AUTOMATIC ANALYSIS OF MAMMOGRAPHY IMAGES

Célia Freitas da Cruz

February 2011
AUTOMATIC ANALYSIS OF MAMMOGRAPHY IMAGES
ANÁLISE AUTOMÁTICA DE IMAGENS DE MAMOGRAFIA

Célia Freitas da Cruz
060519017

Faculdade de Engenharia da Universidade do Porto
Engineering Faculty – Porto University

2010/2011

Master in Bioengineering – Biomedical Engineering

Supervisor: Prof. João Manuel R. S. Tavares (Ph.D)

Porto, February 2011
Acknowledgments

I am grateful to Professor João Manuel R. S. Tavares for support granted to this work.
I would also like to thank Serviço Médico de Imagem Computorizada (SMIC), in Porto, Portugal, and especially Dr. António Paulo Cardoso, for the availability to introduce me some radiology concepts and the clinic reality.
I am also grateful to Professor Artur Cardoso, for the availability to answer our questions and concerns regarding this work.
Abstract

Breast cancer is the utmost usual cancer among the women world population and the most common form of cancer death. However, when early detected, the treatment can be performed earlier and therefore be more efficient.

Mammography is the most common exam to early detect this disease. There are different lesions which are breast cancer characteristic such as microcalcifications, masses, architectural distortions and bilateral asymmetry, which can be detected through this technique. These lesions have some variability, becoming difficult to detect. Furthermore, some other diseases have similar patterns to the breast cancer, which challenges the diagnosis.

Computed aided diagnosis (CAD) intends to provide assistance to the mammography detection, reducing breast cancer misdiagnosis, thus allowing better diagnosis and more efficient treatments. CAD systems result of a collection of computed algorithms which characterize lesions through automatic image analysis. Considerable efforts have been done in this area, and some CAD systems have already been commercialized and approved by Food and Drug Administration in the US. Nevertheless, improvements still need to be done to decrease to the minimal the failure of those systems due to the large variability of the abnormalities and to the difficulty to detect some subtle lesions.

Keywords

- Architectural distortion
- Breast
- Cancer
- Computer-aided detection
- Image analysis
- Image processing
- Mammography
- Medical imaging
- Microcalcification
- Tumoral mass
# Index

1. Introduction .................................................................................................................. 1
   1.1. Overview ................................................................................................................ 1

2. Breast ............................................................................................................................ 2
   2.1. Breast cancer statistics .......................................................................................... 2
   2.2. Breast Anatomy ..................................................................................................... 2
   2.3. Breast Cancer ........................................................................................................ 4
      2.3.1. Breast cancer lesions ..................................................................................... 4
      2.3.2. Types of Breast Cancer ................................................................................. 6
   2.4. Other breast pathologies ....................................................................................... 8
   2.5. Breast Imaging Reporting and Data System ........................................................ 8
   2.6. Summary .............................................................................................................. 9

3. Mammography .............................................................................................................. 10
   3.1. Conventional Mammography Equipment .............................................................. 10
   3.2. Noise and Radiation Dose .................................................................................... 13
   3.3. Full Field Digital Mammography .......................................................................... 13
   3.4. Summary .............................................................................................................. 14

4. Computer Aided Detection .......................................................................................... 15
   4.1. CAD classification .................................................................................................. 15
   4.2. CAD Benefits ........................................................................................................ 17
   4.3. CAD History ......................................................................................................... 19
   4.4. Summary .............................................................................................................. 20

5. Computer Aided Detection Algorithms ........................................................................ 21
   5.1. Preprocessing ......................................................................................................... 21
   5.2. Detection and Classification of Breast Cancer Lesions .......................................... 22
      5.2.1. Image Enhancement ...................................................................................... 22
      5.2.2. Segmentation and Detection ......................................................................... 24
      5.2.3. Classification ................................................................................................. 29
   5.3. Analysis of bilateral asymmetry ............................................................................ 31
   5.4. Summary .............................................................................................................. 31

6. Conclusion .................................................................................................................... 32

7. Work Plan ..................................................................................................................... 33

References ......................................................................................................................... 34
Figure Index

Figure 1 - Cancer incidence among women world population (Ferlay, Shin, Bray, Forman, Mathers, & Parkin, 2010)................................................................. 2

Figure 2 – Anatomy of the breast. From (Seeley, Stephens, & Tate, 2004). .................. 3

Figure 3 - Type of microcalcifications commonly seen on mammographic studies. From (Gunderman, 2006)........................................................................... 5

Figure 4 – Morphologic spectrum of mammographic masses. From (Bruce & Adhami, 1999)................................................................................................. 5

Figure 5 – Mass example with different shape and margin. From (Arnau, 2007). ....... 6

Figure 6 - A craniocaudal view of the right breast demonstrates benign vascular calcifications as well as two well-circumscribed masses containing “popcorn” calcifications classic for involuting fibroadenomas. From (Gunderman, 2006). ......... 6

Figure 7 - Invasive Ductal Carcinoma showing microlobulated borders and microcalcifications. From (Kaushak, 2007). ......................................................... 7

Figure 8 – Mammography equipment diagram. Adapted from (Bronzino, 2000). ....... 11

Figure 9 – Different mammography projections. a) cranio-caudal view, b) mediolateral oblique view. From (Arnau, 2007). ................................................................. 12

Figure 10 – Film-screen receptor. From (Bronzino, 2000). ........................................... 13

Figure 11 - Two plots illustrating receiver operating characteristics (ROC) and free- response receiver characteristics (FROC) curves. The dotted line in the ROC curve represents chance performance. ROC curves are used for diagnosis studies and FROC curves are used for detection studies. From (Sampat, Markey, & Bovik, 2005). .......... 16

Figure 12 – Block diagram of a common CAD system. From (Cheng, Cai, Chen, Hu, & Lou, 2003). ........................................................................................................ 21
Glossary

AEC – Automatic Exposure Control
ANCE – Adaptive Neighborhood Contrast Enhancement
ANN – Artificial Neural Network
CAD – Computer aided detection
BBN – Bayesian belief network
BIRADS – Breast imaging reporting and data system
CC – Cranio-caudal
CR – Computed radiography
FDA – Food and Drug Administration
FFDM – Full Field Digital Mammography
FN – False negative
FNN – Fuzzy Nearest Neighbor
FNSE – Fixed-Neighborhood Statistical Enhancement
FP – False positive
FPI – False positive per image
FROC – Free-response receiver operating characteristic
FSM – Film-screen mammography
HNN – Hybrid Neural Network
KNN – K-Nearest Neighbors
MLO – Mediolateral oblique
ROC – Receiver operating characteristic
ROI – Region of interest
RVM – Relevance vector machine
SVM – Support vector machine
TP – True positive
TN – True negative
1. Introduction

Breast cancer is the most common cancer among the women world population (Autier, et al., 2010). The survival rate and the disease prognosis differ greatly on the cancer stage. When early detected, the treatment is more efficient, because the evolution into a more severe stage is avoided, which implies less mortality risk.

The breast cancer can be detected through imaging exams as mammography, ultrasonography, magnetic resonance imaging, where mammography is the most common exam. Mammography aims to detect characteristic breast cancer lesions.

Computed aided diagnosis intends to provide assistance to the mammography detection, reducing breast cancer misdiagnosis, thus allowing better treatment and prognosis.

The Dissertation ensuing from this research aims to develop computational techniques to enhance and detect lesions in mammography imaging, as well as analyze and apply the methods developed in synthetic and real cases.

This monography aims to provide some information for the comprehension of the further dissertation as well as the state of art in this domain.

1.1. Overview

This report is organized according to the following sections:

Section 2 – Breast: This section intends to explain the overall anatomy and physiology of the breast, as well as the breast cancer statistics in order to demonstrate the coverage of this disease. The breast anatomy and breast cancer biology is approached in order to understand the breast cancer and its imaging. Other breast pathologies are analyzed to explain the differences between their imaging and the breast cancer.

Section 3 – Mammography: In this section, the components and physics of the usual equipment of mammography are explained.

Section 4 – Computer Aided Detection: This section explains the meaning of computer aided detections, its advantages and classification. A history of those systems is also briefly approached.

Section 5 – Computer Aided Detection Algorithms: In this section, there is an explanation of the different algorithms of the various phases of image processing and analysis used to detect lesions in the mammographic images.

Section 6 – Conclusion: In this section are presented the final conclusions of this monography, as well as some other perspectives to the implementation of an algorithm of automatic analysis of mammographic images.

Section 7 – Work Plan: Finally, in this section, the plan of the future work to develop an analysis image system is presented.
2. Breast

This section aims to demonstrate the importance of the breast cancer study and to provide some fundamental knowledge on the breast structure and diseases. Thus, the statistics related to the breast are shown, as well as the anatomic structure of the breast. There is also a description of the different types of breast cancer and some other diseases that affect the breast.

2.1. Breast cancer statistics

Breast cancer affects each year an average of 1.4 million people in the world (Autier, et al., 2010).

Among women, is the most common cancer, comprising 1 in 5 of all new cases of cancer, Figure 1. It is also and the most common form of cancer death, representing 1 in 8 of all deaths from cancer, according to the International Agency of Research on Cancer (Ferlay, Shin, Bray, Forman, Mathers, & Parkin, 2010). It is estimated that more than 150 000 women around the world die of breast cancer annually (Ferlay, Shin, Bray, Forman, Mathers, & Parkin, 2010).

Only 1% of breast cancer cases occur in men (Gunderman, 2006).

In some developed countries, there has been a decrease in breast cancer mortality in the last years. The accepted explanation for this occurrence is the implementation of breast screening programmes, which detects breast cancer in an early stage, decreasing the cancer consequences (Autier, et al., 2010), (Lee, et al., 2010).

2.2. Breast Anatomy

In humans, the breasts are located in left and right sides of the upper ventral region of the trunk and each extends from the second rib above to the sixth rib below. The female breasts correspond to two large hemispherical eminences, which contain the mammary
gland which secretes milk, when stimulated, Figure 2. The mammary glands are sweat glands modified. They exist both in female and male, but in the former is only rudimentary, except in some peculiar circumstances (Gray, 2000), (Seeley, Stephens, & Tate, 2004).

![Figure 2 – Anatomy of the breast](from (Seeley, Stephens, & Tate, 2004)).

The surface of the breast is convex and has, just below the center, a small conical prominence, called papilla or nipple. It is located about the level of the fourth intercostals space. The base of the papilla is surrounded by an areola (Gray, 2000), which has a slightly rough surface due to the presence of rudimentary mammary glands, aerolar glands, just below the surface (Seeley, Stephens, & Tate, 2004).

The adult female breast consists of gland tissue, fibrous tissue, fatty tissue, blood vessels, nerves and ducts. The breast has numerous lobes, usually 15 to 20 (Seeley, Stephens, & Tate, 2004), which are composed of lobules. Those consist of alveoli and lactiferous ducts. These lactiferous ducts enlarge to form a small lactiferous sinus, which accumulates milk during lactation. The milk leaves the breast trough some holes in the nipple. The fibrous tissue lays at the entire surface of the breast and connects the lobes together. The fatty tissue covers the surface of the gland, except for the areola, and is located between the lobes. Usually, this tissue is abundant and determines the form and size of the gland (Gray, 2000), (Seeley, Stephens, & Tate, 2004).

The breast is hold in place as a result of the Cooper’s ligaments support, which extends from fascia over the pectoralis major muscles to the skin over the mammary glands (Seeley, Stephens, & Tate, 2004).

The breast weight and dimension differ between individuals and at different periods of life (Gray, 2000), (Seeley, Stephens, & Tate, 2004). The female breasts start to develop at puberty, stimulated by the hormones estrogens and progesterone of the female sexual menstrual cycle. Higher growth and glands development occurs during pregnancy, when the estrogens levels rise as they are secreted by the placenta and increase even more after delivery, when they are secreting milk to feed the baby. The breasts become atrophied in old age (Gray, 2000), (Guyton & Hall, 2000), (Seeley, Stephens, & Tate, 2004).
A children breast consists principally of ducts with dispersed alveoli, being similar in both female and male. A teenage breast mostly consists on fibrous and gland tissue. When adult, the fatty tissue substituted some of the fibrous and gland tissue. During menopause, the breast is mainly adipose tissue.

The breast is intensely influenced by some hormones. Estrogens stimulate the breast adipose deposition and the growth of the mammary glands, as well as the initial development of lobules and alveoli of the breast. Progesterone and prolactin cause the final growth and are responsible for the function of these structures, and cause the external appearance of the mature female breast (Guyton & Hall, 2000).

During pregnancy, the concentration of estrogens and progesterone increases. This phenomena cause expansion and branching of the breast gland ducts and deposition of additional adipose tissue. Prolactin is responsible for the milk production (Gunderman, 2006), (Seeley, Stephens, & Tate, 2004).

2.3. Breast Cancer

There are many pathologies that affect the breast. Nevertheless, the imagiology of the breast is almost completely addressed to the breast cancer (Gunderman, 2006).

Breast cancer, as the other cancers, corresponds to a malignant growth, which, in this case, begins in the cells of breast tissues. In normal situations, the cell division cycle is controlled and ordered, allowing the formation, growth and tissue regeneration. When this does not occur and there is no reparation of the eventual mutations, there is tumor formation.

After its formation, the evolution depends on the patient. However, an early detection and treatment is essential to stop the cancer evolution and to minimize the damages. The breast cancer, as the majority of other cancers, can have the ability to spread to other tissues, metastasizing, allowing the dissemination of cancer. When the breast cancer is early detected, this phenomenon does not occur, which provides a better prognosis for the patient.

The breast cancer risk is increased with the age, where the majority of patients are over 50 years (Gunderman, 2006). Other risk factors correspond to family history of breast cancer, previous breast cancer, early menarche, late menopause, obesity, nulliparity and chest radiation exposure, abnormal cells in fibrocystic disease and hormone replacement therapy (Gunderman, 2006), (Seeley, Stephens, & Tate, 2004).

Due to these risks, some countries developed the screening programmes, where women over 40 or with higher risk of developing breast cancer perform mammographic exams in a periodic interval.

2.3.1. Breast cancer lesions

Breast cancer has some characteristic lesions such as microcalcifications, masses, architectural distortions. Asymmetry between breasts can also be a breast cancer indicator.

Microcalcifications are small size lesions, typically in the range 0.05 to 1 mm. With these dimensions, microcalcifications are relatively difficult to detect. They are bright and have various sizes, shapes and distributions and in some cases low contrast due to a reduced intensity difference between the suspicious areas and the surroundings. Another reason to their difficult detection is the proximity to the surrounding tissues. In dense tissues, suspicious areas are almost invisible as a result of the tissue superimposition. Some anatomic structures such as fibrous strands, breast borders or hypertrophied
lobules are similar to microcalcifications in the mammographic image (Sankar & Thomas, 2010).

Frequently, the microcalcifications appear in clusters becoming easier to detect (Giger, 2004). There is a high correlation between the presence of microcalcifications and breast cancer, particularly when the microcalcifications are clustered. Therefore, an accurate detection of microcalcifications is essential to an early detection of the majority of breast cancers (Li, Liu, & Lo, 1997). Generally, larger, round and oval shaped calcifications with uniform size have higher probability of being benign, while smaller, irregular, polymorphic and branching calcifications, with heterogeneous size and morphology have higher probability of being malignant (Arnau, 2007), Figure 3.

Masses appear as dense regions of different sizes and properties. They can be circular, oval, lobular or irregular/speculated, Figure 4, and their margins can be (Arnau, 2007), Figure 5:
- circumscribed, which are well-defined and distinctly demarcated borders;
- obscured, which are hidden by superimposed or adjacent tissue;
- micro-lobulated, which have undulating circular borders;
- ill-defined, which are poorly defined scattered borders;
- spiculated, which are radiating thin lines.

Figure 3 - Type of microcalcifications commonly seen on mammographic studies (from (Gunderman, 2006)).

Figure 4 – Morphologic spectrum of mammographic masses (from (Bruce & Adhami, 1999)).
Depending on the morphology, the masses have different malignant probability. The ill-defined and spiculated borders have higher probability of malignancy (Arnau, 2007). A benign process is usually associated with the presence of circular or oval masses. However, the great variability of the mass appearance is an obstacle to a correct mammography analysis (Mini & Thomas, 2003). Some masses can incorporate microcalcifications, as represented in Figure 6.

Architectural distortions refer to the derangement of the normal disposition of the parenchyma in a radiating or arbitrary pattern, without a visible centre or mass. They are very variable and, consequently, very difficult to detect (Mini & Thomas, 2003).

2.3.2. Types of Breast Cancer

Breast cancer can be classified according to the breast tissue where the cancer was originated (glands, ducts, fat tissue or connective tissue) and according to the extent of the cancer spread (non-invasive/in situ or invasive/infiltrating) (Gunderman, 2006).
Carcinoma in situ tumor is an early form of carcinoma (invasive malignant tumor due to muted epithelial cells) detected in an early stage and with absence of invasion of surrounding tissues. A cancer is known as infiltrating when the cells that started in the glands or ducts spread to healthy surrounding tissue. This type of cancer can have a variety of appearances (Eastman, Wald, & Crossin, 2006).

When cancer spreads to other parts of the body through blood and lymph circulation, it is called metastization.

Both in situ and infiltrating cancers can be ductal and lobular, depending on the breast cancer location. Ductal carcinoma arises from the epithelial cells that line the breast milk ducts. In the ductal carcinoma in situ, cancer cells have not penetrated the basement membrane of the ducts. In the mammographic images is characterized by fine microcalcifications; however, the degree of cancer infiltration is not generally visible (Gunderman, 2006). The infiltrating ductal carcinoma is the most frequent type of breast cancer, being responsible for nearly 80% of cases. A tumor irregular mass is characteristic in the mammography of this type of cancer.

Lobular carcinoma begins in the milk glands, in the terminal lobules. Approximately, 10% of breast cancer is lobular carcinoma (Gunderman, 2006). The lobular carcinoma in situ is hardly detected in mammography.

![Figure 7 - Invasive Ductal Carcinoma showing microlobulated borders and microcalcifications (from (Kaushak, 2007)).](image)

Paget’s disease occurs when a ductal carcinoma invades the skin of the nipple.

Inflammatory breast cancer corresponds to an aggressive tumor that invaded the dermal lymphatics (Gunderman, 2006), representing about 1 to 4% of the breast cancer. This cancer usually presents breast inflammation.

Medullary breast carcinoma arises from the stromal cells of the breast (Gunderman, 2006). Mucinous carcinoma is associated with large amounts of cytoplasmic mucin (Gunderman, 2006). The last two types of cancer generally experience lower ability to metastize than the ductal and lobular.
2.4. Other breast pathologies

Some changes in the breast are not malignant. To analyze breast cancer lesions is necessary to regard some other similar lesions caused by different pathologies and benign processes in order to distinguish them.

Fibroadenoma is a benign tumor of the breast developed usually in young women, below 30 years old. This tumor remains in place for some time, but never to a malignant cancer. It can grow rapidly due to the proliferation of the strome and epithelium cells. In mammography, it is characterized as an oval mass with smooth borders, which may have some calcifications (Eastman, Wald, & Crossin, 2006).

A cyst is a closed structure which contains a distinct membrane and may contain air, fluid or semi-solid material. Generally, arises from dilated glandular ducts or lobules. In some rare cases cancer may occur inside the cyst, usually when the inside liquid contains some blood. Some cysts may contain calcium and develop calcification within the walls. Mammographically is a rounded mass with a well defined contour (Eastman, Wald, & Crossin, 2006). After a breast injury with hematoma and fat tissue necrosis, oil cyst may occur, being physically similar to a simple cyst; however, with density equivalent to fat tissue (Eastman, Wald, & Crossin, 2006).

Mastitis is the inflammation of breast tissue due to an infection. In plasma cell mastitis, there are solid, dense, regular rodshape calcifications in the glandular ducts of the breast (Eastman, Wald, & Crossin, 2006).

Mammary dysplasia, also called fibrocystic disease, mastopathy, is a common condition due to excess of estrogen or bigger tissue response to estrogens. It is characterized by three major conditions: formation of fluid filled cysts with fluid, breast duct system hyperplasia and fibrous connective tissue deposition (Eastman, Wald, & Crossin, 2006).

2.5. Breast Imaging Reporting and Data System

The breast imaging resulting of the image analysis can be classified in the level of suspicion of the possibility of breast cancer: breast imaging reporting and data system (BIRADS) score. There are seven categories (Eberl, Fox, Edge, Carter, & Mahoney, 2006):

- Category 0 – assessment incomplete. The mammogram or ultrasound did not provide enough information to a clear diagnosis. Another image exam is required.
- Category 1 – normal. There is an absence of abnormalities.
- Category 2 – benign or negative. There is evidence of benign masses.
- Category 3 – probably benign. The exams are probably normal, but a repeat mammogram should be completed in 6 months.
- Category 4 – possibly malignant. There are suspicious abnormalities. A biopsy is recommended to make a diagnosis.
- Category 5 – malignant. There is indication of malignant lesions. A biopsy is recommended.
- Category 6 – malignant. This category indicates that a malignant diagnosis has already been done.
2.6. **Summary**

The breast cancer affects a large amount of people, particularly women. Additionally, this cancer is the most common form of cancer death. However, when early detected, more possibilities of treatment are promising.

The breasts are composed of gland tissue, fibrous tissue, fat tissue, blood vessels, nerves and ducts. The percentage of these components varies with age and between women.

There are different lesions which are breast cancer characteristic such as microcalcifications, masses and architectural distortions.

Breast cancer can be classified according to the breast tissue where the cancer was originated, usually glands, ducts, fat tissue or connective tissue, and according to the extent of the cancer spread, where it can be non-invasive/in situ or invasive/infiltrating. These lesions have some variability, becoming of challenging detection. Some other diseases have patterns similar to the breast cancer, which difficult the diagnosis.

A breast imaging reporting and data system (BIRADS) score is, generally, used to classify the suspicion of breast cancer.
3. Mammography

Mammography is the most commonly used technique to detect breast cancer at early stages. The goal of this technique is the detection of the disease at a pre-symptomatic phase. When symptoms are developed, the cancer has typically become invasive, and consequently the prognosis is less favorable (Oliver, et al., 2010).

Currently, the mammogram is the most efficient system to detect clinically occult illness, being the only image-based method recommended for breast cancer screening (Chagas, Rodrigues, Tavares, Reis, Miranda, & Duarte, 2007). Mammography can greatly reduce the breast cancer mortality in a well organized screening program over the population, being the breast cancer detection technique that most reduces mortality (Eastman, Wald, & Crossin, 2006). The performance of the mammography decreases as the density of the breast increases. This situation is inconvenient since breast cancer risk increases as the breast density increases (Oliver, et al., 2010).

3.1. Conventional Mammography Equipment

Mammography is a diagnosis exam that uses low-amplitude and high current X-rays to examine the human breast. X-ray is an electromagnetic radiation with high energy: wavelength in the range of $10^{-12}$ m and high frequency ($10^{16} - 10^{19}$ Hz). These characteristics allow the penetration of objects and bodies (Bronzino, 2000), (Nersissian, 2004).

The main X-ray photons interactions with the tissue are photoelectric effect and Compton scattering (Akay, 2006), (Bronzino, 2000). The photoelectric effect occurs when an X-ray photon of short wavelength interacts with the electric field of an atom nucleus and ejects one of its inner electrons. The free electron becomes an ionizing particle (Lima, 1995). In Compton scattering, the X-ray photon interacts with an external electron and becomes free. The incident photon transfers energy to the scattering electron, which is ejected and becomes ionized. The photon changes direction (Lima, 1995). The photoelectric effect is the primary responsible for the radiologic image contrast, while Compton scattering is the primary mechanism for the image resolution limit.

Currently, mammography equipment has an X-ray tube which produces X-rays, Figure 8. This radiation crosses a metal filter and a collimator, which narrows the beam wave. The radiation is transmitted to the breast, which transmits a portion to an anti-scatter grid, passing to the image receptor. In the image receptor, the photons interact and deposit their energy locally, allowing the image formation.

A fraction of X-rays passes through the receiver without interaction, reaching a sensor, which is used to activate the mechanism of automatic exposure control (Bronzino, 2000), (Webster, 2006).

The image formation will depend on the structures’ densities when penetrated with the X-rays, as it absorption is dependent on the structures’ densities. The image must have high spatial resolution to delineate the edges of structures of reduced dimension, as microcalcifications.

Usually, there are two standard image projections: cranio-caudal (CC), which is a view from top, allowing a better imaging of the central and inner breast sectors; and mediolateral oblique (MLO), which is a lateral view from a certain angle, having a better perspective of the glands (Arnau, 2007), Figure 9.
The structures of the conventional mammography will be explained in more detail in the following sections.

\textit{a) X-ray Source}

X-rays used in mammography are originated by the electron bombardment of a hot vacuum tube (cathode) in a metal target (anode), usually molybdenum. The vacuum glass tube heats with the passage of electric current. This current is usually more than 200 mA for short exposures of time (Webster, 2006). The X-ray tube acquires enough thermal energy to leave the cathode (thermoelectric emission), being accelerated toward the anode. X-rays are produced through the de-excitation of the anode element. The resulting photons are transmitted in all directions, so it is necessary the existence of a collimator and filters to limit and direct the output of radiation. Generally, the X-ray tube uses a rotating anode. The cathode electrons reach this anode in a low angle (0° to 16°) of normal incidence (Akay, 2006), (Bronzino, 2000).

The actual focal point corresponds to the anode region involved in the production of X-rays. This region is determined by the width of the electron beam that reaches the anode and the inclination angle. The size of the focal point limits the resolution of the equipment. Small focal points create detailed images with better spatial resolution, allowing detection of microcalcifications, for example. Major focal points allow better heat dissipation (Nersissian, 2004). The angle at which the X-rays hit the target also allows it, but it implies that the effective focal point varies across the image. In modern equipment, the typical size of the focal point for mammography normal contact is 0.3 mm while the small focal point mainly used for the magnification is 0.1 mm (Akay, 2006) (Bronzino, 2000).
b) **X-ray Filter**

X-ray filter, usually molybdenum filter, is needed in order to filter the low energy photons, which creates image artifacts and photons of high energy (higher than 20 keV). This reduction allows the reduction of radiation transmitted to the breast and high contrast images (Haus & Yaffe, 2000).

c) **Compression Unit**

The breast compression allows: dispersion of the dissimilar breast tissues, minimizing the overlap of different breast plans; reduction of the absorbed radiation; breast immobility reducing noise and scattering of the X-rays in the film and homogenization of the radiation in the different tissues, minimizing the noise and improving the image contrast (Akay, 2006), (Bronzino, 2000).

d) **Anti-scatter Grid**

Anti-scatter grids are used to avoid an image contrast decrease produced by scattered radiation when reaches the image receptor. The scattered radiation is due to Compton scattering. Consequently, these grids only allow the passage of primary radiation to create the image. These grids are composed of thin slides made from a non-emitting X-ray material (Akay, 2006), (Bronzino, 2000), (Webster, 2006).

e) **Image Receptor**

The film-screen receptor is usually used as image receptor in the conventional mammography. X-rays get through the light-proof cassette and the film-screen and collide in a phosphor intensifier. The crystals absorb the phosphor energy and produce light with an isotropic distribution. The film emulsion is pressed against the screen, preventing dispersion of photons, which degrade the spatial resolution. The screen is usually treated with chemicals that absorb most of the light, allowing a more accurate picture. Thus, the photons penetrate the film again, forming the image, as schematized in Figure 10 (Bronzino, 2000). Due to this type of image receptor, this mammography is commonly designed film-screen mammography (FSM).
f) Automatic Exposure Control

Proper operations of automatic exposure control (AEC) are essential to obtain mammograms with good image resolution and adjusted amount of radiation. It controls the time of exposure for each examination using sensors adjusting the amount of radiation to the thickness of compressed breast (Akay, 2006), (Bronzino, 2000).

3.2. Noise and Radiation Dose

The noise in the mammogram has origin mainly in two sources: X-ray detector random absorption and granularity associated with the screen-film system. The first, known as quantum noise, depends on the amount of radiation that reaches the image receptor per unit area and on the attenuation coefficient of phosphorous material compared with the thickness of the screen. The granularity of the film increases the higher the speed of film used. Hence, there is a necessity to adjust the speed to maintain a high image quality (Bronzino, 2000).

In mammography high image quality is essential because most of the relevant information of the mammogram corresponds to small details, such as microcalcifications, which can only be identified with a high spatial resolution image.

Although X-rays of low energy allow a better differentiation between tissues, there is a higher dose absorption by tissues and a greater exposure time. Thus, raising the potential of X-ray tube increases the penetration of the beam. Hence, a balance between dose and image quality is necessary. The dose is dependent on several factors such as the speed of receptor-screen film, the existence of anti-scatter grid, the filtration of X-rays, to breast compression, density and composition of breast tissue, the applied current, processing the film and the magnification, i.e. the distance from the source to the image (Akay, 2006).

3.3. Full Field Digital Mammography

A full field digital mammography (FFDM) uses essentially the same equipment and physical principles, but the image receptor is digital and the image is acquired digitally to a computer. This method can overcome disadvantages related with film-screen receptors such as poor image contrast and noise due to the granularity of the film emulsion to record the image.
With digital mammography, the magnification, orientation, brightness and contrast of the image can be adjusted after the exam to allow a better visualization of breast structures. Digital mammography can also make improvements related with more efficient image acquisition since the detector is thin enough to absorb a large fraction of X-rays transmitted by the breast. The digital mammography may improve diagnostic capability and should outweigh the potential reduction in limiting spatial resolution (Akay, 2006).

In digital mammography, digital detectors respond more to a linear increase of absorbed radiation dose than film-screen systems without saturation of high intensities, with more efficient absorption of the radiation beam incident, decreased intrinsic noise and spatial resolution higher (Akay, 2006), (Bronzino, 2000).

The image quality of mammography may be measured in the efficiency with which a detector converts the information from X-ray photons to a signal capable of producing an image. In the case of screen-film system, efficiency is reduced and digital mammography has a higher efficiency for a radiation dose equal to or less (Akay, 2006).

The acquisition system of digital mammography has advantages such as elimination of artifacts from signal processing, contrast enhancement, less time per patient and availability of images. There is the possibility to optimize each of the processes of image acquisition, display and storage, as these are performed independently. However, there is higher cost equipment, there is the need to integrate the equipment in the system, and the images require much processing power from the computer and workstations (Evans, 2007).

Despite the digital mammography being quite promising, some improvements must be done in respect of having a high image resolution with lower cost.

There are several types of integrated digital detector system such as computed radiography, integrated digital detector systems, indirect conversion detectors. To further information see, for example, (NHSBSP Equipment, 2009).

3.4. Summary

Mammography is important to detect early stages of breast cancer, as it detects asymptomatic lesions.

Conventional mammographic equipment has an X-ray tube, which produces X-rays, a metal filter to narrow the beam, an anti-scatter grid, a breast compressor and an image receptor. Additionally, an automatic exposure control is available to adjust the amount of radiation.

The image receptor in the conventional mammography is a screen-film system, while in the full-film digital mammography is a digital receptor. The digital mammography may improve diagnostic capability due to the potential to improve contrast resolution compared with film-screen imaging.
4. Computer Aided Detection

Correct mammograph detection of asymptomatic lesions is essential to discover early breast cancer phases, increasing the treatment options and survival rate (Lee C., 2002).

To properly detect mammogram lesions, radiologists may double read the exams as distinct readers miss different cancers (Blanks, Wallis, & Moss, 1998). However, less costly in man terms, would be the improvement of the performance of individual readers, as the double reading stops being required. In this process, software may be an important assistance (Astley, 2003).

Computer aided detection aims to improve the correct detection of abnormalities in the breast.

Computed aided detection and computer aided diagnosis, both commonly abbreviated as CAD, can be defined as the detection and/or diagnosis made by the radiologist considering the results of a computed algorithm which characterize lesions through automatic image analysis (Masala G., 2006), (Simonetti, Cossu, Montanaro, Caschili, & Giuliani, 1998). CAD systems are used to assist radiologists to locate the lesions, being a “second opinion”, rather than substitute the human diagnosis. This allows the reduction of variability in the radiologists’ mammograms interpretation and the frequency of errors by assuring that suspicious regions are revised and increasing the influence of subtle signs, which may be dismissed otherwise (Akay, 2006).

The use of CAD is supposed to follow the subsequent steps (Rangayyan, Ayres, & Desautels, 2007):
- Initial radiologist mammography reading, marking suspicious areas;
- A CAD system scanning to detect suspicious features;
- Radiologists’ analysis of the prompts given by the CAD system and verification if the suspicious areas were left unchecked in the first reading.

4.1. CAD classification

The efficiency of a CAD system can be classified in four perspectives (Sampat, Markey, & Bovik, 2005):
- True Positive (TP), when the suspected abnormality is in fact malignant;
- True negative (TN), when there is no detection of abnormality in a healthy person;
- False positive (FP), when occurs detection of abnormality in a healthy person;
- False negative (FN), when there is no detection of a malignant lesion.

The last two classifications are critical situations. The false positive requires an invasive examination which implies patient anxiety, stress and unnecessary costs. The false negative is an even worse situation as it compromises the health of the patient and the disease treatment (Sampat, Markey, & Bovik, 2005), (Thangavel, Karnan, Sivakumar, & Mohideen, 2005).

The evaluation of mammography images are analyzed by expert radiologists, by histological examination, in the patological cases and by three years follow-ups in the negative results (Sampat, Markey, & Bovik, 2005), (Thangavel, Karnan, Sivakumar, & Mohideen, 2005).

The performance criteria are evaluated through sensitivity and specificity. The sensitivity is the fraction of the true positive cases over the real positive cases:
High values of sensitivity imply minimal false negative detection. The specificity of the test is the fraction of the true negative cases over the real negative cases:

\[
\text{sensitivity} = \frac{\text{true positives}}{\text{true positives + false negative}},
\]

\[
\text{specificity} = \frac{\text{true negatives}}{\text{true negatives + false positives}}.
\]

High values of specificity imply minimal false positive detection.

Using these two criteria, the results are usually defined in terms of Receiver Operating Characteristic (ROC) curve, Figure 11, which is the tradeoff between the true-positive rate and the false-positive rate inherent in selecting specific thresholds on which predictions might be based (Thangavel, Karnan, Sivakumar, & Mohideen, 2005). ROC also shows the true positive fraction (sensitivity), as a function of the false positive fraction (FP fraction = 1-specificity) obtained varying the threshold level of the region of interest (ROI) selection procedure. Thus, the ROC curve produced allows the detection of massive lesions with predictable performance. The area over the ROC curve represents the error due to the use of the same test. The area under the curve represents the probability that, given a positive and a negative case, the classifier rule will be higher for the positive case, independently of the choice of the threshold decision. The overall performance is evaluated in terms of the area under the ROC curve and the relative errors (Sampat, Markey, & Bovik, 2005), (Thangavel, Karnan, Sivakumar, & Mohideen, 2005).

A perfect classifier would have a true positive rate of 1 (one) and a false positive rate of 0 (zero), for which the rule decision does not fail, as it has no false positive or false negative. Therefore, would have an area under the curve of 1 (one). As the ROC curve is arched towards this point, the better the decisional test. Random guessing would result in an area under the ROC curve of 0.5 (Sampat, Markey, & Bovik, 2005), (Thangavel, Karnan, Sivakumar, & Mohideen, 2005).

![ROC Curve and FROC Curve](image)

**Figure 11** - Two plots illustrating receiver operating characteristics (ROC) and free-response receiver characteristics (FROC) curves. The dotted line in the ROC curve represents chance performance. ROC curves are used for diagnosis studies and FROC curves are used for detection studies (from Sampat, Markey, & Bovik, 2005).

To evaluate true-positive detection, sometimes is also required the localization of the tumor. A better method for this case is Free-Response Receiver Operating Characteristic (FROC), which is a plot of sensitivity versus the false positive per image (FPI), Figure
11. It is typically used to report the performance of the detection algorithm (Sampat, Markey, & Bovik, 2005).

Both FROC and ROC analysis suffer from their limitations. Neither addresses the complexity of images and it is difficult to transform the subjective measurements (radiologists’ observations) to the objective FROC curve (Thangavel, Karnan, Sivakumar, & Mohideen, 2005). ROC analysis has been more developed than FROC curves (Sampat, Markey, & Bovik, 2005).

4.2. CAD Benefits

The human detection of abnormalities in the mammography is often performed subconsciously, without a rule definition, which makes the computer-aided detection a challenging task (Masala G., 2006).

The radiologist analysis of the mammography is fallible, increased by the repetitive and fatiguing task of detection abnormalities, poor image quality, subtlety of some abnormalities, overlap of anatomical structures in the mammogram, low disease prevalence and breast structure complexity. These difficulties can be overcome by approaches such as double reading, which provides double perception and interpretation. Obviously, this procedure is too expensive, complex, and time consuming, particularly in screening programs with a high amount of mammographic images. The development of computerized systems as second readers represents an alternative (Mencattini, Salmeri, Rabottino, & Salicone, 2010).

According to (Ciattò, et al., 2003) CAD had almost the same performance of simulated conventional double reading. However, in (Khoo, Taylor, & Given-Wilson, 2005) is indicated that CAD increases sensitivity of single reading by 1.3%, whereas double reading increases sensitivity by 8.2%.

The use of CAD increases the time taken for an individual reader to review the images. Still, this extra-time taken is not prohibitively slow in practice and the time taken is less than the one taken for double-reading situations (Astley, 2003).

Computers are consistent and indefatigable, and do not require years of practice to acquire the experience need to analyze the mammograph (Simonetti, Cossu, Montanaro, Caschili, & Giuliani, 1998). Hence, the CAD systems are most helpful in those situations and in other circumstances such as screening mammography, when there is large volume of examinations with low disease incidence (up to 30% missed lesions); follow-up examinations, where lesion extraction and quantification are needed in order to measure it (Masala G., 2006).

Consequently, 10 to 30% (Bird, Wallace, & Yankaskas, 1992) of cancers are not detected by radiologists due to misdiagnosis or misinterpretation, where about two/thirds of those are lesions that were posteriorly evident (Sampat, Markey, & Bovik, 2005), (Simonetti, Cossu, Montanaro, Caschili, & Giuliani, 1998).

Studies indicate that radiologists have false-negative rate diagnosis of 21%. CAD has potential to reduce this false-negative rate by 77% (Burhenne, et al., 2000). However, there is some controversy in the efficiency of CAD, when comparing with the radiologists’ performance.

Cancers may also be ignored if the signs are subtle, being wrongly dismissed by the radiologist as being normal. In this case, a correct prompt would add weight to the lesion as abnormal, thus reducing the possibility of misclassification. Many of the very early cancers seen retrospectively show only subtle changes, but there is evidence that
CAD systems are sensitive enough to prompt in such cases (Astley, 2003), (Burhenne, et al., 2000).

Additionally, from the masses referred to surgical biopsies only 10 to 20% are actually malignant (Simonetti, Cossu, Montanaro, Caschili, & Giuliani, 1998).

CAD has, in general, good performance detecting microcalcifications, which can be as high as 99% (Burhenne, et al., 2000), and detecting breast masses, which have been reported to 75 to 89% (Houssami, Given-Wilson, & Ciatto, 2009). Architectural distortion cannot be so accurately detected (Baker, Rosen, Lo, Gimenez, Walsh, & Soo, 2003).

According to (Baker, Rosen, Lo, Gimenez, Walsh, & Soo, 2003), who studied the sensitivity of two commercial CAD systems to architectural distortions, fewer than one half of the cases were detected. Improvements still need to be done in order to increase the detection of this lesion.

The consequences of a benign lesion misdiagnosed as malignant is a biopsy which implies cost and psychological effects such as women anxiety, discomfort and stress. However, the cost and the consequences of a missed cancer are much higher than to a benign lesion misdiagnosed as malignant (Rangayyan, Ayres, & Desautels, 2007), (Schulz-Wendtland, Fuchsjäger, Wackerc, & Hermannnd, 2009), (Simonetti, Cossu, Montanaro, Caschili, & Giuliani, 1998).

CAD needs image digitalization, in case of film-screen mammography, image analysis and characterization of the abnormalities (Simonetti, Cossu, Montanaro, Caschili, & Giuliani, 1998). The use of CAD with digital mammography has advantages when compared with screen-film mammography, which mammograms have to be digitized. Beyond time and money for digitalization, the image quality decreases with this system (Pisano & Yaffe, 2005). Thus, with digital mammography, CAD increases the detections (Akay, 2006). CAD false positive are higher for the digital system when compared with the screen-film system (Pisano & Yaffe, 2005).

Breast cancer CAD has commonly higher sensitivity and positive predictive value than radiologists. However, its false positives need to be reduced in order to increase even further the positive predictive value (Simonetti, Cossu, Montanaro, Caschili, & Giuliani, 1998).

As the radiologists makes the final decision, some of the CAD false prompts are easily dismissed when they are benign calcifications or image artefacts. However, the effect of false prompts high ratio will reduce the potential of CAD to overcome misclassification errors. False prompts may also degrade performance, as they act as distracters, drawing attention away from genuinely abnormal regions. Therefore, successful CAD requires algorithms that are both sensitive and specific (Astley, 2003).

According to (Freer & Ulissey, 2001) the number of cancers detected increased by 19.5% with the use of CAD, and the proportion of early-stage malignancies detected increased from 73 to 78%. The recall rate increased from 6.5 to 7.7%, and the positive-predictive value of biopsy remained unchanged at 38%. Therefore, with this study was concluded that CAD can improve the detection of early-stage malignancies without an excessively adverse effect on the recall rate or the positive-predictive value of biopsy. Another study (Taplin, Rutter, & Lehman, 2006) indicated that CAD increased interpretive specificity but did not affect the sensitivity as unmarked visible non-calcified lesions were less likely to be assessed as abnormal by radiologists. Breast density did not affect CAD’s performance.

However, improvements still need to be done in order to decrease to the minimal the failure of those systems as a consequence of the importance of the diagnosis, due to the
large variability of the abnormal features and to the difficulty to detect lesions in dense breast tissues (Sampat, Markey, & Bovik, 2005).

The consequences of its failures can have serious implications. For these reasons, CAD detection has been quite challenging.

4.3. CAD History

The first paper dealing with computers identifying lesions at mammography was published in 1967 (Winsberg, Elkin, Macy, Bordaz, & Weymouth, 1967). It was based on bilateral comparison, which was recognized as useful in screening mammography with routine viewing of a large number of mostly normal examinations. The concept of computer diagnosis or automated diagnosis in radiology was established at that time (Doi, 2007). Although some interesting results were reported, these early attempts were not successful, because computers were not sufficiently powerful, digital images were not easily accessible and advanced image-processing techniques were not available (Doi, 2007).

By 1980, improvements in computer vision techniques, mammographic quality and digitalization methods started to make clinical CAD possible (Masala G., 2006), (Simonetti, Cossu, Montanaro, Caschili, & Giuliani, 1998). Before this, the concept was that computer would replace radiologists, giving the diagnosis, which was called automated computer diagnosis. Due to this notion, there was some criticism in the early phase to the implementation of computational software to aid diagnosis. By this time, the computer aided detection concept arises (Doi, 2007).

A lot of research has been done from that date on, mainly towards the computer aided diagnosis and the radiologists’ acceptance to this technique started to increase. CAD was first introduced in clinical practice in April 1995, at the University of Chicago, where routine screening mammograms are digitalized and analyzed for masses and calcifications by a clinical workstation (Simonetti, Cossu, Montanaro, Caschili, & Giuliani, 1998).

The United States Food and Drug Administration (FDA) approved the first CAD system in screening mammography in 1998. In 2001, only 130 CAD units were in clinical operation in the U.S. In 2005, this increased to 1600 (Arnau, 2007), (Masala G., 2006).

The first CAD approved by FDA was ImageChecker© of R2 Technology Inc (Hologic, 2010), which system detects potential microcalcifications clusters and masses. This system incorporates a digitizer to convert film mammograms to digital format, detection algorithms and prompts appear on suspicious abnormalities. It has suffered some improvements to strength the evidence and provides detailed examination of the suspicious regions, such as the presence of a threshold to establish whether or not a prompt is displayed. The threshold is set to achieve the optimum balance between sensitivity and specificity. The detection accuracy of calcifications was reported as 98.5% sensitivity at 0.74 false positives per case (set of four images). The detection accuracy of masses was reported as 85.7% at 1.32 false positive marks per case (Sampat, Markey, & Bovik, 2005), (Taylor, Champness, Reddy, Taylor, & Given-Wilson, 2003).

In 2002, two new mammographic CAD systems were approved: MammoReader™ from iCad (iCAD, 2009) and Second Look™ from CADx (CADx, 2003). They have similar principle to the Image Checker©, but with different algorithms, and therefore responding differently to the potential lesions.
MammoReader™ was designed to detect primary signs of breast cancer in mammogram images including microcalcification clusters, well and ill-defined masses, spiculated lesions, architectural distortions, and asymmetric densities. The reported overall sensitivity was 89.3% (91.0% in cases where microcalcifications were the only sign of cancer and 87.4% in the remaining cases where malignant masses were present).

Second Look™ detects mainly microcalcifications and masses. The sensitivity of the system was reported to be 85% for screening detected cancers.

4.4. Summary

Computer aided diagnosis is a computational tool that radiologists can use, which aims to improve the correct detection of abnormalities in the breast. CAD results of a computed algorithm which characterizes lesions through automatic image analysis.

There is still some controversy in this area. However, there are some evidences indicating that this tool, when correctly used, improves the correct detection of microcalcification and masses and consequently the presence of a breast tumor.

The CAD evaluation tools are based on their values of false positives and negatives and true positive and negatives, and thus on the sensitivity and specificity.

Extensive research has been done in this area, and some CAD systems have already been commercialized and approved by FDA. Nevertheless, some improvements still need to be done to decrease to the minimal the failure of those systems due to the large variability of the abnormalities and to the difficulty to detect lesions in dense breast tissues.
5. Computer Aided Detection Algorithms

A CAD unit is generally composed of a scanner, in order to digitize the mammogram when the mammography equipment is not digital, a computer software, to analyze the image and mark the regions with lesions, and a viewer (Sampat, Markey, & Bovik, 2005).

There is a substantial research regarding detection and classification of masses and calcifications. These problems are generally considered well studied, and new developments must meet or exceed the high standards of performance set by the existing algorithms. Moreover, commercial CAD systems have satisfactory effectiveness detecting masses and calcifications. Even so, certain areas of research in CAD of breast cancer still require attention (Rangayyan, Ayres, & Desautels, 2007).

The development of new breast cancer computer-aided detection is an active research field, particularly regarding the detection of subtle abnormalities in mammograms (Rangayyan, Ayres, & Desautels, 2007).

Common CAD systems include image acquisition, followed by enhancement of the image. Segmentation/detection of regions of interest is an essential step of any CAD software. Those regions have a high probability of lesion, thus, this step allows the reduction of the amount of data to process. Following the segmentation, feature extraction is important in order to characterize objects. The features should have similar values for objects in the same categories and different ones for different categories in order to distinguish them. The last step of a common CAD software corresponds to the classification, which is based in the features (Sampat, Markey, & Bovik, 2005). The Figure 12 schematizes these steps. Some CAD software do not integrate all the indicated steps.

![Figure 12 – Block diagram of a common CAD software](from (Cheng, Cai, Chen, Hu, & Lou, 2003)).

5.1. Preprocessing

Some methods are used in CAD softwares to preprocess the original images in order to reduce the noise. These methods include thresholds, low and high pass filters.

A study (Mudigonda & Rangayyan, 2001) uses Gaussian low pass filters and subsampling operations. These subsampling algorithms include thresholding the image by varying levels of intensity to generate a map of iso-intensity contours, and extracting groups of closed contours to represent isolated regions in the image.
5.2. Detection and Classification of Breast Cancer Lesions

5.2.1. Image Enhancement

Mammography lesions such as microcalcifications and masses are usually small and have low contrast regarding to the contiguous breast tissues, thus these abnormalities are hard to detect. Image enhancement can improve the radiologists’ perception to subtle diagnosis, and consequently to more accurate diagnosis (Rangayyan, Ayres, & Desautels, 2007).

Image enhancement includes techniques such as contrast and intensity manipulation, additional reduction of noise, background removal, edges sharpening and filtering. The usual task of mammogram enhancement is to increase the contrast between regions of interest (ROI) and background and to sharpen the edges or borders of regions of interest (Cheng, Cai, Chen, Hu, & Lou, 2003).

However, some image enhancement techniques may distort diagnostic features appearance and shape, leading to misdiagnosis (Kimme-Smith, Gold, Bassett, Gormley, & Morioka). The major problem corresponds to the under-enhancement of some regions and over-enhancement of others. Under-enhancement can cause false negatives, and over-enhancement can cause false positive (Cheng, Cai, Chen, Hu, & Lou, 2003).

With the introduction of digital mammography, there is no need to digitalize film mammograms, which increases the dynamic range, signal to noise, and therefore the need to enhance the image is reduced (Rangayyan, Ayres, & Desautels, 2007).

In this section, some enhancement techniques are introduced.

a) Contrast Stretching

Contrast stretching, also called normalization, aims to improve the image stretching the range of intensity values, rescaling it, usually through the analysis of the image histogram. Usually, is employed when the gray-level distribution is narrow. The goal of this technique is the adjustment of the histogram to achieve a higher separation between the foreground and the background gray-level distribution. However, it is difficult to remove noise whose gray-level are similar to the microcalcification ones (Cheng, Cai, Chen, Hu, & Lou, 2003), (Thangavel, Karnan, Sivakumar, & Mohideen, 2005).

b) Histogram Modeling

Contrast can be increased in a mammogram by adjusting the image histogram in order to separate the foreground and background gray-level distributions.

Histogram modeling modifies the original histogram into a shape as the image gets enhanced. As an example, when the histogram is narrow, it is useful to stretch the low contrast levels. One usual technique corresponds to redistributing the gray levels in order to obtain a histogram as uniform as possible, maximizing the mammogram information, which is designed histogram equalization (Cheng, Cai, Chen, Hu, & Lou, 2003), (Rangayyan, Ayres, & Desautels, 2007). When the equalization is performed locally for every pixel, is called local histogram equalization. Another technique, histogram specification, corresponds to the processing of histogram of the image in order to be similar to a prespecified one. An alternative technique corresponds to the adaptive neighborhood histogram equalization (Rangayyan R., 2005).

However, some normal tissues and noise will still be enhanced with these techniques (Cheng, Cai, Chen, Hu, & Lou, 2003).
**c) Gradient Operators**

Some usual gradient operators are convolution masks, unsharp masks and Sobel gradient (Thangavel, Karnan, Sivakumar, & Mohideen, 2005).

A well-known convolution mask is the unsharp mask. When an image is blurred by some unknown phenomenon, each pixel is composed of its own true value plus fractional components of its neighbors. This technique uses this concept to reduce the blur and improve the image through the reduction of low frequency information and amplification of high frequency detail. It should be noted, that this process can changes dramatically the input image (Cheng, Cai, Chen, Hu, & Lou, 2003).

**d) Fixed-Neighborhood Statistical Enhancement**

This technique, as opposed to the previous ones, is a local-based enhancement approach. For mammograms with no homogeneous background, these techniques may have better performance. Fixed-Neighborhood Statistical Enhancement (FNSE) uses statistical properties in a pixel neighborhood to estimate the background and suppress it. Hence it is possible to increase the local contrast locally (Thangavel, Karnan, Sivakumar, & Mohideen, 2005).

**e) Adaptive Neighborhood Contrast Enhancement Technique**

This technique is similar to the previous one; however, there is adaption of the size of the neighborhood to the local properties. Mammograms have some regions of interest with some image features, which can vary widely in size, shape. With adaptive neighborhood, the detail can be enhanced, without significantly introducing artifacts (Thangavel, Karnan, Sivakumar, & Mohideen, 2005).

According to (Morrow, Paranjape, Rangayyan, & Desautels, 1992), the region-based methods can improve the visibility of microcalcifications clusters and some anatomic details.

The adaptive neighborhood contrast enhancement (ANCE) algorithm has several steps. Each pixel in the digitized mammographic images is a seed pixel in a region growing procedure. This procedure identifies the set of pixels that are similar and connected to the seed pixel, which corresponds to the foreground region. It identifies also the three-pixel wide ribbon of pixels surrounding the foreground region, which corresponds to the background region. Using the contrast value between the foreground and the background, the new value of seed pixel in the contrast enhanced image is obtained (Rangayyan, Ayres, & Desautels, 2007).

A study presented in (Dhawan, Buelloni, & Gordon, 1986) used an optimal adaptive enhancement method and was able to emphasize the features in the image with little enhancement of the noise.

In (Kim, Park, Song, & Park, 1997), an adaptive image enhancement method was developed for mammographic images, based on the first derivative and the local statistics. This method has three steps, where the first one is to remove the artifacts which can be misread as microcalcifications. The second step is the computation of gradient images using first derivative operator and the last step is the enhancement of important features of the mammogram adding adaptively weighted gradient images. Additionally, local statistics of the image are used for adaptive realization of the enhancement, in order to enhance the image details and suppress the noise.
The study described in (Sivaramakrishna, Obuchowski, Chilcote, Cardenosa, & Powell, 2000) compared the performance of several contrast enhancement algorithms: adaptive unsharp masking, contrast-limited adaptive histogram equalization, adaptive neighborhood contrast enhancement, and wavelet-based enhancement) in a preference study. In a majority of the cases with microcalcifications, the ANCE algorithm provided the most-preferred images (58%), followed by the unsharp masking algorithm.

### 5.2.2. Segmentation and Detection

Segmentation is the division of the input image into non-overlapping regions. Usually, it corresponds to the extraction of objects from the background. The segmentation can be done in order to obtain locations of suspicious areas to assist radiologists for diagnosis or to classify the abnormalities as benign or malignant (Cheng, Cai, Chen, Hu, & Lou, 2003). To further information see, for example, (Cheng, Cai, Chen, Hu, & Lou, 2003).

**a) Statistical methods**

These methods include the use of local statistics such as histograms, means and standard deviations. They can be widely used, but they do not work in images without peaks. When the background is homogeneous, the histogram of the image has two distinct peaks: the object and the background separated by a valley. Through a threshold located in the valley, it is possible to segment the image. However, the histogram is not usually bimodal due to the variations in shapes, sizes and intensities of microcalcifications; hence it is difficult to choose an adequate threshold (Cheng, Cai, Chen, Hu, & Lou, 2003).

Local thresholding is an alternative. The threshold is based on an expected bimodal intensity distribution in a selected size window that contains the sub-image to be segmented. The original image is divided into square sub-images. Each sub-image is overlapped by four other sub-images. The level histograms of the sub-images are smoothed by a median filter in order to remove local maxima and minima. Then, the resulting histogram is classified as either bimodal, if there is a valley at the histogram, or unimodal. Once all sub-images have been processed, each unimodal threshold is replaced by a value interpolated from neighboring sub-images (Cheng, Cai, Chen, Hu, & Lou, 2003).

There are other different statistical approaches, such as the one described in (Karssemeijer & Brake, 1996) which is based on statistical analysis of a map of pixel orientations. An important feature of the method is that the way in which an orientation of the image intensity map is determined at each pixel. If is found an increase of pixels pointing to a region, this region is marked as suspicious, especially if such an increase is found in many directions. Around 90% of the malignant cases were detected at rate of one false positive per image.

**b) Region based image processing**

Region-based processing can also be designated as pixel-independent processing, adaptive neighborhood processing or object-oriented processing. It is based on the knowledge that neighbor pixels in a region have similar values.

It may be performed in two perspectives: segmenting the given image and then processing each segment in turn or overlapping regions for each pixel and process each of these regions independently. Overlapping regions are employed to avoid noticeable
edge artifacts and an inferior enhanced mammogram (Morrow, Paranjape, Rangayyan, & Desautels, 1992).

Image processing procedures can then be applied on an image feature basis, rather than pixel by pixel (Morrow, Paranjape, Rangayyan, & Desautels, 1992). This method can also be applied to image enhancement. Region-based method can enhance more anatomical detail without significantly introducing artifacts, and has been demonstrated that it can identify calcifications more effectively in the image of dense breasts where the contrast between calcifications and breast tissue is quite low (Cheng, Cai, Chen, Hu, & Lou, 2003).

A region based segmentation method is region growing. It examines the neighborhood pixels of a “seed” point and groups the pixels with similar properties of it. Two variables need to be specified: the window size and the absolute difference in gray levels between the processed pixel and the seed pixel (Cheng, Cai, Chen, Hu, & Lou, 2003). If the average intensity of the grown region is much greater than the surrounding region, then the pixel is classified as a pixel of the microcalcification. Every pixel in the image is chosen successively as the seed pixel, repeating the overall process (Cheng, Cai, Chen, Hu, & Lou, 2003).

Local thresholding and region growing methods were compared in (Kallergi, Woods, Clarke, Qian, & Clark, 1992), which indicated that local thresholding is more stable, but more dependent on parameter selection.

In (Bankman, Nizialek, Simon, Gatewood, Weinberg, & Brody, 1997) is reported the use of a region-growing-based algorithm for the segmentation of calcifications which do not require threshold or window selection. This method was compared to the multi-tolerance region-growing and to the active contour model, and the results indicated they have similar statistic performance, but the one developed is faster and does not require so computational effort.

c) Edge detection

Edge detection is a common segmentation method. It is based on the property that usually pixel values change rapidly at the boundaries between regions. Many mathematical morphological operations such as erosion, top-hat transformations can be used, and many operators were proposed, such as Roberts gradient, Sobel gradient, Prewitt gradient and Laplacian operator (Cheng, Cai, Chen, Hu, & Lou, 2003). One difficulty that can arise is the knowledge about the resolution of the mammogram that the morphological operation requires to determine the size and shape of the structure elements.

In (Dengler, Behrens, & Desaga, 1993) is presented a systematic method for the detection and segmentation of microcalcifications in mammograms. This technique applies a two stage algorithm to spot detection and shape extraction. The first step uses a weighted difference of Gaussian filter to the detection of spots noise invariant and size-specific. The second stage used a morphological filter to reproduce the shape of the spots.

d) Wavelet approaches

These techniques correspond to image filtering and analysis in the wavelet domain. They can be used to feature enhancement, segmentation and even classification. The mammograms can be examined in a low frequency level of the transform or in a high frequency in order to examine small structures, such as microcalcifications. Commonly, the wavelet transforms reconstructed the original image from transformed coefficients.
modified at each level by local and global nonlinear operators (Cheng, Cai, Chen, Hu, & Lou, 2003).

There are different approaches in the wavelet domain; some of them are analyzed subsequently.

Multiresolution wavelet techniques can show in different levels distinct type of object. This allows the separation of small objects such as microcalcifications, which are included in one level, from large objects such as the background structures, which are included in a different level (Cheng, Cai, Chen, Hu, & Lou, 2003). The advantage of multistage wavelets is that they do not require a priori knowledge of the image or computation of local statistics inside the filter window.

Wavelet theory provides a powerful framework for multiresolution analysis, and it can be used for texture analysis. The discrete wavelet transform is used to map the regions of interest into a series of coefficients, constituting a multiscale representation of the ROIs. To obtain the features reflecting scale-dependent properties, a set of features can be extracted from each scale of the wavelet transform. The most frequently used features are energy, entropy, and norm of the coefficients (Thangavel, Karnan, Sivakumar, & Mohideen, 2005).

In (Strickland & Hahn, 1996), a two-stage method based on wavelet transforms for the detection and segmentation of microcalcifications was proposed. The detected sites, such as microcalcifications, are enhanced in the wavelet domain, before the computation of the inverse wavelet transform. A threshold procedure is done in order to segment the calcifications. A sensitivity of 91% was obtained.

In (Bruce & Adhami, 1999) was performed a multiresolution analysis, specifically the discrete wavelet transform modulus-maxima method to extract mammographic mass shape features. These shape features are used to classify masses as round, nodular, or stellate. These features were compared with traditional uniresolutional shape features in their ability to discriminate between shape classes. These features provided a means of evaluating the shapes at various scales. When utilizing a statistical classification system with Euclidean distance measures determining class membership, the use of multiresolution features significantly increased the classification rates. The classification system when using the multiresolution and uniresolution shape features resulted in classification rates of 83 and 72%, respectively.

Tree-structure wavelet transform is also used to obtain better microcalcification segmentation. Nonlinear multistage tree structured filter suppresses the noise and an edge detection and wavelet transform completed the segmentation. The morphology of the microcalcification and the spatial extent of the cluster were well preserved, which is essential for the later classification (Cheng, Cai, Chen, Hu, & Lou, 2003).

In (Heine, Deans, Cullers, Stauduhar, & Clarke, 1997), a method for identifying clinically normal tissue in mammograms was developed that separates normal regions from potentially abnormal regions. Its first step is the decomposition of the image with a wavelet expansion, which contains a sum of independent images, each one with different levels of image detail. When there are calcifications, there is strong empirical evidence that only some of the image components are necessary for detecting the abnormality. The underlying statistic for each of the selected expansion components can be modeled with a simple parametric probability distribution function. This corresponds to a statistical test that allows the recognition of normal tissue regions. The distribution function depends on only one parameter, which has a statistical distribution and can be used to set detection error rates. Once the summary statistic is determined, spatial filters that are matched to resolution are applied independently to each selected expansion.
image. Regions of the image that correlate with the normal statistical model are discarded, producing an output image consisting only of suspicious areas.

The study presented in (Wang & Karayiannis, 1998) used an approach to detect microcalcifications which employs wavelet-based sub-band image decomposition. The microcalcifications appear usually in small clusters with relatively high intensity when compared with the neighbor pixels. These image features can be preserved by a detection system which uses a suitable image transform that can localize the signal characteristics in the original and the transform domain. As the microcalcifications correspond to high-frequency components of the image spectrum, detection of microcalcifications is achieved through the decomposition of the mammograms into different frequency sub-bands, suppressing the low-frequency sub-band, and, finally, reconstructing the mammogram from the sub-bands containing only high frequencies.

e) Fractal models
Fractals are defined in several different ways, where the most common is a pattern composed of repeated occurrences of a basic unit at multiple scales of detail in a certain order of generation (Rangayyan R., 2005). These models have usually been used to texture analysis.

Mammographic parenchymal and ductal patterns in mammograms possess structures with high local self-similarity which is the basic property of fractals. Tissue patterns can be constructed by fractal models and can be taken out from the original image, and the microcalcification information, which is not similar to the others structures can be enhanced (Sankar & Thomas, 2010), (Thangavel, Karnan, Sivakumar, & Mohideen, 2005). For example, in (Li, Liu, & Lo, 1997) was proposed a fractal model of breast background tissues to enhance the presence of microcalcifications.

The limitation of fractal modeling is the time required for encoding. A modification of the conventional fractal coding was proposed (Sankar & Thomas, 2010) to reduce the encoding time required in the fractal modeling of the mammogram. Hence, instead of searching for a matching domain in the entire domain pool of the image, three methods based on mean and variance, dynamic range of the image blocks, and mass center features are used.

f) Fuzzy approaches
These approaches apply fuzzy operators, properties or inference rules to handle the uncertainty inherent in the image. Due to the variable shapes of microcalcifications, these methods approximate inferences (Cheng, Cai, Chen, Hu, & Lou, 2003), (Thangavel, Karnan, Sivakumar, & Mohideen, 2005). These approaches are very efficient to locate microcalcifications in the mammograms with various densities. In fact, microcalcifications can be accurately detected even in dense breast mammograms. Mammogram enhancement is also more adaptive and robust and the contrast based on fuzzy homogeneity uses both local and global information, which allows to enhance the main feature while suppress the noise (Thangavel, Karnan, Sivakumar, & Mohideen, 2005).

There are different fuzzy approaches. In (Saha, Udupa, Conant, Chakraborty, & Sullivan, 2001) is employed scale-based fuzzy connectivity methods to segment dense regions from fatty regions in mammograms. The segmented dense and fatty regions were quantified through the measurement of the respective area and total density, and the features were derived from these measures. The features were linearly correlated between the two projections: MLO and CC. The method was found to be robust in the segmentation of dense regions.
A novel approach to microcalcification detection based on fuzzy logic and scale space techniques was presented in (Cheng, Wang, & Shi, 2004). First, the images are fuzzyfied through the fuzzy entropy principal and fuzzy set theory. The images are enhanced and then scale-space and Laplacian-of-Gaussian filter techniques are used to detect the sizes and locations of microcalcifications. The major advantage of the method is its ability to detect microcalcifications even in the mammograms of very dense breasts.

**g) Contour detection**

Active contours or “snakes” were introduced by (Kass, Witkin, & Terzopoulos, 1988). This technique seeks for local minimum contours. Placing the contour near the desired image features, the snake essentially seeks for the points, taking a minimum energy measure of all possible points in the neighborhood. In general, the energy measure of a snake contains internal and external forces. The internal forces regulate the ability of the contour to stretch or bend at a specific point. The external forces attract the contour to specific image features.

In (Wirth & Stapinski, 2004), the application of active contours to extract breast region in mammograms was explored. The method is based in the facts that breast-air interface is a very low gradient and may be obscured by noise and that uncompressed fat near the breast-air interface is a gradient, growing as the fat nears the center of the breast. Hence, this method includes noise removal to allow the snake to distinguish the breast contour and the noise. Snakes are designed to fill in gaps that occur in contours.

Right-to-left edge detection picks up the gradient of the breast as an edge when the breast is approaching from the left. As opposite, left-to-right edge detection does not pick up the breast contour, but will pick up noise and other artifacts. A dual threshold would produce a difference in terms of the breast area detected. By taking this difference, it is possible to obtain an approximate location of the breast contour (Thangavel, Karnan, Sivakumar, & Mohideen, 2005).

There are several reasons why active contours are a good approach to breast region extraction. The principal is that the breast is a well defined curve, hence is open to curve approximation characteristics of active contours. In addition, the background in most mammograms is a low intensity and low gradient region, which can be avoided by the active contour due to the search for a local minimum. However, it is necessary some pre-processing techniques to avoid situations such as medium intensity noise which may attract the active contour away from the breast region; the breast-air interface is typically a medium gradient, so energy functional based on edges needs preprocessing; the initial contour will have to be placed relatively close to the desired breast contour (Wirth & Stapinski, 2004).

In (Valverde, Guil, & Muñoza, 2004) was presented an algorithm for the segmentation of vessels in mammograms. This technique is useful in order to eliminate vascular false positives during detection of microcalcifications in mammograms. However, the main problem corresponds to the high level of noise presence in mammograms. An initial theoretical analysis of edge detection is done to select the optimum edge detector and threshold value. Then, a local approach is performed, which corresponds to a segmentation process based on a snake with a new noise energy term to extract the vessel contour and remove particle noise that remained in the image.

**h) Level-set methods**

Level-set methods were introduced by (Osher & Sethian, 1988). These methods can be seen as deformable models. The shape to be recovered is captured through the
propagation of an interface represented by the zero level set of a smooth function (Gelas, Bernard, Friboulet, & Prost, 2007). Hence, the topological changes can be easily handled and the geometric properties of the contour can be implicitly calculated (Ma, Tavares, Jorge, & Mascarenhas, 2009). This approach is a numerical technique for computing and analyzing motion of interfaces, which may develop sharp corners, break apart, merge together and disappear due to significant topologic changes (Wang, Lim, Khoo, & Wang, 2007).

The evolution of the interface is determined by a time-dependent partial differential equation which corresponds to the Hamilton-Jacobi equation. The velocity terms reflect the image features, which characterizes the object to be segmented (Gelas, Bernard, Friboulet, & Prost, 2007). This method can be implemented in two different ways (Gelas, Bernard, Friboulet, & Prost, 2007): narrow-banding, where this method is only applied in narrow bands around the interface, having lower computational cost; reshaping, where the level-set function may develop steep or flat gradients due to the propagation, which yield inaccuracies in the numerical approximation.

This method has been commonly applied to structural shape and topology optimization problems (Wang, Lim, Khoo, & Wang, 2007).

5.2.3. Classification

A great number of features and classification methods have already been developed to detect and classify the lesions as malignant or benign. If the features are adequate, will highlight the differences between the abnormal and normal tissue, and thus the classifier will be more robust.

In the next sections, some classification methods to detect mammographic lesions are introduced.

a) Artificial Neural Networks

The development of artificial neural networks (ANN) was inspired by the biological learning systems. In these systems there is a very complex web of interconnected neurons which possess high information processing abilities of the biological neural systems due to highly parallel processes operations distributed over many neurons. Hence, ANN mimics the highly parallel computation based on distributed representation (Wang, Lederman, Tan, & Zheng, 2010).

Using a set of training data (feature vectors), the ANNs are trained iteratively to minimize the error (Wang, Lederman, Tan, & Zheng, 2010).

The neural network rule extraction algorithms have some general steps: selection and training of the network to meet the pre-specified accuracy requirement; removal of the redundant connections in the network through pruning, while maintaining its accuracy; discretization of the activation values of the pruned network by clustering; extraction of rules that describe the network outputs in terms of the discretized values; generation of the rules that describe the discretized hidden unit activation values in terms of the network inputs. Finally, the two sets of rules generated previously are merged to obtain a set of rules that relates the inputs and outputs of the network (Thangavel, Karnan, Sivakumar, & Mohideen, 2005).

b) Hybrid Neural Network Classification

A hybrid intelligent system to the identification of microcalcification clusters in digital mammograms was presented in (Papadopoulos, Fotiadis, & Likas, 2002). The system has two components: a rule construction and a neural network sub-system. The
rule construction includes the feature identification step and the selection of a threshold value for each feature. For every feature, several threshold values are examined in its range of value. For each threshold value is recorded the number of ROIs below and above the threshold value. The ratio of the number of ROIs that belong to a specific class (normal or pathological) over the total number of the ROIs that belong to the same class should be more than 6%.

c) K-Nearest Neighbors
This class of method classifies objects based on the closest training examples in the feature space. Thus, an object is classified according to the majority of its K-nearest neighbors. Hence, it is instance based learning.

For the K-Nearest Neighbors (KNN) is necessary to have a training set not too small, and a good discriminating distance. KNN performs well in multiclass simultaneous problem solving. The parameter K corresponds to the number of nearest neighbors considered to perform the classification. There is an optimal choice for this value which brings to the best performance of the classifier (Masala G., 2006).

d) Support Vector Machines
Support Vector Machines (SVM) is a machine-learning method, based on the statistical learning theory and the principle of structural risk minimization, which aims to minimize the errors in the data set. Hence, it performs well when applied to data outside the training set. In (Wei, Yang, & Nishikawa, 2009) was investigated an approach based on Support Vector Machines for detection of clusters of microcalcification in digital mammograms. Microcalcifications are detected as a supervised-learning problem and SVM is applied to develop the detection algorithm. SVM is used to detect at each location in the image whether a microcalcification is present or not. The ability of SVM to outperform several well-known methods developed for the widely studied problem of microcalcification detection suggests that SVM is a promising technique for object detection in a medical imaging application.

e) Relevant vector machine
Relevance vector machine (RVM) is another machine learning technique to detect microcalcifications in digital mammograms. RVM is based on Bayesian estimation theory. A distinctive feature of this theory is that it can yield a sparse decision function that is defined by only a very small number of so-called relevance vectors.

In (Wei, Yang, & Nishikawa, 2005) was developed a supervised-learning method through the use of RVM as a classifier to determine at each location in the mammogram if a microcalcification is present or not. To increase the computation a two-stage classification network was developed, in which a computationally simple linear RVM classifier is applied first to quickly eliminate the overwhelming majority non-microcalcification pixels in a mammogram. Comparing with SVM it reduced the computational complexity of the SVM maintaining the detection accuracy.

f) Fuzzy approaches
The fuzzy binary decision tree procedure contains three steps: splitting nodes, determining terminal nodes, and assigning a class to the terminal nodes. A training data set is split into two independent sets and a large tree is grown based on the first training set by splitting until all terminal nodes have pure class membership. Then a pruned sub-
tree is selected by minimizing the second training set misclassification rate. The procedure is then iterated (Thangavel, Karnan, Sivakumar, & Mohideen, 2005).

There are several fuzzy approaches to classify features. For example, in (Seker, Odetayo, Petrovic, & Naguib, 2003) was studied the fuzzy-nearest neighbor (FNN) classifier as a fuzzy logic method. This approach provided a certainty degree for prognostic decision and assessment of the markers. The overall results indicated that the FNN-based method yields the highest predictive accuracy, and that it has produced a more reliable prognostic marker model than the statistical and ANN methods.

On the other hand, in (Grohman & Dhawan, 2001) was described a convex-set-based neuro-fuzzy algorithm for classification of difficult-to-diagnose instances of breast cancer. With its structural approach to feature, it offers rational advantages over the backpropagation algorithm. The training procedure is completely automated-function and parameters are automatically computed from statistical distributions of the data. Two different approaches to construction of fuzzy membership functions were tested: sigmoidal decision surfaces (backpropagation-like approach) and bell-shaped functions cluster-specific approach.

5.3. Analysis of bilateral asymmetry

An additional indicator of the presence of breast cancer is the bilateral asymmetry of the left and right breasts. This is defined by the presence of a greater volume or density of breast tissue without distinct mass or prominent ducts in one breast when compared with the other.

In (Miller & Astley, 1994), a technique to detect breast bilateral asymmetry trough anatomical features was proposed. The method was based on measures of shape, topology, and distribution of brightness in the fibroglandular disk. An accuracy of 74% was obtained. Another method for the detection of breast tumors by analyzing bilateral asymmetry through the measurement of brightness, roughness, and directionality was proposed in (Lau & Bischof, 1991), and a sensitivity of 92% was obtained with 4.9 false positives per mammogram.

Although all work that has been developed, more methods are desirable in this area to analyze asymmetry from multiple perspectives as can improve the detection robustness.

5.4. Summary

There is a substantial literature research regarding detection and classification of masses and calcifications. Commercial CAD systems have satisfactory effectiveness detecting masses and calcifications. However, certain areas of research in CAD of breast cancer still require attention. For example, only a small number of researchers focused on detecting architectural distortions in the absence of mass. And even fewer studies have been done in order to detect bilateral asymmetry. Hence, the development of new breast cancer computer-aided detection is an active research field, particularly regarding the detection of subtle abnormalities in mammograms.

Usually, the CAD softwares integrate common steps: image pre-processing, image enhancement, detection and classification of lesions. There are plenty dissimilar approaches to the different phases. These approaches can still be improved and new approaches or even distinct combination of techniques can be used in order to create better algorithms for more robust and efficient computer aided detection of breast tumors.
6. Conclusion

Breast cancer is the most common cancer among the women world population (Autier, et al., 2010). When early detected, the treatment is more efficient, because the evolution into a more severe stage is avoided, implying less mortality risk.

The breast cancer can be detected through imaging exams, where the most common one is the mammography. This technique aims to detect breast asymptomatic lesions.

Computed aided diagnosis intends to provide assistance to the mammography detection, reducing breast cancer misdiagnosis, thus allowing better treatment and prognosis.

There is a considerable amount of research concerning detection and classification of masses and calcifications. Some commercial CAD systems have satisfactory results on detecting those lesions. However, some subtle abnormalities in mammograms are not successfully detected, as well as when the breast density is higher, the detection of these lesions decreases, increasing the false negatives rates. Besides, there are few researches regarding other common indicators of breast cancer: architectural distortion and bilateral asymmetry.

The effectiveness of CAD in assisting mammogram readers has been the subject of intense debate. However, some of there is no indication that is not a useful tool. Refinement of CAD algorithms would increase its sensitivity and specificity and consequently, would increase the trust from the radiologists in these systems, as the role of CAD is still emerging. The development of the digital mammography may also increase the CAD effectiveness due to the potential to improve contrast resolution compared with film-screen imaging.

Concluding, the development of new breast cancer computer-aided detection is an active research field, particularly regarding the detection of subtle abnormalities in mammograms.
7. Work Plan

The dissertation ensuing from this monography concerns the computational analysis and characterization of mammographic images. Therefore, during this Master project, computational techniques regarding the robust and efficient enhancement and detection of breast lesions in mammographic images will be developed considering cues of architectural distortion and bilateral asymmetry. The techniques will be applied to real cases.

This plan can be divided in the following tasks:

<table>
<thead>
<tr>
<th>Task</th>
<th>Date</th>
</tr>
</thead>
<tbody>
<tr>
<td>Study and development of techniques of mammographic image enhancement</td>
<td>25th March</td>
</tr>
<tr>
<td>Study and development of techniques of mammographic image detection</td>
<td>16th April</td>
</tr>
<tr>
<td>Application of the techniques developed to real cases and further improvements</td>
<td>20th May</td>
</tr>
<tr>
<td>Dissertation writing</td>
<td>Middle June</td>
</tr>
</tbody>
</table>
References


Automatic Analysis of Mammography Images


