Thèse de Doctorat

pour obtenir le grade de docteur délivré par:

l'Université de Lille 1 – Sciences et Technologies

en cotutelle avec

l'Université de Yaoundé I

Spécialité doctorale "Automatique, Génie Informatique, Traitement du Signal et des Images"

présentée et soutenue publiquement par

Cyrille D. FEUDJIO KOUGOUM

le 5 Octobre 2016

Segmentation of mammographic images for computer aided diagnosis

Directeur de thèse : **Olivier Colot, Alain Tiedeu**
Co-encadrant de thèse : **John Klein**

devant le **jury** composé de :

Olivier Colot,	Pr. (Univ. Lille)	Examinateur
Christine Fernandez Maloigne,	Pr. (Univ. Poitiers)	Examinateur
Ludovic Macaire,	Pr. (Univ. Lille)	Examinateur
John Klein,	Mcf. (Univ. Lille)	Examinateur
Michèle Rombaut,	Pr. (Univ. Grenoble Alpes)	Examinateur
Su Ruan,	Pr. (Univ. Rouen)	Rapporteur
Alain Tiedeu,	Pr. (Univ. Yaoundé I)	Examinateur
Didier Wolf,	Pr. (Univ. Lorraine)	Rapporteur

Ecole Doctorale régionale Sciences pour l'Ingénieur Lille Nord-de-France
Université de Lille
Centre de Recherche en Informatique Signal et Automatique de Lille (CRIStAL)
UMR CNRS 9189, France

to Sophie & André KOUGOUM

to Ismaelle
Jane, Irina, Liz & Jason.

Acknowledgments

The research work presented in this document was carried out partially at the Centre de Recherche en Informatique, Signal et Automatique de Lille (CRIStAL) and partially at the Laboratoire d'Electronique et de Traitement du Signal (LETS) under a co-tutorial agreement. This convention is the fruit of the cooperation between the Universities of Yaoundé I and Lille1 and put in place through the exchange students program under ERASMUS ACP framework who partially sponsored this work with the grant no. MA10D1862UY. During these past years of intense research activities, I have interacted with so many people that I would like to seize this opportunity to express my profound gratitude to them.

I would first of all like to sincerely acknowledge the members of jury of this thesis for the interest they showed in examining this research work. A special thanks to Prof. Su RUAN and Prof. Didier WOLF who have kindly accepted to report this thesis in a reduced timescale despite the pressure of their own activities. I also want to express my gratitude to Prof. Ludovic MACAIRE, it is an honour for me to have him as chairperson of this jury. I would also acknowledge Prof. Michèle ROMBAUT as well to Prof. Christine-Fernandez MALOIGNE who generously took some of their time and energy at reviewing this work. I am particularly impressed and grateful as well to all your critics and remarks that contributed to improve the quality of this manuscript and thus enlighten the work achieved in this research.

I would like to express my greatest appreciation to my thesis supervisor Prof. Alain TIEDEU from University of Yaoundé I who trustfully assigned me this research topic thereby giving the opportunity to start this adventure. His constant support and encouragement enabled me to face with determination the obstacles that were stewed all along the achievement of this thesis. I am truly grateful to Prof. Olivier COLOT my thesis supervisor from Université Lille1 and also director of CRIStAL lab. The impact of this thesis had would have been impossible without his brave and brilliant advice. Deep thanks to Maître de Conférence John KLEIN who devoted a special attention in the follow up of my research activities. His consistent support and experience sharing providing insightful contributions and technical proficiency throughout this work have significantly improved my research and intellectual capacities. Without his generous presence and help, I would not have been able to write up this dissertation successfully. It was a honour for me to have worked very closed to him during this period. I also wish to acknowledge Prof. Samuel DOMNGANG and Prof. Martin KOM for providing their support and very wise counsels in many difficult situations.

I would also like to express my gratitude to the researchers I meet in Lille especially those with who we evolved in the same office during my stays there. Your counsels and advice have been very helpful for me to solve some difficulties I faced and your presence

have always contributed to create a lively environment for my research work. I will not forget to sincerely appreciate the friendship of other PhD and Master students of the lab for the discussions, comments and lots of laughs we had. I would like to acknowledge the International Office of Lille1 for devotion in addressing foreign students issues related to administrative matters and especially to its former manager Béatrice DELPOUVE for her devotion and her kind attention paid on foreign students' difficulties. Without her active implication, my co-tutorial agreement would have not existed. I also thanks Nicolas POTTIEZ for his help in addressing administrative matters linked to students exchange.

I would like to acknowledge at colleagues and all relatives at College of Technology in University of Buea for their continual support and motivations more precisely for their friendship habit and team spirit. I am very grateful to them.

I would like to extend my deepest love and thanks to my sweetheart *Ismaelle* for her endless encouragement, her cooperation and patience with me and kids during my long periods of absence at home. Thanks to my kids whose smiles always help me see the difference between the reality and the dream.

This work could not have been achieved without the unconditional love, encouragement and support of my parents, sisters, brothers and others family members. They were all a source of extra motivation and strength in moving forward in the completion of this research work. I truly appreciate all what you have done or achieved in my life.

Contents

Contents

List of figures

Abstract

Computer-aided diagnosis (CAD) systems are currently at the heart of many clinical protocols since they significantly improve diagnosis and therefore medical care. However, designing a CAD tool for early breast cancer detection remains a difficult and challenging task. In fact, it is hard to conceptualize an expert radiologist's judgment. This research work therefore puts forward a hierarchical architecture for the design of a robust and efficient CAD tool for breast cancer detection. More precisely, it focuses on the reduction of false alarms rate through the identification of image regions of foremost interest (dense breast tissues). Adapted strategies for breast cancer pattern identification can then be applied in priority. The approach hereby introduced relies on two macro-steps. Firstly, raw mammographic images are gotten rid of poorly informative image regions (background and muscle tissues) impairing automatic breast tissue analysis and cancer signs identification. Then, a more advanced analysis is performed on the remaining image to characterize dense breast tissues with respect to their density in order to identify potential cancerous areas.

This PhD manuscript starts with useful insights into mammograms followed by a number of image processing developments to carry out the two macro-steps mentioned above.
In the first macro-step, the dynamic range of gray level intensities in dark regions is stretched to enhance the contrast between tissues and background. This enhancement process favors accurate breast region extraction and suppression of all unwanted patterns in the background image region. A second segmentation follows background suppression. Indeed, some muscle tissues regularly tampering breast tissue analysis remains inlaid in the foreground region *i.e* pectoral muscle tissues. Extracting pectoral muscle tissues is both hard and challenging due to mammogram peculiarities such as the overlap between dense breast and pectoral muscle tissues. In such conditions, even exploiting spatial information during the clustering process of the *fuzzy C-means* algorithm does not always produce a relevant segmentation of this image region. To overcome this difficulty, a new validation process followed by a refinement strategy are proposed to detect and correct the segmentation imperfections and thus enabling accurate pectoral muscle region extraction.

The second macro-step is devoted to breast tissue density analysis. To address the variability issues observed in gray levels distributions with respect to mammographic density classes, we introduce an optimized gray level transport map for mammographic image contrast standardization. Despite the lack of a target histogram distribution, useful parameters can be derived allowing an easier discrimination of mammogram density classes. Thanks to this technique, dense tissues regions are segmented using simple thresholding. We prove that the dense region areas computed from segmented images are highly correlated to density classes from an annotated dataset.

Résumé

Les outils d'aide au diagnostic sont de nos jours au coeur de plusieurs protocoles cliniques car ils améliorent la qualité du diagnostic posé et des soins médicaux. Cependant, la conception d'outils de détection assisté par ordinateur (OAD) pour la recherche des signes précurseurs du cancer du sein deumeure une tâche difficile et laborieuse. En effet, il est difficile de conceptualiser le raisonnement d'un expert radiologue. Ce travail de recherche met en exergue une architecture hiérarchique pour la conception d'un outil d'aide à la détection (OAD) du cancer du sein robuste et performant. Il s'intéresse plus précisément à la réduction des fausses alarmes en identifiant les régions probables de l'image où les techniques plus adaptées doivent être appliquées pour la recherche des indices de cancer du sein de manière prioritaire. L'approche ainsi introduite repose sur deux macro-étapes. L'image mammographique brute est d'abord débarrassée des régions peu informatives et néfastes à l'analyse automatique des tissus mammaires et l'identification des indices de cancer. Ensuite, une analyse plus approfondie est réalisée sur la région de l'image restante pour caractériser les tissus mammaires en fonction de leur densité en vue de déterminer les zones potentiellement cancérogènes.

Ce manuscrit de thèse commence par fournir des éléments nécessaires à l'analyse des mammogrames. Ensuite, de nombreux développements autour du traitement d'images mammographiques sont présentés afin de mener à bien les deux macro-étapes évoquées ci-dessus.
Concernant la première macro-étape, la gamme dynamique des niveaux de gris des zones sombres est étirée pour améliorer le contraste entre la région rassemblant les tissus organiques et l'arrière plan. Cette amélioration de contraste permet une meilleure estimation de la region du sein et la suppression de toute entité superflue présent dans l'image de fond. Toutefois, un tissu musculaire capable d'interférer avec l'analyse des tissus mammaires deumeure incrusté dans la region du sein: le muscle pectoral. Son extraction est à la fois difficile et complexe à mettre en oeuvre à cause de diverses particularités liés aux vues de mammogrames et son chevauchement avec les tissus denses du sein. Dans de telles conditions, même en exploitant l'information spatiale pendant le processus de *clusterisation* par un algorithme tel que *fuzzy C-means* ne permet pas toujours de produire des résultats de segmentation pertinents. Pour s'affranchir de cette difficulté, une nouvelle étape de validation suivi d'un ajustement de contour est mis sur pied pour détecter et corriger les imperfections de segmentation afin de produire une meilleure extraction de la région du muscle pectoral.

La seconde macro-étape est consacrée à la caractérisation de la densité des tissus mammaires. Pour faire face au problème de variabilité des distributions de niveaux de gris constatée en fonction des classes de densités mammographiques, nous introduisons une modification de contraste basée sur un transport optimisé des niveaux de gris est appliqué aux images. Cette manoeuvre a pour but d'uniformiser le contraste d'un

mammogramme à l'autre. Malgré que la distribution cible des niveaux de gris soit mal connue, des paramètres pertinents peuvent être estimés afin que le transport garantisse un pouvoir discriminant accru vis à vis des classes de densité. Cette technique permet ainsi de segmenter les zones de tissus denses par des méthodes de segmentation classiques comme un simple seuillage. Nous montrons que la surface relative des tissus denses estimée à partir des images segmentées est très fortement corrélée à la classe de densité pour des mammogrammes issus d'un jeu de données étiquetées.

List of symbols

i, j, k : discrete variables
a, b, γ : real variables
m, n, p : constant parameters
$\mathcal{G}$: image grid
u : mapping function of an image pixel coordinate to gray level
$\mathbf{p}$: pixel's coordinate of pair (p_x, p_y)
r : radius delineating neighboring pixel of a given pixel
n_w : size of a local window of neighboring pixels
I : intensity of X-rays radiation that have crossed a body
I_0 : initial intensity of X-rays radiation before crossing a body
$\eta(t)$: mapping function of X-rays attenuation crossing a material of depth t
L : number of possible gray level in an image u
$|\cdot|$: cardinality of a set
$\|\cdot\|$: Euclidean distance
v : modified or enhanced image
v_{log} : logarithmic enhanced version of an image u
v_{seg} : segmented image
f, g : mapping function
l_k : number of occurrences of gray level k
τ : threshold value
$p_{a|b}$: conditional probability
$P(l)$: probability of an event l
F_u : cumulative empirical distribution of gray level in an image
x : continuous variable
$\tilde{\mu}$: modified version of the membership function
S_i : weighted membership in a class i
$\mathcal{P}_{\mathbf{p}}$: path around a contour pixel $\mathbf{p}$
σ : standard deviation
$M_{b,\tau}$: gray level mean of background region defined by the threshold τ
$M_{f,\tau}$: gray level mean of foreground region defined by the threshold τ
$\sigma_{b,\tau}$: gray level standard deviation of background region defined by the threshold τ
$\sigma_{f,\tau}$: gray level standard deviation of foreground region defined by the threshold τ
$\sigma^2_{w,\tau}$: within group variance of two image regions defined by a threshold τ
$\mu_i(u(\mathbf{p}))$: membership function of a pixel $\mathbf{p}$ in the image region class i
ν_i : centroid value of the image region class i
J : objective function
ϵ : convergence criteria
$d(c, r)$: Euclidean distance between two point c and r

d_H : Hausdorff distance
θ : angle
R_θ : Radon transform
A_C : area of an image region obtained by segmentation
A_R : area of an image region delineated by a radiologist
A_D : area of dense tissue region
ε : area of a small fractal element of square size
α : fractal scaling constant
φ : fractal dimension
λ : contrast standardization procedure scaling constant
R : region of an image
$\mathbf{h}$: histogram of an image
h_i : i^{th} coordinate of an histogram $\mathbf{h}$
$\mathbf{H}$: cumulative distribution of an histogram
$\mathbf{H}^{-1}$: pseudo-inverse histogram of $\mathbf{H}$
C_{Δ_p} : co-occurrence matrix compute with an of offset Δ_p
Δ_p : offset between two pixels defined by the pair (Δ_x, Δ_y)
$\boldsymbol{\theta}$: vector of model of parameters
sgm : sigmoid function
$\mathbf{x}$: input feature vector of a neural network
$\mathbf{y}$: output feature vector of a neural network
ψ : wavelet function
$\mathcal{W}_\psi$: wavelet transform of an image u with the wavelet function ψ
ω : discrete measure over $\mathcal{G}$
$\delta_{\mathbf{p}}$: Dirac measure on the singleton $\mathbf{p}$
s_{lga} : signature of local gray level appearance
N_g : number of gray levels
$\mathbb{R}$: real number domain
$\mathcal{L}^2$: vector space of square integrable functions over $\mathbb{R}^2$
g_σ : gaussian filter
s_i : subset of the image grid $\mathcal{G}$
$\mathcal{T}$: topographic map collections of level set s_i
Tr : tree whose nodes are connected sets s_i'
ϑ : density function mapping
f_γ : gamma correction transport mapping
d_w : Wasserstein distance
β : relative dense tissue score
ρ : Pearson correlation coefficient
X : multivariate random variable
π^* : optimal transport plan
$E_\pi[\cdot]$: expectation under the law defined by π
$\mathcal{C}$: finite co-domain of a random variable
$\mathcal{X}$: co-domain of the standardized image
$\hat{\mathbf{x}}$: estimate of an input vector $\mathbf{x}$ at output of an auto-encoder
$\hat{x}_i$: component of the output vector $\hat{\mathbf{x}}$
$\hat{c}$: most probable class of an unseen data
χ^2 : Chi square

List of acronyms

CAD	: computer aided diagnosis
WHO	: world health organization
HPV	: human papillomavirus
FN	: false negative
FP	: false positive
MRI	: magnetic resonance imaging
CT	: computed tomography
CCD	: couple charged device
CC	: cranio caudal
MLO	: Medio lateral oblique
LM	: lateral median
2D	: two dimensions
3D	: three dimensions
CADe	: computer aided detection
ROI	: region of interest
BIRADS	: breast imaging reporting and data system
N	: Wolfe's grade representing fatty radiolucent tissue
P_1	: Wolfe's grade indicating fibrous tissue
P_2	: Wolfe's grade indicating fibro-glandular tissue
DY	: Wolfe's grade indicating dense sheets of fibro-glandular tissue
LBP	: local binary pattern
MR8	: maximum response filters (8)
POA	: percent overlap area
FCM	: fuzzy C-means
DWT	: discrete wavelets transform
DCT	: discrete cosine transform
SIFT	: scale invariant feature transform
LBP	: local binary pattern
LGA	: local gray level appearance
PCA	: principal components analysis
BIF	: basic image feature
SVM	: support vector machine
kNN	: k-nearest neighbors
ACR	: American college of radiology
APs	: adaptive pyramids
MST	: minimum spanning trees
V_B	: background region mode
V_P	: pectoral muscle region mode
V_G	: glandular region mode
AG	: average gradient

MIAS	: mammographic Image Analysis Society
HoG	: histogram of oriented gradients
DICOM	: digital imaging and communications in medicine
MNIST	: mixed national institute of standards and technology
SFS	: sequential feature selection
CNN	: convolutional neural network
pSLA	: probabilistic semantic latent analysis
ROC	: receiver operating curve
SMF	: standard mammogram form
SFM	: screen-film mammogram
DM	: digital mammogram
FFDM	: full field digital mammogram
CSP	: contrast standardization procedure
A, B, C, D	: Visual rating of mammographic dense tissues segmentation

Chapter 1

General introduction

Breast cancer and need for computer-aided detection

Cancer is a major public health problem worldwide. The number of people suffering from cancer related pathologies has increased dramatically during these last decades. According to the World Health Organization (WHO), about 13.7% of the global mortality in 2012 was due to cancer. Furthermore, this rate has an upward tendency and is expected to exceed 22 millions of deaths by 2030. Among top lethal cancers, breast cancer appears only at the fifth position of this ranking but it is the second-most common type of cancer among women.

Cancer mainly arises from damages or mutations occurring on genes which contain the core information of cell functioning. In fact, cells multiply through cellular division, grow and die or are destroyed when they can no longer fulfill optimally their role in an organ. This process is controlled by the genes which ensure a certain equilibrium between generation and destruction of cells in activity in an organ. However, due to some unknown factors, a normal cell can slowly and gradually mutate into a pre-cancerous cell and later becomes a potential malignant cancerous cell. In such case, the growth process which is controlled by the genes is no longer respected and therefore results into an anarchic development of aggregated cells called *tumor*. A tumor is benign if the cancerous cells stand over and do not tend to invade and destroy surrounding tissues or organs. In the opposite case, the tumor is said to be malignant.

Current ongoing research efforts attempt to address the problem of cancer symptoms identification at their very early stage because it has been proved that cancerous lesions detected in the early stage can be treated efficiently. The detection process relies on methods that enable to unveil cancerous tumor activity in an organ. As regards breast cancer, such signs of cancer activity are masses, microcalcifications and distortions in parenchymal tissues. Various methods have been developed so far to detect breast cancer including in one hand visual inspection and palpation which are referred to as traditional methods and on the other hand reactive tests and medical imaging techniques also known as modern methods.

Traditional methods rely on basic techniques of detecting breast cancer. However, they constitute the primary examinations often carried out by the patient himself or the physician before going on for further advanced examinations. The main weakness of these approaches is their relative low detection rates. Furthermore, when the detection

is successful, the disease is most often in a serious and advanced state.

Modern methods take advantage of progress achieved in sensors technology to develop imaging techniques that can produce image of inner structure of the body. The most common imaging methods used for breast cancer detection are X-rays imaging, ultrasound imaging and magnetic resonance imaging (MRI). Each of these imaging techniques put forward specific aspects of tissues belonging to organs according to their operating principle. However, among all these imaging methods used for breast cancer detection, X-ray imaging *i.e* mammography, remains the standard method used in detecting early signs of breast cancer on a screening basis. It is the cheapest one and easily accessible with moderately complex examinations protocol and is therefore the most suitable imaging method for screening campaigns.

Mammography is typically used for two purposes: screening and diagnosis. A screening mammogram is prescribed for women who have an empty breast disease history. A diagnosis mammogram is for abnormalities evaluation in women experiencing symptoms. In this case, a particular view or region of the breast may be preferred.

Regarding screening mammogram, it consists of two views of each breast *i.e* one craniocaudal (CC) view and one medio-lateral oblique (MLO) view. The objective of such a strategy is to minimize as much as possible the chances of cancer signs occultation that can occur with tissues overlapping. However, MLO view is most preferred view by radiologists since it allows a better projection of breast tissues with less overlapping than the CC view.

Once the mammograms have been produced, they are to be analyzed by expert radiologists. The detection method is mainly based on human eye inspection. It consists of careful scrutiny of mammograms in order to identify suggestive patterns featuring breast cancer. Such an approach produces good results provided that the radiologist is given good quality images and enough time to visually analyze these latter. In practice, such conditions are rarely met especially during screening campaigns where a huge amount of mammograms are generated and are to be analyzed by few radiologists in a given timescale. In such conditions, sometimes, radiologists may experience fatigue, stress, dazzle and oversight or may have to deal with poor quality image and make radiologists dubious with their own diagnoses. The consequences of an incorrect diagnosis are of important consequences for the by patients. In general, it induces either carelessness or anxiousness to the patients or additional expensive examinations.

To reduce the chances of incorrect diagnosis, double reading has been recommended. The drawback of this approach is the increase of radiologists' workload while it is already difficult to have at disposal two expert radiologists at the same place. Computer aided diagnosis (CAD) has been recently introduced to reduce radiologists' workload while improving significantly the accuracy of their diagnoses. CAD is a system based on artificial intelligence whose output coming from a computerized analysis of medical images is used by radiologists as preliminary opinion in detecting lesions or symptoms patterns and diagnosis making.

Problem studied in this research work

CAD systems have drawn much attention to both researchers and radiologists during past years because of the joint research challenges and potential clinical applications. It is under this scope that the work carried out throughout this PhD thesis has been done. A deep analysis of the literature reveals that the performances of CAD systems introduced to date suffer from many limitations. Several issues are still to address in CAD systems in order to come up with an effective one with comparable performances to human expertise.

In fact, most CAD systems are designed to process either a whole mammogram or a ROI. However, as mammogram is concerned, it is made of a mixture of tissues regions and some of these regions do not contain any valuable information for cancer sign search. More precisely, some examination notes and artifacts may be found on a mammographic image. These patterns and poorly informative image regions can be seen as parasite information harmful for an automatic breast tissue analysis and cancer pattern recognition since they can only induce waste of computation resources and increase of false negative rates.

Another aspect influencing CAD system performances is inherent to the nature of mammograms. Indeed, mammogram is a 2D projected image of a 3D compressed and deformed object. In addition to tissue overlapping producing patterns resembling or altering the appearance of real lesions, tissues shifts induce deformation in shape of masses or distribution of microcalcifications which are important parameters for cancer malignancy assessment.

Nowadays, many imaging methods are available and easily accessible. Each of them put forward some specific aspects of tissues of an organ and provide complementary information in comparison to others. A proper integration of complementary information derived from many imaging sources will enable to improve the accuracy of CAD tools. However, the prime impediment to overcome is to represent the various image modalities in a common reference space. This difficulty, related to the image modality acquisition principle, constrains most CAD systems to mono-modal usage. In fact, the severe compression applied to breast during mammogram acquisition makes it hard to jointly use this imaging modality with other imaging methods such as ultrasound or MRI where almost no compression is applied.

Finally, early signs of breast cancer *i.e* microcalcifications or masses are hardly visible on mammographic images by untrained eyes. Worse than this, these cancer signs have similar texture appearance as dense breast tissues and are located within these tissues most of the time. Looking for cancer signs in such conditions is tedious and complex to carry out even for expert radiologists. On the other hand, medical research studies have shown that women with dense breast tissues have four to sixfold higher breast cancer risks. Besides being an index for cancer risk, mammographic density also implies a higher lesion misdetection risk due to masking by dense tissue and it justify additional computation efforts for advanced analysis.

From the above analysis, it can be noticed that there are still several difficulties and obstacles to address in order to build up a robust and highly accurate CAD tool for

breast cancer detection. In this work, we intend to assess mammographic density in order to facilitate and improve the cancer sign identification task. We therefore adopt a multi-stage strategy to efficiently handle the different peculiarities found in a mammogram that are harmful for the automatic analysis of breast tissue density. In this manuscript, we present a coarse to fine strategy for mammographic density analysis and cancer risk scoring. From raw mammograms, we start by suppressing poorly informative image regions in a fully automatic manner to subsequently address the issues of breast dense tissues segmentation and density assessment.

Contributions of this thesis

The goal of the research work developed in this manuscript is twofold. First of all, we provide methods allowing the identification of mammographic image regions with a very small probability to contain potential cancerous patterns. Second, we introduce a procedure allowing an estimation of breast density to identify high risky population for prioritized medical care. To successfully address the problem of mammographic density assessment, other issues need to be solved. Our contributions lie in several improvements at various steps of this work and can be summarized as follows:

- The first contribution of this thesis is the implementation of a simple, fast and unsupervised approach for accurate suppression of background in mammogram. We first applied an enhancement technique to mammogram to improve boundary identification. Then, follows image segmentation with a boundary refinement strategy to correctly delineate breast image region.

- The second major contribution deals with an *ad-hoc* strategy to automatically identify and segment out the pectoral muscle which appears in MLO view mammograms. It relies on two main steps:

  - a coarse prediction of pectoral muscle boundary,
  - a validation of the estimated boundary followed by an iterative contour refining to improve extraction performances.

- Another contribution of this work is the introduction of a mammogram contrast standardization procedure to address the issue of contrast variability which makes ineffective the use of classical segmentation methods to estimate dense breast tissue proportion in mammographic density assessment. Besides providing a cancer risk index, such an approach helps identifying highly probable cancerous regions *i.e* dense tissue regions where further investigations for cancer signs should take place.

Organization of dissertation

The manuscript is organized in four chapters whose highlights are summarized in the following paragraphs.

Chapter 2 provides an overview on breast cancer as well as the background on mammography and discusses CAD performances and presents possible directions for CAD performance improvement. It also presents the difficulties and challenges to overcome

for an automatic breast cancer sign detection and then introduces a broad algorithmic architecture for mammogram analysis in a step-by-step approach.

Although the core problem in this work is image partitioning (more specifically dense breast tissue segmentation), an effort was made to develop tools for mammogram visual understanding allowing characterization of image regions in raw mammogram in order to first of all suppress parasite regions prior to a discriminating analysis of breast tissues with respect to tissue density.

Chapter 3 and 4 present the algorithms introduced for parasite image regions (background and pectoral muscle) suppression. The strategies developed to tackle each of these image region extraction globally follows the same methodology. A rough estimation of the region of interest is firstly done using clustering techniques, then follows a refinement strategy based on search path technique to produce accurate extraction results of the given image region. However, in the case of pectoral muscle extraction, to accurately address some difficulties such as non uniform gray level distribution in this image region, the classical *fuzzy c-means* algorithm was modified so that it initializes at robust region modes and includes spatial information in the clustering process in order to produce better segmentation results.

The second part of this work is mainly devoted to mammographic density analysis. Chapter 5 first presents the general facts on breast density and establishes the relationship that exists between breast density and cancer risk. Then follows a review of image features used in breast tissue type description and various methods introduced so far in the literature for assessing mammographic density. Finally, to address the contrast variability issues observed in mammographic images, we present the contrast standardization procedure allowing an effective estimation of dense breast tissue proportions through a segmentation by a simple thresholding method.

Chapter 2

Breast cancer and screening process: Overview

Contents

2.1 Introduction

Breast cancer is considered as a major worldwide public health problem. The number of cases diagnosed in western countries as well as in developing ones has risen up dramatically during these last decades. In response to this issue, some efforts have been made to understand the risk factors of this disease though its origins are still unclear. The only way to increase and guarantee better healing chances is to detect this disease at its early stage. The development of medical imaging has significantly improved detection rates of pathologies and therefore enhanced medical care. Among all imaging techniques used for breast cancer detection, X-ray imaging remains the most commonly used one. Breast X-ray imaging is frequently called mammography. This imaging method is widely used in screening campaigns organized in many countries to facilitate the early detection of breast cancer. Consequently, the amount of images generated by usual diagnosis processes and screening campaigns are increasing continuously while the number of radiologists able to analyze such images does not increase at the same rate. Computer Aided Diagnosis (CAD) has therefore been introduced to provide radiologists with artificial intelligence tools allowing to reduce their workload significantly.

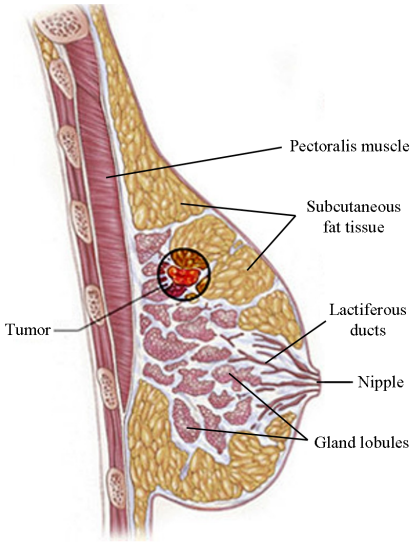


Figure 2.1: Anatomy of a breast gland having a cancer (Tumor)

2.2 Overview on cancer

According to World Health Organization (WHO) statistics, cancer is a major cause of death worldwide. It was the cause of 8.2 million fatalities in 2012, which represented about 13.7% of the global mortality worldwide [1]. The number of deaths caused by cancer will continue to increase in the future and is expected to exceed 22 millions in 2030 [2, 1].

The main lethal cancer types are:

- Lung cancer (1.59 million of deaths),
- Liver cancer (745 000 of deaths),
- Stomach cancer (723 000 of deaths),
- Colorectal cancer (694 000 of deaths),
- Breast cancer (521 000 of deaths).

Breast cancer affects almost exclusively women whereas the others occur in both genders. It is the second-most common type of cancer among women [3].

Cancer generally arises from a single cell. The transformation from a normal cell into a cancerous one is a slow and long process that takes place in multiple stages across time. Usually, a cell undergoing pre-cancerous lesions turns into a potential malignant cancer cell. Cancers arise when the mechanisms of the cellular division which is controlled by the genes are no longer respected. The growth process is completely out of control and results into an anarchic development of aggregated cells called **tumors**. Figure 2.1 shows an illustration of breast gland anatomy with a tumor located in the upper region of the gland between subcutaneous fat and gland lobules. Malignant cancerous tumors grow up, invade and destroy all surrounding normal tissues. A tumor is non-cancerous (benign) if the tumor stands over and does not tend to invade surrounding tissues.

The causes of pre-cancerous lesions are related to genetic, physical, chemical and biological factors. In most cases, cancerous cells appear because of unexpected gene mutations. These mutations are due to external factors such as: smoking, overweight or obesity, low fruit and vegetable diet, physical inactivity, alcohol, sexually transmitted diseases (like human papillomavirus HPV infection), urban air pollution or smoke inside houses produced by faulty heating systems.

Today, a large amount of on-going research efforts address the problem of cancer symptom identification at very early stages. Studies have shown that only cancerous lesions detected in their early stage can be treated efficiently. Therefore, cancer mortality can be significantly reduced if detected or diagnosed early enough. The detection consists in looking for abnormalities featuring a particular cancer inside a population at risk before the disease becomes too grave for quick medical care and treatment.

Depending on the location of the cancerous cells, the detection methods may vary from one cancer type to another. In general, the detection process relies on methods which help to shed light on the activity of cancerous tumors inside an organ. In the case of breast cancer, such signs of activity usually appear as masses[1] or irregularities observed in the tissues. More precisely, some thin spots in mammograms corresponding to microcalcifications[2] can be found.

Various detection methods have been developed. The detection can be done by visual inspection, palpation, reactive tests or modern medical imaging methods. All these detection methods generally require a good interpretation of the physician which is often subjective. For these last two decades, cancer detection is a prominent topic in the field of medical imaging. A physician's diagnosis can be incorrect due to the differences in appearance of images from one patient to another or to the variability of parameters of imaging acquisition systems. Consequently, it becomes difficult to guarantee a relevant diagnosis for all patients. For these reasons and according to the difficulties to face, an appropriate image processing algorithms are needed.

2.3 Screening process

Some types of diseases can be detected before theirs symptoms come to light. Checking for a disease within a group of asymptomatic (apparently healthy) people in an attempt to detect a disease as early as possible is called screening. In the case of breast cancer screening, the precursory signs to track are microcalcifications which are tiny deposits of calcium resulting of pre-cancerous cells activity. Their identification therefore constitutes the early diagnosis. The goal of this process is to detect as soon as possible the beginning of the disease and apply specific treatments which have been proved to be more efficient at this stage of the disease.

[1] Cysts formed by fluid or solid entities in breast tissues

[2] small calcium mineral deposit in the breast tissue. Microcalcifications are of size less than 1/50 of an inch, if they are greater than this, they are called *macrocalcifications*

2.3.1 Definition

The screening process aims at the identification of a disease or an unknown illness by applying tests, examinations or other detection procedures that are efficient, simple and easy to carry out. Screening is therefore a search operation within a target population that does not present prior symptoms of a disease before the symptoms of this latter become obvious.

To build up a screening tool, it is necessary to have a good knowledge of the detection procedure. Also, screening campaigns are very expensive and can be afforded only if the target disease has a major impact on the public health in the society.

2.3.2 Efficiency of screening

A screening process is efficient if it allows to completely detect the precursory signs of a disease when it can still be treated. To be efficient, detection methods should:

- identify every early signs of the disease,
- produce similar diagnosis to that of physician or at least allow him to quickly perform his diagnosis.

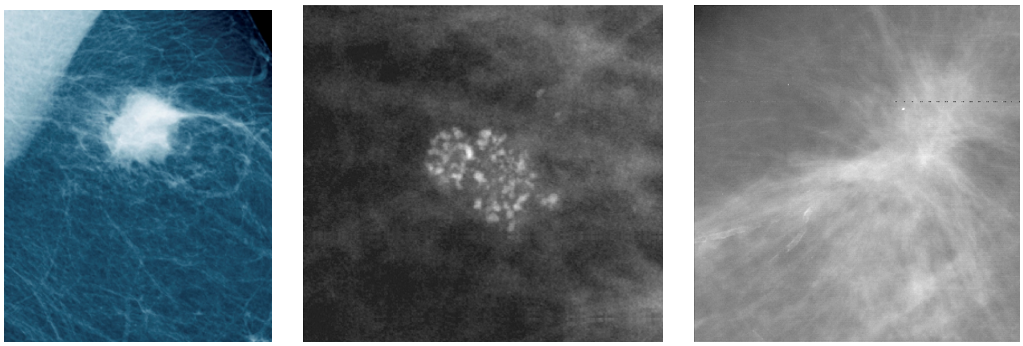
As regards cancer, the goal of the screening process is therefore to unveil the activity of cancerous cells in their early stage.

In the case of breast cancer, precursory signs such as microcalcifications are often very subtle and hardly discernible to untrained eyes. There are different types of screening tests. The most common ones are:

- Physical examination which consists in a general examination of the body looking for signs of illness like nodules or anything that sounds abnormal. The medical history must also be taken into account.
- Biological or histological tests which aim at analyzing tissue samples, blood, urine or other substances of the body.
- Imaging methods which consist in acquiring and analyzing internal images of body parts.
- Genetic analysis that aims at looking for genetic mutation responsible for particular types of cancer.

It should be noted that a detection procedure is not always risk free. For instance, long and repetitive exposure to X-rays can cause cancer. Some tests likes biopsies can induce pain, bleeding and create some complications. It is therefore noteworthy to know the risks and the danger of a screening test to which a patient is exposed before applying it.

Besides, a radiologist's diagnosis can be incorrect due to external factors (fatigue, overseeing, stress, etc.) or poor image quality. Erroneous diagnoses are divided in two sub-categories:



(a) (b) (c)

Figure 2.2: Common signs of breast cancer: (a) masse, (b) microcalcifications, (c) distortion in parenchymal tissue.

- *False Negative (FN) diagnoses*, when patients having the disease are declared unaffected by the disease after the detection procedure. The consequences of such diagnoses are that further medical visits are often postponed even though the symptoms of the disease are coming out.

- *False Positive (FP) diagnoses*, when patients have been diagnosed as having the disease but are healthy. Such diagnoses usually induce additional and expensive examinations tests which make patients becoming very anxious.

2.4 Breast cancer screening methods

The screening methods developed so far to address breast cancer detection are in general goal oriented. Related information as risk factors, heredity, weight, menopause, number of child deliveries, contraceptive usage, alcohol, cigarette and so on are in general not of prime interest but attention is paid on breast gland examination through visual, manual or image analysis. The main purpose of this process is to state whether a patient is potentially having the disease or not before carrying out further examination that may enable to produce a more complete diagnosis. Typical signs commonly put forward in the detection of breast cancer are skin change, presence of nodule inside the breast, microcalcifications, masses and distortion in parenchymal tissue. The three latter which are the most important indices tracked in the field of image processing for breast cancer detection are depicted on figure 2.2.

The detection methods of breast cancer can be sorted into two main groups:

- traditional methods which consists of visual inspection and manual palpation,

- modern methods that, in particular, encompass medical imaging techniques.

2.4.1 Traditional methods

These methods correspond to examinations usually carried out during clinical analysis either by the patient himself or by a physician. Visual inspection is the preliminary approach used in traditional methods. It enables to appraise any change on the skin, nipple, shape of breast gland or discharge.

Palpation consists in a haptic examination of the breast in a preferential position. Despite that this approach seems basic, it is an important step for breast cancer detection. Some studies have revealed that 10% of the tumors detectable in this way do not show up on mammograms.

In the mean time, the main weakness of traditional methods is their low detection rates. Furthermore, when the detection is successful the cancer is rather in a serious and advanced state.

2.4.2 Modern methods

Modern methods mainly refer to medical imaging techniques. Recent progress in sensor technologies have enabled to develop imaging methods that can produce images of the inner structure of a body. Medical imaging therefore allows to show the composition and structure of the soft tissues of an organ. The most common imaging techniques used for breast cancer detection are X-ray imaging, magnetic resonance imaging (MRI) and ultrasound imaging also known as echography.

X-ray imaging uses a low-dose of X-ray radiations to produce a picture depicting the inner constitution of a body or material. The image is produced from various levels of attenuation undergone by X-rays which are related to the physical properties and density of objects they passed through.

Ultrasound imaging uses ultrasonic sound waves to create a picture of a body they reflect on. The particularity of this imaging method is to give information on some physical properties of the body. For instance, it allows to distinguish whether a lump is solid or filled with fluid. A cyst is a fluid-filled benign vesicle whereas solid masses may be cancer.

Magnetic resonance imaging uses a powerful magnet linked to a computer to produce a detailed picture of a body section. This examination enables to reveal the internal functional structure of many organs.

Each of these methods put forward specific aspects of tissues belonging to organs according to their operating principles. However, each of these imaging methods have some limitations. The choice of a given method is therefore based on criteria such as:

- cost of the imaging device and examinations,
- availability of radio-technicians and expert physicians able to analyze such images,
- efficiency of the imaging method to detect the disease,
- examinations prerequisites.

Among all these medical imaging methods used for breast cancer detection, X-ray imaging remains the standard method used in detecting small tumors on a screening basis. It is the cheapest one with moderate complexity owing to examination protocols.

It should also be noticed that the two other imaging methods may be used along with mammography.

In this work, the algorithms for breast cancer detection will be developed and tested only on mammographic image databases. The next section introduces X-ray imaging for breast cancer detection and reviews the principle of mammogram acquisition and analysis.

2.5 Mammography

Mammography is a medical imaging method that is based on X-rays emission to produce an image of the inner structure of the breast. Mammograms *i.e* X-ray images therefore provide a possibility of discovering abnormalities inside breast glands. Mammography is thus a relevant modality for early breast cancer detection as part of screening campaigns because it is cheap and easily accessible. In addition, its examination process is simple and fast as compared to other imaging types and requires almost no special conditioning for patients.

Mammography is typically used for two purposes: screening and diagnosis. A screening mammogram is ordered for women who have an empty breast disease history. It consists of two views of each breast. A diagnostic mammogram is for evaluation of abnormalities in women experiencing symptoms. In this case, a particular view or region of the breast may be preferred.

2.5.1 Mammography principle

Mammography is a specific type of imaging that uses a low-dose X-ray system to examine the breast gland. X-rays is a form of electromagnetic radiation with wavelength in the range of 0.001 to 10 nanometer (nm) and energy in the range of $100eV$ to $100keV$. This radiation can go through a solid object and enables to reveal the inner structure of this object thanks to the various levels of attenuation undergone by X-rays. During mammogram acquisition, the breast is firmly compressed between the film holder and a rectangular plastic paddle. The compression can be uncomfortable but it is essential. First of all, it minimizes the movement of breast tissue during acquisition in order to lower blur effect on the mammogram. Secondly, it reduces the thickness of the breast to create uniform density and minimizes tissue overlapping to produce good quality images. Finally, it allows a lower exposure time and radiation dose to undergo for the patient while producing a detailed image.

The intensity of X-rays that cross a human body and reach the film depends on the physical property and density of materials the body is made of. The intensity of a ray that reaches the film is given by the following equation:

$$I(x) = I_0 \exp\left(-\int_0^x \eta(t)dt\right) \tag{2.1}$$

where I_0 is the initial intensity of the emitted X-rays, $\eta(t)$ is the function representing the X-ray attenuation when crossing a material of depth t. The intensity variations on the film produce an image revealing the inner constitution of the body. Image gray level values are obtained by integrating $I(x)$ over a span of time called exposure time.

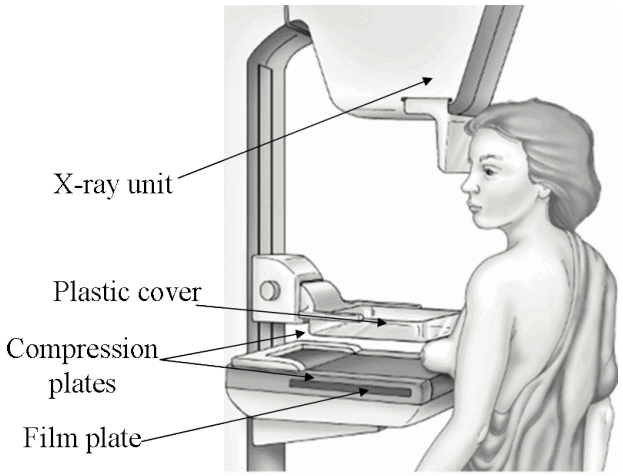
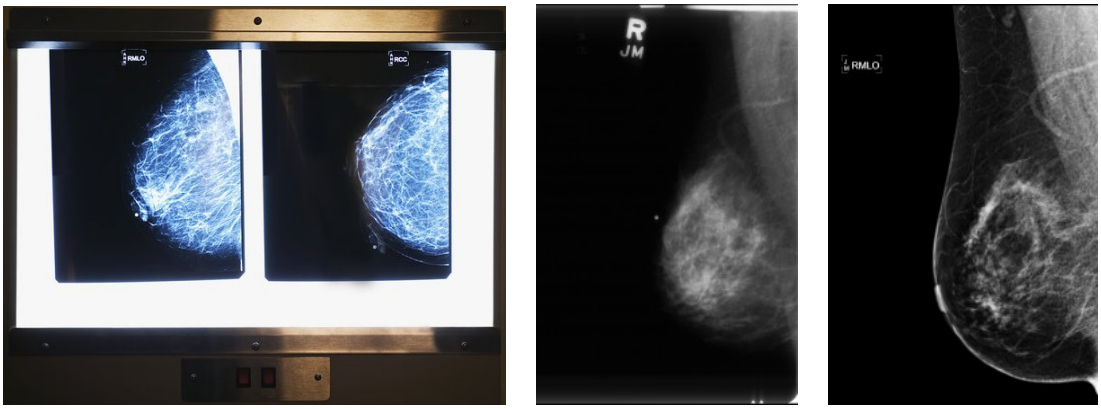


Figure 2.3: Mammography acquisition principle



(a) (b) (c)

Figure 2.4: Type of mammographic image formats. (a) two mammographic films hanged on a ligth-box for reading purpose, (b) digitized version of a screen-film mammogram, (c) digital mammogram

During the mammography examination, breast is squeezed between the paddle and a plate placed under the breast which is made of a cassette containing the film where the resulting picture is formed. The X-rays crossing a human body can either be impressed on a film to produce a mammogram or be sensed by a CCD (couple charged device) and stored on a computer as a digital mammogram. Figure 2.3 shows a mammography machine and illustrates its operating mode.

2.5.2 Mammographic image formats

As previously mentioned, mammograms are nowadays available into two main formats: screen-film mammograms (SFM) and digital mammograms (DM). Figure 2.4 shows typical mammographic image formats. It should be recalled that the same acquisition technique is applied to obtain both image formats and only differs in their recording medium. However, image format may increase or reduce complexity in the automatic analysis of mammograms. The subsequent paragraphs first introduce each of mammographic image format and then discuss their relative advantages and disadvantages.

Screen-film mammogram

Screen-film mammographic images are obtained with radiographic techniques using single-emulsion film to produce high resolution images which are read on intensifying screen (light-box). This technique enables to produce fine image quality at low radiation exposure level. Figure 2.4(a) shows two SFM views (CC and MLO) hanged on a light-box for reading purpose. The light-box is used to enlighten the mammogram in order to magnify the contrast for better visualization of breast structures. This image format has always been considered as the gold standard for breast cancer detection.

It should be noticed that mammographic images in this format need to be digitized prior to their processing with CAD tools. An example of digitized mammogram is shown in figure 2.4(b). The digitization process is a simple mean to produce suitable image format for CAD. However, it should be noticed that the image quality of this image format type is subject to degradation across time. Several defects can alter the image quality such as super-impression from exposure to radiations due to poor storage conditioning, scratches from damage occurred on the film or failure during digitization process. All these defects, produce undesirable patterns on the digitized version of the mammographic image which, in return, increase processing burdens with noticeable effects on CAD system performances. These issues are discussed in the last paragraph of this subsection.

Digital mammogram

A digital mammography is a unit that takes advantage of sensors to directly produce mammogram in digital format from X-rays that have crossed the breast gland. Figure 2.4(c) shows an example of DM mammogram. According to the sensor technology used, digital mammography can be subdivided into two main groups: computer radiography (CR) and direct radiography (DR).

Computer radiography can be considered as an adaptive technique designed to produce a digital mammogram from SFM mammography unit. In fact, in this technique, the emulsion-film contained in the cassette of the SFM device is replaced with a photo-luminescent plate. During the acquisition process, the latent image produced on the photo-luminescent plate is then digitized with a special equipment which uses a laser ray to read the latent image in order to produce a digital version of this latter. The advantage of such a technology is that both mammographic image formats are basically produced from the same mammography unit though an additional device is needed to produced the final digital mammogram.

Direct radiology is also known as full field digital mammography (FFDM). In this type of mammography unit, the cassette is simply replaced with a CCD sensor device that directly records the mammographic image in digital format. It should be noticed that there is no significant difference in mammographic images produced with CR and DR technique. The level of performances achieved by each of these image techniques are almost the same in terms of image quality and cancer detection rates. Nevertheless, acquisition complexity and maintenance issues due to use of additional equipment in CR are real disadvantages in comparison to DM.

Screen-film versus digital mammography

Screen-film and digital mammography each have certain strengths and weaknesses. A clinical study on this topic was carried out by Faridah [4] in order to assess and compare their performances in breast cancer detection. A good review on this concern was introduced in [5]. It is clear that differences in imaging acquisition systems and usage constraints enlarge the panels of comparison criteria. In this work, instead of carrying out a comprehensive comparison between SFM and DM, we rather summarize the important findings reported in the aforementioned publications which are relevant to the present discussion. These findings can be gathered into the following scopes:

- clinical point of view,
- image quality and contrast,
- image manipulation and storage,
- impact on CAD system.

Clinical point of view: From 10 studies reviewed in [5], only two reported a lower cancer detection rates for DM whereas the all subsequent studies have shown higher detection rates for DM. However, this higher detection rate is achieved at the cost of higher recall rates. As a result, there is no significant difference in the positive predictive value between DM and SFM. On the other hand, it was found that the overall diagnostic accuracy of both technologies in detecting breast cancer was similar [6, 7]. Nevertheless, the DM was found to be more accurate in women under 50, women with dense breast and, in premenauposal and perimenauposal women [6]. This is likely due to the wide dynamic range of DM that offers an advantage over SFM. Finally, the most noticeable findings from these clinical studies were a non significant lower cancer detection rate and a significantly lower recall at DM. It was also reported that there is no significant difference in diagnostic performance between SFM and DM. In conclusion, these clinical trials have shown that the overall diagnostic accuracy levels on current digital and screen film mammograms are similar when used in breast cancer screening [6].

Image quality and contrast: The breast is a difficult organ to image as it consists of tissues of various densities and whose amount vary in different women with age. SFM has a high spatial resolution of approximately 16 lines per mm which enables detection of fine structures such as microcalcifications. The spatial resolution of DM is limited by the size of the CCD array. However, this is compensated in DM by the increased contrast dynamics which enhances its ability to visualize small high contrasted structure such as microcalcifications [8]. As compared to SFM, digital systems exhibits a greater dynamic range enabling a better visualization of dense breast tissue, chest wall and the peripheral area of the breast in comparison to screen-film systems.

Unlike SFM in which higher radiation dose produce darker images that are clinically less desirable, DM actually improves the ability to visualize low contrasted structures in the breast. Moreover, because of the difference in thickness and breast consistency, the technologist may fail to determine the optimal settings which could as result yield a poor image quality. In addition, advancement of digital imaging now allows new techniques of breast cancer detection such as use of contrast agent (iodine-based) which

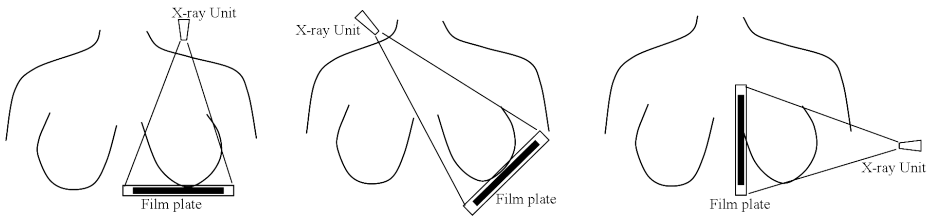


Figure 2.5: Projection angle of mammogram views. (a) Cranio-caudal projection, (b) Medi-olateral oblique projection and (c) Lateral median projection

significantly magnifies the post-contrasted images and other geometric image operations.

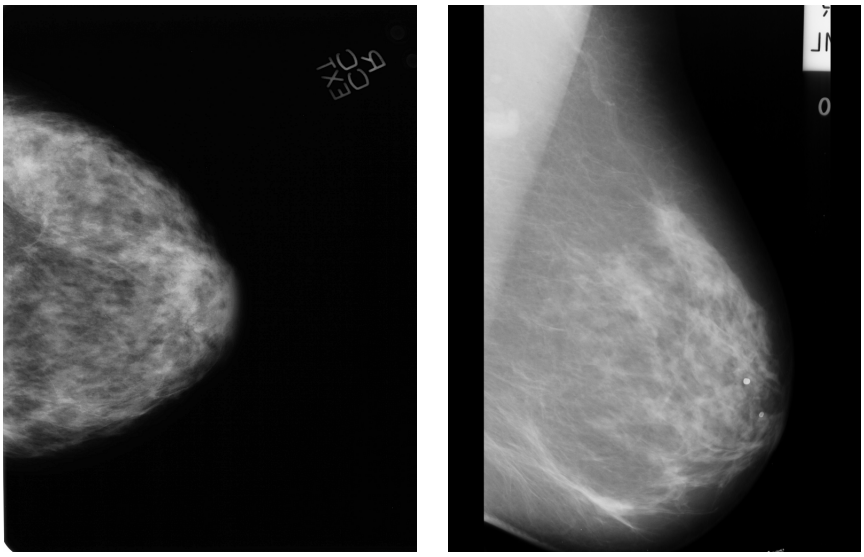
Image manipulation and storage: DM provides opportunities for image post-processing whereas with SFM, an image that has been under/over exposed will lose its diagnostic quality and/or would need to be repeated. DM enables to easily delineate soft tissues especially subcutaneous fat which is not well seen in SFM. However, radiologists need workstation monitors to fully utilize the ability to manipulate DM images. On the other hand, they need training in order to familiarize themselves from hard copy to soft copy reporting.

In terms of image archival, storage and retrieval, it clearly appears that DM better conserves image quality as it is not subject to any damage due to poor storage conditions as it is often the case with SFM. In addition, DM offers several potential benefits as compared to SFM such as elimination of technical failure, recalls, simplified archival, transmission of images, higher patient work-flow, improved diagnostic accuracy especially in dense breast parenchymal tissues due to higher contrast resolution, implementation of straightforward coupling with CAD and tomosynthesis and potential for tele-mammography and tele-consultation [5].

Impact on CAD systems: DM are directly obtained in digital format and need almost no pre-processing prior to their use in CAD tools. Although DM offers more flexibility in terms of image storage archival and retrieval, this image format is still not widely available. On the other hand, screen film mammograms must be digitized prior to their use in a CAD tool. Moreover, some additional pre-processing may be required to handle undesirable patterns due to digitization failure or damages occurring on films. For these reasons, in presence of these two image formats, digital mammograms will be preferred over screen film mammograms for computer aided detection of breast cancer [2]. Nevertheless, the CAD tool wrought out in this research work was developed and tested on a publicly accessible database of digitized mammograms.

2.5.3 Mammographic views

During mammography examination many views of mammograms can be realized for each breast. Depending on the projection angle between the X-ray source unit and the vertical direction (head to toe), three main views are distinguished. Those views are illustrated on figure 2.5 and are referred to as follows:



(a) (b)

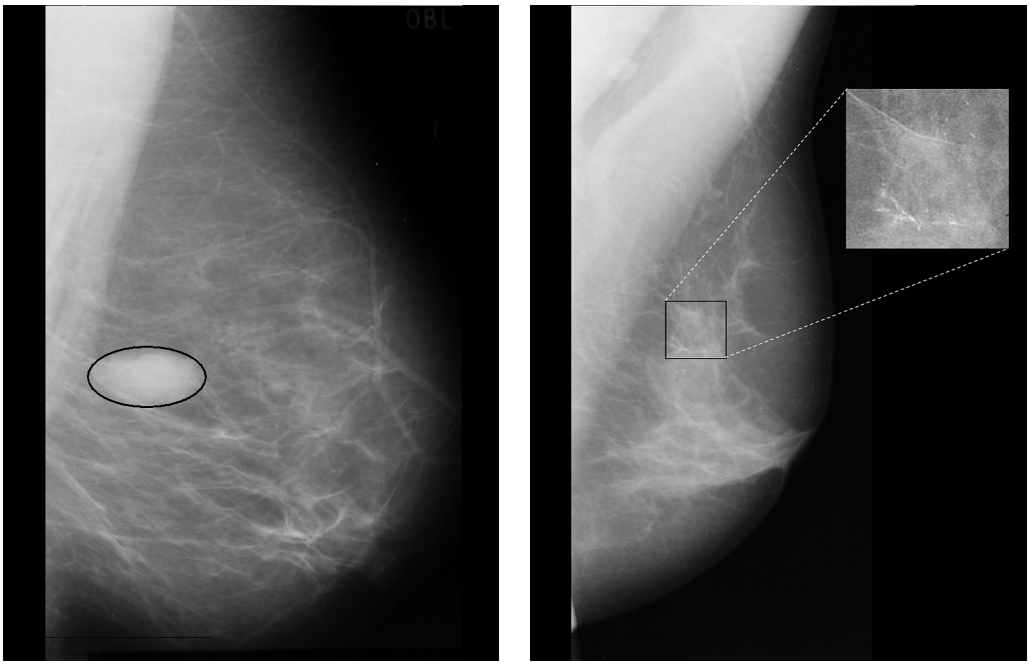
Figure 2.6: Mammographic images obtained under different projection angles of the X-rays system. (a) Cranio-caudal view, (b) Mediolateral oblique view.

- **Cranio-caudal (CC) view**: This view is obtained with the cone beam oriented downward in the vertical direction in such a way that the angle between the vertical direction and the X-rays system tends to 0° (see figure 2.5(a)).

- **Medio-lateral oblique (MLO) view**: this view is obtained with the cone beam of X-rays oriented in the direction parallel to the pectoral muscle so that the angle between the direction of projection and the vertical axis is from 30° to 60° (see figure 2.5(b)).

- **Lateral Median (LM) view**: this view is obtained with the cone beam of the X-ray unit oriented in the perpendicular direction to the vertical direction from head to toe (see figure 2.5(c)).

The CC view is quite simple in realization in practice but the MLO view is more toilsome for lab technologist and patient due to orientation angle of X-rays system and patient's breast positioning between compression plates. Typical images obtained during mammography are represented on figure 2.6. Figure 2.6(a) depicts a CC view while figure 2.6(b) shows a MLO view. LM view is usually used in diagnostic mammograms. Regarding screening mammograms, these latter are usually prescribed to women who are most likely to develop breast cancer. It consists of two views of X-rays images of each breast (one CC view and one MLO view). The objective of such an approach is to minimize as much as possible the chances of obstruction of breast cancer signs that can occur with overlapping tissues. However, the MLO view is mostly preferred by radiologists since it produces a better projection of inner breast tissues with less overlapping than the CC view. In the sequel of this work, the approaches discussed or introduced assume that mammograms are exclusively MLO views unless explicitly stated otherwise.



(a) (b)

Figure 2.7: Example of breast cancer signs identified in mammograms. (a) manually circumscribed mass with a bold black ellipse, (b) microcalcifications appearing as bright spots in the focused ROI.

2.5.4 Mammogram limitations

Screening mammograms are performed to detect breast cancer in its early stage and to guarantee medical care efficiency. Once mammograms are performed, they are to be analyzed by an expert radiologist. This latter looks for suggestive patterns of breast cancer. Figure 2.7 shows MLO views of mammograms where areas containing suspicious signs of cancer have been manually delineated. Such patterns in mammographic images that must draw a radiologist's attention are masses (figure 2.7(a)), microcalcifications (figure 2.7(b)) but also distortion or asymmetry in parenchymal tissues. However, early signs of breast cancer are hardly visible on mammographic images by untrained eyes. The difficulty in identifying cancer signs is due, in general, to their similarity in textural appearance to that of dense tissues.

Besides, mammography has intrinsic limitations impairing analyses aiming at detecting cancer signs. First of all, a mammogram is a 2D image of a 3D compressed and deformed object. Breast tissues overlapping are likely to produce suspicious patterns or alter the appearance of real mammographic lesions [9]. Breast tissues shift during compression resulting in deformations in shape of masses or microcalcifications foci which are important parameters in the assessment of breast cancer malignancy. More specifically, the occlusion of cancer signs are due to the fact that in most cases they are located within dense tissues.

Moreover, a major inconvenience of mammograms is their low contrast [10]. Consequently, edges and boundaries of regions are less discernible. As a result, when two regions that are close in texture appearance overlap, as it is often the case of dense tissues with masses or microcalcifications, these latter become hard to identify. Radi-

ologists may fail to produce accurate diagnoses in such situations.

Also, the deformation undergone by breast during mammogram acquisition makes it difficult to spatially localize entities in the breast. This results in unnecessary biopsies when analyzing suspected areas identified in mammogram.

Another difficulty pertaining to mammography is the impossibility for medical staff to produce images with identical quality from one patient to another because of differences in their breast consistency. Consequently, cancer signs are slightly different from one image to the other and thus require more attention of radiologists or robustness of CAD tools to handle all these disparities successfully.

Finally, some studies [11, 12] have shown that mammography analysis is subject to false positives as well as false negatives, causing a high proportion of women without cancer to undergo further clinical evaluation or breast biopsy, and some other women to miss the best time interval for cancer treatment.

2.6 Visual breast cancer detection

It should be noted that masses and microcalcifications are detectable with the help of mammography long before they become clinically palpable [13]. For this reason, mammography is considered as a valuable screening procedure that can detect breast cancer early. The detection of breast cancer is mainly based on an interpretation and analysis of medical data (imaging, personal information, medical history, etc. . .). Although the risk group of a patient can be easily derived from personal data, the most difficult task in breast cancer detection is to analyze a mammogram and state whether this latter contains suspicious patterns of breast cancer or not. Two approaches of image analysis are to be considered in this particular case: human eye inspection and CAD analysis.

2.6.1 Human eye analysis

As previously mentioned, visual analysis consists of scrutinizing a mammogram in order to identify patterns featuring breast cancer. This approach yields good results provided that the expert radiologist is given good images and enough time to visually analyze mammograms. In practice, these conditions are not always met, especially during screening campaigns where a huge amount of mammograms are to be analyzed by few radiologists in a given timescale. In general, analyzing mammograms in such conditions induces fatigue, stress, oversight which in turn alter radiologists' diagnoses. A complication comes from the uncertainty of radiologists in their own judgements.

As mentioned before, incorrect diagnoses may have serious consequences. For healthy ones, it creates anxiousness and they start a long and expensive examination process to finally find out that they have no cancer. Conversely, it creates in non-healthy patients carelessness which leads them to postpone all examinations (even those normally scheduled) to later discover they have cancer in an advanced stage.

To reduce the chances of incorrect diagnosis and thus the proportion of missed cancers, double reading of mammograms has been advocated [14, 15]. However, retrospective

studies of mammograms have shown in one hand a high disparity in results with previous reading and, on the other hand, an increasing number of missed tumors [9]. In addition, double reading increases radiologist's workload.

2.6.2 CAD image analysis

The goal of CAD approaches is to reduce radiologist workload and improve the accuracy of their diagnoses especially during screening campaigns. They consist in developing algorithms that can perform a defined task; that is, analyzing a mammogram with or without a human interaction. All approaches developed in this domain rely on digital image processing. The image analysis algorithms developed in CAD systems aim at extracting relevant patterns in mammograms or analyzing patterns identified by an experienced radiologist. Relevant reviews and recent advances on breast cancer detection and diagnosis using mammography can be found in the following works [2, 9, 16].

Computer aided diagnosis can be defined as a diagnosis that is made by a radiologist who uses the output from computerized analysis of medical images as a preliminary opinion in detecting lesions or symptom patterns [17]. It is an artificial intelligence tool that has recently been introduced to reduce radiologists' workload while improving their diagnoses [18]. CAD is a class of approaches that integrates imaging with, computer sciences, image processing, pattern recognition, and artificial intelligence technologies [16]. In the past years, CAD systems have drawn much attention of both researchers and radiologists because of joint research challenges and potential clinical applications [2]. The number of papers published on CAD issues illustrates the level of interest carried on this area of research.

The main issues addressed in related works reported in the literature deal with microcalcifications detection [19, 20, 21, 22], masses detection [23, 24, 25, 26, 27], parenchymal distortion [28, 29, 30], microcalcifications foci reconstruction [31] and more recently, breast tissue characterization [32, 33, 34, 35]. Another issue in the literature that authors attempted to tackle deals with the classification of suspicious areas in breast into benign and malignant groups [19, 36, 37, 38, 39]. All the methods introduced in these publications have been developed in order to improve CAD tools for breast cancer detection. Although, they have met various degrees of success, none of them can claim to have entirely solved all mammogram visual understanding problems. This topic remains thus a vivid research topic.

The performances obtained by CAD systems are promising since they drastically reduce radiologists' workload and are reasonable alternative to double reading. Although modeling expert radiologist reasoning seems hard and complex to conceptualize, CAD systems use artificial intelligence approaches in an attempt to produce similar level of image understanding to that of radiologist or better. To achieve this task, a CAD system usually resorts to data annotated by radiologists. Based on a radiologist analysis procedure of a mammogram, one can derive two main steps in designing a CAD system:

- localization and delineation of suspicious area or region of interest (ROI) which usually involves image segmentation,
- analysis and identification of potential patterns of breast cancer in the region of

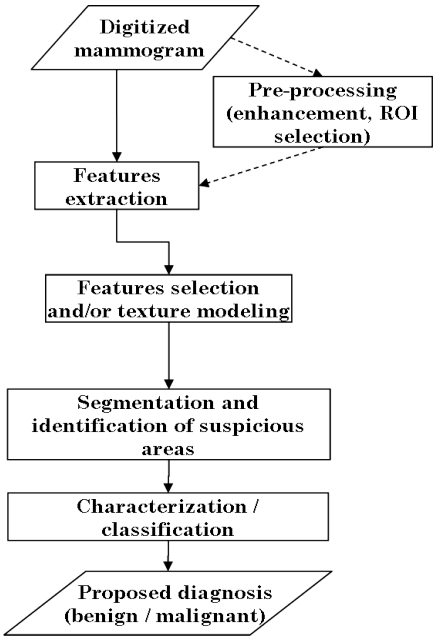


Figure 2.8: Typical algorithm flowchart of a CAD system for breast cancer signs detection

interest. In general, this step involves a classification task whose output can be regarded as a preliminary diagnosis.

A typical flowchart for a CAD system is depicted on figure 2.8. The main steps of the algorithm for mammographic image analysis can be identified as image pre-processing, segmentation, features extraction, features selection and classification.

2.6.3 Image manipulation effects on CAD performances

The development of computer aided diagnosis techniques has motivated increased efforts in the acquisition and derivation of high quality mammographic images. Though the performances of CAD systems are mainly dependent on the efficiency of the approach elaborated and the optimality of their implementation, they are also sensitive to indirect effects related to image manipulation.

CAD systems usually focus on region of interest analysis for the purpose of masses or microcalcifications detection and classification. However, the texture pattern derived from the spatial variations of gray levels in the ROI neighborhood is often weaker in comparison to those of cancer area and may consequently tamper somehow ROI feature analysis. This therefore raises the issue of optimal neighborhood size estimation for texture analysis. In addition, mammogram size is usually of thousands of pixels which increase computational burden. On the other hand, some preprocessing operations are often applied to mammograms to increase indicative breast structures visibility. All these image manipulations have significant influence on CAD's performances. The subsequent paragraphs therefore investigate on the most important of these effects which are due to:

- pixel resolution and integration scale,
- preprocessing.

Effect of pixel resolution and integration scale: Digital or digitized mammograms are usually acquired at about $50\mu m$ per pixel using 12 bits depth. However, such a resolution would be excessively large for practical computation. Concerning bit quantization, some gray levels will occur with low or negligible rate of incidence that may not permit to derive reliable statistics for cancer identification [40]. Adequate pixel bit depth and resolution thus play a significant role in feature quality and consequently in classification accuracy.

Several authors [40, 41] have studied the influence of these effects on feature discriminating breast tumors as benign and malignant masses. The motive underlying these studies was to derive a good trade-off between performance accuracy and pixel resolution. Shape feature was used in these studies as it is a powerful feature in discrimination of malignant and benign masses. The boundaries of malignant masses usually have irregular shape while that of benign masses have regular ones. Pixel resolution and integration are in this case a critical factor because downsampling may remove some fine details from the image.

In both studies, the feature performance accuracy in classification was assessed at various resolutions and scales. It was noticed that pixel resolution significantly influences the performances. It was also observed that a resolution as from $200\mu m$ per pixel is a good trade-off for better performances. Obviously, when resolution is very low, the classification performance degrades as the shape of the boundary of benign and malignant masses will be very similar.

Preprocessing: The development of CAD tools has raised a growing interest among researchers during this last decade especially in the area of biomedical applications. Yet, one of the major difficulties in the design of robust CADs is the lack of image quality standards (*i.e* image quality which satisfies both clinical and CAD purposes). As a result, some images acquired under particular settings will not be appropriate for CAD use. Some preprocessings are therefore needed to be done on this later in order to fulfill the necessary criteria for CAD systems.

The goal of preprocessing approaches is to make small scale structures in an image more visible in order to increase the discriminative power of an image analysis method. The effect of preprocessing on classification performances was studied in [41]. In this study, three processing methods including contrast limited adaptive histogram equalization, median filter and image sharpening were used to assess their impact on the performances of some texture analysis methods and classification. The authors found out a coherent relation between the efficiency of some texture analysis methods and the preprocessing used. Nonetheless, the real difficulties in such conditions is always to find the preprocessing approach that will in average improve the performances of a given classification method.

2.6.4 CAD limitations

Following recent progress in machine learning, CAD systems can reach remarkable performances. However, failures may occur and one should always keep in mind the limitations of such systems. The following paragraphs discuss some of the limitations in common CAD systems.

Most CAD systems process either the whole mammogram or a ROI. In both cases, the processed image contains a mix of undesirable tissues or entities that may tamper the automated analysis. Tissues that do not belong to breast can be regarded as parasite information impairing cancer pattern recognition and wasting computation resources leading to high false negative rates. Indeed, some muscle tissues have unfortunately similar gray level as some dense breast tissues. Although some morphological and shape information can help to discriminate these tissues, this similarity in gray level is a source of confusion for algorithms.

Another aspect influencing CAD system performances is the number of image modalities that they are able to process. Nowadays, many imaging methods are available and each of them put forward some peculiarities of tissues contained in a body. An appropriate integration of various information derived from many imaging sources will enable to produce a more robust CAD tool. Information fusion techniques can provide useful algorithms to adequately combine information coming from many sensors. A valid multi-modal CAD should contain a data fusion step allowing better performances than each mono-modal CAD individually. Designing such a data fusion is a challenging task beyond the scope of this manuscript.

Finally, another limitation of CAD systems has its roots in the breast compression operated during mammogram acquisition. Since breast is a soft organ, compression induces displacement and deformation of the breast structure. Some superposition occurring between breast tissues often yields patterns that appear like suspicious signs of cancer or alter the appearance of real mammographic lesions [9]. In addition, breast deformation is likely to make it harder for data fusion with ultrasound or MRI modalities for which almost no compression is applied.

Spatial layout of lesions which is also a useful index in assessing cancer malignancy is also hard to represent from a mammogram. A possible solution consists in 3D reconstruction from multiple views of mammogram based on a physical model of compression and tissue deformation. This is also a mean to allow an accurate localization of lesions inside a breast.

2.7 Difficulties and challenges in automatic analysis of mammograms

Many obstacles are to overcome in the automatic analysis of mammograms. These impediments have their roots in low contrast quality, mammograms layout and digitization. In addition, mammographic images usually contain parasite image regions for CAD systems such as examination protocol related textual data and background. Likewise, some muscle tissues regularly appear in MLO view. These tissues belong to the pectoral muscle and do not bring any valuable information for breast tissues analysis. Worse than this, these muscle tissues have similar density as glandular ones, thereby impairing dense breast tissue processing. Such specific characteristics of mammogram thus require adequate pre-processing to meet CAD requirements in terms of image quality and to save some computation efforts for more interesting tasks. Intuitively, a multi-stage approach appears to be necessary for designing robust CADs.

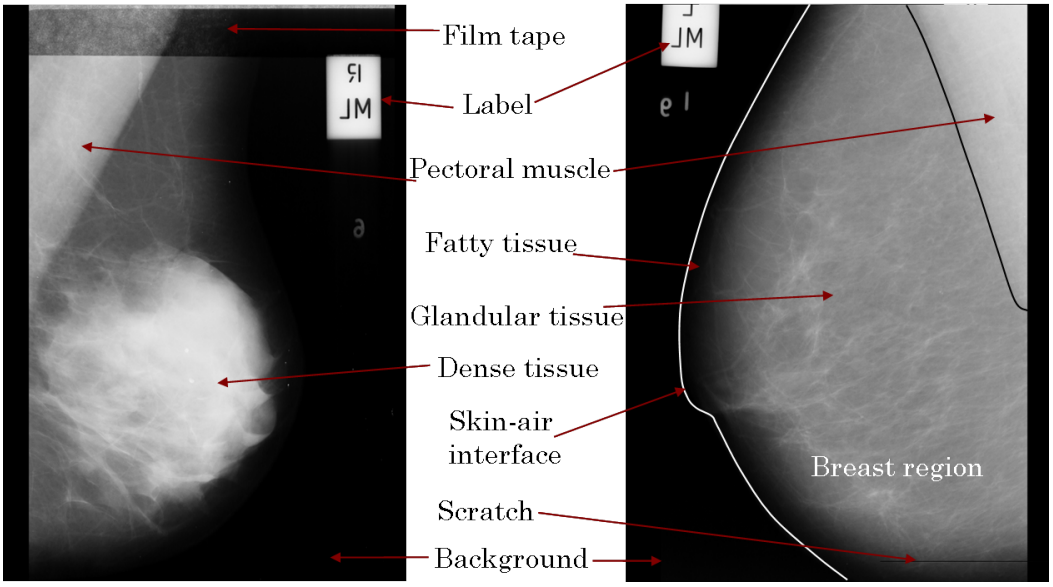


Figure 2.9: Mammographic images contents with some undesirable patterns for the automatic analysis. On the left, a raw mammographic image and on the right a mammographic image where the pectoral muscle region (**black**) and the breast region (**white**) were manually delineated from background (**dark region**). The different breast tissue types are indicated with arrows as well as view labels and digitization failures most often encountered.

This section reviews common difficulties and challenges arising at different levels of an automatic analysis of mammograms. First, we describe mammogram contents and then discuss the influences of some regions on mammographic image analysis. Finally, this section gives an insight of important steps to implement in order to design an efficient CAD tool.

2.7.1 Difficulties in mammographic images

Mammographic image contents

Figure 2.9 shows two MLO view mammograms. On the right image, the main tissue types have been delineated manually. The scratch at the bottom of the image is the result of a digitization failure. A damaged film was used to record the left image, thus creating a parasite grayish stripe near the upper image border. Both images also contain textual data indicating the types of the mammogram views and other examination details. These examples belong to a first class of difficulties called acquisition shortcomings.

A mammogram is basically made of two main regions: foreground and background (see figure 2.9). X-rays with obstacle-free trajectories yield pixels belonging to the background region while X-rays with tissue-crossing trajectories yield the foreground region. Unlike most images, pixels receiving the largest amount of photons are the darkest ones and those receiving the fewest are the brighter ones. Roughly speaking, the background region contains pixels with no valuable information for breast tissue feature extraction, therefore textual annotations are also belonging to this region. As opposed to background, the foreground region contains all pixels corresponding to mammary tissues or to the pectoral muscle. Mammary tissues are in turn divided into

three sub-regions:

- fatty tissue region,
- glandular tissue region,
- dense tissue region.

The frontier between regions are less discernible in X-ray images to due to their low contrast and the slight variation in intensity across boundaries. Good contrast in mammograms (producing a better visualization of breast tissues) depends on a judicious selection of the range of X-rays intensity and exposure time. However, breast consistency is uncertain and a radio-technician can only presume that a given setting will produce a contrasted image. This accounts for the variability in mammographic image quality.

One can also notice that dense tissues and pectoral muscle appear as brighter areas and can be easily differentiated from fatty and glandular tissues which are radiolucent. Conversely, it is difficult to clearly delineate the frontier between fatty and glandular tissues as well as the boundary between pectoral muscle and dense tissues. In addition, the amount, the size and layout of these tissue types vary from one mammogram to another.

Mammogram digitization issues

In the past, mammograms were essentially produced in screen film format. The conservation of such films is delicate as they should be kept away from any radiation that can super-impress parasite patterns or, in general, any damage making mammograms useless for CAD purposes. With the development of CAD systems for breast cancer detection, the need of digital mammograms has emerged. Screen film mammogram digitization is a simple mean to obtain digital mammograms in radiology department that cannot afford digital mammographic sensors yet. However, digitization is not harmless to informative content. First, an appropriate sampling frequency or image resolution must be chosen. Second, a quantization factor must be selected in order to grasp a sufficient range of image intensity variations without generating too massive data for computer memory. The number of bits to encode pixel values is often called pixel depth. In addition, any defect on the screen film mammogram (artifact, scratches and so on) will be reproduced on the digitized one. These digitization issues are well-illustrated on mammograms shown on figure 2.9.

These inconveniences often impose additional pre-processing to reduce or annihilate their effects on CAD system performances. Fortunately, digital mammography is becoming increasingly affordable and is progressively replacing screen film ones. Although clinical trials showed that the overall diagnostic accuracy is similar for these two types of mammograms when used for breast cancer detection [6], digital mammography features some potential advantages for CAD purposes.

2.7.2 Challenges in automatic tissue density analysis

Many patterns such as breast contour, nipple, pectoral muscle or tissue density are useful primitives in breast cancer sign detection but cannot be extracted simultaneously in

mammograms. For instance, breast contour is a primary pattern used in nipple localization and pectoral muscle identification tasks. These patterns are likely to influence CAD performances if they are not handled separately or in a meaningful order. Given that dense tissues are a common area for cancer to develop, detection of cancer signs should follow a coarse to fine approach by rejecting all poorly informative regions (non breast tissue regions) while focusing only those of high probability to contain cancer (dense tissues).

Poorly informative regions removal

Regarding cancer sign detection, two disjoint regions of mammograms are poorly informative: background and pectoral muscle. A prior extraction of these regions has proved to provide improved CAD system performances [42, 43, 44, 45, 46].

Concerning background extraction, a dual problem consists in breast contour detection. Indeed, this contour is always the boundary of the background region. Addressing breast contour detection is more interesting in the sense that such approaches are far less sensitive to occasional textual annotations that may, or may not, be present in the background region. Despite being easier, breast contour detection is still a rather hard task. Occasional artifacts mainly impair contour detection algorithm initialization. Poor (and often irregular) contrast impairs contour detection algorithms on the whole. It is also noteworthy that the transition from breast region to background is very smooth since the tissue quantity is decreasing progressively.

The second image region that must be discarded from further processing is the pectoral muscle region. Not only does this region have no relevant information regarding breast cancer but it has similar texture features as dense breast tissues. This similarity hinders cancer detection. Running mammogram analysis algorithms without removing the pectoral muscle region before is known to lead to both false negatives and positives [45, 47, 34]. In addition, cancerous lesions are more frequent along the pectoral muscle surface. Consequently, removing the pectoral muscle helps highlighting such lesions.

Efficient design of CAD system

As discussed in the previous section, poorly informative regions removal are prerequisite steps for an efficient cancer detection approach. Such pre-processing steps are illustrated on figure 2.10. This process enables to obtain an image representing only breast tissues. Furthermore, breast tissues can be subdivided into non dense and dense tissues. It should be noticed that non dense tissues are fatty and glandular tissues and are less likely to contain cancer. Conversely, dense breast tissues are areas with a high probability to contain cancer. The rightmost picture of figure 2.10 depicts the targeted image sub-region corresponding to dense tissues only.

The most challenging steps in the process of cancer detection are breast tissue analysis and suspicious cancer pattern extraction and classification into, for instance, malignant and benign classes. Breast tissue density analysis implies looking for a relevant feature space to describe breast tissue texture. The primary difficulty at this level is to build adequate models with low level features that globally describe the most relevant and common patterns of each breast tissue type. However, because of the variability observed in mammographic image contrast quality, most of usual image analysis features

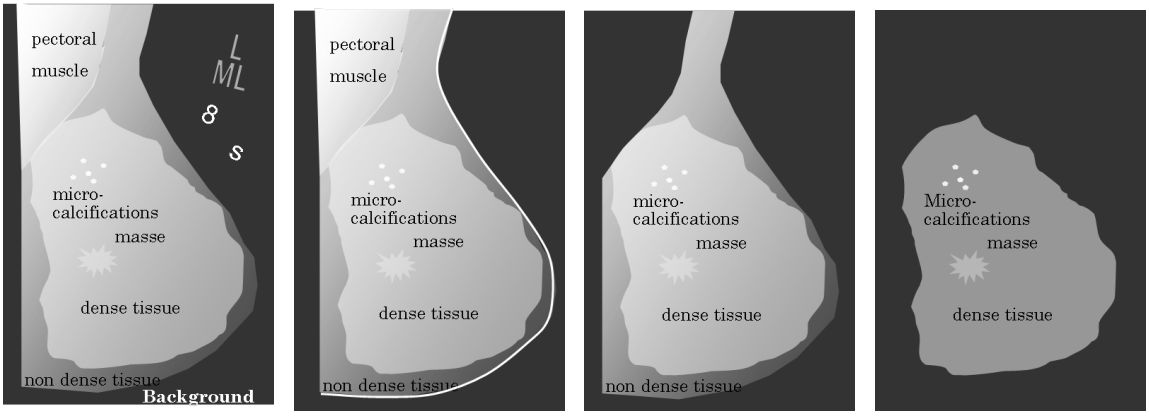


Figure 2.10: Sketch of an efficient strategy for mammogaphic tissue density assessment and cancer sign search. The leftmost image depicts an initial mammogram model with harmful artifacts for a CAD tool present in the background. Then follows a model where background region has been identified and removed with breast contour delineated (**in white**). The third image represents a model with the pectoral muscle segmented out. At this level the image is constituted of breast tissues only which is basically the region of interest for breast tissue density assessment. The rightmost image shows a model where mammographic density analysis has been performed so that only dense tissue area is represented on the image whereas the glandular tissue is discarded. Dense tissue regions represent areas where cancer signs are most likely to develop. Consequently, they feature an increased concern and require a careful analysis for cancer signs search.

will fail to yield conditional class distributions given the input image with compact supports. As a result, these distributions overlap in the feature space therefore leading to moderate CAD performances.

Although breast tissue density analysis itself can provide a cancer risk index associated to a given mammographic image, cancer detection is preferably performed by looking for specific patterns like masses or microcalcifications which correspond to a higher abstraction as compared to tissue types. Such pattern recognition requires its own appropriate feature selection as well, suggesting again a hierarchical (coarse to fine) approach for this machine vision task.

There are several difficulties in retrieving higher level features for cancer patterns. First, they are scattered over the dense breast tissue sub-region and are consequently poorly contrasted. A ghost texture of dense tissues is super-impressed on them because of X-ray imaging acquisition principle. In addition, the shape of cancer signs is quite random, especially microcalcifications whose surfaces are counted in a few pixel units.

As conclusions, to successfully overcome difficulties arising in automatic analysis of mammograms, a coarse to fine design of CADs should be preferred. In view of all the above, one can reasonably design a CAD system with respect to the architecture represented on figure 2.11. In this work, we first of all address mammogram visual understanding tasks to remove uninformative regions of the image (chapters 3 and 4) and subsequently, we carry out an histologic analysis of this latter to assess cancer risk (chapter 5).

2.8 Conclusion

In short, breast cancer is a worldwide health problem that affects almost exclusively women. There is still no way to prevent this disease and the only mean to fight against it relies on early detection. To encourage this early detection, many countries regularly organize screening campaigns which always generate an increasing amount of mammograms to be analyzed in most of the cases by few radiologists in a defined timescale.

To reduce incorrect diagnosis rates of radiologists, double reading of mammogram has been prescribed which in turn increases radiologists' workload. To cope with these difficulties, CAD systems have been introduced to provide radiologists with an artificial intelligence tool allowing them to improve their diagnosis accuracy while reducing drastically their workload. However, like for any intelligent system, there is still room for improvements in detection rates and computation time so that CAD systems may one day outperform experts in cancer pattern recognition.

Several perspectives for significant improvements to CAD systems have been pointed out in the previous section. Among the issues reviewed, two of them are noteworthy for extended investigations: multi-modal data combination and breast tissue analysis. The first improvement consists of developing CAD tools that combine information derived from various imaging modalities which are nowadays more easily accessible. The second improvement is oriented towards the use of breast histological information[3] in the CAD process.

Beside the fact that histological information can help to reduce false alarm rates, it also appears to be a relevant source of information for cancer risk assessment and cancer sign detection. In turn, deriving such histological information relies on efficient breast tissue characterization. This point constitutes the main interest of our research activities and is thoroughly investigated and addressed in this work.

More precisely, some prior information on mammographic image layout must be exploited. Roughly speaking, a mammogram is divided into three regions: mammary gland, pectoral muscle and background. The background and pectoral muscle contain no relevant information regarding breast cancer and must be removed using appropriate segmentation algorithms.

Finally, a hierarchical image analysis task must be performed on the remaining image region, *i.e.* breast region. This region must be first represented in a relevant feature space for tissue type analysis and subsequently for cancer pattern detection.

In summary, CAD systems for digital mammograms rely on two macro-steps : poorly informative region removal and breast region analysis. Each of these steps are themselves divided in sub-steps as can be seen on figure 2.11. Each of the sub-steps of informative region retrieval is itself a challenging image processing task and are developed throughout chapters 3 and 4. Chapter 5 deals with the next step, *i.e.* breast

[3]Histology is the study of organic tissues. Breast histological information is therefore information derived from the study and analysis of breast tissues. In the framework of this thesis, it mainly refers to tissues density information.

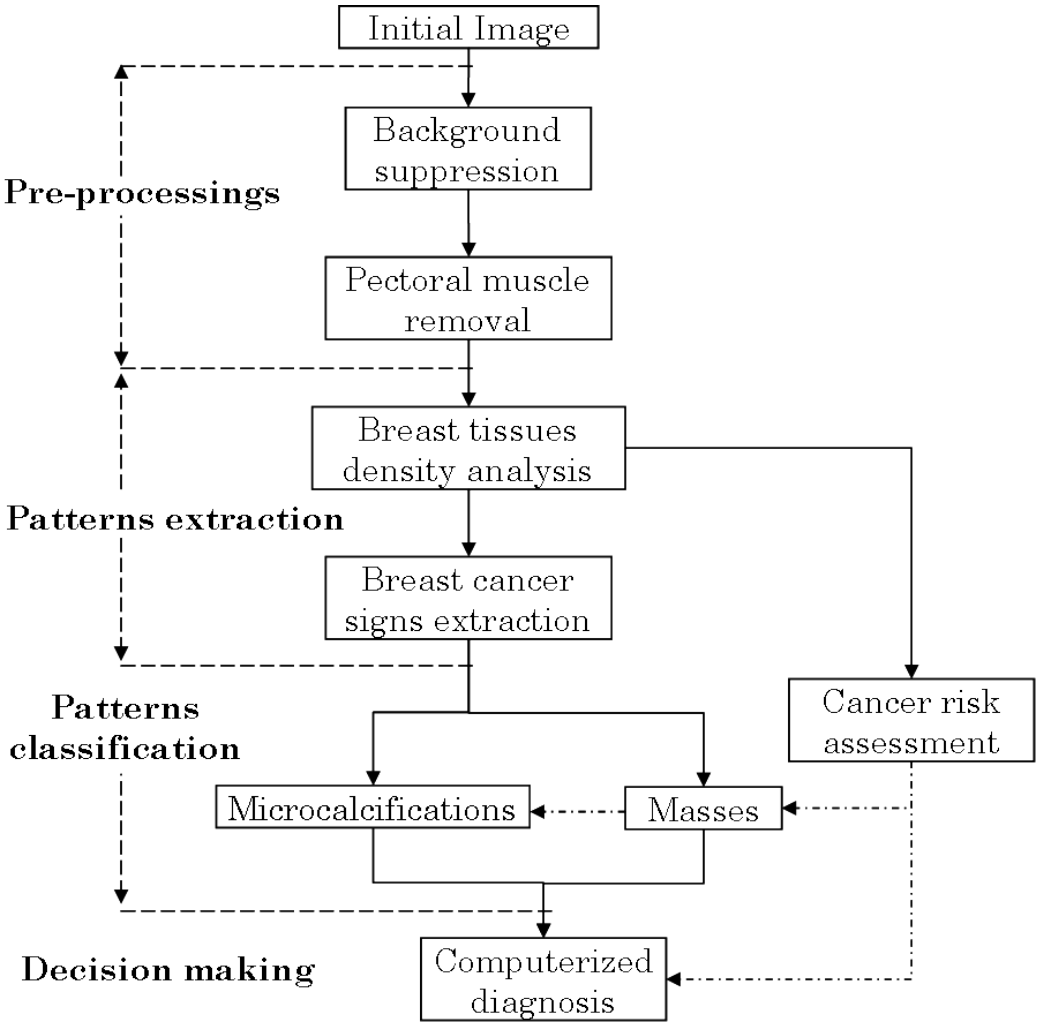


Figure 2.11: Flowchart of a recommended architecture for breast cancer detection approaches in mammograms. Solid arrows stand for steps of mammographic image processing and dash-dot arrows indicate integration of cancer risk information in the process of cancer signs classification and diagnosis making.

tissue density computerized analysis

2.9 References

[1] W. H. Organisation, "World cancer report, WHO cancer fact sheets." IARC Press, 2014. 8

[2] J. Tang, R. M. Rangayyan, J. Xu, I. E. Naqa, and Y. Yang, "Computer-aided detection and diagnosis of breast cancer with mammography: Recent advances," *IEEE Transactions on Information Technology in Biomedicine*, vol. 13, pp. 236 – 251, March 2009. 8, 17, 21

[3] M. Althuis, J. Dozier, W. Anderson, S. S. Devesa, and L. A. Brinton, "Global trends in breast cancer incidence and mortality 1973−1997," *International Journal of Epidemiology*, vol. 34, no. 2, pp. 405 – 412, 2005. 8

[4] Y. Faridah, "Digital versus screen film mammography: a clinical comparison," *Biomedical Imaging and intervention Journal*, vol. 4, no. 4, p. e31. doi:10.2349/biij.4.4.e31., 2008. 16

[5] P. Skaane, "Studies comparing screen film mammography and full-field digital mammograpy in breast cancer screening: Updated review," *Acta Radiologica*, vol. 50, no. 1, pp. 3–14, 2009. 16, 17

[6] E. D. Pisano, C. Gatsonis, E. Hendrick, M. Yaffe, J. Baum, S. Acharyya, E. Conant, L. Fajardo, L. Bassett, C. D. R. Jong, and M. Rebner, "Diagnostic performance of digital versus filmmammography for breast cancer screening," *New England Journal of Medicine*, vol. 353, no. 17, pp. 1773 – 1783, 2005. 16, 26

[7] J. M. Lewin, C. J. D'Orsi, E. Hendrick, and et al., "Clinical comparison of full-field mammography and screen film mammography for detection of breast cancer," *American Journal of Roentgenoly*, vol. 179, no. 3, pp. 671–677, 2002. 16

[8] U. Fisher, F. Baum, S. Obenauer, and et al., "Comparative study in patients with microcalcifications: full-field digital mammography vs screen-film mammography," *European Journal of Radiology*, vol. 12, no. 11, pp. 2679–2683, 2002. 16

[9] M. Elter and A. Horsch, "CADx of mammographic masses an clustered microcalcifications: A review," *Medical Physic*, vol. 36, no. 6, pp. 2052 – 2068, 2009. 19, 21, 24

[10] T. Wang and N. Karayiannis, "Detection of microcalcifications in digital mammograms using wavelets," *IEEE Transactions on Medical Imaging*, vol. 17, pp. 498 – 509, August 1998. 19

[11] R. Bird, T. Wallace, and B. Yankaskas, "Analysis of cancers missed at screening mammography," *Radiology*, vol. 184, no. 3, pp. 613 – 617, 1992. 20

[12] K. Kerlikowske, P. Carney, B. Geller, M. Mandelson, S. Taplin, K. Malvin, V. Ernster, N. Urban, G. Cutter, R. Rosenberg, and R. Ballard-Barbash, "Performance of screening mammography among women with and without a first-degree relative with breast cancer," *Annual of International Medicine*, vol. 133, no. 11, pp. 855 – 863, 2000. 20

[13] K. H. Ng and M. Muttarak, "Advances in mammography have improved early detection of breast cancer," *Journal of Hong Kong College of Radiology*, vol. 6, no. 3, pp. 126 – 131, 2003. 20

[14] J. Brown, S. Bryan, and R. Warren, "Mammography screening: An incremental cost effectiveness analysis of double versus single reading of mammograms," *British Medical Journal*, vol. 312, no. 7034, pp. 809 – 812, 1996. 20

[15] R. Warren and S. Duffy, "Comparison of single and double reading of mammograms and change in effectiveness with experience," *British Journal of Radiology*, vol. 68, no. 813, pp. 958 – 962, 1995. 20

[16] R. M. Rangayyan, F. J. Ayres, and J. E. L. Desautels, "A review of computer-aided diagnosis of breast cancer: Towards the detection of early signs," *Journal of Franklin Institute*, vol. 344, no. 3/4, pp. 312 – 348, 2007. 21

[17] M. Giger, "Computer-aided diagnosis of breast lesions in medical images," *Computers in Science and Engineering*, vol. 2, no. 5, pp. 39 – 45, 2000. 21

[18] M. Reddy and R. Given-Wilson, "Screening for breast cancer," *Women's Health Medicine*, vol. 3, no. 1, pp. 22 – 27, 2006. 21

[19] H. D. Cheng, X. Cai, X. Chen, L. Hu, and X. Lou, "Computer-aided detection and classification of microcalcifications in mammograms: A survey," *Pattern Recognition*, vol. 36, no. 12, pp. 2967 – 2997, 2003. 21

[20] L. Zhang, R. Sankar, and W. Qian, "Advances in microcalcification clusters detection in mammography," *Computers in Biology and Medicine*, vol. 32, no. 6, pp. 515 – 528, 2002. 21

[21] S. Yu and L. Guan, "A CAD system for the automatic detection of clustered microcalcifications in digitized mammograms films," *IEEE Transactions on Medical Imaging*, vol. 19, pp. 115 – 126, February 2000. 21

[22] E. Regentova, L. Zhang, J. Zheng, and G. Veni, "Microcalcification detection based on wavelets domain hidden markov tree model: Study for inclusion to computer aided diagnosis prompting system," *Medical Physic*, vol. 34, no. 6, pp. 2206 –2219, 2007. 21

[23] G. Kom, A. Tiedeu, M. Kom, C. Nguemgne, and J. Gonsu, "Détection automatique des opacités dans les mammographies par la méthode de minimisation de la somme de linertie," *Innovation et Technologie en Biologie et Médecine*, vol. 26, no. 5-6, pp. 347 – 356, 2005. 21

[24] G. Kom, A. Tiedeu, and M. Kom, "Automated detection of masses on mammograms by local adaptative thresholding," *Computers in Biology and Medicine*, vol. 37, no. 1, pp. 37 – 48, 2007. 21

[25] M. P. Sampat, M. K. Markey, and A. C. Bovik, *Computer-aided detection and diagnosis in mammography*. New York: Academic: Handbook of Image and Video Processing, 2nd ed., 2005. 21

[26] R. Campanini, D. Dongiovanni, E. Iampieri, N. Lanconelli, M. Masotti, G. Palermo, A. Riccardi, and M. Roffilli, "A novel featureless approach to mass detection in digital mammograms based on support vector machines," *Physics in Medicine and Biology*, vol. 49, no. 6, pp. 916 – 975, 2004. 21

[27] N. H. Eltonsy, G. D. Tourassi, and A. S. Elmaghraby, "A concentric morphology model for the detection of masses in mammography," *IEEE Transactions on Medical Imaging*, vol. 26, pp. 880 – 889, June 2007. 21

[28] L. Hadjiiski, B. Sahiner, H. P. Chan, N. Petrick, M. A. Helvie, and M. Gurcan, "Analysis of temporal changes of mammographic feature: Computer-aided classification of malignant and benign breast masses," *Medical Physic*, vol. 28, no. 11, pp. 2309 – 2317, 2001. 21

[29] S. Timp, C. Varela, and N. Karssemeijer, "Temporal change analysis for characterization of mass lesions in mammography," *IEEE Transactions on Medical Imaging*, vol. 26, pp. 945 – 953, July 2007. 21

[30] S. Timp, S. van Engeland, and N. Karssemeijer, "A regional registration method to find corresponding mass lesions in temporal mammogram pairs," *Medical Physic*, vol. 32, no. 8, pp. 2629 – 2638, 2005. 21

[31] C. Daul, P. Graebling, A. Tiedeu, and D. Wolf., "Three-dimensional-reconstruction of microcalcification clusters using stereo imaging: algorithm and mammographic unit calibration," *IEEE Transactions on Biomedical Engineering*, vol. 52, pp. 2058 – 2073, December 2005. 21

[32] A. Oliver, J. Freixenet, R. Martì, J. Pont, E. Pérez, E. R. E. Denton, and R. Zwiggelaar, "A novel breast tissue density classification methodology," *IEEE Transactions on Information Technology in Biomedicine*, vol. 12, pp. 55 – 64, January 2008. 21

[33] Z. Chen, E. Denton, and R. Zwiggelaar, "Local feature based breast tissue appearance modelling for mammographic risk assessment," *Annals of BMVA*, vol. 2013, no. 3, pp. 1–19, 2013. 21

[34] P. K. Saha, J. K. Udupa, E. F. Conant, D. P. Chakraborty, and D. Sullivan, "Breast tissue density quantification via digitized mammograms," *IEEE Transactions on Medical Imaging*, vol. 20, pp. 792 – 803, August 2001. 21, 27

[35] T. Subashini, V. Ramalingam, and S. Palanivel, "Automated assessment of breast tissue density in digital mammograms," *Computer Vision and Image Understanding*, vol. 114, pp. 33–43, 2010. 21

[36] R. M. Rangayyan, N. M. El-Faramawy, J. E. Desautels, and O. A. Alim, "Measures of acutance and shape for classification of breast tumors," *IEEE Transactions on Medical Imaging*, vol. 16, pp. 799 – 810, June 1997. 21

[37] L. M. Bruce and R. R. Adhami, "Classifying mammographic mass shapes using the wavelet transform modulus-maxima method," *IEEE Transactions on Medical Imaging*, vol. 18, pp. 1170 – 1177, December 1999. 21

[38] A. J. Méndez, P. G. Tahoces, M. J. Lado, M. Souto, and J. J. Vidal, "Computer-aided diagnosis: Automatic detection of malignant masses in digitized mammograms," *Medical Physic*, vol. 25, no. 6, pp. 957 – 964, 1998. 21

[39] S. Pohlman, K. A. Powell, N. A. Obuchowski, W. A. Chilcote, and S. Grundfest-Broniatowski, "Quantitative classification of breast tumors in digitized mammograms," *Medical Physic*, vol. 23, pp. 1337 – 1345, August 1996. 21

[40] R. M. Rangayyan, T. M. Nguyen, F. J. Ayres, and A. K. Nandi, "Effect of pixel resolution on texture features of breast masses in mammograms," *Journal of Digital Imaging*, vol. 23, pp. 547–553, 2010. 23

[41] M. Abdel-Nasser, J. Melendez, A. Moreno, and D. Puig, "The impact of pixel resolution, integration scale, preprocessing, and feature normalization on texture analysis for mass classification in mammograms," *International Journal of Optics*, pp. 1–12, 2016. 23

[42] C. Feudjio, A. Tiedeu, M. Gordan, S. Domngang, and A. Vlaicu, "Computerized detection and smoothing of contour in mammograms," in *11th African Conference on Research in Computer Science and Applied Mathematics*, pp. 299–306, 13-16 October 2012. 27

[43] M. Karnan and K. Thangavel, "Automatic detection of breast border and nipple position on digital mammograms using genetic algorithm for asymmetry approach to detection of microcalcifications," *Computer Methods and Programs in Biomedicine*, vol. 87, pp. 12 – 20, 2007. 27

[44] L. Liu, J. Wang, and T. Wang, "Breast and pectoral muscle contours detection based on goodness of fit measure," in *5th International Conference on Bioinformatics and Biomedical Engineering*, 2011. 27

[45] R. J. Ferrari, R. M. Rangayyan, J. E. L. Desautels, R. A. Borges, and A. F. Frère, "Automatic identification of the pectoral muscle in mammograms," *IEEE Transactions on Medical Imaging*, vol. 23, pp. 232 – 245, February 2004. 27

[46] S. M. Kwok, R. Chandrasekhar, Y. Attikiouzel, and M. T. Rickard, "Automatic pectoral muscle segmentation on mediolateral oblique view mammograms," *IEEE Transactions on Medical Imaging*, vol. 23, pp. 1129 – 1140, September 2004. 27

[47] N. Karssemeijer, "Automated classification of parenchymal patterns in mammograms," *Physics in Medicine and Biology*, vol. 43, pp. 365–378, February 1998. 27

Chapter 3

Background suppression in mammographic images

Contents

3.1 Introduction

Mammographic density is clearly established to provide useful index for breast cancer risk assessment. A neat integration of breast histological information in the CAD process will certainly improve cancer signs detection. To achieve this goal, the first steps to carry out deal with identification and characterization of dense tissues in mammographic images since these latter are known to be common areas for cancer to develop. A prior removal of all non breast patterns is thus required in order to derive useful histological information. This means identifying the breast region in mammograms. However, low contrast quality and slight variation of pixel intensity at breast edge impair for accurate breast region extraction. The situation grows worse for cases of mammograms with noisy background.

As a result, in many cases, accurate breast region cannot be extracted just on a segmentation basis. To cope with all these difficulties, a contrast enhancement is first of all applied prior to the segmentation for a good estimation of the breast region. Then, the true breast boundary points are searched on paths computed on the estimated breast contour. Finally, smoothing the true edge points allows to delineate the breast region accurately.

3.2 Overview of background suppression

A mammogram is an X-ray projection of breast components. Unlike other imaging methods such as Computed Tomography (CT) and MRI, mammograms have an inherent *fuzzy* or diffuse appearance. This is due to the differential attenuation characteristic associated to various breast tissue densities. In addition, mammographic images often contain non-breast patterns in the background such as examination data or noise resulting from the digitization process. Such patterns regularly tamper cancer signs search and consequently have a significant influence on the CAD system performances. It is therefore necessary to handle background objects separately during CAD analysis.

3.2.1 Breast region importance in CAD

The use of CAD system for mammograms analysis is becoming an effective strategy to reduce radiologists' workload and assist them in efficient detection of breast cancer [1]. Identifying breast region or suppressing background in mammograms is of great importance in CADs since the search for cancer signs needs to be carried out only in relevant portions of the image. This implies segmenting and identifying meaningful image regions from irrelevant ones. Such an operation improves CAD systems by reducing false alarm rates, area to analyze for calcification and lesion detection [2] and thus computation time [3].

The precise delineation of the breast region in mammograms is an important step in the process of automatic detection of breast cancer signs. Several important tasks in the computerized analysis of mammograms are strongly dependent on proper and reliable segmentation of the breast region. The most common ones encompass

- label and artifact suppression accounting for false positive diagnoses,
- breast nipple identification which is used as reliable reference point on a mammogram [4],
- pectoral muscle suppression which is depicted with similar texture appearance as dense breast tissue and cancer signs as well [5, 6],
- breast tissue density assessment known to be an index for cancer risk estimation [7, 8],
- registration purpose for multi-view analysis, lesion characterization, and alignment procedure for bilateral comparison of mammograms [9].

Extracting breast region in mammography can be reduced to a simple two-class image region partitioning *i.e* foreground and background. However, most popular segmentation methods and classical edge detection algorithms such as those of Prewitt, Sobel and Roberts which are based on gradient fail to produce a proper extraction of breast region in every situation due to the following grounds [1, 10, 11, 6]:

- decrease of contrast near breast edge caused by lack of compression during acquisition process,
- presence of irrelevant objects in the background such as examination notes, labels or frames which interfere with the segmentation of the breast region,

- noise and non uniform distribution of pixel intensity in the background especially in digitized mammograms.

An accurate breast region segmentation is vital in the framework of mammographic density characterization in the sense that it reduces the bias introduced by breast region under or over estimation.

3.2.2 Related works on breast region extraction

Several approaches for breast region extraction in mammogram have been developed in the literature with various level of success. A good review of approaches introduced in the literature so far to address this issue is presented in [12]. Generally speaking, approaches introduced in the literature to address the problem of breast region identification rely on the following image processing tasks: thresholding [13, 14, 15, 16, 17], morphology operations [18], region growing [11], active contour [19], Gabor filters and edge linking [1] and pixel features [6, 10].

Thresholding based segmentation

Breast region identification has been addressed by some authors as a two-class image partitioning using image thresholding. One of the earliest attempts to suppress background in mammograms was based on a simple setting of threshold [20]. In some similar approaches, a single threshold is automatically computed to partition breast region and background [14, 15]. Others authors, Czaplicka and Wlodarczyk [17] developed a combined strategy using both a global thesholding based on minimization of measures of fuzzyness of a mammogram and Sobel edge detector to estimate breast skin-line. A similar approach was also implemented by Kus and Karagoz [16] in which breast region is first of all roughly estimated through global thresholding. Then, the image was enhanced prior to border pixel extraction with a gradient based algorithm. Observing that the decrease in contrast near breast edge is the main failure cause of thresholding segmentation, Maitra *et al.* [13] first applied an adaptive histogram equalization to enhance image contrast prior to an iterative thresholding based on pixel feature homogeneity. The breast boundary was then delineated with a gradient edge detector algorithm.

Using a single threshold to partition a mammographic image usually results in misclassification of some background pixels as breast region and vice-versa. In opposite, the use of local thresholding [21] has been shown to produce better results since each pixel is thresholded using local information of its neighborhood. However such an approach is time consuming.

Pixel features based segmentation

One inherent limitation with thresholding techniques is their sensitivity to noise and contrast variation. On the other hand, gradient edge detector are subject to inaccurate results in unsharp intensity transition and noisy edge as it is usually the case at the vicinity of the breast boundary.

To cope with these limitations, some approaches make use of local information in conjunction with gray level features. In this way, Boss *et al.* [18] developed a two-step

strategy by first applying gray level normalization and noise reduction with a median filter to subsequently segment breast region by using connected component labeling. Chen *et al.* [6] used local features such as local binary pattern (LBP), histogram of gradient (HoG) and intensity histogram and a support vector machine (SVM) classifier to build feature vector models. These were later fed in a neural machine of two-layer committee to vote for which class label of mammogram regions is associated to each pixel in the image. Touil *et al.* [10] developed an iterative fuzzy segmentation combined with some local pixel feature information to gradually integrate the area in the vicinity of breast edge that is under-segmented with a classical fuzzy C-means segmentation technique. However, selecting and modeling ideal features for classification is difficult especially when a huge amount of data is available. A set of features who produces accurate result with a given image may fail on another image due to variation in texture, contrast or brightness.

texture-based segmentation

Some texture-based approaches have also been used to tackle the difficulties faced in breast region extraction in mammograms. For instance, Casti *et al.* [1] used a multidirectional Gabor filtering to extract edge feature in image and then apply an edge linking algorithm to extract breast skin-line. However, to be successful, they first applied a logarithmic transformation to the image to improve the contrast in breast boundary area.

Ojala and Liang [22] described a semi automated method where a boundary traced interactively was used to initialize a snake algorithm. Later, they introduced an active contour method for extracting and smoothing breast contour in mammograms [23]. Wirth and Stapinski [24] and McLoughlin and Bones [25] implemented the active contours to identify breast region in mammograms. They first of all performed a global threshold and then modeled the background noise using Poisson approximation. The binary mask obtained from global thresholding was used as a seed for the snake algorithm. Ferrari *et al.* [19] also used an enhancement technique prior to use of an active contour model for identifying breast boundary. An approximate mask of breast region is first of all obtained using Lloyd-Max quantizer. The contour extracted on this mask is used as seed in an adapted active contour model to produce the final breast boundary. Likewise, Chen and Zwiggelaar [11] roughly estimated breast region using global thresholding, then, with orthogonal lines on the extracted contour, they estimated the true edge points used as seed in an active contour model to refine breast boundary. The major inconvenience with active contour is that they are initialization dependent and, in some cases, proper convergence may be computationally demanding.

No algorithm can be considered 100% robust, especially due to the heterogeneous nature of mammograms. Problems with mammogram acquisition such as digitization, induced artifacts, excessive background noise, scratches and dust artifacts could all influence the reliability of any algorithm [24]. Background suppression in mammograms thus remains an open problem in CAD design. In short, accurate extraction of breast region in mammographic images cannot be obtained using a straightforward segmentation. Generally, contrast enhancement is applied for accurate breast boundary extraction or a multi-stage segmentation performed with some fine tuned parameters to address these difficulties at various levels. Surprisingly, it appears that to address this problem, an appropriate enhancement combined with a classical segmentation

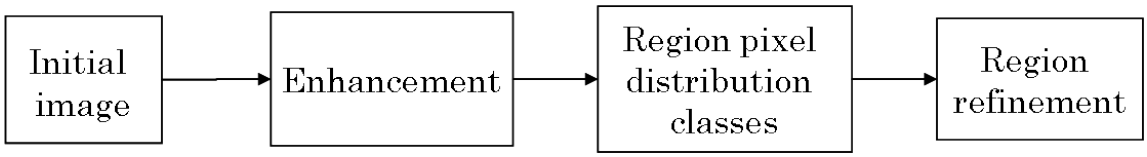


Figure 3.1: Flowchart of the strategy to perform for background suppression in mammograms

methods can produce not only proper results for CAD systems but will also reduce its computational burden.

3.2.3 Proposed approach for background suppression

From related works, one can notice that various methods have been implemented to suppress background in mammograms with different levels of success. For most of those approaches, the tricky issue is the trade off between accuracy and computation cost. Elaborating a method to successfully process all peculiarities found in background of mammograms is hard to achieve without increasing computational burdens for a CAD system. In this work, an effort was done to develop an approach less coercive for CAD systems in terms of accuracy and computation time. To achieve such a result, rather than working out a new image processing method, we make use of simple mathematical tools to accurately address this problem.

The following considerations highlight the main actions carried out in this work to suppress the background in mammographic images.

1. The primary difficulty to overcome in suppressing background in mammograms is related to their low contrast quality especially at the breast edge which does not facilitate boundary detection. To cope with this inconvenience, a contrast enhancement is applied to stretch the dynamic range in dark region and facilitate boundary identification.

2. Basically, mammograms are assumed to be made of two classes of objects: patient tissue region (foreground) and background. Hopefully, class conditional empirical distributions of grey level values are well separated, meaning that gray level values are informative enough to discriminate the two image regions. A global thresholding and a fuzzy C-means clustering segmentation method were implemented and relevant results are obtained when these latter are fitted together with an adequate contrast enhancement.

3. In general, the segmentation produces good results. However, because of noise in the background coupled to the smooth variation of gray level intensities along breast edge, some of these pixels are incorrectly labeled as background pixels. To cope with this inconvenience, a refinement step is performed to estimate the true boundary point.

The block diagram of figure 3.1 summarizes the main steps here above introduced to suppress background in mammographic images.

3.3 Mammographic contrast enhancement

In the remainder of this chapter, some image processing developments are presented. We therefore introduce some notations that will be used throughout the next sections. A (mammographic) image is a mapping $u : \mathcal{G} \longrightarrow \{0, .., L-1\}$ with $\mathcal{G}$ a domain often called image grid and L an integer which stands for the number of possible gray levels. An element $\mathbf{p} \in \mathcal{G}$ is called a pixel. A pixel $\mathbf{p}$ is a pair of coordinates (p_x, p_y) on the image grid. Gray level pixel intensities are thus given by $u(\mathbf{p}) = u(p_x, p_y)$. Black pixels are such that $u(\mathbf{p}) = 0$ while white pixels are such that $u(\mathbf{p}) = L-1$.

Background suppression (as far as mammograms are concerned) is an operation consisting in the removal of non breast entities contained in the background of a mammographic image. Ideally, this task is easy to achieve if class conditional empirical pixel value distributions are non-overlapping. However, this is not always the case with mammographic images where contrast is poor especially across the breast edge.

The compression applied during mammogram acquisition aims at reducing X-ray attenuation across similar tissues, thus producing a uniform contrast for a given tissue density. However, non uniform compression at breast edge produces a decrease of contrast on the mammogram. This fading effect diminishes the visibility along the peripheral region of the breast making it difficult to perceive the breast boundary and identify the nipple position. To ensure a good visualization of boundaries and improve segmentation results, image contrast enhancement is required.

Image enhancement techniques are used to improve contrast objectively if it produces an increase of the signal-to-noise ratio or subjectively if it makes certain features easier to perceive by modifying gray level intensities. There is a wide range of contrast enhancement techniques [26] the majority of which are obtained by composition with a relevant function $f : \{0, .., L-1\} \longrightarrow \{0, .., L-1\}$. If v denotes the enhanced image, we have $v = f \circ u$. The most widely used functions f for such a purpose are:

- Linear functions that enhance the dynamic range by stretching the original gray level range. These functions are defined as:

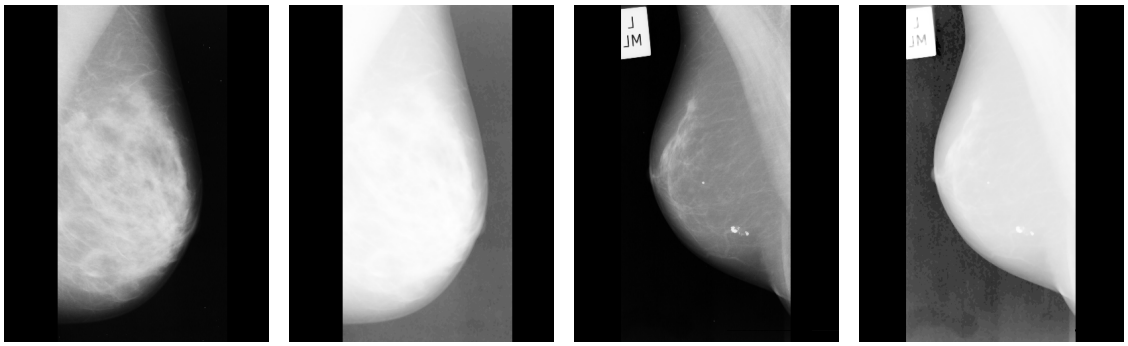
  $f(x) = ax + b$, where a and b are real constants. Note that a rounding function must also be applied because f may not map x directly to $\{0, .., L-1\}$. This is also true for other types of functions and we do not explicitly mention it for brevity.

- Piece-wise linear functions that apply various stretching in K desired intervals of gray level ranges. These functions are defined by:

  $f(x) = a_k x + b_k$, for $l_k \leq x \leq l_{k+1}$, with $l_k \in \{0, .., L-1\}$ and $k = 0, 1, \ldots, L-1$. Note that coefficients a_k and b_k are generally chosen so that f is monotonic.

- Parametric families of non-linear functions subject to:
  $\min_{0 \leq l \leq L-1} f(l) = 0$ and $\max_{0 \leq l \leq L-1} f(l) = L-1$.
  Example of popular non-linear function families are:

  - Logarithmic functions which stretches dark regions and suppress bright ones. They are defined as:
    $f(x) = b \log(ax + 1)$, where a and b are real constants.



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

Figure 3.2: Two examples of mammograms with shaperned breast contours. (a) and (c) Initial mammograms, (b) and (d) mammograms after applying a logarithmic contrast correction.

  - Exponential functions which expands bright regions and are defined as: $f(x) = b(e^{ax} - 1)$, where a and b are real constants.
  - Gamma correction functions which are power transforms defined as: $f(x) = a(x^{\gamma} + b)$, where a, b and γ are real constants.

- Histogram equalization functions that map an image with an arbitrary histogram to one with a flat histogram. In this case f is dependent on u. A solution to this problem is to choose $f = F_u$ where F_u is the cumulative empirical distribution of gray levels in u. Note that this operation can be performed locally on smaller image regions. Indeed image region histograms are very often significantly different from the global image histogram.

In the following, we are interested in methods that help to sharpen gray level transitions in mammographic images especially at breast edge pixels. The logarithmic transformation appears to be the best suited one for this task as it stretches dark regions.

The image contrast is therefore significantly enhanced by applying the logarithmic operation as defined in [26, 27]. The contrast-correction is performed on an initial image u to produce its logarithmic version defined by the equation:

$$v_{log} = \frac{L-1}{\max\limits_{\mathbf{p}\in\mathcal{G}} \log\left(u\left(\mathbf{p}\right)+1\right)} \times \log(u+1) \tag{3.1}$$

Note that the intensity levels of the enhanced image v_{log} are normalized so that their values belong to the interval $[0, L-1]$. Figure 3.2 shows the effect of the logarithmic enhancement technique on some mammographic images. This dynamic range expansion is applied to the whole image and significantly improves contrast especially near the breast edge and therefore facilitates breast boundary identification (see figure 3.2(b) and (d)). This technique also unfortunately amplifies background noise which was not visible on the initial image. The main idea underlying this enhancement step is to enable breast region extraction to be as close as possible to the true breast boundary. The next sections introduce the methods developed to segment breast regions in mammograms.

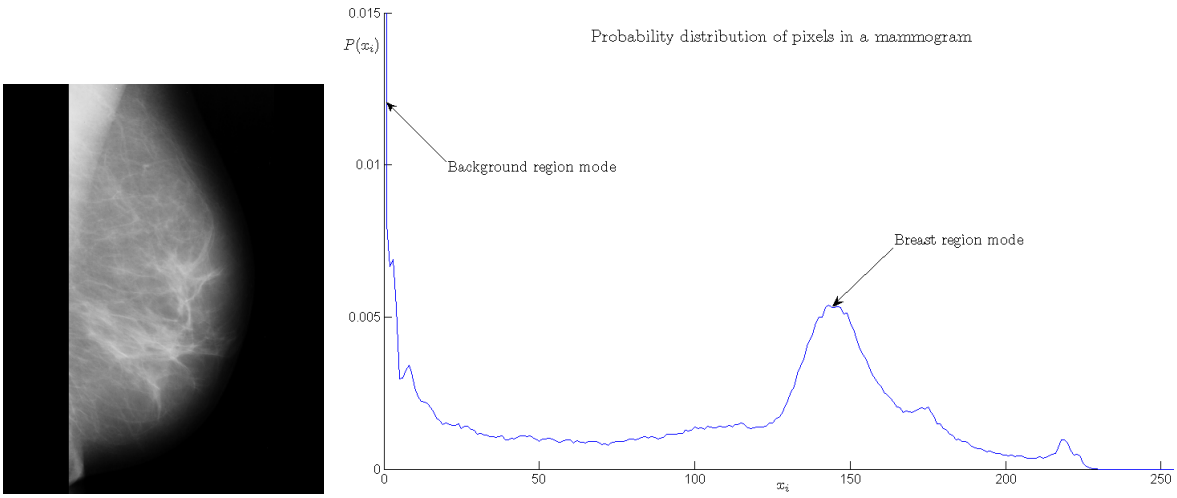


Figure 3.3: A mammographic image along with its corresponding histogram expressed as probability distribution. The two modes identifiable in the histogram represent the breast region and the background in the image.

3.4 Mammogram background segmentation

3.4.1 Global threshold method

The most famous global thresholding technique is that of Otsu [28] which sets a single threshold value to partition an image into two classes. Each pixel of the image is assigned to a class with respect its gray level intensity value and the threshold value.

Definition

A histogram determines the occurrences of a gray level value l_k in an image. If a histogram is normalized so that it sums to one, it is also the empirical probability distribution p_u of gray level values in the image u.

Figure 3.3 shows a mammographic image along with its corresponding histogram. The modes observed on the histogram are commonly belonging to the same image region which correspond to a given object present in the image. Global thresholding techniques attempt to make use of information from histograms to compute the optimal threshold that would provide a good partition of the image. Otsu's method [28] performs a minimization of the within group variance to compute the optimal threshold defining the frontier between the two modes of the histogram.

Otsu's threshold calculation

Assume that the image contains only two objects corresponding to two pixel classes: foreground and background. Let us also suppose that class conditional distributions $p_{u|f}$ and $p_{u|b}$ are non overlapping, *i.e.* there exists a threshold value $\tau \in \{1, .., L-2\}$ such that $p_{u|b}(l > \tau) = p_{u|f}(l \leq \tau) = 0$. Under these assumptions, we have

$$p_{u|b}(l) = \frac{p_u(l)}{F_u(\tau)}, \tag{3.2}$$

$$p_{u|f}(l) = \frac{p_u(l)}{1 - F_u(\tau)}. \tag{3.3}$$

consequently, the mean intensities in the foreground and in the background are respectively defined as follows:

$$M_{b,\tau} = \sum_{l=0}^{\tau} l\, p_{u|b}(l), \tag{3.4}$$

$$M_{f,\tau} = \sum_{l=\tau+1}^{L-1} l\, p_{u|f}(l). \tag{3.5}$$

The standard deviations of each class are denoted by $\sigma_{b,\tau}$ and $\sigma_{f,\tau}$ respectively are given by the following equation:

$$\sigma^2_{b,\tau} = \sum_{l=0}^{\tau} (l - M_{b,\tau})^2 p_{u|b}(l), \tag{3.6}$$

$$\sigma^2_{f,\tau} = \sum_{l=\tau+1}^{L-1} (l - M_{f,\tau})^2 p_{u|f}(l). \tag{3.7}$$

Finally, the within-group variance is expressed as follows:

$$\sigma^2_{w,\tau} = F_u(\tau)\, \sigma^2_{b,\tau} + (1 - F_u(\tau))\, \sigma^2_{f,\tau}. \tag{3.8}$$

Global segmentation

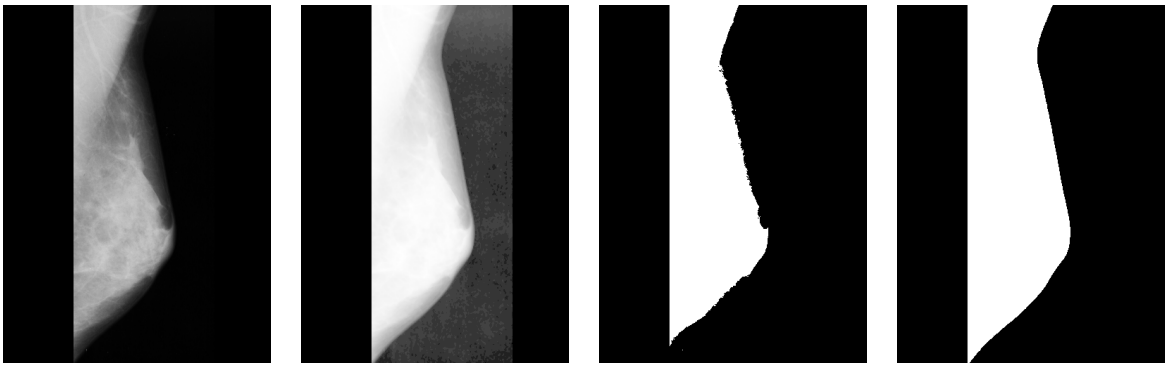
A global thresholding technique computes a single threshold τ to which all pixels contained in the image are compared. Using Otsu's method [28], the best threshold τ is obtained by minimizing the within group variance $\sigma^2_{w,\tau}$. In our setting, once the value of τ is found, the segmented image v_{seg} is evaluated at a given pixel (p_x, p_y) by the following equation:

$$v_{seg}(\mathbf{p}) = \begin{cases} 0 & \text{if } v_{log}(\mathbf{p}) \leq \tau \\ 1 & \text{otherwise} \end{cases}. \tag{3.9}$$

Image v_{seg} is often referred to as the breast region binary mask. Some breast region segmentation results using Otsu's global thresholding method are presented on figure 3.4. It can be noticed that the segmentation performed on the initial mammogram results in incorrect assignments for pixels located in the vicinity of the breast edge. The low contrast on this area makes it difficult to accurately partition pixels. On the other hand, the segmentation performed on the logarithmic version of the mammogram produces better results thanks to the adjustment of contrast around the breast edge.

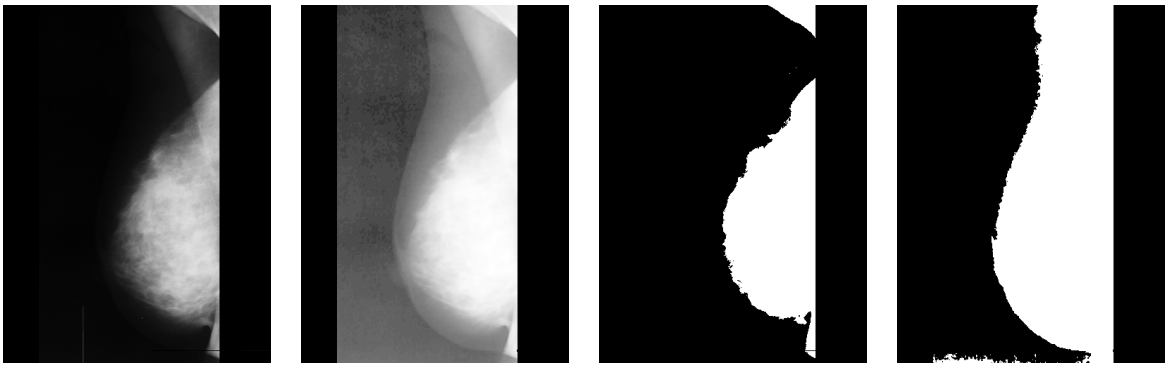
Figure 3.5 shows another segmentation of a poorly contrasted mammogram. In this mammographic image, the breast is made of dense tissues and some noise in the background can be observed after logarithmic contrast correction. Segmenting the initial image with Otsu's technique produces this time an inconsistent breast region. Again, applying the same technique after contrast correction produces a better segmentation result thereby proving that the logarithmic contrast correction increases the robustness of the segmentation. However, even the result displayed in figure 3.5 (d) is not completely satisfying.

Indeed, a major inconveniency of screen film mammograms is that there is a stronger noise in the background which usually comes from the digitization process. Such a noise is illustrated in figures 3.4 (b) and 3.5 (b) by the shaded areas in the background. This



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

Figure 3.4: Segmentation of breast region in mammogram using a global thresholding. (a) Initial image, (b) logarithmic enhanced image, (c) initial image segmentation, (d) Enhanced image segmentation



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

Figure 3.5: Breast region segmentation in a noisy mammogram. (a) Initial image, (b) logarithmic enhanced image, (c) initial image segmentation, (d) enhanced image segmentation

noise is too strong in figure 3.5 (b) and accounts for to the inaccurate segmentation displayed in 3.5 (d), where the detected breast has irrelevant fluctuations.

From the histogram presented on figure 3.3 (b), it can be noticed that there is no clear cut separation between the two modes of the histogram. In such conditions, Otsu's technique (and more generally global thresholding methods) cannot perform an accurate segmentation of objects. On the other hand, the Fuzzy C-Means (FCM) algorithm has been proven to be robust in clustering overlapping data. The next section therefore investigates an alternative segmentation method based on the FCM algorithm.

3.4.2 Fuzzy C-Means (FCM) method

The FCM algorithm is an unsupervised data clustering technique commonly used in image processing for segmentation tasks. This is an appropriate method for clustering overlapping data.

Definition

For each pixel class (background and foreground), one membership function is defined over $\mathcal{G}$. The value of the membership function noted $\mu_i(\mathbf{p})$ depicts the possibility for a pixel $\mathbf{p}$ to belong to the i^{th} region. Membership functions are given by:

$$\mu_i(\mathbf{p}) = \left(\sum_{k=1}^{c} \left(\frac{d(u(\mathbf{p}), \nu_i)}{d(u(\mathbf{p}), \nu_k)} \right)^{\frac{2}{m-1}} \right)^{-1} \tag{3.10}$$

with $m > 1$ a fuzzification parameter, d the Euclidean distance, $c = 2$ the number of classes, and ν_i is the centroid of the i^{th} class which is given by:

$$\nu_i = \frac{\sum_{\mathbf{p}\in\mathcal{G}} \mu_i(\mathbf{p})^m u(\mathbf{p})}{\sum_{\mathbf{p}\in\mathcal{G}} \mu_i(\mathbf{p})^m}. \tag{3.11}$$

The FCM algorithm main principle consists in minimizing the inter-class distance through an objective function J_m defined by:

$$J_m(\nu_1, \nu_2) = \sum_{\mathbf{p}\in\mathcal{G}} \mu_1(\mathbf{p})^m d^2(u(\mathbf{p}), v_1) + \mu_2(\mathbf{p})^m d^2(u(\mathbf{p}), v_2). \tag{3.12}$$

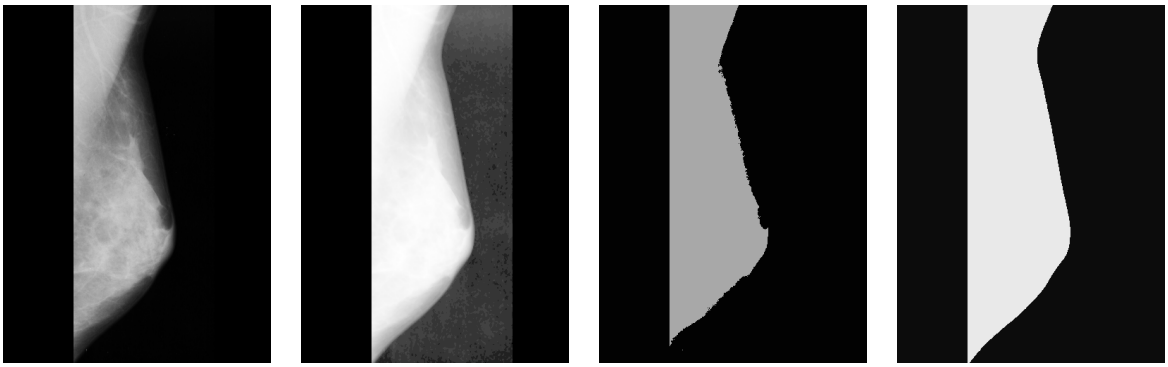
Moreover, the membership functions are subject to the following constraint :
$\forall \mathbf{p}, \sum_{i=1}^{c} \mu_i(\mathbf{p}) = 1$.
Algorithm 1 summarizes the FCM procedure which converges iteratively to centroid values minimizing J_m. The membership function $\mu_i^{(k)}$ and the centroid $\nu_i^{(k)}$ are those obtained at the k^{th} iteration.

Algorithm 1 FCM Algorithm

Require: $c = 2$, $m = 2$, $\epsilon = 0.01$
  $k \leftarrow 1$
  $\boldsymbol{\nu}^{(0)} = \left\{\nu_1^{(0)}, ..., \nu_c^{(0)}\right\} \leftarrow$ Randomly select region modes
  for i from 1 to c **do**
    $\mu_i^{(0)}(\mathbf{p}) \leftarrow$ Evaluate membership function with equation (3.10)
  end for
  while $\|\boldsymbol{\nu}^{(k+1)} - \boldsymbol{\nu}^{(k)}\| > \epsilon$ **do**
    for i from 1 to c **do**
      $\nu_i^{(k+1)} \leftarrow$ Update regions mode with equation (3.11)
      for i from 1 to c **do**
        $\mu_i^{(k+1)}(\mathbf{p}) \leftarrow$ Update membership function with equation (3.10)
      end for
    end for
    $k \leftarrow k + 1$
  end while

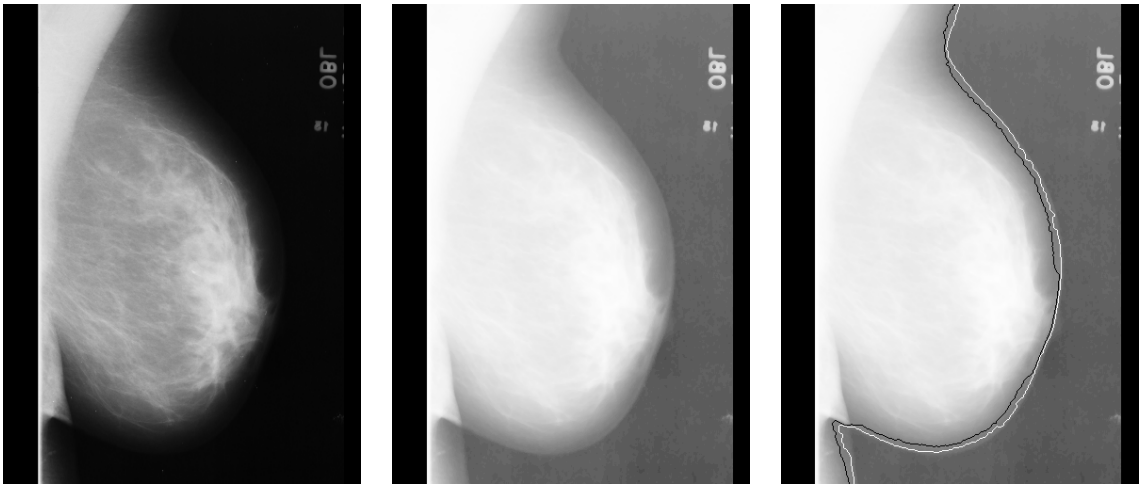
FCM segmentation

Once the FCM algorithm has converged, the image partitioning is performed by assigning each pixel to the class of highest membership function value. Figure 3.6 shows some segmentation results of foreground/background regions in mammograms using



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

Figure 3.6: Segmentation of breast region in mammogram using the FCM algorithm. (a) Initial image, (b) logarithmic enhanced image, (c) initial image segmentation, (d) Enhanced image segmentation



(a) (b) (c)

Figure 3.7: Comparison of breast region segmentation using the FCM algorithm and global thresholding method (a) Initial image, (b) logarithmic enhanced image, (c) enhanced image super-impressed with the breast contour extracted using FCM (white) and global threshold (black) segmentation respectively.

the FCM algorithm. Similarly as for Otsu's segmentation, one can notice that the contrast corrected image produces better segmentation results than the original image.

Figure 3.7 shows a comparison of segmentation results between FCM and global thresholding method. One can see that FCM performs better than the thresholding method in segmenting the background region of this noisy mammogram. More precisely, the extracted contour and in particular, the skin-air interface curve obtained from FCM method is smoother than that of the thresholding method. This shows that FCM is robust to overlapping data which in this case corresponds to breast edge pixels with similar gray level intensities as those of noisy background pixels.

In most cases, background region segmentation results are accurate when FCM is used. However, for mammograms made of dense tissues and with strong noise in the background, the results are not always as accurate. Unlike Otsu's thresholding based segmentation, the contour is smooth but its localization is shifted. Figure 3.8 shows an

(a) (b) (c)

Figure 3.8: Inaccurate foreground region segmentation using FCM method (a) Image segmentation result: breast and muscle tissue region and examination notes are assigned to the foreground class, (b) foreground region obtained after suppression of non breast entities in the image, (c) Enhanced image super-impressed with the skin-air interface obtained from segmentation.

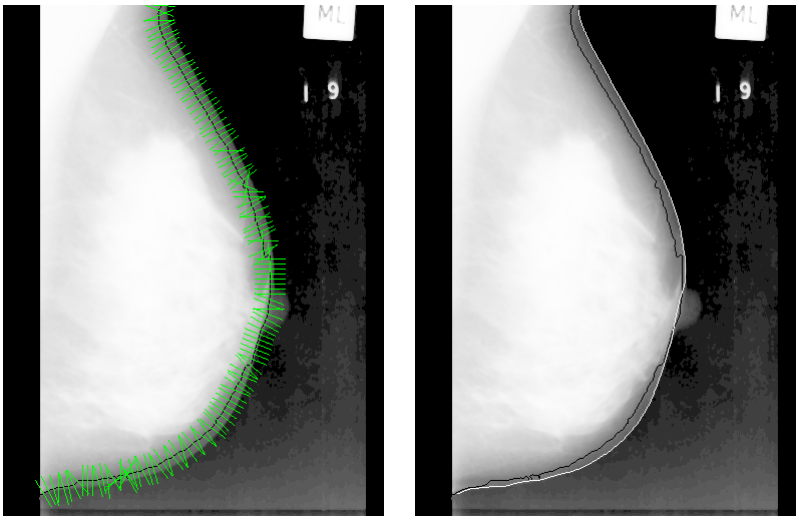
example of a mammogram where such a problem is encountered. This drawback is due to presence of dense tissues who tend to shift the cluster center of the foreground region towards an excessively high value. As a result, pixels with low gray level intensities around breast edge are treated as background pixels during the segmentation process.

The results yielded by the FCM based segmentation are good enough if the goal is just to reduce the area of search for breast cancer signs. Obviously, early signs of breast cancer are hardly found at the breast edge and this latter is exclusively made of fatty tissues. On the other hand, if the motive is to derive histological information, then an inaccurate breast region estimation may significantly bias histological data. The next section therefore introduces a post-processing strategy for an increased background region segmentation accuracy.

3.5 Accurate breast edge estimation

3.5.1 Artifacts removal from the foreground region mask

Figure 3.8(a) shows the segmentation result of a typical mammogram having some examination notes in the background. One can see that these artifacts have been classified into the foreground class. In addition, some isolated pixels at the vicinity of the breast contour are sporadically also classified as foreground class members. To clean up segmentation results, a simple morphological processing is performed. An opening followed by a closing operation are applied to the mask to remove thin objects while preventing large objects from area modifications. To select definitively the foreground region, all objects whose surface is less than that of the larger object are considered as background patterns and suppressed. Figure 3.8(b) shows an image of the binary mask of foreground region with the background suppressed.



(a) (b)

Figure 3.9: Accurate breast contour extraction using search path technique. (a) Enhanced image super-impressed with the contour extracted from segmentation and search path segment represented only at every 5 points of the contour. (b) Enhanced image super-impressed with estimated contour (in black) and final contour (in white)

As mentioned previously, noisy mammograms often result in inaccurate breast region extraction. Next subsections introduce the strategy developed to address this other issue.

3.5.2 Estimating true breast edge point

According to the contour obtained from foreground segmentation displayed in figure 3.8(b), the foreground region is underestimated. The approach developed in this work is meant to recover the true breast boundary from the shifted one that we obtained earlier. Following the idea introduced in [29, 30], this approach relies on the definition of a search path for each point of the estimated contour. A search path is defined as the orthogonal line segment centered on its corresponding contour point. In most of the cases, the contour obtained from the segmentation step is close to the breast edge. A search path of 15 pixels length is sufficient to cover the true boundary.

Assuming that the correct breast contour point belongs to the search path, our problem consist now in finding out which one it is for each path. Figure 3.9(a) shows search paths represented only at every 5 points of the contour for visualization purpose. Due to noise in the background, searching the true breast edge points inside a path using only gray level information as done in [29, 30] is computationally exhausting as pixels intensity varies slowly in noisy area. We propose to use pixel gradients in the path direction to search the true edge pixel because the gradient is likely to be maximal at this pixel. More comments on this are given in the sequel.

3.5.3 Contour refinement

The edge point of the breast contour is the one whose gradient value is maximal among pixels in the search path. The gradient $G(\mathbf{p})$ is computed as the norm of the first derivatives of the image u in the xy-direction as:

$G(\mathbf{p}) = \|u'(\mathbf{p})\|$ with $u'(\mathbf{p}) = \left(\frac{\partial u(\mathbf{p})}{\partial p_x}, \frac{\partial u(\mathbf{p})}{\partial p_y}\right)$
where:

- the derivative in the x-direction $\frac{\partial u(\mathbf{p})}{\partial p_x} = u(p_{x+1}, p_y) - u(p_x, p_y)$,
- the derivative in the y-direction $\frac{\partial u(\mathbf{p})}{\partial p_y} = u(p_x, p_{y+1}) - u(p_x, p_y)$,

It was observed that iteratively shortening the search path during boundary refinement as described in [31, 30] is time consuming and not necessary in the case of breast contour refinement.

Our assumption is that small search paths contain the true edge points because the segmentation result is already quite good. Yet, this assumption may occasionally be untrue, especially in noisy mammograms or when the breast contour curvature is too high. To cope with this difficulty, the contour refinement step is iterated to make sure that each edge point derived from a search path stabilizes at a given pixel location. From our observations, only two iterations are necessary to converge to true breast edge points.

After refining the contour, a set of points detected as true breast edge pixels are found. The contour points are smoothed within a sliding window to remove outliers. The smoothing method applied is based on locally weighted least square fitting method. Figure 3.9(b) show the final breast contour obtained after refinement and smoothing.

3.6 Performances metrics and evaluation

3.6.1 Dataset

The methods implemented to identify background region in mammograms were tested on images selected from the mini-MIAS database [32] which is made of 322 images. This open database was chosen because it is the most widely used in the literature dealing with background region extraction. It therefore allows an easy comparison of our results with those of other methods. All images in this database are MLO view mammograms digitized at 200 μm and 8-depth resolution. The size of these images is 1024×1024 pixels. Images were further sub-sampled to a 512×512 resolution.

Among the images of the database, 14 are poorly contrasted with a strong noise in the background. Those images were discarded in the segmentation evaluation process because such mammograms do not meet our 2 gray level class assumption. These images have generally three classes. They can be handled efficiently with minor adaptations of our approach as discussed in subsection 3.6.6.

On other hand, 12 images have artifacts (band tape) but were used without any preprocessing since those artifacts are not harmful for background region extraction. However, these latter might be the source of difficulties to handle in other steps of the CAD process. All in all, only 14 images of the database were discarded in the following evaluation study.

3.6.2 Radiologist's study

To evaluate the performances of the background region extraction, a manually delineated background region by an expert radiologist was used as reference standard. The radiologist manually drew the breast contour of all mammograms in the database. It should be noticed that the logarithm enhancement was applied to the images to facilitate the radiologist's identification of the breast boundary in the images. The coordinates of the radiologist's hand-drawn boundaries of background regions were used to compute several performance metrics in order to enable a fair comparison of our results with those of related works. The performance metrics are defined on several criteria as introduced in the next section.

3.6.3 Performance metrics

The accuracy of the background region extraction is assessed through the following performance ratings:

- False Positive (FP) and False Negative (FN) rates, which evaluate pixel assignment errors. A FP pixel is one assigned by the algorithm as belonging to the foreground region but assigned by the radiologist outside of the foreground region. A FN pixel is one assigned outside of the foreground region by the algorithm but assigned inside by the radiologist. The FP and FN rates for an image are computed as follows:

$$FP = \frac{\mid A_C \cup A_R \mid - \mid A_R \mid}{\mid A_R \mid}$$

$$FN = \frac{\mid A_C \cup A_R \mid - \mid A_C \mid}{\mid A_R \mid}$$

  where A_C and A_R are the areas of the foreground regions obtained by the algorithm and the radiologist respectively.

- Percent Overlap Area (POA), which expresses the accuracy of the region delineated by the algorithm to the reference one drawn by the radiologist. It is defined as:

$$POA = \frac{|A_C \cap A_R|}{|A_C \cup A_R|}$$

- The Hausdorff distance d_H, which is a metric assessing the dissimilarity of the detected boundary by the algorithm $\mathcal{C}_{\text{algo}}$ with the one drawn by the radiologist $\mathcal{C}_r$. It is defined as:

$$d_H\left(\mathcal{C}_{\text{algo}}, \mathcal{C}_r\right) = \max\{\max_{c_i \in \mathcal{C}_{\text{algo}}}\{\min_{r_j \in \mathcal{C}_r}\{d(c_i, r_j)\}\}; \max_{r_j \in \mathcal{C}_r}\{\min_{c_i \in \mathcal{C}_{\text{algo}}}\{d(r_j, c_i)\}\}\} \quad (3.13)$$

  where c_i and r_j are contour points obtained by the algorithm and the radiologist respectively, $d(c_i, r_j)$ is the Euclidean distance between points c_i and r_j.

These metrics are widely used and thereby allow a fair comparison with other works found in the literature. These metrics are also complementary and can thus catch every aspect of image segmentation performances.

The next sections present the results and performances of breast region extraction implemented in this work. It also carries out a comparison and discussion of these results with related works.

3.6.4 Breast region extraction results

Figure 3.10 shows some background region extractions obtained with the implemented methods in this work. In comparison with radiologist's manually drawn breast region boundaries, the following performances were obtained:

- The POA mean and the standard deviation are 96.72±2.28% and 96.72±2.37% using FCM method and global thresholding method respectively.
- The FN mean and the standard deviation are 2.63 ± 2.27% and 2.71 ± 2.39%, while the FP mean and the standard deviation are 0.68 ± 1.09% and 0.59 ± 0.89% for FCM method and global thresholding method respectively.
- The Hausdorff distance mean and the standard deviation are 19.12±18.77 pixels and 19.32±22.22 pixels for FCM method and global thresholding method respectively.

From these results, it can be noticed that both methods perform almost identically in extracting background regions in mammograms. Highly accurate results are achieved thanks to the boundary refinement step. It is difficult to decide on which method outperforms the other, but FCM seems to produce slightly more accurate results.

To assess the impact of the refinement on the breast region detected, a quantitative analysis was conducted. Figure 3.11 shows the evolution of the performances thanks to boundary refinement with respect to each method. The histogram giving the performances before and after refinement process clearly shows that the accuracy of the methods implemented has been significantly improved. For instance, the number of images of high accuracy (POA>95%) has raised from 188 images for both methods to 248 images and 247 images for global threshold and FCM respectively.

However, it can be noticed that the number images of highest accuracy (POA>99%) has decreased for each methods. This is a drawback of the refinement step which tends to reject noisy pixels around nipple when this latter is in the breast profile. The reason is that nipple is represented by pixels of low intensity, thus low gradient values and situated in the neighborhood of some dense tissues such as ductal tissues of high gray level intensities and gradient values. As higher gradient pixels are those lying at the frontier between nipple and ductal tissues, they are selected as true breast edge points and thus rejecting the nipple (see figure 3.9(b)). The non detection of the nipple region was the most common cause of FN pixels. In addition, it was also observed that the smoothing was one of the main cause of high FP rates. Similar observations have also been reported by Ferrari *et al.* [19].

The main asset accounting for the methods implemented to extract background region in mammograms lies in their ability to deal efficiently with low noise corrupted images.

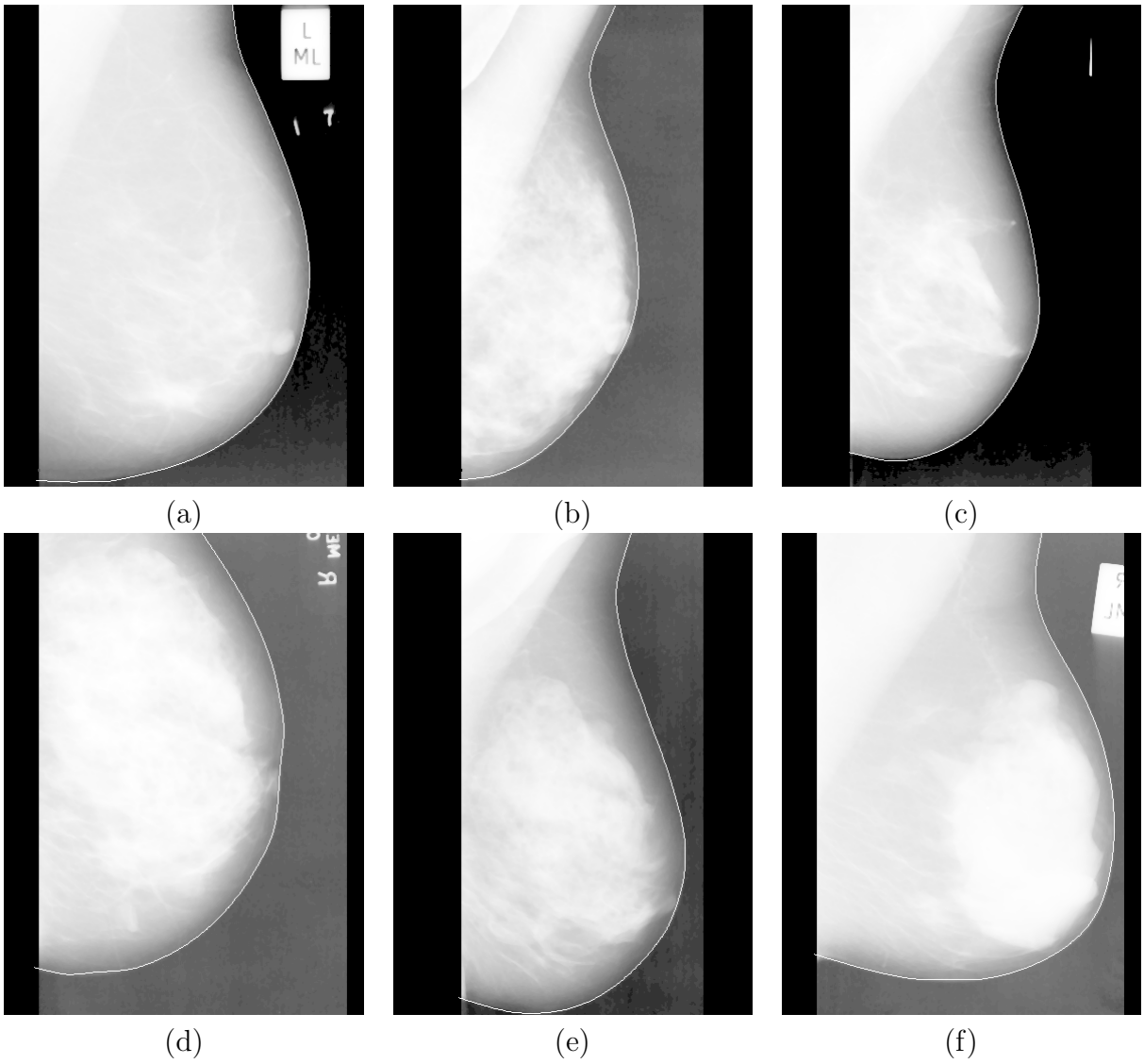


Figure 3.10: Some examples of background region boundaries or skin-air interface contour extracted by the implemented approach and superimposed on the enhanced mammograms. (a) mdb083, (b) mdb123, (c) mdb127, (d) mdb236, (e) mdb286, (f) mdb320

This performance is also noticeable in figure 3.11 as a POA lower than 90% is reported for less than 5 images. In addition, it is should be noted that the lowest POA value recorded in both approaches was greater than 87%. These results globally show that the breast region is accurately identified in the mammograms by the proposed approach.

3.6.5 Discussion

Although, background region extraction in mammograms is a ground work for CAD system, in a number of cases, the background region is only roughly estimated. This shows that accurate background region extraction is difficult to achieve. However, to derive useful histological insights and improve CADs, an accurate background region extraction is required. No algorithm can be considered 100% robust, especially due to the heterogeneous nature of mammograms. Problems with image acquisition such as scanner-induced artifacts, excessive background noise, scratches and dust artifacts influence the reliability of this type of algorithms [24].

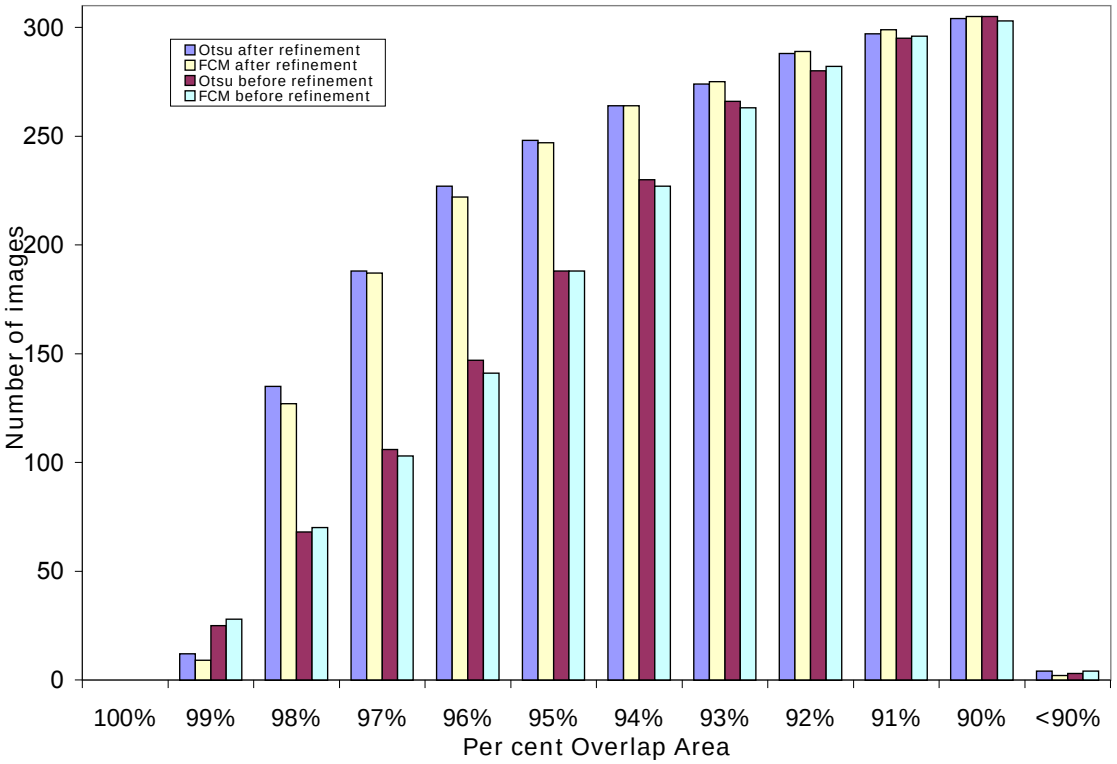


Figure 3.11: Histogram showing the number of images for corresponding percent overlap area between the computed and the reference standard background region before and after boundary refinement for the FCM method and the global thresholding method respectively.

To be coherent in performance comparison, only related works tested on images of the same database are taken into account for this study. Nevertheless, it is still difficult to perform a straight comparison because many authors expressed their results using different performance metrics with or without regard to a reference standard. For instance, although Nagi *et al.* [33] used a reference standard in their work but they did not provide quantitative measure of the performances of their approach. Other authors [34, 21] provided their results based on only visual assessment in comparison to other methods. In addition, all these approaches were tested on small data sets usually containing less than 100 images.

Other authors [35, 3, 24] tested their method on a set of images and reported their results using metrics such as accuracy, correctness and quality computed from FN, FP and True positive (TP). Only the metric "quality" meets the definition of POA. Despite other metrics are computed using FP and FN values, they do not provide useful insights when the values of FP and FN rates are unknown therefore making any comparison difficult. The first group of authors reported an accuracy of 96.0% on a set of 120 images. The second group used those metrics to demonstrate the robustness of their approach but they did not report performances on the entire data set. The last group reported 0.98 of correctness and 0.99 of accuracy on a data set of 120 images. Considering these performances, ones can notice that the performances of the proposed method are in the same range or better. In addition, it was almost tested on the entire

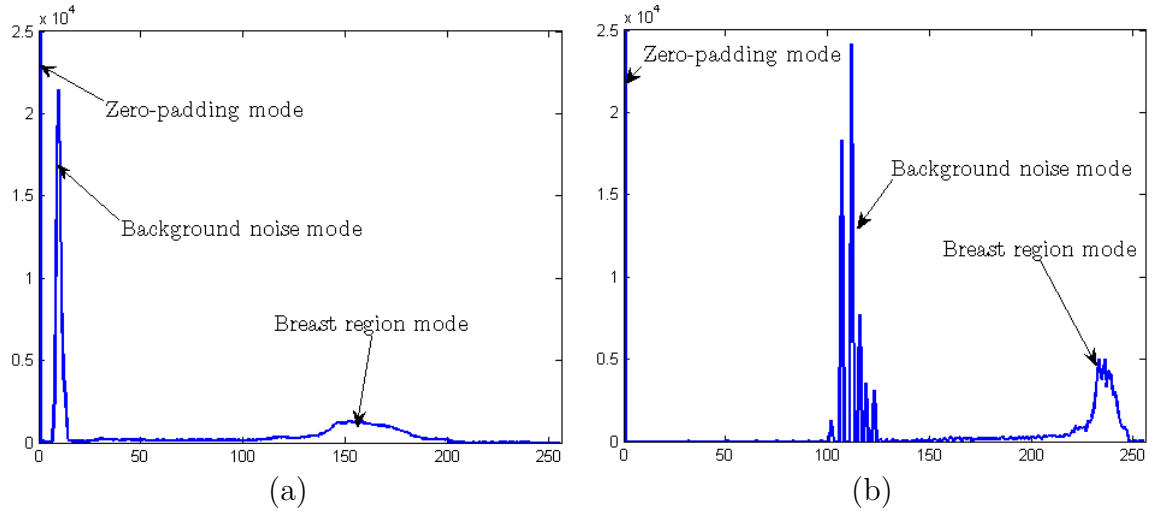


Figure 3.12: Histograms of a noisy mammogram showing the different modes of pixel classes present in the image. (a) Histogram of the initial mammogram (b) Histogram of the mammogram after logarithmic enhancement.

MIAS database.

To the best of our knowledge, the only related work evaluating their background segmentation approach with respect to widely accepted quantitative performance criteria is presented in [19]. They tested their method on a set of 84 images and reported the following performances: 0.41 ± 0.25 and 0.58 ± 0.67 for FP and FN rates respectively. In comparison with ours, FP rates are almost identical. The gap observed on FN rates is certainly due to noisy mammograms. Their influence increases with respect to the number of test images. Their approach which was based on deformable contours requires a prior estimation of breast region and was shown to be time consuming whereas the proposed approach is very simple in implementation and less time consuming as compared to those of related works.

3.6.6 Difficult cases and limitations

The main difficulty in extracting background regions in mammograms is noise or non uniformity of gray level intensity in the background created by digitization and zero-padding. It should be noted that the zero-padding was performed to standardize the size of images in the database. Zero-padding consists in assigning a zero intensity value to pixels in the blank areas of a mammogram. However, these padded pixels may create a third mode the gray level intensity empirical distribution as they do not blend with pre-existing background pixels whose mode is not necessarily centered on zero.

Figure 3.12 shows the histogram of a noisy mammogram. One can notice that this histogram violates our two class assumption and has three distinct modes representing foreground region, pre-existing background of the mammogram and zero-padded pixel region. It is clear that trying to partition such an image into two classes will produce inconsistent segmentation results. Consequently, good partition of such cases of mammogram can only be done into three classes of object instead of two. Figure 3.13 shows the segmentation of noisy mammograms using FCM with two and three classes respectively. One can see that using three classes of object, the segmentation produces

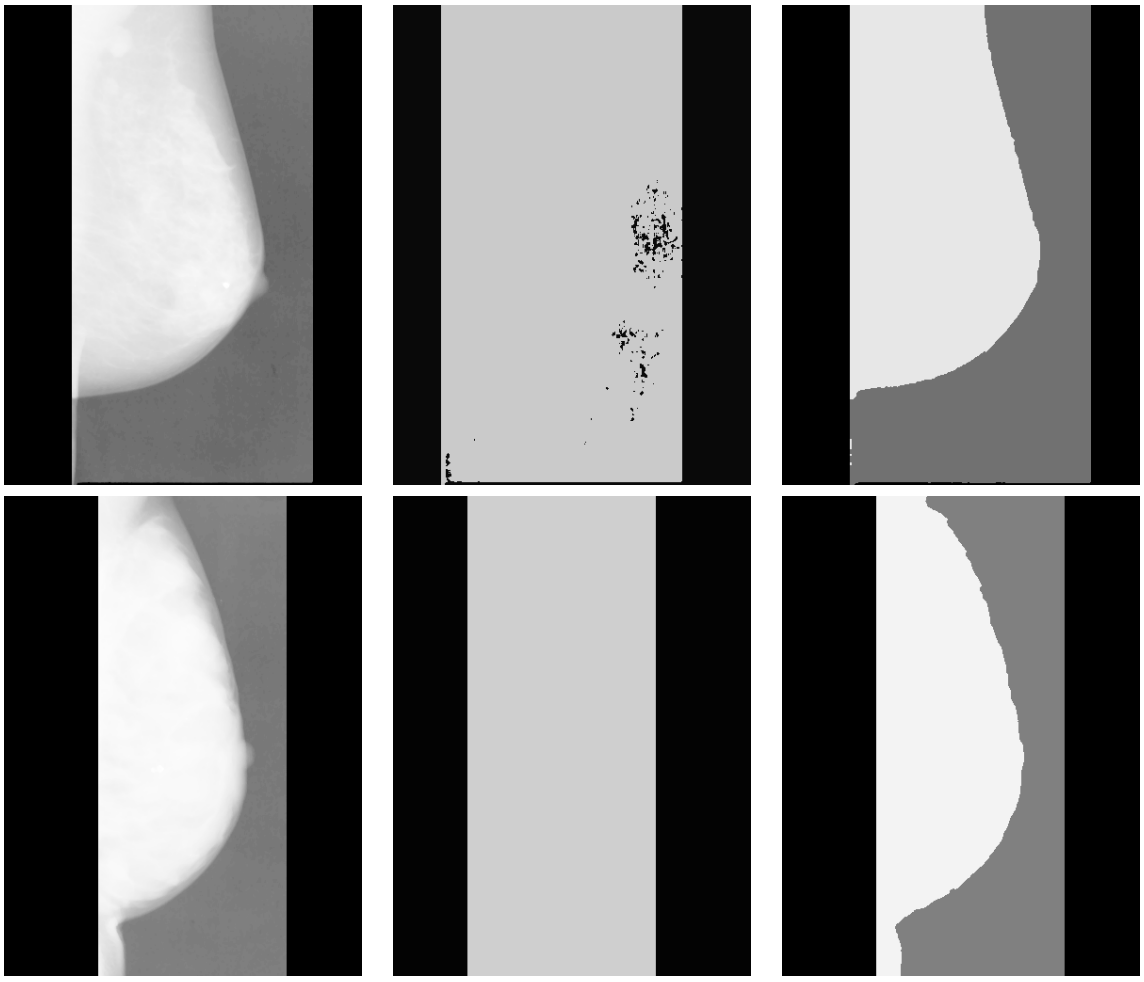


Figure 3.13: Segmentation of noisy mammograms into two and three classes of object respectively Left column shows the enhanced version of the mammograms. The shaded area represent the blank region of the mammogram that contains strong noise. Dark region is the zero-padded area. Middle column presents the results of segmentation in an attempt of partitioning the image into two classes. The right column shows segmentation results when partitioning the image into three classes.

consistent results as opposed to those obtained when only two classes are considered.

From these observations, it can be deduced that inaccurate or poor segmentation of mammogram commonly happens when there is a mismatch between the number of modes present in the histogram and the *a priori* number of classes. A way to address this issue is to derive the true number of modes for segmentation from the histogram of the image. An automatic estimation of number of modes in the histogram of an image is not an easy task. Peaks overlapping in histogram profile is one of the main problem to handle and may required to design complex strategy to tackle it. Another possibility would be to analyze the distribution of zero value pixels are on $\mathcal{G}$. Zero-padding correspond to situations where the distribution is deterministic (big stripes on each vertical image border).

Yet, one could easily think of addressing this issue under a supervised classification framework. Recent developments in supervised approaches have shown to be powerful in many segmentation tasks of significant complexity. In this situation, the problem

consists of learning discriminative features describing foreground and background regions from an annotated data-set. Unfortunately, contrast variability and unsharp edge transitions at breast boundary vicinity are strong impediments for such supervised algorithms to produce accurate results in all cases because there is not enough training examples to capture all these variables aspects. Besides, rather than using a great deal of effort to implement a supervised solution with reasonable complexity to address a simplistic problem such as background suppression in mammograms, for sake of computation effectiveness, it is justified to use a classical segmentation approach that can produce similar results.

On the other hand, keeping in mind that this difficulty is linked to the digitization of screen film mammograms, use of digital mammograms is a straightforward solution to circumvent these impediments as they simply will not occur.

3.7 Conclusion

Background suppression in mammographic images is the primary task to achieve in a CAD system. However, poor contrast and smooth variation of gray level intensities across breast edge do not facilitate accurate boundary identification. In addition, artifacts, excessive noise, scratches and band tape are impediments that influence algorithm reliability in suppressing mammogram backgrounds. Complex strategies that usually require heavy computations can be designed and tuned up to successfully deal with these peculiarities. However, keeping in mind that this is a pre-processing step, such efforts should be saved for the more challenging CAD tasks.

To overcome the difficulties encountered in mammogram background suppression, a logarithmic contrast correction operation is first of all applied to stretch the dynamic range in dark regions. This operation enhances mammogram contrasts especially at breast edge and thus facilitates boundary identification. A global threshold technique and a fuzzy C-means clustering algorithm were used on gray level intensities to partition pixels into two subsets: bright and dark. Assuming dark pixels belong to the background region while bright ones belong to the foreground region, the targeted segmentation is achieved. Nevertheless, noise in the background often induces incorrect classification of pixels near the breast edge. Consequently, a search path technique was developed to recover true breast boundary points. Finally, accurate results were obtained after applying this refinement step.

One difficulty that severely affects the methods introduced in this chapter is mismatch that sometimes occurs between the *a priori* number of classes and the actual number of modes in the image histogram. The source of this drawback lies in some excessive noise in the background of mammograms. This latter is created by the digitization process. A solution to this problem is to automatically estimate the true number of classes into which a given image should be partitioned using mixture models statistical analysis. It should be noted that addressing this issue left for future work.

Anyway, the foreground region extracted at this level still contains a non breast tissues such as pectoral muscle tissues which regularly tamper breast tissue characterization. The next chapter therefore introduces a strategy to segment out the pectoral muscle region in mammograms.

3.8 References

[1] P. Casti, A. Mencattini, M. Salmeri, A. Ancona, F. Mangeri, M. L. Pepe, and R. M. Rangayyan, "Estimation of the breast skin-line in mammograms using multidirectional gabor filters," *Computer in Biology and Medicine*, vol. 43, pp. 1870–1881, November 2013. 36, 37, 38

[2] R. Chandrasekhar and Y. Attikiouzel, "Mammogram-attribute database: A tool for mammogram segmentation and analysis," in *IASTED International Conference of Signal Processing, Pattern Recognition and Applications* (A. Press, ed.), (Rhodes, Greece), pp. 143–148, 3-6 July 2001. 36

[3] I. K. Maitra, S. Nag, S. K. Bandyopadhyay, and T.-H. Kim, "A novel approach to detect accurate breast boundary in digital mammogram using binary homogeinity enhancement algorithm," in *International Conference on Ubiquitous Computing and Multimedia Applications*, pp. 71–75, 2011. 36, 53

[4] C. Olsén and F. Georgsson, "Problems related to automatic nipple extraction," in *Image Analysis* (H. Kalviainen, J. Parkkinen, and A. Kaarna, eds.), vol. 3540 of *Lecture Notes in Computer Science*, pp. 161–176, Springer Berlin, Heidelberg, 2005. 36

[5] C. K. Feudjio, J. Klein, A. Tiedeu, and O. Colot, "Automatic extraction of pectoral muscle in the mlo view of mammograms," *Physics in Medicine and Biology*, vol. 58, pp. 8493–8515, 2013. 36

[6] C. Chen, M. Nielsen, N. Karssemeijer, and S. S. Brandt, "Breast tissue segmentation from x-rays radiographs," *Physics in Medicine and Biology*, vol. 59, no. 10, pp. 2445–2456, 2014. 36, 37, 38

[7] A. Oliver, J. Freixenet, R. Martì, J. Pont, E. Pérez, E. R. E. Denton, and R. Zwiggelaar, "A novel breast tissue density classification methodology," *IEEE Transactions on Information Technology in Biomedicine*, vol. 12, pp. 55 – 64, January 2008. 36

[8] Z. Chen, A. Oliver, E. Denton, and R. Zwiggelaar, "Automated mammographic risk classification based on breast density estimation," in *Lectures Notes in Computer Science* (J. Sanches, L. Micó, and J. Cardoso, eds.), vol. 7887, pp. 237–244, Springer-Verlag Berlin Heidelberg: IbPRIA, 2013. 36

[9] S. Pertuz, C. Julia, and D. Puig, "A novel mammography image representation framework with application to image registration," in *22nd international Conference on Pattern Recognition*, 24 - 28 August 2014. 36

[10] A. Touil and K. Kalti, "Iterative fuzzy segmentation for an accurate delimitation of breast region," *Computer Methods and Programs in Biomedicine*, vol. 132, pp. 137–147, 2016. 36, 37, 38

[11] Z. Chen and R. Zwiggelaar, "A combined method for automatic identification of the breast boundary in mammograms," in *5th International Conference on BioMedical Engineering and Informatics*, pp. 121–125, 16 -18 October 2012. 36, 37, 38

[12] M. Mustra, M. Grgic, and R. M. Rangayyan, "Review of recent advances in segmentation of the breast boundary and the pectoral muscle in mammograms," *Medical and Biological Engineering and Computing*, vol. 54, pp. 1003–1024, July 2016. 37

[13] I. K. Maitra, S. Nag, and S. K. Bandoypadhyay, "Accurate breast contour detection algorithms in digital mammogram," *International Journal of Computer Applications*, vol. 25, no. 5, pp. 1–13, 2011. 37

[14] C. Feudjio, A. Tiedeu, M. Gordan, S. Domngang, and A. Vlaicu, "Computerized detection and smoothing of contour in mammograms," in *11th African Conference on Research in Computer Science and Applied Mathematics*, pp. 299–306, 13-16 October 2012. 37

[15] U. Bick, M. L. Giger, R. A. Schmidt, R. M. Nishikawa, D. E. Wolverton, and K. Doi, "Automated segmentation of digitized mammograms," *Academic Radiology*, vol. 2, no. 2, pp. 1–9, 1995. 37

[16] P. Kus and I. Karagoz, "Fully automated gradient based breast boundary detection for digitized x-ray mammograms," *Computers in Biology and Medicine*, vol. 42, pp. 75–82, 2012. 37

[17] K. Czaplicka and J. Wlodarczyk, "Automatic breast-line and pectoral muscle segmentation," *Schedae Informaticae*, vol. 20, pp. 195–209, 2011. 37

[18] R. S. C. Boss, K. Thangavel, and D. A. P. Daniel, "Automatic mammogram image breast region extraction and removal of pectoral muscle," *International Journal of Scientific and Engineering Research*, vol. 4, pp. 1722–1729, May 2013. 37

[19] R. J. Ferrari, A. F. Frère, R. M. Rangayyan, J. E. L. Desautels, and R. A. Borges, "Identification of the breast boundary in mammograms using active contour models," *Medical and Biological Engineering and computing*, vol. 42, pp. 201–208, 2004. 37, 38, 51, 54

[20] A. Hoyer and W. Spiesberg, "Computerized mammogram processing," *Phillips Technical review*, vol. 38, pp. 347–355, 1979. 37

[21] B. K. M. Shahedi, R. Amirfattahi, F. T. Azar, and S. Sadri, "Accurate breast region detection in digital mammograms using local adaptive thresholding method," in *8th International Workshop on Image Analysis for Multimedia Interactive Services*, (Santorini, Greece), pp. 26–29, 6-8 June 2007. 37, 53

[22] T. Ojala and J. Liang, "Interactive segmentation of the breast from digitized mammograms with united snakes," Tech. Rep. 315, Turku Centre for computer Science, 1999. 38

[23] T. Ojala, J. Nappi, and O. Nevalainen, "Accurate segmentation of the breast region from digitized mammograms," *Computerized Medical Imaging and Graphics*, vol. 25, no. 1, pp. 47–59, 2001. 38

[24] M. Wirth and A. Stapinski, "A segmentation of the breast region in mammograms using snakes," in *1st Canadian Conference on Computer and Robot Vision*, (Ontario, Canada), pp. 385–392, 17-19 May 2004. 38, 52, 53

[25] K. J. McLoughlin and P. J. Bones, "Location of the breast-air boundary for a digital mammogram image," in *Image and Vision Computing*, 2000. 38

[26] R. C. Gonzalez and R. E. Woods, *Digital image processing*. Addison-Wesley, Reading, MA, 1992. 40, 41

[27] E. Wharton, S. Agaian, and K. Panetta, "Comparative study of logarithmic enhancement algorithms with performance measure," in *SPIE Electronic Imaging*, (San Jose, CA,), pp. 606 412–606 413, 17 January 2006. 41

[28] N. Otsu, "A threshold selection method from gray level histograms," *IEEE on Systems, Man, and Cybernetics*, vol. 9, pp. 62–66, January 1979. 42, 43

[29] Z. Chen and R. Zwiggelaar, "Segmentation of breast region with pectoral muscle removal in mammograms," in *Medical image understanding and analysis* (A. H. Bhalerao and N. M. Rajpoot, eds.), (Coventry, UK), July 2010. 48

[30] S. M. Kwok, R. Chandrasekhar, Y. Attikiouzel, and M. T. Rickard, "Automatic pectoral muscle segmentation on mediolateral oblique view mammograms," *IEEE Transactions on Medical Imaging*, vol. 23, pp. 1129 – 1140, September 2004. 48, 49

[31] J. Chakraborty, S. Mukhopadhyay, V. Singla, N. Khandelwal, and P. Bhattacharyya, "Automatic detection of pectoral muscle using average gradient and shape based feature," *Journal of Digital Imaging*, vol. 25, no. 3, pp. 387–399, 2012. 49

[32] J. Suckling, J. Parker, D. Dance, S. Astley, I. Hutt, C. Boggis, I. Ricketts, E. Stamatakis, N. Cerneaz, S. Kok, P. Taylor, D. Betal, and J. Savage, "The mammographic images analysis society digital mammogram database," *Experta Medica International Congress Series*, vol. 1069, pp. 375–378, 1994. 49

[33] J. Nagi, A. S. Kareem, F. Nagi, and S. K. Ahmed, "Automated breast profile segmentation for roi detection using digital mammograms," in *IEEE Conference on Biomedical Engineering and Science*, (Kuala Lumpur, Malaysia), pp. 87–92, 30 November-2 December 2010. 53

[34] L. Wei, L. Y. sheng, R. X. Yi, and D. Peng, "A new contour detection in mammogram using sequential edge linking," in *Intelligent Information Technology Application, 2008. IITA '08. Second International Symposium on*, vol. 1, pp. 197–200, December 2008. 53

[35] G. Liasis and S. Petroudi, "Estimation of the breast boundary in mammograms using level sets," *Journal of Medical Imaging and Health Informatics*, vol. 1, pp. 199–206, 2011. 53

Chapter 4

Pectoral muscle extraction in mammograms

Contents

4.1 Introduction

Even though one has succeeded in getting rid of the background region, advanced analysis tasks (*e.g.* dense breast tissue characterization) are impaired by the presence of remaining non-breast tissues, *i.e.* the pectoral muscle. Therefore, it is necessary to first identify the pectoral muscle region separately before any further analysis of breast tissues. One of the major difficulties to overcome when segmenting out the pectoral muscle is its strong overlapping with dense glandular tissues which tampers its extraction. Like for the previous segmentation problem exposed in chapter 3, it is preferable to solve the dual boundary detection problem instead of the region segmentation one. Unfortunately, the slight variation of gray level intensities across pectoral muscle boundary make this problem much more demanding. In addition, mammographic MLO views layout is a real problem to handle for the automatic extraction of the pectoral muscle. In such conditions, applying straightforwardly classical segmentation or contour detection methods leads to poor performances. In this chapter, we introduce an *ad hoc* approach [1] relying on two main steps:

1. a coarse contour pre-detection,
2. followed by an iterative contour fitting.

The performances of this approach are demonstrated on a mammographic dataset using a comparison to reference standard provided by radiologists.

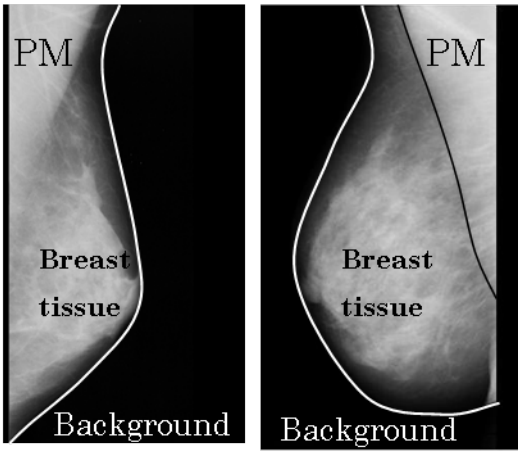


Figure 4.1: Pectoral muscle location in the left and right views of mammograms. According to the mammogram view, the pectoral muscle may appear at the upper left or right corner of the image respectively.

4.2 Pectoral muscle segmentation: problem statement and state-of-the-art

In this section, we begin with a detailed presentation of difficulties in pectoral muscle segmentation. Afterwards, an overview of the literature on this issue is given. Finally a general architecture of the proposed strategy to tackle this problem is introduced and justified.

4.2.1 Automatic pectoral muscle extraction challenges

As already justified in section 2.7, a prior extraction of patterns like breast contour [2, 3, 4], nipple [3] and pectoral muscle [5, 6, 4] have to be carried out to allow accurate analysis of breast tissues. This chapter focuses on the pectoral muscle which appears almost only in MLO views. Roughly speaking, it is assumed to be a triangular region with high gray level intensities located at the upper left corner of the breast region, provided that the breast is right oriented in the image.

Another interest of extracting pectoral muscle is that its contour is used as landmark for registration [7] in mammogram comparison or for breast reconstruction from multi-views of mammograms [8, 6]. Consequently, the justifications for performing pectoral muscle segmentation are not restricted to the scope of the approach described in Section 2.7.

According to the mammogram layouts shown in figure 4.1, several major difficulties to overcome for automatic extraction of pectoral muscle in mammograms can be noticed:

- pectoral muscle appears in different locations with respect to **breast orientation** in mammograms,
- there can be a strong **overlap** with fibro-glandular tissues that makes it difficult to reliably characterize the pectoral muscle image region,
- some **skin folding** tampering with the pectoral muscle contour can be found,

- X-ray images tend to produce **blur** edges for pectoral muscles especially in its lower part,
- there is a high **variability of region surface**; in some extreme cases, the pectoral muscle covers a very small area or is completely missing.

These various and complex factors illustrate how difficult it is to automatically extract pectoral muscle in mammograms [8].

4.2.2 Related works on pectoral muscle extraction

Many approaches dealing with pectoral muscle extraction have been introduced in the literature. Certain methods used for breast contour detection presented in section 3.2.2 were applied also to pectoral muscle contour detection. Therefore there are some redundancies between the current section and section 3.2.2. These works can be sorted into three main groups as presented in the following paragraphs:

Pixel intensity based segmentation

Pixel intensity based segmentation approaches use only gray level information to partition an image. In this way, Suckling *et al.* [9] used a multiple linked self-organizing neural network to extract pectoral muscle by segmenting a mammogram into four main types of components: background, pectoral muscle, fibro-glandular region and adipose region. On the other hand, Saha *et al.* [10] introduced a semi-automatic method relying on an input from an operator to locate the pectoral muscle prior to its automatic delineation. Raba *et al.* [11] used region growing method to extract pectoral muscle. However, their results were assessed only visually and rated on a two scale grade (adequate and quite adequate).

Texture based segmentation

Texture based segmentation approaches use information derived from image texture analysis to segment an image into regions. Using such a strategy, many algorithms developed to extract pectoral muscle were based on texture-field orientation [8], wavelets decomposition [12], Gabor wavelets [5] and non linear filtering [13]. In these approaches, it is assumed that homogeneous regions have their texture orientation feature in the same direction. Location where texture orientation feature are of opposite direction are identified as probable traces or line segment of boundary. Finally, the pectoral muscle contour is obtained by thresholding the line segments or candidate regions using shape and size criteria. To take advantage of both intensity and texture information, Li *et al.* [14] introduced an approach based on homogeneous texture and intensity deviation to identify an initial pectoral muscle edge. They subsequently employed a Kalman filter to refine the ragged initial edge. On the other hand, Chakraborty *et al.* [15] introduced a method based on the average of gradient to the extract pectoral muscle. They used a weighted average gradient and adaptive band selection to approximate the pectoral muscle boundary as a straight line and local gradients to fit to the true edge of the pectoral muscle.

Texture features have been also intensively used to analyze mammograms. In this way, Chen *et al.* [16] implemented a two-layer neural classifier to segment a mammogram into three main regions corresponding to breast tissues, pectoral muscle and background.

Features include local binary patterns, histogram of oriented gradients and intensity histogram. In the first layer, individual experts each formed by a feature vector and an SVM classifier vote the local class of the mammogram. The second layer which is a committee vote machine combines the previous labels in association with prior maps to produce final segmentation. Although promising results are obtained, this approach is less effective at producing an accurate pectoral muscle contour extraction and at dealing with dense breast tissues overlapping with pectoral muscle.

A more detailed presentation of texture features for mammographic images is given in the next chapter (see subsection 5.4.1).

Contour based segmentation

Another commonly used approach consists of estimating the boundary between pectoral muscle and mammary tissues as a straight line by the use of Hough transform [5, 17]. The main drawback of this method is that the pectoral muscle edge is not always straight and may sometimes have concavities. To cope with this difficulty Kwok *et al.* [18] firstly estimated the straight line delineating the pectoral muscle using iterative thresholds and then refined this line by cliff detection to be aligned on the pectoral muscle boundary curvature.

Recently, a new image segmentation approach based on graph theory has been introduced. This approach relies on the assumption that the pectoral muscle region boundary has its two extremities in the first row of the image and the first non blank column respectively. These extremities are called end-points and the shortest path between two end-points is searched by minimization of a cost-function. Domingues *et al.* [19] estimated the pectoral muscle contour as the shortest path throughout the two end-points in the image gradient. The image is represented as a weighted graph where nodes are pixels and edges are connecting neighbor pixels. Finally, the shortest path is found by mean of a SVM classifier previously trained to minimize the cost function of the path between two end-points. Ma *et al.* [7] introduced another graphical approach based on adaptive pyramids (APs) and minimum spanning trees (MST) to extract pectoral muscle. Camilus *et al.* [20] implemented a graph-cut based segmentation technique followed by a refining step using Bezier curves for identifying the pectoral muscle edge in mammograms.

It is noteworthy that most of these works were tested on small data-sets and that the accuracy of their algorithms was based on visual inspection. Moreover, only few performance metrics are provided. Good reviews on pectoral muscle segmentation approaches as well as discussions on the overall performances achieved by various strategies introduced in the literature to address this issue are presented in [21, 22]. The analysis of this literature review shows how complex it is to extract the pectoral muscle and that its boundary cannot be retrieved through a straightforward method. In this PhD, we follow one of the most popular strategy [15, 18, 7] which relies on two main steps. The pectoral region frontier is roughly estimated based on gray level homogeneity or *a priori* information on the pectoral muscle location. The boundary is then refined by selection of potential contour line segments, search of the shortest path [7, 8], or cliff detection using gray level variation [18].

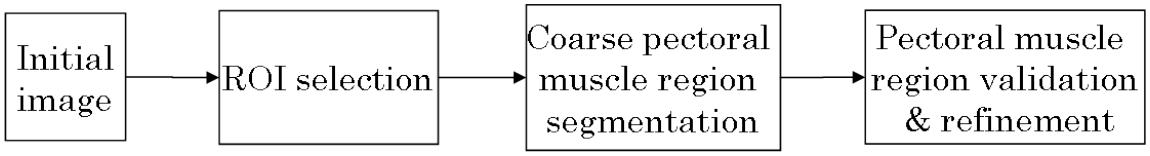


Figure 4.2: Flowchart of the strategy to perform for pectoral muscle removal in mammograms

4.2.3 Strategy for pectoral muscle segmentation

Regarding pectoral muscle layouts in mammograms in figure 4.1, a number of reasonable assumptions can be made to design an efficient approach to extract the pectoral muscle in mammograms.

Assumption 1: The pectoral muscle appears at only two different locations in mammograms depending on which breast (left or right) is screened. Therefore, a pre-processing is needed so that the pectoral muscle lies at the same location (upper left corner) in all images from the dataset. By convention, only images with left oriented breast have to be flipped sideways.

Assumption 2: Assuming that the pectoral muscle now appears only in the upper left corner in a mammogram, a region of interest (ROI) is selected so that it roughly covers almost all the pectoral muscle area in a mammogram. The ROI parameters (width and height) are given by a heuristic based on *a priori* knowledge on women anatomy and mammography protocol. Defining a ROI allows to reduce the contour search area and the computation time while ensuring better segmentation results.

Assumption 3: Once the ROI is obtained, it can be segmented into sub-regions: pectoral muscle, glandular tissues and background. Using the approach presented in the previous chapter, all pixels belonging to background are known and forced to zero. It would also be possible to consider the intersection of the foreground region with the ROI but this is a little bit more memory demanding.

Given these assumptions, two initial steps arise: conditional image flipping and ROI selection. Afterwards, a coarse segmentation can be performed followed by a final refinement as in most state-of-the-art approaches. Figure 4.2 summarizes this global strategy using a flowchart.

4.3 Pectoral muscle segmentation: a new comprehensive approach

This section presents one of the main contribution introduced as part of this PhD manuscript, *i.e.* an efficient pectoral muscle contour estimation. In the sequel, each step from the flowchart of figure 4.2 is presented in details.

4.3.1 Breast orientation detection

Using breast contour extracted in the previous chapter, the chest wall in the image is assumed to be, in first approximation, a straight line. In figure 4.3, the chest wall corresponds to segment OB. Its position is thus determined by applying a Radon

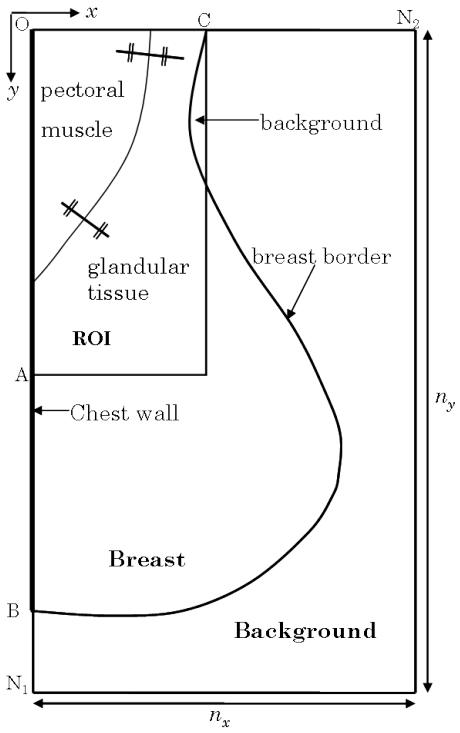


Figure 4.3: Schematic view of a mammogram with a right orientation and a pectoral muscle at the upper left corner of the image. The coordinate axis are directed as shown with the origin at top left. N_1 and N_2 are respectively the number of rows and columns in the image. ROI is selected as a rectangular window of size OAxOC, where OA is half height of the image and C is the upper endpoint of the breast contour.

transform to the breast contour image. The Radon transform is closely related to Hough transform which is a common function used in computer vision to detect straight lines in images. The Radon transform computes the integral of multiple, parallel beam projections of the image at different angles by rotating the source around the center of the image.

Using notations in compliance with the previous chapter, let u denote an input image which is a function over a discrete grid $\mathcal{G}$. Suppose $\mathcal{G}$ is obtained by evenly sampling the unit square of $\mathbb{R}^2$. Let x and y denote the axes of $\mathbb{R}^2$. A pixel is a pair $(p_x, p_y) \in \mathcal{G}$ and $u\,(p_x, p_y)$ is its gray level intensity in the image. The Radon transform is computed at each line of u parallel to a $y'-$axis as:

$$R_\theta(x') = \int_{-\infty}^{\infty} u(x'\cos\theta - y'\sin\theta, x'\sin\theta + y'cos\theta)\; dy', \tag{4.1}$$

where

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix},$$

and x' and y' are two axes obtained by applying a rotation of angle θ to x and y. Note that equation (4.1) does not take discretization and rounding effects into account.

Figure 4.4 shows an illustration of straight line position detection in an image of breast contour. The coordinates of the brightest spot found in the Radon transform plot

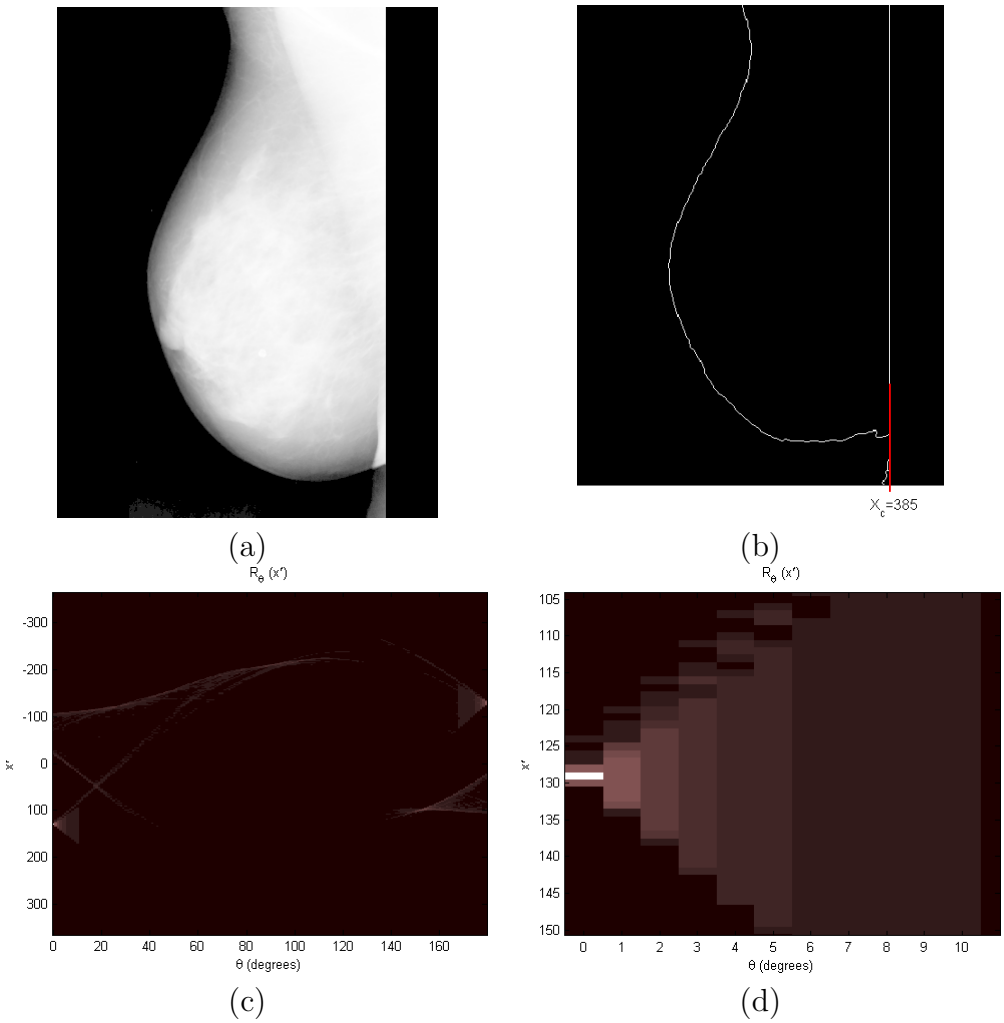
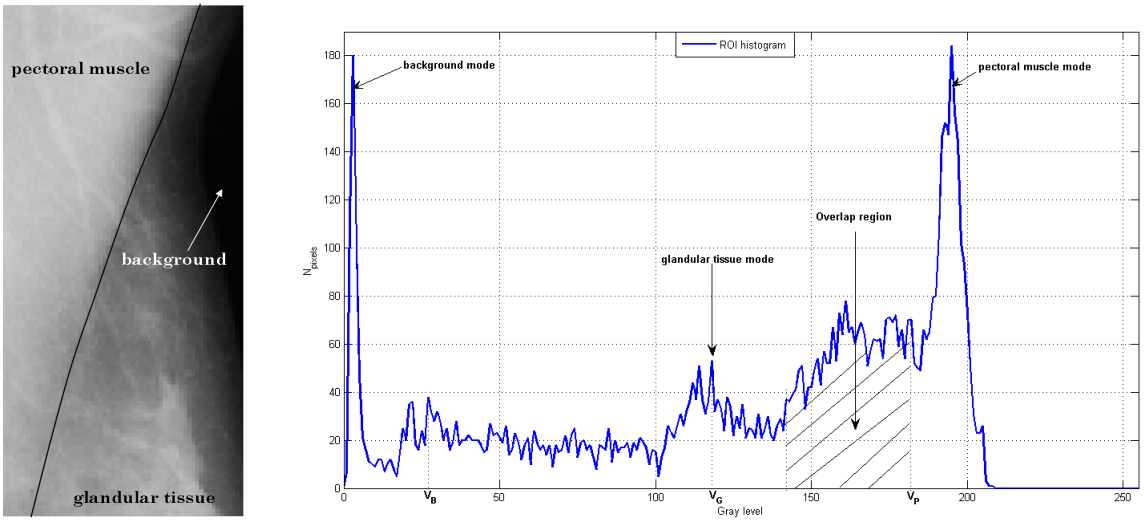


Figure 4.4: Breast orientation detection. The chest wall position in a mammogram is an indication of its breast orientation. Radon transform is applied to an image of breast contour to derive its position in a mammogram. (a) Initial image enhanced, (b) contour extracted after a rough segmentation, (c) plot of Radon transform projections of the image of breast contour, (d) zoomed-in area containing the brightest spot of image (c). Spots in the plot of Radon transform define probable lines in the image. Spot intensity varies with respect to the length and width of lines in the image. The coordinates x' and θ give the position and the orientation of the lines in the original image with respect to Radon reference whose origin is taken at the center of the image. The brightest spot angle value $0°$ means that the corresponding straight line is vertical and the value of x' indicates that the line is in the right half of the image.

enable to derive the position of the chest wall position and therefore breast orientation in the screened mammogram. Left-oriented breast mammograms are then vertically mirrored thereby allowing to process images with identical region layouts.

4.3.2 ROI selection

Once a mammogram has been processed in such a way that the pectoral muscle appears in the upper left corner, the point O in figure 4.3 is necessarily the upper left corner of the target ROI. A relevant choice for the ROI height is the length of the line segment OA where A is the half height of the image. This allows to reduce the amount of dense



(a) (b)

Figure 4.5: ROI extracted from a mammogram (mdb018). (a) ROI with pectoral muscle, breast tissues and background. Pectoral muscle boundary was traced out manually. The lower portion of the ROI contains an overlap between pectoral muscle and breast tissues. (b) Intensity histogram of the ROI. For better visualization, the histogram was plotted using the raw mammogram. V_B, V_G and V_P are modes of the following regions: background, breast glandular tissues and pectoral muscle respectively.

glandular tissue in the ROI in order to guarantee good segmentation results. On the other hand, the ROI width should be large enough to fully contain the pectoral muscle. A simple way to ensure this is to set the ROI width equal to OC where C is the top end-point of the breast border (see again figure 4.3).

4.3.3 ROI segmentation

Ideally, the ROI should entirely cover the pectoral muscle while rejecting as much as possible other regions in order to ensure good extraction results. However, in practice such a hypothesis is not always met since there is a high variability in shape and size from one patient pectoral muscle to another. The concavity of the breast border in the vicinity of point C (see figure 4.5) implies that the ROI comprises a small part of the background. Therefore, the selected ROI is partitioned into three regions:

- pectoral muscle,
- breast glandular tissue,
- background (whose pixel values are zeros).

The choice of the segmentation method applied in this stage is justified from an analysis of gray levels of regions contained in the ROI. Figure 4.5(a) shows the ROI extracted in a mammogram where the pectoral muscle has been delineated from breast tissues manually. Along with this ROI, the corresponding gray level histogram is given by figure 4.5(b). It should be noticed that for better visualization purpose of region modes, the intensity histogram of the ROI was plotted using raw mammogram image. In many MLO mammograms, the lower portion of the pectoral muscle is spatially superimposed

on some glandular tissues. The impact of the overlap between glandular tissues and pectoral muscle can easily be observed on the ROI histogram (see figure 4.5(b)). There is no clear cut separation between the modes of the histogram. Thus, it is not always possible to find thresholds that can completely separate the pectoral muscle from other tissues [6].

The only pieces of information available for segmentation are textures and gray level intensities [8]. However, texture and gray level features are not very discriminating. In addition, the feature distributions of the regions are overlapping. In such conditions, one can think of two solutions: either changing the feature space hoping for a less inter-class overlapping or resorting to robust algorithms. Given that the ROI content is not a complex scene, the latter solution was retained.

The FCM algorithm, which is an unsupervised classification method, can be used to optimally cluster overlapping data. Similarly as in the background region segmentation presented in the previous chapter, it was decided to use the FCM algorithm for a coarse segmentation of the pectoral muscle region. This segmentation is then refined using an iterative contour fitting. The next subsections present this approach.

4.3.4 Limitations of the FCM segmentation

The FCM algorithm was already presented in the previous chapter, see subsection 3.4.2, algorithm 1. As can be observed from equations 3.10 and 3.11, one needs the region modes to compute membership functions and conversely. Therefore, the FCM algorithm has to successively update the cluster centers and membership functions in order to minimize the objective function J_m. It should be noticed that the FCM algorithm starts by randomly initializing the cluster centers and evaluating the membership functions. At each iteration, the membership function and the cluster centers are updated. An optimal partition of clusters is obtained when from one iteration to another the Euclidean distance between the previous and the current cluster centers is less than a convergence parameter denoted by ϵ. Segmentation of the image is then performed by assigning each pixel to the class corresponding to the highest membership function value.

A major weakness of the classical FCM algorithm is the random initialization of cluster centers which penalizes fast convergence of the algorithm. Furthermore, as the FCM algorithm relies on a gradient descend, such a random initialization of cluster centers may lead the algorithm to converge to a local minimum thus resulting in less accurate segmentation results.

To better serve our purpose, two modifications will be added in order to speed up computation and ensure better segmentation performances. These modifications address the following limitations of FCM:

1. random initialization of region modes which penalizes fast convergence of the algorithm,

2. non-use of spatial information leading to pixel clusters made of unconnected pixel aggregates scattered on the image grid.

4.3.5 Robust cluster centers initialization

It has been shown that neatly estimating region modes significantly reduces the number of iterations while enhancing accuracy in the results. One appealing method for the estimation of region modes is the block density approach [23]. In the latter, the image (normalized) histogram $\mathbf{h} = [h_j]_{0 \leq j \leq L-1}$ is subdivided into c blocks B_i of equal size, one block for each region involved in the segmentation. The j^{th} component of $\mathbf{h}$ is h_j and represents the empirical probability to observe the gray level j in u. If the image has $L = 256$ gray levels, then $B_i = \left\{(i-1) \times \frac{256}{c}, ..., i \times \frac{256}{c} - 1\right\}$. A random variable X is defined to represent the probability for a pixel to have a given gray level value. The distribution of variable X corresponds to gray level frequencies of occurrence in the image, *i.e.* the histogram: $P(X = j) = h_j$. The initial mode of the first region denoted by $\nu_1^{(0)}$ and is the most probable gray level in the first block B_1. Its probability is denoted by h_{ν_1}. The subsequent region modes $\nu_2^{(0)}, ..., \nu_c^{(0)}$ are then computed by finding the gray level of the block that maximizes the following expression:

$$\nu_i^{(0)} = \arg\max_{j \in B_i} h_j \times d(j, \nu_{i-1}^{(0)}), \tag{4.2}$$

with $d(j, \nu_{i-1}^{(0)})$ a distance given by:

$$d(j, \nu_{i-1}^{(0)}) = \left((j - \nu_{i-1}^{(0)})^2 + (h_j - h_{\nu_{i-1}})^2\right)^{1/2}. \tag{4.3}$$

Robust cluster center initialization algorithm

Using equation (4.2), the FCM was modified in such a way that the cluster centers are initialized at robust region modes and subsequently updated in the same way as in the classical FCM setting. The pseudo-code presented in algorithm 2 shows an outline of the mode initialization process.

Algorithm 2 Robust cluster modes initialization

Require: $c = 3$, L, $\mathbf{h}$
  for i from 1 to c **do**
    $B_i \leftarrow$ Split gray level range into c blocks of equal size
  end for
  for i from 1 to c **do**
    for all $j \in B_i$ **do**
      $d(j, \nu_{i-1}^{(0)}) \leftarrow$ Estimate distance between current block samples and previous mode with equation (4.3)
    end for
    $\nu_i^{(0)} \leftarrow$ Find the sample maximizing expression (4.2)
  end for
  $\boldsymbol{\nu}^{(0)} = \left\{\nu_1^{(0)}, ..., \nu_c^{(0)}\right\}$: Vector of cluster centers

4.3.6 Exploiting spatial information in FCM clustering

As explained previously, the FCM algorithm often misclassifies some pixels because it processes each pixel as a point living in the gray level feature space thereby completely dropping topological information related to its position in the image grid. This section

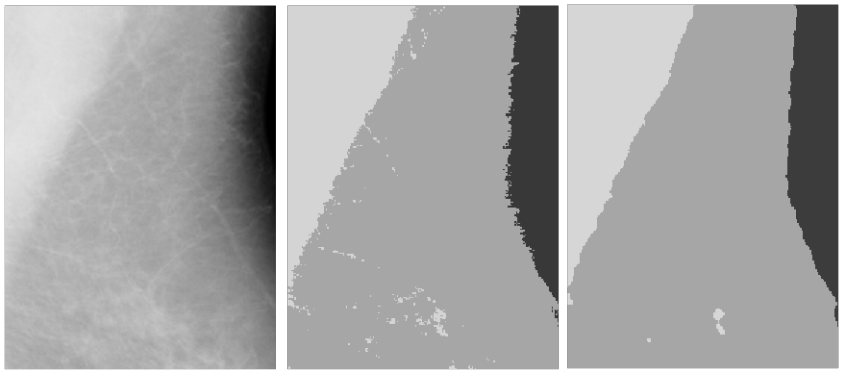


Figure 4.6: Segmentation of the ROI into 3 regions: pectoral muscle, breast tissue and background. On the left column lies the initial image, the two next columns are results of segmentation using FCM algorithm without and with local information respectively.

introduces a new modified FCM algorithm that allows better pixel classification results. The idea of using spatial information inside the FCM procedure when processing images has already been proposed by other authors [24, 25]. It has been shown that integrating local information when updating pixel membership functions significantly improves the segmentation performances in MRI brain images [26]. The neighborhood information is integrated in the clustering process by computing a weighted version of the membership function defined as followed:

$$\tilde{\mu}_i(\mathbf{p}) = \frac{\mu_i^m(\mathbf{p})\, S_i^n(\mathbf{p})}{\sum_{k=1}^{c} \mu_k^m(\mathbf{p})\, S_k^n(\mathbf{p})} \tag{4.4}$$

$$S_i(\mathbf{p}) = \sum_{\mathbf{p}' \in N(\mathbf{p})}^{c} \mu_i(\mathbf{p}') \tag{4.5}$$

$N(\mathbf{p})$ is the neighborhood of pixel $\mathbf{p}$ in the image grid, *i.e.* a square window of size n_w. m, n are weighting parameters and c is the number of image regions or equivalently the number of modes in the empirical gray level distribution.

The modified version of the membership function $\tilde{\mu}_i$ can be regarded as class-wise product of the initial function μ_i with a local average of μ_i. The functions S_i act on the membership functions of a given pixel by leading its values towards classes with many representatives in the local window. Consequently, the membership functions are now not only depending on pixel gray level intensities but also on neighboring pixels. Figure 4.6 shows the effects of integrating local information on segmentation.

One can see that some misclassified pixels in the ROI are now well classified thanks to neighborhood information. This modified FCM algorithm can still be criticized to increase computation load since the membership function has to be computed for every pixel locally and within a window. However, this increased computation load is partially compensated by the robust modes initialization introduced in subsection 4.3.5.

modified FCM algorithm

Using equation 4.2, the membership functions are initialized at robust regions modes and then updated using local information. The modified version of the FCM algorithm

will be referred to as the **mFCM** algorithm in the remainder of this manuscript. The outline of the mFCM (with default parameter values) is given in algorithm 3.

Algorithm 3 modified FCM (mFCM)

Require: $c = 3$, $m = 2$, $n = 3$, $\epsilon = 0.01$, n_w=5x5
  $k \leftarrow 1$
  $\boldsymbol{\nu}^{(0)} = \left\{\nu_1^{(0)}, ..., \nu_c^{(0)}\right\} \leftarrow$ Estimate region modes using sample block density with equation (4.2)
  for i from 1 to c **do**
    $\mu_i^{(0)} \leftarrow$ Evaluate membership function with equation (3.10)
  end for
  while $\|\boldsymbol{\nu}^{(k+1)} - \boldsymbol{v}^{(k)}\| > \epsilon$ **do**
    for i from 1 to c **do**
      $\tilde{\mu}_i^{(k+1)} \leftarrow$ Update weighted membership function with equation (4.4)
    end for
    for i from 1 to c **do**
      $\nu_i^{(k+1)} \leftarrow$ Update regions mode with equation (3.11)
    end for
    $k \leftarrow k + 1$
  end while

Once the mFCM algorithm has converged, the segmentation is performed by assigning each pixel to the region with the highest membership value:

$$v_{seg} = \arg\max_{1 \leq i \leq c} \tilde{\mu}_i. \tag{4.6}$$

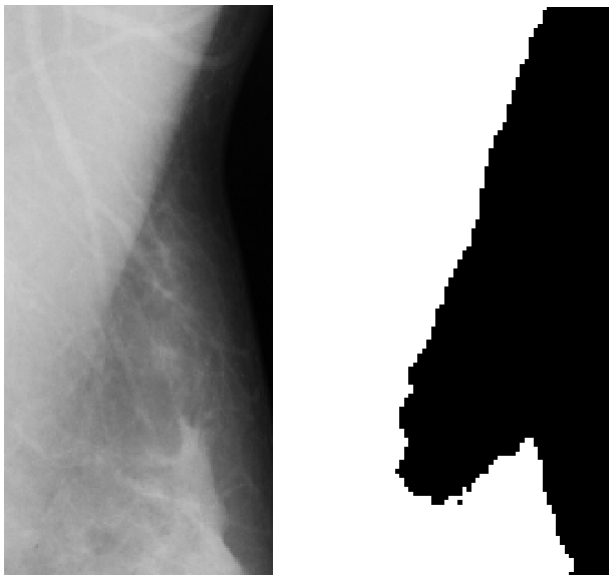
The pectoral region is then identified as the area with highest gray levels and containing the upper left corner of the image.

4.3.7 Pectoral muscle region validation and refinement

The mFCM segmentation results match our purpose and allow the estimation of a relevant approximation of the pectoral muscle boundary. However, in a few cases, an undesirable drift is observed when breast tissues have a density close to the pectoral muscle one. Figure 4.7 shows an example of this phenomenon. In this example, a strong tissue overlapping results in an over-segmented pectoral muscle region. In the upper half of the ROI, the boundary estimation is satisfying but in the lower half, the estimated boundary drifts apart from the actual one, thereby creating an irrelevant concavity in the pectoral muscle region.

A quick validation post-processing is necessary to solve that specific problem. We propose to exploit the curvature angles of contour points to that end. The curvature angle θ_j of the j^{th} contour point is defined as the angle between the tangent line of preceding contour points and the tangent line of following contour points at each point of the contour (see figure 4.8(a)). The angle between these two tangent lines is given by the formula:

$$\theta_j = \tan^{-1}\left(\frac{a_1 - a_2}{1 + a_1 a_2}\right). \tag{4.7}$$



(a) (b)

Figure 4.7: Incorrect pectoral muscle region: (a) Initial image, (b) Incorrect pectoral muscle region obtained from segmentation.

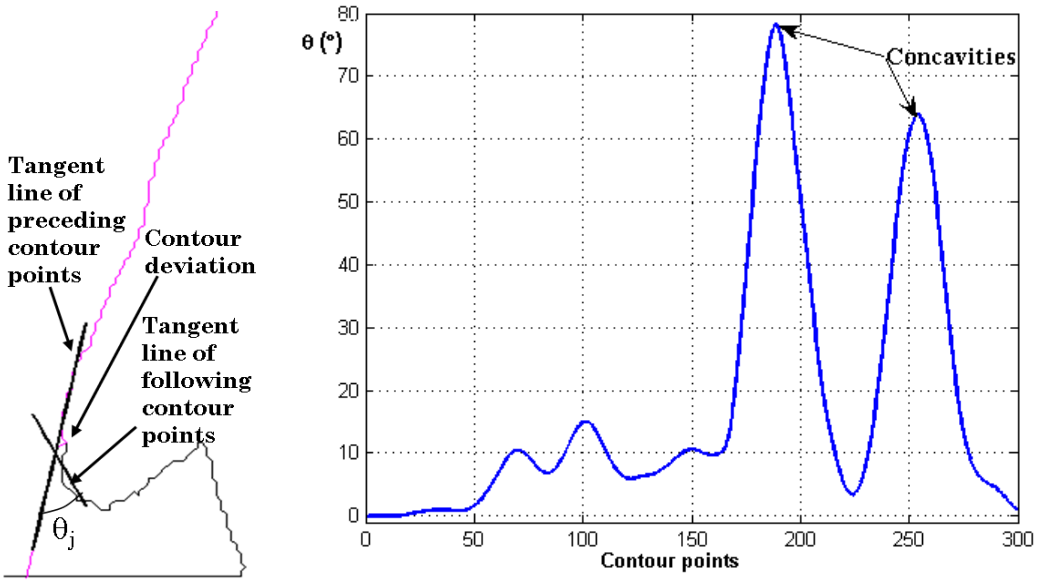
where a_1 and a_2 are the respective slopes of the tangent lines.

It can be seen from figure 4.8 that the estimated pectoral muscle contour is very noisy (quick variation of contour pixel positions). Consequently it is necessary to compute the tangent lines on a rather wide range of contour points to smooth the absolute values of curvature angles. We obtained reliable curvature angle estimates by using a span of 25 pixels for tangent lines and least square regression for smoothing the angle curve.

Abnormal deviations in the contour yield large peaks in the graph of the function mapping contour pixels to their curvature angles. Conversely, if there is no abnormal deviation in the contour, the angle of curvature will remain quite flat. Figure 4.8(b) shows the graph of the angle curvature function after smoothing in the case where the pectoral region was overestimated due to glandular tissue overlapping. The function computation is robust enough to allow the use of naive techniques for finishing the job. On the datasets investigated in this PhD, it was observed that a thresholding at 30° suffices to detect all abnormal deviation occurrences.

When the angle of curvature is greater than the threshold, the rest of the contour is extrapolated from the tangent line of points preceding the deviation. This is illustrated by figure 4.8(a). This approximation allows to cope with the cases of pectoral region overestimation especially where there is an overlap between pectoral muscle and glandular tissues. In cases where the pectoral muscle is not entirely contained in the ROI, the rest of its contour in the image is estimated by a straight line on the basis that pectoral muscle is in first approximation triangular.

Now, the contour obtained from the estimated pectoral region roughly corresponds to the true pectoral boundary but does not perfectly fit it. To improve the accuracy of the algorithm, a refinement of the extracted contour is carried out. The next paragraphs explain the strategy carried out to refine the contour in order to improve the accuracy



(a) (b)

Figure 4.8: Incorrect pectoral muscle contour detection and correction (a) Extracted contour with tangent lines (**bold black segments**) at the deviation point of the contour and the final estimated contour (**magenta**) with the lower part fitted to straight line after the point of deviation. (b) Angle curvature of the extracted contour. The peaks indicate the concavities due to contour deviation.

of the pectoral muscle extraction.

Search path technique

Similarly as for the breast contour refinement (see subsection 3.5), the pectoral muscle contour fitting approach developed in this work is based on the idea introduced in [6]. This approach relies on the definition of a search path for each point of the estimated pectoral contour. A search path is defined as the orthogonal line segment centered on its corresponding contour point. The search path slope is obtained straightforwardly from the previously computed curvature angle by adding 90°. The length of the search path is an odd integer denoted by M. In our experiments, it was set to 15 pixels. For contour points near ROI borders, search paths exceed the ROI limits. In these cases, they are rotated accordingly so that they lie within the breast area (see figure 4.9(b)). Having obtained these paths, we make the following reasonable assumption: the correct pectoral muscle contour point belongs to the search path, one just has to find out which one it is. In [6], the intensity profile along the search path is modeled as a sigmoid function. The pectoral muscle edge point is supposed to lie at the inflection point of the sigmoid. However, it is hard to model the intensity profile using a sigmoid in the area where the pectoral muscle overlaps glandular tissues because there is no sharp gray level intensity variation within these areas. Consequently, the refined contour may deviate from the true edge.

Average gradient

Instead, we propose to use a refinement method based on local maximum average gradient search that was introduced in [15]. This method is based on the fact that the gradient is known to be higher at the edge of two regions. Let the path around

contour pixel $\mathbf{p}$ be denoted by $\mathcal{P}_{\mathbf{p}}$. Suppose we have $\mathcal{P}_{\mathbf{p}} = \{\mathbf{p}_1, .., \mathbf{p}_M\}$ and for any $j \in \{1, .., M-1\}$, $\mathbf{p}_j$ and $\mathbf{p}_{j+1}$ are in 8-connectivity. In particular, this implies that $\mathbf{p}_{\frac{M-1}{2}} = \mathbf{p}$. The average gradient AG is computed for each element of $\mathcal{P}_{\mathbf{p}}$ along the search path as follows:

$$AG(\mathbf{p}_j) = \frac{2}{M-1} \sum_{1 \leq k \leq \frac{M-1}{2}} \frac{u(\mathbf{p}_{j-k}) - u(\mathbf{p}_{j+k})}{2k} \tag{4.8}$$

Figure 4.9 displays a mammogram in which two profiles (in red) were used to compute the standard gradient (c and d) and average gradient (e and f) respectively.

Thanks to an underlying smoothing effect, the average gradient is robust to spike structures observed in the gradient of a profile. It clearly features the location of edges with a prominent peak even in boundary areas with low pixels intensity variations and thus make it easier to retrieve actual contour points.

Pectoral muscle contour refinement

The selected contour point of the pectoral muscle is the one whose average gradient value is maximal among pixels in the search path. The refinement process is iterative. Actual contour points are replaced by those obtained with search paths. It was observed in some cases that shortening the search path during boundary refinement as described in [15, 6] does not produce accurate contour and more iterations were needed to obtain a good one. Therefore, instead of shortening the search path we used a fixed length search path and performed iterations until a convergence of contour points. The convergence is obtained when, from one iteration to another, the Hausdorff distance between the previous contour and the current one is less than a threshold (2 pixels).

At the convergence of the refining process, the set of points detected as true pectoral edge pixels is obtained. Figure 4.10 depicts the steps of the refinement process starting with the search path technique (a), followed by estimated edge pixels (b) and the final pectoral muscle contour (c). The contour points are smoothed within a sliding window to filter outlier points. The smoothing is performed by a local weights least square fitting. This least square fitting is implemented as follows:

- only data points within the the sliding window are used to calculate the coefficient of the fitting curve,
- a linear least square regression is performed using a first degree polynomial model.

After the smoothing process, it may happen that a row in the area of pectoral muscle has more than one edge point or no edge point at all. If two contour points are detected in a ROI row, then the pixel with the maximum gradient is chosen as pectoral edge pixel.

For no edge pixel in a row, the pectoral pixel is estimated to lie at the same column as the one in the previous row. The pectoral region is finally completely segmented after an extrapolated (tangent) straight line is added at the lower extremity of the pectoral muscle contour in case it does not reach the chest wall.

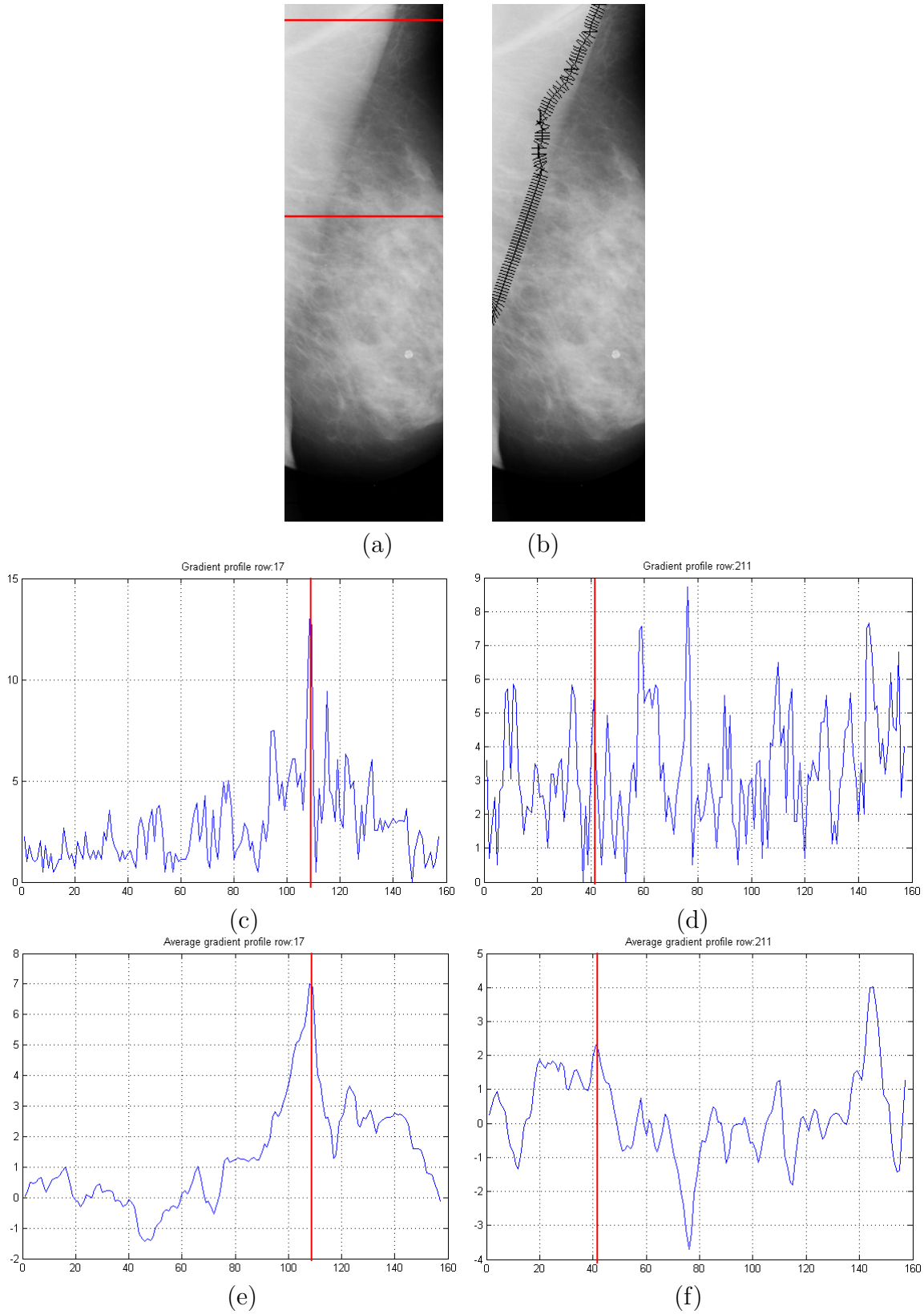
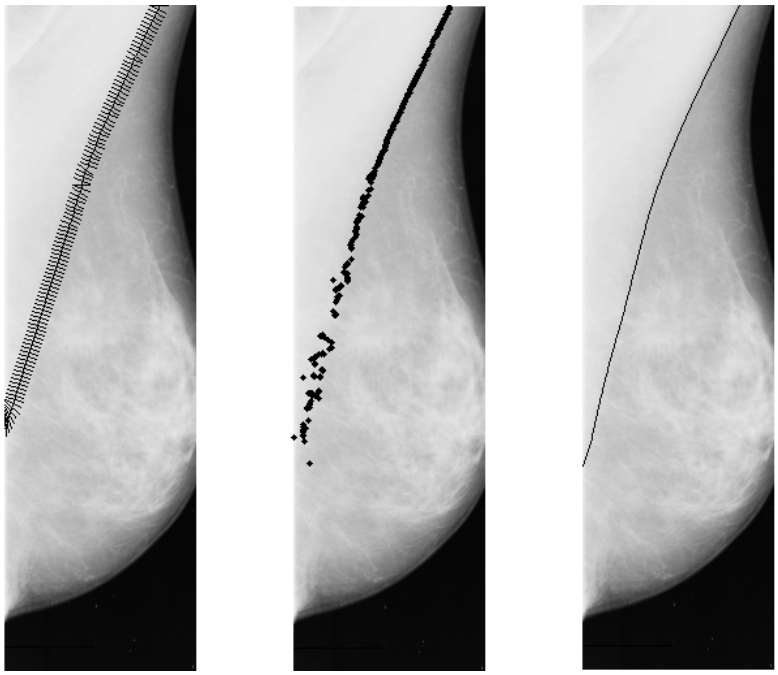


Figure 4.9: Search path technique and gradient plots for accurate boundary detection. (a) Initial image, red lines indicate profile selected to compute standard gradient and average gradient. (b) Validated contour with search paths drawn every three contour points, (c-d) Standard gradient and (e-f) average gradient plots computed along the selected profiles. Vertical red lines indicate the coordinate of the true boundary position on each profile.



(a) (b) (c)

Figure 4.10: Pectoral muscle contour refinement. (a) Validated contour with search paths drawn every three contour points, (b) True pectoral edge points detected after boundary refinement, (c) Final pectoral muscle contour obtained after smoothing the true pectoral edge points.

Figure 4.11 presents a detailed flowchart of the approach introduced in this PhD manuscript to tackle the problem of automatic pectoral muscle segmentation. This flowchart shows the different steps carried out to address this issue and which can be sorted into the following main tasks: Pre-processings, segmentation and refinement.

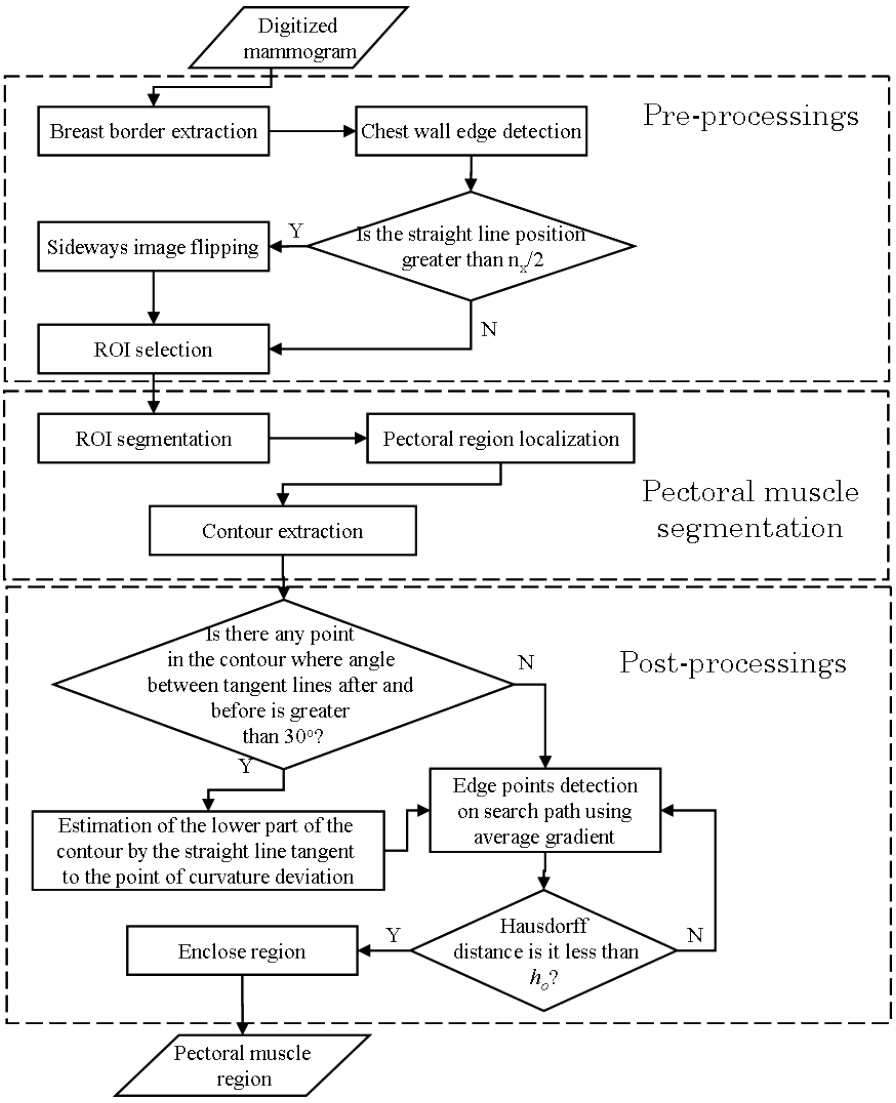


Figure 4.11: Detailed flowchart presenting the various steps of the strategy perfomed to successfully segment out the pectoral muscle in mammograms

4.4 Results and performances evaluation

Now that we have presented our comprehensive approach for pectoral muscle segmentation, it is necessary to validate its performances through several experiments on digital mammograms. This section starts with some general comments on the dataset and experimental protocol. A number of quantitative and qualitative performances then follows.

4.4.1 Dataset

The algorithm developed to segment out the pectoral muscle on mammograms was tested on images selected from the same dataset as the one used in the previous chapter: the mini-MIAS dataset [27].

This dataset is made of 322 images. This open database was chosen because it is the most widely used in the literature dealing with pectoral muscle extraction. It therefore allows an easy comparison of our results with those of other methods. All images

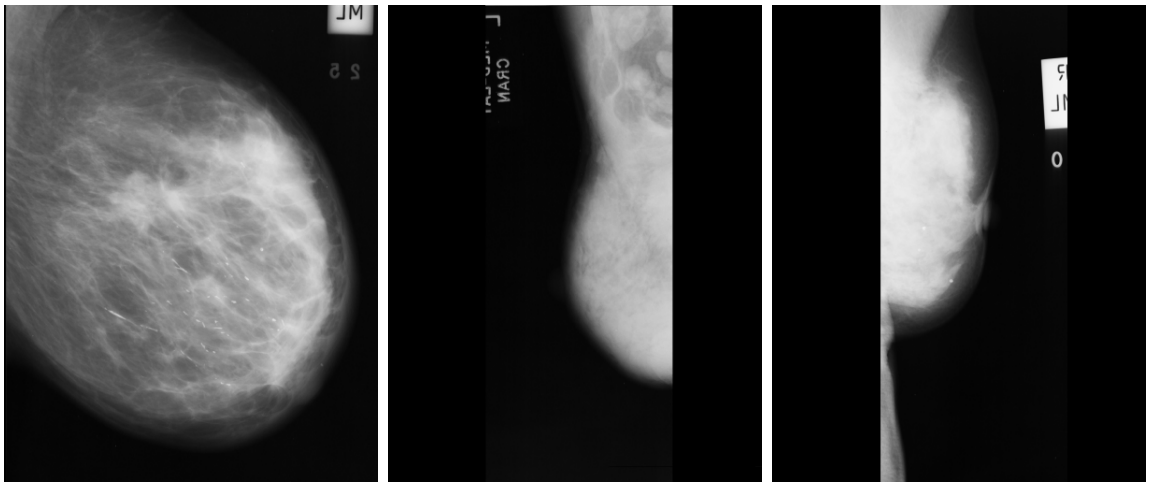


Figure 4.12: Some examples of mammographic images of poor quality. On the left, case mbd148 depicts a pectoral muscle with gray level intensity similar to those of breast tissue In the middle case mdb061 is an illustration of poor mammogram contrast quality with high variability in texture of the pectoral muscle. On the right, case mdb288 shows strong overlap of pectoral muscle tissues with breast dense tissues.

in this database are MLO view mammograms digitized at 200 μm with an 8-depth resolution. The images are 1024×1024 and were further sub-sampled to 512×512. Among the images of the database, the radiologist identified three of them (mdb098, mdb137, mdb236) as not containing pectoral muscle tissues. On other hand, two others (mdb313 and mdb314) are such that pectoral muscle region covers a little area made of only few pixels. The pectoral muscle in these cases do not induce a significant influence on the CAD and accordingly these cases were discarded.

Images having a super-impressed area or band tape (12 cases) were pre-processed to remove such areas which are harmful to the automatic extraction of the pectoral muscle. The super-impressed area was clipped out and the bottom of the image was zero-padded so that the final image size remains identical to the initial size. However, some images of the dataset are of very poor contrast quality and do not fulfill the assumptions made to track down pectoral muscle such as high gray level intensity or showing strong overlap with dense breast tissues. Figure 4.12 shows some examples of poor images quality. Following this pre-selection policy, only 22 images out of 322 were discarded.

4.4.2 Radiologist's expertise as reference standard

To evaluate the performances of the pectoral muscle extraction, a manually delineation of the pectoral muscle by an expert radiologist was used as reference standard. The radiologist manually drew the pectoral muscle contour of all mammograms in the dataset. It should be noticed that although the pectoral muscle was present in some images, the radiologist did not feel confident with his manually delineated pectoral boundary due to poor contrast quality or strong overlapping of the pectoral muscle with breast glandular dense tissues. This consequently induces a variability of the reference standard with respect to the observer. To take into account the influence of such a variability of the reference standard on the performances of the algorithm, an inter-observer analysis was carried out.

The coordinates of the radiologist's hand-drawn boundaries of pectoral muscle of images used in [5] as well as coordinates of pectoral muscle obtained by their algorithms were kindly provided by the authors. These data enable us to carry out a comparison of their results to our own on the same reference standard. Subsection 4.4.5 highlights the strong correlation between the two hand-drawn contours. The performances of the proposed method is assessed on several criteria as defined in the next section.

4.4.3 Performance criteria

The same criteria as in subsection 3.6.3 of chapter 3 are used. If the reader has kept these definitions in mind, this subsection can be skipped.
The accuracy of the pectoral muscle region extraction is assessed through the following performance ratings:

- False Positive (FP) and False Negative (FN) rates, which evaluate pixel assignment errors. A FP pixel is one assigned by the algorithm as belonging to the pectoral muscle region but assigned by the radiologist outside of the pectoral muscle region. A FN pixel is one assigned outside of the pectoral muscle region by the algorithm but assigned inside by the radiologist. The FP and FN rates for an image are computed as follows:

$$FP = \frac{\mid A_C \cup A_R \mid - \mid A_R \mid}{\mid A_R \mid}$$

$$FN = \frac{\mid A_C \cup A_R \mid - \mid A_C \mid}{\mid A_R \mid}$$

  where A_C and A_R are the areas of the pectoral muscle regions obtained by the algorithm and the radiologist respectively.

- Percent Overlap Area (POA), which expresses the accuracy of the region delineated by the algorithm to the reference one drawn by the radiologist. It is defined as:

$$POA = \frac{|A_C \cap A_R|}{|A_C \cup A_R|}$$

- The Hausdorff distance d_H, which is a metric assessing the dissimilarity of the detected boundary by the algorithm $\mathcal{C}_{\text{algo}}$ with the one drawn by the radiologist $\mathcal{C}_r$. It is defined as:

$$d_H\left(\mathcal{C}_{\text{algo}}, \mathcal{C}_r\right) = \max\{\max_{c_i \in \mathcal{C}_{\text{algo}}}\{\min_{r_j \in \mathcal{C}_r}\{d(c_i, r_j)\}\}; \max_{r_j \in \mathcal{C}_r}\{\min_{c_i \in \mathcal{C}_{\text{algo}}}\{d(r_j, c_i)\}\}\} \tag{4.9}$$

  where c_i and r_j are contour points obtained by the algorithm and the radiologist respectively, $d(c_i, r_j)$ is the Euclidean distance between points c_i and r_j.

These metrics are widely used and thereby allow a fair comparison with other works found in the literature. These metrics are also complementary and can thus catch every aspect of image segmentation performances.

In [5, 7], the segmentation was evaluated using performance metrics computed using area normalized errors. Normalized errors can be considered as partially biased to some extent. Normalized errors based criteria will thus be computed only for comparing our approach with theirs. In the sequel of this manuscript, unnormalized errors will be used.

The next sections present the results yielded by the proposed method tested on images from the MIAS dataset, the discussion and comparison of results with those of related works and the limitations of the proposed method in face of difficult cases.

4.4.4 Pectoral muscle segmentation: detailed results

Figure 4.13 shows some pectoral muscle extractions obtained with the proposed method. In comparison with radiologist's manually drawn pectoral muscle boundary, the following results can be highlighted:

- The POA mean and standard deviation are 86.30±13.19.
- The FN mean and standard deviation are 11.07 ± 12.49, while the FP mean and standard deviation are 3.58 ± 9.20.
- The Hausdorff distance mean and standard deviation are 15.00±16.00 mm.
- 55.93% (165/295), 80.34% (237/295) and 89.15% (263/295) of the computed pectoral muscle edges had a POA greater than 90%, 80% and 70% respectively.
- 27.46% (81/295) and 52.20% (154/295) of the computed boundaries had Hausdorff distances within 5 mm and 10 mm from the reference boundaries respectively.

Identifying the pectoral muscle becomes difficult if dense tissues appear near the pectoral muscle. Consequently, as explained in subsection 4.3.7, a strategy to cope with contour deviation and refine the estimated contour is proposed to track down the true pectoral muscle boundary. The evolution of POA performances thanks to boundary refinement is presented on figure 4.14.

One can notice that the accuracy of the proposed algorithm has been significantly improved through boundary refinement. For instance, the rate of images with inaccurate ($POA < 80\%$) estimation of pectoral muscle drops from 41.02%(121/295) to 19.66%(58/295) while the one with high accurate estimation ($POA > 95\%$) raises from 0.31%(1/295) to 18.98%(56/295). In addition, the Hausdorff distance drops from 23.32 ± 18.30 mm before refinement process to 15.00±16.00 mm after refinement.

The inaccurate images obtained before boundary refinement were mainly those having strong overlapping between pectoral muscle and glandular tissues or those with some artifacts or having non uniform gray level intensities in the pectoral muscle area. The above results prove that the pectoral muscle contour fitting step improves the accuracy of the pectoral muscle extraction in those cases.

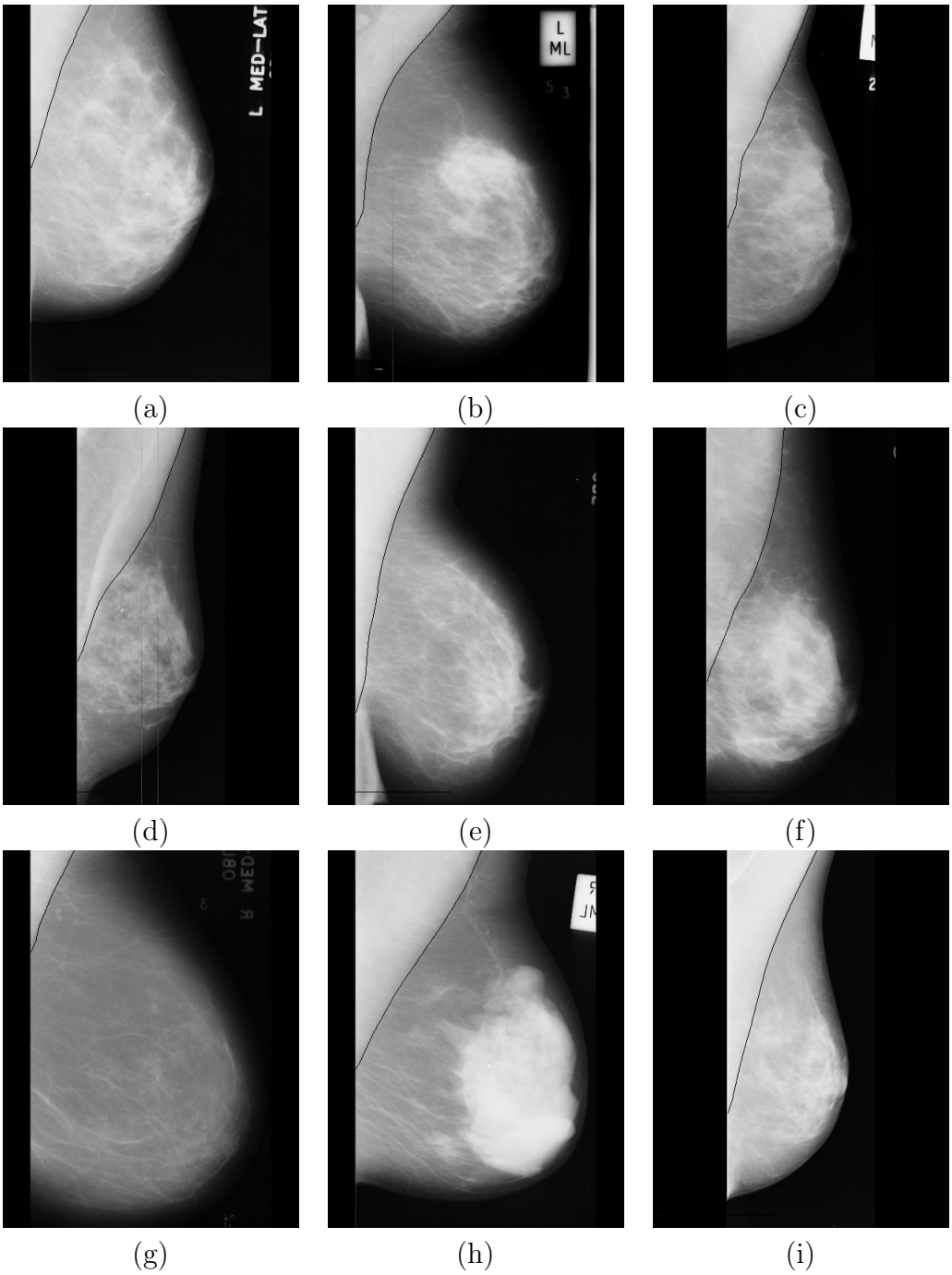


Figure 4.13: Some examples of pectoral muscle boundaries detected by the proposed method and superimposed on mammograms. (a) mdb029, (b) mdb111, (c) mdb113, (d) mdb185, (e) mdb217, (f) mdb221, (g) mdb272, (h) mdb320 and (i) mdb191.

4.4.5 Observer variability influence

Regarding the subset of 84 images used in [5], two hand-drawn pectoral boundaries were available. In the first case, the pectoral boundaries coordinates from a radiologist R2 were those used in [5] and provided by Rangayyang while in the second case, the pectoral boundaries correspond to that drawn by our radiologist R1. These two reference standards were used to study the effect of the reference standard variability on the performance evaluation. The same experiment was carried out on both our results and those of [5] as the outputs of their algorithm were provided. Note that in one very

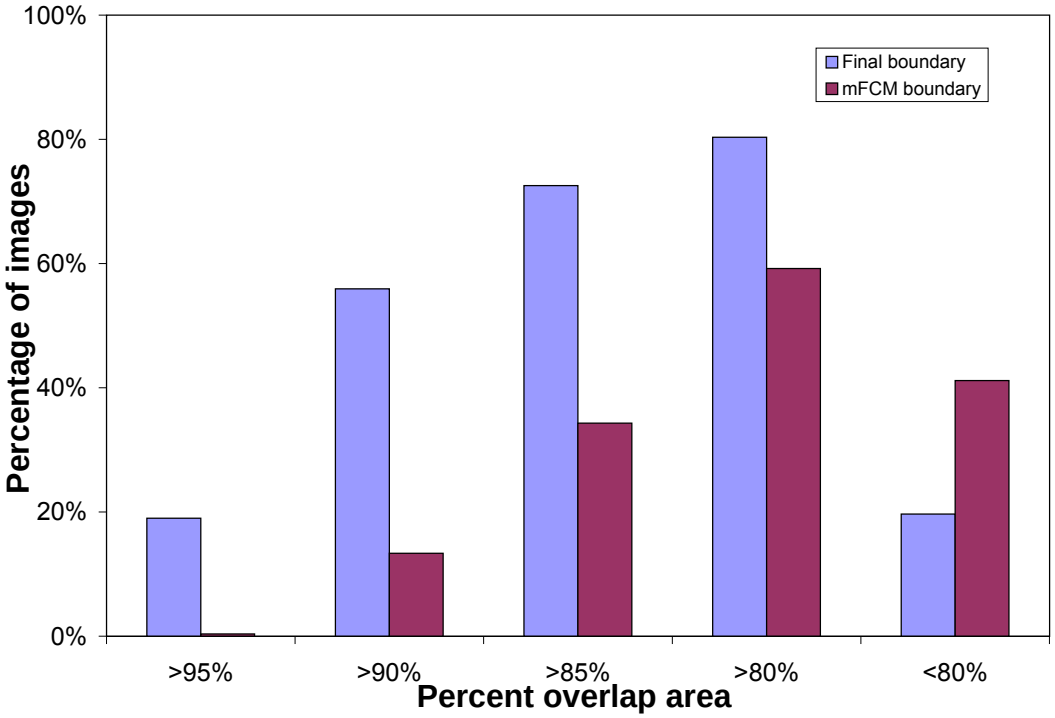


Figure 4.14: Histogram showing the proportion of images with percent overlap area (POA) between the computed and the reference standard pectoral muscle region before and after boundary refinement.

particular image, the radiologist R2 identified a very small region as pectoral muscle while the radiologist R1 did not identify any. This is explained by the fact that the pectoral muscle region is composed of a very few pixels and consequently radiologist R1 considered it unworthy to delineate. This image was therefore discarded from the following study.

Table 4.1: Effect of the variation of reference standard on the performances of our algorithm (mFCM) and the approach of [5] (Gabor)

		R1vsR2	R1vsmFCM	R1vsGabor	R2vsmFCM	R2vsGabor
POA	Mean	95.58	**89.12**	81.98	**87.95**	84.27
(%)	Std	2.12	**14.09**	15.13	**14.66**	16.02
d_H	Mean	3.89	**9.30**	10.51	**8.50**	12.47
(mm)	Std	7.37	**8.60**	19.46	**14.06**	21.97

Table 4.1 gives pairwise performances computed from hand drawn pectoral muscle boundaries of radiologist R1 and R2, the computed boundaries of our method and the ones obtained in [5]. From this table, it can be noticed that the proposed approach yields better results for both POA and Hausdorff distance criteria in comparison to that introduced in [5]. This preliminary observation tends to show that the method introduced in this PhD manuscript is the less susceptible to observer influence. The Mc Nemar's test was further used to assess the statistical significance of the performances

of these two approaches. The images fulfilling the accuracy criterion (POA>90%) for both reference standards in one method but not fulfilling it with the other are 45 (54.22%) for mFCM and only 4 (4.82%) for the method introduced in [5]. The Mc Nemar's value is 34.31 which compared to χ^2 indicates that the mFCM improvement is statistically significant with a p-value less than 0.01.

The Wilcoxon's test was used to evaluate observer variability of each method given that in this case data are paired. Using POA criteria, the test values are 2.697 for mFCM and 6.320 for the method introduced in [5] which are compared to the threshold test value 2.58 having p-value 0.01. These results indicate that both methods have some significant variation of their mean to each reference standard though mFCM seems to be less sensitive than Ferrari *et al.* method [5].

In figure 4.15, our approach and [5] approach are further compared. For each method and each radiologist, we give the cumulative percentages of images with performance measures greater than a given value x. These percentages are represented as functions of x. For the POA criterion, the proposed method performs better than Ferrari et al. [5] method for both reference standards. However, for d_H criterion, the performances of the two methods are quite close with minor advantage to the method introduced in [5].

4.4.6 Comparison to related works

We will now compare our approach to a wider spectrum of related works. The approaches selected for this comparative study are based on:

- Hough Transformed and Gabor filter introduced in [5],
- adaptive pyramids (AP) and minimum spanning trees (MST) presented in [7].

In their papers, the authors reported only FN and FP related performance rates. Also, these values were computed on the same data-subset as initially selected in [5]. This subset is made of 84 images from the MIAS dataset.

To allow a fair comparison, we computed the same performance measures on the same data-subset. Only one image was removed from this data-subset in our case. This image is the one without any pectoral muscle tissue reported by radiologist R2. All results are given in table 4.2.

From Table 4.2, one can notice that the approach using Gabor wavelets exhibits the lowest rates of FP and FN. The proposed method exhibits competitive performances as compared to methods based on APs and Hough transform. Absolute performances values are not the only quality criterion that matters. Indeed, the robustness of a pectoral muscle segmentation algorithm is also very important. The robustness of an algorithm is its ability to avoid producing very bad results. Robustness can be assessed by the counts of images for which both FP and FN are below a given threshold value. The rate of images for which both FP and FN are large (rows 5 and 6 of Table 4.2) is smallest for the proposed method. In conclusion, our approach is the most robust method among the four compared ones.

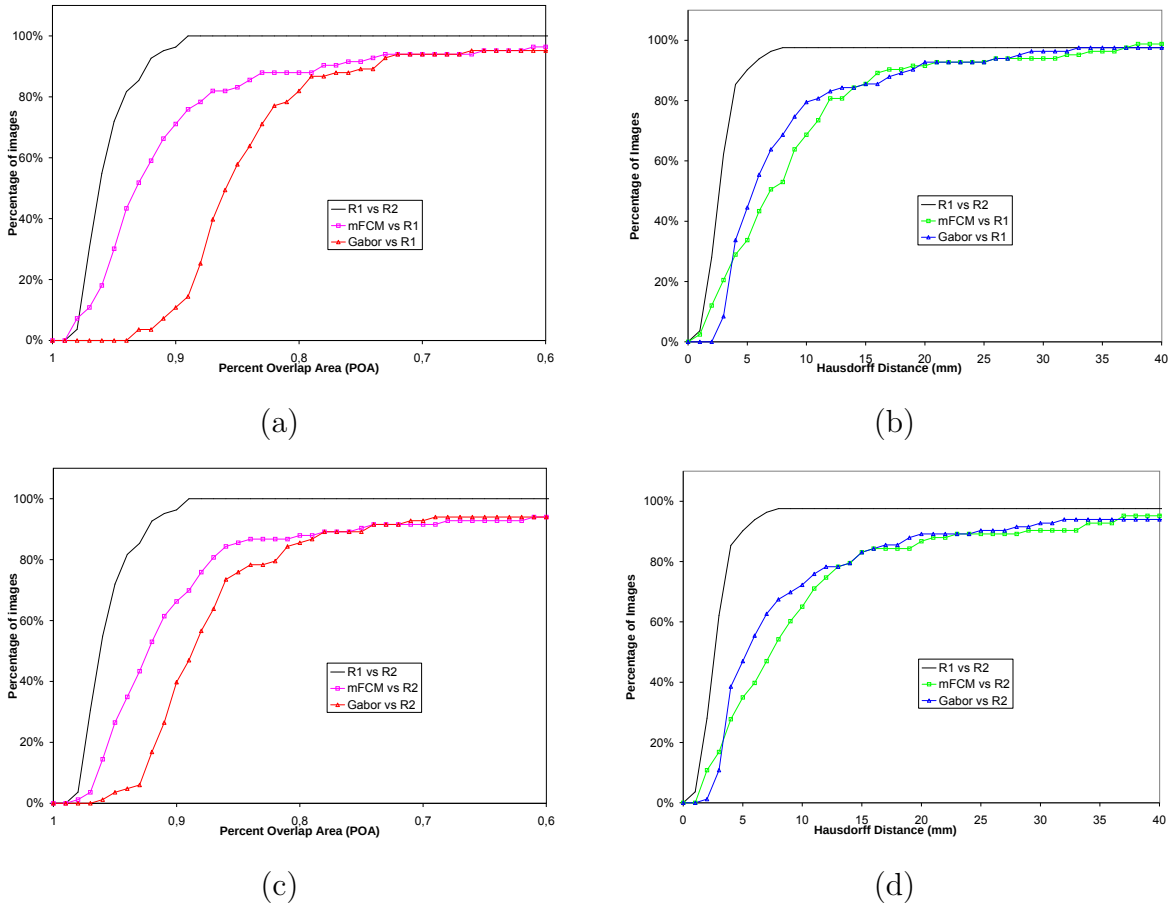


Figure 4.15: Cumulative percentages of images with performance metrics greater than a given value. Percent overlap area (a), (c) and Hausdorff distance (b), (d) performances of pectoral extraction using Gabor wavelets [5] and the mFCM compared to two expert reference standards R1 and R2 respectively.

The above conclusions drawn from this comparative study are unfortunately limited to a selected set of images. There is some uncertainty regarding the ability of these results to generalize to other mammographic images. Due to lack of time, it was preferred to focus on the development of advanced analysis tools presented in the next chapter rather than re-programming each of these methods in order to run them on other datasets.

4.4.7 Difficult cases and limitations

Like any image analysis approach, our algorithm is subject to some limitations and it is important to report them. Unsurprisingly, the worst performances are observed for images that do not match some of the assumptions made to justify our approach. These assumptions on the pectoral muscle region are concerned with:

- its size which is supposed to cover a significant proportion of the ROI,
- its average gray level intensity which is supposed to be higher than that of breast tissues,
- its texture which is supposed to be uniform.

Table 4.2: Comparison of reported studies using several performance metrics: mean and standard deviation and the percentage of images with the conditions set in the reported studies, where FP and FN rates are considered to be accurate ($< 5\%$), acceptable (5% to 10%) and unacceptable ($> 10\%$).

Performance criterion	Hough	Gabor	AP	MST	mFCM
FP (Mean $\pm$ Std Dev)	1.98 ± 6.09	$\mathbf{0.58 \pm 4.11}$	3.71	2.55	2.58 ± 6.43
FN (Mean $\pm$ Std Dev)	25.19 ± 19.14	$\mathbf{5.77 \pm 4.83}$	5.95	11.68	8.78 ± 13.95
$FP < 5\%$ and $FN < 5\%$	10	45	**50**	40	38
$5\% < FP < 10\%$ and $5\% < FN < 10\%$	8	22	**0**	**0**	1
$FP > 10\%$ and $FN > 10\%$	66	17	5	3	**1**

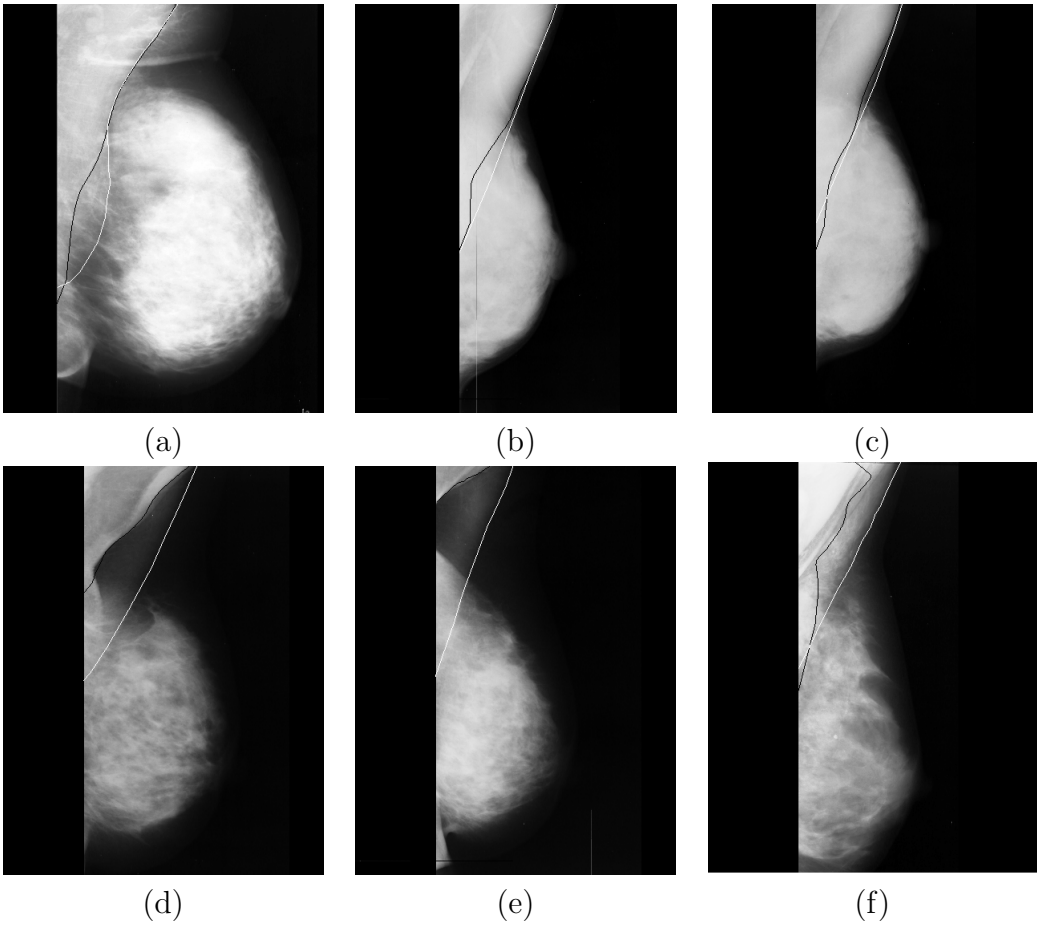


Figure 4.16: Some examples of pectoral muscle boundaries detected by the proposed method (black) and hand-drawn by the radiologist (white) superimposed on the mammograms. Cases with strong overlap of glandular tissue (a) mdb240, (b) mdb053, (c) mdb054 and cases with skin bulges or non uniform texture (d) mdb066, (e) mdb065, (f) mdb034.

Figure 4.16 presents the outputs of our approach on very difficult cases. In some cases with strong overlap with glandular tissues or non uniform contrast, the proposed method gives moderately good pectoral muscle extraction results (see figure 4.16(a-c)). This is mainly due to the contour validation step that relies on the upper part of the muscle contour where there is no overlapping. The refinement of contour in these cases

is effective only in the area where the overlap is not very strong or where the contrast is acceptable.

In presence of severe artifacts like prominent skin bulges or poor contrast, the algorithm fails to produce accurate results (see figure 4.16(d-f)). The main cause of this failure lies in the segmentation step. In these cases, the contour obtained after segmentation mainly fits the skin bulge contour instead of the pectoral muscle contour.

Although some significant deviations between the computed boundary and the actual one can be seen, the proposed solution is never completely absurd. Indeed, even in these cases, a very large amount of mammary tissues are correctly recognized. Our primary goal is to detect such tissues as part of the global CAD approach depicted in section 2.7.

Besides, the limitations discussed in the paragraphs above are mostly observed in particular cases of images where mammograms were not performed under optimal conditions. Carelessness on patient positioning during mammogram acquisition may result in blurred images with patterns like pectoral muscle partially occluded (see figure 4.16(d-e)). In addition, the difficult cases represent less than 5% of processed images on the whole MIAS dataset. The robustness of our approach is thus satisfying. Better results on such difficult cases could be achieved by using an additional computation efforts or by imposing more stringent conditions into mammographic examination protocols. In general, experts' knowledge could be incorporated inside the pectoral segmentation step using (semi)-supervised machine learning techniques. In this PhD, it was decided to stick to fast learning-free solutions in order to save as much computation resources as possible for the final stage of our CAD system, *i.e.* advanced dense breast tissues analysis.

4.5 Conclusion

Identifying the pectoral muscle region in mammograms is a challenging task due to various disrupting factors that make its automatic segmentation complicated. The main difficulties to overcome are the variability in shape, size, contrast and texture of the pectoral muscle region and the overlapping with dense glandular tissues. An automated approach of pectoral muscle extraction was designed by using some basic *a priori* knowledge and gray level intensity information to accurately delineate the pectoral muscle boundary. The motivation of this step is to discard pectoral muscle tissues from further analysis since they hinder CAD tasks like for instance mammographic density analysis and cancer risk assessment.

To achieve good pectoral muscle segmentation, we introduced a sequential approach which can be summarized as follows:

1. static ROI selection,
2. ROI segmentation using an improved version of the FCM clustering algorithm,
3. contour curvature correction,
4. contour fitting using average gradient and search paths that are orthogonal to the contour.

Our contribution lies in the combination of each of these steps as well as in their fine tuning.

Several experiments have been carried out on the MIAS dataset. Two radiologist hand-drawn pectoral muscle contours were used as reference standards to compute a wide range of performance measures. Regarding these measures, it can be concluded that our approach outperforms state-of-the-art approaches in terms of robustness.

Now as the pectoral muscle has been successfully removed in the mammographic image, the remaining image is only constituted of breast tissues ready for advanced tasks like tissue density characterization. The strategy presented in this chapter can also be useful as a pre-processing step of applications dealing with mammogram analyses like registration and breast deformation modeling which are other key steps in working out accurate CAD systems.

4.6 References

[1] C. K. Feudjio, J. Klein, A. Tiedeu, and O. Colot, "Automatic extraction of pectoral muscle in the mlo view of mammograms," *Physics in Medicine and Biology*, vol. 58, pp. 8493–8515, 2013. 61

[2] C. Feudjio, A. Tiedeu, M. Gordan, S. Domngang, and A. Vlaicu, "Computerized detection and smoothing of contour in mammograms," in *11th African Conference on Research in Computer Science and Applied Mathematics*, pp. 299–306, 13-16 October 2012. 62

[3] M. Karnan and K. Thangavel, "Automatic detection of breast border and nipple position on digital mammograms using genetic algorithm for asymmetry approach to detection of microcalcifications," *Computer Methods and Programs in Biomedicine*, vol. 87, pp. 12 – 20, 2007. 62

[4] L. Liu, J. Wang, and T. Wang, "Breast and pectoral muscle contours detection based on goodness of fit measure," in *5th International Conference on Bioinformatics and Biomedical Engineering*, pp. 1–4, 2011. 62

[5] R. J. Ferrari, R. M. Rangayyan, J. E. L. Desautels, R. A. Borges, and A. F. Frère, "Automatic identification of the pectoral muscle in mammograms," *IEEE Transactions on Medical Imaging*, vol. 23, pp. 232 – 245, February 2004. 62, 63, 64, 80, 81, 82, 83, 84, 85

[6] S. M. Kwok, R. Chandrasekhar, Y. Attikiouzel, and M. Rickard, "Automatic pectoral muscle segmentation on mediolateral oblique view mammograms," *IEEE Tran. Med. Imaging*, vol. 23, no. 9, pp. 1129–1140, 2004. 62, 69, 74, 75

[7] F. Ma, M. Bajger, J. P. Slavotinek, and M. J. Bottema, "Two graph theory based methods for identifying the pectoral muscle in mammograms," *Pattern Recogn.*, vol. 40, no. 9, pp. 2592 – 2602, 2007. 62, 64, 81, 84

[8] C. Zhou, J. Wei, H.-P. Chan, C. Paramagul, L. M. Hadjiiski, B. Sahiner, and J. A. Douglas, "Computerized image analysis: Texture-field orientation method for pectoral muscle identification on mlo-view mammograms," *Med. Phys.*, vol. 37, no. 5, pp. 2289–2299, 2010. 62, 63, 64, 69

[9] J. Suckling, D. R. Dance, E. Moskovic, D. J. Lewis, and S. G. Blacker, "Segmentation of mammograms using multiple linked self-organizing neural networks," *Med. Phys.*, vol. 22, no. 2, pp. 145–152, 1995. 63

[10] P. K. Saha, J. K. Udupa, E. F. Conant, D. P. Chakraborty, and D. Sullivan, "Breast tissue density quantification via digitized mammograms," *IEEE Transactions on Medical Imaging*, vol. 20, pp. 792 – 803, August 2001. 63

[11] D. Raba, A. Olivier, J. Martí, M. Peracaula, and J. Espunya, "Breast segmentation with pectoral muscle suppression on digital mammograms," in *Lectures Notes in Computer Science* (P. recognition and image analysis, eds.), 2005. 63

[12] M. Mustra, J. Bozek, and M. Grgic, "Nipple detection in craniocaudal digital mammograms," in *International Symposium, ELMAR '09*, pp. 15–18, 2009. 63

[13] H. Mirzaalian, M. R. Ahmadzadeh, and S. Sadri, "Pectoral muscle segmentation on digital mammograms by nonlinear diffusion filtering," in *IEEE Pacific Rim Conference on Communications, Computers and Signal Processing, PacRim 2007.*, 22-24 August 2007. 63

[14] Y. Li, H. Chen, and N. Yang, "Pectoral muscle segmentation in mammograms based on homogeneous texture and intensity deviation," *Pattern Recognition*, vol. 46, pp. 681 – 691, 2013. 63

[15] J. Chakraborty, S. Mukhopadhyay, V. Singla, N. Khandelwal, and P. Bhattacharyya, "Automatic detection of pectoral muscle using average gradient and shape based feature," *J. Digit. Imaging*, vol. 25, no. 3, pp. 387–99, 2012. 63, 64, 74, 75

[16] C. Chen, M. Nielsen, N. Karssemeijer, and S. S. Brandt, "Breast tissue segmentation from x-rays radiographs," *Physics in Medicine and Biology*, vol. 59, no. 10, pp. 2445 – 2456, 2014. 63

[17] N. Karssemeijer, "Automated classification of parenchymal patterns in mammograms," *Physics in Medicine and Biology*, vol. 43, pp. 365–378, February 1998. 64

[18] S. M. Kwok, R. Chandrasekhar, Y. Attikiouzel, and M. T. Rickard, "Automatic pectoral muscle segmentation on mediolateral oblique view mammograms," *IEEE Transactions on Medical Imaging*, vol. 23, pp. 1129 – 1140, September 2004. 64

[19] I. Domingues, J. S. Cardoso, I. Amaral, I. Moreira, P. Passarinho, J. S. Comba, R. Correia, and M. J. Cardoso, "Pectoral muscle detection in mammograms based on the shortest path with endpoints learnt by SVMs," in *Annual International Conference of the IEEE on Engineering in Medicine and Biology Society (EMBS)*, 31 August - 4 September 2010. 64

[20] K. S. Camilus, V. K. Govindan, and P. S. Sathidevi, "Computer-aided identification of the pectoral muscle in digitized mammograms," *Journal of Digital Imaging*, vol. 23, pp. 562 – 580, October 2010. 64

[21] K. Ganesan, U. R. Archarya, K. C. Chua, L. C. Min, and K. T. Abraham, "Pectoral muscle segmentation: A review," *Computer Methods and Programs in Biomedicine*, vol. 110, pp. 48 – 57, 2013. 64

[22] M. Mustra, M. Grgic, and R. M. Rangayyan, "Review of recent advances in segmentation of the breast boundary and the pectoral muscle in mammograms," *Medical and Biological Engineering and Computing*, vol. 54, pp. 1003–1024, July 2016. 64

[23] R.-C. G. Guo, S. sheng Ye, M. Quan, and H. xia Shi, "Modified fast fuzzy c-means algorithm for image segmentation," in *Second International Symposium on Electronic Commerce and Security*, vol. 2, pp. 39–43, 2009. 70

[24] K.-S. Chuang, H.-L. Tzeng, S. Chen, J. Wu, and T.-J. Chen, "Fuzzy c-means clustering with spatial information for image segmentation," *Computerized Medical Imaging and Graphics*, vol. 30, no. 1, pp. 9–15, 2006. 71

[25] F. Zhao, L. Jiao, and H. Liu, "Fuzzy c-means clustering with non local spatial information for noisy image segmentation," *Frontiers of Computer Science in China*, vol. 5, no. 1, pp. 45–56, 2011. 71

[26] T. Wang and N. Karayiannis, "Detection of microcalcifications in digital mammograms using wavelets," *IEEE Transactions on Medical Imaging*, vol. 17, pp. 498 – 509, August 1998. 71

[27] J. Suckling, J. Parker, D. Dance, S. Astley, I. Hutt, C. Boggis, I. Ricketts, E. Stamatakis, N. Cerneaz, S. Kok, P. Taylor, D. Betal, and J. Savage, "The mammographic images analysis society digital mammogram database," *Experta Medica International Congress Series*, vol. 1069, pp. 375–378, 1994. 78

Chapter 5

Breast density scoring using mammographic image processing

Contents

5.1 Introduction

Although mammography remains the standard modality used for breast cancer detection, carrying an automatic visual characterization of cancerous patterns remains a challenging task. Indeed, some cancer signs have similar textures as dense breast tissues and most of the time are located within these tissues. Looking for cancer signs in such conditions is tedious and complex to carry out even for expert radiologists. Medical research studies have shown that women with dense breast tissues have four to sixfold higher breast cancer risks. Since there is a strong correlation between breast tissue density and cancer risk, this feature is at the heart of many clinical protocols. In the next sections of the chapter, we start with a definition and a presentation of breast density classification systems. After this, we move to computerized analysis of mammograms for breast density assessment and present a state-of-the-art of such CAD systems as well as a preliminary novel approach to detect dense tissues using image processing.

5.2 Breast tissue density: General facts

Recent studies have shown that women having breast with a dense tissue ratio greater than 60% have an increased risk of breast cancer. Their cancer risk ranges from four

to sixfold higher to those of women with a dense tissue ratio[1] lower than 25% [1, 2, 3, 4]. This shows that there is a strong positive correlation between breast density in mammograms and the risk of developing breast cancer [4, 2, 5]. Other studies have shown that the number of cancer detected when reviewing the mammograms previously read by radiologists is always increasing [6]. This rate of missed cancers expresses the difficulties a radiologist faces to identify cancer signs when these latter are located within dense breast tissues.

Similarly, it has been shown that the efficiency of a CAD system to detect mammographic abnormalities is decreasing significantly as the density of the breast increases [7, 8]. The roots of these drawbacks lie in the design of CAD systems. The approaches commonly found in CAD systems are implemented to:

- process an entire mammographic image and identify breast cancer signs,
- compare pairs of mammograms and look for asymmetry in parenchymal tissue,
- analyze specific image patches selected by a radiologist and classify these latter as having cancer signs or not,
- integrate information from multimodal images and improve cancer signs detection.

Given that cancer signs are commonly depicted with similar texture as normal dense tissues, designing a CAD to search for cancer signs at broad image scale is bound to produce moderate results due the disparity in mammograms informative contents.

5.2.1 Definition and scope

Mammographic density analysis consists in characterizing the different types of tissues found in a mammogram according to their density related features. However, it should be made clear that segmentation or characterization of dense breast tissues allows cancer risk assessment but not the detection or the classification of abnormalities such as masses and microcalcifications. [7, 9, 10].

Noticing that dense tissues contain critical information on breast cancer, advanced strategies should be developed to address cancer sign detection conditionally to tissue density (with special care for highly dense breast tissue areas). Although CAD systems have been introduced to improve cancer detection rates, their performances to identify or discriminate cancerous tissues from normal tissues are still to be improved. In fact, to improve breast cancer detection, the design of a CAD system should take into account histological information during the process of cancer signs tracking.

In this chapter, we focus on breast tissue density characterization only. The outcome of density characterization approaches is a stepping stone toward the design of coarse to fine detection of breast cancer signs conditionally to tissue information. This approach sounds relevant to address afterwards the problem of cancer signs detection since they are hard to identify in dense regions when the whole mammogram is considered.

[1] The dense tissue ratio is the volume of dense tissues over breast volume. In general, this ratio is approximated from 2D mammograms as the surface (in pixel counts) of high intensity image region over the breast region surface.

5.2.2 Clinical assessment of the correlation between cancer risk and mammographic density

It is only during this last decade that computerized analysis of mammographic parenchymal patterns have been developed to automatically assess breast cancer risk whereas clinical studies have been carried out to examine the correlation between mammographic density and cancer risk for more than four decades. Clinical researches can be gathered into qualitative and quantitative studies. It should be noticed that some of these studies were carried out with respect to qualitative descriptions of mammographic parenchymal tissues that are subsequently introduced in section 5.3.

Qualitative studies

Boyd *et al.* [2] carried out a case control study to examine the relationship between mammographic patterns and breast cancer. They randomly collected mammograms of non-cancerous breast with mammograms of breast cancer control. This set of mammograms was submitted to radiologists for analysis and classification.

Gram *et al.* [11] used a meta-analysis framework to study the correlation between tissue density classification with respect to Wolfe and Tabár schemes and the following cancer risk factors:

- age,
- age at menarche,
- age at menopause,
- age at first child delivery,
- number of children,
- body mass index.

They obtained a strong correlation coefficient between tissue density and cancer risk.

McCormack and Silva [5] also carried out a meta-analysis on papers published on this issue and fulfilling some predefined criteria. The goal of this approach was to attempt to assess the origin of the disparity of the correlation coefficient between quantitative and qualitative studies of mammographic density with respect to cancer risk. They showed that the heterogeneity observed in the results is due to the difference in tissue density estimation. They therefore concluded that breast density is a marker that must be taken into consideration in etiology and prevention of breast cancer.

Quantitative studies

The earlier attempts to clinically estimate breast density quantitatively were carried out using planimetry approach [12, 13]. This method consists in delineating dense tissue regions in mammograms with an instrument called the *planimeter* and evaluate the surface of enclosed areas. The ratio of dense tissue is therefore derived with respect to the total breast region surface. Although this method seems straightforward to perform, difficulties arise in presence of "*islands*" of dense tissue as it is often the case with mammograms.

To cope with such difficulties a semi-automated approach was introduced by Byng *et al.* [14]. In this method, an observer selects gray level thresholds that produces a good segmentation into dense tissues areas and other breast tissues respectively. The surface of each region is computed and the dense tissue ratio is then derived. However, this approach relies on operator appraisal of dense regions and undoubtedly induce observer variability.

On other hand, volumetric methods attempt to determine the volume of dense tissue in the breast using a range of angular projections. Unfortunately, the heterogeneity observed from one breast to another is a major problem to overcome for a proper modeling. Consequently, such model is bound to produce results of moderate accuracy owing to the assumptions made. However, it should be noticed that promising new measures of mammographic density, including volumetric density, which can be standardized using full-field digital mammography, will likely result in a stronger risk factor and improve accuracy of risk prediction models [15]. More insights on quantitative and qualitative studies are presented in [16, 15].

These clinical studies show that there is a strong correlation between mammographic density and cancer risk. In addition, they pointed up the difficulties in estimating dense tissue ratio in mammograms. However, computerized approaches can provide consistent, quantitative, observer independent estimations of mammographic parenchymal patterns and therefore facilitate cancer risk assessment.

5.2.3 Justification with respect to cancer risk

In order to assess breast cancer risk, it is necessary to characterize and classify breast tissue with respect to their density. The availability of a breast cancer suspicion index as function of tissue density will thus allow to evaluate cancer risk level on the basis of breast tissue analysis. The determination of such a suspicion index relies on the characterization of dense tissues in mammograms. Although this task can be done manually, observer variability is a real problem when some radiologist's subjective appraisal is used to assess the amount of dense tissues in mammographic image [17].

An objective tool allowing tissue density assessment will be not only valuable for cancer risk evaluation but will also be helpful for monitoring tissues density evolution across time. In addition, tissue density is of a great asset in the search of cancer signs. It indicates cases with a high probability to develop cancer and therefore enables to implement advanced strategies to accurately detect cancer signs in such mammary tissues configuration [7].

To reduce observer variability issues, some classification schemes of mammographic density have been defined and are currently used as standard reference for mammographic risk scoring [9, 3, 10]. Such approaches are meant to provide radiologists with accurate and repeatable risk assessment. These clinical rating systems are not automated. They are presented in greater details in the next section.

5.3 Dense tissue related classification systems of mammograms

In this section, we introduce qualitative approaches that are commonly used as reference scale for classification of dense parenchymal tissues in mammographic images.

In this document, we use frequently use the terms "breast tissue density" and "dense tissues". In some literatures, the same concepts are sometimes referred to as "parenchymal patterns", "fibro-glandular disk" or "parenchymal density".

The classifications of mammographic density in order to assess cancer risk are based on a number of patterns that may not describe the same mammographic features [7, 10, 9, 4, 2]. The most common classification schemes found in the literature are those defined by Wolfe, Tabár and BIRADS systems which are introduced in the next paragraphs.

5.3.1 Wolfe's classification system

Wolfe was the first to advocate in 1976 for a strong correlation between parenchymal patterns seen in the breast and the risk of breast cancer [3, 4]. He therefore defined four classes of patterns to characterize mammograms of similar breast tissue textures. These classes are actually known as Wolfe's grades and are defined as follows:

- the **N** class pattern which represents fatty radiolucent tissues in the breast,
- the P_1 and P_2 class patterns which stand for a progressively increase fibrous tissues surrounding ducts,
- the **DY** class pattern which stands for a breast containing dense sheets of fibro-glandular tissues.

It should be noticed that the **N** grade is related to breast with the lowest breast cancer risk while P_1 and P_2 denote progressive higher risk and **DY** indicates the highest risk.

Some results in [2] show that some radiological appearance in mammograms are strongly related to breast cancer presence and this discrimination is compliant with Wolfe's mammographic pattern description. However, Wolfe's classification is sometimes criticized for its great disparity in breast cancer risk estimates and for its low reproducibility [18, 19].

5.3.2 Tabár's classification system

The Tabár's classification system is based on anatomic-mammographic correlations, following three dimensional histopathological comparison rather than a simple pattern reading like Wolfe's system. The mammographic patterns are associated with three selected breast cancer risk factors: parity, number of children and age of first child delivery. Tabár's system classifies mammographic features in patterns I to V.

- **Pattern I** is commonly made of three breast features which are:
  - scalloped contours and Copper's ligaments,

  - evenly scattered terminal ductal lobular units,
  - oval shaped translucent area corresponding to fatty replacement.

- **Pattern II** is usually formed from pattern I evolution and represents complete fatty replacement.
- **Pattern III** represents the combination of the retroareolar prominent duct pattern due to periductal elastosis and fatty involution.
- **Pattern IV** exhibits extensive nodular and linear densities throughout the breast.
- **Pattern V** consists of homogeneous, ground glass like, structureless fibrosis with convex contour.

In Tabár's classification system, patterns I to III indicate a low cancer risk while pattern IV and V point out to a high cancer risk.

In [11], Gram *et al.* noticed that the correlation coefficient between tissue density and cancer risk is different using Wolfe and Tabár classification schemes. Brisson *et al.* [20] obtained similar results when analyzing the relationship between Wolfe's mammographic parenchymal patterns and cancer risk. They showed that using Wolfe's grade description, mammograms might provide more information on breast cancer rather than what is provided by simply measuring the dense tissue ratio.

Another comparative study of Tabár's and Wolfe's system was carried out in [21]. Although no significant correlation was noted between the two classification schemes, the results derived from this comparison study indicate that Tabár's classification scheme can be considered as an evolution and improvement of Wolfe's classification [21]. However, the superiority of Tabár's system over Wolfe's system may be due to the integration of the three selected breast cancer risk factors into the classification process.

5.3.3 BIRADS classification system

The American College of Radiology (ACR) has defined a system of assessment of mammographic images density [9]. The Breast Imaging Reporting and Data System (BIRADS) was developed as quality assurance tool, and indicates the significant relationship between increasing breast density and decreasing efficiency in detecting cancer [22]. This system classifies mammographic images into four groups according to their ratios of dense tissues. Figure 5.1 gives an illustration of mammographic images from MIAS dataset with variable breast densities, each of them belonging to a given BIRADS category.

- BIRADS I represents almost entirely fatty breasts. In this category, the proportion of glandular tissue is less than 25%.
- BIRADS II contains breasts with scattered fibroglandular tissues. In this case, the proportion of dense tissues is between $25 - 49\%$.
- BIRADS III is the class of breasts with a ratio of glandular tissue in the range of $50 - 74\%$.

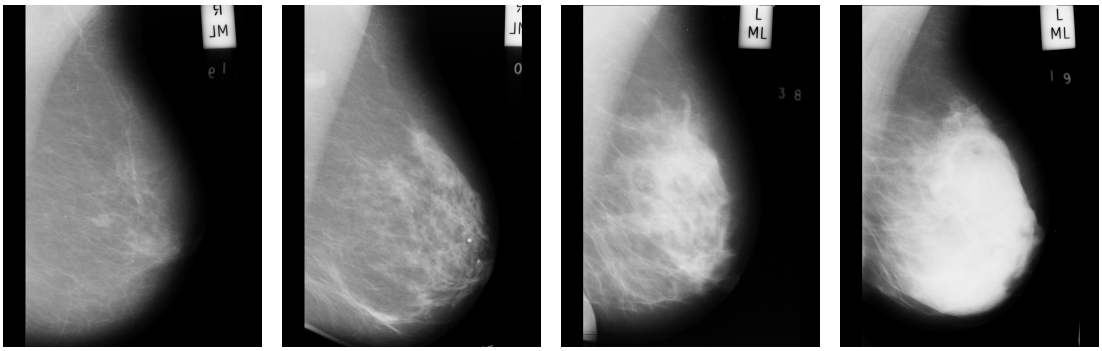


Figure 5.1: Mammograms with different amount of dense tissues from the 4 BIRADS categories. From left to right sample of mammograms of BIRADS I to IV and representing respectively low to highest mammographic risk.

- BIRADS IV is the class of breasts containing extremely dense tissue. In this case, the proportion of glandular tissue is greater than 75%.

It should be noticed that breasts belonging to BIRADS III group may tamper the sensitivity of mammography in detection of breast cancer signs whereas those of BIRADS IV configuration could obscure a cancerous lesion.

BIRADS I and II categories stand for a low cancer risk while BIRADS III and IV categories indicate higher risks of cancer. This qualitative system was not developed to quantify breast cancer risk, but to allow a radiologist to indicate his self-confidence with respect to false negatives due to lesions masked by dense tissues [16]. High BIRADS scores should imply further tests with modalities that are less affected by density [16].

Nowadays, the BIRADS system is becoming the radiological gold standard for breast density characterization and is intensively used in the evaluation of the automatic mammographic density and cancer risk assessment. This classification system will be therefore preferred for some developments presented in the next section.

5.4 Computerized dense tissue characterization: a state-of-the-art overview

Since it has been established that tissue density plays a major role in cancer risk assessment, designing automatic dense tissue characterization algorithms using image processing techniques is a relevant problem to address. This section presents an overview of dense tissue image analysis related works. Note that a greater research effort has been dedicated to another task: suspicious pattern detection and classification into abnormality classes or into benign or malignant types. A brief review on this concern is presented in [23]. This task is somehow related to the one we investigate in this section but it is widely agreed that they require very different image analysis solutions. As a consequence, we do not present mammographic abnormality detection approaches in greater details.

Mammographic density analysis mainly aims at scoring the amount of dense tissues for a given input image. Density scoring methods can be sorted into two categories:

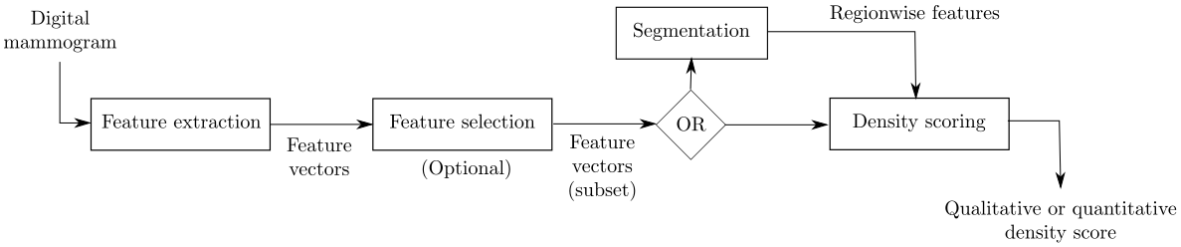


Figure 5.2: Main steps of a typical approach for mammographic density assessment

- segmentation free approaches: when general features are extracted from the image allowing some density related pattern recognition,

- joint segmentation and scoring approaches: when each pixel in the breast region is first sorted into dense and non-dense regions before making more general conclusions at the global image scale. These methods deliver more precise information on patients but are more computationally demanding.

The flowchart represented on figure 5.2 shows the main steps commonly found in related works dealing with mammographic density issues. Both types of density scoring approaches start with a feature extraction step followed by an optional feature selection. Feature vectors are then used for classification into density related classes or for regression if a numerical density score is sought. We briefly review this literature in the sequel, starting with an overview of feature extraction methods followed by a presentation of density scoring approaches for mammograms.

In the sequel of this document, any image is supposed be a function $u : \mathcal{G} \longrightarrow [0; 1]$ with $\mathcal{G}$ a domain often called image grid. An element $\mathbf{p} \in \mathcal{G}$ is called a pixel. A pixel $\mathbf{p}$ is a pair of coordinates (p_x, p_y) on the image grid. Gray level pixel intensities are thus given by $u(\mathbf{p})$. Black pixels are such that $u(\mathbf{p}) = 0$ while white pixels are such that $u(\mathbf{p}) = 1$.

5.4.1 Frequently used image features in mammograms

Feature extraction is a fundamental step in image analysis. In general, features are another representation of the same information as the one contained in the raw image. This new representation is usually more compact and allows an easier discrimination of the input data with respect to a given criterion, like class labels for instance. When a feature extraction is performed on a set of gray level values drawn from the neighborhoods of pixels, such features are called texture features. The simplest definition of the neighborhood of a pixel $\mathbf{p}$ is the set of pixels $\mathbf{p}'$ such that $\|\mathbf{p}' - \mathbf{p}\|_\infty < r$. Such neighborhoods are square subsets of the input image and are often called patches of radius r.

Raw patches can be considered themselves as feature vectors, in which case they are also called textons. Textons are used for mammographic image processing in [24, 25] for example. Note that in [25], patches are normalized by subtracting their means and dividing them by their standard deviations. The texture feature literature is very large and an exhaustive state-of-the-art is beyond the scope of this manuscript. In

this subsection, we only review those frequently used in mammographic (or sometimes medical) image analysis.

Raw features

The crudest form of texture feature are image patches as we have just explained. When patches are processed to yield another texture representation, the obtained features can be called raw features as opposed to representations where several cascading steps are performed. We present such raw features in the following paragraphs.

In mammographic image processing, one of the earliest texture features investigated are fractal features. Cadwell *et al.* [26] characterized mammographic parenchymal tissue using fractal analysis. In their approach, a mammogram is regarded as a surface with valleys in dark pixel regions and peaks in bright pixel regions. It was found empirically that the area A of this surface is given by the following equation:

$$A(\varepsilon) = \alpha \varepsilon^{2-\varphi} \tag{5.1}$$

where ε is the area of a small square, α is a scaling constant and φ is the *fractal dimension* parameter. The surface area A is a piece-wise affine approximation using constant squares whose size is ε. Intuitively, the surface approximation is better as ε decreases since one is able to capture finer variations. The fractal dimension φ does not depend on ε and thus captures some intrinsic properties of the surface ruggedness.

The fractal dimension can be numerically approximated by linear regression on a set of pairs $\{(\log(\varepsilon_i), \log(A(\varepsilon_i)))\}_{i=1}^{n}$. Several methods are available to compute values of $A(\varepsilon_i)$, see [27] for more details on this point. Note that if one is interested in segmenting mammographic images with respect to tissue types, the fractal dimension has to be computed locally on each patch. Consequently, one way to characterize texture is to analyze the local distribution of the fractal dimension using local histograms [28].

From Caldwell *et al.* [26] analysis, it has been argued that fractal dimension alone is not enough to accurately describe mammographic textures. A similar conclusion was draw by Chen *et al.* [29] to improve medical image texture classification. Among other possibilities, Lee *et al.* [30] investigate a multi-resolution approach based on wavelet decomposition of echographic liver images. The fractal dimension is computed for each scale considered in the decomposition thereby yielding a feature vector.

In general, multi-dimensional analysis of fractal properties is called multi-fractal analysis. Lopes and Betrouni [27] give a review of multi-fractal techniques for medical image processing.

Textbook examples of texture features are gray level histograms and co-occurrence matrices [31]. Both of these objects are obtained by counting occurrences of pixel intensity related events. The histogram $\mathbf{h}$ of a ROI $R \subseteq \mathcal{G}$ is a vector of size n whose i^{th} coordinate h_i is given by:

$$h_i = \frac{\left| \left\{ \mathbf{p} \in R \middle| \frac{i-1}{n} \leq u(\mathbf{p}) < \frac{i}{n} \right\} \right|}{|R|}. \tag{5.2}$$

The profile of vector $\mathbf{h}$ encodes some textural information.

A co-occurrence matrix is a distribution that characterizes the frequency of pairs of gray level values (u_1, u_2) whose ante-image pixels $\mathbf{p}_1$ and $\mathbf{p}_2$ are separated by a given offset $\Delta\mathbf{p} = (\Delta_x, \Delta_y)$. Let $\mathbf{C}_{\Delta_\mathbf{p}}$ denote a $n \times n$ co-occurrence matrix extracted from image u. We have:

$$C_{i,j,\Delta_\mathbf{p}} = \frac{|\left\{\mathbf{p} \in R | \frac{i-1}{n} \leq u(\mathbf{p}) < \frac{i}{n} \text{ and } \frac{j-1}{n} \leq u(\mathbf{p}+\Delta\mathbf{p}) < \frac{j}{n}\right\}|}{|\{\mathbf{p} \in R | \mathbf{p}+\Delta\mathbf{p} \in R\}|}. \tag{5.3}$$

Note that common choices for $\Delta\mathbf{p}$ are those compliant with 8-connectivity. Consequently, several co-occurrence matrices need to be computed to characterize textures. In [7, 32, 33] both histograms and/or co-occurrence matrices are used on mammographic images. Note that histograms or co-occurrence matrices are rarely considered as a final data representation (see next subsubsection).

Also, similarly as for fractal features, histograms and co-occurrence matrices can be computed for multiple image scales. He *et al.* [34] developed a multi-resolution approach to capture both small and large local texture structures.

Speaking of multi-resolution, another type of features emerged in the late 90's for texture analysis: wavelet coefficients. Using wavelet decomposition of images, multi-scale properties are naturally captured. Let ψ denote a function in $\mathcal{L}^2$ the vector space of square integrable functions over $\mathbb{R}^2$. ψ is a wavelet if the set of functions $\{\psi_{i,j,k}\}$ obtained by dyadic translations and dilations of ψ is a basis of $\mathcal{L}^2$. For any pixel $\mathbf{p}$, one has:

$$\psi_{i,j,k}(\mathbf{p}) = 2^{\frac{k}{2}}\psi\left(2^k\mathbf{p} - (i,j)\right). \tag{5.4}$$

The wavelet transform of image u is denoted by $\mathcal{W}_\psi\{u\}$ and is given by:

$$\mathcal{W}_\psi\{u\}(a,\mathbf{b}) = \frac{1}{\sqrt{a}}\int_\mathcal{G} \psi^d_{a,\mathbf{b}} u d\omega, \tag{5.5}$$

where $\psi^d_{a,\mathbf{b}}$ is a discretized version of a child wavelet function defined as $\psi_{a,\mathbf{b}}(\mathbf{x}) = \psi\left(\frac{\mathbf{x}-\mathbf{b}}{a}\right)$ for any $\mathbf{x} \in \mathbb{R}^2$ and ω is the discrete measure over $\mathcal{G}$ defined as $\omega = \sum_{\mathbf{p}\in\mathcal{G}} \delta_\mathbf{p}$. $\delta_\mathbf{p}$ is the Dirac measure on the singleton $\{\mathbf{p}\}$.

The value of this integral is a texture feature and, by computing it for several values of a and $\mathbf{b}$, a set of features (known as wavelet coefficients) is obtained. Of course, these features are depending on the chosen mother wavelet ψ. Common choices are Haar, Daubechies or Gabor wavelets. A review on wavelet analysis for biomedical signals is given in [35]. This review contains a few paragraphs on microcalcification detection in mammograms using wavelet coefficients. An example of dense tissue texture analysis using wavelets is [36]. A common difficulty with such coefficients is the dimensionality of the feature space which is potentially very high. Selecting a relevant subset of wavelet coefficients (or equivalently relevant set of values for a and $\mathbf{b}$) is a feature selection problem. This more general problem is briefly evoked in subsection 5.4.1.

Similarly as co-occurrence matrices, some other empirical distributions appear to be relevant local features for images because they jointly process spatial and gray level information. Histograms of oriented gradients (HoG) [37] are such distributions. This

feature extraction method starts with the computation of gradients using usual digital filters like Prewitt [38]. The grid is then partitioned into small patches like 4×4 ones for instance. For each patch, a histogram of local gradient orientations is computed. The orientation levels are limited generally to 8 or 9 possible levels in the interval $[0, 2\pi]$.

The HoG features are related to others belonging to the scale invariant feature transform (SIFT) [39]. This operator is meant to detect interest points inside an image. Typical interest points in an image are corner or edge pixels. In the SIFT setting, each interest point is associated to a texture feature vector very much alike HoGs. More precisely, SIFT feature vectors are the concatenation of HoGs extracted from neighbor patches.

Bosch *et al.* [24] compare those SIFT features with textons for a classification task into BIRADS grades. Surprisingly, textons outperforms HoGs for this task. The authors argue that dense and non-dense tissues cannot be disambiguated by HoG features because line orientations in such image region are rather anarchic.

Another kind of histograms were introduced by Zwiggelaar [40]: gray appearance (LGA) histograms. This approach relies on the computation of local signature s_{lga} of patches. This signature is obtained as follows:

$$s_{lga} = 1 + \sum_{\mathbf{p} \in \text{ patch}} (N_g)^{ind_{\mathbf{p}}} \, u_q(\mathbf{p}) \tag{5.6}$$

where $ind_{\mathbf{p}}$ is the index of pixel (i, j) in the patch and u_q is a quantized version of u in N_g gray levels. The way pixels of a patch are indexed is arbitrary but the same indexes must be used for each patch. Finally, the histograms of s_{lga} are computed over a given region of the image. For once, this feature extraction method was specifically introduced for mammogram analysis.

LGA signature can be viewed as the generalization of a popular texture signature known as local binary patterns (LBP) introduced by Ojala *et al.* [41]. The LBP signature is a LGA signature with only two levels, *i.e.* $N_g = 2$, computed on a shifted patch. A shifted patch is obtained by subtracting the gray level of the center pixel to all the other ones. This latter modification allows LBP signature to be rather invariant to illumination changes.

More recently, Crosier and Griffin [42] introduced a new empirical distribution for texture representation. In order to build this distribution, the authors propose to first compute the basic image feature (BIF) of each pixel. This feature is obtained by quantizing the responses of six filters. This set of filters is made of:

- a gaussian discrete filter g_σ,
- the two discretized first derivatives of g_σ: $\frac{\partial g_\sigma}{\partial p_x}$ and g_σ: $\frac{\partial g_\sigma}{\partial p_y}$,
- the three discretized second derivatives of g_σ: $\frac{\partial^2 g_\sigma}{\partial p_x^2}$, g_σ: $\frac{\partial^2 g_\sigma}{\partial p_y^2}$ and $\frac{\partial^2 g_\sigma}{\partial p_x \partial p_y}$.

Using this filter bank, each pixel is thus embedded in $\mathbb{R}^6$. This space is then divided into seven subsets for quantization. This quantization is performed using an interesting

geometric principle. Indeed, it can be shown that each of the six filter has a large output value when the image is locally invariant under a group of transformations. The authors investigate isometry with respect to both space and gray level intensities. Consequently, the filter responses allow to select the closest symmetry type among a set of seven predefined symmetric patterns. The authors comment on the fact that a crude histogram with respect to BIF categories is not precise enough for texture classification. They recommend using a vector of BIF labels for each pixels. Each component of these vectors is the BIF category for a given scale which is selected using parameter σ from the Gaussian filter g_σ. Histograms computed in this larger space produce better results.

Both BIF and LGA features are experimented for dense tissue characterization by Chen *et al.* [43].

Features are not always arranged in vectors. They can be organized in more complex structures. Chen *et al.* [44] developed a quantitative estimation of breast density using topographic maps. A topographic map $\mathcal{T}$ is a collection of level sets: $\mathcal{T} = \{s_1, .., s_n\}$. The i^{th} level set s_i is the subset of $\mathcal{G}$ containing pixel with intensities greater than $i \times \frac{\max_{\mathbf{p}\in\mathcal{G}} u(\mathbf{p})}{n+1}$. We thus have $s_0 = \mathcal{G}$ and $s_{n+1} = \emptyset$. Some level sets are not connected, in which case they are separated into connected subsets of maximal cardinalities. The family of subsets thus obtained is a first degree of feature representation denoted by $\{s'_i\}_{i=1}^{m}$.

Since the set-inclusion is a total order relation for such sets, a convenient representation for feature are directed trees. Let Tr denote the tree whose nodes are connected sets s'_i. Two nodes s'_i and s'_j are connected if $s'_i \subseteq s'_j$ and there is no other set s'_k such that $s'_i \subseteq s'_k \subseteq s'_j$. By detecting sets whose shape are significantly different from tree-ancestors but similar as tree-descendants, a sub-tree Tr' can be extracted in order to keep only the most interesting sets.

Trees Tr' are global feature for the whole image. Local features can be retrieved from these trees by computing a map $\vartheta : \mathcal{G} \longrightarrow \mathbb{R}$. This map is called density map and is such that $\vartheta(\mathbf{p}) = \frac{1}{|R|}\sum_{\mathbf{p}'\in R} u(\mathbf{p}')$ with R the smallest set in Tr' containing $\mathbf{p}$.

The feature extraction methods mentioned in the preceding paragraphs do not take into account the specificity of the X-ray imaging and mammogram acquisition protocol. Highnam and Brady [45] followed in a different path and introduced an algorithm that is meant to retrieve the actual depth (in millimeters) of dense tissues for every pixel of the mammographic image. Given acquisition related meta-data (X-ray tube voltage and current, exposure time and breast thickness), one can use physical laws to write an equation relating the photonic energy flowing to a pixel with dense tissue depth. Since the energy can be estimated from pixel intensity values [46], the only unknown variable left is the dense tissue depth. The main problem in this setting is that such meta-data are not always available. Images in DICOM format enclose such data in the file header. This feature provides radiologists with a physically sound measure.

Features of features

Feature extraction can be viewed as the action of a mapping from the space of raw images to the feature space. Sometimes, it may appear relevant to use function composition in order to access gradually higher degrees of abstraction in the data representation. In this case, several intermediate levels of features need to be computed. Such features can be called features of features, or compositional features.

An illustration of compositional features are descriptive statistics computed from histograms or co-occurrences matrices of features. [34, 7, 32, 33] Among such statistics, one can cite for instance:

- for intensity histograms: mean, standard deviation, skewness and kurtosis,
- for co-occurrence matrices: contrast, energy, correlation, sum average, sum entropy, difference average, difference entropy and homogeneity.

The mathematical expression of these features as well as the texture information they evaluate are given in table 5.1. They are used for dense tissue texture characterization in [34, 7, 32, 33].

There are of course many other approaches falling in the category of compositional features. Saha *et al.* [1] compute a fuzzy membership function based on several similarity functions between neighborhoods. Unlike many other approaches, the neighborhoods have variable shapes depending on which pixel is the neighborhood center. However, a function for determining neighborhood boundaries needs to be known *a priori*. Two different similarities are computed with respect to either neighbor pixel intensities or to intensities of reference labeled pixels. A third sub-component of the fuzzy membership function is given by pixel connectivity in the grid $\mathcal{G}$. The three sub-components are then aggregated using a multiplicative rule to yield the membership function. The membership value of a pixel can be viewed as a compositional feature.

Mapping features into another space can serve some more specific purpose than reaching more abstract representations. In particular, a first feature extraction step usually yields texture representation with a high dimensionality. High dimensionality is a problem as many algorithms (like classifiers) do not scale easily to very large inputs.

The most popular dimensionality reduction technique is the principal component analysis (PCA). This method consists in computing eigenvectors of the co-variance matrix of feature vectors. Feature vectors are then projected onto an eigen-subspace whose base vectors are those with the highest eigenvalues. This ensures that dimensions along which the data have little variance are eliminated. PCA is very efficient in the sense that usually many eigenvalues are negligible as compared to the largest one. PCA is used in a density scoring context in [33, 25, 36].

Intuitively, one needs to select the most discriminative features with respect to tissue types. Such approaches are called feature selection methods. PCA is not a feature selection method because dimensions with large variance are not necessarily discriminative ones. In the worst case, PCA may even delete the most valuable features.

One way to guide a selection method towards discriminative sub-spaces is to use la-

Table 5.1: Statistical features used in texture characterization derived from histograms and co-occurrence matrices (non-exhaustive list).

	Feature name	Mathematical expression	Feature interpretation
Features from histogram $\mathbf{h}$	Mean	$\mu = \sum_{i=1}^{n} i h_i$	Mean pixel intensity
	Variance	$\sigma^2 = \frac{1}{N}\sum_{i=1}^{n}(i-\mu)^2 h_i$	Heterogeneity of pixel intensities
	Skewness	$\mu_3 = \sum_{i=1}^{n}\left(\frac{i-\mu}{\sigma}\right)^3 h_i$	Asymmetry of pixel values around the mean
	Kurtosis	$\mu_4 = \sum_{i=0}^{L-1}\left(\frac{i-\mu}{\sigma}\right)^4 h_i$	Indicates if histogram is peaked or flat
	Smoothness	$R = 1 - \frac{1}{(1+\sigma^2)}$	Relative smoothness of a region
	Energy	$E = \sum_{i=1}^{n} h_i^2$	Gray levels uniformity
	Entropy	$H = -\sum_{i=1}^{n} h_i \log(h_i)$	Pixel intensity disorder
Features from co-occurrence matrix $\mathbf{C}_{\Delta_\mathbf{p}}$	Sum of squares	$VAR = \sum_{1\leq i,j\leq n}(i-\mu_x)^2 C_{ij,\Delta_\mathbf{p}}$	Texture heterogeneity
	Energy	$E = \sum_{1\leq i,j\leq n}(C_{ij,\Delta_\mathbf{p}})^2$	Texture uniformity
	Entropy	$H' = -\sum_{1\leq i,j\leq n} C_{ij,\Delta_\mathbf{p}} \log(C_{ij,\Delta_\mathbf{p}})$	Texture disorder in a image
	Contrast	$CST = \sum_{1\leq i,j\leq n}(i-j)^2 C_{ij,\Delta_\mathbf{p}}$	Variation in texture
	Correlation	$C = \sum_{1\leq i,j\leq n}\frac{ijC_{ij,\Delta_\mathbf{p}} - \mu_x\mu_y}{\sigma_x\sigma_y}$	Similarity in texture
	Absolute value	$A_V = \sum_{1\leq i,j\leq n}\lvert i-j\rvert C_{ij,\Delta_\mathbf{p}}$	Variation in texture
	Inverse difference moment	$IDM = \sum_{1\leq i,j\leq n}\frac{1}{1+(i-j)^2} C_{ij,\Delta_\mathbf{p}}$	Texture homogeneity
	Max probability	$\max_{i,j} C_{ij,\Delta_\mathbf{p}}$	Maximum probability for a given texture
	Sum average	$M_S = \sum_{i=1}^{2n-1} C_{i,x+y}$	Texture uniformity aggregate
	Sum entropy	$E_S = \sum_{i=1}^{2n-1} C_{i,x+y}\log(C_{i,x+y})$	Texture disorder aggregate
	Sum variance	$S_V = \sum_{i=1}^{2n-1}(i-SE)C_{i,x+y}$	Aggregate in texture heterogeneity
	Diff entropy	$\delta_H = -\sum_{i=1}^{2n-1} C_{i,x-y}\log(C_{i,x-y})$	Texture disorder variation

n is the size of histogram $\mathbf{h}$ or the number of rows/columns of matrix $\mathbf{C}_{\Delta_\mathbf{p}}$. μ_x, μ_y, σ_x and σ_y are the means and standard deviations of co-occurrence matrix marginals in the x and y-direction respectively. $C_{x\pm y}(k) = \sum_{1\leq i,j\leq n \text{ s.t. } \lvert i\pm j\rvert = k} C_{ij,\Delta_\mathbf{p}}$.

beled data. Indeed the very notion of discrimination is always defined with respect to a given underlying pattern. Feature selection can be done using a very wide range of methodologies although the mainstream approach is probably to optimize a given criterion. A typical criterion one wants to maximize is the mutual information between classes and feature values. Conversely, a typical criterion one wants to minimize is the mutual information between distinct features which stands for information redundancy among features.

Another possible feature evaluation method is to apply χ^2 test between a discretized version of the feature distribution and the class distribution. Mustra *et al.* [47] rank co-occurrence statistical features with respect to that criterion. They also compare different rank based methods to select a subset of features. Genetic search gives the highest recognition rate with only 9 features out of 18. Oliver *et al.* [7] applied a similar method known as sequential feature selection (SFS) to select statistical co-occurrence features. The selected features are used to train a classifier so as to predict BIRADS classes. In SFS, the criterion used to evaluate features is usually the correct classification rate.

Feature learning

A wide variety of mammographic texture descriptors have been presented in the preceding paragraphs. A full characterization of breast tissues seems difficult to attain by picking only some of these features. Using naively all these features together is also not acceptable as it leads to an extremely vast feature space making further computation untractable.

Recently, new approaches to automatically learn features from images have been introduced. This framework is very appealing in the sense that there is no arbitrary choices among feature extraction methods to make anymore. An adequate texture representation is driven from the data itself.

In a supervised context, a growing interest has been paid to convolutional neural networks (CNN) in the past years. CNNs were initially introduced by Lecun [48] to recognize hand-written decimal numbers in small 28×28 images of the MNIST dataset. This approach is defined within the usual neural networks formalism where a cost function $J(\boldsymbol{\theta})$ needs to be minimized with respect to the vector of model parameters $\boldsymbol{\theta}$.

CNNs are organized into several pairs of layers. Each pair of layers is made a convolution layer followed by a pooling layer:

- In a convolutional layer an input vector $\mathbf{x}$ of size n is mapped to an output vector $\mathbf{y}$ whose components are obtained by applying the sigmoid function sgm to an affine combination of the input vector components: $y_j = \text{sgm}\left(\sum_{i=1}^{n} \theta_{ij} x_i + \theta_{n+1}\right)$. The parameters that we want to learn are θ_{ij}, *i.e.* those involved in the affine combination. Concatenating all such parameters together yields the vector $\boldsymbol{\theta}$. Since discrete convolution resembles affine combination, this layer is called a convolutional layer. Parameters of this combination can be regarded as filter parameters, that is why we are learning feature representation automatically.

- The pooling layer is meant to reduce the dimensionality of output vectors. Usually, the one with the highest component value is selecting from a set of neighbor component values. Indeed, another interesting aspect of CNNs is that spatial information is propagated through the layers.

A CNN final layer is usually a softmax layer which allows to map the output of the preceding layer to a cost function which can be interpreted as a negative log-likelihood of training examples given $\boldsymbol{\theta}$.

On top of that, the sigmoid function allows to grasp non-linear structures. CNNs have demonstrated a very high potential on many visual classification tasks provided that a substantial number of layers are used. They are highly compositional models belonging to the family of deep learning methods.

Petersen *et al.* [49] implemented a CNN to analyze mammographic tissues. Note that the feature extraction part of a CNN is optimized with respect to the targeted classification task. It is consequently hard to say if the features thus obtained are a general representation of the tissue texture or a specialized one.

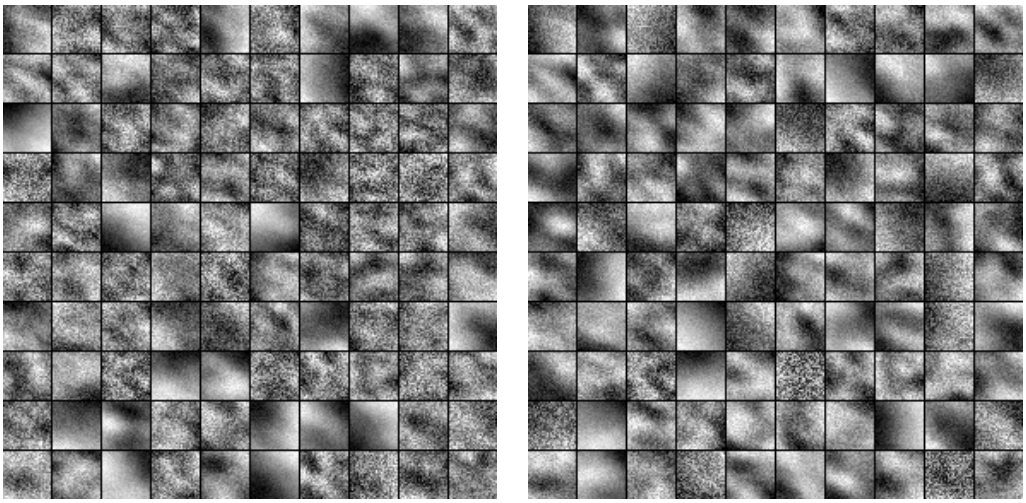
Another popular approach in the deep learning community are auto-encoders [50]. Unlike CNNs, auto-encoders are unsupervised algorithms. Roughly speaking an auto-encoder is made of two convolutional layers whose parameters are trained so that the output vector is the same as the input one. Let $\hat{\mathbf{x}}$ denote the output vector of an auto-encoder. $\hat{\mathbf{x}}$ is meant to be an approximation of $\mathbf{x}$ and they have identical sizes: n. Each component of $\hat{\mathbf{x}}$ is given by:

$$\hat{x}_k = \text{sgm}\left(\sum_{j=1}^{m} \theta_{j,k}^{(2)}\, \text{sgm}\left(\sum_{i=1}^{n} \theta_{i,j}^{(1)} x_i + \theta_{n+1,j}\right) + \theta_{m+1,k}^{(2)}\right), \tag{5.7}$$

where m is the dimensionality of the vectors in the intermediate representation, $\theta_{i,j}^{(1)}$ are the parameters of the first layer (coding) and $\theta_{j,k}^{(2)}$ are the parameters of the second layer (decoding). The objective function is then constructed so as to minimize $\|\hat{\mathbf{x}} - \mathbf{x}\|$ for all training examples.

If one succeeds to do so, then the values obtained in the intermediate representation can be viewed as coefficients of a decomposition of the input with respect to patterns. These patterns are formed by the trained parameter values of the coding layer. In an image processing context, input vectors are patches. By preserving pixel topology during the convolution like operation performed by the auto-encoders, the learnt patterns are also patches which can be visualized. Figure 5.3 (a) shows such patterns when 28×28 patches that were evenly sampled in breast regions of MIAS images are used as input vectors when $m = 100$.

One difficulty in training auto-encoders is to avoid convergence to a trivial solution like simply copy inputs. One solution is to add a sparsity term in the cost function. This term will force the auto-encoder to learn patterns that are very rarely jointly involved in the decomposition of an input patch. Another possibility to improve convergence is to use denoising auto-encoders [51]. This method consists in "zeroing out" some pixels of the input patches while still asking for the auto-encoder to fully reconstruct them. The idea is that auto-encoders should be robust to partial occlusions like our visual



(a) (b)

Figure 5.3: Patterns learnt from regularly sampled 28×28 patches from images of the MIAS dataset: (a) using auto-encoders, and (b) using denoising auto-encoders. 100 units are used as intermediate representation. In (b) 30% of pixels in input patches are set to 0.

cortex is. Figure 5.3 (b) shows learnt patterns when the same input patches as in Figure 5.3 (a) are used but 30% of their pixels were set to 0.

An auto-encoder alone is a shallow network. Usually auto-encoders are piled up in which case they are referred to as stacked auto-encoders [52] and become deep networks. After convergence, it is also possible to retain only the coding part of auto-encoders and remove the decoding part. The coding layers can then be plugged with a neural network or just a softmax function in order to perform supervised learning. In such a case, the convergence of auto-encoders is called pre-training, while the final convergence for supervised learning is called fine tuning [53].
Petersen *et al.* [54] also used such a stacked auto-encoder in another paper.

Surprisingly, the patterns learnt by CNNs or auto-encoders are close to classic feature extraction techniques. In particular, in figure 5.3, many of the displayed patches correspond to edge or corner detectors. This achievement tends to show that deep learning methods learn relevant features. Also, in higher level layers, learnt patterns are gradually matching more abstract representations like object parts for instance [55].

5.4.2 Segmentation free density scoring

Density scoring can be addressed directly using some reference models without classifying each pixel of the breast region into dense or non-dense regions. A majority of such approaches resort to supervised machine learning techniques. Indeed, there are some online public mammogram datasets with labels corresponding to some of the clinical classification system presented in section 5.3. Most of the time, BIRADS labels are available. Since those systems use a finite number of classes, predicting the class of an unseen example is a classification problem.

One of the simplest classification approach is the k nearest neighbor (kNN) algorithm. Using a distance, each feature vector extracted from an unseen example is compared to

all feature vectors generated from training set examples. Of course, kNN only works if one can define a metric in the feature space. Using distance values, the estimated label of the unseen example is obtained by majority voting among the k closest examples in the training set. Although kNN is usually outperformed by other classifiers, it has the advantage to have only one hyperparameter[2]: k. However, when the training set is large, computing all distance values is time consuming.

This type of classifier was used for density characterization in mammograms in [24, 25, 44]. The content of these papers are further commented in the sequel of this subsection.

The Bayesian framework is another powerful tool to train classifiers. Let c denote a random variable representing a mammogram density class. Let C denote the finite co-domain of this random variable. Feature vectors can also be considered as realizations of a multivariate random variable X. In this setting, the classification problem is solved by finding the most probable class $\hat{c}$ of an unseen mammogram given the observed value $\mathbf{x}$ of the feature vector that was extracted from this image:

$$\hat{c} = \arg\max_{c_i \in C} p\left(c = c_i | X = \mathbf{x}\right), \tag{5.8}$$

$$p\left(c = c_i | X = \mathbf{x}\right) = \frac{p\left(X = \mathbf{x} | c = c_i\right) p\left(c = c_i\right)}{\sum_{c_i \in C} p\left(X = \mathbf{x} | c = c_i\right) p\left(c = c_i\right)}, \tag{5.9}$$

with p a probability measure defined on the domain of the random variables. Training a Bayesian classifier is tantamount to estimating the probability distribution of c (prior) and the class-conditional densities[3] $p\left(X = \mathbf{x} | c = c_i\right)$ using training examples.

Caldwell *et al.* [26] used the Parzen technique to estimate these probabilities. Feature vectors in their approach have two components: average fractal dimension of the entire breast and the absolute value of the difference between the average fractal dimension and the fractal dimension of a pre-selected region near the nipple. This second parameter evaluates the discrepancy in texture between the periareola tissue texture against all other breast tissues. Their Bayesian classifier is trained to predict Wolfe's class of unseen mammograms.

Support vector machines (SVM) [56] are another widely used family of classifiers for density characterization in mammograms. Alleging two classes are linearly separable in the feature space, an SVM attempts to find the separating (affine) hyperplan that maximizes the margin with the closest points of each class. The margin is the distance from one such point to the hyperplan. To alleviate the linear separability condition, the canonical dot product is replaced with non-linear kernel functions.

Subashini *et al.* [32] used an SVM with statistical co-occurrence and histogram input feature vectors to classify mammograms from MIAS dataset into three breast density categories: fatty, glandular and dense. These public mammogram labels are those established during the very creation of the MIAS dataset but they do not really match any

[2] A hyperparameter in machine learning is a parameter that is not estimated during the learning step but still influences classification/regression results. They are termed this way as opposed to model parameters, *i.e.* those that we want to learn. Typical hyperparameters are those governing the convergence of the training step.

[3] If the feature space is discrete then class-conditional distributions are used instead of densities.

clinical rating system definition. The authors achieve 95.44% of correct classification for this task.

Sometimes, unsupervised and supervised methods are jointly used. Chen *et al.* [43] use a dictionary learning approach a on several texture features extracted from patches. These features are LBP, LGA, BiF and textons. The dictionary is generated using unsupervised learning. Patches are collected from training images and clustered. Cluster centers are used as mammographic texture prototypes and added to the dictionary. Patches from an unseen image are then used to produce histograms containing prototype occurrences in this image. A k-NN algorithm is applied on histograms for the supervised classification task of BIRADS class prediction.

Diamant *et al.* [25] used a very similar setting with only textons as feature vectors. Prior to k-means clustering, a number of pre-processing steps (histogram equalization, filtering, patch normalization) are applied followed by PCA for dimensionality reduction. Like Subashini *et al.* [32], they train a SVM using MIAS labels. They also use a k-NN which is outperformed by the SVM.

Bosch *et al.* [24] also used dictionary learning from SIFT attributes and textons. Instead of k-means, they use probabilistic semantic latent analysis (pSLA). In this framework, a latent (unobserved) random variable is introduced. The joint probability of feature vectors and images is decomposed with respect to the latent variable using Chapman-Kolmogorov equation. The conditional distributions of the latent variable given an image serve as a compact representation of the mammogram content. These distributions are used as inputs for a k-NN and an SVM classifier. Reported results indicate that best performances are obtained when building the dictionary using textons and learning BIRADS classes with an SVM.

Qualitative scoring such as classification with respect to a given clinical rating system is a very valuable information for cancer risk assessment. However, one may also desire quantitative measures allowing more refined cancer risk ranking.

Using the density map feature, Chen *et al.* [44] further compute two global quantitative measures (dense area and average density). Dense area alone can be viewed as a density score. However, the authors use both features and train a kNN classifier to predict BIRADS categories. Depending on the dataset, reported recognition rates are ranging from 74.61% to 81.22%.

5.4.3 Joint dense tissue segmentation and density scoring

Supposing one is able to identify the image region corresponding to dense tissues, then density scoring is immediately obtained using very simple attributes of the region, like the relative region surface as compared to breast surface for instance. Consequently, a number of authors have addressed dense tissue segmentation which provides more information to the user than segmentation free approaches do.

Saha *et al.* [1] use a threshold to segment feature map obtained using a fuzzy membership function (see subsection 5.4.1). The optimal threshold is obtained by energy minimization. The energy function defined in their article characterizes spatial and

feature value homogeneity of both dense and non-dense regions. Although a linear correlation is reported between dense regions extracted by their algorithm and those defined by expert radiologists, the computation of the feature map necessitates a significant amount of *a priori* knowledge.

Geeraert *et al.* [57] focus on pixel-wise estimation of the volumetric breast density (VBD). This feature is similar as the dense tissue depth of Highnam and Brady [45] that was evoked in subsection 5.4.1. Breast thickness is the only additional parameter to know in order to compute VBD as compared to dense tissue depth. By integrating VBD on the whole breast ROI, a very precise density score is obtained. The authors show that the actual dense tissue volume is correctly estimated with a relative error lower than 5% on breast phantoms[4]. The dense region is fairly segmented using VBD alone except for pixels near the breast contour whose density seems to be overestimated.

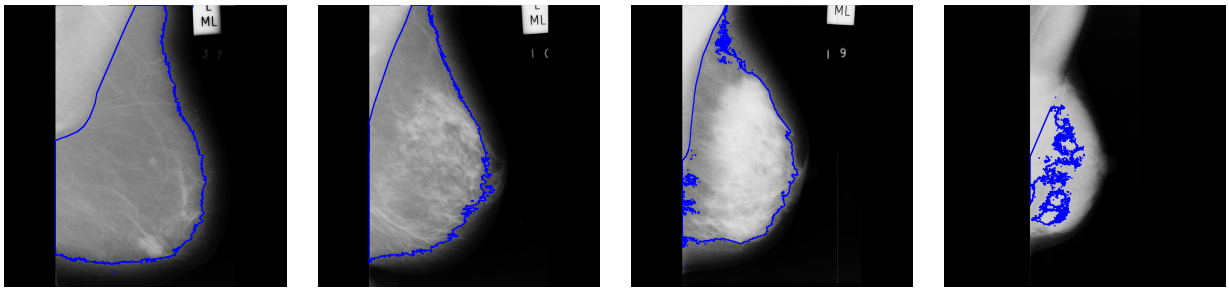
Zwiggelaar [40] exploited more *a priori* information to achieve density related segmentation of mammograms. They select four subsets of MIAS images, one for each BIRADS category. For each of these subsets, they compute a global LGA histogram, thereby creating four prototypes. Pixel-wise segmentation of an unseen image is then performed by computing a local LGA histogram in the neighborhood of each pixel and find the closest prototype. Mammograms are consequently segmented into four regions. The region associated with the BIRADS I prototype is supposed to be made of non-dense tissues while the one associated to the BIRADS IV prototype is allegedly made of maximally dense tissues. Various distance metrics are used to compare LGA histograms: Euclidean, transportation and hybrid transportation.

The remainder of dense tissue segmentation approaches found in the literature are relying on machine learning techniques. Supervised segmentation of mammograms is more rarely used because necessary annotated data are not public and very tedious for expert radiologists to produce.

Georgsson *et al.* [28] build local histograms of the fractal dimension conditionally to several tissue types. Each class appears to have characteristical peaks in their histograms. Using a similarity measure detecting the presence of such peaks, it is possible to classify unseen image patches into one of the tissue types. This method is supervised and the reported ROC curves prove its efficiency. However the labeled dataset is made of image regions extracted from a subset of the MIAS dataset. There is consequently a possibility that the method does not generalize well to the whole MIAS dataset.

He *et al.* [34] use a set of 643 patches extracted from MIAS images. Each patch is labeled according to some tissue type information: nodular, linear, homogeneous and radiolucent. For each class, the authors use k-means clustering on multi-resolution gray level histogram features to obtain class dependent prototypes. The tissue type at a given pixel can then be identified by finding the closest prototype from the feature vector extracted at this pixel location. By using several histogram sizes, different predictions are observed. The set of such predictions are used in the training phase to estimate empirical *a priori* and conditional distributions. These distributions are those pertaining to the Bayes classifier setting. During the test phase, the Bayes classifier

[4]In medical imaging, a phantom is an artificial object mimicking human tissue properties. Images acquired from phantoms are used to validate medical image processing algorithms.



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

Figure 5.4: Dense tissue segmentation using FCM clustering on raw pixel intensities. (a) image mdb005 (BIRADS I), (b) mdb007 (BIRADS II), (c) mdb003 (BIRADS III) and (d) mdb054 (BIRADS IV).

is used on each pixel of an unseen image, thereby yielding a segmented image. The relative proportion of the dense tissue region in this segmented image is used as density scoring.

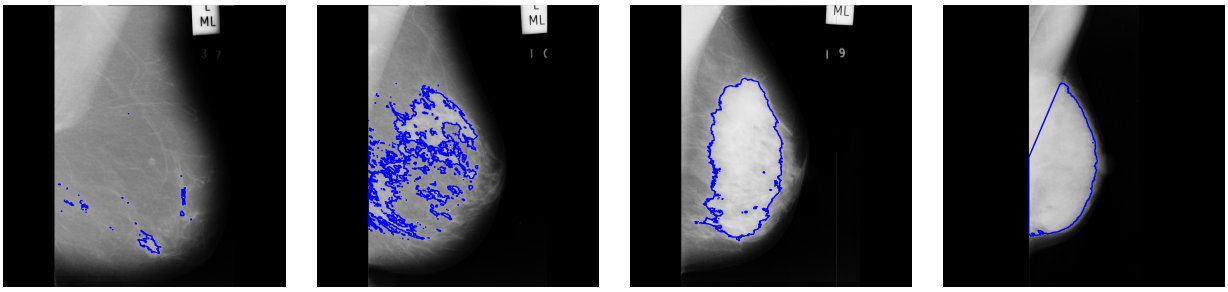
Unsupervised learning can also be used for image segmentation. Oliver *et al.* [7] use a fuzzy c-means (FCM) clustering to segment the breast region into two sub-regions. These two sub-regions contain respectively fatty and dense tissues. Although clustering techniques for mammogram segmentation produce very good results in a number of cases, they unfortunately often rely on the assumption that there is exactly two clusters to find. In practice, some breasts only made of dense tissues (or conversely only made of fatty tissues) can be observed. In such cases, clustering algorithms will generally converge to irrelevant clusters. Figure 5.4 gives FCM segmentation into two regions for four images of the MIAS dataset. Each image has a different BIRADS class and segmentation results are obviously very poor for all of them.

Tortajada *et al.* [33] also used FCM to segment mammograms into two regions. By applying a multiplicative correction term to gray level intensities for pixels near the breast contour, better segmentation results can be obtained. Indeed, this pre-processing step will prevent FCM from clustering such pixels together like in figure 5.4 (a) to (c). Yet this does not solve the problem of the estimation of the actual number of clusters which can still be equal to 1 (as in figure 5.4 (d)) or to 2 (as in figure 5.4 (a) to (c)).

Another interesting point in [7, 33] is that they use a Bayesian combination rule to merge classifiers trained to predict BIRADS classes. Indeed, choosing *a priori* a given type of a classifier is not easy and an alternative solution consists in training several types of classifiers and then use data fusion to aggregate their decisions. Best performances are obtained when combining a kNN classifier with a decision tree. Each of these classifiers are trained using relative region area and histogram and co-occurrence statistical features extracted from both segmented regions.

5.4.4 Concluding remarks on computerized dense tissue characterization

In this section, a great deal of dense tissue texture characterization methods have been reviewed, and there is no obvious texture feature extraction technique leading to improved performances in all circumstances. Surprisingly, the dense tissue region



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

Figure 5.5: Dense tissue segmentation examples using *ad hoc* raw pixel intensities threshold. (a) image mdb005 (BIRADS I), (b) mdb007 (BIRADS II), (c) mdb003 (BIRADS III) and (d) mdb054 (BIRADS IV). The thresholds used for segmentation are respectively 0.63, 0.63, 0.60 and 0.50.

can be accurately segmented by simply applying a global gray level intensity threshold. Figure 5.5 presents some dense tissue regions obtained after directly *ad hoc* thresholding on gray level intensities for pixels inside the breast region only.

In this figure, the segmentation results are very good even if a few post processing steps (like small connected region deletion) might be necessary. In the context of mammographic image processing, segmenting the dense tissue region by a threshold operation has many advantages. The main one is obviously complexity reduction. In general, the region that needs to be found is a subset of a bigger region or of the entire image. Threshold values live usually a small finite discrete space while the set of subsets of the breast ROI is also finite but has a huge number of elements. Since many images have gray level intensities encoded on 8 bits, 254 different threshold values are by far enough.

In addition, cases where the breast region is entirely made of dense tissues are no problem for threshold based segmentation. One just has to select a sufficiently small threshold value to select the entire breast region. Consequently, we will not be facing the same difficulty as the one reported for clustering algorithms (see figure 5.4).

Another advantage is that, for some mammograms, several threshold values produce good results, which tends to show that the segmentation quality is not too sensitive with respect to the threshold value.

Finally, image thresholding is a very fast operation. Some computation time can thus be saved for advanced segmented region analysis allowing to choose the best threshold value for instance.

If the threshold based segmentation precision does not match one's requirements, it is always possible to use a contour fitting method afterwards. The threshold based segmentation will provide a very good initialization for such techniques. Similar techniques as those already used in subsection 4.3.7 can be employed. Another interesting framework for such purpose are active contours. Rahmati and Ayatollahi [58] use an active contour segmentation technique to characterize abnormalities in mammograms. The energy function used for optimization relies on the analysis of gray level distributions inside and outside the active contour. Although their method is not designed for

dense tissue region segmentation, it could be adapted to that end.

The fact that gray level intensities seem to contain sufficient information to separate dense and non-dense tissue regions is justified by the very principle of mammogram acquisition. Indeed, dense tissues will absorb more X-rays yielding brighter pixels than non-dense tissues. As a consequence, significant gaps between gray level intensities are expected between dense and non-dense regions.

Nonetheless, this does not mean that gray level intensity is the only feature allowing to discriminate dense tissues as demonstrated by several authors cited in subsection 5.4.1. Additional texture features should be used to decide on the best threshold value to choose.

An arguable point is the following: should the threshold be global or local? Intuitively, a local threshold seems necessary because there is frequently some drift in pixel intensity from the breast region centroid to the peripheral area. In practice, the drift is very slight except for pixels which are really close to the breast contour, where dense tissues are more rarely found. That is why a global threshold works pretty well.

In the wake of approaches presented in subsection 5.4.3, the next section presents a first attempt to enhance mammogram contrast prior to segmentation using a global threshold. By modifying the contrast of mammograms, the same threshold value could be used for all input images. Finding an appropriate contrast modification or an appropriate threshold are two problems of relatively similar complexity. However, contrast enhancement will additionally provide a complementary visualization of breast patterns.

The relative area of dense tissue is used as a quantitative density scoring. Measured dense tissues areas are compared to qualitative BIRADS classes in order to evaluate the relevance of our approach.

5.5 Mammogram contrast standardization for improved dense tissue region segmentation

In this section, we introduce a mammogram contrast standardization algorithm. Indeed, there is a high variability in contrast among mammographic images, especially when breasts have different types of tissues. For example, fatty tissues are in average darker in BIRADS IV images than in BIRADS I ones. This variability can be explained by the fact that breast volume and exposure time during mammogram acquisition are not constant.

Contrast standardization aims at providing mammographic images such that their histograms have a desirable shape. More specifically, we would like breast region histograms to have either only one mode or two modes centered at significantly different gray level intensities. Indeed, breast tissue ROIs in mammograms are approximately made of bright and dark regions corresponding respectively to dense and non-dense tissues and, more rarely, of only one of these two categories. Let f denote a mapping from an original mammogram u to a contrast standardized one $v = f \circ u$. If f is well

chosen, the dense tissue region A_D should be segmented using a global thresholding: $A_D = \{\mathbf{p} | f(u(\mathbf{p})) > \tau\}$. If the contrast standardization works perfectly well, τ should be constant for all u with $\frac{1}{2}$ as default value.

Among previous attempts to standardize mammograms, the most cited piece of work is the Standard Mammogram Form (SMF) introduced by Highnam and Brady [45]. This approach has been already discussed in subsection 5.4.1 on texture features for mammogram analysis. The authors propose to estimate a physical feature: the actual dense tissue thickness. This can be regarded as a standardization method compliant with physics which is a completely different goal from ours. In our problem, standardized images should be such that (non-)dense regions always have the same texture rendering and similar gray level histograms. Using SMF, pixels corresponding to maximally dense tissues can have different values because the thickness is patient and acquisition dependent. Besides, SMF requires meta-data to estimate dense tissue thickness while we prefer not to make the assumption that such information is available.

Finding a good mapping f is, however, a challenging task. Optimal transport is a framework where the problem underlying such a task is formalized in an elegant and efficient manner. In the next subsection, we will give a brief introduction to this framework. We will see that some classical contrast or histogram modification techniques fall within the scope of optimal transport.

A major difficulty with mammograms is that the desired histogram is ill-known in the sense that the proportion of pixels belonging to the dense tissue region is unknown. In a first attempt to attack this problem, we will formalize it as a cost function minimization problem. The set of constraints (to which the cost function is subject to) is designed to take that lack of information into account. Finally some preliminary experimental results will be given in order to assess the segmentation and density scoring qualities.

5.5.1 Preprocessing: breast mask erosion

A fast pre-processing is necessary before trying to standardize contrast in mammograms. Indeed, our goal is to fight against inter-mammogram contrast variability but, unfortunately, there is also an intra-mammogram contrast variability. As already mentioned in the previous section, it is easy to notice that gray level intensities are gradually vanishing as one gets closer to the breast contour. The reason behind this phenomenon is simply that there are less tissue in this area due to breast compression which creates a turgidity along the breast contour (see figure 2.3).

Tortajada *et al.* [33] use a peripheral pixel correction to deal with this problem. The peripheral pixels are detected using Otsu automatic thresholding. Although this method works in a large number of cases, Otsu's threshold may converge to a value separating other regions than the peripheral one. Figure 5.6 gives an example of such a situation using a mammogram from the MIAS dataset.

Interestingly, the regions segmented in this example are the dense and non-dense ones. Anyway, the peripheral region in this case is overestimated which calls for some more careful processing.

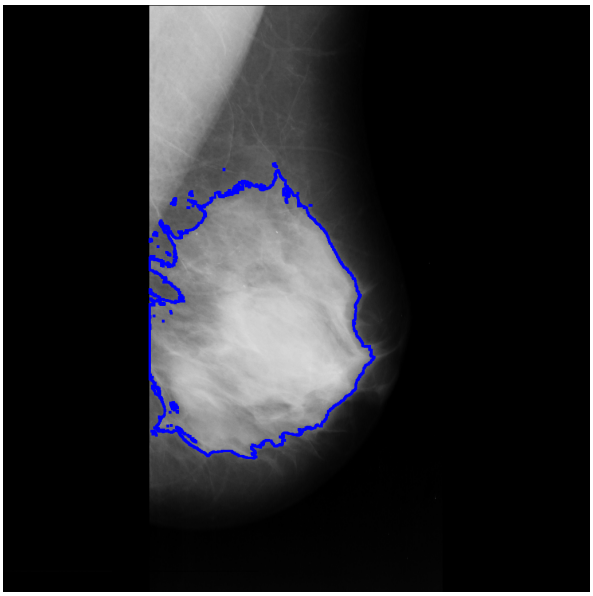


Figure 5.6: A segmentation example using automatic Otsu thresholding on image mdb001 (BIRADS IV). The threshold to which the method converges is approximately 0.41.

In order to avoid such cases and to focus on inter-mammogram contrast correction, we chose to apply a crude erosion of the breast region. Image eroding is a binary operation belonging to the class of mathematical morphology methods. In a few words, two basic operations are defined in this framework: dilation and erosion. In both operations, an input binary image is convolved with a pre-defined binary pattern, called structuring element. Since the output image is not binary anymore, it has to be thresholded:

- in case of dilation, all strictly positive convolutions are set to 1 while null convolutions are left unchanged,
- in case of erosion, all convolutions whose result is smaller than the size[5] of the structuring element are set to 0. All other pixel values are set to 1.

A lot more morphological operations are defined as well as some extensions to grayscale or color images, see [59] for a review on this topic.

For simplicity, we applied an erosion using a 100×100 rectangular structuring element. This element is sufficiently large to ensure that almost all pixels of the peripheral area are removed. This crude erosion will produce an excessively small eroded breast region but it will not impair the selection of an appropriate mapping f. Mammograms have a rather large image resolution and there is far enough pixels left for texture feature extraction in the eroded breast region. Also, once it has been estimated, f can be applied to the entire breast region afterwards anyway.

The erosion is performed on the breast region image mask but only borders corresponding to the breast contour should be eroded. Those corresponding to imaging sensor boundaries should not be eroded. Those are easily identified because they are either horizontal or vertical large segments. Figure 5.7 gives an illustration of this erosion on four images of the MIAS dataset.

[5]The size of the structuring element should be understood as the number of positive pixel values it has.

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

Figure 5.7: Peripheral breast tissue estimation with morphological erosion. Examples of breast region eroded in (a) image mdb005 (BIRADS I), (b) mdb007 (BIRADS II), (c) mdb003 (BIRADS III) and (d) mdb054 (BIRADS IV). The eroded regions are in white and the removed peripheral area is in gray.

Adapting the size of the structuring element to the input image or performing intra-mammogram contrast correction is left for future work.

5.5.2 Mammogram contrast standardization

For the sake of equation concision, we will consider that any input mammogram u is evenly quantized into n bins in the remainder of this chapter. From now on, the co-domain of images is denoted by $\mathcal{X}$ and we suppose that $\mathcal{X} = \left\{0, \frac{1}{n-1}, .., \frac{n-2}{n-1}, 1\right\}$.

We will now present a first attempt to automatically modify the contrast of mammograms so that non-dense and dense regions always have a similar visual rendering. As explained in the introduction of this section, this can be done by applying a relevant mapping $f : \mathcal{X} \longrightarrow \mathcal{X}$ to an input mammogram u. To mitigate contrast discrepancies between mammograms, a different mapping f needs to computed for each u but we do not explicitly underline it in our notation for brevity.

In most contrast correction techniques, the target histogram $\mathbf{h}_T$ is known. In this case, that problem is called histogram specification. The histogram of the input image is denoted by $\mathbf{h}_u$. Under such circumstances, finding an appropriate mapping f^* can be formalized through an optimization problem:

$$f^* = \arg\min_{f \in \mathcal{L}} J(f), \tag{5.10}$$

where $\mathcal{L}$ is some function space and J is a cost function to minimize. Typically, the cost function J comprises a model fitting term of the form $\|\mathbf{h}_T - \mathbf{h}_{f \circ u}\|$.

A famous sub-class of this problem are optimal transport problems. In this setting, $\mathbf{h}_T$ and $\mathbf{h}_u$ are viewed as two probability distributions on $\mathcal{X}$. These probability distributions will be denoted by h_T and h_u because they are functions over $\mathcal{X}$ instead of just vectors. Their related measures are denoted by μ_T and μ_u respectively. Note that the elements of $\mathcal{X}$ are the bins generated when computing the histogram (see equation 5.2). Under a measurability condition for f, our optimization problem then translates into

the following one:

$$f^* = \underset{f \text{ s.t. } \mu_T = \mu_u \circ f^{-1}}{\arg\min} \sum_{i=0}^{n-1} \left| f\left(\frac{i}{n-1}\right) - \frac{i}{n-1} \right| h_u\left(\frac{i}{n-1}\right), \tag{5.11}$$

$$= \underset{f \text{ s.t. } \mu_T = \mu_u \circ f^{-1}}{\arg\min} \sum_{x \in \mathcal{X}} |f(x) - x| \, h_u(x). \tag{5.12}$$

The above problem is known as Monge problem [60] and it does not always have a solution. This problem was later generalized by Kantorovich [61] by considering all joint probabilities whose marginals are h_T and h_u. The class of problems thus defined by Kantorovich are known as optimal transport problems. They can be stated as follows:

$$\pi^* = \underset{\pi \text{ s.t. } \mu_T \text{ and } \mu_u \text{ are marginals}}{\arg\min} E_\pi[c], \tag{5.13}$$

where distribution π^* is called the optimal transport plan, $c : \mathcal{X} \times \mathcal{X} \longrightarrow \mathbb{R}$ is a cost function and E_π is the expectation with respect to law π. Transport plans are finer transport strategies than transport maps because they allow to break the initial masses of h_u and tell you how to spread them.

For 1D histograms, like gray level ones, an approximate solution to the Monge problem is known. This transport map is given by $f = H_T^{-1} \circ H_u$, where H_u and H_T are the cumulative distributions of h_u and h_T and H_T^{-1} is the pseudo-inverse of H_T:

$$\begin{aligned} H_T^{-1} : [0;1] &\longrightarrow \mathcal{X}, \\ r &\longrightarrow \inf\{x \in \mathcal{X} | H_T(x) > r\}. \end{aligned} \tag{5.14}$$

This solution is in general a quite good approximation of the optimal transport and, additionally, it is very easy to implement and to compute. Note that when the target histogram is uniform, this operation boils down to histogram equalization.

Although the optimal transport framework is appealing, it cannot be applied in a straightforward fashion to our situation because, in our case, the target distribution h_T is unknown. In order to obtain harmonized mammograms, we would like our post-transport distribution $h_{f \circ u}$ to have two modes centered at significantly different gray levels.

To simplify the problem, we chose to restrict the set of possible transport maps. We investigate gamma correction transports only as a first attempt to solve our contrast standardization problem. A transport map f_γ is called a **gamma correction transport** if we have:

$$\begin{aligned} f_\gamma : \mathcal{X} &\longrightarrow \mathcal{X}, \\ x &\longrightarrow x^\gamma. \end{aligned} \tag{5.15}$$

This is a crude simplification as finding the best transport map is now tantamount to finding the best value of parameter γ. Gamma correction transport are good candidates for our task because when $\gamma > 1$, such transports will stretch the input image distribution h_u. Typical mammogram histograms have two modes: one for non-dense tissues related pixels and another one for dense tissue related pixels. The center of the first one is always smaller than the center of the second one but there is a strong

overlap between them. The variability in visual rendering is mainly caused by the first mode. Starting from $\gamma = 1$ and by gradually increasing the value of γ, the first mode will slide downward while the second one will not move too much.

Following this principle, a procedure can be written but we need a stopping criterion. One possibility is to compare the empirical distribution $h_{f_\gamma \circ u}$ with a predefined unimodal distribution centered at a sufficiently low gray level. Let g denote one such unimodal distribution. Distributions $h_{f_\gamma \circ u}$ and g can be compared using the Wasserstein distance d_w, a.k.a EMD distance, which is very relevant for histogram comparison. This metric is defined as:

$$d_w\left(h_{f_\gamma \circ u}, g\right) = \inf_{\pi \text{ s.t. } h_{f_\gamma \circ u} \text{ and } g \text{ are marginals}} \left(E_\pi\left[d^p\right]\right)^{\frac{1}{p}}, \tag{5.16}$$

with p a positive integer and d a metric on reals, like Euclidean distance. The reader will immediately notice the analogy with equation 5.13. Indeed, the value of this distance is the optimal transport cost between $h_{f_\gamma \circ u}$ and g.

In practice, we used a discretized gaussian distribution for g centered around $\frac{1}{4}$ and with $\sigma = 0.065$ as standard deviation. From our experience, the average value of non-dense tissue related pixel is never below $\frac{1}{4}$ and the value of σ does not have a big impact on the estimation of parameter γ. However, considering that mammograms have only two modes is not a realistic hypothesis. In many cases, it can have more than two. Consequently, we stop increasing γ as soon as a first local minimum of $d_w\left(h_{f_\gamma \circ u}, g\right)$ is reached. This ensures that least dense tissues related pixels will always appear in a same way after contrast standardization.

The above precaution does not solve issues related to situations where the input mammogram distribution has only one mode as in figure 5.5 (d). To circumvent this difficulty, we need to add a regularization term to our objective function $J(\gamma)$. We chose the Euclidean distance d_e between HoG features extracted from u and HoG features extracted from $f \circ u$. Since HoG are concatenated small histograms, using Wasserstein distance between them could be justified. However, Euclidean distance works well enough for our purpose. This is probably due to the fact that histograms are small (9 orientation bins in our experiments). The choice of HoG as texture features is justified by the fact that they are invariant to illumination changes so the action of the transport will not modify them too much, unless the transport introduces patterns significantly different from the initial image. In the case of image (d) of figure 5.5, this term will rapidly increase with respect to γ.

The approach introduced in the preceding paragraphs is summarized in algorithm 4. This algorithm is called Contrast Standardization Procedure (CSP). The efficiency of CSP is demonstrated on several experiments in the next subsection.

5.5.3 Segmentation performance evaluation

In this section, we present several experiments to assess the efficiency of the CSP algorithm. We only use mammograms of the MIAS datset in these experiments as it is the only public dataset for which we have ground truth breast and pectoral muscle ROIs. We also have BIRADS ratings for this datasets. These ratings were kindly provided by Pr. Arnau Oliver who is with the University of Girona (Spain). These

Algorithm 4 Contrast Standardization Procedure (CSP)

Require: image u, distribution g, parameters λ and ϵ
  $\gamma \leftarrow 1$
  Compute gray level histogram from u and store it as h_u.
  Compute HoG features from u and store them as HoG_u.
  while objective function $J(\gamma)$ decreases **do**
    Apply transport f_γ to input image u and store it as $v \leftarrow f_\gamma \circ u$
    Compute gray level histogram from v and store it as h_v.
    Update contrast fitting term as $E \leftarrow d_W(h_v, g)$.
    Compute HoG features from v and store them as HoG_v.
    Update regularization term $R \leftarrow d_e(HoG_u, HoG_v)$.
    Update objective function $J(\gamma) \leftarrow E + \lambda R$.
    Increment parameter $\gamma \leftarrow \gamma + \epsilon$
  end while
  return γ

ratings were generated by majority voting from three different radiologists. Pr. Oliver notably used these ratings in [7].

Parameter settings and implementation details

Before unveiling experiment results, a few comments on parameter settings and implementation are necessary.
The most influential parameter in the CSP procedure is λ which controls the trade-off between the contrast modification term and the regularization term in the expression of the objective function J. If this parameter needs to be hand-tuned for each input image, then the interest of CSP is questionable as one should spend time on directly hand-tuning an appropriate threshold value applicable directly on the input image. Consequently, we set this parameter to a constant value $\lambda = 230$ for all input images and all experiments presented in this section.

Two other inputs need to be specified in order to run CSP: distribution g and parameter ϵ. CSP is far less sensitive to these two than to λ. As already explained in the previous section, g is a discretized version of a gaussian density function. We used the following function in all the experiments reported in this section: $\frac{1}{\sqrt{2\pi}\sigma} e^{\frac{(x-0.25)}{2\sigma^2}}$ with $\sigma = 0.065$. The parameter ϵ is just the increment used to increase progressively γ. We chose $\epsilon = 0.1$ and this value was kept for all experiments.

Finally, some feature extraction parameters also need to be chosen for computing the contrast modification term and the regularization term of J. For the contrast modification term, the number of gray level bins for histogram computation need to be specified. We chose $n = 256$ bins. Also, from equation 5.16, computing the Wasserstein distance is obviously time consuming. We used an approximation algorithm to estimate this distance [62]. A Python toolbox written by Rémi Flamary (associate professor at the University of Nice) provides a function corresponding to this reference.

For the regularization term, the extraction of HoG features is parameterized as follows: we used 9 orientations on 28×28 patches. Each of these patches is divided into 3×3

sub-patches and therefore generates 81 orientation histograms which are concatenated to yield a feature vector. In addition, the regularization term is smoothed using a mean filter over the 10 preceding values. This filtering removes small irrelevant local minima from function J.

Detailed results on selected mammograms

We begin CSP performance assessment by commenting the results obtained on four mammograms of the MIAS dataset. These mammograms are mdb005, mdb007, mdb003 and mdb054 whose BIRADS classes are respectively I, II, III and IV. These mammograms have already been discussed in the previous section and are good candidates to describe our algorithm because they have different characteristics:

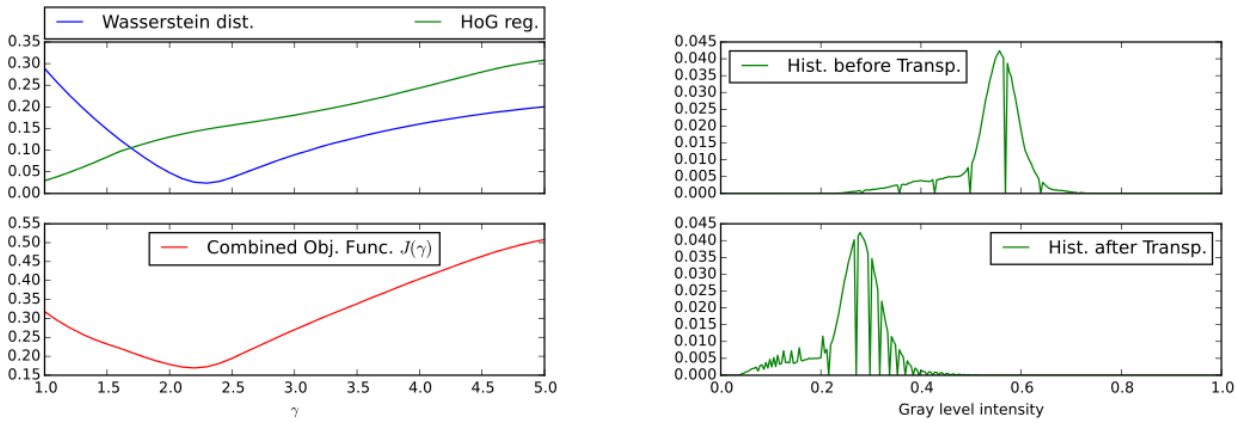
- In image mdb005, non-dense tissues are prominent. There are two classes of pixels but the numbers of class members is highly unbalanced. Consequently, only one mode is visible in the gray level histogram of mdb005 breast region[6].
- In image mdb007, there are a lot more pixels corresponding to dense tissues but the transitions between image regions is obviously very smooth. The breast region histogram of mdb007 has two strongly overlapping modes.
- Image mdb003 is probably the easiest of the four images to process. Dense and non-dense regions are easily distinguished and its histogram has two weakly overlapping modes.
- In image mdb054 the breast is entirely made of dense tissues and the histogram has only one mode.

Figure 5.8 to 5.11 give the evolution of the function J and its two component terms (contrast modification and regularization) with respect to γ for these four images. These curves are given for $\gamma \in [1; 5]$. Observe that the dynamics of these curves are smooth which implies that CSP converges without difficulty to an estimate of γ. These estimates are respectively 2.2, 2.7, 2.8 and 1. Indeed, in the last image, $J(\gamma)$ increases whenever γ grows and CSP stops at the first iteration.
The same figures also give the histograms of the four mammograms before and after transport f_γ is applied. As expected, the dark pixel mode is slid toward $\frac{1}{4}$.

In figure 5.12, the four initial mammograms are displayed in the first column. The second column contains contrast standardized images obtained by transport f_γ where γ is estimated using CSP. The last column contains the contours of dense tissue regions estimated from the standardized images using a thresholding with constant default value $\tau = \frac{1}{2}$. These contours (in blue) are superimposed on the initial images.

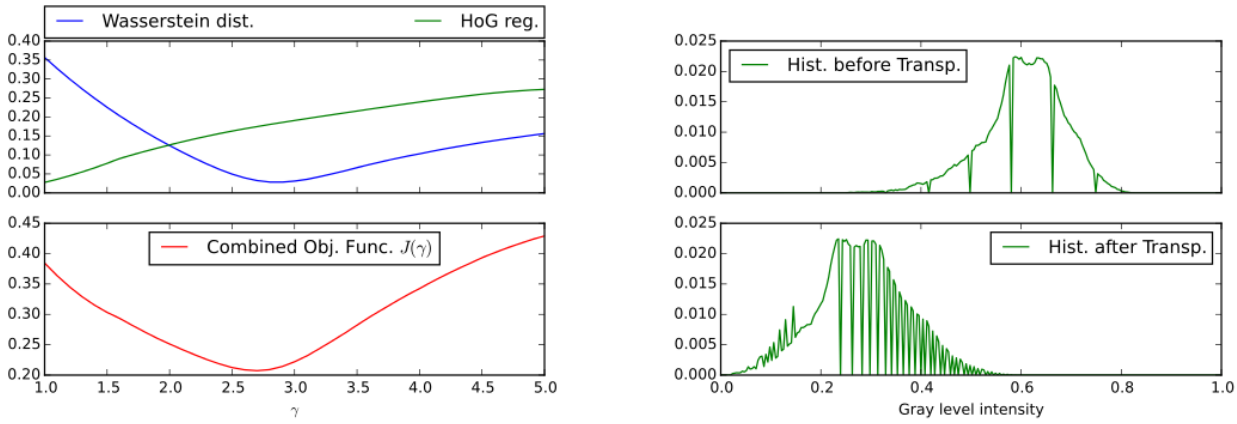
Satisfying segmentation results are obtained. The quality of the contour obtained for image mdb007 is more arguable. The dense tissue region is slightly under-segmented because the transport f_γ has moved both modes of the initial histogram. This is mainly explained by the fact that there a few microcalcifications in this image. These patterns contain very bright pixels. Cropping very bright pixels or performing an advanced analysis of the initial image histogram is one of the main perspectives for future work.

[6]Unless stated otherwise, all mammogram histograms mentioned or displayed in this section are computed from the breast region only, not from the whole mammographic image.



(a) (b)

Figure 5.8: Curves derived from mdb005 image (BIRADS I). (a) Evolution of objective function and its components with respect to γ, (b) Histograms of the initial and contrast standardized images.



(a) (b)

Figure 5.9: Curves derived from mdb007 image (BIRADS II). (a) Evolution of objective function and its components with respect to γ, (b) Histograms of the initial and contrast standardized images.

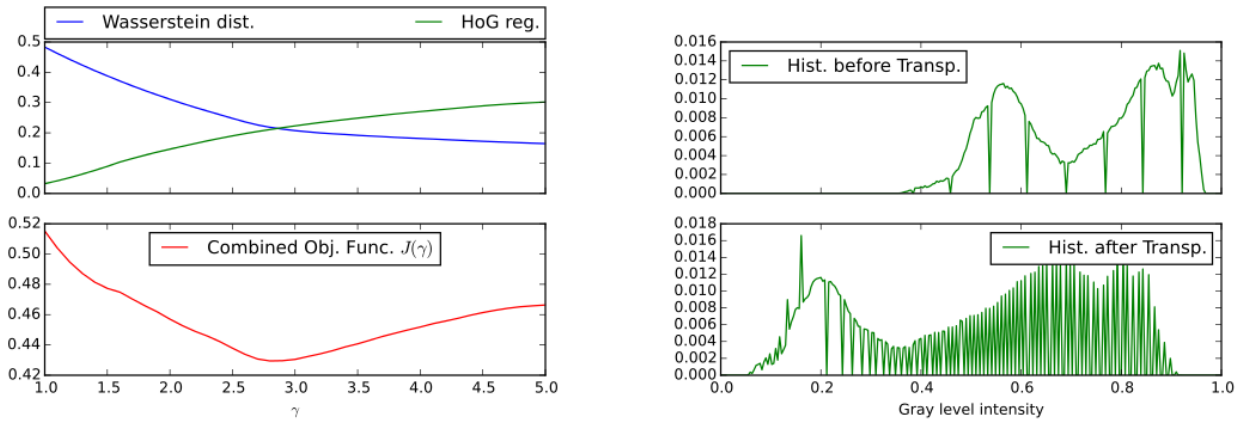
Quantitative evaluation of density scoring on MIAS dataset

Although, CSP has met our expectations on four very different mammograms, one cannot conclude that such performances will generalize to the whole dataset. Recall that the main objective of this chapter is to compute automatically a density score from an input mammogram u.

As explained in the previous section, once the dense tissue region is segmented, some numerical macroscopic density score can be computed easily. We chose to focus on the relative dense tissue area score β which is given by:

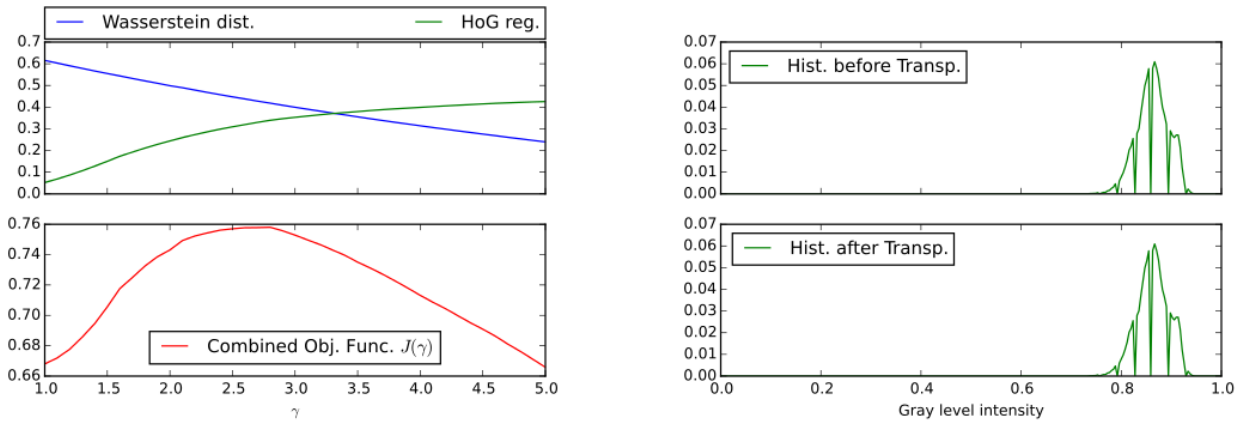
$$\beta(u) = \frac{|\text{dense region}|}{|\text{breast region}|}. \tag{5.17}$$

There is no public ground truth available for β values, however, if the estimation of β values is correct, they should be highly correlated with BIRADS classes. Indeed, the



(a) (b)

Figure 5.10: Curves derived from mdb003 image (BIRADS III). (a) Evolution of objective function and its components with respect to γ, (b) Histograms of the initial and contrast standardized images.



(a) (b)

Figure 5.11: Curves derived from mdb054 image (BIRADS IV). (a) Evolution of objective function and its components with respect to γ, (b) Histograms of the initial and contrast standardized images.

proportion of dense tissues is one of the most important features that radiologists use to classify images according to the BIRADS rating system.

Figure 5.13 gives the Pearson correlation coefficients ρ obtained between β values and BIRADS classes with respect to the threshold τ. We compare correlations when thresholded images are original MIAS mammograms and when thresholded images are contrast standardized mammograms.

A significantly higher correlation is obtained when image contrast is enhanced by transport f_γ prior to thresholding. This proves that the estimate of γ provided by CSP makes sense. Without transport the maximal correlation is 0.65 (for $\tau = 0.7$) while the maximal correlation is 0.75 (for $\tau = 0.38$) when transport is used. In addition, since 0.75 is considered to be a very strong correlation score, one can be confident in the relevance of β as a density score.

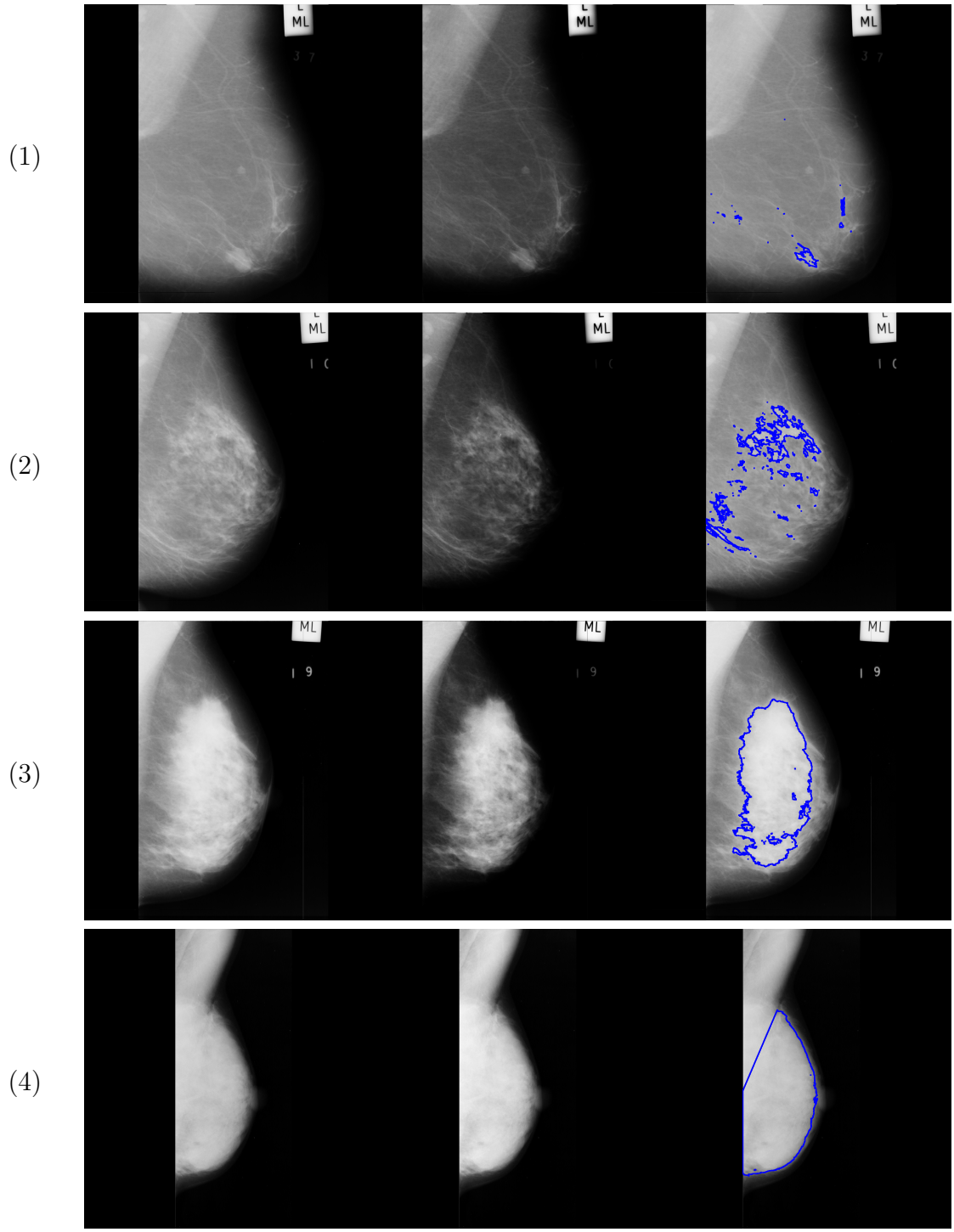


Figure 5.12: Examples of contrast standardized mammograms. Original images in column (a), contrast standardized images in column (b), thresholded area contours ($\tau = \frac{1}{2}$) superimposed to original images in column (c). Results for mdb005 image (BIRADS I) on line (1), results for mdb007 image (BIRADS II) on line (2), results for mdb003 image (BIRADS III) on line (3), results for mdb054 image (BIRADS IV) on line (4).

Note that the observed correlation score does not prove that dense regions are correctly segmented. This is a necessary condition but not a sufficient one. Other tests are necessary to make sure about that. This point is discussed in the next paragraphs.

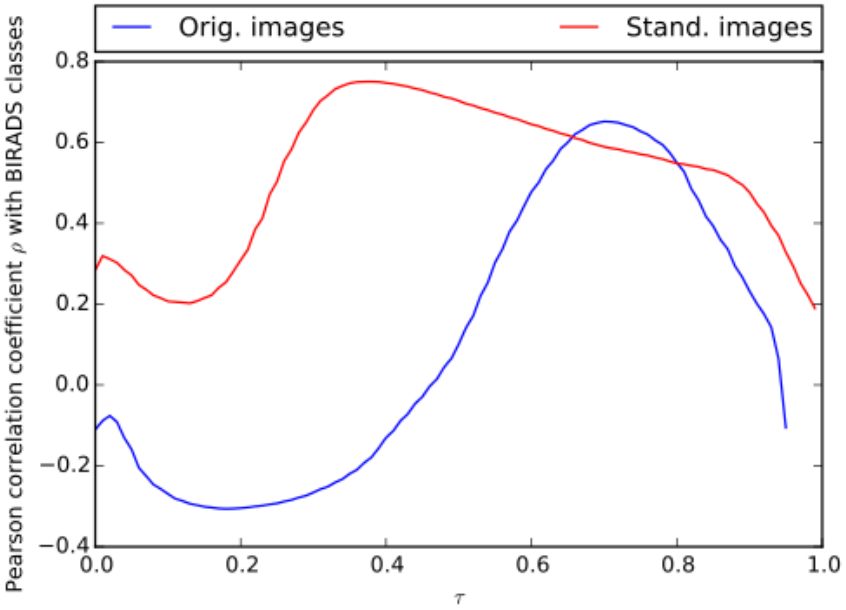


Figure 5.13: Pearson correlation coefficient ρ between dense tissue relative area β and BI-RADS classes with respect to τ. The blue curve is the one obtained when thresholding original mammograms. The red curve is obtained when thresholding contrast standardized mammograms.

Qualitative evaluation of segmentation on MIAS dataset

Unfortunately, there is no public ground truth allowing us to evaluate objectively the quality of the dense tissue segmentation results for MIAS images. Such ground truth are indeed very tedious for radiologists to produce as they must manually delineate dense regions which have very complex shapes.

Instead, we have visually inspected all 322 segmented images computed from the entire MIAS dataset and we have sorted them into 4 segmentation quality categories:

- A : very good segmentation,
- B : fair / good segmentation,
- C : partially incorrect segmentation,
- D : severe mismatch.

Whenever an image is graded C or D, we also reported if the dense region is either under or over-segmented. Table 5.2 contains the number of images falling into one these categories when the threshold τ is set to default value ($\frac{1}{2}$) or to an optimized value (0.7).

For comparison, the same statistic is given for segmentation obtained by constant thresholding of original mammograms. An optimized threshold value (0.38) is selected for that transport free approach. The optimized threshold value is the one maximizing the correlation coefficient between β values and BIRADS classes (see previous experiment).

From table 5.2, it can be seen that segmentation results can still be improved but,

Table 5.2: Grading statistics for dense region segmentation results obtained after thresholding all MIAS images with a constant threshold value.

Method	Original image thresholding (optimized threshold, $\tau = 0.38$)	Standardized image thresholding (default threshold, $\tau = 0.5$)	Standardized image thresholding (optimized threshold, $\tau = 0.7$)
Grade A	95	131	124
Grade B	118	103	132
Grade C (under-segmented regions)	46	45	14
Grade C (over-segmented regions)	6	5	28
Grade C (total)	52	50	42
Grade D (under-segmented regions)	50	33	14
Grade D (over-segmented regions)	7	5	10
Grade D (total)	57	38	24

again, using CSP and applying transport f_γ enhances segmentation results which confirms the conlusions drawn from the previous experiments. In particular, the number of severe mismatch is significantly reduced.

Note that the optimized threshold values need to be known for better results in average. This value can be retrieved in the same way as in the previous experiment provided that one has a set of BIRADS annotated mammograms.

5.6 Conclusion

Dense tissues are some of the most important ones to detect in a mammogram. CAD system as well as expert radiologists' ability to detect abnormalities is significantly impaired as dense breast tissue proportion increases. This problem is mainly explained by the fact that dense tissues and cancer patterns have similar gray level intensities and textures. On top of that, it has been shown that dense tissues are more likely to host cancer lesions than non-dense ones. Computerized dense tissue detection is consequently a major goal in order to improve breast cancer detection rates. At least, associating mammograms with relevant density scores would help radiologists to prioritize patients.

Even if the breast region of a mammogram has been precisely segmented (using some approaches presented in previous chapters), detecting dense tissues in this region is a challenging task. Indeed, there is a high contrast variability among mammograms which implies that dense and non-dense tissues cannot be disambiguated easily unless parameters (like a threshold value) are hand-tuned.

To alleviate this difficulty, we propose an automatic procedure to estimate a parameter γ. Applying gamma correction with this parameter allows to standardize the visual rendering of mammograms. Consequently, dense tissue segmentation method parameters are far less dependent on the processed image.

We have proved that the segmentation results we obtain allows to produce a relevant numerical density score. However, this is a preliminary approach and there is room

for improved performances. In particular, the peripheral area pre-processing that was presented in subsection 5.5.1 deserves more attention. Active contours could be used to detect this area and some other contrast correction techniques could also be applied to it. Secondly, our analysis of input image breast region histograms is too coarse. Statistical techniques could be used to detect modes prior to the estimation of γ. If the mode corresponding to bright pixels was precisely known, histograms would be stretched by gamma correction in a more useful way.

5.7 References

[1] P. K. Saha, J. K. Udupa, E. F. Conant, D. P. Chakraborty, and D. Sullivanv, "Breast tissue density quantification via digitized mammograms," *IEEE Transactions on Medical Imaging*, vol. 20, pp. 792 – 803, August 2001. 92, 103, 109

[2] N. F. Boyd, J. W. Byng, R. A. Jong, E. K. Fishell, L. E. Little, A. B. Miller, G. A. Lockwood, D. L. Tritchler, and M. J. Yaffe, "Quantitative classification of mammographic densities and breast cancer risk: Results from the Canadian national breast screening study," *Journal of National Cancer Institute*, vol. 87, pp. 670 – 675, 1995. 92, 93, 95

[3] J. N. Wolfe, "Breast pattern as an index of risk for developing breast cancer," *American Journal of Radiology*, vol. 126, pp. 1130 – 1139, 1976. 92, 94, 95

[4] J. N. Wolfe, "Risk for breast cancer development determined by mammographic parenchymal pattern," *Cancer*, vol. 37, pp. 2486 – 2492, 1976. 92, 95

[5] V. A. McCormack and D. S. Silva, "Breast density and parenchymal patterns as markers of breast cancer risk: A meta-analysis," *Cancer Epidemiology Biomakers Prevention*, vol. 15, pp. 1159 – 1169, 2006. 92, 93

[6] R. L. Birdwell, D. M. Ikeda, K. D. O'Shaughnessy, and E. A. Sickles, "Mammographic characteristics of 115 missed cancers later detected with screening mammography and the potential utility of computer-aided detection," *Radiology*, vol. 219, pp. 192 – 201, 2001. 92

[7] A. Oliver, J. Freixenet, R. Martì, J. Pont, E. Pérez, E. R. E. Denton, and R. Zwiggelaar, "A novel breast tissue density classification methodology," *IEEE Transactions on Information Technology in Biomedicine*, vol. 12, pp. 55 – 64, January 2008. 92, 94, 95, 100, 103, 105, 111, 119

[8] W. T. Ho and P. W. Lam, "Clinical performance of computer-assisted detection (CAD) system in detecting carcinoma in breasts of different densities," *Clinical Radiology*, vol. 58, pp. 133 – 136, 2003. 92

[9] L. W. Bassett, S. A. Feig, V. P. Jackson, D. B. Kopans, M. N. Linver, and E. A. Sickles, "American college of radiology," in *Illustrated breast imaging reporting and data system BIRADS* (V. Reston, ed.), Philadelphia, PA: American College of Radiology, 3rd ed., 1998. 92, 94, 95, 96

[10] L. Tabár, T. Tot, and P. B. Dean, *Breast cancer - The art and science of early detection with mammography: Perception, Interpretation, Histopathologic correlation*. Stuggart, Germany: Georg Thieme Verlag, 1st ed., December 2004. 92, 94, 95

[11] I. T. Gram, Y. Bremnes, G. Ursin, G. Maskarinec, N. Bjurstam, and E. Lund, "Percentage density, Wolfe's and Tabár's mammographic patterns: agreement and association with risk factors for breast cancer," *Breast Cancer Research*, vol. 7, no. 5, pp. R854 – R861, 2005. 93, 96

[12] J. N. Wolfe, A. F. Saftlas, and M. Salane, "Parenchymal patterns and quantitative evaluation of mammographic densities: a case-control study," *American Journal of Roentgenoly*, vol. 148, pp. 1087 – 1092, 1987. 93

[13] A. F. Saftlas, R. N. Hoover, L. A. Brinton, M. Szklo, D. R. Olson, M. Salane, and J. N. Wolfe, "Mammographic densities and risk of breast cancer," *Cancer*, vol. 67, pp. 2833 – 2838, 1991. 93

[14] J. W. Byng, N. F. Boyd, E. Fishell, R. A. Jong, and M. J. Yaffe, "The quantitative analysis of mammographic densities," *Physics in Medicine and Biology*, vol. 39, pp. 1629 – 1638, 1994. 94

[15] C. M. Vachon, C. H. van Gils, T. A. Sellers, K. Ghosh, S. Pruthi, K. R. Brandt, and V. S. Pankratz, "Mammographic density, breast cancer risk and risk prediction," *Breast Cancer Research*, vol. 9, no. 6, pp. 217 – 225, 2007. 94

[16] M. J. Yaffe, "Mammographic density review: Measurement of mammographic density," *Breast Cancer Research*, vol. 10, no. 3, pp. 209 – 218, 2008. 94, 97

[17] M. Moscowitz, P. Gartside, and C. McLaughlin, "Mammographic patterns as markers for high-risk benign breast desease and incident cancers," *Radiology*, vol. 134, pp. 293 – 295, 1980. 94

[18] A. F. Saftlas and M. Szlko, "Mammographic parenchymal patterns and breast cancer risk," *Epidemiology Revue*, vol. 9, pp. 146 – 174, 1987. 95

[19] P. Toniolo, A. R. Bleich, C. Beinart, and K. L. Koenig, "Reproducibility of wolfe's classification of mammographic parenchymal patterns," *Preventive Medicine*, vol. 21, pp. 1 – 7, January 1992. 95

[20] J. Brisson, C. Diorio, and B. Mâsse, "Wolfe's parenchymal pattern and percentage of the breast with mammographic densities: Redundant or complementary classifications?," *Cancer Epidemiology Biomakers and Prevention*, vol. 12, pp. 728 –723, 2003. 96

[21] I. T. Gram, E. Funkhouser, and L. Tabár, "The Tabár classification of mammographic parenchymal patterns," *European Journal of Radiology*, vol. 24, no. 2, pp. 131 – 136, 1997. 96

[22] E. A. Sickles, "Wolfe mammographic parenchymal patterns and breast cancer risk," *American Journal of Roentgenoly*, vol. 188, no. 2, pp. 301 – 303, 2007. 96

[23] G. Zhang, W. Wang, J. Moon, J. K. Pack, and S. I. Jeon, "A review of breast tissue classification in mammograms," in *RACS*, (Miami, Florida, USA), pp. 232 – 237, November 2-5 2011. 97

[24] A. Bosch, X. Munoz, A. Oliver, and J. Martì, "Modeling and classifying breast tissue density in mammograms," in *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2006. 98, 101, 108, 109

[25] I. Diamant, H. Greenspan, and J. Goldberg, "Breast tissue classification in mammograms using visual words," in *27th IEEE Convention of Electrical and Electronics Engineers in Israel*, 2012. 98, 103, 108, 109

[26] C. B. Caldwell, S. J. Stapleton, D. W. Holdsworth, R. A. Jongt, W. J. Weiser, G. Cooke, and M. J. Yaffe, "Characterisation of mammographic parenchymal pattern by fractal dimension," *Physics in Medicine and Biology*, vol. 35, no. 2, pp. 235–247, 1990. 99, 108

[27] R. Lopes and N. Betrouni, "Fractal and multifractal analysis: A review," *Medical Image Analysis*, vol. 13, no. 4, pp. 634–649, 2009. 99

[28] F. Georgsson, S. Jansson, and C. Olsén, "Fractal analysis of mammograms," in *Lectures Notes in Computer Science* (B. K. Ersboll and K. S. Pedersen, eds.), vol. 4522, pp. 92 – 101, Springer-Verlag, Berlin Heildelberg: SCIA, 2007. 99, 110

[29] C. Chen, J. Daponte, and M. Fox, "Fractal feature analysis and classification in medical imaging," *IEEE Transactions on Medical Imaging*, vol. 8, no. 30, pp. 133 – 142, 1989. 99

[30] W. L. Lee, Y. C. Chen, Y. C. Chen, and K. S. Shich, "Unsupervised segmentation of ultrasonic liver images by multiresolution fractal feature vectors," *Information Sciences*, vol. 175, pp. 177 – 199, 2005. 99

[31] R. M. Haralick, K. Shanmugan, and I. Dinstein, "Textural features for image classification," *IEEE transactions on Systems, Man and Cybernetics*, vol. 3, no. 6, pp. 610–621, 1973. 99

[32] T. Subashini, V. Ramalingam, and S. Palanivel, "Automated assessment of breast tissue density in digital mammograms," *Computer Vision and Image Understanding*, vol. 114, pp. 33–43, 2010. 100, 103, 108, 109

[33] M. Tortajada, A. Oliver, R. Martì, M. Vilagran, S. Ganau, L. Tortajada, M. Sentis, and J. Freixenet, "Adapting breast density classification from digitized to full-field digital mammograms," in *IWDM* (A. Maidment, P. Bakic, and S. Gavenonis, eds.), vol. 7361, pp. 561–568, Springer-Verlag Berlin Heidelberg: Lecture Notes in Computer Science, 2012. 100, 103, 111, 114

[34] W. He, E. R.E.Denton, and R. Zwiggelaar, "Mammographic segmentation and risk classification using a novel binary model based bayes classifier," in *IWDM* (A. Maidment, P. Bakic, and S. Gavenonis, eds.), vol. 7361, pp. 40 – 47, Lecture Notes in Computer Science, springer-verlag berlin heidelberg ed., 2012. 100, 103, 110

[35] M. Unser and A. Aldroubi, "A review of wavelets in biomedical applications," *Proceedings of the IEEE*, vol. 84, pp. 626–638, April 1996. 100

[36] B. N. Prathibha and V. Sadasivam, "An analysis on breast tissue characterization in combined transform domain using nearest neighbor classifiers," in *IEEE International Conference on Computer, Communication and Electrical Technology*, pp. 50 – 54, 18 -19 March 2011. 100, 103

[37] N. Dalal and B. Triggs, "Histograms of oriented gradients for human detection," in *Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on*, vol. 1, pp. 886–893, June 2005. 100

[38] J. Prewitt, "Object enhancement and extraction," *Picture Processing and Psychopictorics*, pp. 75–149, 1970. Academic Press, New York. 101

[39] D. Lowe, "Object recognition from local scale-invariant features," in *IEEE Int. Conf. on Computer Vision (ICCV'99)*, vol. 2, (Kerkyra (Greece)), pp. 1150–1157, 1999. 101

[40] R. Zwiggelaar, "Local greylevel appearance histogram based texture segmentation," in *IWDM* (J. Martí and et al., eds.), vol. 6136, pp. 175–182, Springer-Verlag berlin Heidelberg: Lecture Notes in Computer Science, 2010. 101, 110

[41] T. Ojala, M. Pietikäinen, and T. Mäenpää, "Multiresolution gray-scale and rotation invariant texture classification with local binary patterns," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 24, pp. 971 – 987, March 2002. 101

[42] M. Crosier and L. D. Griffin, "Using basic image features for texture classification," *International Journal of Computer Vision*, vol. 88, no. 3, pp. 447 – 460, 2010. 101

[43] Z. Chen, E. Denton, and R. Zwiggelaar, "Local feature based breast tissue appearance modelling for mammographic risk assessment," *Annals of BMVA*, vol. 2013, no. 3, pp. 1–19, 2013. 102, 109

[44] Z. Chen, A. Oliver, E. Denton, and R. Zwiggelaar, "Automated mammographic risk classification based on breast density estimation," in *Lectures Notes in Computer Science* (J. Sanches, L. Micó, and J. Cardoso, eds.), vol. 7887, pp. 237–244, Springer-Verlag Berlin Heidelberg: IbPRIA, 2013. 102, 108, 109

[45] R. Highnam, M. Brady, and B. Shepstone, "A representation for mammographic image processing," *Medical Image Analysis*, vol. 1, no. 1, pp. 1–18, 1996. 102, 110, 114

[46] R. Highnam, J. Brady, and B. Shepstone, "Computing the scatter component of mammographic images," *IEEE Transactions on Medical Imaging*, vol. 13, pp. 301–313, June 1994. 102

[47] M. Mustra, M. Grgic, and K. Delac, "Feature selection for automatic breast density classification," in *ELMAR, 2010 PROCEEDINGS*, pp. 9–16, September 2010. 105

[48] Y. Lecun, B. Boser, J. S. Denker, R. E. Howard, W. Habbard, L. D. Jackel, and D. Henderson, "Advances in neural information processing systems 2," ch. Handwritten Digit Recognition with a Backpropagation Network, pp. 396–404, San Francisco, CA, USA: Morgan Kaufmann Publishers Inc., 1990. 105

[49] K. Petersen, M. Nielsen, P. Diao, N. Karssemeijer, and M. Lillholm, "Breast tissue segmentation and mammographic risk scoring using deep learning," in *Lecture Notes in Computer Science* (H. Fujita, T. Hara, and C. Muramatsu, eds.), vol. 8539, pp. 88–94, Springer International Publishing Switzerland: IWDM, 2014. 106

[50] D. E. Rumelhart, G. E. Hinton, , and R. J. Williams, "Learning representations by back-propagating errors," *Nature*, vol. 323, pp. 533–536, 1986. 106

[51] P. Vincent, H. Larochelle, Y. Bengio, and P.-A. Manzagol, "Extracting and composing robust features with denoising autoencoders," in *Proceedings of the 25th International Conference on Machine Learning*, ICML '08, pp. 1096–1103, 2008. 106

[52] P. Vincent, H. Larochelle, I. Lajoie, Y. Bengio, and P.-A. Manzagol, "Stacked denoising autoencoders: Learning useful representations in a deep network with a local denoising criterion," *Journal of Machine Learning Research*, vol. 11, pp. 3371–3408, December 2010. 107

[53] Y. Bengio, P. Lamblin, D. Popovici, and H. Larochelle, "Greedy layer-wise training of deep networks," in *Advances in Neural Information Processing Systems 19* (B. Schölkopf, J. Platt, and T. Hoffman, eds.), pp. 153–160, MIT Press, 2007. 107

[54] P. K. Petersen, K. Chernoff, M. Nielsen, and A. Y. Ng, "Breast density scoring with multiscale denoising autoencoders," in *15th International Conference on Medical Image Computing and Computer Assisted Intervention*, 2012. 107

[55] M. D. Zeiler and R. Fergus, "Visualizing and understanding convolutional networks," in *Computer Vision - ECCV 2014* (D. Fleet, T. Pajdla, B. Schiele, and T. Tuytelaars, eds.), vol. 8689 of *Lecture Notes in Computer Science*, pp. 818–833, Springer International Publishing, 2014. 107

[56] C. Cortes and V. Vapnik, "Support-vector networks," *Machine Learning*, vol. 20, no. 3, pp. 273–297, 1995. 108

[57] N. Geeraert, R. Klausz, L. Cockmartin, S. Muller, H. Bosmans, and I. Bloch, "Comparison of volumetric breast density estimations from mammography and thorax ct," *Physics in Medicine and Biology*, vol. 59, no. 15, p. 4391, 2014. 110

[58] P. Rahmati and A. Ayatollahi, "Maximum likelihood active contours specialized for mammography segmentation," in *2nd International Conference on Biomedical Engineering and Informatics, BMEI '09.*, pp. 1–4, October 2009. 112

[59] V. Ćurić, A. Landström, M. J. Thurley, and C. L. L. Hendriks, "Adaptive mathematical morphology - a survey of the field," *Pattern Recognition Letters*, vol. 47, pp. 18–28, 2014. Advances in Mathematical Morphology. 115

[60] G. Monge, "Mémoire sur la théorie des déblais et des remblais," *Histoire de l'Académie Royale des Sciences*, 1781. 117

[61] L. Kantorovich, "On the transfer of masses (in russian)," *Translated in Management Science, Vol. 5, pp. 1-4, 1959*, vol. 37, no. 2, pp. 227–229, 1942. 117

[62] M. Cuturi, "Sinkhorn distances: Lightspeed computation of optimal transport," in *Advances in Neural Information Processing Systems 26* (C. Burges, L. Bottou, M. Welling, Z. Ghahramani, and K. Weinberger, eds.), pp. 2292–2300, Curran Associates, Inc., 2013. 119

Chapter 6

General conclusions

Medical imaging methods are currently at the heart of many clinical practices as they have been established to provide useful information to physicians and therefore contribute to ameliorate health care conditions. However, for an effective use of these latter, objective, qualitative and quantitative information have to be extracted from them. In the mean time, new medical imaging methods are continuously introduced thanks to progresses achieved in sensor technology which in consequence increase the quantities of data to be analyzed by experts in reduced timescale. Undoubtedly, reliable CAD tools become imperative for pre-analysis of such a volume of data in order to relieve physician workloads and ameliorate diagnoses of these latter. The interest of CAD tools for potential clinical applications is continuously growing but they are expected to demonstrate higher levels of performances in order to be integrated in clinical protocols.

In this work, we were particularly interested in the development of tools for improving breast cancer detection. We introduced in this manuscript an approach for mammographic density analysis based on breast tissue characterization in order to assess cancer risk related to a mammogram. It addresses the first main steps of the broad algorithmic architecture presented in Section 2.7 (see figure 2.11). The most significant outcomes of this research work are poorly informative mammographic image regions removal followed by an advanced breast tissues analysis in order to identify mammographic image areas most likely to contain cancerous lesions, *i.e* dense breast tissue regions.

The goal of such a strategy relying on breast tissue characterization in mammogram was to:

- estimate the proportion of dense tissue in breast and derive a cancer risk index which can help identify high risky patients for quick and preventive medical care,
- reduce false alarms rates through a prior identification of potential cancer areas where further advanced investigations should be carried out.

Mammographic image analysis is a hard task to achieve both by expert radiologists and CAD systems. In fact, there are several factors that hinder the automatic analysis of mammograms thereby leading to inaccurate results. Amongst these factors, the most harmful ones are: presence of examination notes in the background of the mammograms, presence of some muscle tissues not belonging to mammary gland, poor

contrast of X-ray images and texture similarities between some cancer signs and breast dense tissues.

To successfully handle these peculiarities and come up with an efficient analysis of breast tissue density, we introduced in this work a strategy based on two macro-steps. We first get rid of poorly informative image regions and subsequently perform an advanced analysis of the breast region. Each of these macro-steps were themselves divided into challenging sub-steps.

In the first macro-step, we addressed the issues of background suppression in mammograms followed by the problems related to pectoral muscle extraction. It is noteworthy that the algorithms implemented in this first macro-step are learning-free ones and solve the problems encountered in a fully automatic manner. As second step, we developed an approach to estimate dense breast tissues proportion in mammographic images. The MIAS database was used to test the performances the algorithms introduced at each step of this work.

Although background suppression in mammograms is a preliminary step to achieve in CAD systems, it is far from being a simple task because of low contrast and smooth variation of gray level intensities across breast edge which do not facilitate boundary identification. A logarithmic contrast enhancement technique is therefore applied on mammograms to stretch the dynamic range in the dark regions. More precisely, the grey level dynamics are greatly amplified at the vicinity of the breast edge therefore enabling boundary identification. Mammograms are subsequently segmented using gray level information and non-breast patterns are identified and suppressed. Reliable results have been obtained and compared favorably to those of a reference standard. It should be noted that cases where our algorithm yielded improper results are those having excessive noise or parasite defects resulting from mammograms digitization process.

Following background suppression, we also addressed in this work the segmentation of another poorly informative image region *i.e* pectoral muscle. Such tissues regularly appear in the MLO views of mammograms and are known to significantly tamper breast tissue analysis. However, complex and disrupting factors such as view layout, tissues overlapping, gray level intensity variation in the region and blur edges make the design of an automatic process for the pectoral muscle extraction very challenging. In such conditions, even using some basic *a priori* knowledge and gray level intensity information do not enable to reliably estimate pectoral muscle boundary.

To tackle some of these issues, we modified the FCM clustering algorithm to exploit spatial information during cluster center update. Indeed the pectoral region is made of connected pixels. Spatial information helps discarding irrelevant unconnected blobs. Nonetheless, the pectoral muscle region segmentation is still irrelevant in tricky cases exhibiting strong overlap with dense breast tissues. We introduce a fast validation process to solve this specific problem and implement a refinement strategy to accurately estimate the pectoral muscle boundary. We compared our extraction results to reference standard ones and also studied the effects of observer variability. This latter study shows the subjectivity of observer perception's to visually delineate pectoral muscle boundary which consequently induces a certain dependency of results with respect to

the observer. Nevertheless, the pectoral muscle has been successfully removed in the mammographic image so that the remaining image region is only constituted of breast tissues ready for tissue density characterization.

Concerning the second macro-step, we focused on breast tissue density assessment since a strong correlation between cancer and breast density has been clearly demonstrated in the literature. In this work, we attempted to segment dense tissue image regions which allow the estimation of the relative proportion of dense breast tissues. A numerical density score is thus automatically derived from an input mammogram. As previously stated, the motivations of such a strategy are:

- identifying potential image regions containing cancer lesions where next investigations should be done,
- evaluating the risk of cancer associated to a mammogram which can justify additional computation efforts to search for breast cancer signs.

The real difficulty in segmenting dense breast tissues in mammographic images is that, in some cases, the amount of dense tissue is almost negligible or completely non-existent in comparison to glandular tissues. In addition, the contrast quality is highly variable from one image to another therefore making dense tissue regions difficult to prototype based gray level intensity information only. This shortcoming has its roots in the inability for radiologist technicians to evaluate breast consistency to derive ideal settings to produce identically contrasted images at each examination.

When gray levels are not informative enough, texture features need to be extracted. Despite the wide variety of texture based approaches introduced in the literature to solve the problem of breast tissue density assessment, none can claim to perfectly describe breast tissue densities. In this work, we propose a new alternative approach relying on a contrast modification. This modification is operated by gamma correction transport mapping gray level intensities. This transport allows to obtain desirable histogram shapes for each kind of breast density category. Although this is not our primal goal, such contrast standardized mammograms offer a complementary visualization of breast tissues for radiologists. More importantly, this contrast modification allows to compute the proportion of dense breast tissue through a simple image thresholding. Indeed, since the images are standardized the value of the threshold is kept constant for all input mammograms. Our experiments show that our density score is highly correlated to density related class labels. These labels were produced by expert radiologists for all mammograms belonging to the MIAS dataset.

While the outcomes of mammographic density analysis as presented throughout this manuscript are many, several perspectives and issues are still to be addressed to come up with a complete and efficient CAD tool for breast cancer detection. Indeed, in this work, we have laid down the basis for an automated and efficient approach for breast cancer detection. The next investigations coming right away at the end of this work should allow the completion the CAD architecture introduced in this PhD. The next stage of this architecture deals with cancer pattern identification. It mainly consists of deriving relevant features to highlight cancerous patterns within dense tissues. Such operations may be computationally demanding if high resolution images such as mammograms are considered. Fortunately, this step will take advantage on the fact that potential

cancerous image areas have already been delimited to not only speed up the process but to also reduce false alarm rates.

On the other hand, our mammographic density score can be considered as an expression of the level of difficulty one faces to identify cancer signs. It could therefore serve as a criterion for selection of an efficient analysis method with respect to breast density.

On top of those already evoked in the conclusion of chapter 5, the results achieved in this work do naturally give rise to several perspectives for further improvements of CAD systems. For instance, mammograms are 2D images of a compressed and deformed 3D object. As consequence, one can not reliably use mammographic image to derive the position of an entity (*e.g* microcalcification or masse) in the breast. One may think of performing mammograms decompression in order to model a 3D reconstruction of breast. However, a proper modeling relies on an appropriate biomechanics tissues description which in turns depends on breast tissue characterization. Other investigations worthy of interest deal with image registration and fusion. In fact, the knowledge on breast tissues composition will facilitate an appropriate registration of mammograms with others modalites and therefore pave the way for an efficient aggregation of images. Such complementary exploitation of information from many sources has proved to produce better results.

Publications of the author

Publications related to this thesis

1. C K Feudjio, J Klein, A Tiedeu, O Colot, "Automatic extraction of pectoral muscle in the MLO view of mammograms", *Physics in Medicine and Biology*, vol.58, no.23, pp. 8493-8515, 2013.

2. Cyrille K Feudjio, John Klein, Alain Tiedeu, Olivier Colot, "Pectoral muscle extraction and breast density estimation in MLO views mammograms", in 3^{rd} *International Conference of the Cameroon Physics Society*, Yaoundé, Cameroon, November 2013.

Other publications

1. Cyrille K Feudjio, Alain Tiedeu, Mihaela Gordan, Marie-Laure Noubeg, Aurel Vlaicu, "Extracting and smoothing contours in mammograms using Fourier descriptors", *Journal of Biomedical Science and Engineering*, vol.7, pp. 119-129, 2014.

2. Cyrille Feudjio, Alain Tiedeu, Mihaela Gordan, Samuel Domngang, Aurel Vlaicu, "Computerized detection and smoothing of contour in mammograms", in 11^{th} *African Conference on Research in Computer Science and Applied Mathematics*, Algiers, Algeria, 13-16 October 2012, pp. 299-306.

3. Kom Guillaume, Alain Tiedeu, Cyrille Feudjio, John Ngundam, "Computerized detection of masses on mammograms by entropy maximization thresholding", in *ICTP-Trieste preprint*, IC/2010/009, march 2010.