop on Digital Mammography,* Jun. 9-12, Chicago, pp. 475-478, 1996.

[121] C. Li, C. Xu, C. Gui and M. D. Fox, "Level set evolution without re-initialization: a new variational formulation," in *Proc. of the 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)*, 2005.

[122] L. A. Zadeh, *Fuzzy Logic and Its Applications,* New York: Academic Press, 1965.

[123] S. Singh and R. Al-Mansoori, "Identification of regions of interest in digital mammograms," Journal of Intelligent Systems, 2000.

[124]   Z. Tu and S.-C. Zhu, "Image segmentation by data-driven Markov chain Monte Carlo," *IEEE Trans. Pattern Analysis and Machine Intelligence,* vol. 24, no. 5, pp. 657–673, May. 2002.

[125]   J. Shi and J. Malik, "Normalized cuts and image segmentation," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 22, no. 8, pp. 888–905, Aug. 2000.

[126]   S. X. Yu and J. Shi, "Multiclass spectral clustering," in *Proc. IEEE Int. Conf. Computer Vision*, 2003, vol. 1, pp. 313–319.

[127]   G. Winkler, *Image Analysis, Random Fields and Markov Chain Monte Carlo Methods: A Mathematical Introduction (2nd edition)*, Springer, Feb., 27, 2006.

[128]   H. Shen, L. Zhang, B. Huang, and P. Li, "A MAP approach for joint motion estimation, segmentation, and super resolution," *IEEE Trans. on Image Processing*, vol. 16, no. 2, pp. 479-490, 2007.

[129]   W. Lin, "Computational models for just-noticeable difference," in *H. R. Wu and K. R. Rao (Editors), Digital Video Image Quality and Perceptual Coding*, ,Taylor & Francis, 2005.

[130]   Y. Jia, W.Lin, and A. A. Kassim, "Estimating just-noticeable distortion for video," *IEEE Trans. Circuits and System for Video Technology*, vol. 16, no. 7, pp. 820-829, Jul. 2006.

[131]   X. Zhang, W. Lin and P. Xue, "Just-noticeable difference estimation with pixels in images," *Journal of Visual Communication and Image Representation (ELSEVIER)*, vol. 19, pp. 30-41, 2008.

[132]   W. Lin*,* L. Dong and P. Xue*,* "Visual distortion gauge based on discrimination of noticeable contrast changes," *IEEE Trans. Circuits and Systems for Video Technology*, vol., 15, no. 7, pp. 900-909, Jul. 2005.

[133]   Z. Liu, L. J. Karam and A. B. Watson, "JPEG2000 encoding with perceptual distortion control," *IEEE Trans. Image Processing*, vol. 15, no. 7, pp. 1763-1778, Jul. 2006.

[134]   X. Yang, W. Lin, Z. Lu, E. P. Ong and S. Yao, "Just-noticeable-distortion profile with nonlinear additivity model for perceptual masking in color images," in *Proc. IEEE International Conference on* Acoustics, Speech and Signal Processing (ICASSP), vol. 3, pp. 609-612, 2003.

[135] Y. Boykov, O. Veksler and R. Zabih, "Fast approximate energy minimization via graph cuts," IEEE Trans. Pattern *Analysis and Machine Intelligence,* vol. 23, no. 11, pp. 1222–1239, Nov. 2001.

[136] N. Otsu, "A threshold selection method from gray-level histograms," *IEEE Trans. Sys., Man., Cyber, v*ol. 9, pp.62–66, 1979.

[137] P. Liao, T. Chen and P. Chung, "A fast algorithm for multilevel thresholding," *J. Inf. Sci. Eng.*, vol. no. 5, pp.713–727, 2001.

[138] C. Li, C. Kao, J. C. Gore and Z. Ding, "Minimization of region-scalable fitting energy for image segmentation," *IEEE Trans. Image Processing*, vol. 17, no.10, pp.1940-1949, 2008.

[139] T. Cour, F. Benezit and J. Shi, "Spectral segmentation with multi-scale graph decomposition," in *Proc. IEEE International Conference on Computer Vision and Pattern Recognition (CVPR),* 2005.

[140] J. Besag, "On the statistical analysis of dirty pictures" (with discussions), *Journal of the Royal Statistical Society, Series B*, vol. 48, pp. 259-302, 1986.

[141] S. M. Ali and S. D. Silvey, "A general class of coefficients of divergence of one distribution from another," *J. Royal Statistics Soc.*, vol. 28, pp. 132-142, 1966.

[142] M. J. Nadenau, J. Reichel, and M. Kunt, "Wavelet-based color image compression: exploiting the contrast sensitivity function," *IEEE Trans. on Image Processing,* vol. 12, no. 1, pp. 58-70, Jan. 2003.

[143] C. S. Won and R. M. Gray, *Stochastic Image Processing*, Kluwer Academic/Plenum Publishers, NY, 2004.

[144] H. D. Li, M. Kallergi, L. P. Clarke, V. K. Jain and R. A. Clark, "Markov random field for tumor detection in digital mammography," *IEEE Trans. on Medical Imaging*, vol. 14, no. 3, 1995.

[145] T. Hastie and R. Tibshirani, "Varying-coefficient models," *Journal of the Royal Statistical Society*, series B, vol. 55, no. 4, pp. 757-796, 1993.

[146]   F. Salzenstein and W. Pieczynski, "Parameter estimation in hidden fuzzy Markov random fields and image segmentation," *Graphical Models and Image Processing*, vol. 59, no. 4, pp. 205–220, Jul., 1997.

[147]   B. M. Sadler, R. J. Kozick and T. Moore, "Bounds on bearing and symbol estimation with side information," *IEEE Trans. on Signal Processing*, vol. 49, no. 4, Apr. 2001.

[148]   M. Rivera, O. Ocegueda and J. L. Marroquin, "Entropy-controlled quadratic Markov measure field models for efficient image segmentation," *IEEE Trans. on Image Processing*, vol. 16, no. 12, pp. 3047-3057, Dec. 2007.

[149]   J. L. Marroquin, E. A. Santana and S. Botello, "Hidden Markov measure feld models for Image segmentation," *IEEE Trans. on Patter Analysis and Machine Intelligence*, vol. 25, no. 11, pp. 1380-1387, Nov. 2003.

[150]   H. L. V. Trees and K. L. Bell (ed.), *Bayesian Bounds for Parameter Estimation and Nonlinear Filtering/Tracking*, Wiley-IEEE Press, 2007

[151]   B. Z. Bobrovsky, E. Mayer-Wolf and M. Zakai, "Some classes of global Cramér–Rao bounds," *Ann. Stat.*, vol. 15, pp. 1421-1438, 1987.

[152]   P. Stoica and B.C.Ng, "On the Cramèr-Rao bound under parametric constraints," *IEEE Signal Processing Lett.*, vol. 5,pp. 177-179, Jul., 1998.

[153]   D. R. Farenick and F. Zhou, "Jensen's inequality relative to matrix-valued measures," *Journal of Mathematical Analysis and Applications*, vol. 327, no. 2, pp. 919-929, 2007.

[154]   S. Kay and Y.C. Eldar, "Rethinking biased estimation," *Signal Processing Magazine*, vol. 25, no. 3, pp. 133-136, May, 2008.

[155]   Y. C. Eldar, "Minimum variance in biased estimation: bounds and asymptotically optimal estimators," *IEEE Trans. Signal* Processing, vol. 52, pp. 1915-1930, Jul., 2004.

[156]   Y. C. Eldar, "MSE bounds with affine bias dominating the Cramèr-Rao bound," *IEEE Trans. on Signal Processing*, vol. 56, no. 8, Aug., 2008.

[157] P. Chatterjee and P. Milanfar, "Is denoising dead?" *IEEE Transactions on Image Processing*, vol. 19, num. 4, pp 895-911, Apr., 2010.

[158] A. C. Davison and D.Hinkley, *Bootstrap Methods and Their Application* (8th ed.), Cambridge: Cambridge Series in Statistical and Probabilistic Mathematics, 2006.

[159] H. J. Adèr, G. J., Mellenbergh and D. J. Hand, *Advising on research methods: A consultant's companion*, Huizen, The Netherlands: Johannes van Kessel Publishing, 2008.

[160] J. C. Bezdek, *Pattern Recognition with Fuzzy Objective Function Algorithms*, Plenum Press, New York, 1981.

[161] D. Dubois, and H. Prade, *Fundamentals of Fuzzy Sets*, Boston, Kluwer, 2000.

[162] I. Gath and A.B. Geva, "Unsupervised optimal fuzzy clustering," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 7, pp. 773-781, 1989.

[163] A. Wiesel, Y. C. Eldar and A. Yeredor, "Linear regression with Gaussian model uncertainty: algorithms and bounds," *IEEE Trans. on Signal Processing*., vol. 56, no. 6, pp. 2194-2205, Jun., 2008.

[164] Y. C. Eldar, "Minimax MSE estimation of deterministic parameters with noise covariance uncertainties," *IEEE Trans. Signal Processing*, vol. 54, no. 1, pp. 138-145, Jan., 2006.

[165] Y. C. Eldar, "Minimax estimation of deterministic parameters in linear models with a random model matrix," *IEEE Trans. Signal Processing*, vol. 54, no. 2, pp. 601-612, Feb., 2006.

BIOGRAPHY

Renbin Peng received the B.S. and M.S. degrees in electrical engineering from Xidian University, Xi'an, Shaanxi, China, in 2001 and 2004, respectively. He is currently pursuing the Ph.D. degree in the Department of Electrical Engineering and Computer Science at Syracuse University, Syracuse, NY. He joined Adobe Systems Inc. as a Computer Scientist in May 2011.

His research interests are in the areas of image, video and signal processing, computer vision, as well as their applications in natural and medical image and video, and mobile and wireless communications.