



00362

UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

POSGRADO EN CIENCIAS FÍSICAS

INSTITUTO DE FÍSICA

CENTRO DE CIENCIAS APLICADAS Y
DESARROLLO TECNOLÓGICO

CARACTERIZACIÓN DE MAMOGRAFÍAS DIGITALES

T E S I S

*Que para obtener el grado académico de
Maestro en Ciencias (Física Médica)
presenta:*

FLAVIO ERNESTO TRUJILLO ZAMUDIO

Directores:

Dra. María Ester Brandan Siqués

Dr. Jorge Alberto Márquez Flores

Miembro de Comité Tutoral:

Méd. Yolanda Villaseñor Navarro



posgrado en ciencias físicas
u n a m

MÉXICO, D.F.

2005

m. 345651



Universidad Nacional
Autónoma de México

Dirección General de Bibliotecas de la UNAM

Biblioteca Central



UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.

DEDICO ESTE TRABAJO A:

Mis queridos padres:

María Esther Zamudio Blancas y Juan Francisco Trujillo Amozorrutia

Por su cariño y apoyo a lo largo de mi vida, así como en las decisiones que he tomado y en la carrera que elegí.

Mis adoradas hermanas:

Gisela Trujillo Zamudio y Rosa María Trujillo Zamudio

Gracias por su comprensión y oportuna ayuda en momentos de indecisión. Gracias también a **Sashi** quien me compartió de su alegría y cariño.

Mis amados abuelitos:

*Esther Blancas de Zamudio† y Luis Zamudio†
Angelina Amozorrutia de Trujillo† y Juan Trujillo†*

Por su cariño en la infancia y durante mi crecimiento, así como por los consejos recibidos.

Todos mis tíos y primos

*Familias: Trujillo, Zamudio, Amozorrutia,
Guerrero, Flores, Barnard y González*

Por su gran apoyo durante mi formación.

M. en C. *César Gustavo Ruiz Trejo*

Gracias por su apoyo incondicional tanto académico como personal, más allá de un simple profesor. Quien con sus valiosos consejos y amistad supo encaminarme hasta este logro.

La Universidad Nacional Autónoma de México

Gracias por la educación y formación académica recibida.

*“Sólo en el silencio la palabra, sólo en la oscuridad la luz, sólo en la
muerte la vida; el vuelo del halcón brilla en el cielo vacío”*

*La Creación de Ea (An Mago de Ferramar)
Hrsula K. Le Guin*

AGRADECIMIENTOS

A la Dra. María Ester Brandan Siqués y al Dr. Jorge Alberto Márquez Flores, tanto por su ayuda en la dirección de este trabajo de tesis como por su apoyo personal y consejos recibidos a lo largo de la carrera de maestría.

A los integrantes del jurado Dra. María Ester Brandan Siqués, Dr. Fernando Barrios Álvarez, Dr. Neil Charles Bruce Davidson, Dr. Jorge Alberto Márquez Flores y Dra. Yolanda Villaseñor Navarro, por sus imprescindibles comentarios y correcciones en lo correspondiente a la presentación de este trabajo.

Al grupo de Dosimetría de Radiaciones del Instituto de Física de la UNAM: Dra. María Ester Brandan Siqués, M. en C. Ana Elena Buenfil Burgos, Dra. Mercedes Rodríguez Villafuerte, Dra. María Isabel Gamboa de Buen, M. En C. César Gustavo Ruiz Trejo, por haberme recibido en el grupo y por la formación académica recibida. Les estoy enormemente agradecido por toda su ayuda y apoyo que desinteresadamente me brindaron.

A la Universidad Nacional Autónoma de México y al Instituto de Física (IFUNAM) junto con el Centro de Ciencias Aplicadas y Desarrollo Tecnológico (CCADET) por los conocimientos que se me brindaron y que trataré de emplearlos siempre para el bien del país.

Al Instituto Nacional de Cancerología y en especial a la Dra. Yolanda Villaseñor Navarro por su amistad y su disposición en permitirme realizar esta tesis en el Departamento de Radiología. A Erika Islas por la atención brindada. A las técnicas radiólogas Silvia A. Calderón y Rosalía Rosario Miranda, por su valiosa ayuda durante mi trabajo con el equipo de mamografía.

Agradezco al CONACYT (no. de expediente de becario 171932) y a la Dirección General de Asuntos de Personal Académico de la UNAM (proyecto IN101399) por su apoyo económico durante los 2 años de cursos de la maestría. También agradezco al proyecto PAPIIT (no. IN109302) por la ayuda y apoyo económico recibido durante la elaboración de este trabajo de tesis y su correspondiente defensa.

A mis incomparables amigos César Chávez Alfaro, Jesús Arturo Nava Mendoza, Juan José Gajón Bolaños, Juan Alberto Morales García, Mario Jesús Abel Martínez Monte de Oca y Yuri Emilio Villanueva Arreguin, gracias por su ayuda y consejos constantes, por su apoyo tanto en los buenos momentos como en los malos tiempos, por los gratos momentos de juegos de rol al disfrutar de la convivencia y la catarsis experimentada, en fin, gracias por su soporte leal y amistad sincera e incondicional.

A mi buen y sincero amigo de la infancia, Olivier François Marti Velásquez quién siempre ha estado dispuesto a ayudarme y a ofrecerme su apoyo en todo momento. También a su familia quien siempre me ha brindado una mano de apoyo y total amistad.

A mis grandes amigos Jorge Ferrer, Miguel Iburgüengoita, Rogelio Rojas (la comunidad del anillo), por su amistad constante durante la carrera.

A mis compañeros y amigos de la maestría Alejandro Rodríguez Laguna, Amanda García, Eurídice Rioja, Héctor Galván, Héctor Alva, Marvin Rodríguez, Miguel Rodríguez Ponce, Porfirio Díaz Aguirre y Verónica Rodríguez, por sus consejos, apoyo, tanto malos como buenos momentos, indicaciones de la realidad, en fin, por su verdadera y sincera amistad.

Por último y no por eso menos importantes, a mis preciados amigos George Jiménez, David Priego y Elizabeth Cano quienes a lo largo de los años voy perdiendo el contacto visual, pero que su amistad y apoyo sinceros siempre están presentes en todo momento. Les agradezco de todo corazón sus imprescindibles consejos y su más profunda amistad.

"...No conozco a la mitad de ustedes, ni la mitad de lo que querría, y lo que yo querría es menos de la mitad de lo que la mitad de ustedes merece...."

*El Camino sigue y sigue
desde la puerta.
El Camino ha ido muy lejos,
y si es posible he de seguirlo
recorriéndolo con pie decidido
hasta llegar a un camino más ancho
donde se encuentren senderos y cursos.
¿Y de ahí a dónde iré? No podría decirlo.*

*Bilbo Baggins
Lord of the Rings – La Comunidad del Anillo
J. R. Tolkien*

TABLA DE CONTENIDO

CARACTERIZACIÓN DE MAMOGRAFÍAS DIGITALES

INTRODUCCIÓN.....	i
-------------------	---

CAPÍTULO I

PRINCIPIOS BÁSICOS DE MAMOGRAFÍA

1.1. <u>Utilidad de la mamografía</u>	1
1.2. <u>Anatomía de la mama</u>	2
1.3. <u>Hallazgos mamográficos y la estandarización de su informe</u>	2
1.4. <u>Esquemas utilizados para evaluar imágenes mamográficas registradas en película radiográfica</u>	4
1.5. <u>Ensayos de análisis objetivos y cuantitativos en imágenes mamográficas registradas en película radiográfica</u>	7

CAPÍTULO II

FUNCIONAMIENTO DEL EQUIPO DE MAMOGRAFÍA CONVENCIONAL Y DIGITAL

2.1. <u>Equipo de mamografía convencional (analógico)</u>	9
2.1.A. Tubo de rayos X.....	9
2.1.B. Producción de rayos X para mamografía.....	10
2.1.C. Colimación.....	12
2.1.D. Compresión.....	12
2.1.E. Rejilla antidispersora.....	12
2.1.F. Control Automático de Exposición (CAE).....	12
2.2. <u>Descripción general del equipo de mamografía digital</u>	13
2.2.A. Generación de rayos X.....	13
2.2.B. Optimización Automática de Parámetros (OAP).....	13
2.2.C. Receptor de imagen.....	14
2.2.D. Visualización de las imágenes.....	15
2.3. <u>Proceso de adquisición de las imágenes digitales (Senographe 2000D)</u>	15
2.3.A. Detección radiográfica.....	15
2.3.B. Adquisición digital de las imágenes y procesamiento en el equipo.....	17

CAPÍTULO III

ANÁLISIS DE IMÁGENES

3.1. <u>Procesamiento y análisis de imágenes – Teoría básica</u>	19
3.1.A. Imagen digital.....	19
3.1.B. Resolución.....	20
3.1.C. Nitidez.....	20
3.1.D. Contraste.....	20
3.1.E. Brillo.....	21
3.1.F. Formatos de archivos de imagen.....	21
3.2. <u>Realce de imágenes</u>	22
3.3. <u>Restauración de imágenes</u>	23
3.4. <u>Segmentación de imágenes</u>	24

3.4.A. Análisis global.....	25
3.4.B. Análisis mediante bordes.....	27
3.4.C. Análisis por regiones.....	28
3.5. Análisis de características básicas.....	29
3.5.A. Morfología.....	30
3.5.B. Textura.....	30

CAPÍTULO IV MÉTODO EXPERIMENTAL

4.1. <u>Presentación</u>	33
4.2. <u>Selección de pacientes</u>	34
4.3. <u>Análisis mediante procesamiento digital de imágenes</u>	35
4.3.A. Cociente promedio alto/ancho de extremos (ρ).....	35
4.3.B. Compacidad de estructuras (κ).....	37
4.4. <u>Análisis estadístico propuesto</u>	40

CAPÍTULO V RESULTADOS Y ANÁLISIS DE RESULTADOS

5.1. <u>Análisis de histogramas</u>	43
5.2. <u>Fase de entrenamiento</u>	45
5.2.A. Resultados del cociente promedio alto/ancho de extremos (ρ).....	46
5.2.B. Resultados de la compacidad (κ).....	48
5.3. <u>Análisis de la fase de entrenamiento</u>	49
5.3.A. Resultados con el cociente promedio alto/ancho de extremos (ρ).....	50
5.3.B. Resultados con la compacidad (κ).....	52
5.4. <u>Fase de validación</u>	55

CAPÍTULO VI CONCLUSIONES

6.1. <u>Conclusiones</u>	59
6.1.A. Validación del modo de lectura del formato DICOM de la imagen.....	59
6.1.B. Análisis de histogramas (caracterización en intensidades).....	59
6.1.C. Parámetro textural del cociente promedio alto/ancho de extremos (ρ).....	59
6.1.D. Compacidad de estructuras (κ).....	60
6.1.E. Discriminación entre los grupos de imágenes con cáncer y con lesiones benignas.....	60
6.1. <u>Perspectivas a futuro</u>	61

REFERENCIAS

Referencias.....	63
------------------	----

ANEXOS

Anexo A – Aspectos clínicos de la imagen mamográfica

A.1. <u>Anatomía de la mama</u>	A-1
A.1.A. Glándula mamaria.....	A-1
A.1.B. Pezón y areola.....	A-1
A.1.C. Morfología de la mama con la edad.....	A-2
A.2. <u>Evaluación clínica y estandarización de su informe mamográfico</u>	A-3
A.3. <u>Validez de una prueba de diagnóstico</u>	A-4
A.3.A. Sensibilidad.....	A-4
A.3.B. Especificidad.....	A-5
A.3.C. Valor predictivo positivo.....	A-6
A.3.D. Valor predictivo negativo.....	A-6
A.4. <u>Hallazgos mamográficos normales</u>	A-6
A.5. <u>Hallazgos mamográficos anormales</u>	A-8
A.5.A. Lesiones ocupantes.....	A-8
A.5.B. Calcificaciones.....	A-10
A.6. <u>Ejemplos de lesiones benignas y malignas</u>	A-10

Anexo B – Aspectos técnicos en el análisis de imágenes digitales

B.1. <u>Consideraciones referentes al análisis de las mamografías digitales</u>	B-1
B.1.A. Matlab v7.....	B-1
B.1.B. Mamografías digitales en formato DICOM.....	B-1

Anexo C – Obtención de la mascarilla binaria

C.1. <u>Proceso empleado para la obtención de la mascarilla binaria</u>	C-1
C.1.A Dilatación (\oplus) y erosión (\ominus) morfológicas.....	C-1
C.1.B. Apertura (\circ) y cerradura (\bullet) morfológicas.....	C-2

INTRODUCCIÓN

El cáncer de mama es una de las causas más importantes de mortalidad y morbilidad por neoplasia maligna en Norte América. Se estima que aproximadamente 1 de cada 8 mujeres desarrollará cáncer de mama a lo largo de su vida [By96]. En México, el cáncer mamario ocupa el segundo lugar en la mortalidad de mujeres por este tipo de padecimientos, superado tan sólo por el cáncer cérvico-uterino [WHO]. Debido a que las causas y/o medios de prevención para este tipo de cáncer no son aún conocidos, la identificación de las características sospechosas de malignidad en la imagen mamográfica se vuelve de particular importancia. La información que se pueda extraer de estas características es de interés para mejorar la evaluación por imagen mamográfica.

La mamografía es un examen radiográfico que puede ser dividido en dos categorías, escrutinio (o tamizaje) y diagnóstico. El escrutinio sirve para la detección temprana de cáncer a través de la observación de lesiones no palpables en mujeres sanas. El examen de diagnóstico arroja información acerca de la localización, el número y las características de una lesión palpable. Un examen mamográfico típico consiste en la obtención de 2 proyecciones radiográficas de cada mama, resultando 4 imágenes en total.

La mama se compone principalmente de tejido fibroglandular, tejido adiposo y de sostén. El tejido fibroglandular es una mezcla de tejido conectivo fibroso (estroma) y tejido glandular (células epiteliales y ductos lactíferos). El tejido adiposo es más radiolúcido que el tejido fibroglandular; de esta forma este tipo de tejido aparece más oscuro en una mamografía. Las regiones más claras se asocian con el tejido fibroglandular y se les refiere como “de alta densidad mamográfica”, o tejido “mamográficamente denso”.

La variación en la composición de la mama entre mujeres resulta en un amplio intervalo de apariencias visuales en la imagen mamográfica. La composición de la mama tiene un impacto considerable en la mamografía, ya que es muy difícil de obtener un buen contraste en las mamas de las mujeres con mucho tejido mamográficamente denso, el cual puede enmascarar signos sutiles de anormalidades mamarias. La mama densa es considerada como un factor que contribuye en la interpretación mamográfica de falsos-negativos (ver anexo A.3).

Hay estudios actuales que discuten el hecho de que el incremento de la densidad radiográfica de la mama esté asociado con un elevado riesgo de cáncer mamario; sin embargo los mecanismos de esta relación no se han explicado. Por otra parte una alta densidad radiográfica de la mama está presente en una gran proporción de los casos de cáncer mamario. El interés de aseverar la posible existencia de una correlación entre la densidad radiográfica de la mama y el riesgo de cáncer de mamario, radica en la posibilidad de comprender mejor las causas de cáncer en la mama. Esto, identificando los factores asociados al tejido mamográficamente denso y determinando la forma en que tales tejidos

cambian con factores como la edad, la densidad mamográfica, la región de nacimiento del paciente, el estatus socioeconómico, la edad de la primera concepción, el previo desarrollo de cáncer de mama, la vida sexual, la no concepción, la edad de la aparición de la menopausia y la menarca, el peso (postmenopausia) y la previa aparición de cáncer (ovarios o endometrio), entre otros. Más aún, dado que la densidad radiográfica de la mama puede ser modificada, ésta puede ser una forma para el desarrollo de estrategias de prevención tales como la liberación de hormonas inhibitoras de gonadotropina, el uso de tamoxifen y la adaptación de una dieta baja en grasas y alta en carbohidratos [Bo96].

A partir de la evaluación de imágenes analógicas (película radiográfica), se han determinado ciertas características de la morfología de la mama, considerando alteraciones anatómicas. Desde los años 70 se ha intentado caracterizar un factor de probabilidad de aparición de cáncer mamario (riesgo de cáncer), a través de la evaluación de la densidad radiográfica de la mama. Se han probado diversos esquemas para obtener una medición de la densidad mamográfica y así poder asociar un factor de riesgo adecuado. En un principio se utilizaron clasificaciones subjetivas haciendo varias divisiones de los casos en que la imagen mamográfica se visualizaba poco densa o muy densa, asignando valores de riesgo a cada una [Wo76a, Wo76b]. Sin embargo la clasificación dada por un observador para evaluar una mamografía puede llevar a considerables variaciones entre observadores (60%) e incluso con él mismo al utilizar estas divisiones subjetivas. De esta forma, la posibilidad de medir las pequeñas diferencias en una imagen mamográfica sería de interés en la investigación de cambios con la edad o con una intervención potencial que pudiera alterar a la mama.

Otros factores de la imagen mamográfica influyen en la dificultad de poder detectar el cáncer mamario. Debido a la naturaleza de la imagen radiográfica, la mamografía representa la proyección de un objeto con volumen en el plano de la imagen. De esta forma los tejidos mamaros se superponen en la imagen mamográfica resultando en diferentes configuraciones radiográficas. Así, se han determinado configuraciones características de anormalidades mamaras con probabilidad de ocurrencia de cáncer de mama. Estas configuraciones pueden presentar microcalcificaciones (aparecen en la imagen como pequeños grupos de motas blancas no pertenecientes a la anatomía regular de la mama) o masas tumorales (zonas de mayor densidad mamográfica observadas en las dos proyecciones). Las microcalcificaciones se pueden presentar en ductos lactíferos, y se pueden visualizar en la imagen mamográfica a través de cúmulos pequeños. Las masas tumorales presentan 4 formas características, esférica, oval, lobular e irregular (sin forma visible definida) [Ko98].

Para sugerir la malignidad de una anormalidad se observan las configuraciones que se presentan en la imagen mamográfica, así como las diferencias en simetría entre ambas mamas. Sin embargo estas medidas resultan ser subjetivas al depender completamente del observador para evaluar la mamografía, lo que puede llevar a considerables variaciones entre observadores. Se observa que estas variaciones o diferencias de diagnóstico disminuyen si se tiene la cooperación de un segundo observador durante la evaluación mamográfica, lo que resulta en general impráctico. Además, en muchas ocasiones las configuraciones que se desean analizar pueden estar enmascaradas debido a un aumento en la densidad mamográfica o a la superposición de tejidos en la imagen. Estas dificultades

contribuyen a diagnósticos falsos-negativos. Además, se debe tomar en consideración que el diagnóstico definitivo para malignidad requiere de un estudio histopatológico.

La posible aplicación de mediciones cuantitativas en la imagen mamográfica, aprovechando la utilidad de un equipo de mamografía digital, durante los exámenes de escrutinio pueden servir a las mujeres de seguimientos más estrictos en un programa de tamizaje (escrutinio o “screening” en inglés). Utilizar las herramientas de procesamiento de imágenes digitales puede ayudar también a la reducción de diagnósticos falsos-negativos y a la reducción de la morbilidad del cáncer de mama.

En propuestas de estudios automatizados para obtener evaluaciones cuantitativas de la densidad mamográfica se observa que el análisis de una sola imagen mamográfica por paciente resulta en un valor representativo. Gracias a eso se reduce el espacio de almacenamiento y el análisis de las otras imágenes del estudio mamográfico [By96].

Con el desarrollo de las herramientas de cómputo se propusieron otros tipos de análisis cuantitativos, semiautomáticos y automáticos, al tratar de analizar las regiones de densidad mamaria y alteraciones en la anatomía de la mama. Estos métodos [Bo95, Bo99, By94, By96, By97, Ya98], han utilizado imágenes digitalizadas a partir de las obtenidas en película radiográfica. Además, se han centrado en el análisis de técnicas del procesamiento de imágenes como realce de imágenes, segmentación por umbralaje, análisis de áreas y medición de dimensiones fraccionarias o fractales. Otros trabajos [Ch98, Ch01, Sa01], han contribuido utilizando otras técnicas como segmentación de objetos de interés a través de diferentes parámetros morfológicos y texturales. Todos estos trabajos han proporcionado resultados en el cálculo de factores de riesgo (basados en el aumento de la densidad mamográfica) y caracterización de la mama, pero algunos de ellos obteniendo resultados parcialmente significativos en comparación con los métodos subjetivos y sin llegar aún a conclusiones satisfactorias.

Sin embargo, en la actualidad se han desarrollado equipos de obtención de imágenes directamente en formato digital [GE00] sin tener que pasar por un proceso analógico (película radiográfica). Esto trae consigo una nueva etapa en la aproximación del estudio de las imágenes mamográficas, al poder aplicar directamente en el archivo digital de la imagen las herramientas del procesamiento de imágenes.

Con esto en mente, este trabajo de tesis propone obtener parámetros característicos (a partir de las imágenes resultantes de un estudio de mamografía digital), de las imágenes de pacientes diagnosticadas con cáncer, a través de un estudio histopatológico, con respecto de las que, presentando una imagen radiográficamente similar, su biopsia indica que no tienen cáncer.

El objetivo de este trabajo es obtener parámetros característicos de diagnóstico positivo de cáncer de mama en imágenes mamográficas digitales.

Para cumplir este objetivo las metas a lograr son:

- i. Obtención de imágenes de mama de pacientes informadas como BIRADS 4 Y 5 en un estudio digital.

- ii. Caracterización de las mamografías utilizando técnicas de procesamiento de imágenes. Proponemos en particular el análisis de texturas y formas.
- iii. Análisis estadístico de las características de estas imágenes considerando grupo de estudio las pacientes con estudio histopatológico positivo y grupo de control aquéllas con biopsia negativa.

En el capítulo I se describen los principios básicos de la mamografía, la anatomía regular de la mama y las diferentes anomalías mamarias. Se presentan y comentan los medios utilizados anteriormente para el análisis médico de las mamografías a través de observaciones subjetivas. Finalmente se exponen los diferentes intentos por obtener un análisis cuantitativo y objetivo de las imágenes mamográficas para diagnosticar con mayor certeza el cáncer mamario.

En el capítulo II se explica el funcionamiento del mastógrafo convencional (analógico) y el mastógrafo digital con el que se obtienen las imágenes mamográficas de este estudio.

En el capítulo III se expone, en general, la teoría del tratamiento de imágenes digitales. Se comentan las características y formatos de las imágenes digitales, se presentan las formas de análisis mediante estudio global, regional y local de las imágenes digitales, así como ejemplos de los parámetros que se pueden extraer.

En el capítulo IV se describe el método experimental empleado para la obtención de las imágenes mamográficas de pacientes que cumplan los criterios establecidos y la forma de su evaluación digital.

En el capítulo V se presentan los resultados de las pruebas, obtenidos siguiendo los procedimientos presentados en el capítulo anterior. Se analizan los resultados para la evaluación de la eficiencia de la caracterización realizada a las imágenes digitales.

En el capítulo VI se ofrecen las conclusiones y perspectivas del presente trabajo de tesis.

CAPÍTULO I

PRINCIPIOS BÁSICOS DE MAMOGRAFÍA

A continuación se presentan los conceptos y principios básicos del estudio de mamografía. Se describe la anatomía regular de la mama y las diferentes anomalías mamarias, se presentan y comentan los métodos de análisis subjetivos establecidos, y los intentos por obtener un análisis cuantitativo. El contenido de este capítulo se basa principalmente en las siguientes referencias: [Bl00], [Br00], [By94], [Ca90], [Ki00], [Ko98], [Sa01] y [Zh01].

1.1. UTILIDAD DE LA MAMOGRAFÍA

La formación de una imagen mamográfica tradicional se logra mediante un haz de rayos X que atraviesa la mama de la paciente y que finalmente incide sobre una película radiográfica. Un estudio regular de mamografía consiste de 4 imágenes. Una, craneocaudal (CC), la cual consiste de una proyección vertical superior a inferior de la mama con el eje del haz de rayos X paralelo a la normal del plano de la imagen. Otra, mediolateral-oblicua (MLO), la cual consiste de una proyección a 45° con respecto del eje de simetría del cuerpo de la paciente. La proyección MLO, abarca una mayor cantidad de tejido mamario, incluyendo la totalidad del cuadrante superoexterno y la cola axilar de la mama. En comparación con la proyección MLO, la incidencia de la proyección CC permite una mejor visualización del aspecto medial de la mama y una imagen más detallada en la medida en que posibilita una mayor compresión de la glándula. Estas vistas se toman a cada mama de la paciente resultando en 4 imágenes por estudio, sin embargo, se pueden tomar proyecciones adicionales según lo dictamine el médico radiólogo para una correcta evaluación mamográfica.

La mamografía es un estudio radiográfico, que puede tener dos objetivos: escrutinio y diagnóstico. El escrutinio (o tamizaje) es útil en la detección temprana de cáncer a través de la observación de lesiones no palpables en mujeres asintomáticas. El escrutinio se lleva a cabo a través de programas de revisión o referidos por su nombre en inglés como “screening”, los cuales consisten de un estudio regular de mamografía una vez al año a partir de los 40 años de edad, o pudiendo variar de acuerdo a la evaluación del médico radiólogo. El examen de diagnóstico brinda información sobre la localización de un posible tumor canceroso en mujeres a las cuales se les ha encontrado algún tipo de afección mamaria, como la presencia de un nódulo mamario o secreciones a través del pezón. El examen de diagnóstico generalmente es adaptado a los hallazgos clínicos o a una anomalía detectada durante el escrutinio. En mamografía, la sensibilidad se define como la probabilidad de detección de un resultado anormal en presencia de cáncer mamario y la especificidad como la probabilidad de un resultado normal en ausencia de enfermedad. La mamografía es un importante instrumento de diagnóstico en mujeres con síntomas mamarios, dado que permite definir características de las lesiones o anomalías mamarias, orientando a un diagnóstico de benignidad o malignidad, donde este último debe ser corroborado por una biopsia. La mamografía también se utiliza para dirigir diversas intervenciones mamarias, como la localización con aguja previa a la biopsia, la aspiración con aguja y la ductografía.

El llamado “estándar de oro” en mamografía se refiere a tomar siempre en cuenta que el estudio mamográfico sirve exclusivamente para la detección de anomalías mamarias en mujeres asintomáticas y a la localización de las lesiones (hallazgos malignos o benignos). Sin embargo no se puede servir exclusivamente de la mamografía para emitir un diagnóstico completo, y se requiere de métodos adicionales como un estudio por ultrasonido. Para determinar con certeza la malignidad de un tejido se realiza un estudio histopatológico. Para esto último se recurre a efectuar una biopsia, la cual consiste en la toma de una muestra del tejido mamario por medio de una incisión en la mama y en la evaluación de la patología de la muestra.

1.2. ANATOMÍA DE LA MAMA

La mama se compone principalmente de tejido fibroglandular y tejido adiposo. El tejido fibroglandular es una mezcla de tejido conectivo fibroso (estroma) y tejido glandular (células epiteliales y ductos lactíferos). El tejido adiposo se encuentra principalmente recubriendo los conductos galactóforos cerca de la dermis y como grasa intersticial entre ellos. La mama se fija a la piel por los ligamentos suspensorios de Cooper y es separada de la aponeurosis de revestimiento del músculo pectoral mayor por la bolsa retromamaria. Los ligamentos de Cooper forman tabiques fibrosos en la estroma que brindan apoyo al parénquima mamario. De 15 a 20 conductos galactóforos se extienden desde los lobulillos compuestos por epitelio glandular hasta los orificios situados en el pezón. Una dilatación del conducto, el seno galactóforo, se encuentra cerca del orificio del conducto en el tejido subareolar. El tejido subcutáneo y el tejido adiposo distribuidos alrededor de los lobulillos de la glándula confieren a la mama su contorno uniforme y en el período sin lactancia representan la mayor parte de su volumen. Los vasos linfáticos atraviesan la estroma que rodea a los lobulillos de la glándula y transportan la linfa hasta los conductos colectores. En la figura 1.1 se muestra una vista tangencial de la mama sobre la pared torácica y vista de un corte (sagital) de la mama y la pared torácica asociada (para mayor información sobre la anatomía de la mama ver anexo A.1).

1.3. HALLAZGOS MAMOGRÁFICOS Y LA ESTADANDARIZACIÓN DE SU INFORME

Para la evaluación de las imágenes mamográficas se ha implementado una terminología estandarizada: el “Breast Imaging Reporting and Data System” (BI-RADS) del Colegio Estadounidense de Radiología, con la finalidad de evitar la confusión durante la interpretación de los informes y facilitar la vigilancia ulterior de las pacientes [AC98], [Do93]. En el presente trabajo de tesis se emplearán más adelante imágenes evaluadas como BI-RADS 4 y 5, las cuales significan, hallazgo probablemente maligno y hallazgo altamente sugestivo de malignidad, respectivamente (para mayor información sobre la terminología BI-RADS empleada, ver anexo A.2).

Para la evaluación de los métodos de detección (subjetivos o cuantitativos), se suele hacer referencia a esquemas de análisis que cuentan el número de casos falsos-positivos (el

sistema detecta cáncer de mama en una mujer “sana”), falsos-negativos (FN), verdaderos-positivos (VP) y verdaderos-negativos (VN). La asignación positiva (P) o negativa (N) de cáncer de mama que se asocia a la imagen está dada por el sistema o método de evaluación de la imagen mamográfica, mientras que la asignación falsa (F) o verdadera (V) con la que se califica la asignación P o N esta dada por el dictamen del patólogo (previo análisis histopatológico de una muestra del tejido mamario afectado). Un ejemplo de este tipo de resultados se muestra en la tabla 1.1 para un estudio de 5000 mamografías practicadas en USA a mujeres entre 50 y 54 años [Va98]. Para mayor información sobre validez de una prueba diagnóstica mediante este tipo de evaluaciones referirse al anexo A.3.

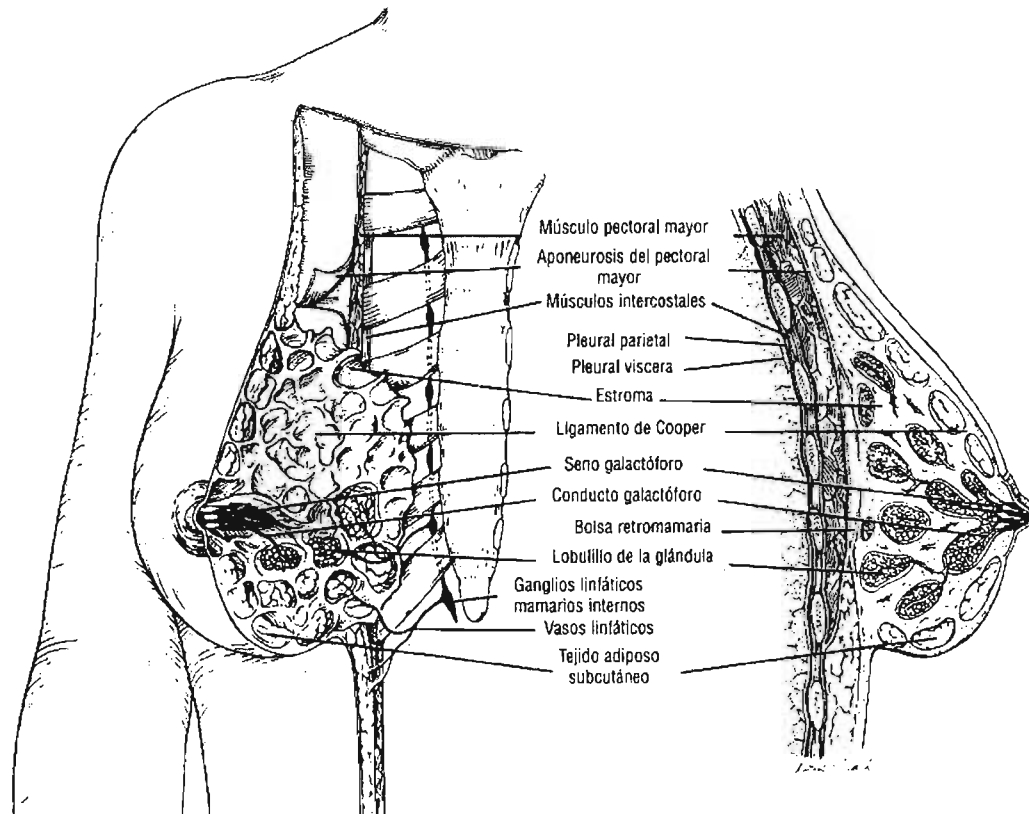


Figura 1.1. Vista tangencial de la mama sobre la pared torácica y vista de un corte (sagital) de la mama y la pared torácica asociada (ver texto).

Existe una amplia gama de hallazgos mamográficos normales en lo que concierne al tamaño, la configuración y la composición de los tejidos mamaros. La composición del tejido mamario puede variar entre la presencia casi exclusiva de tejido adiposo y un tejido fibroglandular extremadamente denso, y estas características se correlacionan con la sensibilidad de la mamografía. Algunos signos del cáncer de mama se muestran por la aparición de microcalcificaciones o por masas tumorales, las cuales tienen coeficientes de atenuación tales que en la imagen radiográfica se observan blancas (mayor atenuación del haz de radiación), por lo que se les denomina radiopacas. El tejido adiposo con coeficientes de atenuación menores que los anteriores, se le dice radiolúcido pues se observa de color gris oscuro o negro en la mamografía. El tejido adiposo representa un buen fondo sobre el cual puede detectarse un cáncer de mama pequeño, pero el tejido fibroglandular denso (de

color blanco en la mamografía) puede enmascarar un cáncer de mama (para mayor información sobre hallazgos mamográficos normales, ver anexo A.4).

Tabla 1.1. Sistema de calificaciones a través de evaluaciones: VP, VN, FP y FN.

CANCER MAMARIO:		
	Presente	Ausente
Positiva	9 (Verdaderos-Positivos)	500 (Falsos-Positivos)
MAMOGRAFIA:		
Negativa	1 (Falsos-Negativos)	4490 (Verdaderos-Negativos)
Sensibilidad : $VP / (VP + FN)$	→	$9 / (9 + 1) = 90\%$
Especificidad: $VN / (VN + FP)$	→	$4490 / (4490 + 500) = 89.97\%$
Valor Predictivo Positivo: $VP / (VP + FP)$	→	$9 / (9 + 500) = 1.77\%$
Valor Predictivo Negativo: $VN / (VN + FN)$	→	$4490 / (4490 + 1) = 99.97\%$
Exactitud Diagnóstica: $(VP + VN) / (VP + VN + FP + FN)$	→	$9 + 4490 / (9 + 4490 + 500 + 1) = 89.98\%$

Para mayor información sobre valor predictivo positivo o negativo, ver anexo A.3.

Las alteraciones más frecuentes observadas en las mamografías consisten en lesiones ocupantes y calcificaciones, y las características radiográficas de estas alteraciones pueden aportar indicios etiológicos (para mayor información sobre lesiones ocupantes y microcalcificaciones, ver anexo A.5). En el informe mamográfico estandarizado los términos utilizados para describir las lesiones ocupantes y las calcificaciones indican la probabilidad de malignidad. Otros hallazgos significativos consisten en la presencia de lesiones neoformadas o evolutivas, la distribución asimétrica bilateral del tejido fibroglandular, distorsiones de la arquitectura glandular, aumento del espesor o retracción de la piel, retracción del pezón y aumento de tamaño de los ganglios linfáticos (para más ejemplos de lesiones benignas y malignas, ver anexo A.6).

1.4 ESQUEMAS UTILIZADOS PARA EVALUAR IMÁGENES MAMOGRÁFICAS REGISTRADAS EN PELÍCULA RADIOGRÁFICA

A mediados de 1970, los estudios realizados por Wolfe [Wo76a, Wo76b], sugirieron que existe una asociación entre los patrones del parénquima mamario y el riesgo de desarrollar cáncer de mama. Desde entonces ha habido muchos estudios buscando la relación entre la densidad mamográfica y el riesgo de desarrollar cáncer de mama. En los términos utilizados por los radiólogos “alta densidad mamográfica” se usa como sinónimo de atenuación de la radiación, es decir, tejidos que aparecen blancos en la imagen mamográfica. Esto es lo contrario del término físico “densidad óptica”.

Los patrones del parénquima mamario se han utilizado como un indicador del riesgo de desarrollar cáncer de mama. Mientras que otros métodos se han sugerido, es común para los médicos radiólogos clasificar las mamografías con las cuatro “categorías de Wolfe”, con cada categoría correspondiente a un patrón diferente de densidad mamográfica, tal como se describe en la tabla 1.2.

Tabla 1.2. Esquema de Wolfe.

Categoría	Descripción
N1	Principalmente grasa, con muy pocas densidades y sin ductos visibles
P1	Principalmente grasa, pero con ductos ocupando menos del 25% de la mama
P2	Tiene ductos prominentes ocupando más del 25% de la mama
DY	La imagen se visualiza difusa, con regiones irregulares presentando un aumento de la densidad, pero sin tener ductos visibles

Varios estudios se han realizado para analizar la correlación entre los patrones del parénquima mamográfico y el riesgo de cáncer de mama. Incluso se ha sugerido que el incremento del riesgo de cáncer de mama en mujeres con la clasificación P2 y DY es un artefacto que aparece a partir del prejuicio en la detección de cáncer debido a la variación en la densidad radiográfica del tejido mamario en la mujer con diversos patrones. Mientras que la controversia aún existe, se estima que la mujer con mamografías P2 y DY tiene un mayor riesgo de desarrollar cáncer de mama que las mujeres con mamografías N1 y P1.

Además de los patrones de Wolfe, muchos estudios [Sa91] han demostrado que la extensión de la alta densidad mamográfica se asocia con el riesgo de desarrollar cáncer de mama. En estos estudios, el porcentaje del área proyectada de la mama ocupada por densidades mamográficas fue subjetivamente clasificado en grandes categorías. Las estrategias de categorización variaron, tanto en el número de categorías como en la extensión de la densidad que representan. Una clasificación típica de seis categorías (SCC) se ilustra en la figura 1.2. En este esquema, las categorías se definen por: “nada”, menor que 10%, 10-25%, 25-50%, 50-75% y mayor que 75% del área proyectada por la apariencia de la mama como densidad mamográfica. Warner et al [Wa92] en un primer análisis, compararon diferentes métodos de clasificación subjetiva para el parénquima mamográfico; concluyeron que técnicas como las SCC proveen un mejor índice de riesgo que las clasificaciones descritas por Wolfe.

Como un indicador de riesgo, las categorías de Wolfe presentan grandes problemas. Una es la dificultad para un observador humano definir las 4 categorías con precisión. Aún más, las bases físicas o biológicas para el sistema de clasificación es problemático dado que se determina solamente con dos factores; la cantidad de tejido fibroso denso presente y el grado de prominencia ductal. Los límites entre las categorías de Wolfe se basan en elecciones arbitrarias observando la importancia relativa de estos dos factores. Además, la evaluación de las categorías de Wolfe es subjetiva y tiene un alto grado de variabilidad entre observadores e incluso con uno mismo. Dada esta variabilidad, las recomendaciones para el seguimiento de pacientes basado en las categorías de Wolfe muestran ser altamente dependientes del observador.

Debido a su naturaleza discreta, las clasificaciones de Wolfe no son útiles para asignar cambios (otros que no sean grandes), en mamografías de pacientes tomadas en lapsos de

meses o años. Sin embargo, la adición de más categorías al sistema de clasificaciones de Wolfe haría la tarea del médico radiólogo más difícil y susceptible de mayor variabilidad.

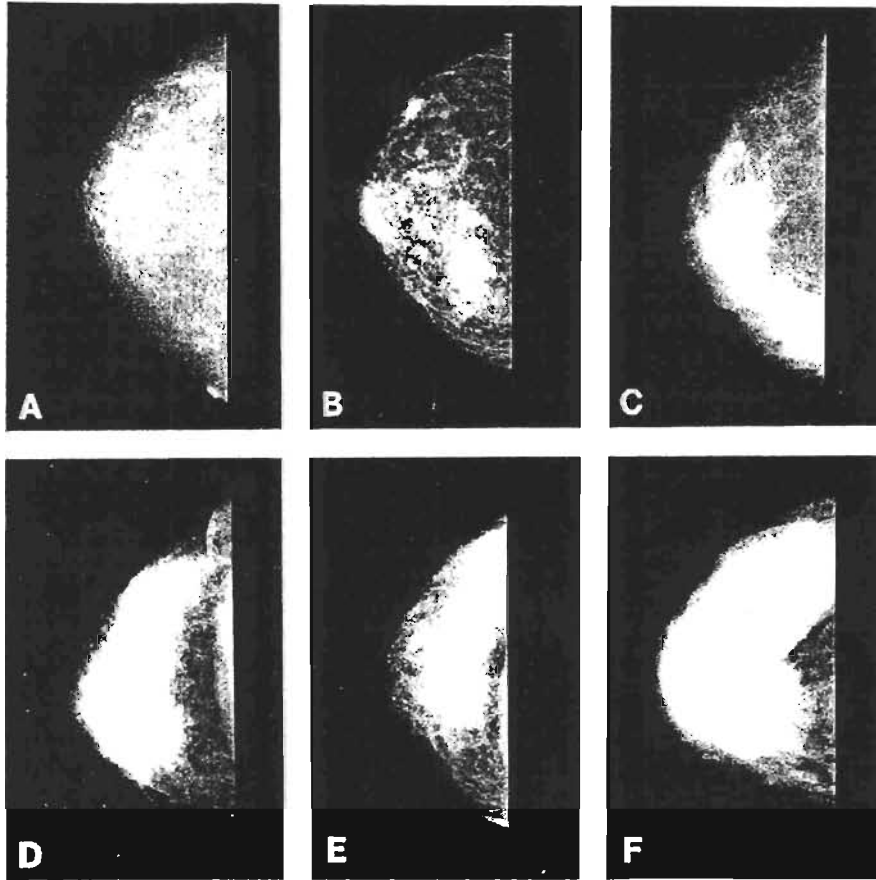


Figura 1.2. Mamografías obtenidas por película-pantalla, en la proyección craneocaudal, correspondientes a las categorías de densidad en la clasificación subjetiva de seis categorías de densidad mamográfica.

En la práctica clínica, los médicos radiólogos rutinariamente estiman la densidad mamográfica en las mamografías a través de la terminología BI-RADS como lo recomienda el Colegio Estadounidense de Radiología [AC98] para proveer una referencia para la sensibilidad mamográfica. Debido a la falta de un método cuantitativo para la estimación de la densidad mamográfica, los investigadores usualmente emplean la terminología BI-RADS para la vigilancia de las respuestas en tratamientos preventivos o intervencionistas y en los cambios asociados con el riesgo de cáncer de mama [Wh98, Bo96]. Se ha encontrado una gran variabilidad entre observadores en la interpretación del BI-RADS entre médicos con experiencia [Br04, Zh01]. Una estimación automática y cuantitativa, proveería un medio eficiente para medir la densidad mamográfica y también una estimación reproducible que reduciría la variación entre observadores en las mediciones de la densidad mamográfica.

Un hecho atractivo que se supone es que el riesgo asociado con los patrones mamográficos varía continuamente con los cambios en el patrón, en lugar de sólo cuatro pasos discretos. Siguiendo esta idea, se han desarrollado técnicas cuantitativas para evaluar cambios en los patrones del parénquima mamario en series de mamografías. Tales evaluaciones se requieren en estudios que involucran, por ejemplo, dietas o intervención hormonal, y los

cuales usan los patrones del parénquima vistos en series de mamografías como indicador de riesgo [Bo96].

1.5 ENSAYOS DE ANÁLISIS OBJETIVOS Y CUANTITATIVOS EN IMÁGENES MAMOGRÁFICAS, REGISTRADAS EN PELÍCULA RADIOGRÁFICA

Con las estimaciones cuantitativas de la densidad mamográfica, la diferencia de riesgo entre la más alta y la más baja categoría de riesgo es substancial y es mayor que los otros factores de riesgo para cáncer de mama.

Actualmente, las mamografías son analizadas visualmente por los médicos radiólogos. Sin embargo, un método computacional para analizar las características mamográficas sería útil como un suplemento en la evaluación del médico radiólogo. Investigaciones previas en el diagnóstico asistido por computadora (CAD) para la detección del cáncer de mama se han concentrado principalmente en la detección de masas tumorales y microcalcificaciones en las mamografías utilizando técnicas de visión computacional. Se ha demostrado que un algoritmo efectivo de CAD puede aumentar la precisión del diagnóstico de cáncer de mama en las mamografías, lo que reduce a su vez las biopsias innecesarias.

Numerosos grupos han trabajado en el problema de clasificar estructuras sospechosas en tipos normales o anormales. Varios grupos aplican algoritmos para delimitar (segmentar) una región sospechosa [Ka98, Br99, Sa96, Po97, Ke94, Yi93], pero en ocasiones las características de las imágenes (texturas en la imagen, por ejemplo), se computan sobre una gran región conteniendo la estructura sospechosa. La segmentación de la región sospechosa es útil para separar tejido anormal y normal, ya que esto permite el cálculo de características relacionadas al borde de la región, tanto como características de contraste y forma.

El tipo de características que serán útiles en la etapa de rechazo de falsos positivos (imágenes dictaminadas como conteniendo tejido normal, siendo que en realidad hay tejido anormal), depende del tipo de región que se seleccione en la etapa de detección inicial. Por el momento, si la etapa de detección genera muchos falsos positivos en una región con ductos galactóforos, una característica que detecte ductos podría ser útil. Por otra parte, si todas las áreas brillantes se señalan, el análisis de forma de esas regiones podría ser una aproximación útil. El rechazo de falsos positivos no es lo mismo que clasificar entre masas tumorales benignas y malignas. Para discriminar entre masas malignas y benignas, el análisis de bordes es muy importante. La remoción de falsos positivos podría requerir de otros tipos de características, ya que varios son debidos a la proyección o son inducidos por el borde del músculo pectoral.

Análisis de patrones basados en la idea de la “dimensión fractal” se han aplicado en diversos campos. El uso del concepto fractal se basa principalmente en el trabajo de Mandelbrot, quien mostró que muchos sistemas físicos en la naturaleza muestran “invariancias en escala” o rasgos de “autosimilitud” a distintas escalas. Tales sistemas pueden considerarse que tienen una “dimensión fraccionaria” particular, característica del

sistema físico en cuestión. Otros han aumentado el trabajo de Mandelbrot describiendo sistemas biológicos y el problema de describir la textura de las imágenes, como en [Sa01] donde se describe un método automático para segmentar tejido radiopaco de tejidos grasos en una mamografía convencional. La dimensión fractal ha mostrado correlacionarse fuertemente con la observación humana subjetiva de la textura en las imágenes.

El análisis asistido por computadora en las imágenes mamográficas proveería una medición objetiva y cuantitativa en la detección del cáncer de mama. Los esfuerzos en el procesamiento de imágenes encaminado hacia esta meta parecieran ser escasos, y métodos automáticos y eficientes para generar una medición de la densidad mamográfica parecieran ser aún menores. Boyd et al [Bo95] estudiaron la relación entre las densidades mamográficas y el riesgo de cáncer de mama usando a la vez la clasificación del médico radiólogo y la medición de la densidad asistida por computadora. La medición asistida por computadora estuvo basada en el umbralaje interactivo de las densidades usando dos selecciones de valores umbrales. Se observó estadísticamente un aumento significativo en el riesgo de cáncer de mama asociado con el incremento de la densidad mamográfica en ambos métodos.

Boone et al. [Bo98], desarrollaron y evaluaron un método computarizado para calcular un índice de densidad mamario y comparó este índice con el índice de densidad mamario provisto por los médicos radiólogos. Byng et al. [By96] hicieron un análisis cuantitativo de la simetría entre las mamografías de mamas diferentes de una misma paciente y las mamografías en proyecciones diferentes de la misma mama en vías de una clasificación subjetiva, umbralaje interactivo, medición de la asimetría regional, y análisis de textura. Ursin et al. [Ur98] estudiaron el cambio de las densidades mamográficas en las mujeres a través de una prueba de hormona agonista liberadora de gonadotropina (GnRHA) en régimen base para la prevención de cáncer de mama. Las imágenes las evaluaron usando evaluación simultánea, delimitación de las áreas de densidad por un médico radiólogo con experiencia, y métodos de segmentación por umbralaje computacionales no expertos. Los tres métodos anteriores observaron una reducción estadísticamente significativa en las densidades iniciales hasta los 12 meses de seguimiento. Se encontró una alta correlación entre los resultados computacionales por delimitación regional y los basados en delimitar regiones por expertos. Huo et al. [Hu98] estudiaron la habilidad de la extracción computacional de características, a través de analizar el parénquima mamario en un área de interés seleccionada a partir de la región central de la mama, junto con la edad, para identificar a las mujeres en riesgo. Se encontró que una caracterización computacional de los patrones del parénquima podrían asociarse con el riesgo de cáncer de mama. Un método automático para la segmentación de la región del parénquima de una mamografía y el análisis de las densidades mamográficas usando el tercer momento en los histogramas de niveles de gris se presenta en [By96*]. Otra aproximación se presenta en [Hi96] para determinar el volumen de tejidos no grasos en las mamografías. Saha et al. [Sa01], proponen un método automático y reproducible para cuantificar densidades mamográficas y estudiar la precisión de los parámetros relacionados. El método empleado utiliza una escala base de conectividad borrosa. Un método de usuario interactivo asistido por computadora para cuantificar la densidad mamográfica se publicó en [Bo95], el cual concluyó que la clasificación cuantitativa de densidades permite la determinación de gradientes de riesgo más específicos que los usados en las clasificaciones de Wolfe.

CAPÍTULO II FUNCIONAMIENTO DEL EQUIPO DE MAMOGRAFÍA CONVENCIONAL Y DIGITAL

A continuación se describe el equipo de mamografía analógico, y digital principalmente en la forma de adquisición de la imagen. El contenido de este capítulo se basa principalmente en las siguientes referencias: [GE00], [Ky03], [Ta03], [Ve99], [Ya97].

2.1. EQUIPO DE MAMOGRAFÍA CONVENCIONAL (ANALÓGICO)

Para la producción de una mamografía se utiliza un equipo exclusivo para tal efecto; durante 50 años se ha utilizado un equipo analógico como el que se ilustra en la figura 2.1, i.e. donde la imagen final se obtiene en una película radiográfica. A continuación se describen las partes esenciales del equipo de mamografía, necesarias para la producción de una imagen mamográfica.

2.1.A. Tubo de rayos X

El tubo de rayos X consiste de un tubo de vidrio sellado y evacuado, que en su interior contiene dos electrodos que se mantienen a una diferencia de potencial (ver figura 2.2). Un electrodo es el cátodo el cual se encuentra a un potencial negativo y contiene un filamento que se calienta y emite electrones (emisión termoiónica). El otro electrodo es el ánodo; éste se encuentra a un potencial positivo y cuenta con un blanco metálico. En este blanco chocan los electrones emitidos por el cátodo al ser acelerados por la diferencia de potencial entre los electrodos, para producir los rayos X. La energía de los rayos X es proporcional a la tensión aplicada al tubo; la intensidad del haz de rayos X depende de la diferencia de potencial y de la corriente en el tubo (flujo de electrones). Sin embargo, para las tensiones utilizadas en mamografía, algunos de los electrones emitidos por el cátodo no son atraídos por el ánodo y forman una carga espacial. Este incremento de electrones en la vecindad del cátodo repele a los electrones que se encuentran en la superficie del filamento reduciéndose su emisión. Bajo estas condiciones se dice que la corriente del tubo de rayos X está limitada por la carga espacial; este efecto puede reducirse acercando ambos electrodos e instalando circuitos de compensación que minimicen la presencia de la carga espacial.

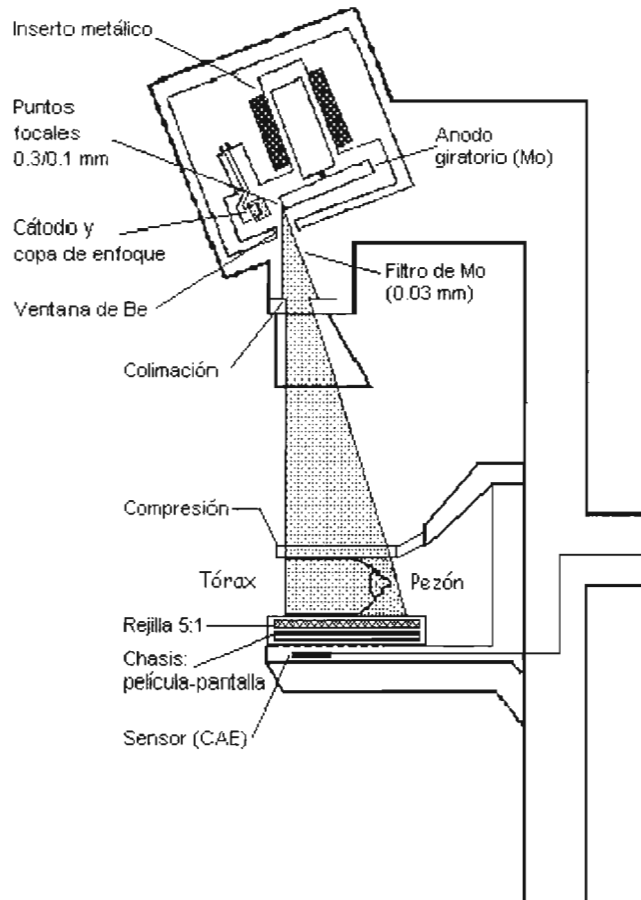


Figura 2.1. Sistema dedicado a mamografía. Se aprecia en detalle el tubo de rayos X en la parte superior del sistema, y demás componentes y accesorios. Se presenta una mama comprimida.

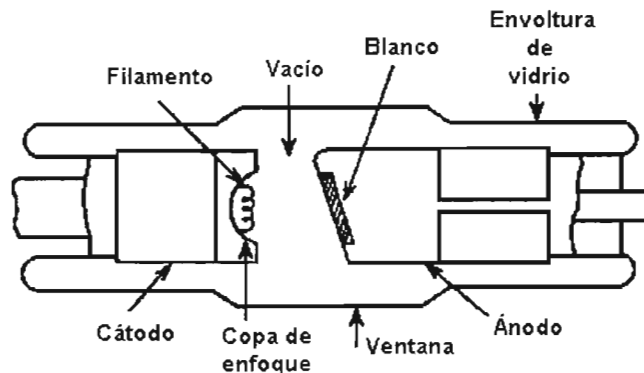


Figura 2.2. Diagrama de un tubo de rayos X.

2.1.B. Producción de rayos X para mamografía

Los tejidos glandular y adiposo, así como los tumores cancerosos tienen coeficientes de atenuación muy similares, que presentan su mayor diferencia en la película radiográfica cuando son irradiados con rayos X de bajas energías, entre 15 y 22 keV. Para producir estos últimos se utilizan tubos de rayos X con blancos de molibdeno (Mo) o rodio (Rh), los que al ser irradiados con electrones generan rayos X característicos con energías adecuadas. El

espectro de emisión de un tubo de rayos X con blanco de molibdeno se muestra en la figura 2.3, donde puede verse que existe también una componente de radiación de frenado.

En la mayoría de los equipos de mamografía se utiliza un blanco de molibdeno con filtro de molibdeno o rodio (estos haces los denominaremos Mo/Mo y Mo/Rh, respectivamente). El filtro se utiliza para atenuar la parte de los rayos X de frenado de baja energía, ya que éstos no contribuyen satisfactoriamente a la formación de la imagen y sí a la dosis impartida al tejido. Además, los filtros mencionados anteriormente atenúan también parte de los rayos X de frenado con energías superiores a las de la radiación característica; lo que implica la disminución de la cantidad de rayos X para los cuales no existe una diferenciación significativa en los coeficientes de atenuación de los diferentes tejidos. En la figura 2.4 se presenta el espectro del molibdeno el cual muestra rayos X característicos cercanos a lo deseado y cómo se modifica al utilizar un filtro de molibdeno.

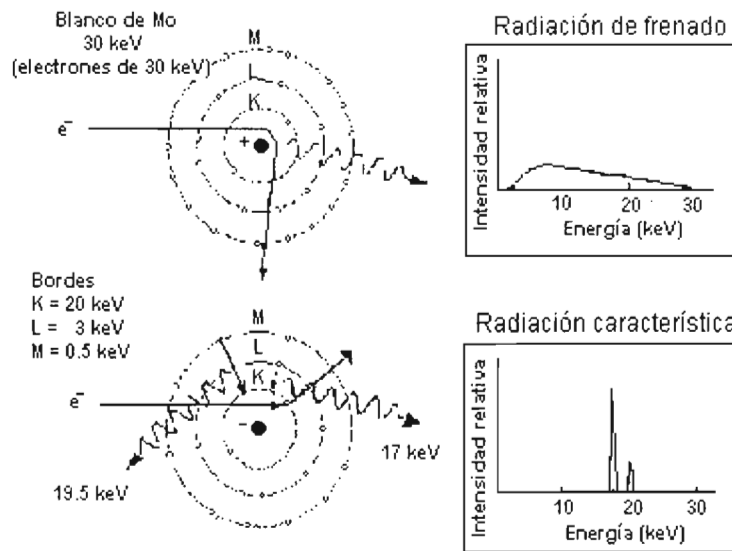


Figura 2.3. Rayos X producidos por un equipo de mamografía, lo que incluye rayos X de frenado y característicos. Los rayos X característicos del molibdeno tienen energías muy cercanas a las óptimas necesarias para la detección de lesiones de bajo contraste.

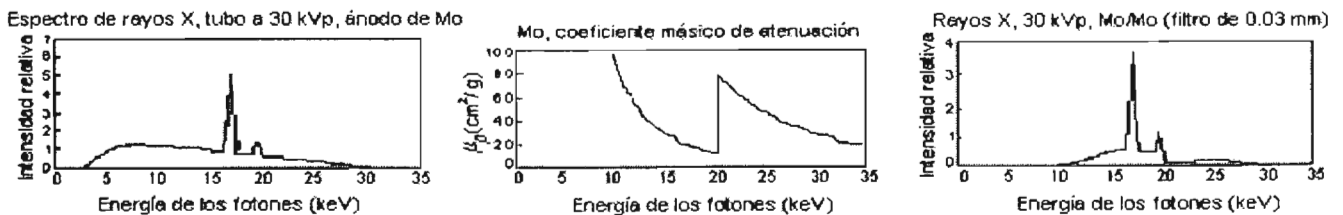


Figura 2.4. Espectro del Mo y Mo/Mo.

Una vez que se obtiene el haz de rayos X adecuado se debe utilizar otra serie de accesorios para obtener la imagen, tales como los que se pueden apreciar en la figura 1.1 y que se describen a continuación.

2.1.C. Colimación

Es necesario colimar el haz de rayos X producido para que éste incida sobre toda el área de la película mamográfica, y que a la vez coincida con el haz de luz. El haz de luz sirve para visualizar el área de irradiación y proporcionar ayuda a las técnicas radiólogas al momento de colocar correctamente la mama de la paciente dentro del área de irradiación. También, la colimación de los rayos X para que incidan sobre toda la placa ayuda al diagnóstico médico, proporcionando un oscurecimiento de la película en toda el área externa a la imagen de la mama, lo que mejora la visualización de la imagen al realizar su interpretación.

2.1.D. Compresión

La mama debe ser comprimida para “esparcir” el tejido mamario y ayudar a la diferenciación visual de la anatomía de la mama, ya que una radiografía es simplemente una proyección de un objeto tridimensional sobre un plano. En caso de no comprimir adecuadamente a la mama se puede ocultar algún tejido o lesión. Además, la compresión ayuda a prevenir el movimiento de la paciente lo que originaría una imagen borrosa. Así, la compresión cumple con las siguientes funciones: 1) mantiene la mama inmóvil y, en consecuencia, evita distorsiones de la imagen asociadas con el movimiento, 2) acerca la mama al plano de la imagen, aumentando la nitidez de la imagen, 3) separa los tejidos superpuestos que pueden enmascarar lesiones subyacentes y 4) reduce la dosis impartida a la mama (al reducir la radiación necesaria para obtener una mamografía de buena calidad), y la radiación dispersa (que no contribuye a la imagen y que se deposita en el cuerpo de la paciente).

2.1.E. Rejilla antidispersora

Se utiliza una rejilla antidispersora localizada justamente arriba del chasis que guarda y protege de la luz a la película mamográfica. La función de la rejilla antidispersora es impedir la incidencia sobre la película mamográfica de radiación dispersa. Esto reduce la cantidad de radiación que incide sobre la película. Lo anterior hace necesario aumentar la cantidad de rayos X proporcionada por el equipo para lograr obtener una imagen mamográfica aceptable, es decir, con el contraste y la densidad óptica adecuados para emitir el diagnóstico médico. Esto implica un aumento de hasta tres veces en la dosis impartida a la mama de la paciente, pero el aumento de dosis como consecuencia de la presencia de la rejilla antidispersora se ve justificado, ya que gracias a ella se obtienen imágenes mamográficas de mejor calidad al mejorar el contraste y la nitidez.

2.1.F. Control automático de exposición (CAE)

Existe un sistema electrónico en los equipos de película/pantalla, llamado Control Automático de Exposición (CAE), que es el que regula el tiempo que dura una exposición y la finaliza, de acuerdo, entre otras cosas, con los valores de espesor de mama comprimida y el control de densidad que las técnicas radiólogas programan en el panel de control. El objetivo del CAE es asegurar que la película recibe el mínimo adecuado de fotones para generar una imagen de buena calidad.

2.2. DESCRIPCIÓN GENERAL DEL EQUIPO DE MAMOGRAFÍA DIGITAL

Un equipo de mamografía digital es aquel equipo que produce una imagen digital mamográfica en forma de un archivo digital, i.e. obtiene directamente una imagen que se despliega en un monitor. La principal diferencia con un equipo analógico es la manera de adquisición de la imagen, ya que no se utiliza película radiográfica y el control automático de exposición (CAE) se cambia por la optimización automática de parámetros (OAP). Tiempo atrás se desarrolló un equipo digital de mamografía (“full field digital mammography”), resultando el equipo Senographe 2000D (GE), el primero en utilizarse clínicamente en EUA, y también el primero en emplearse en México. El equipo que se utiliza en este trabajo de tesis es el Senographe 2000D, del Instituto Nacional de Cancerología. A continuación se describe de manera general el funcionamiento del equipo de mamografía digital Senographe 2000D (GE) y una imagen de éste se muestra en la figura. 2.5.

2.2.A. Generación de rayos X

La generación de rayos X en el equipo de mamografía digital Senographe 2000D (GE) no difiere significativamente de uno analógico y para su descripción puede verse la sección 2.1.A. En general, el Senographe 2000D cuenta con un blanco de molibdeno (Mo) y otro de rodio (Rh). Los filtros que utiliza son de los mismos materiales (Mo con 30 μm de espesor y Ro con 25 μm de espesor), puede emplear cualquier combinación de blanco y filtro durante un estudio para obtener la mejor calidad de imagen mamográfica, por ejemplo Mo/Mo o Mo/Rh (blanco/filtro). La selección del blanco y filtro se puede realizar automáticamente, el equipo emite un haz de rayos X de corta duración y de acuerdo a su atenuación en la mama, registrada por el detector de imagen, éste selecciona la combinación de blanco y filtro más adecuados para la mamografía. Se emplea típicamente molibdeno para una mama pequeña (<4cm de compresión) y rodio para una mama mayor ó más densa (radiográficamente). El intervalo de voltaje en que puede emplearse el equipo va de 22 kV a 35 kV y de 25 kV a 49 kV para las combinaciones de blanco y filtro Mo/Mo y Rh/Rh, respectivamente. El de corriente va de 20 mA a 130 mA.

2.2.B. Optimización Automática de Parámetros (OAP)

El modo de OAP activa un sistema de detección electrónica que regula la cantidad de radiación y permite al equipo la selección de los principales parámetros de exposición, como la combinación blanco/filtro, la tensión, la carga y el tiempo de exposición. Antes de cada exposición realizada bajo el modo de OAP (μs antes del disparo principal), se realiza un breve disparo para evaluar la absorción de radiación en la mama. Esto sirve para determinar el espesor y la densidad de la mama examinada. Esta información junto con la configuración elegida por el técnico radiólogo permite la optimización de los parámetros de exposición. Una vez realizado lo anterior, se efectúa el disparo usando la combinación óptima de blanco/filtro y de tensión/carga, las cuales proporcionan la menor dosis posible con respecto a las características seleccionadas por el operador para la imagen.

2.2.C. Receptor de imagen

El receptor de imagen contiene el detector digital, el cual consiste de una capa plana de silicio amorfo (transistor de capa delgada acoplado a un fotodiodo discreto sobre un sustrato de vidrio), sobre la cual se deposita un centelleador de ioduro de cesio dopado con talio, CsI(Tl) (el centelleador se desarrolla en una especie de capa de fibras columnares con un diámetro de cada fibra de aproximadamente 3 a 6 μm y largo de 3 mm), para transformar los rayos X en luz visible (ver figura 2.6). El receptor provee una detección para un campo del tamaño de $19 \times 23 \text{ cm}^2$, con un tamaño de matriz de pixeles de 1914 renglones y 2294 columnas (cada pixel puede almacenar hasta 16 bits de información). La superficie superior del receptor puede removerse para un examen y puede utilizarse un soporte con o sin rejilla antidispersora (ver sección 2.1.D) [Ve99, Ky03, Ta03].

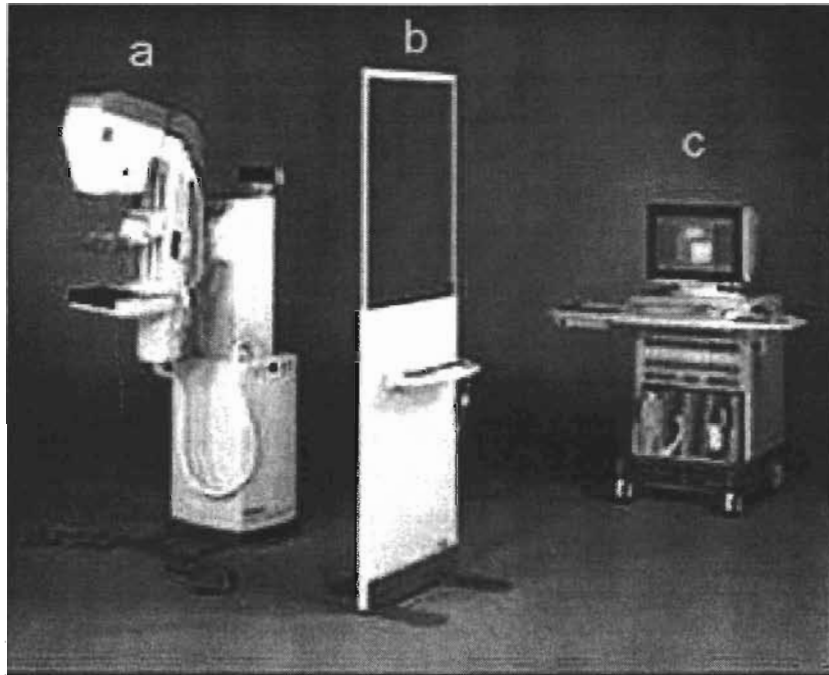


Figura 2.5 Equipo Senographe 2000D (GE). (a) Equipo de mamografía, (b) Panel de control y mampara de protección para el técnico radiólogo y (c) Estación de adquisición o por sus siglas en inglés, AWS (acquisition workstation).

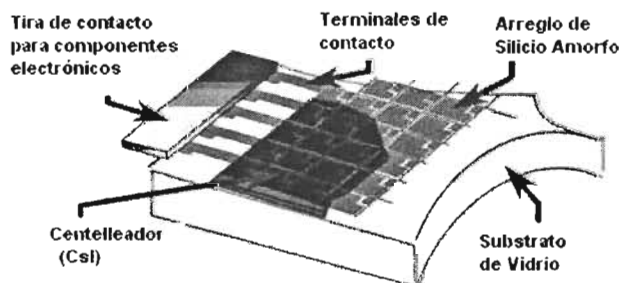


Figura 2.6. Esquema del detector ($19 \times 23 \text{ cm}^2$)

2.2.D. Visualización de las imágenes

Una vez que la imagen ha sido adquirida y procesada ésta se visualiza en la estación de revisión, la cual cuenta con monitores adecuados (figura 2.7) para su correcta evaluación clínica por el médico radiólogo.

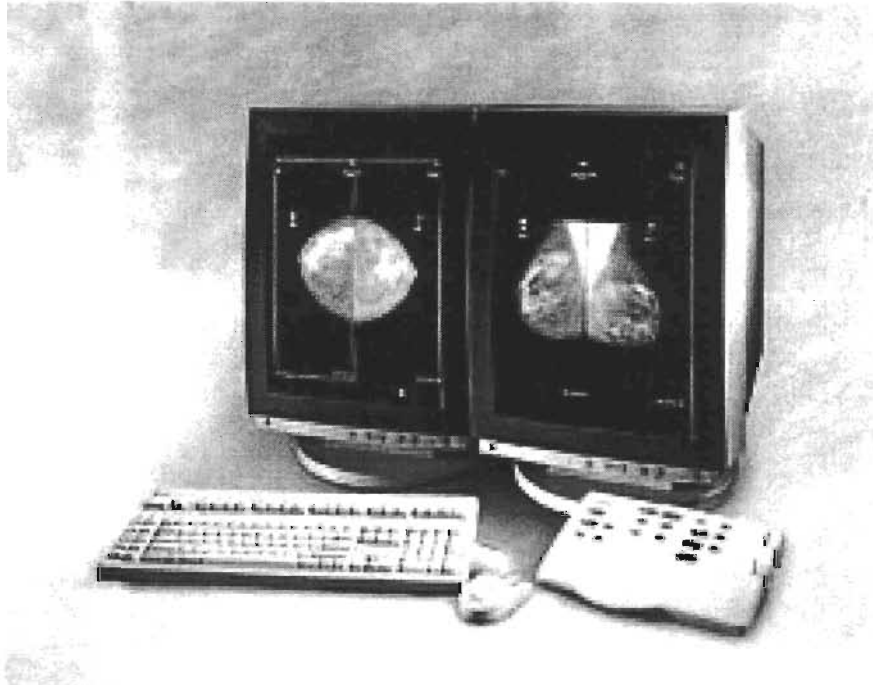


Figura 2.7. Monitores de la estación de revisión. Se muestra el despliegue de un estudio mamográfico.

2.3. PROCESO DE ADQUISICIÓN DE LAS IMÁGENES DIGITALES (SENOGRAPHE 2000D, GE)

En esta sección se describe el proceso de adquisición de la imagen radiográfica y cómo ésta es registrada por el equipo.

2.3.A. Detección radiográfica

La imagen radiográfica se obtiene al detectar los rayos X (primarios y secundarios), que atraviesan la mama de la paciente (generalmente hay un convertidor de rayos X a luz visible y lo que se detecta finalmente es luz). En la mama, el haz se atenúa y sufre principalmente las siguientes interacciones: efecto Compton y fotoeléctrico [At86]. La información que se utiliza para la formación de la imagen mamaria es la atenuación de los rayos X, razón por la cual el efecto Compton se trata de reducir para tener la menor dispersión de radiación posible. Esto se logra al tener un espectro de radiación con energía máxima lo más pequeña posible, de tal forma de reducir la probabilidad del efecto Compton y aumentar la del efecto fotoeléctrico. Las bases de la formación de la imagen se deben a la diferencia en los coeficientes de atenuación de las estructuras de la mama, lo que produce la diferencia en intensidades en la imagen (contraste radiográfico). Típicamente,

las zonas más blancas en la imagen obtenida (película/pantalla), se deben a una mayor atenuación, mientras que las zonas más oscuras se deben a una mayor intensidad del haz de radiación. En la radiografía digital sucede justamente lo contrario al nivel de la detección de los rayos X; sin embargo en la imagen desplegada se invierten los valores para que éstos sean iguales a los observados en un equipo de película/pantalla.

Cuando el haz de rayos X ha atravesado la mama de la paciente, éste interacciona con la capa de ioduro de cesio emitiendo fotones luminosos y, luego, éstos inciden en la capa de silicio amorfo para producir una carga eléctrica. En la tabla 2.1 se resumen las características del detector. La señal eléctrica producida se utiliza para formar la imagen digital. El detector, básicamente detecta el lugar donde se produce la interacción en el detector y la intensidad de ésta, para indicar un valor en la imagen.

En la figura 2.8 se muestra la sección transversal de un fotodiodo de silicio típico. Silicio como material tipo-N es el material inicial en el fotodiodo. Una capa delgada de material “p” se forma en la superficie frontal por difusión térmica o implantación iónica del material de dopación adecuado. La interfaz entre la capa del material “p” y del silicio “n” se le dice zona pn. Contactos pequeños se aplican a la superficie frontal y a la superficie total opuesta se le coloca un contacto metálico. Este contacto es el cátodo, mientras que el contacto frontal es el ánodo. Cerca de la zona pn el silicio se vacía de cargas eléctricas (depleción). Esta zona se le conoce como zona de depleción. La profundidad de la región de depleción puede variarse al aplicar un voltaje a través de la zona pn. Cuando la región de depleción llega a la parte trasera del fotodiodo se dice que se halla “totalmente vaciado”. La mayoría de la sensibilidad a la radiación luminosa se origina en esta región. El área activa se recubre con nitrato de silicio, monóxido de silicio o dióxido de silicio para protección y para funcionar como un recubrimiento antirreflejante. La superficie de este recubrimiento se optimiza para una longitud de onda particular. Cuando la luz se absorbe en el área activa un par electrón-agujero se forma. Los electrones y los agujeros se separan: los electrones emigrando a la región “n” y los agujeros a la región “p”. El resultado es una corriente generada por la luz incidente. La migración de los electrones y los agujeros a su región respectiva se le llama “efecto fotovoltaico”.

Tabla 2.1. Características del detector del Senographe 2000D, GE.

Tamaño del pixel	Tamaño del detector	Cristal que utiliza
100x100 μm^2	19x23 cm^2	CsI(Tl)

Los pixeles que integran el detector son el equivalente del tradicional grano de una emulsión fotográfica. La matriz de células (receptores de luz), produce una carga eléctrica proporcional a la cantidad de luz que cae sobre la superficie de los receptores. La información analógica que registra es transformada en un registro binario (digital), que se aloja en una memoria digital.

En vez de tomar una imagen analógica sobre una película radiográfica en donde las partículas de plata metálica o colorante forman una imagen final, un fotodiodo captura una imagen digitalmente. Esta adquisición digital de la imagen con base en pixeles (elementos discretos uniformemente espaciados), tiene la característica de poder controlar y procesar los valores que se obtienen en la imagen (representación numérica), por medio de un

programa de cómputo. Los millones de bits de información binaria que se generan en el momento de que la luz cae sobre la superficie del detector digital, organizan la imagen pixel por pixel; en el Senographe 2000D el tamaño del pixel es de aproximadamente $100 \times 100 \mu\text{m}^2$, i.e. una máxima resolución espacial de 5 pares de líneas por milímetro (5 pl/mm).

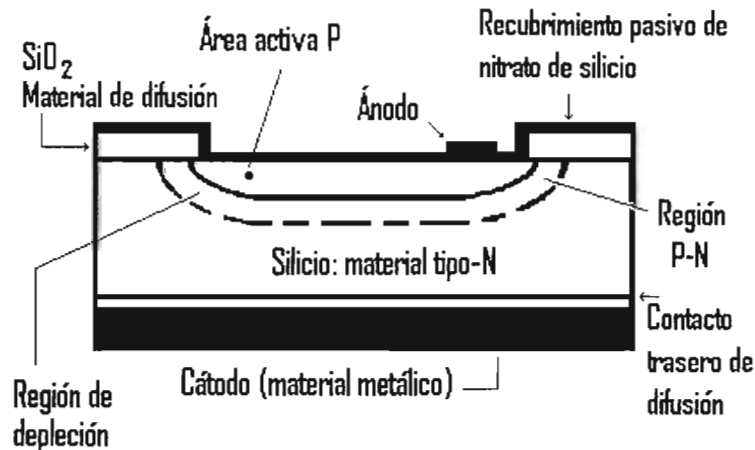


Figura 2.8. Sección transversal de un fotodiodo típico de silicio. [PHO]

Los valores de todos los millones de píxeles en la imagen pueden manipularse, es decir, se puede cambiar su brillantez, su contraste y sus valores cromáticos (características de color), entre otras características. Esta manipulación trae consigo una ventaja importante en cuanto a la calidad de la imagen (la cual se establece por la densidad de datos provenientes del registro original). Se debe tomar en cuenta que, generalmente, entre mayor número de píxeles se encuentren en el detector más datos por área existirán; por lo que se obtendrá una mejor resolución espacial. (una imagen de alta calidad, por lo tanto, es una imagen con una resolución espacial grande, véase sección 3.1.D).

2.3.B. Adquisición digital de las imágenes y procesamiento en el equipo

Después de la exposición, el archivo de la imagen adquirida por el detector digital se transfiere automáticamente a la estación de adquisición, AWS. Primero, la imagen obtenida es una imagen digital cruda, "RAW" (detección simple por el detector sin "ningún tipo de procesamiento"), de bajo contraste, permitiendo al ojo humano sólo la clara observación del borde de la piel y el pezón. Segundo, inmediatamente después de la adquisición de la imagen RAW, se aplican un número de cálculos para desplegar una imagen procesada, "PROCESSED". Al visualizar las imágenes en el monitor de las estaciones de adquisición y revisión, se despliega en el monitor información general, como la técnica radiológica utilizada para la exposición (kVp, mAs, dosis glandular promedio), la proyección mamográfica y los datos de la paciente.

La imagen cruda se guarda en formato DICOM (16 bits) [WIK] para su posterior procesamiento digital, por lo que es este tipo de imagen la que se analiza en nuestra investigación. La imagen procesada se guarda en formato DICOM (12 bits), para su

presentación y lista para visualizarse en un estudio clínico (a esta imagen se le puede variar el brillo y contraste para mejorar su presentación).

Principales cálculos que se aplican a la imagen RAW para transformarla en “procesada” [GE00]:

- Detección de colimador: éste aplica una máscara negra alrededor del área útil de la imagen, cubriendo las áreas que de otra manera serían blancas. Esto permite un modo más cómodo de visualización, similar al de una imagen analógica.
- Transformación pseudo-log (función de transformación de los valores originales de la imagen para mejorar el contraste final): ésta facilita la manipulación del brillo y del contraste durante su visualización. Después de la transformación el intervalo dinámico se reduce a 12 bits, sin pérdida de información clínica. La forma exacta en que trabaja esta transformación no se brinda en el manual del equipo por la compañía General Electric. Sin embargo, el Dr. A. Burgess publicó en la revista Medical Physics en julio del 2004 [Bu04] un artículo donde propone una versión simplificada para pasar de la imagen RAW a la procesada al aplicar una transformación de pseudo-Beers para niveles de gris superiores a 300 y una transformación lineal para los inferiores.
- Ecuilibración de espesor: los tejidos subcutáneos son difíciles de visualizar en la pantalla debido a las limitaciones del monitor. Esto se corrige al aplicar un algoritmo de ecualización de espesor para disminuir el intervalo dinámico. La información de la imagen en tonos de gris seleccionados por encima de un nivel umbral de gris, se modifican para corresponder a tejidos subcutáneos y así permitir una mejor visualización. La imagen resultante muestra una mejor visibilidad de la información clínica en todas las regiones de la mama. Cuando este algoritmo se aplica al procesamiento de la imagen se le refiere en el Senographe 2000D como Proc 1 y cuando no se utiliza el algoritmo se le refiere como Proc 0.
- Contraste-automático: éste mejora la calidad de la imagen al optimizar los niveles de brillo y contraste en la imagen.
- Tan pronto como este algoritmo se aplica a la imagen cruda, ésta es remplazada por la imagen procesada, la cual se despliega con el contraste estándar en la pantalla.

CAPÍTULO III ANÁLISIS DE IMÁGENES

El capítulo presenta la teoría del análisis y procesamiento de imágenes digitales, enfocándose en su aplicación a mamografía. Las referencias principales del contenido en cuanto al análisis de las imágenes son: [Ja02], [Ke93], [Mu01], [Qu01].

3.1. PROCESAMIENTO Y ANÁLISIS DE IMÁGENES – TEORÍA BÁSICA

La imagenología médica proporciona importantes herramientas de diagnóstico. Existen una gran variedad de técnicas para la formación (fenómeno físico) y producción (técnicas de adquisición y pre-procesamiento) de imágenes; cada una está basada en una propiedad física o fisiológica del cuerpo. Por ejemplo, las imágenes por rayos X, en dos dimensiones (2D), son imágenes de la distribución de los coeficientes de atenuación lineal a una energía promedio de los rayos X en el cuerpo humano. Esta es una propiedad dependiente de la densidad y composición del tejido. Así, un hueso, por ejemplo, tiene un coeficiente de atenuación mayor, lo cual puede explicarse debido a que éste contiene cantidades significativas de elementos con un número atómico mayor (comparándolo con los tejidos blandos).

Para la obtención de una imagen digital, en nuestro caso de una mama femenina, se requiere proyectar un objeto tridimensional (3D) con diversas características, a una imagen plana 2D. La imagen mamográfica (en (x,y)), se forma a partir de las diferencias en los coeficientes de atenuación de los diferentes tejidos mamarios, en el camino de los rayos X en (z) . El procesamiento de la imagen mamográfica es el conjunto de las transformaciones (T) aplicadas a la imagen original (I_{in}) para la obtención de la imagen final (I_{out}), como se describe en la ecuación 3.1.

$$I_{out}(x, y) = T(I_{in}(x,y)) \quad (3.1)$$

En las imágenes de salida (I_{out}), es posible medir o detectar “a ojo” o por programas de cómputo, características de interés. Sin embargo el resultado final de éstas, es un conteo o cuantificación, ya sea manual o asistido.

3.1.A. Imagen digital

Una imagen digital está caracterizada por diferentes valores o condiciones, algunos de éstos los citamos a continuación:

- Primitiva de representación: punto, pixel.
- Localización espacial: enteros (i,j) corresponden a la posición (x,y) (retícula ortogonal discreta de $N \times M$). Esta estructura permite tratar la imagen como arreglo matricial.
- Atributo: escalar como la intensidad, o vectorial como el color en espacio R,G,B. Los valores del atributo también son discretos, representados usualmente en 8 ó 16 bits.

3.1.B. Resolución

La resolución en una imagen cuantifica la capacidad de apreciar dos objetos pequeños y cercanos como dos entidades ajenas entre sí. Mientras mayor sea la resolución, más cerca se pueden colocar dichos objetos sin que sus imágenes se traslapen (ver figura. 3.1). El tamaño de la mancha es la función de dispersión puntual del sistema. La resolución del sistema de imagen está influenciada por el desempeño del equipo de rayos X y el comportamiento del receptor de imagen.

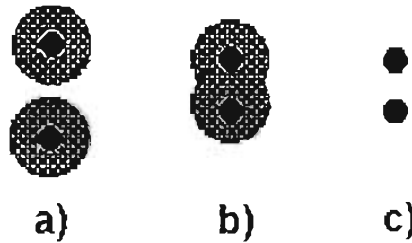


Figura 3.1. a) Imágenes de puntos pobremente definidas se resuelven si los puntos están bastante alejados (la parte “sombreada” es la dispersión del objeto negro causada por el sistema de adquisición). b) Con puntos menos separados las imágenes no se resuelven. c) Con la misma separación que en b), las imágenes de los mismos puntos se resuelven si su definición es mayor.

Debido al proceso de detección de la imagen, un objeto “puntual” será degradado, por lo que el objeto se verá esparcido o disperso en una región alrededor del punto central, en la imagen resultante. La respuesta de un sistema de imagen a un punto se llama “función de dispersión puntual” (ó PSF, por sus siglas en inglés: “Point Spread Function”) y describe la dispersión lateral de información del sistema. La dispersión de la luz ocurre en los ejes x e y y entonces la PSF se expresa usualmente como una función espacial, i.e. $PSF(x,y)$. Una forma de cuantificar la resolución del sistema de producción de la imagen es a través de medir en la PSF el ancho total a media altura de su máximo o FWHM, por sus siglas en inglés (full width at half maximum).

3.1.C. Nitidez

La nitidez es la capacidad de un sistema de despliegue o producción de imágenes, para que un borde aparezca más o menos delineado. La falta de nitidez de un sistema de imagen proviene principalmente de cuestiones geométricas, radiográficas y por su PSF. En un equipo digital, la imagen es procesada para obtener mejoras en su visualización.

3.1.D. Contraste

Contraste significa diferencia relativa entre valores extremos. En una imagen digital el contraste se presenta en la forma de diferencias en los tonos de gris o en la brillantez (ver figura 3.2). El contraste es la característica más importante de una imagen y en la ecuación 3.2 se muestra la medida normalizada del contraste, cuyo máximo es 1. Sin embargo, si una imagen está demasiado contrastada generalmente no es útil y en la mayoría de los casos es indeseable, por ejemplo una imagen binaria divide en 2 clases toda la información de los valores de forma arbitraria.

$$C_{normalizado} = \frac{I_{Max} - I_{min}}{I_{Max} + I_{min}} \quad (3.2)$$

El contraste físico de un objeto puede representar una o varias diferencias en las características de los tejidos. En radiografía los objetos pueden distinguirse de sus alrededores si cuentan con la diferencia adecuada ya sea en densidad, número atómico (composición), o espesor. La relación entre el grado de contraste requerido para la visualización y para el brillo (ver sección más abajo) del entorno está influenciado por el tamaño del objeto. Los objetos pequeños requieren un alto nivel de contraste.

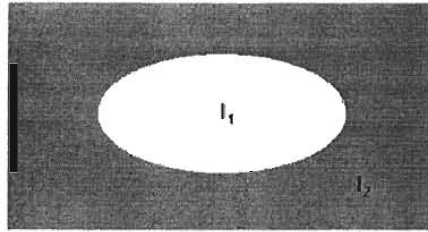


Figura 3.2. Contraste. I_1 e I_2 son los niveles de gris de las regiones 1 y 2, respectivamente.

3.1.E. Brillo

El brillo es un atributo de la percepción visual en el cual una fuente parece emitir una determinada cantidad de luz. Puede verse como el valor inicial del nivel de gris más bajo de una imagen. El concepto de “brillo” fue utilizado en fotometría como un sinónimo para el término “luminiscencia”. El “brillo” debería usarse solamente para referencias no cuantitativas de la percepción puramente visual. En espacios de color el brillo puede pensarse como la media aritmética de cada una de las componentes de color, o sea las coordenadas en el espacio de color empleado (v.g. coordenadas R, G, B).

3.1.F. Formatos de archivos de imagen

Existen una variedad de formatos de archivos para guardar una imagen. TIFF (Tagged Image File Format), GIF (Graphics Interchange Format), JPEG (Joint Photographic Expert Group) o PGM (Portable Graymap Format), son algunos ejemplos de formatos de imagen, sin embargo, éstos proveen mecanismos de compresión de la imagen, lo cual puede resultar en diferencias en el momento de realizar su análisis. Esto hace necesario un formato de imagen que registre adecuadamente la información de la imagen y permita su distribución entre usuarios. En el presente trabajo se analizan las imágenes mamográficas en formato DICOM (para más detalle del formato DICOM ver anexo B.1.B).

Una imagen puede clasificarse como:

- 1-bit, **binaria**, “duotono”, o “Blanco y negro”: imágenes de este tipo contienen pixeles que se codifican como blanco o negro. Éstas son imágenes binarias. La calidad visual es pobre y las estructuras internas de la mama no se distinguen del fondo mamario. Se usan sin embargo como “mascarillas” para aislar zonas de interés y confinar un procesamiento o análisis en otra imagen.

- 8-bits, **escala de grises**: éste es el formato típico que se utiliza en diversos estudios. Cada pixel en la imagen puede ser uno de los 256 niveles de gris, donde 1 representa el tono negro y el $2^8=256$ el blanco. Este formato proporciona una imagen clara.
- 12-bits, **escala de grises**: muchos estudios enfocados en mamografía digital usan la codificación de 12-bits en escala de grises, donde 1 representa el tono negro y el $2^{12}=4096$ el blanco. Mientras mayor sea el intervalo en la escala de grises, mayor será la información que se pueda registrar. Esta alta resolución permite transformar los datos usando funciones no-lineales, por ejemplo en escala logarítmica.

Una vez que se ha obtenido la imagen digital, I , ésta representa la proyección de la mama en el plano de detección y los valores en tonos de gris de la imagen para cada punto (x, y) .

Las etapas típicas posteriores a la obtención de la imagen digital son:

- **Realce**: se obtiene una mejora en su visualización, por ejemplo en contraste y tamaño.
- **Restauración**: la imagen se filtra para eliminar el ruido y completar información en la imagen.
- **Segmentación**: se seleccionan y separan las áreas específicas para su análisis, a fin de no analizar todas las estructuras de la imagen. Así se evita analizar regiones con baja probabilidad de presentar alguna anomalía y optimizar el estudio de la imagen.
- **Análisis**: se cuantifican factores característicos de la imagen o de los objetos de interés, al estudiar su número y distribución, su morfología, textura y relaciones espaciales.
- **Interpretación y entendimiento**: se evalúa la imagen o los objetos de interés para clasificarlos en una categoría determinada de acuerdo a los análisis efectuados.

A continuación se describen las etapas que se utilizarán en el procesamiento y análisis de las imágenes mamográficas digitales. Cabe mencionar que aunque otros trabajos citados anteriormente en el área de procesamiento digital realizan un realce de la imagen en primer lugar, en este trabajo no se hará. La razón es que al efectuar el realce de la imagen éste afecta los valores de los píxeles que forman la imagen para resaltar alguna característica particular o región en la imagen, pudiendo así perder información sobre la configuración original de los valores, representando otra característica o región de la imagen. En este trabajo de tesis es importante conservar los valores originales pues una de las metas es caracterizar una imagen mamográfica a través de su textura y morfología original.

Sólo por completez se hace una breve descripción de lo que significa el realce y la restauración de imágenes.

3.2. REALCE DE IMÁGENES

El realce de imágenes sirve para obtener una mejor visualización de la imagen (ver figura 3.3). Esto puede hacerse, al mejorar el cociente de señal a ruido: $SNR = h/\sigma^2$, con h la diferencia en el nivel de gris entre el objeto y el fondo, y σ la desviación estándar del ruido. El cociente señal a ruido SNR es un término utilizado para el cociente entre la magnitud de una señal (información con significado de la señal) y la magnitud del ruido del fondo. SNR se expresa usualmente en términos de una escala logarítmica: $SNR = \log(\text{señal/ruido})$

[Ki81]. La calidad en la imagen puede deteriorarse como resultado del proceso de producción de la mamografía, por ejemplo, debido al movimiento de la paciente cuando se le posiciona durante la compresión. Uno de los principales objetivos del realce de la imagen es mejorar el contraste entre las regiones de interés y el fondo. La calidad del realce repercute en la calidad de la segmentación. El realce de la imagen puede hacerse con una variedad de métodos en el dominio espacial o espectral. Métodos en el dominio espacial se basan entre otros, en la manipulación del histograma (redistribución de los niveles de gris en la imagen, i.e., en la ecualización (a veces conocida como “igualación”) del histograma y sus variantes), [un ejemplo de histograma se muestra en la figura 4.3. el cual no es más que la gráfica de la frecuencia relativa de los niveles de gris más probables en la imagen], o en el uso de máscaras o ventanas para aislar zonas de interés y confinar un procesamiento aplicado a la imagen original para mejorar su calidad. Los métodos espectrales usan apropiadamente las funciones de transferencia en el dominio espectral, las cuales realizan operaciones de convolución mediante la transformada de Fourier de la imagen, y luego se obtiene su inverso, para lograr el realce de la imagen. Algunos de estos métodos incluyen filtros pasa-bajos, pasa-banda y pasa-altos.

3.3. RESTAURACIÓN DE IMÁGENES

El ruido que se encuentra en una imagen es información no deseada, que puede ser resultado del proceso de adquisición de la imagen. Tradicionalmente las aplicaciones de procesamiento de imágenes utilizan varias técnicas para suprimir el ruido al aplicar una variedad de técnicas de filtrado espaciales o espectrales. Por ejemplo, filtros espaciales trabajan directamente en los píxeles para atenuar o remover el ruido al convolucionar con máscaras espaciales resultando en una imagen más suavizada. En general, filtros en el dominio de frecuencias operan en la representación de la transformada de Fourier de una imagen, como en el caso de [Mu01] el cual emplea un método de suavizado “gausiano” junto con operaciones de submuestreo como fases de pre-procesamiento en la película mamográfica digitalizada. En el presente trabajo de tesis, no se realiza ningún proceso de restauración y se analizan las imágenes digitales RAW tal como las entrega el equipo Senographe 2000D.

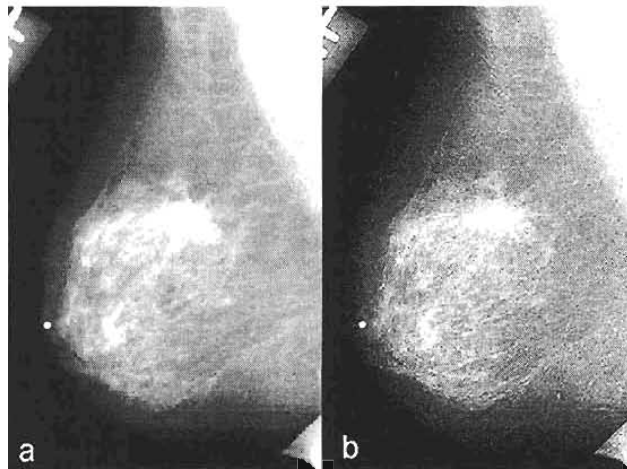


Figura 3.3. (a) Mamografía original y (b) mamografía después del realce.

3.4. SEGMENTACIÓN DE IMÁGENES

La segmentación de imágenes es una etapa fundamental hacia la identificación y análisis de regiones de interés en un objeto. Es importante diferenciar entre lo que se está deseando analizar y los otros detalles que no son de interés (ver figura 3.4). Como ejemplo se puede pensar en un punto o línea negra sobre un fondo blanco. Una de las suposiciones en las aplicaciones de segmentación es que el objeto o región de interés tiene características diferentes del fondo que pueden extraerse usando criterios de diferencias en intensidad, textura, procesos estadísticos, entre otros. Estas diferencias permiten separar o segmentar las regiones de interés.

Después del realce de la imagen, la siguiente etapa es encontrar y aislar regiones de interés que necesitan más investigación para determinar si éstas presentan alguna forma de anomalía. Algunos métodos comunes para aislar masas en las mamografías consisten en la substracción bilateral y en la descomposición y análisis de una sola imagen. Las técnicas de substracción bilateral alinean la mama izquierda y la derecha para una misma proyección (por ejemplo, usando la información de la posición del pezón), y encuentran las diferencias de simetría entre las dos mamas al sustraer una imagen a la otra. La descomposición de una sola imagen asume que las regiones uniformes en una imagen requieren investigación más detallada. La mayoría de las masas y regiones de alta densidad mamográfica usualmente muestran regiones de homogeneidad en niveles de gris, i.e. son tan densas en el centro como en la periferia y se observan en las dos proyecciones estándar.

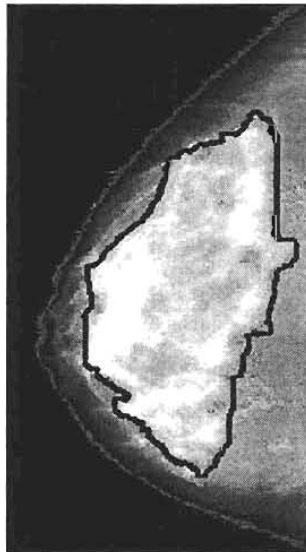


Figura 3.4. Se muestra un ejemplo de la segmentación (línea negra) de un área de alta densidad mamográfica en una mamografía.

Los métodos de segmentación por sí mismos no permiten juzgar las características de una región (normal o anormal). Otras técnicas de forma y textura deben aplicarse para encontrar lo que se requiere. Sin embargo, las regiones de interés deben detectarse primero para computar sus características. Desde una perspectiva computacional, un buen sistema de segmentación de imágenes debería mostrar un número pequeño de regiones con alta

probabilidad de ser los casos que se están buscando, en vez de un gran número de regiones que tengan una probabilidad menor de ser anormales.

La detección de regiones de densidad o masas tumorales de alta malignidad muestran problemas en su segmentación como, a) son de tamaño, forma y densidad variable, b) éstas muestran poco contraste, c) las regiones de interés pueden estar ligadas al tejido parenquimatoso circundante (particularmente mayor para lesiones espiculadas), d) las regiones de los tejidos del fondo rodeando las regiones de interés no son uniformes y ambas tienen características similares. La taxonomía de la segmentación puede representarse con cuatro clases o técnicas: técnicas de umbralaje, técnicas basadas en las fronteras, técnicas basadas en las regiones y métodos híbridos [Ad94]. Otra clasificación más simple puede ser: métodos globales, métodos basados en los bordes y métodos basados en las regiones. En los métodos globales, el objetivo es analizar las imágenes como un todo e investigar la distribución de probabilidad de los niveles de gris de los píxeles para encontrar un umbral apropiado para segmentar (separar) las zonas que se desean analizar. Los métodos basados en los bordes utilizan información de los cambios de gradientes en una imagen para localizar los bordes o fronteras que separan a los objetos del fondo. Los métodos basados en las regiones se enfocan en diferenciar las regiones en base a alguna característica particular.

3.4.A. Análisis global

Los métodos globales de segmentación involucran el uso de la información de las distribuciones de los niveles de gris de una imagen para su segmentación. En el presente trabajo de tesis se utiliza el análisis global de cada imagen mamográfica digital en el caso de la extracción de sus características texturales (ver secciones 3.5.B y 4.3.A).

Métodos de umbralaje: se basan en encontrar un valor umbral óptimo T como nivel de gris. Todos los píxeles con un nivel de gris mayor que T se definen como parte del objeto y aquellos por abajo forman el fondo. Si los objetos en la imagen tienen fronteras definidas, entonces el nivel de gris de los píxeles en las fronteras pueden servir como nivel umbral. Si sólo existe un objeto superpuesto al fondo, y los niveles de gris son considerablemente diferentes que los del fondo, entonces el histograma resultante será bimodal. Otra alternativa es graficar el histograma de los píxeles que tengan el mayor gradiente en la imagen. El histograma será unimodal y el nivel de gris correspondiente al pico será el umbral T , o el promedio de N valores rodeando al pico pueden utilizarse también para estimar el valor de T . La figura 3.5 muestra un ejemplo del uso de múltiples umbrales para un histograma multimodal a fin de dividirlo en muchas regiones de interés [Ja02].

Umbralaje óptimo: puede usarse como otra alternativa al método por umbralaje citado anteriormente. En este caso se realiza el promedio de los diferentes picos en el histograma multimodal de la imagen aplicando diferentes pesos a cada pico (ver figura 3.6). El umbral así definido es el nivel de gris más cercano que corresponde a la mínima posibilidad entre los máximos de dos o más distribuciones normales [Qu01].

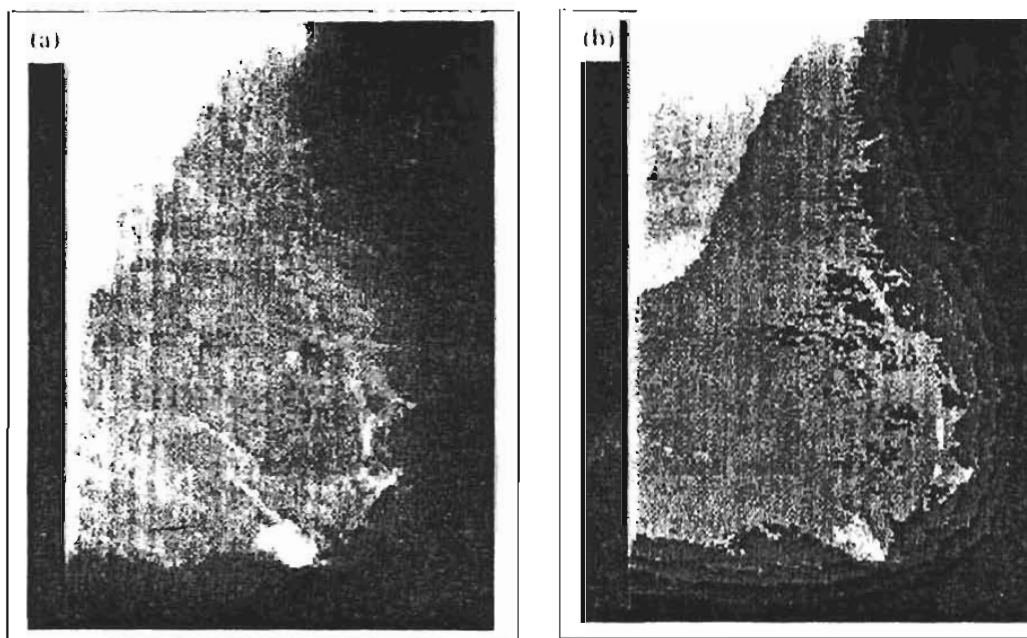


Figura 3.5. (a) Mamografía original y (b) mamografía por umbralaje escalonado a través de un histograma multimodal.

Umbralaje adaptable: la idea básica es dividir la imagen en un número de partes de igual tamaño, por ejemplo dividirla en ventanas de 4×4 o de 3×3 . Para cada ventana puede aplicarse un proceso de umbralaje basado en histogramas o umbralaje óptimo. Si para cada ventana v , se detecta un umbral T_v , entonces los píxeles con valor mayor a T_v son vistos como parte del objeto, mientras que el resto pasan a ser parte del fondo. Para resultados óptimos el tamaño de las ventanas usadas debe ser mayor que el objeto más grande que se desea detectar para que el umbralaje tenga sentido (ver figura 3.7) [Ja02].

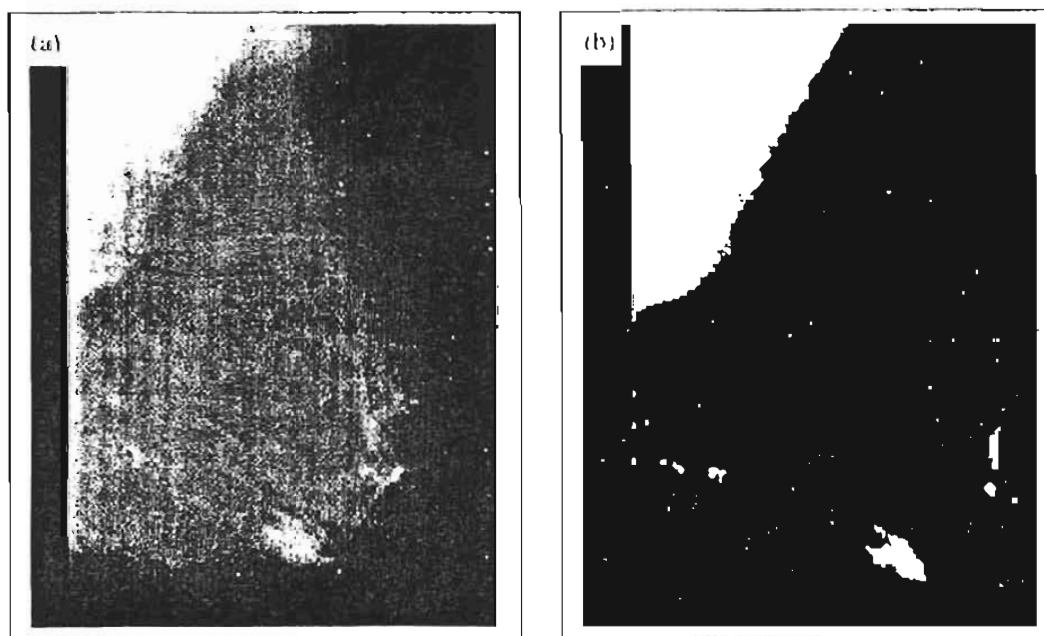


Figura 3.6. (a) Mamografía original y (b) mamografía por umbralaje óptimo.

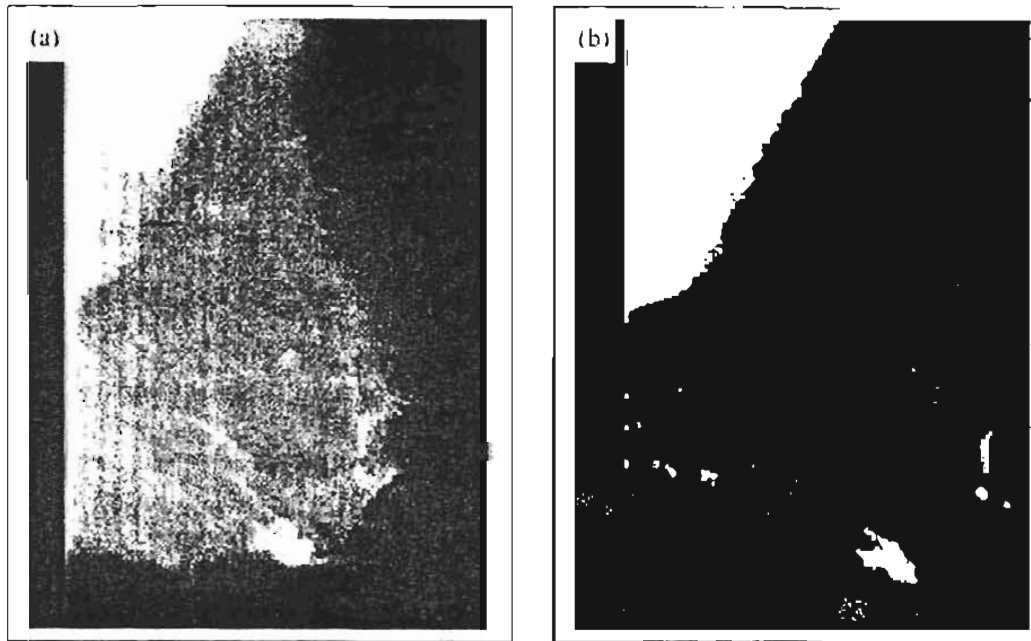


Figura 3.7. (a) Mamografía original y (b) mamografía por umbralaje adaptable.

3.4.B. Análisis mediante bordes

Los métodos basados en el análisis de bordes para la segmentación de imágenes son técnicas efectivas empleadas para separar las regiones de interés de su fondo en las mamografías. La suposición fundamental es que la separación entre dos regiones, objeto y fondo, es realizable usando la información del gradiente. El gradiente en una imagen representa la tasa de cambio con la cual los niveles de gris de los píxeles cambian en una dirección dada. La detección del gradiente se realiza a través de usar ciertas máscaras que se convolucionan con la imagen original para dar una imagen gradiente. La magnitud de la imagen gradiente en las direcciones vertical y horizontal puede combinarse para proporcionar un promedio de las magnitudes del gradiente y sus direcciones. La magnitud del gradiente representa la intensidad del borde (la forma abrupta o no en que se presenta la frontera entre dos regiones). Dado que todos los bordes con intensidades mayores que algún umbral pueden ser detectados, los análisis subsiguientes pueden enfocarse en la región de interés. Por ejemplo Kegelmeyer [Ke93], desarrolló un método para detectar lesiones espiculadas en las mamografías (digitalizadas), computando características texturales a partir de un mapa de las orientaciones de los bordes locales de cada lesión.

Uno de los métodos más sencillos usando la información de los bordes en mamografía es simplemente remover de la imagen original aquellos píxeles que tengan una magnitud de gradiente mayor que un umbral. Si se escoge un valor umbral pequeño, se puede eliminar un gran número de píxeles. Las bases de esta suposición es que las regiones densas son relativamente homogéneas en su nivel de gris, por lo que contienen un número pequeño de tipos de bordes, la región resultante será considerada como región de interés. Otro método puede basarse en el umbralaje de los bordes de la imagen para encontrar bordes intensos y bordes ligados para determinar regiones cerradas. Técnicas más sofisticadas basadas en la

manipulación de los bordes identifican primero las fronteras y entonces determinan las regiones que se encuentran en ellas [Ja02].

3.4.C. Análisis por regiones

Este es el método más popular de segmentación de imágenes en mamografía digital. La segmentación basada en las regiones involucra el análisis de las propiedades regionales en la imagen. Algunas técnicas populares para tal segmentación incluyen la fusión de regiones, la separación de regiones, una combinación de éstas, el método del momento y de agrupamientos, entre otros [Ja02]. Este tipo de análisis el utilizado en el presente trabajo para la extracción de las características morfológicas de cada imagen mamográfica digital analizada (ver secciones 3.5.A y 4.3.B).

Fusión de regiones: se basa en el concepto de agrupar los píxeles que satisfagan el criterio de fusión. Este criterio se basa en que la región formada después de la combinación de dos o más píxeles sea homogénea. Mientras que la región crece, la homogeneidad de las regiones resultantes debe mantenerse. Los criterios para verificar la homogeneidad incluyen el promedio de los niveles de gris, la varianza de los niveles de gris, propiedades de los histogramas y propiedades de la textura. Mientras que estas propiedades se mantengan estables antes y después de que nuevos píxeles sean añadidos a la región en crecimiento, el proceso puede continuar. La definición de una región comienza como un simple píxel y se convierte en una región de agregados de píxeles con el tiempo. El píxel inicial para el inicio del proceso es especificado por el usuario y dependiendo de esto, pueden realizarse segmentaciones diferentes. Por ejemplo Huo et al. [Hu95], segmentaron las regiones de las masas tumorales empleando métodos por crecimiento de regiones y propusieron 2 medidas de espicularidad obtenidas a partir de un análisis de la información radial del gradiente de borde alrededor de la periferia de las regiones analizadas.

Separación de regiones: sigue una forma diferente en la segmentación de la imagen a través de separar la imagen en pequeñas regiones homogéneas. Se aplica el mismo criterio que en el método anterior para vigilar la homogeneidad de las regiones. Muy seguido se utiliza una combinación de los métodos de *fusión* y *separación* los cuales se basan la representación como estructura de datos piramidales de la información de la imagen y un criterio de homogeneidad. Un ejemplo de esta técnica se muestra en la figura 3.8.

Método del momento: la segmentación se basa esencialmente en la interpretación como superficie topográfica de la imagen donde los niveles de gris del gradiente de la imagen representa la altitud, de tal forma que los bordes de tales regiones corresponden a un alto momento y las regiones interiores de bajo gradiente corresponden a “pozos” (“catching bassins” en inglés). Los “pozos” representan las áreas de la imagen con niveles de gris de bajo gradiente y su identificación se vuelve el principal objetivo de un algoritmo de segmentación por momentos [Be97]. En el presente trabajo se realiza un método muy parecido para el análisis morfológico, solo que no buscamos las regiones de bajo o alto gradiente, si no las curvas de isonivel (de gris) que sean características de una lesión mamográfica (ver sección 4.3.B).

Método por agrupamiento: una de las técnicas más conocidas para identificar regiones [Ja88]. Los métodos de agrupamientos pueden aplicarse directamente en los píxeles de la imagen para separar regiones que son homogéneas. El agrupamiento puede aplicarse usando aproximaciones estadísticas o de conectividad. Aproximaciones estadísticas incluyen técnicas de agrupamiento simples y difusas. Al usar la conectividad se incluyen redes neuronales.

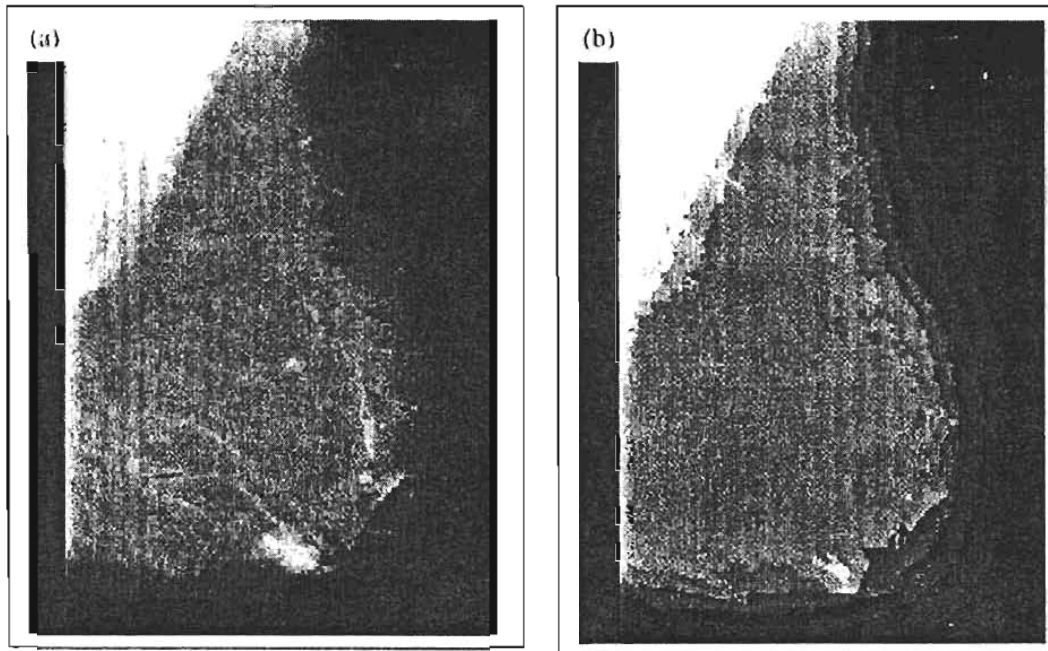


Figura 3.8. (a) Mamografía original y (b) mamografía por umbralaje de fusión y separación.

3.5. ANÁLISIS DE CARACTERÍSTICAS BÁSICAS

Después de pasar por el realce de la mamografía y la segmentación de las imágenes, la siguiente etapa es la extracción de características globales y regionales. A partir de las masas o regiones de densidad se obtienen vectores de características, los cuales se asumen como diferentes del tejido normal. Basándose en la colección de ejemplos, a partir de muchos casos, se puede entrenar un sistema para diferenciar cada una de las características correspondientes a distintas condiciones en las imágenes mamográficas analizadas. El objetivo principal es que estas características debieran ser sensibles y precisas para reducir los falsos-positivos. Típicamente, un conjunto de vectores de características se extrae para una cierta región segmentada. Las características pueden clasificarse a groso modo como los tipos que se listan a continuación:

- Características morfológicas relacionando los rasgos de forma y geometría, como los contornos de una región.
- Características texturales caracterizando las texturas subyacentes de la región.
- Características clínicas extraídas manualmente en base de observaciones clínicas.
- Características de la distribución en la escala de grises relacionando el histograma de escala de grises de la imagen.

- Otras características incluyendo mediciones de la asimetría, patrones lineales e intensidades de bordes.

A continuación se presentan dos características que han mostrado una relevancia importante en el análisis de las mamografías y que se utilizan en el presente trabajo de tesis.

3.5.A. Morfología

En el procesamiento de imágenes digitales, un estudio morfológico es una colección de técnicas basadas en la morfología matemática. como forma, tamaño, geometría, apariencia y configuración (la relación que guarden entre si las regiones analizadas) [WIK].

Muchos estudios han usado las características morfológicas que se relacionan con la forma y geometría de una región. Un método es extraer características del área (número de píxeles que comprende la región) y una característica del borde (la distancia en píxeles a partir de la región y el borde de la mama). Yin et. al. [Yi93], utilizan estas características para eliminar resultados falsos-positivos de la sustracción bilateral de pares de mamas (ver anexo A.3 y tabla A.2, para definición de falsos-positivos). En este trabajo de tesis las regiones obtenidas fueron objeto de un análisis adicional de pruebas de su área y perímetro. Algunos autores [Qi99] definen una medición de circularidad C , como se muestra en la ecuación 3.3, donde p es el número de píxeles delimitando el perímetro de la región y a el área de la región.

$$C = \frac{p^2}{4\pi a} \quad (3.3)$$

Otros autores [Ng92], comentan técnicas de diferenciación de tumores para discriminar entre lesiones espiculadas y microsegmentadas (ver figura A.4. en el anexo A.5.A). Una medición de la espicularidad se define como el número de espículas conectadas a una región. Estas espículas se detectan durante el proceso de segmentación. Dos condiciones deben tenerse para considerar a una espícula como conectada: a) no estar colocada en la masa central, y b) la dirección de la espina (eje de la espícula) debe cruzar la frontera de la masa. Usando estos dos criterios, los autores clasificaron correctamente 90% de las regiones en su conjunto de estudio como regiones espiculadas o microsegmentadas.

Estos son sólo unos ejemplos de las características morfológicas que pueden extraerse de una región segmentada en una mamografía, pero existen muchas más como descriptores de Fourier [Ra97], mediciones de excentricidad, tamaño, densidad promedio [Ch98], entre otros.

3.5.B. Textura

La textura es la propiedad de las sensaciones causadas por la superficie externa de los objetos a través de los sentidos como el tacto. En el caso del procesamiento de imágenes digitales, la textura se usa para describir la sensación de experiencias no táctiles, como la apariencia superficial rugosa o lisa de una imagen. La textura puede también emplearse

como un patrón que se organiza en escala ascendente o descendente (especialmente en el caso de experiencias sensoriales no táctiles) donde los elementos individuales más pequeños forman el patrón no distinguible visualmente [WIK]. Por ejemplo en la figura 3.9 se muestran dos imágenes, una con características “rugosas” (izquierda) y otra con características “suaves” (derecha).

Las características texturales proveen una medición de la textura subyacente en una región dada. Una variedad de métodos se usan para extraer características texturales. En general existen dos grandes clasificaciones de los métodos de análisis textural: una primera, estructural, donde se analiza la imagen en el dominio espacial por medio del estudio de sus primitivas de representación, y una segunda, estadística, a través de las características de las distribuciones de niveles de gris en el histograma de la imagen y su correlación espacial.

Las características texturales de una imagen dependen del tamaño espacial de las primitivas de textura (representaciones geométricas utilizadas para la descripción o análisis de una imagen). Primitivas texturales grandes dan origen a texturas rugosas (como la superficie de las rocas), y primitivas pequeñas dan texturas finas (como la superficie de la seda). Una función de autocorrelación puede evaluar las medidas de rugosidad. Esta función evalúa las relaciones lineales espaciales entre primitivas. Si la primitiva es grande la función decrece lentamente conforme la distancia se incrementa y decrece rápidamente si la textura consiste de primitivas pequeñas. Sin embargo, si las primitivas son periódicas, entonces la autocorrelación se incrementa y decrece periódicamente con la distancia.

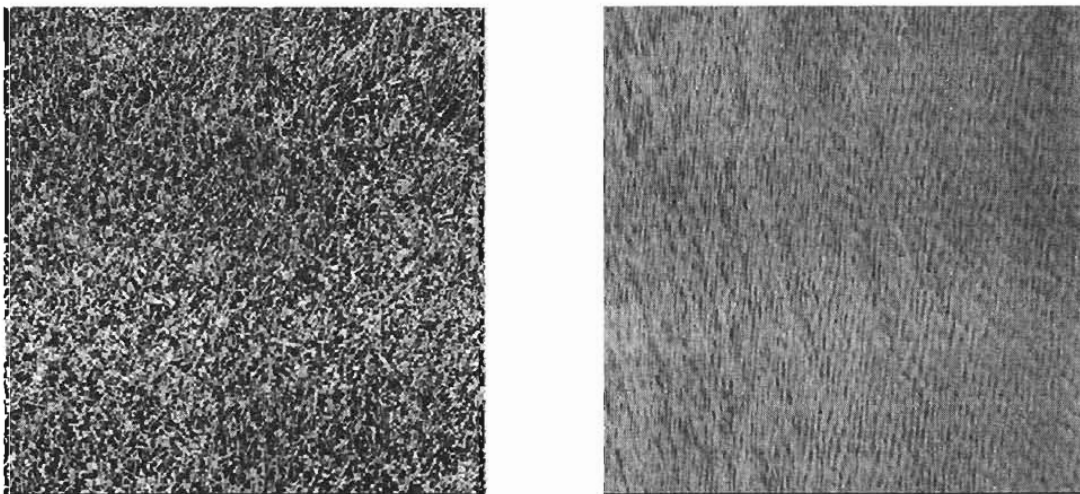


Figura 3.9. Izquierda: ejemplo de una imagen donde se observa una superficie “rugosa” del pasto. Derecha: ejemplo de una imagen donde se observa una superficie “suave” del agua (cortesía: SIPI DataBase Textures free domain).

Algunas aproximaciones comúnmente utilizadas incluyen el uso de frecuencias espaciales, frecuencias de bordes, longitud, distribución de probabilidad de unión de píxeles y máscaras especiales como las máscaras de Law. Se ha demostrado en varios estudios que diferentes métodos de extracción de características pueden llegar a diferentes resultados basados en la aplicación del dominio de estudio [Gu95].

CAPÍTULO IV MÉTODO EXPERIMENTAL

Este capítulo presenta la metodología de trabajo empleada para la selección de las mamografías digitales incluidas en el estudio, el análisis utilizado (textural y morfológico) y el análisis de los resultados que se presentan en el capítulo V. Las referencias principales del contenido en cuanto al análisis de las imágenes son: [Bu04], [Co98], [Ja02].

4.1. PRESENTACIÓN

De acuerdo a lo presentado en los capítulos anteriores, podemos resumir brevemente que el cáncer de mama se visualiza radiográficamente bajo diversas apariencias morfológicas (por ejemplo, la forma y bordes de masas tumorales) y configuraciones (como en el caso de grupos de microcalcificaciones). Debido a esto, se han propuesto diversos esquemas para estandarizar el dictamen de las imágenes mamográficas (clasificaciones de Wolfe, por ejemplo). En un principio se calificaba la imagen obtenida en película radiográfica, sin embargo este tipo de análisis resultó ser altamente subjetivo y se recomendó efectuar una lectura independiente por un segundo evaluador. Con el desarrollo de las herramientas de cómputo se propusieron otros tipos de análisis cuantitativos, semi-automáticos y automáticos (análisis objetivos), al tratar de analizar las regiones de alta densidad mamaria y alteraciones en la anatomía de la mama a través de una imagen mamográfica (película radiográfica), previamente digitalizada.

Estos métodos [Bo95, Bo99, By94, By96, By97, Ya98], han utilizado imágenes digitalizadas a partir de las obtenidas en película radiográfica. Además, se han centrado en el análisis de técnicas del procesamiento de imágenes como realce de imágenes, segmentación por umbralaje, análisis de áreas y medición de la dimensión fractal local. Otros trabajos [Ch98, Ch00, Sa01], han contribuido utilizando otras técnicas como segmentación de objetos de interés a través de diferentes parámetros morfológicos y texturales. Todos estos trabajos han proporcionado resultados en el cálculo de factores de riesgo (basados en el aumento de la densidad mamográfica) y caracterización de la mama, pero algunos de ellos obteniendo resultados parcialmente significativos en comparación con los métodos subjetivos y sin llegar aún a un resultado de utilidad clínica.

En la actualidad se han desarrollado equipos de obtención de imágenes directamente en formato digital sin tener que pasar por un proceso analógico (película radiográfica). Esto trae consigo un nuevo enfoque en el estudio de las imágenes mamográficas. De esta forma se pueden aplicar directamente en el archivo digital las más diversas herramientas del procesamiento digital de imágenes, reduciendo las probabilidades de degradación de la imagen.

Con esto en mente, este trabajo de tesis propone obtener parámetros característicos, a partir de las imágenes resultantes de un estudio de mamografía digital, de las imágenes de la mama de pacientes diagnosticadas con cáncer con respecto de las que, presentando una

imagen radiográficamente similar, su biopsia indica que no tienen cáncer. Este estudio es un análisis global de la mama, en base a parámetros de textura y morfología, no limitado a la región radiográficamente sospechosa.

Para este propósito se utiliza la metodología descrita en las secciones 4.2, 4.3 y 4.4.

4.2. SELECCIÓN DE PACIENTES

Debido a la necesidad de contar con mamografías digitales de pacientes que pudieran presentar cáncer en la evaluación radiográfica clínica, se seleccionaron aquellas pacientes cuyos estudios mamográficos son evaluados como BI-RADS 4 o 5 (ver sección 1.3 y Apéndice A.2). Esta selección obedece al hecho que las categorías BI-RADS 4 y 5 representan hallazgos probablemente malignos y altamente sugestivos de malignidad, respectivamente. De esta forma todas las mamografías digitales que se obtienen corresponden a pacientes con algún signo de anormalidad que haga sospechar que existen características de cáncer de mama. Además, los criterios empleados para la obtención de las mamografías digitales son:

- Criterios de inclusión: Mujeres con BIRADS 4 o 5 en un estudio digital.
Lesiones menores que, o del orden de, 1 cm.
- Criterios de exclusión: Mama con tratamiento previo y/o biopsia.
Lesiones palpables.
Embarazo y lactancia.
Tratamiento hormonal.

Las pacientes participantes fueron informadas y expresaron por escrito su anuencia a participar en este proyecto.

El banco de imágenes así formado por las mamografías digitales seleccionadas, puede contener mamografías de pacientes con mama principalmente densa, principalmente grasa o mixta. El tipo de alteraciones observadas en las mamografías digitales puede ser de cualquier tipo (ver anexo A.5 y A.6), siempre y cuando hagan que la imagen sea evaluada como BI-RADS 4 ó 5.

Para el diagnóstico positivo o negativo de cáncer de mama, las pacientes son sometidas a la toma de una muestra del tejido afectado (biopsia), y éste es sometido a un examen histopatológico.

El grupo de estudio se constituye de aquellas imágenes seleccionadas que corresponden a un resultado histopatológico positivo de cáncer de mama. 24 pacientes recibieron un diagnóstico de cáncer y se seleccionó la imagen cráneo-caudal (sección 1.1) de la mama afectada. Se tienen entonces 24 imágenes. El grupo control se forma con las imágenes correspondientes a un estudio histopatológico negativo de cáncer de mama. 78 pacientes pertenecen a esta situación y se seleccionó la imagen CC de la mama sospechosa, resultando en un total de 78 imágenes.

Las imágenes fueron obtenidas en el Instituto Nacional de Cancerología, Distrito Federal, México, en el periodo enero 2004- enero 2005.

4.3. ANÁLISIS MEDIANTE PROCESAMIENTO DIGITAL DE IMÁGENES

El banco de imágenes cuenta con mamografías digitales con una amplia variedad de afecciones mamarias malignas y benignas, razón por la cual no podemos proponer un método específico para un análisis particular de algún tipo de anormalidad. Sin embargo, es posible discernir texturas caracterizadas por componentes filamentosas, en particular en tejido fibroglandular, con propiedades de lacunaridad (de “lacuna”, huecos de diversos tamaños o estructuras), y existencia de masas con contornos irregulares. Las proyecciones en un plano de estructuras tridimensionales en red combinan diversas componentes y persisten los problemas de oclusión en particular en mamas densas. Ello impide obtener mediciones separadas de componentes puramente filamentosas, lacunares o de masas densas. De esta manera se propusieron extraer 2 parámetros, uno textural (ver sección 3.5.B), y otro morfológico (ver sección 3.5.A), que se supone están poco correlacionados: variaciones debidas a texturas parcialmente estructuradas (densidad de picos o valores extremos), y medidas de irregularidad en contornos de isodensidad (isocontornos). En el primer caso se emplea el *cociente promedio alto-ancho de extremos*, y en el segundo, factores de forma o índices de *compacidad*.

4.3.A. Cociente promedio alto/ancho de extremos (ρ)

Los parámetros de rasgos texturales caracterizados por contener elementos parcialmente estructurados, combinan información de intensidad y de extensión espacial. Uno de los parámetros más simples y efectivos, es el propuesto inicialmente por Haralick et.al. [Ha73]: el cociente de densidad de extremos (Ratio of Extrema Density). Una versión “sintonizable” para caracterizar enjambres de cromosomas en imágenes de microscopía a baja resolución (probada con éxito por [Co98]), es el *cociente promedio alto/ancho de extremos* o MHWRE por sus siglas en inglés (Mean Height-Width Ratio of Extrema).

Para calcular este parámetro textural en una región, se consideran barridos o perfiles lineales en una dimensión (1D) de una imagen. Una “firma textural” puede obtenerse a partir del cociente entre el alto y el ancho de todos los picos de intensidad en un perfil sencillo de la imagen. Un pico puede caracterizarse por su profundidad o altura (intensidad en niveles de gris) y por su ancho (separación entre extremos del pico). Se miden las proyecciones vertical (alto) y horizontal (ancho) de vectores ξ y ζ en cada pico del perfil (ver figura 4.1). La altura del pico i , ($i = 0, \dots, N-1$) se define como el promedio $(\xi^{\perp} + \zeta^{\perp})/2$ y el ancho del pico i , como la suma $(\xi^{\parallel} + \zeta^{\parallel})$; N es el número de picos en un perfil. Se realiza el cociente de la altura y el ancho y se promedia para los N picos del perfil y M , como el número de perfiles en la imagen. El promedio altura/ancho para un perfil es MHWRE y el promedio de éstos para la imagen es ρ .

El cálculo anterior sólo se efectúa para picos que cumplen las siguientes condiciones:

- El cociente entre el mínimo y máximo de las proyecciones verticales para un pico debe ser mayor que 0.6 (aproximadamente el valor para el cual el ojo humano distingue una característica “rugosa” en las imágenes texturales).
- La altura del pico debe ser mayor que un umbral de ruido η preestablecido.
- Todos los perfiles muestreados corresponden a perfiles dentro de la imagen de la mama.

La proyección vertical y horizontal de cada pico representa el alto y el ancho ilustrados en la figura 4.1.

La medición se repite para un número predeterminado de muestras de perfiles por imagen (aprox. 500). El valor ρ de una imagen (o de una región de interés ROI) es el valor promedio de todos los MHWRE obtenidos en cada perfil [Co98].

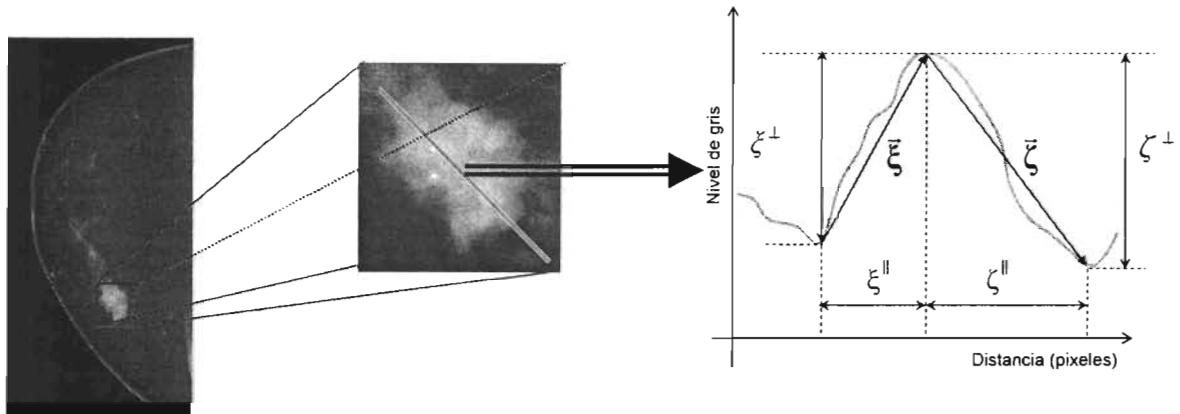


Figura 4.1. Izquierda: Ejemplo de mamografía. Centro: Ejemplo de una región de la mamografía. Derecha: Descripción del cociente altura/ancho de un pico.

El parámetro textural ρ se define como se muestra en las ecuaciones 4.1 a 4.3.

$$\rho = \frac{1}{M} \sum_{i=1}^M MHWRE_i \quad (4.1)$$

$$MHWRE = \frac{1}{N} \sum_{i=1}^N w_i R_i \quad (4.2)$$

donde:

$$R_i = \frac{(\xi_i^\perp + \zeta_i^\perp)/2}{\xi_i^\parallel + \zeta_i^\parallel} \quad (4.3)$$

y

$$w_i = \begin{cases} 1 & \text{si } \min(\xi_i^\perp, \zeta_i^\perp) > 0.6 \text{ y } (\xi_i^\perp + \zeta_i^\perp)/2 > \eta \\ 0 & \text{otros casos} \end{cases}$$

En estas ecuaciones se ha usado una notación usual en la literatura de física de superficies: $\xi_i^{\perp}, \zeta_i^{\perp}$ representan alturas características (componentes perpendiculares a la superficie) y $\xi_i^{\parallel}, \zeta_i^{\parallel}$ representan anchos característicos (componentes tangenciales a la superficie).

4.3.B. Compacidad de estructuras (κ)

Anteriormente en la sección 3.5.A se citó el parámetro C de circularidad; en la literatura este parámetro se utiliza de dos maneras diferentes: tal como se definió en 3.5.A o se toma su inverso. En este trabajo de tesis se utiliza una medición de compacidad discreta κ (igual al inverso de la circularidad C) de una estructura, dada por la ecuación 4.4 (ver por ejemplo [Qi99]).

$$\kappa = \frac{4\pi A}{P^2} \quad (4.4)$$

donde P es el perímetro de la estructura (aproximado por el número de píxeles que la delimita) y A es el área de la estructura (en número de píxeles).

Con la compacidad así definida en la ecuación 4.4, se tiene como valor máximo 1, correspondiente a un círculo. La motivación en la aplicación de este análisis es precisamente que ciertos tipos de lesiones malignas se clasifican, dependiendo del grado de malignidad y de su infiltración en el tejido mamario, a través de la evaluación visual de sus bordes o contornos cada vez más irregulares (i.e., menos compactos). Dado que las estructuras mamarias de interés tienen mayor intensidad (niveles de gris) que el fondo y tejidos grasos (éstos no son tejidos malignos), se aplicó un valor de umbral característico individual para seleccionar sólo las estructuras mamarias posiblemente anormales. Posteriormente, en las zonas seleccionadas se encuentran aquellos píxeles que tienen un mismo nivel de gris, tales que formen estructuras cerradas que representan contornos de isonivel. Un ejemplo de la escala de valores y formas que se obtienen al usar el parámetro κ se muestra en la figura 4.2.

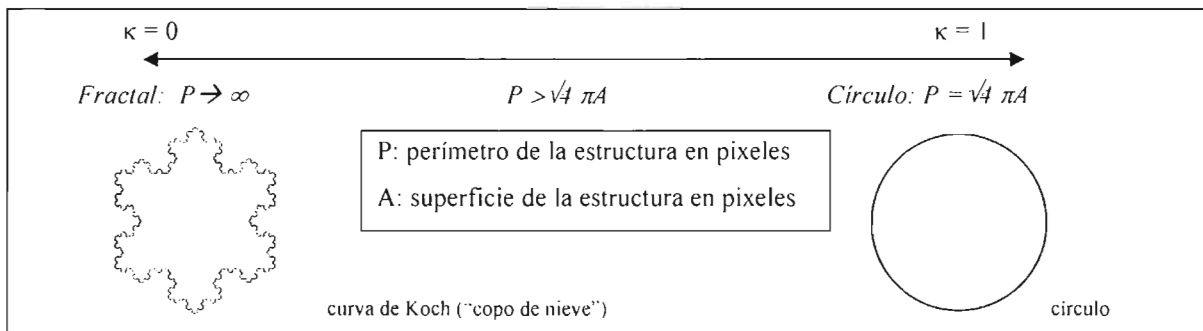


Figura 4.2. Ejemplo de la escala de valores que se obtiene al utilizar el factor de compacidad

El análisis de la imagen digital por medio de la compacidad puede entenderse interpretando la imagen como un relieve tridimensional donde las coordenadas (i,j) para un punto son las mismas que las del píxel en la imagen bidimensional y la coordenada $k = I(i,j)$ es el nivel de gris del píxel (i,j) interpretada como una altura. De esta forma tenemos un objeto tri-

dimensional (3D) parecido a los mapas topográficos al cual se le realizan cortes a una altura fija para obtener las estructuras.

La determinación del valor de umbral característico antes mencionado, requiere de un estudio simple de los valores de gris obtenidos para cada una de las imágenes analizadas. Se utiliza el histograma individual de valores de gris de la mama, que no es otra cosa que la gráfica de la frecuencia con que cada nivel de gris aparece en la imagen mamaria. La figura 4.3 muestra un ejemplo de histograma para una imagen típica en que los tonos de gris van desde 453 hasta 8695, siendo un valor cercano a 1000 el más frecuente. El pico alrededor del valor máximo del histograma parece corresponder principalmente a una mezcla entre tejido graso (niveles de gris altos) y tejido fibroglandular (niveles de gris bajos). Obviamente, su anchura dependerá de la composición de cada mama. La cola del histograma que se extiende hacia valores altos de tonos de gris representa al tejido graso. A la izquierda del pico máximo se aprecia claramente una estructura angosta que corresponde a regiones mamarias de alto coeficiente de atenuación (identificadas por el radiólogo como radiopacas) probablemente originadas en el tejido glandular mamario.

Se seleccionó el nivel de umbral característico para cada imagen como el nivel de gris medio entre el nivel de gris del pico máximo del histograma y el nivel de gris del pico (u hombro) a la izquierda (ver figura 4.3). Con el nivel característico así definido se pretende entonces ofrecer, en una primera aproximación, una separación entre los tejidos (probablemente malignos) que queremos analizar, del resto de la mama. Es de notar que este nivel varía con cada imagen al tener éstas diferentes histogramas.

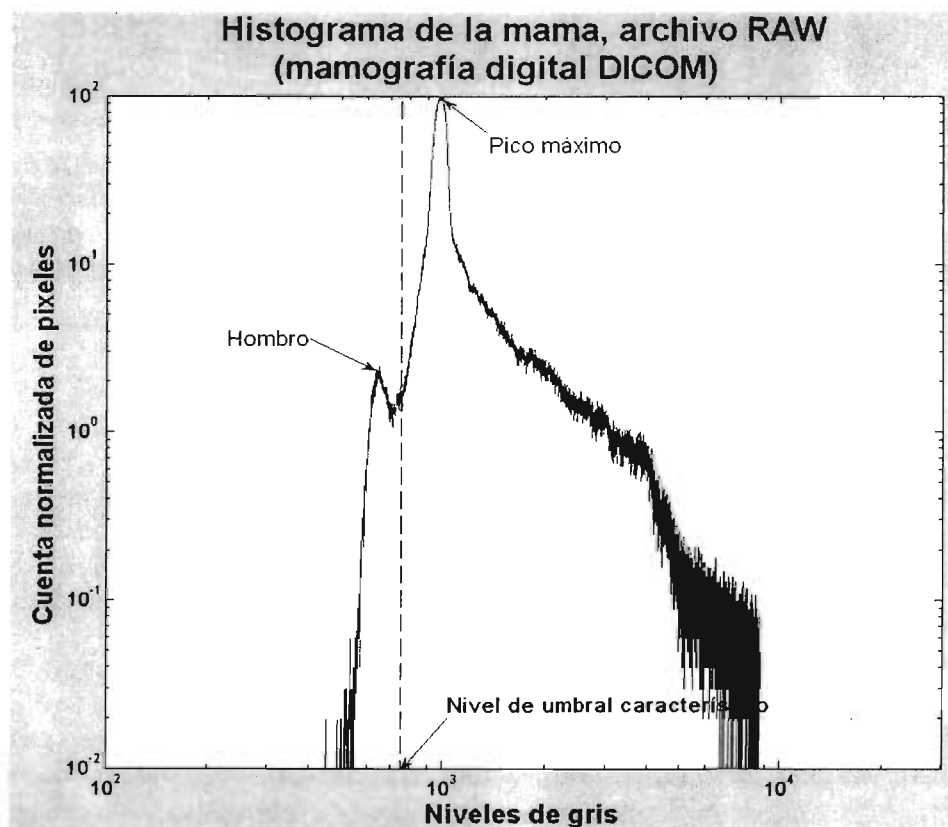


Figura 4.3. Definición del nivel de umbral característico. Se presenta el histograma de una imagen RAW.

El nivel de umbral característico se aplica para segmentar (Sección 3.4) la imagen. A las estructuras obtenidas de este ejercicio (imagen binaria) se les extrae el factor de compacidad κ definido anteriormente. Esto puede visualizarse en el mapa topográfico de la mama como un corte realizado justo en el nivel umbral característico, obteniéndose estructuras de isonivel a las cuales se les calcula κ . Debido a que el nivel umbral característico es propio de cada imagen y es posible que las composiciones de tejido mamario aparezcan traslapadas, y en este caso el umbral no establezca la diferencia que se pretende al aplicarlo, se aplican varios valores cercanos al nivel umbral, promediándose las κ y calculando su desviación estándar. Con esto se pretende obtener una medida de la compacidad global de las estructuras mamarias. Un ejemplo de esta situación se muestra en la figura 4.4.

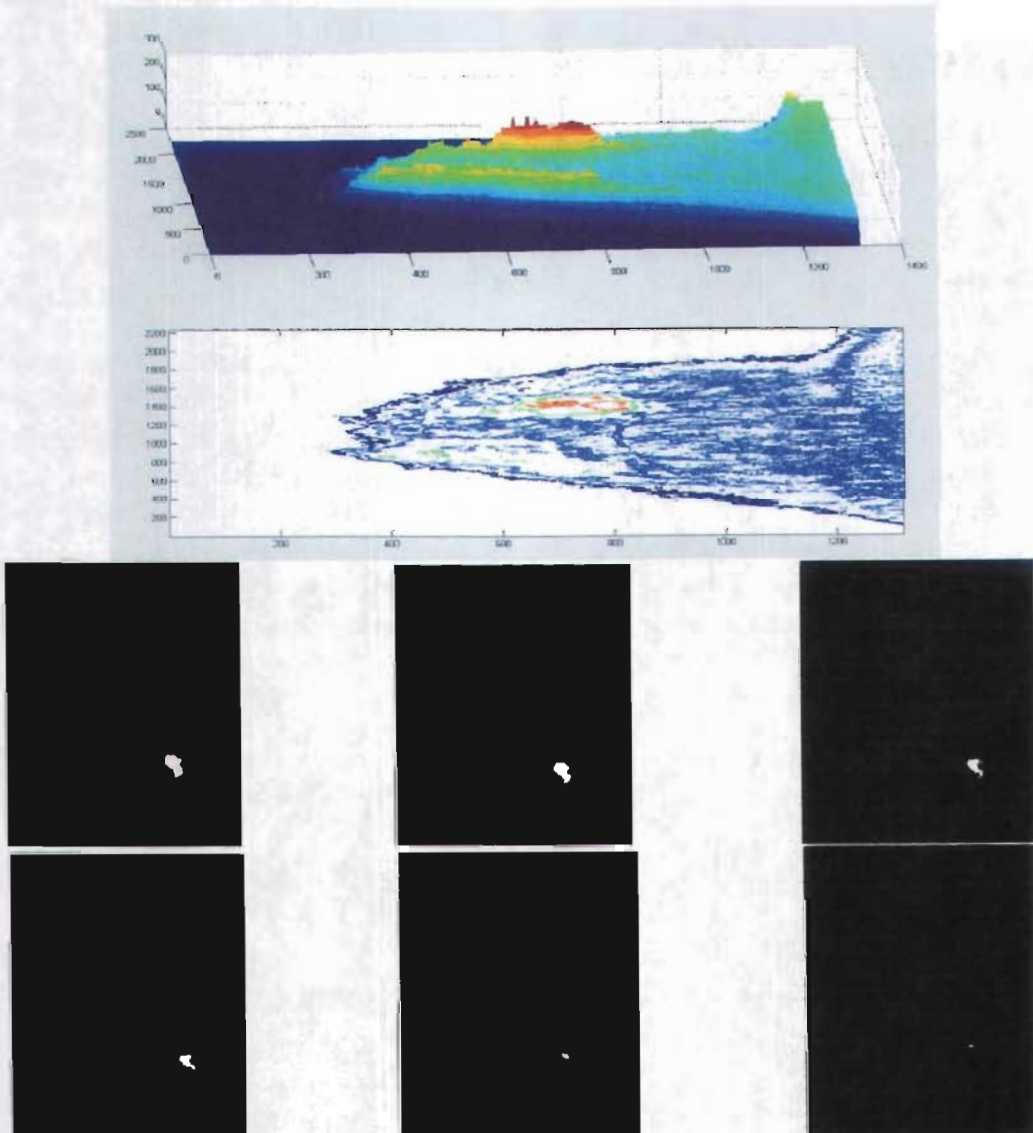


Figura 4.4. Cortes múltiples a diferentes alturas (niveles de gris) del mapa topográfico de una imagen procesada de la mama. Imagen superior. Ejemplo de una mamografía digital como objeto tridimensional, se visualizan en rojo las zonas más radiopacas de la mama. Imágenes inferiores: De la imagen superior izquierda a inferior derecha se muestran ejemplos (imágenes binarias) de las estructuras obtenidas al realizar cortes de nivel progresivamente ascendentes al objeto tridimensional de la mama en la imagen superior. La estructura en blanco corresponde a la zona roja de la imagen superior.

Al considerar los parámetros anteriormente propuestos, es de notar que el parámetro ρ caracteriza la distribución de intensidades en niveles de gris a lo largo de líneas horizontales o verticales y permite describir el carácter “rugoso” de texturas diversas. El parámetro κ caracteriza la forma irregular de un contorno de isodensidad de una estructura de alta densidad, y por lo tanto, posiblemente sospechosa.

Con los parámetros ya propuestos para el análisis de la textura (ρ) y la morfología (κ) global de las imágenes, se desarrollaron rutinas (llamadas scripts) en Matlab (plataforma de desarrollo empleada) para la evaluación de las imágenes. Se sugiere ver el Anexo B.1 para encontrar más consideraciones acerca de Matlab y el análisis de las mamografías digitales.

4.4. ANÁLISIS ESTADÍSTICO PROPUESTO

Con los grupos de estudio y control ya definidos se propuso una primera fase de entrenamiento donde se seleccionaron 18 imágenes de cada grupo. En la fase de entrenamiento se caracterizó cada subgrupo con los parámetros propuestos y se trató de definir una metodología de detección o diferenciación de las mamografías digitales de pacientes con cáncer, con respecto de las que no tienen.

En una primera aproximación, se estableció comparar en los dos grupos de imágenes el promedio de los parámetros ρ y κ y su desviación estándar. Esta primera comparación serviría para evaluar si existen intervalos para ρ y κ significativamente diferentes en las imágenes seleccionadas del grupo de estudio comparados con los del grupo control (clasificación “lineal”). Si los intervalos fueran significativamente diferentes se podría definir un intervalo para ρ y κ al que se le asociarían todas las imágenes correspondientes a un resultado histopatológico positivo de cáncer de mama. Cabe remarcar que incluso en distribuciones normales puede existir un traslape entre las diferentes poblaciones como se ejemplifica en la figura 4.5.

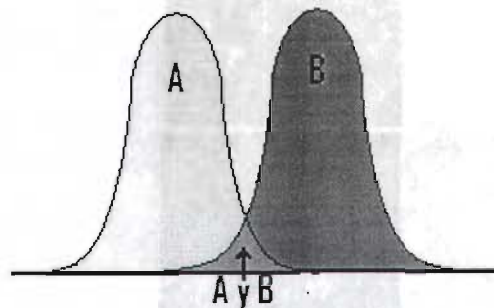


Figura 4.5. Ejemplo de dos distribuciones A y B. Cada distribución tiene un valor medio significativamente distinto del otro, pero aún así existe un traslape entre ambos (A+B) en la región intermedia de las distribuciones.

Además, con los valores promedios de ρ y κ obtenidos para cada imagen de ambos grupos (estudio y control en la fase de entrenamiento), se propuso visualizarlos en una gráfica de κ contra ρ , para ofrecer una diferenciación visual entre los grupos de “nubes” (conjuntos de puntos) que se formaran a partir de haber graficado puntos con coordenadas (ρ_i, κ_i) . Tal

gráfico es conocido en inglés como “scatter plot”. De esta forma, si las nubes de puntos no se confundieran para los dos grupos, se podría ofrecer también una región donde todos los puntos contenidos en ésta fueran asociados con imágenes de pacientes con cáncer. En la figura 4.6 se muestran dos ejemplos de distribuciones para 2 tipos de nubes. En el primer caso se observa una clara separación entre los grupos de puntos, por lo que es posible realizar una separación paramétrica lineal (propuesta en este trabajo). En el segundo caso se tiene una separación no lineal de los grupos de puntos, por lo que es necesario aplicar una separación elaborada usando herramientas de reconocimiento de patrones, como redes neuronales.

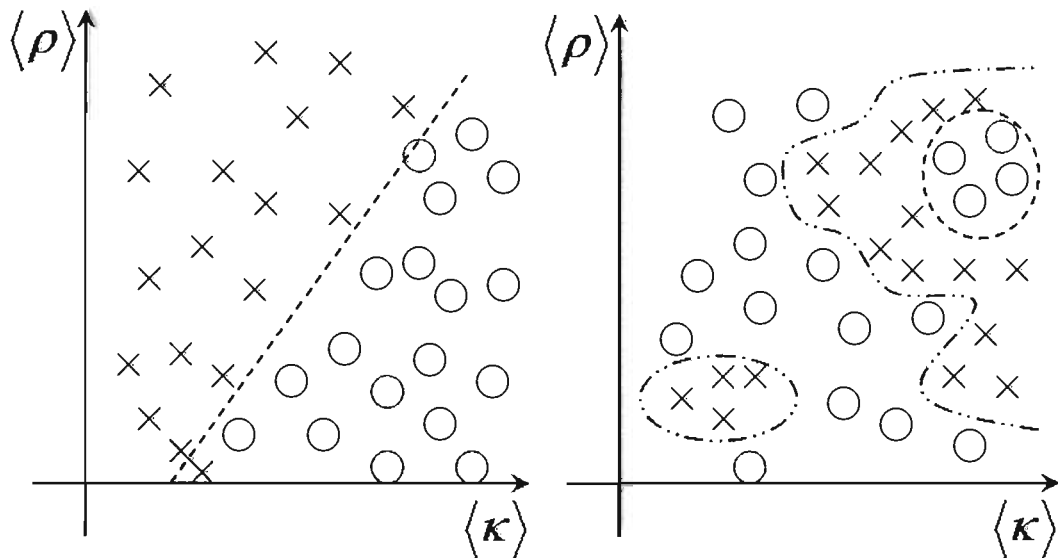


Figura 4.6. Izquierda: Ejemplo de 2 tipos de poblaciones diferentes (18 cruces y 18 círculos), separadas mediante una clasificación lineal. Derecha: Ejemplo de 2 tipos de poblaciones (18 cruces y 18 círculos), las cuales forman grupos de nubes y su separación se realiza mediante un proceso no lineal.

En una segunda fase, con los resultados obtenidos de las regiones (ρ, κ) asociadas a imágenes de pacientes con cáncer, se alimentaron las rutinas propuestas para la evaluación (“tipo ciego”, ver sección 5.4) de los parámetros ρ y κ , luego estas rutinas se aplicaron a todas las imágenes restantes. Esta etapa representó la fase de validación para evaluar el porcentaje de discriminación efectiva entre las imágenes del grupo de estudio y del grupo control. El resultado de la fase de validación ofrece entonces a cada imagen una sugerencia de malignidad o benignidad, de acuerdo con la evaluación de los resultados para ρ y κ en la fase de entrenamiento.

CAPÍTULO V

RESULTADOS y ANÁLISIS DE RESULTADOS

A continuación se exponen los resultados obtenidos al utilizar la metodología del capítulo anterior a las 102 mamografías digitales obtenidas.

5.1. ANÁLISIS DE HISTOGRAMAS

En un principio se utilizó el programa de dominio público ImageJ [IMJ], para pasar las mamografías digitales del formato DICOM a TIFF, a fin de que Matlab 6.5(R13), pudiera reconocer las imágenes adecuadamente. Sin embargo en julio 2004 el Dr. A. Burgess publicó un artículo [Bu04], donde presenta un histograma normalizado de 884 mamografías digitales clínicas adquiridas con un Senographe 2000D. En este trabajo de tesis se comparó, en primer lugar, el histograma global de la región mamaria de nuestras imágenes mamográficas digitales con el histograma publicado por Burgess, con la finalidad de validar nuestro proceso de lectura de las mamografías digitales. En la figura 5.1 se muestra la comparación del histograma publicado en [Bu04], con el histograma obtenido usando nuestras mamografías digitales una vez convertidas a TIFF con ImageJ.

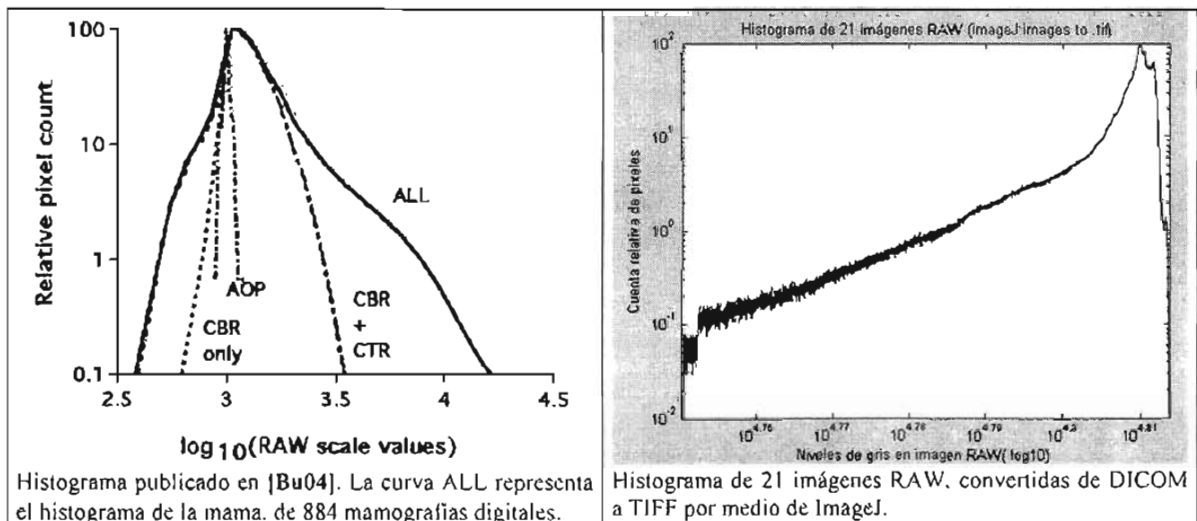


Figura 5.1. Comparación del histograma publicado por el Dr. Burgess (izquierda) en [Bu04], contra el histograma global de la mama (derecha) al utilizar 21 imágenes RAW convertidas de formato DICOM a TIFF por medio de ImageJ.

Tal como puede apreciarse en la figura 5.1, la curva ALL (izquierda) y la curva de la derecha no son iguales. Esto puede deberse a que las mamografías digitales son diferentes o que el programa ImageJ introduce algún cambio a la imagen al convertir la imagen del formato DICOM a TIFF. Sin embargo, todas las mamografías fueron obtenidas en un equipo de la misma marca y modelo, por lo que se deduce que lo más probable es que el programa ImageJ sea el causante de la diferencia en los histogramas, pues se esperaba que éstos fueran muy similares.

Una vez que salió al mercado el programa Matlab vers. 7(R14), donde se indicaba en la página web que muchos de los errores y problemas de la versión anterior de Matlab habían sido corregidos al trabajar con imágenes en formato DICOM, se decidió utilizar esta versión para el análisis de las mamografías digitales. De esta forma, no fue necesario utilizar el programa ImageJ para convertirlas mamografías digitales de DICOM a TIFF, ya que éstas podrían analizarse directamente en Matlab 7(R14). En la figura 5.2 se presenta el histograma normalizado de la región mamaria de 46 imágenes de muestras.

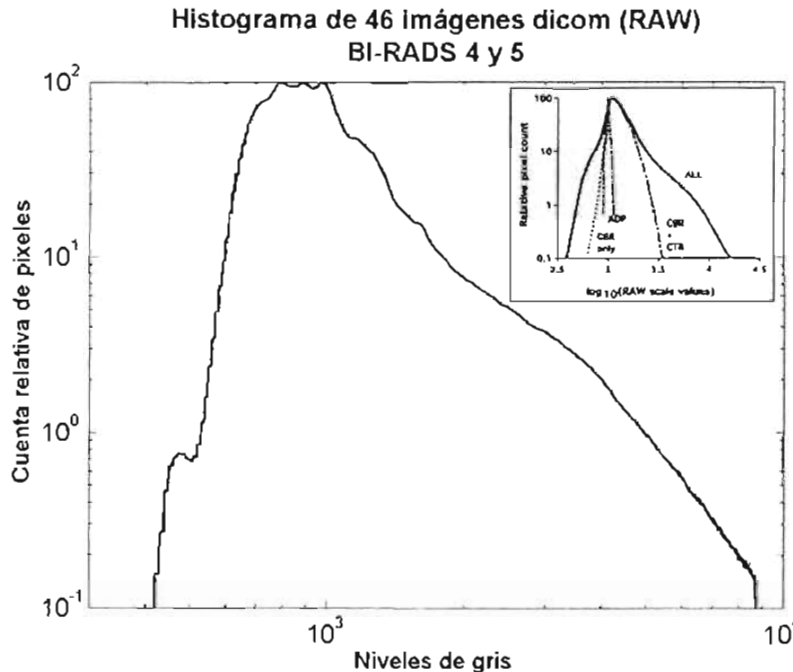


Figura 5.2. Histograma normalizado de la región mamaria para 46 mamografías digitales (18 cáncer, 28 benignas). En la esquina superior derecha se muestra una imagen del histograma publicado por el Dr. Burgess.

Como puede apreciarse, la gráfica de la figura 5.2 es similar a la curva ALL de la figura 5.1 (gráfica izquierda), por lo que concluimos que efectivamente el programa ImageJ introducía algún cambio al archivo de la imagen (posiblemente una inversión de escala), al convertir una imagen del formato DICOM al formato TIFF. Por esta razón se decidió utilizar solamente el programa Matlab7(R14), para el análisis de las imágenes, ya que ahora estamos confiados en que estamos leyendo adecuadamente los datos de las mamografías digitales. Los detalles en el pico alrededor del valor 1000, que no aparecen en el diagrama de Burgués (Figura 5.1, izquierda), son consecuencia del tamaño de la muestra: 46 imágenes en nuestro análisis y 884 en la de Burguess.

En la figura 5.3 se presenta un ejemplo de la mascarilla adaptable (binaria) utilizada para analizar únicamente la región interna de la mama (ésta cambia de acuerdo a cada imagen mamográfica). La figura de la izquierda muestra la imagen mamográfica RAW que despliega el programa de Matlab 7(R14), mientras que a la derecha se muestra la mascarilla resultado de un proceso binario de cerradura morfológica seguido por un proceso de apertura morfológica de la imagen (ver anexo C.1 para más detalle de esta operación). Además se subtrae una parte del borde de la mama (aproximadamente 10 píxeles), y aproximadamente 19 píxeles correspondientes al borde de la mama cerca del tórax y al

borde superior e inferior de la imagen; esto para evitar cualquier irregularidad o efecto de borde en la imagen que pudiera traer un cambio en el análisis de cada imagen. De esta forma, se extraen los parámetros ρ y κ únicamente en la parte de la imagen original delimitada por la mascarilla binaria (en la figura 5.3 cada imagen se enmarca en un recuadro negro, para que se aprecie la sustracción de píxeles realizada a la imagen izquierda).

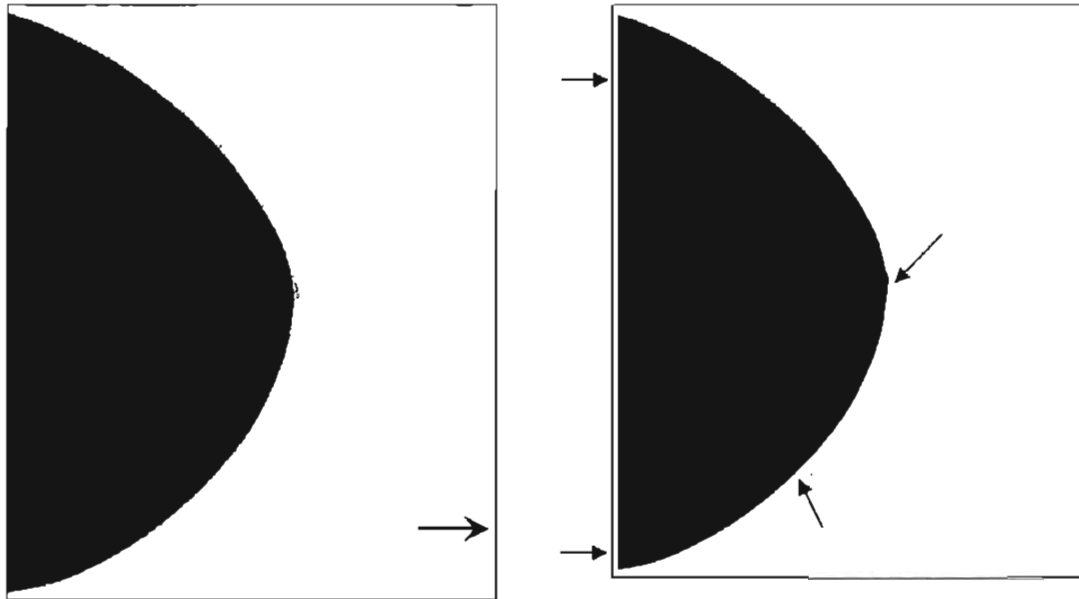


Figura 5.3. Izquierda: Imagen desplegada por Matlab 7(R14) de una mamografía digital RAW en formato DICOM; nótese que en la parte derecha de la imagen (donde apunta la flecha), se observa una franja con niveles de gris bajos como los que aparecen en la mama. Derecha: Mascarilla binaria obtenida de un proceso binario de cerradura morfológica seguido de una apertura morfológica de la imagen original. Nótese que la mascarilla delimita efectivamente el área de la mama descontando aproximadamente 10 a 19 píxeles del borde de la mama (19 píxeles del borde de la imagen (izquierdo, superior e inferior para esta imagen) y 10 píxeles en el resto del borde mamario), como se indica por las flechas pequeñas.

5.2. FASE DE ENTRENAMIENTO

Una vez programadas las rutinas necesarias para el análisis (scripts) de texturas por medio del parámetro ρ y de la morfología por medio de la compacidad κ , éstas se aplicaron a un banco de imágenes para observar sus resultados con estas herramientas.

La fase de entrenamiento consistió de:

- Seleccionar 18 mamografías de pacientes con cáncer y 18 con lesiones benignas, como banco de imágenes de entrenamiento.
- Aplicar el script de cociente promedio alto/ancho de extremos (ρ), en la región de la mama, a las imágenes seleccionadas y analizar los resultados.
- Aplicar el script del análisis por compacidad (κ), en la región de la mama, a las imágenes seleccionadas y estudiar los resultados.
- Estudiar en conjunto los resultados anteriores a fin de poder obtener un programa que los correlacione y pueda discriminar lo mejor posible las mamografías de pacientes con cáncer de las que tienen lesiones benignas.

5.2.A. Resultado del cociente promedio alto/ancho de extremos (ρ)

Como se indica en la metodología, este parámetro analiza la imagen por medio de barridos horizontales a lo largo de la dirección vertical de la imagen (no se observan cambios significativos al usar barridos verticales). Con cada barrido se obtiene su promedio del cociente alto/ancho de extremos junto con su desviación estándar. En la tabla 5.1 se muestra el significado de cada parámetro obtenido.

Tabla 5.1. Significado de los parámetros extraídos: ρ

Parámetro	Asignación
ρ (de la imagen)	Promedio de todos los promedios alto/ancho de extremos de todos los barridos unidimensionales efectuados en la imagen
σ_ρ	Desviación estándar de los promedios alto/ancho de extremos de todos los barridos efectuados en la imagen
$\bar{\sigma}$	Promedio de las desviaciones estándar de todos los barridos efectuados en la imagen
$\sigma_{\bar{\sigma}}$	Desviación estándar de las desviaciones estándar de todos los barridos efectuados en la imagen

Debido a la naturaleza del análisis del ρ es posible hacer que este parámetro analice solamente las variaciones de intensidad en los niveles de gris más importantes en la imagen. Esto se realiza al seleccionar un estudio sólo para variaciones en niveles de gris mayores que un cierto valor designado, tal como se observa en el ejemplo de la figura 5.4. Este valor en el script es denominado como maxgap y se observó que su variación trae diferencias no despreciables en el resultado final. De esta forma, se seleccionaron 2 valores para el maxgap: uno pequeño, igual a 10, que llamaremos ρ_{10} (se analizan prácticamente todas las variaciones en la textura de la imagen) y uno grande, igual a 100, que denominaremos ρ_{100} (se analizan únicamente variaciones importantes en niveles de gris debido principalmente a tejidos con alto coeficiente de atenuación).

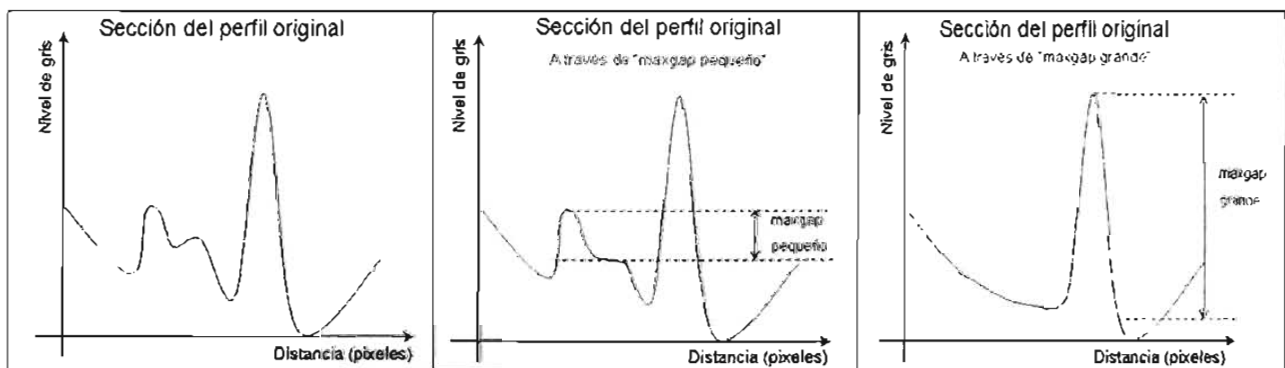


Figura 5.4. Ejemplo del funcionamiento de la selección de picos a analizar de acuerdo al criterio de maxgap (ver texto). Izquierda: Sección de un perfil en la imagen, se observan 3 picos. Centro: Se analizan solamente dos picos, superiores al "maxgap_pequeño" elegido. Derecha: Se analiza solamente un pico superior o igual al "maxgap_grande" elegido.

Los resultados obtenidos en el análisis del ρ se muestran en la tabla 5.2 para las mamografías de pacientes con cáncer (grupo de estudio en la fase de entrenamiento) y en la

tabla 5.3 para las mamografías de pacientes con lesiones benignas (grupo control en la fase de entrenamiento).

Tabla 5.2. Resultados del ρ para imágenes del grupo de estudio en la fase de entrenamiento: $\maxgap10$ y $\maxgap100$.

Imagen	ρ_{10}				ρ_{100}			
	ρ	σ_{ρ}	$\bar{\sigma}$	$\sigma_{\bar{\sigma}}$	ρ	σ_{ρ}	$\bar{\sigma}$	$\sigma_{\bar{\sigma}}$
PA5_IZQ	11.85	2.28	6.23	1.23	12.08	6.59	7.68	3.57
PA29_DER	10.52	2.62	5.13	0.98	10.31	6.00	5.53	3.82
PA32_DER	7.16	0.68	3.71	0.77	14.69	8.73	4.53	2.04
PA33_IZQ	9.26	1.30	6.52	1.20	27.82	12.65	5.30	3.10
PA35_DER	11.99	6.66	5.68	1.86	12.22	9.00	6.10	---
PA43_IZQ	12.98	4.02	7.16	1.63	27.83	10.64	7.40	3.82
PA51_DER	19.19	10.61	7.80	5.50	22.41	14.33	7.02	6.58
PA66_DER	9.55	1.08	4.70	0.54	5.73	7.18	5.70	7.32
PA71_IZQ	11.33	2.11	6.33	1.17	12.72	6.50	8.50	5.85
PA73_IZQ	8.01	0.89	5.17	1.09	24.21	18.32	5.76	4.64
PA74_IZQ	8.91	1.58	4.52	0.87	13.98	8.48	8.90	6.45
PA85_DER	9.65	0.78	5.59	0.75	11.11	9.38	8.19	5.09
PA104_DER	11.21	2.38	4.01	1.09	12.58	2.16	---	---
PA106_DER	8.51	1.09	3.26	0.88	10.05	---	---	---
PA108_DER	10.01	3.37	5.27	1.31	17.94	8.75	5.51	4.86
PA117_DER	14.40	3.70	3.98	1.43	---	---	---	---
PA119_DER	12.12	6.54	7.00	2.95	32.67	17.69	5.36	3.48
PA57_DER	10.58	2.81	4.89	1.30	25.74	6.61	---	---
Promedio	10.96	3.03	5.39	1.48	17.30	9.56	6.53	4.66
σ	2.75	2.60	1.27	1.14	7.87	4.31	1.39	1.55

--- Estas imágenes presentan variaciones menores que 100 niveles de gris. Es decir, en promedio toda la imagen presenta una textura con poca variación.

Tabla 53. Resultados del ρ para imágenes del grupo control en la fase de entrenamiento: $\maxgap10$ y $\maxgap100$.

Imagen	ρ_{10}				ρ_{100}			
	ρ	σ_{ρ}	$\bar{\sigma}$	$\sigma_{\bar{\sigma}}$	ρ	σ_{ρ}	$\bar{\sigma}$	$\sigma_{\bar{\sigma}}$
PA2_DER	11.59	2.04	6.57	1.93	27.09	14.89	19.93	8.22
PA11_IZQ	9.82	0.86	6.64	1.24	28.32	21.21	0.48	0.00
PA12_DER	11.50	2.16	5.78	1.23	17.90	9.62	8.80	6.14
PA13_DER	13.59	3.79	6.59	1.18	19.08	9.43	11.49	7.88
PA14_DER	16.98	6.88	7.78	2.36	23.01	10.88	10.13	6.75
PA15_DER	13.89	7.73	7.93	2.31	27.25	16.45	11.23	7.36
PA16_IZQ	9.30	2.13	5.15	0.69	18.69	7.56	7.56	4.62
PA17_DER	9.36	2.56	4.96	0.69	22.74	13.12	8.61	10.90
PA18_IZQ	9.42	1.22	4.21	0.92	20.94	8.89	1.91	0.00
PA19_IZQ	17.21	10.30	6.73	3.12	17.00	9.76	7.83	9.01
PA21_IZQ	7.60	0.77	3.95	0.57	16.88	13.04	3.07	3.46
PA23_IZQ	7.70	1.44	2.97	0.78	18.89	---	---	---
PA24_IZQ	13.47	4.61	7.42	1.53	22.62	13.38	7.27	5.92
PA26_DER	10.89	2.75	5.40	1.22	20.01	5.55	5.64	5.07
PA28_IZQ	8.56	0.98	4.98	0.73	23.31	9.93	6.97	5.60
PA39_DER	11.53	3.22	4.71	1.25	12.10	6.75	2.78	3.20
PA45_IZQ	10.97	2.43	5.38	1.06	10.03	7.26	6.61	5.54
PA50_IZQ	12.11	4.24	6.66	1.33	32.49	7.04	10.02	6.66
Promedio	11.42	3.34	5.77	1.34	21.02	10.87	7.67	5.67
σ	2.79	2.61	1.37	0.69	5.60	4.06	4.52	2.87

--- Estas imágenes presentan variaciones menores que 100 niveles de gris. Es decir, en promedio toda la imagen presenta una misma textura con poca variación.

En las tablas 5.2 y 5.3 hay algunas casillas vacías (denotadas por ---), al analizar la imagen con un maxgap de 100 niveles de gris. Esto muestra que la imagen analizada tiene variaciones menores que 100 niveles de gris (si la casilla correspondiente a ρ esta vacía), o que las variaciones texturales globales en la imagen son muy pequeñas para analizar la imagen con este valor umbral (si sus demás casillas están vacías).

Al comparar los promedios y desviaciones estándar de ρ_{10} y ρ_{100} en la tabla 5.2, observamos diferencias significativas: $\langle \rho_{10} \rangle = 10.96 \pm 2.75$ contra $\langle \rho_{100} \rangle = 17.30 \pm 7.87$ y $\langle \sigma_{\rho_{10}} \rangle = 3.03 \pm 2.60$ contra $\langle \sigma_{\rho_{100}} \rangle = 9.56 \pm 4.31$. Si tomamos en cuenta que mientras más grande es el parámetro de ρ las estructuras son visualmente más “rugosas” (los perfiles lineales tienen picos más grandes y estrechos), entonces ρ_{10} capta las variaciones más finas en la textura de la imagen para las mamografías analizadas. Además, el promedio de la desviación del ρ nos indica la variación de la textura a lo largo de la imagen, por lo que concluimos que las características más suaves de la textura de la imagen no tienen mucha variación, mientras que las características más rugosas cambian drásticamente a lo largo de la imagen.

En la tabla 5.3, se observa el mismo comportamiento que en la tabla 5.2, en cuanto a la comparación entre los parámetros ρ_{10} y ρ_{100} .

Al comparar los promedios y desviaciones estándar de ρ_{10} y ρ_{100} en las tablas 5.2 y 5.3 cada uno por separado, no se obtienen diferencias significativas que permitan diferenciar los dos grupos de imágenes (pertenecientes a pacientes con cáncer y con lesiones benignas), puesto que los intervalos obtenidos en cada caso ($\langle \rho \rangle \pm \sigma_{\rho}$) se confunden.

5.2.B. Resultado de la compacidad (κ)

De acuerdo con la metodología, debido a que el nivel umbral característico (definido en la sección 4.3.B), es propio de cada imagen y es posible que las composiciones de tejido mamario aparezcan traslapadas (y en este caso el umbral no establezca la diferencia que se pretende al aplicarlo), se aplican varios valores cercanos al nivel umbral, promediándose las κ y calculando su desviación estándar. En la tabla 5.4 se muestra el significado de cada parámetro obtenido en este análisis. El script de compacidad entrega varios valores más relacionados con el perímetro de las regiones y su superficie, entre otros, pero sólo se presentan los de la tabla 5.4 ya que son los valores principales que se deseaba obtener. Los resultados de la compacidad global de las mamografías del grupo de estudio (fase de entrenamiento) se muestran en la tabla 5.5 y las del grupo control (fase de entrenamiento) en la tabla 5.6.

Al comparar los promedios de compacidad y su desviación estándar no se observa una diferencia significativa. Para las imágenes de pacientes del grupo de estudio (fase de entrenamiento) $\langle \kappa \rangle_{\text{estudio_entrenamiento}} = 0.04558 \pm 0.01482$, contra la compacidad para las imágenes del grupo control (fase de entrenamiento) $\langle \kappa \rangle_{\text{control_entrenamiento}} = 0.04454 \pm 0.2979$. Sin embargo, si tomamos en cuenta los promedios de las desviaciones estándar de cada

imagen, tenemos aproximadamente que éste es el doble para las imágenes del grupo de estudio (fase de entrenamiento) comparado al de las imágenes del grupo control (fase de entrenamiento) $\langle\sigma_{\kappa}\rangle_{\text{estudio_entrenamiento}}=0.0202\pm 0.0136$ contra $\langle\sigma_{\kappa}\rangle_{\text{control_entrenamiento}}=0.0135\pm 0.0130$, esto indica que aunque los valores promedios de compacidad para los casos de cáncer y lesiones benignas son muy similares, la variación en la distribución de éstos es mayor para el caso de cáncer.

Tabla 5.4. Significado de los parámetros extraídos: κ

Parámetro	Asignación
Compacidad (global de la imagen)	Promedio de los valores de compacidad obtenidos de todas las estructuras de cada nivel analizado
$\sigma_{\text{compacidad}}$	Desviación estándar de los promedios de compacidad obtenidos de cada nivel analizado

Tabla 5.5. Resultados del análisis de compacidad para imágenes de pacientes con cáncer.

Imagen	κ	σ_{κ}
PA5_IZQ	0.06028	0.020306
PA29_DER	0.05961	0.023332
PA32_DER	0.03497	0.023631
PA33_IZQ	0.06921	0.038274
PA35_DER	0.04884	0.019443
PA43_IZQ	0.06204	0.051139
PA51_DER	0.04775	0.009405
PA66_DER	0.04231	0.014992
PA71_IZQ	0.04837	0.011891
PA73_IZQ	0.02315	0.002831
PA74_IZQ	0.02633	0.010053
PA85_DER	0.04294	0.017257
PA104_DER	0.02984	0.012996
PA106_DER	0.05936	0.049770
PA108_DER	0.04602	0.017990
PA117_DER	0.05643	0.009673
PA119_DER	0.01546	0.006694
PA57_DER	0.04746	0.023698
Promedio	0.04558	0.02019
σ	0.01482	0.01365

Tabla 5.6. Resultados del análisis de compacidad para imágenes de pacientes con lesiones benignas.

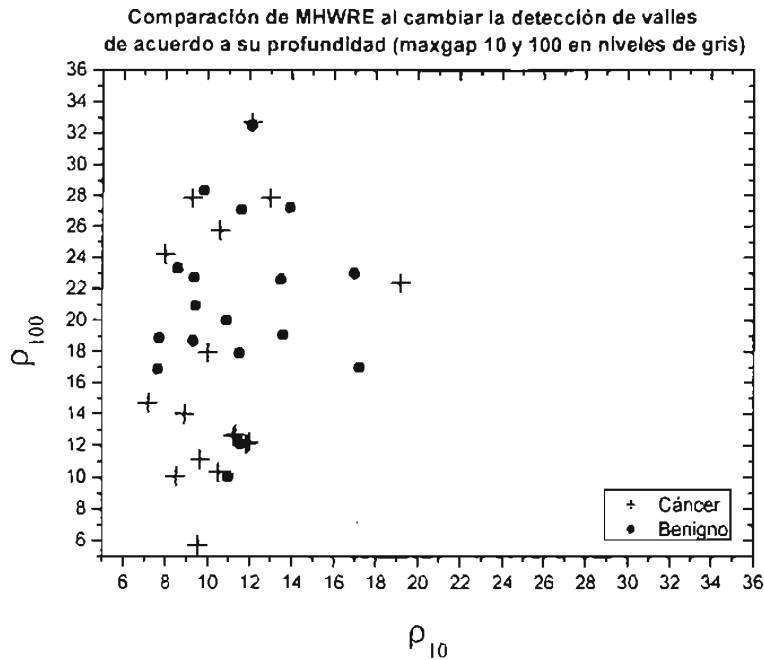
Imagen	κ	σ_{κ}
PA2_DER	0.04419	0.009146
PA11_IZQ	0.04878	0.024397
PA12_DER	0.06793	0.028574
PA13_DER	0.07940	0.011633
PA14_DER	0.03936	0.016106
PA15_DER	0.02707	0.006543
PA16_IZQ	0.01904	0.010464
PA17_DER	0.12188	0.023740
PA18_IZQ	0.01827	0.003464
PA19_IZQ	0.07485	0.055946
PA21_IZQ	0.01198	0.002048
PA23_IZQ	0.02930	0.007992
PA24_IZQ	0.03784	0.007092
PA26_DER	0.08354	0.005871
PA28_IZQ	0.01408	0.004318
PA39_DER	0.01964	0.006698
PA45_IZQ	0.03586	0.012579
PA50_IZQ	0.02878	0.006960
Promedio	0.04454	0.01353
σ	0.02979	0.01300

5.3. ANÁLISIS DE LA FASE DE ENTRENAMIENTO

Con los resultados obtenidos de la sección anterior al aplicar los scripts de ρ y κ , se elaboran las gráficas que se presentan a continuación y se discute cada una de ellas.

5.3.A. Resultado con el cociente promedio alto/ancho de extremos (ρ)

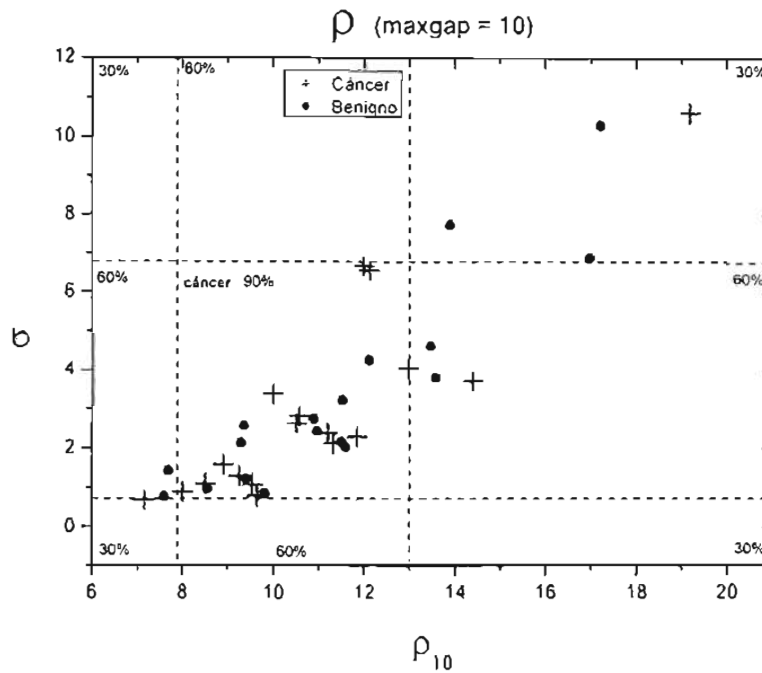
La gráfica 5.1 compara las mediciones del ρ , para cada estudio, extraídas usando los parámetros ρ_{10} y ρ_{100} , comentados anteriormente en la sección 5.2.A.



Gráfica 5.1. Comparación de los parámetros maxgap10 (ρ_{10}) y maxgap100 (ρ_{100}) para el script de ρ .

Tal como se observa en la gráfica 5.1 los puntos correspondientes tanto para las imágenes de pacientes con cáncer como con lesiones benignas, no caen completamente en la diagonal, ni tienen un comportamiento (progresión) similar, lo que implica que cada parámetro captura características texturales diferentes de la imagen. ρ_{10} captura características más suaves en la imagen mientras que ρ_{100} extrae las más rugosas. Esta gráfica representa así la justificación de la elección de los dos parámetros ρ_{10} (bajo maxgap) y ρ_{100} (alto maxgap).

En las gráficas 5.2 y 5.3 se presentan los resultados del script de ρ empleando los parámetros ρ_{10} y ρ_{100} , respectivamente. El tipo de gráficas que se presentan muestran la desviación estándar como función de ρ , es decir, nos indican la variación de la textura para cada grupo de imágenes analizadas. Debido a la naturaleza del análisis del ρ , éste analiza un cierto número de líneas de la mama entregando al final de cada una el valor promedio unidimensional. De esta forma, el tipo de gráfica de desviación estándar contra valor promedio nos indica la variación del parámetro en la imagen y es posible observar grupos de imágenes para las cuales la textura es muy similar en toda la mama (grupos de puntos o “nubes de puntos” muy cercanos entre sí con desviación estándar pequeña), o bien, grupos para los cuales existe una gran diferencia de ρ en la mama, determinada por una mayor desviación estándar.

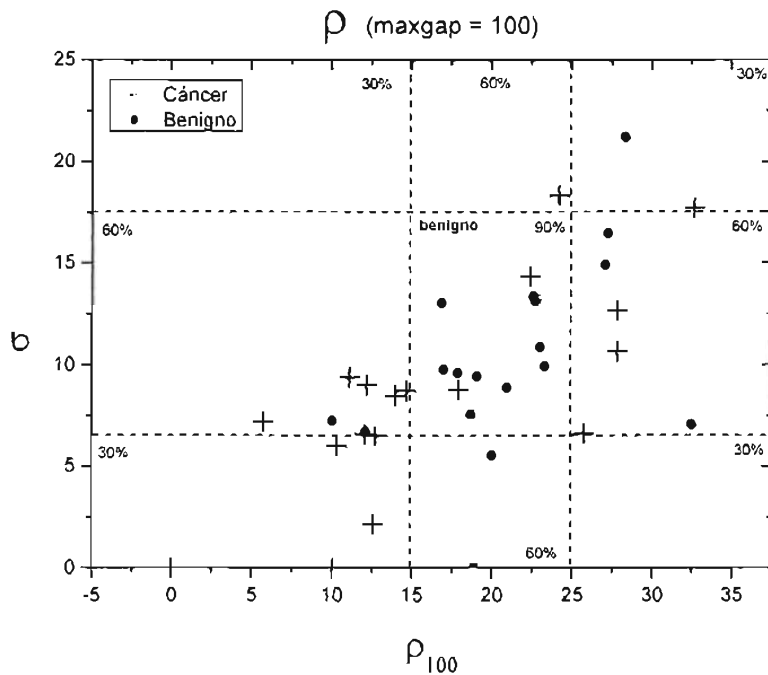


Gráfica 5.2. $\sigma_{\rho_{10}}$ como función de ρ_{10} , para imágenes del grupo de estudio y control en la fase de entrenamiento.

En la gráfica 5.2 se observan los agrupamientos de las imágenes de pacientes con cáncer y con lesiones benignas al comparar la desviación estándar del parámetro ρ_{10} con el valor promedio del ρ_{10} de cada imagen. Tratando de ofrecer una diferenciación simple de los grupos, se encerró en un recuadro un grupo característico de imágenes. En este caso, el grupo central (o característico) es el referente a imágenes de pacientes con cáncer, resultando los intervalos de $[8, 13]$ para ρ_{10} y de $[0.7, 6.8]$ para su desviación estándar. Al trazar estos intervalos en la gráfica resulta una división en 9 secciones de la misma, a las cuales se les asigna una calificación que trate de reflejar la probabilidad de que los grupos de imágenes de estas secciones se parezcan a la sección central. De esta forma, para la intersección de los intervalos del ρ_{10} y su desviación estándar, se le asigna de forma intuitiva (i.e. la sección central debe ser altamente significativa de malignidad y a la vez dejar una probabilidad de contener imágenes con lesiones benignas), la calificación de 0.9, i.e. 90% de probabilidad de que las imágenes en esa sección sean de pacientes con cáncer. Aquellas secciones que no sean el resultado de la intersección de al menos un intervalo de ρ_{10} $[8,13]$ o su desviación estándar $[0.7,6.8]$, se les asigna la calificación de 0.3, i.e., 30% de probabilidad de que las imágenes contenidas en esas secciones sean de pacientes con cáncer (poca probabilidad de malignidad y gran probabilidad de contener imágenes de pacientes con lesiones benignas). Al resto de las secciones se les asigna la calificación de 0.6, i.e. 60% de probabilidad de que las imágenes contenidas en esas secciones sean de pacientes con cáncer (aproximadamente misma probabilidad de contener imágenes de ambos grupos, pero ésta calificación está ligeramente más orientada a considerar que estas secciones contengan más imágenes de pacientes con cáncer).

Lo que observamos de la gráfica 5.2, es que los grupos de imágenes de pacientes con cáncer y de lesiones benignas se traslapan en la sección central. Esto implica que el

parámetro textural ρ_{10} no capta una diferencia significativa en la textura suave entre ambos grupos de imágenes y que en general, en ambos grupos es muy similar la variación de la textura (suave) en la mama junto con su valor promedio de ρ_{10} .



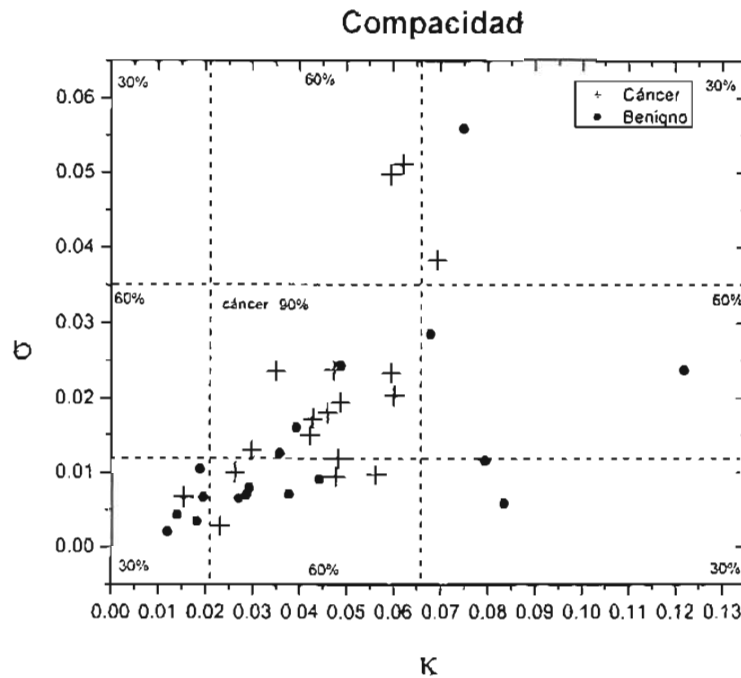
Gráfica 5.3. $\sigma_{\rho_{100}}$ como función de ρ_{100} , para imágenes del grupo de estudio y control en la fase de entrenamiento.

En la gráfica 5.3 se observan los agrupamientos de las imágenes de pacientes con cáncer y con lesiones benignas al comparar la desviación estándar del parámetro ρ_{100} contra el valor promedio del ρ_{100} de la imagen. Procediendo de igual forma que en el caso de la gráfica 5.2, obtenemos que en este caso la sección central $[15,25]$ para ρ_{100} y $[6.5,17.5]$ para su desviación estándar, tiene una calificación de 0.9, i.e. 90% de probabilidad de que las imágenes sean de pacientes con lesiones benignas. Las secciones restantes que se observan en la gráfica 5.3 tienen una calificación similar a la definida en la gráfica 5.2, pero ahora la calificación hace referencia a la probabilidad de que las imágenes sean de pacientes con lesiones benignas. Lo que observamos de la gráfica 5.3, es que existe una mejor separación de los grupos de imágenes de pacientes con cáncer de las de lesiones benignas en la sección central. Esto conlleva a una mejor separación por textura rugosa entre ambos grupos de imágenes que la observada por características suaves en la textura, presentada en la gráfica 5.2.

5.3.B. Resultados con la compacidad (κ)

En la gráfica 5.4 se presentan los resultados del script de κ . Se repite el mismo proceso de división por secciones que en las gráficas anteriores, resultando los intervalos de $[0.0210,0.0659]$, para κ y de $[0.0119,0.0351]$ para su desviación estándar. En este caso la

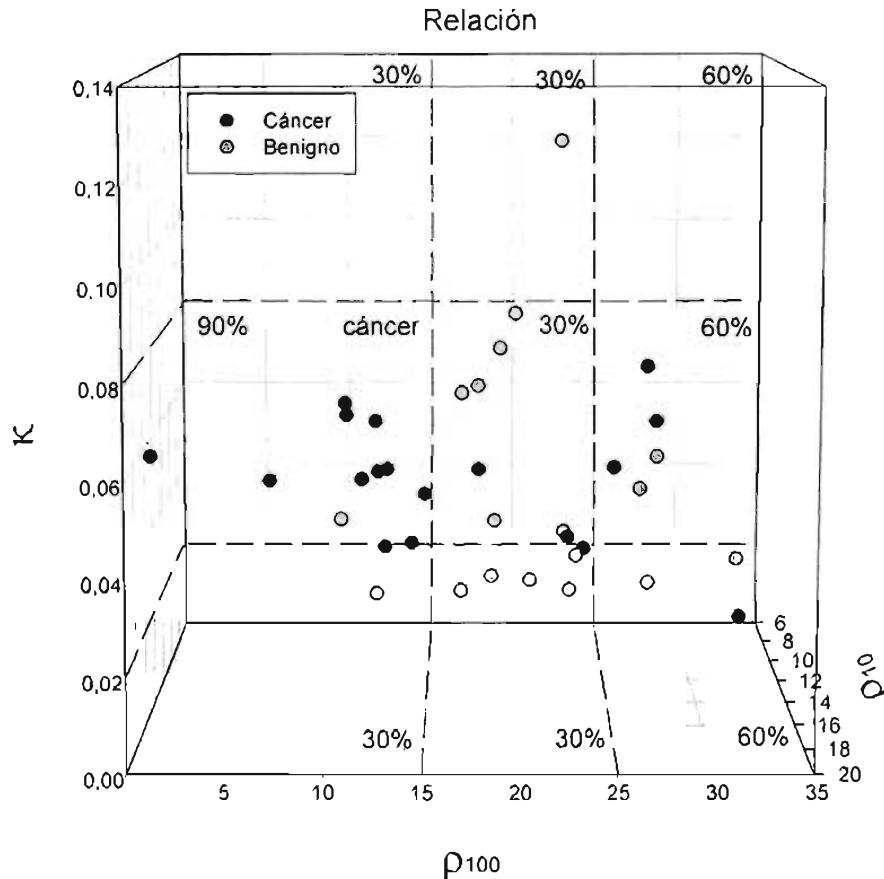
sección central tiene una calificación de 0.9, i.e. 90% de probabilidad de que las imágenes sean de pacientes con cáncer.



Gráfica 5.4. σ_{κ} como función de κ , para imágenes del grupo de estudio y control en la fase de entrenamiento.

De la gráfica 5.4, tenemos que los valores de compacidad para los grupos de imágenes de pacientes con cáncer y de lesiones benignas se traslapan para valores bajos de la compacidad. A partir de $\kappa \approx 0.07$, se observan sólo imágenes de pacientes con lesiones benignas, lo cual confirma un poco lo esperado en la sección 4.3.B, donde se estimó que las estructuras provenientes de las lesiones benignas podrían tener valores de compacidad mayores que en el caso de las estructuras malignas.

Los resultados de los parámetros ρ_{10} , ρ_{100} y κ , se presentan en un espacio tridimensional en la gráfica 5.5, a fin de poder visualizar la relación que guardan éstos con los grupos de imágenes analizados. La gráfica 5.5 presenta el mejor ángulo de vista que se visualizó, para ofrecer una discriminación simple entre los grupos de imágenes. Se observó que el parámetro ρ_{10} no ayudaba a la discriminación simple (a través de intervalos lineales), de los grupos de imágenes por lo que se da preferencia a los otros parámetros en la visualización de la gráfica 5.5 (ver gráfica 5.1, para apreciar la proyección bidimensional de la gráfica 5.5). Tal como puede observarse, en la parte izquierda del gráfico se encuentran concentradas en mayor proporción las imágenes de pacientes con cáncer, en la parte central se encuentran las imágenes de pacientes con lesiones benignas, y en la parte derecha la misma proporción de imágenes de ambos grupos.



Gráfica 5.5. Relación de los parámetros, ρ_{10} , ρ_{100} y κ en un espacio tridimensional para las imágenes del grupo de estudio y control en la fase de entrenamiento.

Debido a la disposición de los grupos de imágenes observada en la gráfica 5.5 se realiza la siguiente división por secciones (todas las calificaciones representan una probabilidad estimada de que las imágenes sean de pacientes con cáncer):

$$\kappa=[0.02,0.8] \text{ y } \rho_{10}=[0,15] \rightarrow \text{calificación } 0.9.$$

$$\kappa=[0.02,0.8] \text{ y } \rho_{100}=(15,25] \rightarrow \text{calificación } 0.3.$$

$$\kappa=[0,0.02) \cup (0.08,1] \text{ y } \rho_{100}=[0,25) \rightarrow \text{calificación } 0.3.$$

$$\rho_{100}=(25,50] \rightarrow \text{calificación } 0.6.$$

Una vez evaluadas las gráficas anteriores (5.1 a 5.5), se decide no utilizar las calificaciones de la gráfica 5.2, pues en la sección central de ésta los grupos de imágenes se traslapan demasiado. De esta forma, se toman en cuenta solamente las calificaciones propuestas de las gráficas 5.3, 5.4(calificación \times 2) y 5.5(calificación \times 2), luego entonces se realiza un programa general que utiliza los scripts creados anteriormente (ρ y κ) y se le alimenta con las calificaciones obtenidas en cada gráfica. El programa así formado deberá ofrecer al final una sugerencia de la naturaleza de la imagen analizada, ya sea de cáncer o de lesión benigna.

5.4. FASE DE VALIDACIÓN

La fase de validación corresponde a la evaluación del programa general que se creó en la fase anterior para tratar de diferenciar entre las mamografías de pacientes con cáncer de las de lesiones benignas. Se remarca el hecho de que esta fase de validación se realiza por un evaluador que no tiene conocimiento de mamografía ni de los resultados histopatológicos de las pacientes. Una vez que el evaluador obtuvo los resultados sugeridos por el programa en cuanto a una imagen característica de cáncer o de lesión benigna, éstos se comparan por el autor del método y del programa con los resultados histopatológicos de las pacientes

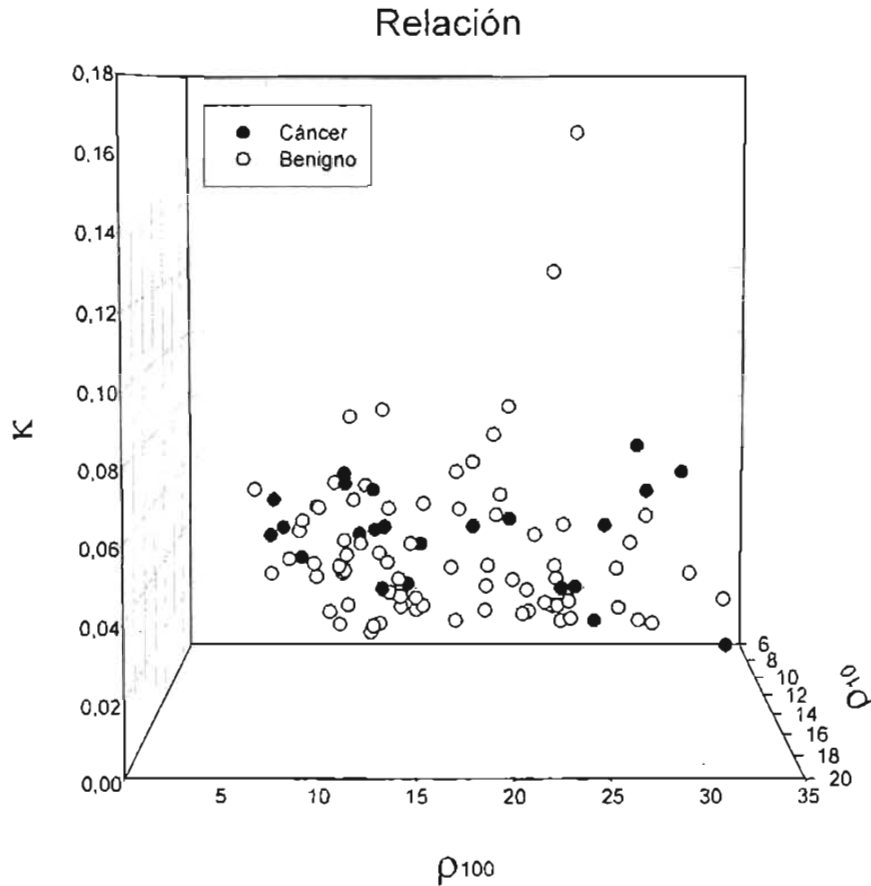
La diferenciación que ofrece el programa general es a través de la evaluación de cada imagen por medio de un sistema de puntuación. Es decir, mientras más alta sea la calificación de la imagen, más probable será que ésta sea característica de cáncer. Con el sistema de asignación de calificaciones descrito en la sección 5.3, la calificación más alta es 4.5 (90% de probabilidad de cáncer) y la más baja 1.5 (30% de probabilidad de cáncer). Calificaciones intermedias alrededor de 50% evalúan la imagen como indeterminada, i.e., no se ofrece sugerencia de cáncer o de lesión benigna.

En la gráfica 5.6 se presentan los resultados obtenidos de los scripts para calcular y desplegar ρ y κ en un espacio tridimensional, para todas las mamografías digitales de este estudio (i.e. 102 imágenes).

Comparando las gráficas 5.5 y 5.6, éstas se observan muy similares, salvo que en la gráfica 5.6 el grupo de las imágenes de pacientes con lesiones benignas se incrementó relativo al de pacientes con cáncer.

En la tabla 5.7 se presentan los resultados obtenidos por el autor de los scripts (ρ y κ), al comparar las sugerencias del programa general (usado por el evaluador sin conocimiento de los estudios histopatológicos de las pacientes), con los resultados histopatológicos de las pacientes. En esta tabla se muestran los porcentajes de aserción, equivocación e indeterminación del programa general, tomando en cuenta por separado los grupos de imágenes de entrenamiento (36 imágenes) y los usados en la validación (66 imágenes).

En la fase de entrenamiento, la división realizada de los grupos de imágenes por el programa general, utilizando el sistema de calificación propuesto, alcanza una aserción de 61% en las 18 imágenes de pacientes con cáncer y un 39% para las imágenes de pacientes con lesiones benignas. Además, se dio mayor peso a que las calificaciones propuestas asignaran de manera correcta un mayor número de imágenes con cáncer. Por esta razón el porcentaje de aciertos para las imágenes de pacientes con lesiones benignas no es muy elevado. Además, se trató que la evaluación efectuada diera el menor número de asignaciones incorrectas en ambos grupos de imágenes. El número de imágenes sugeridas como indeterminadas, es elevado en ambos grupos (mayor en el caso de las lesiones benignas), debido al traslape de los grupos visualizado en las gráficas 5.3, 5.4 y 5.5.



Gráfica 5.6. Relación de los parámetros ρ_{10} , ρ_{100} y K , en un espacio tridimensional, para las 102 imágenes de las mamografías digitales en este estudio.

Tabla 5.7 Porcentajes de asignación correcta, incorrecta e indeterminada en las imágenes de pacientes con cáncer y lesiones benignas, para los grupos de entrenamiento y validación.

RESULTADOS porcentuales				
Asignación	Entrenamiento		Validación	
	Cáncer	Benigno	Cáncer	Benigno
Correcta	61	39	50	43
Incorrecta	11	17	17	32
Indeterminada	28	44	33	25

Fase de entrenamiento: 36 mamografías en total (18 cáncer y 18 benignas).

Fase de validación: 66 mamografías en total (6 cáncer y 60 benignas).

En la fase de validación los porcentajes de aciertos para las imágenes de cáncer disminuyen, mientras que las sugerencias incorrectas e indeterminadas que entrega el programa, aumentan en relación con la fase de entrenamiento. Sin embargo, no podemos asegurar que éste sea un resultado confiable pues sólo se contó con 6 imágenes más de pacientes con cáncer. En cuanto a las imágenes de pacientes con lesiones benignas, el porcentaje de sugerencias indeterminadas disminuye, mientras que el de aciertos aumenta. Sin embargo, el número de sugerencias incorrectas aumenta casi al doble que en la fase de

entrenamiento, debido principalmente a que las calificaciones propuestas fueron determinadas para asignar de manera correcta un mayor número de imágenes con cáncer.

En la tabla 5.8 se encuentra el resultado global de las 102 imágenes, al comparar las sugerencias del programa general con los estudios histopatológicos de las pacientes. Aquí se observa el comportamiento del programa general al emplear todas las imágenes disponibles, resultando en la asignación correcta de 58% de las 24 imágenes con cáncer y una asignación errónea del 13%. Además, al considerar las 78 imágenes de pacientes con lesiones benignas, se tiene que el porcentaje de aserción disminuye (42%) y el de equivocación aumenta (28%), pero la explicación se debe a que en la fase de entrenamiento se dio preferencia a la aserción de las imágenes de pacientes con cáncer, resultando una menor asignación correcta para las lesiones benignas, en las sugerencias del programa general.

Tabla 5.8. Porcentajes de asignación correcta, incorrecta e indeterminada en todas las imágenes de pacientes con cáncer y lesiones benignas.

RESULTADO GLOBAL porcentual		
Asignación	Cáncer	Benigno
Correcta	58	42
Incorrecta	13	28
Indeterminada	29	29

102 mamografías: 24 cánceres y 78 benignas.

Al analizar la tabla 5.8 obtenemos que considerando el grupo entero de imágenes (102) se tiene un porcentaje de asignación correcta (ya sea de cáncer o benigna) del 46%, asignación incorrecta del 25% y una asignación indeterminada del 29%. Es decir, que aproximadamente de cada 10 imágenes analizadas por el programa general, éste sugerirá un diagnóstico acertado, ya sea de cáncer o de lesión benigna, para 4 (ó 5) imágenes, un diagnóstico incorrecto para 3 (ó 2) imágenes y un diagnóstico indeterminado para 3 imágenes (igual probabilidad de que estas 3 imágenes sean características de cáncer o de lesión benigna).

Una forma de mejorar los resultados presentados anteriormente sería modificar la forma de análisis de los scripts de ρ y κ . En el presente trabajo de tesis se analizaron las mamografías digitales mediante un estudio global de la mama, para obtener un número textural y morfológico característico de la mama completa. Debido a esto, al obtener los valores promedios de los parámetros extraídos, sin tomar en cuenta la posición relativa de cada región en la mama, se perdió información que hubiera podido servir para evaluar de manera más precisa las características texturales y morfológicas de la mama. Sin embargo, si el análisis se realizara mediante un estudio regional, éste podría extraer de manera más precisa la textura y morfología (matemática) de la mama, ya que evitaría el aplanamiento de los resultados al obtener un promedio global de todas las mediciones realizadas. De esta forma, los scripts podrían ofrecer resultados más significativos para regiones particulares que fueran notablemente diferentes del resto de la mama.

Además, en cuanto a los scripts desarrollados: 1) para la caracterización de la rugosidad textural mediante ρ , podría complementarse al tomar en cuenta no sólo el alto y ancho de cada pico en cada perfil muestreado en la imagen, sino también el ancho neto a media altura del pico máximo (Full Width at Half Maximum, FWHM), lo que extraería diferencias entre picos angostos y anchos con el mismo valor de ρ , tal como se observa en la figura 5.5, derecha). La caracterización de forma mediante la compacidad κ , podría mejorarse al tomar en cuenta no sólo las variaciones morfológicas en la frontera externa de las estructuras obtenidas, si no también según el caso, las variaciones en la frontera interna de las estructuras (a través de la característica topológica de Euler, genus o número de Euler), lo que extraería diferencias entre estructuras sin agujeros (estructuras completamente radiopacas, sugestivas de cáncer) y estructuras con agujeros (centros radiolúcidos inmersos en estructuras radiopacas, sugestivas de benignidad).

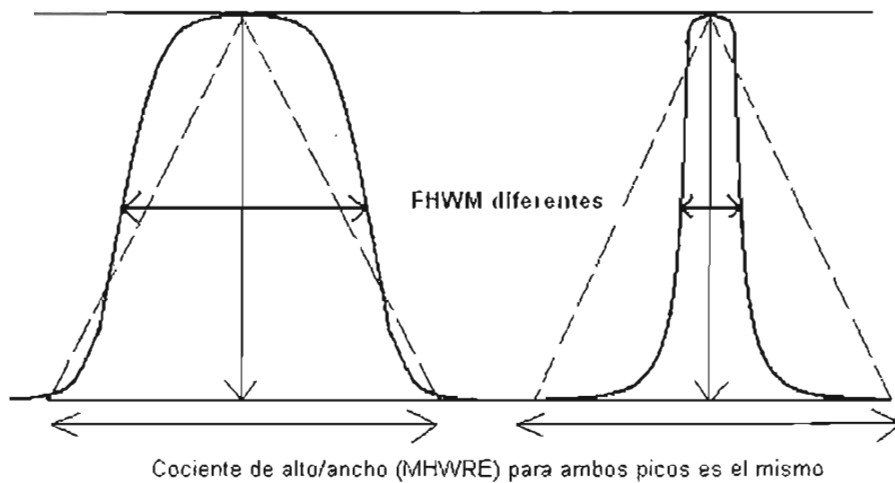


Figura 5.5. Ejemplo de mejora al análisis del script de ρ , para extraer más información y diferencias en los perfiles de intensidad muestreados en la imagen.

En resumen, el programa general desarrollado utilizando el parámetro textural (Cociente promedio alto/ancho de extremos) y el morfológico (Compacidad de estructuras), resulta en una primera aproximación a lo que debiera ser un programa automático de evaluación de imágenes mamográficas digitales. En la actualidad aún no se cuenta en el área de investigación con un programa que trate de evaluar de esta forma una mamografía digital, y menos aún en el ámbito comercial. Esto justifica que al programa se le considere una primera versión y que en futuros trabajos se le tome como base para mejorar la correcta diferenciación de los grupos de imágenes, utilizando mejoras de los parámetros empleados o incorporando otros nuevos. Además, se remarca la necesidad de contar con un mayor número de mamografías digitales de cáncer para la evaluación de los parámetros que se utilicen en el análisis.

CAPÍTULO VI CONCLUSIONES

El capítulo expone un breve comentario de los análisis de resultados y las conclusiones del trabajo de tesis.

6.1. CONCLUSIONES

Se programaron en Matlab las rutinas de análisis textural (ρ) y morfológico (κ) para evaluar 102 mamografías digitales resultando un porcentaje de 46% de resultados correctos (correcta asignación positiva o negativa de cáncer), 25% de resultados incorrectos (incorrecta asignación positiva o negativa de cáncer) y 29% de resultados indeterminados (el programa no realiza asignación positiva o negativa de cáncer). A continuación se presenta en detalle las conclusiones parciales de acuerdo a los rubros siguientes.

6.1.A. Validación del modo de lectura del formato DICOM de la imagen

Se validó la lectura de los archivos de imágenes RAW, en formato DICOM, al comparar los resultados de los histogramas de la región mamaria obtenidos en este trabajo de tesis y los presentados por el Dr. A. Burgess en el artículo [Bu04], encontrándose un buen acuerdo entre ambos. Una muestra de alrededor 50 mamografías digitales puede reproducir aproximadamente la función de distribución probabilística de los niveles de gris en la mamografía digital para el sistema de mastografía empleado.

6.1.B. Análisis de histogramas (caracterización en intensidades)

Se observó que una parte de los histogramas de la mama, contaban con un hombro (aproximadamente para niveles de gris 650-750) a la izquierda del pico máximo (aproximadamente para niveles de gris 1000). Este hombro y pico fueron asociados tentativamente a 2 diferentes tipos de poblaciones, la primera asociada a estructuras radiopacas y la segunda a una combinación de estructuras grasas y fibroglandulares más probables en la mama. Esta asociación, sirvió para determinar el nivel umbral característico de cada imagen para realizar los cortes de nivel en la rutina de compacidad de estructuras.

6.1.C. Parámetro textural del cociente promedio alto/ancho de extremos (ρ)

Se programó en Matlab la rutina deseada para analizar los perfiles unidimensionales de niveles de gris en la región mamaria. Se obtuvieron resultados más significativos al analizar los perfiles con niveles de gris mayores que 100 que aquellos con niveles mayores que 10. Al comparar los resultados promedios del ρ y su desviación estándar en los grupos de imágenes de pacientes con cáncer y con lesiones benignas, no se obtuvieron diferencias significativas entre ambos, $\langle \rho_{100} \rangle_{\text{cáncer}} = 17.30 \pm 7.87$ y $\langle \rho_{100} \rangle_{\text{benigno}} = 21.02 \pm 5.60$.

6.1.D. Compacidad de estructuras (κ)

Se programó en Matlab la rutina deseada para extraer la compacidad de las estructuras obtenidas al realizar cortes de nivel a la representación tridimensional de la densidad mamográfica (coordenada (i,j) como coordenada del pixel y coordenada k , como valor de gris del pixel). Al comparar los resultados promedios de la compacidad ($\kappa_{\text{cáncer}}=0.04558\pm 0.01482$ y $\kappa_{\text{benigno}}=0.04454\pm 0.2979$), para las imágenes de pacientes con cáncer y con lesiones benignas, sólo se observan diferencias significativas en las desviaciones estándar de cada grupo de imágenes: $\sigma_{\kappa_{\text{cáncer}}}=0.02019\pm 0.01365$ y $\sigma_{\kappa_{\text{benigno}}}=0.01353\pm 0.01300$.

6.1.E. Discriminación entre los grupos de imágenes de pacientes con cáncer y con lesiones benignas

Se implementó un sistema de calificaciones óptimo para cada parámetro extraído de la imagen, para que por medio de una evaluación de puntaje, se ofreciera una sugerencia de malignidad o benignidad a la imagen analizada. Para el caso de las 24 imágenes de pacientes con cáncer, el programa da una sugerencia acertada de malignidad para un 58% de las imágenes, una sugerencia incorrecta (sugiere benignidad) para un 13% de las imágenes y no se ofrece ninguna sugerencia (evaluación indeterminada) para un 29% de las imágenes. Para el caso de las 78 imágenes de pacientes con lesiones benignas el programa da una sugerencia acertada de benignidad para un 42% de las imágenes, una sugerencia incorrecta (sugiere malignidad) para un 28% de las imágenes y no se ofrece ninguna sugerencia para un 29% de las imágenes.

Finalmente al considerar a todas las 102 imágenes como un solo grupo, el programa sugiere un diagnóstico acertado (cáncer o lesión benigna) para el 46% de las imágenes, un diagnóstico incorrecto para el 25% de las imágenes y una asignación indeterminada para el 29% restante. Es decir, que aproximadamente de cada 10 imágenes analizadas por el programa general, éste sugerirá un diagnóstico acertado, ya sea de cáncer o de lesión benigna, para 4 (ó 5) imágenes, un diagnóstico incorrecto para 3 (ó 2) imágenes y un diagnóstico indeterminado para 3 imágenes (igual probabilidad de que estas 3 imágenes sean características de cáncer o de lesión benigna).

En conclusión, el presente trabajo de tesis ofrece un método y un programa para computadora personal que evalúa las características texturales (por medio de ρ) y morfológicas de la mama (por medio de κ), para tratar de diferenciar en una primera aproximación grupos de imágenes de pacientes con cáncer, de otro grupo de imágenes de pacientes con lesiones benignas. El resultado final de esta diferenciación no es utilizable en la práctica debido al bajo porcentaje de aserción correcta. Sin embargo este trabajo de tesis resulta pionero en el enfoque en que se analizan las mamografías digitales, que no se había empleado hasta el momento. Por lo tanto esta tesis aporta un inicio en la investigación de nuevos métodos y criterios texturales o morfológicos que desemboquen en un mejor porcentaje de diferenciación efectivo entre las imágenes de pacientes con cáncer y las que no tienen.

6.2. PERSPECTIVAS A FUTURO

En primer lugar, se sugiere la modificación de la rutina textural para que ésta tome en cuenta el valor del ancho neto a media altura del pico máximo (FWHM), para discriminar entre estructuras con igual ρ , pero con características visuales rugosas (o suaves) diferentes, como se observa en la figura 5.5. Igualmente, modificar la rutina de compacidad para que ésta discrimine entre estructuras puramente radiopacas (sugestivas de cáncer) y estructuras radiolúcidas inmersas en un medio radiopaco (sugestivas de benignidad). Estas modificaciones extraerían más información (o características), correspondiente a la textura y morfología de cada imagen, para ofrecer una mejor diferenciación entre los grupos de imágenes (estudio y control).

Para mejorar el porcentaje de diferenciación efectiva entre los grupos de imágenes es necesario contar con más mamografías digitales a fin de obtener una mejor estadística de trabajo. Así mismo, es imprescindible contar con un mayor número de imágenes de pacientes con cáncer ya que la fracción de éstas en nuestro estudio fue de aproximadamente un tercio de las pacientes que no tienen cáncer.

Finalmente, para mejorar las rutinas desarrolladas en el presente trabajo de tesis, se sugieren los métodos citados en la sección 5.4. Para tal efecto se proponen y prueban otros parámetros que se sumen a los anteriores para ayudar a una mejor discriminación de las imágenes. Además, una vez que se obtengan los resultados de los parámetros, proponer una metodología de análisis comparativo de éstos distinta a la discriminación lineal, para la caracterización de los grupos de resultados, como por ejemplo una clasificación en el espacio de parámetros con métodos de reconocimiento de patrones no-lineales tales como las redes neuronales. Al aumentar el número de parámetros, se deberá proceder a identificar combinaciones “ortogonales” o altamente descorrelacionadas de los parámetros que resulten más significativos, mediante un *análisis de componentes principales*.

REFERENCIAS

[AC93]

American College of Radiology. Breast Imaging-Reporting and Data System (BI-RADS), American College of Radiology, Reston, VA, 1993.

[AC98]

American College of Radiology. Breast Imaging-Reporting and Data System (BI-RADS), 3rd ed. (American College of Radiology, Reston, VA, 1998).

[Ad87]

D.D. Adler, M.Rebner, D.R. Pennes, *Accesory breast tissue in the axilla: Mamographic appearance*, Radiology, 1987; **163**:709-711.

[Ad94]

R. Adams, and L. Bischof, *Seeded region growing*, IEEE Trasn. On Patter Analysis and Machine Intelligence, 1994; **16(6)**:641-647.

[Al94]

D.G. Altman, J.M Bland, *Statistics Notes: Diagnostic tests I: sensitivity and specificity*, BMJ 1994; **308(6943)**:1552.

[At86]

F.H. Attix, *Introduction to radiological physics and radiation dosimetry*, ISBN: 0-471-01146-0, John Wiley & Sons, Inc., cap 3,7,8, 1986.

[Bl00]

K.I. Bland, E.M. Copeland, *La mama: manejo multidisciplinario de las enfermedades benignas y malignas, tomo I*, edit. Médica panamericana; cap. 2 y 35, 2000

[Be97]

D. Betal, N. Roberts, G:H: Whitehouse, *Segmentation and numerical analysis of microcalcifications using mathematical morphology*, Brit.J.Radiol., 1997; **70(837)**:903-917.

[Bo95]

N.F. Boyd, J.W. Byng, R.A. Jong, E.K. Fishell, L.E. Little, A.B. Miller, G.A. Lockwood, D.L. Trichler, and M.J. Yaffe, *Quantitative classification of mammographic densities and breast cancer risk: results from Canadian national breast screening study*, J.Nat.Cancer Inst., 1995; **87**:670-675.

[Bo96]

N.F. Boyd, C. Greenberg, G. Lockwood, et al., *The effects at two years of a low-fat high-carbohydrate diet on radiological features of the breast: results from a randomized trial*. J. Natl. Cancer Prevent, 1996 Apr 2; **5**: 281-286.

[Bo98]

J.M. Boone, K.K. Lindfors, C.S. Beatty, and J.A. Slebert, *A breast density index for digital mammograms based on radiologist ranking*, J.Digital Imag., 1998; **11**:101-115.

[Bo99]

N.F. Boyd, G.A. Lockwood, L.J. Martin, J.A. Knight, R.A. Jong, E. Fishell, J.W. Byng, M.J. Yaffe, D.L. Tritchler, *Mammographic densities and risk of breast cancer among subjects with a family history of this disease*, J. Natl Cancer Inst. 1999 Aug 18; **91(16)**: 1404-1408.

[Br89]

R.J. Brenner, E.A. Sickles, *Acceptability of periodic follow-up as an alternative to biopsy for mammographically detected lesions interpreted as probably benign*, Radiology, 1989; **171**:645-646.

[Br99]

G. Brake and N. Karssemeijer, *Single and multi-scale detection of masses in digital mammograms*, IEEE Trans.Med.Imaging, 1999; **18**:628-639.

[Br00]

G.M. te Brake, N. Karssemeijer, and H.C.L. Hendriks, *An automatic method to discriminate malignant masses from normal tissue in digital mammograms*, Phys.Med.Biol., 2000; **45**:2843-2857.

[Br04]

M.E. Brandan, C. Ruiz-Trejo, M. Verdejo-Silvia, M. Guevara, H. Lozano-Zalce, L. Madero-Preciado, J. Martín, L.M. Noel-Etienne, J.L. Ramírez-Arias, J. Soto, and Y. Villaseñor, *Evaluation of equipment performance, patient dose, imaging quality, and diagnostic coincidence in five Mexico city Mammography Services*, Archives of Medical Research, 2004; **35**:24-30.

[Bu04]

A. Burgess, *On the noise variation of a digital mammography system*, Med.Phys., 2004; **31(7)**:1987-1995.

[By94]

J.W. Byng, N.F. Boyd, E. Fishell, R.A. Jong, and M.J. Yaffe, *The quantitative analysis of mammographic densities*, Phys.Med.Biol., 1994; **39**:1629-1638.

[By96]

J.W. Byng, N.F. Boyd, L. Little, G. Lockwood, E. Fishell, R.A. Jong, and M.J. Yaffe, *Symmetry of projection in the quantitative analysis of mammographic images*, Eur.J.Cancer Prevent., 1996; **5**:319-327.

[By96*]

J.W. Byng, N.F. Boyd, E. Fishell, R.A. Jong, and M.J. Yaffe, *Automated analysis of Mammographic densities*, Phys.Med.Biol., 1996; **41**:909-923.

[By97]

J.W. Byng, J.P. Critten, M.J. Yaffe, *Thickness-equalization processing for mammographic images*, Radiol.; 1997, **203**: 564-658.

[Co98]

G. Corkidi, L. Vega, J. Márquez, E. Rojas, P. Ostrovsky-Wegman. *A Roughness Feature of Metaphase Chromosome Spreads and Nuclei for Automated Cell Proliferation Analysis*. *Medical and Biological Engineering and Computing*, 1998; **36(6)**:679-685.

[Ca90]

B.C. Caldwell, S.J. Stapleton, D.W. Holdsworth, R.A. Jong, W.J. Weisser, G.C. Cooke, and M.J. Yaffe, *Characterisation of mammographic parenchymal pattern by fractal dimension*, Phys.Med.Biol., 1990; **35**:235-247.

[Ch98]

H.P. Chan, S.C.B. Lo, B. Sahiner, K.L. Lam, M.M. Helvie, M.M. Goodsitt, and D.D. Adler, *Computerized analysis of mammographic microcalcifications in morphological and texture feature space*, Med.Phys., 1998; **25(10)**:2007-2019.

[Ch00]

R. Chandrasekhar and Y. Attikiouzel, *New range-based neighbourhood operator for extracting edge and texture information from mammograms for subsequent image segmentation and analysis*. Medical Imaging Processing, IEEE Proc.-Sci Meas. Technol., 2000 Nov 6; **147(6)**: 408-413.

[DI03]

Digital Imaging and Communications in Medicine (DICOM), *Part 1: Introduction and Overview*, Published by National Electrical Manufacturers Associations, 1300N. 17th street, Rosslyn, Virginia 22209, USA, 2003.

[Do93]

C.J. D'Orsi, D.B. Kopans, *Mammographic feature analysis*, Semin. Roentgenol, 1993; **28**:204-230.

[Eg88]

R.L. Egan, *Breast imaging: diagnosis and morphology of breast disease*, ISBN: 0-7216-2320-4, W.B. Saunders, Philadelphia, 1988.

[Ga96]

P. Gamagami, *Atlas of mammography: new early signs in breast cancer*, ISBN: 0-8654-2481-0, Blackwell Science, Cambridge, Massachusettes, 1996.

[GE00]

Operator manual of Senographe 2000D acquisitions system, General Electric, Medical Systems; chap. 2,3,5,10,11; 2000.

[Gl93]

B.S. Gluck, D.D. Dershaw, L. Liberman, B.M. Deutch, *microcalcifications on postoperative mammograms as an indicator of adequacy of tumours excision*, Radiology, 1993; **188**:469-472.

[Gu95]

R. Gupta, and P.E. Udndrill, *The use of texture analysis to delineate suspicious masses in mammography*, Phys.Med.Biol., 1995; **40(5)**:835-855.

[Ha73]

R. M. Haralick and K. Shanmugam and I. Dinstein. Textural features for image classification. IEEE Trans. Syst., Man, Cybern. 1973, pp. 610-621.

[Ha88]

F.M. Hall, J.M. Storella, D.Z. Silverstone, G. Wyshak, *Nonpalpable breast lesions: Recommendations for biopsy based on suspicion of carcinoma at mammography*, Radiology, 1988; **167**:353-358.

[Hi96]

R. Highnam, M. Brady, and B. Shepstone, *A representation for mammographic image processing*, Med.Image Anal, 1996; **1**:1-18.

[Hu95]

Z. Huo, M.L. Giger, C.J. Vyborny, U. Bick, P. Lu, D.E. Wolverton, and R.A. Schmidt, *Analysis of spiculation in the computerised classification of mammographic masses*, Med. Phys., 1995; **22(10)**:1569–1579.

[Hu98]

Z. Huo, M.L. Giger, O.I. Olopade, and S.A. Cummings, *Computerized analysis of parenchymal patterns for the assessment of breast cancer risk*, Suppl.Radiology, RSNA, 1998; **209**:354.

[IMJ]

W. Rasband, National Institute of Health, USA, *ImageJ 1.33m*.

<http://rsb.info.nih.gov/ij/> (Java 1.4.1_02)

[Ja88]

A.K. Jain, and R.C. Dubes, *Algorithms for clustering data*, ISBN: 0-13022-2278-x, Prentice Hall, Englewood Cliffs, N.J., 1988.

[Ja91]

V.P. Jackson, L.W. Bassett, R.H. Gold, H.E. Reynolds, *diagnostic importance of radiographic density of noncalcified breast masses: Analysis of 91 lesions*, AJR AQm. J. Roentgenol, 1991; **157**:25-28.

[Ja02]

S.S. Jasjit, S. Kamaledin-Setarehdan and Sameer Singh (eds), *Advances in Pattern Recognition: Advanced Algorithmic Approaches to Medical Image Segmentation (State-of-the-Art Applications in Cardiology, Neurology, Mammography and Pathology)*, ISBN: 1-85233-389-8, Springer-Verlag London, chap. 9, 2002.

[Je99]

Kyung-Wook Jee, L. E. Antonuk, El-Mohri Youcef and Q. Zhao, *System performance of a prototype flat-panel imager operated under mammographic conditions*, Med. Phys., 2003; **30(7)**: 1874-1890.

[Ka98]

N. Karssemeijer, *Automated classification on parenchymal patterns in mammograms*, Phys.Med.Biol., 1998; **43**:365-378.

[Ke93]

W. P. Kegelmeyer Jr., *Evaluation of stellate lesion detection in a standard mammogram data set*, Int. J. Pattern Recogn. Artificial Intell., 1993; **7(12)**:1477–1493.

[Ke94]

W.P. Kegelmeger, J.M. Pruneda, P.D. Bourland, A. Hillis, M.W. Riggs, and M.L. Nipper, *Computer-aided mammographic screening for spiculated lesions*, Radiology, 1994; **191**:331-337.

[Ki81]

L. Kitchen, and A. Rosenfeld, *Edge evaluation using local edge coherence*, IEEE Trans. on Systems, Man and Cybernetics, Vol. 11, no.9, pp. 597-605, 1981.

[Ki00]

I. Kirby, M. Copeland, *La Mama, manejo multidisciplinario de las enfermedades benignas y malignas*, 2ª. Edición, Tomo I, ISBN: 950-06-5076-2, edit. Médica Panamericana, 2000.

[Ko98]

D. Kopans, *Breast imaging*, ISBN: 0-3975-1302-X, Lippincott-Raven, Philadelphia; 1998.

[La86]

M. Lanyi, *Pathogenesis, pathophysiology, and composition of breast calcification. In diagnosis and differential diagnosis of breast calcifications*, Springer-Verlag; 1986:87.

[Mc94]

M.M. McNicholas, J.P. Heneghan, M.H. Milner, et al, *Pain and increased mammographic density in postmenopausal women*, AJR Am. J. Roentgenol, 1994; **163**:311-315.

[Mu01]

N.R. Mudigonda, R.M. Rangayyan, *Fellow, IEEE*, and J.E. Leo-Desautels, *Detection of Breast Masses in Mammograms by Density Slicing and Texture Flow-Field Analysis*, IEEE Transactions on Medical imaging, 2001; **20(2)**:1215-1227.

[Ng92]

S.L. Ng, and W.F. Bischof, *Automated detection and classification of breast tumours*, Computers and Biomedical Research, 1992; **25(3)**:218-237.

[Pi03]

S. Pita-Fernández, S. Pértegas-Díaz,, *Pruebas diagnósticas*, Cad Aten. Primaria 2003; **10**:120-124.

(http://www.fisterra.com/mbe/investiga/pruebas_diagnosticas/pruebas_diagnosticas.htm#Sensibilidad).

[Po97]

W.E. Polakowski, D.A. Cournoyer, S.K. Rogers, M.P. DeSimio, D.W. Ruck, J.W. Hoffmeister, and R.A. Raines, *computer-aided breast cancer detection and diagnoses of masses using difference of Gaussians and derivative-based feature saliency*, IEEE Trans.Med.Imaging; 1997; **16**:811-819.

[PHO]

Photodiode technology

<http://www.centrovision.com/tech2.htm>

[Qi99]

W. Qian, L. Li, and L.P. Clarke, *Image feature extraction for mass detection in digital mammography: influence of wavelet analysis*, Med.Phys., 1999; **26(3)**:402-408.

[Qu01]

S. Quadres, A. Sacristán, *Automated extraction of microcalcifications BI-RADS numbers in mammograms*, IEEE, 2001; **0-7803-6725-1**:289-292.

[Ra97]

R.M. Rangayyan, N.M. El-Faramawy, J.E. Leo Desautels, and O.A. Alim, *Measures of acutance and shape for classification of breast tumors*, IEEE Trans. On Med.Imag., 1997; **16(6)**:799-810.

[Sa91]

A.F. Saftlas, R.N. Hoover, et al., *Mammographic densities and risk of breast cancer*, Cancer, 1991; **67**:2833-2838.

[Sa96]

B. Sahiner, H.P. Chan, N. Petrick, D. Wei, M.A. Helvie, D.D. Adler and M.M. Goodsitt, *Classification of mass and normal breast tissue: a convolution neural network classifier with spatial domain and texture images*, IEEE Trans.Med.Imaging, 1996; **15**:598-610.

[Sa01]

P.K. Saha, J.K. Udupa, E.F. Conant, D.P. Chakraborty, D. Sullivan, *Breast tissue density quantification via digitized mammograms*, IEEE Trans Med. Imaging 2001 Aug; **20(8)**: 792-803.

[Ta03]

A Taibi, S Fabbri, P Baldelli, C di Maggio, G Gennaro, M Marziani, A Tuffanelli and M Gambaccini, *Dual-energy imaging in full-field digital mammography: a phantom study*, Phys. Med. Biol. 2003; **48**:1945–1956.

[Ur98]

G. Ursin, A. Melvin, A. Astrahan, M. Salane, Y.R. Parisky, J.G. Pearce, J.R. Daniels, M.C. Pike, and D.V. Spincer, *The detection of changes in mammographic densities*, Cancer Epidemiol. Biomarkers Prevent, 1998; **7**:43-47.

[Va98]

I. Valdivia Bernstein, *Resultados falsos positivos en programas de detección de cáncer mamario: Resultados a 10 años*, Bol. Soc. Chil., 1998; **3(1)**:5-9.

[Ve99]

S. Vedantham, A. Karellas and S. Suryanarayanan, *Full breast digital mammography with an amorphous silicon-based flat panel detector: Physical characteristics of a clinical prototype*, Med. Phys., 1999; **27(3)**:558-567.

[Ya97]

J. Yaffe and J. A. Rowlands, *X-ray Detectors for Digital Radiography*, Phys. Med. Biol.; 1997, **42**: 1-39.

[Ya98]

M.J. Yaffe, N.F. Boyd, J.W. Byng, R.A. Jong, E. Fishell, G.A. Lockwood, L.E. Little, D.L. Tritchler, *Breast cancer risk and measured mammographic density*. Eur. J. Cancer Prev. 1998 Feb; **7 Suppl 1**: S47-55.

[Yi93]

F.F. Yin, M.L. Giger, C.J. Vyborny, K.Doi, and R.A. Schmidt, *Comparison of bilateral-substraction and single-image processing techniques in the computerized detection of mammographic masses*, Invest.Radiol., 1993; **6**:473-481.

[Wa92]

E. Warner, G. Lockwood, M. Math, D. Tritchler and N.F. Boyd. *The risk of breast cancer associated with mammographic parenchymal patterns: a meta-analysis of the published literature to examine the effect of method of classification*, Cancer Detection Prevention, 1992; **16**:67-72.

[WHO]

World Health Organisation (WHO), mortality database.

<http://www-depdb.iarc.fr/> y <http://www.who.int/whosis/>.

[Wh98]

E.White, P. Velentgas, M.T. Mandelson, C.D. Lehman, J.G. Elmore, P. Porter, Y. Yasui, and S.H. Taplin, *Variation in mammographic breast density by time in menstrual cycle among women aged 40-49 years*, J.Natl.Cancer Inst., 1998; **90**:906-910.

[WIK]

Wikipedia, the free-content encyclopedia.

<http://en.wikipedia.org/wiki/DICOM>, <http://en.wikipedia.org/wiki/Morphology>,

<http://en.wikipedia.org/wiki/Texture>.

[Wo76a]

J.N. Wolfe, *Breast patterns as an index of risk of developing breast cancer*. Am. J. Roentgenol, 1976; **126**: 1130-1139.

[Wo76b]

J.N. Wolfe, *Risk for Breast Cancer Development determined by mammographic parenchymal pattern*. Cancer, 1976; **37**: 2486-2492.

[Wo00]

W.H. Wolberg, *Benign breast disease and breast cancer tutorial*, <Http://www.surgery.wisc.edu/wolberg>, 2000.

[Zh01]

C. Zhou, H-P. Chan, N. Petrick, M.A. Helvie, M.M. Goodsitt, B. Sahiner and M. Hadjiiski, *Computerized image analysis: Estimation of breast density on mammograms*, Med. Phys.; 2001, **28**(6): 1056-1069.

ANEXOS

A. ASPECTOS CLÍNICOS DE LA IMAGEN
MAMOGRAFICA

B. ASPECTOS TÉCNICOS EN EL ANÁLISIS
DE IMÁGENES DIGITALES

C. OBTENCIÓN DE LA MASCARILLA
BINARIA

ANEXO A

ASPECTOS CLÍNICOS DE LA IMAGEN MAMOGRAFICA

ANEXO A

ASPECTOS CLÍNICOS DE LA IMAGEN MAMOGRÁFICA

A.1. ANATOMÍA DE LA MAMA.

La mama se compone principalmente de tejido fibroglandular y tejido adiposo. El tejido fibroglandular es una mezcla de tejido conectivo fibroso (estroma) y tejido glandular (células epiteliales y ductos lactíferos). El tejido adiposo se encuentra principalmente recubriendo los conductos galactóforos cerca de la dermis y como grasa intersticial entre ellos. La mama se fija a la piel por los ligamentos suspensorios de Cooper y es separada de la aponeurosis de revestimiento del músculo pectoral mayor por la bolsa retromamaria. Los ligamentos de Cooper forman tabiques fibrosos en la estroma que brindan apoyo al parénquima mamario. De 15 a 20 conductos galactóforos se extienden desde los lobulillos compuestos por epitelio glandular hasta los orificios situados en el pezón. Una dilatación del conducto, el seno galactóforo, se encuentra cerca del orificio del conducto en el tejido subareolar. El tejido subcutáneo y el tejido adiposo distribuidos alrededor de los lobulillos de la glándula confieren a la mama su contorno uniforme y en el período sin lactancia representan la mayor parte de su volumen. Los vasos linfáticos atraviesan la estroma que rodea a los lobulillos de la glándula y transportan la linfa hasta los conductos colectores. En la figura A.1 se muestra una vista tangencial de la mama sobre la pared torácica y vista de un corte (sagital) de la mama y la pared torácica asociada. A continuación se describe con más detenimiento la estructura y composición regular de la mama [B100].

A.1.A. Glándula mamaria

La mama se ubica dentro de la aponeurosis superficial de la pared torácica anterior y consiste de 15 a 20 lóbulos de tejido glandular de tipo tuboalveolar. El tejido conectivo fibroso forma un marco que sostiene los lóbulos y el tejido adiposo llena los espacios entre dichos lóbulos. El tejido conectivo subcutáneo rodea la glándula y se extiende en forma de tabiques entre los lóbulos y lobulillos, de modo que proporciona sostén a los elementos glandulares, pero no forma una cápsula separada alrededor de los componentes de la mama. Los engrosamientos fibrosos del tejido conectivo se mezclan en el tejido parenquimatoso de la mama, se extienden desde la capa profunda de la aponeurosis superficial (hipodermis) y se fijan a la dermis de la piel. Estas estructuras suspensorias, que se denominan ligamentos de Cooper, se insertan en forma perpendicular con respecto de las delicadas capas de la aponeurosis superficial de la dermis, permitiendo movilidad a la mama y proporcionando sostén. La mitad superior de la mama, y en particular el cuadrante superoexterno, contiene más tejido glandular que el resto de la glándula. En la figura A.1 se puede observar una vista tangencial y un corte sagital de la mama.

A.1.B. Pezón y areóla

La epidermis del pezón y de la areóla es muy pigmentada y algo arrugada. El pezón está revestido por un epitelio escamoso, estratificado y queratinizado. La superficie profunda de la epidermis está invadida por papilas dérmicas largas que permiten que los capilares

aporten sangre en la vecindad de la superficie. En la profundidad de la areola y el pezón existen haces de fibras musculares lisas dispuestas en forma radial y circunferencial en el tejido conectivo denso y de manera longitudinal a lo largo de los conductos galactóforos que se extienden hacia arriba del pezón. La areola contiene glándulas sebáceas, glándulas sudoríparas y glándulas areolares accesorias (de Montgomery), que presentan una estructura intermedia entre las glándulas mamarias verdaderas y las glándulas sudoríparas.

A.1.C. Morfología de la mama con la edad

La morfología de la porción secretoria de la glándula mamaria varía con la edad y durante el embarazo y la lactancia (ver figura A.2). En la mama inactiva el componente glandular es escaso y consiste principalmente en elementos ductales. Durante el ciclo menstrual la mama inactiva presenta cambios cíclicos leves. Al comienzo del ciclo los conductillos aparecen como cordones con una luz escasa o inexistente. Durante la estimulación estrogénica, aproximadamente en el momento de la ovulación, la altura de las células secretorias aumenta y se acumulan líquidos y lípidos en el tejido conectivo. Luego en ausencia de una estimulación hormonal continua la glándula regresa a un estado inactivo durante todo el resto del ciclo.

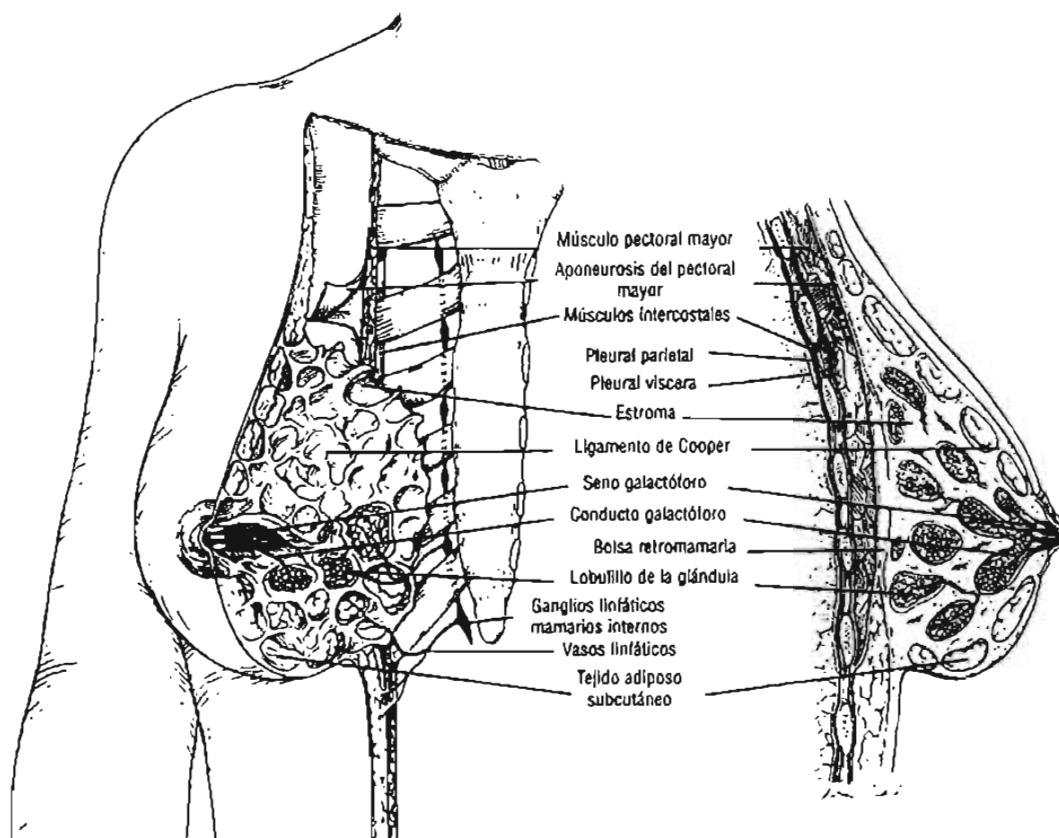


Figura A.1. Vista tangencial de la mama sobre la pared torácica y vista de un corte (sagital) de la mama y la pared torácica asociada (ver texto).

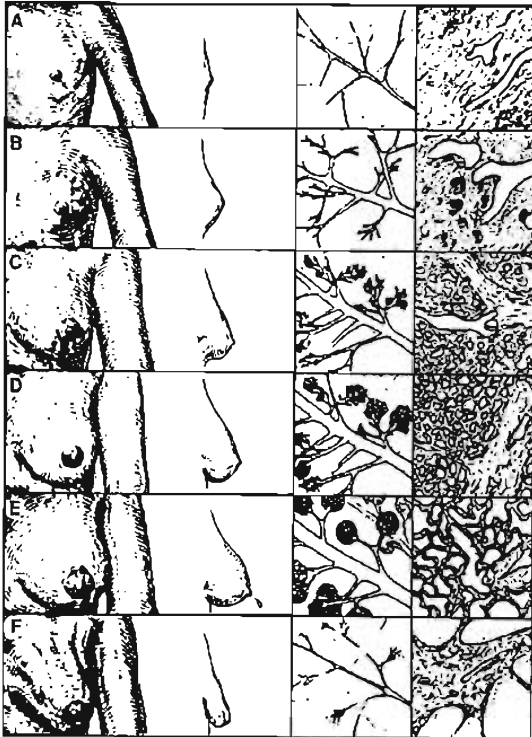


Figura A.2. Dibujo esquemático que muestra el desarrollo de la glándula mamaria. En las columnas 1 y 2 se muestran vistas anteriores y laterales de la mama. Los aspectos microscópicos de los conductos y los lobulillos se ilustran en las columnas 3 y 4, respectivamente. Paneles: (A) prepubertad (niñez); (B) pubertad; (C) madurez (edad de procreación); (D) embarazo; (E) lactancia; (F) estado posmenopáusico (senil). [BI00].

A.2. EVALUACIÓN CLÍNICA Y ESTANDARIZACIÓN DE LOS INFORMES MAMOGRÁFICOS (BI-RADS).

La principal aplicación de la mamografía es la detección de lesiones no palpables. Para esto se sigue una revisión sistemática de los siguientes rubros:

- Material: negatoscopio (en el caso de mamografía analógica), luz, lupa, visores.
- Método: comparación de la simetría de las proyecciones mamográficas de la mama izquierda y derecha, estudios previos, datos clínicos, antecedentes clínicos de la paciente.
- Hallazgos: se basan en la observación de la alteración del patrón mamario normal, como patrones diferentes, incluso en la misma mama.
- Analizar el área de la mama por regiones: región subareolar y retroglándular.
- Posicionamiento de la paciente: colocar en la posición adecuada la mama de la paciente de acuerdo al tipo de proyección radiográfica utilizada.
- Valores de exposición: asegurarse de utilizar valores adecuados de exposición radiográfica.
- Marcaje: La imagen mamográfica debe llevar la fecha del estudio, el nombre de la paciente, el tipo de proyección utilizado y el nombre del técnico radiólogo.

Para la evaluación de las imágenes mamográficas se ha implementado una terminología estandarizada por el Colegio estadounidense de Radiología (American College of Radiology) denominada "Breast Imaging Reporting and Data System" (BI-RADS) con la finalidad de evitar la confusión durante la interpretación de los informes y facilitar la vigilancia ulterior de las pacientes [AC98], [Do93]. BI-RADS se basa en la utilización de

un léxico estandarizado para facilitar la uniformidad de los informes radiológicos provenientes de distintas instituciones.

La estandarización del informe mamográfico también abarca la evaluación global de la probabilidad de las afecciones, la cual se incluye en la "Conclusión" ubicada al final del informe mamográfico. Existen 6 categorías de evaluación, cada una asociada con una recomendación para el manejo de la paciente. La categoría BI-RADS "0" identifica casos en los que se requiere un estudio por imágenes adicional para poder establecer una conclusión. Una vez obtenido este examen, el caso es asignado a una de las cinco categorías de "Evaluación final" (tabla A.1).

Tabla A.1. Categorías de evaluación con mamografía (BI-RADS).

Categoría	Resultado	Descripción, recomendación
0	Incompleto.	Requiere proyecciones adicionales.
1	Normal (hallazgo negativo de anomalía mamaria)	Ninguno, estudio de escrutinio de rutina.
2	Hallazgo benigno	Hallazgo inequívocamente benigno. Estudio de escrutinio de rutina.
3	Hallazgo probablemente benigno	Muy alta probabilidad de benignidad. Se recomienda un seguimiento a corto plazo para garantizar la estabilidad de la lesión.
4	Hallazgo de aspecto sospechoso, probablemente maligno	No definitivo pero con probabilidades razonables de malignidad. Debe considerarse la posibilidad de biopsia.
5	Altamente sugestivo de malignidad	Muy alta probabilidad de malignidad. Se recomienda profundizar el diagnóstico. Se requiere biopsia.

A.3. VALIDEZ DE UNA PRUEBA DE DIAGNÓSTICO

Cuando se estudia una muestra de pacientes, los datos obtenidos permiten clasificar a los sujetos en cuatro grupos según una tabla 2x2 como la que se muestra en la tabla A.2. En ella, se enfrenta el resultado de la prueba diagnóstica (en filas) con el estado real de los pacientes (en columnas) o, en su defecto, el resultado de la prueba de referencia o "estándar de oro" que se vaya a utilizar. El resultado de la prueba puede ser, correcto: verdadero-positivo (VP) y verdadero-negativo (VN), o incorrecto: falso-positivo (FP) y falso-negativo (FN). El análisis de su validez puede obtenerse calculando los valores de sensibilidad y especificidad [A194].

A.3.A. Sensibilidad

Es la probabilidad de clasificar correctamente a un individuo enfermo, es decir, la probabilidad de que para un sujeto enfermo se obtenga en la prueba un resultado positivo. La sensibilidad es, por lo tanto, la capacidad de la evaluación (o prueba) para detectar la

enfermedad. Cuando los datos obtenidos a partir de una muestra de pacientes se clasifican en una tabla como la que se muestra en la tabla A.2, es fácil estimar a partir de ella la sensibilidad como la proporción de pacientes enfermos que obtuvieron un resultado positivo en la prueba diagnóstica. Es decir:

$$Sensibilidad = \frac{VP}{VP + FN} \tag{A.1}$$

De ahí que también la sensibilidad se conozca como “fracción de verdaderos-positivos (FVP)”.

A.3.B. Especificidad

Es la probabilidad de clasificar correctamente a un individuo sano, es decir, la probabilidad de que para un sujeto sano se obtenga un resultado negativo. En otras palabras, se puede definir la especificidad como la capacidad para detectar a los individuos sanos. A partir de una tabla como la tabla A.2, la especificidad se estimaría como:

$$Especificidad = \frac{VN}{VN + FP} \tag{A.2}$$

De ahí que también sea denominada “fracción de verdaderos-negativos (FVN)”.

Tabla A.2. Relación entre el resultado de una prueba diagnóstica y la presencia o ausencia de una enfermedad.

Resultado de la prueba	Diagnóstico verdadero	
	Enfermo	Sano
Positivo	Verdaderos-Positivos (VP)	Falsos-Positivos (FP)
Negativo	Falsos-Negativos (FN)	Verdaderos-Negativos (VN)

Los conceptos de sensibilidad y especificidad permiten, por lo tanto, valorar la validez de una prueba de diagnóstico. Su utilidad es para quienes desarrollan o fabrican sistemas, métodos o pruebas, o diseñan protocolos, entre otros. Sin embargo, carecen de utilidad en la práctica clínica. Tanto la sensibilidad como la especificidad proporcionan información acerca de la probabilidad de obtener un resultado concreto (positivo o negativo) en función de la verdadera condición del enfermo con respecto a la enfermedad. Sin embargo, cuando a un paciente se le realiza alguna prueba, el médico carece de información a priori acerca de su diagnóstico verdadero, y más bien la pregunta se plantea en sentido contrario: ante un resultado positivo (negativo) en la prueba, ¿cuál es la probabilidad de que el paciente esté realmente enfermo (sano)?. Esta información puede completarse mediante los valores predictivos que se mencionan a continuación [Pi03].

A.3.C. Valor predictivo positivo

Es la probabilidad de padecer la enfermedad si se obtiene un resultado positivo en la evaluación. El valor predictivo positivo puede estimarse, por tanto, a partir de la proporción de pacientes con un resultado positivo en la prueba que finalmente resultaron estar enfermos:

$$VPP = \frac{VP}{VP + FP} \quad (A.3)$$

A.3.D. Valor predictivo negativo

Es la probabilidad de que un sujeto con un resultado negativo en la prueba esté realmente sano. Se estima dividiendo el número de verdaderos negativos entre el total de pacientes con un resultado negativo en la prueba:

$$VPN = \frac{VN}{FN + VN} \quad (A.4)$$

El concepto de valores predictivos, a pesar de ser de enorme utilidad a la hora de tomar decisiones clínicas y transmitir a los pacientes información sobre su diagnóstico, presenta la limitación de que dependen en gran medida de lo frecuente que sea la enfermedad a diagnosticar en la población objeto de estudio. Cuando la prevalencia de la enfermedad es baja, un resultado negativo permitirá descartar la enfermedad con mayor seguridad, siendo así el valor predictivo negativo mayor. Por el contrario, un resultado positivo no permitirá confirmar el diagnóstico, resultando en un bajo valor predictivo positivo.

A.4. HALLAZGOS MAMOGRÁFICOS NORMALES

Existe una amplia gama de hallazgos mamográficos normales en lo que concierne al tamaño, la configuración y la composición de los tejidos mamarios. La composición del tejido mamario puede variar entre la presencia casi exclusiva de tejido adiposo y un tejido fibroglandular extremadamente denso y se correlaciona con la sensibilidad de la mamografía. Algunos signos del cáncer de mama se muestran por la aparición de microcalcificaciones o masas tumorales, las cuales tienen coeficientes de atenuación tales que en la imagen radiográfica se observan blancas (mayor atenuación del haz de radiación), por lo que se les denomina radiopacas. Al tejido adiposo, con coeficientes de atenuación menores que los anteriores, se le dice radiolúcido pues se observa de color gris oscuro o negro en la mamografía. El tejido adiposo representa un buen fondo sobre el cual puede detectarse un cáncer de mama pequeño, pero el tejido fibroglandular denso (de color blanco en la mamografía) puede enmascarar un cáncer de mama [B100].

En el sistema BI-RADS la composición del tejido mamario se divide en cuatro categorías (ver figura A.3), una mama compuesta casi exclusivamente por tejido adiposo (A); la

presencia de islotes dispersos de densidades fibroglandulares que pueden enmascarar lesiones en la mamografía (B y C) y un tejido mamario sumamente denso que disminuye la sensibilidad de la mamografía (D) [AC93]. Las mujeres jóvenes generalmente presentan una mayor cantidad de tejido fibroglandular, en este subgrupo de mujeres las mamas, por lo general, son más radiopacas que en las mujeres de mayor edad. Sin embargo, es importante tener presente que existen amplias variaciones de densidad del tejido mamario en mujeres de la misma edad; en algunas mujeres jóvenes las mamas están compuestas casi exclusivamente por tejido adiposo y en algunas mujeres de mayor edad las mamas son extremadamente densas.

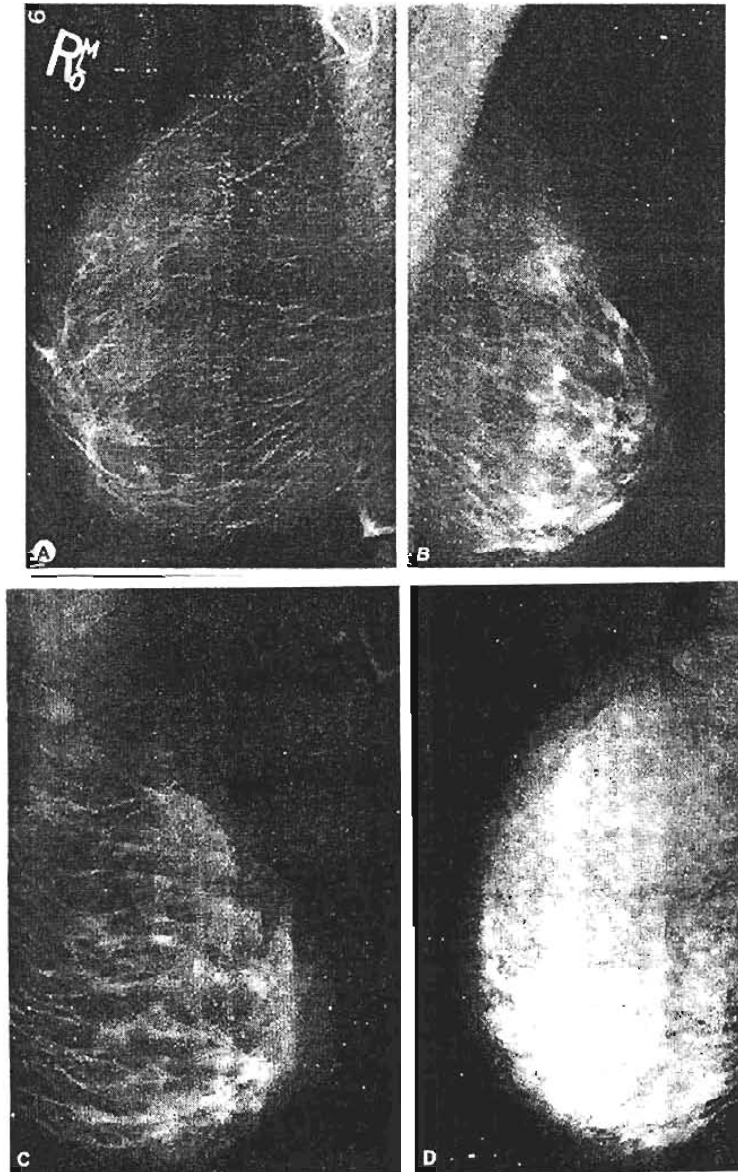


Figura A.3. Cuatro categorías de composición tisular mamaria en la mamografía. A. El tejido adiposo representa casi la totalidad del tejido mamario. B. Se aprecian islotes dispersos de tejido fibroglandular que pueden enmascarar una lesión en la mamografía. C. Se observa una mama de densidad heterogénea. Este factor puede reducir la sensibilidad de la mamografía. D. Se observa una densidad mamaria muy pronunciada que reduce la sensibilidad de la mamografía.

A medida que la mujer envejece o procrea, el tejido fibroglandular es reemplazado por tejido adiposo. El reemplazo del tejido fibroglandular por el tejido graso se produce generalmente desde la parte posterior hacia la parte anterior y desde la parte medial hacia la parte lateral de la mama. Por lo tanto, la mayoría de las mujeres presentan una mayor cantidad de tejido fibroglandular en la región subareolar que en la vecindad de la pared torácica y en la parte lateral que en la parte medial de la mama. El proceso de reemplazo del tejido fibroglandular por tejido adiposo generalmente es simétrico. Algunas mujeres presentan tejido mamario accesorio en la axila; este hallazgo es normal y puede ser unilateral o bilateral [Ad87].

Durante el embarazo la densidad mamaria aumenta. El aumento de la densidad también se asocia con la administración de hormonas exógenas en algunas mujeres posmenopáusicas [Mc94].

A.5. HALLAZGOS MAMOGRÁFICOS ANORMALES

Las alteraciones más frecuentes observadas en las mamografías consisten en lesiones ocupantes y calcificaciones, y las características radiográficas de estas alteraciones pueden aportar indicios etiológicos. En el informe mamográfico estandarizado los términos utilizados para describir las lesiones ocupantes y las calcificaciones indican la probabilidad de malignidad. Otros hallazgos significativos consisten en la presencia de lesiones neoformadas o evolutivas, la distribución asimétrica bilateral del tejido fibroglandular, distorsiones de la arquitectura glandular, aumento del espesor o retracción de la piel, retracción del pezón y aumento de tamaño de los ganglios linfáticos [BI00].

A.5.A. Lesiones ocupantes

Una lesión ocupante es una lesión que se aprecia como mínimo en dos proyecciones mamográficas. Las lesiones ocupantes, o *masas* se definen por la *forma*, las características de los bordes y la densidad. La forma puede ser redonda, ovalada, segmentada o irregular (ver figura A.4). Las lesiones ocupantes ovaladas y redondas generalmente son benignas. Una morfología irregular sugiere una mayor probabilidad de malignidad.

Los *bordes* de las lesiones ocupantes también son importantes como indicadores de malignidad [GI93]. Los bordes pueden ser descritos como bien delimitados, microsegmentados (numerosas lobulaciones diminutas), enmascarados (parcialmente ocultos por los tejidos circundantes), indistintos o espiculados (ver figura A.5). Los bordes *bien delimitados* sugieren una etiología benigna y se asocian con una muy baja probabilidad de malignidad (menos del 2%) [Ha88]. Una lesión ocupante quística no requiere nuevos estudios diagnósticos. Una lesión ocupante sólida puede requerir una mamografía con aumento para confirmar que todos los bordes de la lesión sólida estén claramente delimitados. Una lesión ocupante sólida solitaria bien delimitada generalmente es observada durante 6 meses para determinar si es estable (sin crecimiento). Si la lesión es estable, se recomienda una vigilancia mamográfica continua durante 2 años como mínimo [Br89]. La presencia de varias lesiones ocupantes bien delimitadas es una evidencia aún

más firme de benignidad o ganglios linfáticos intramamarios benignos; en estos casos, por lo general, es suficiente la reevaluación de la paciente después de 1 año.

Los bordes *microsegmentados* aumentan la probabilidad de malignidad. Si la lesión ocupante mantiene una relación de proximidad con el tejido fibroglandular de densidad similar, es posible que el borde sea *enmascarado*, en cuyo caso se recurre a la composición del área con la finalidad de visualizar con mayor precisión los bordes de la lesión, el hallazgo de bordes *indistintos* sugiere la posibilidad de malignidad. Una lesión ocupante con bordes *espiculados* se caracteriza por líneas que irradian desde el borde y se asocia con una elevada probabilidad de cáncer. Un área espiculada en ausencia de una lesión ocupante asociada se define como una *distorsión de la arquitectura*.

La *densidad* de una lesión ocupante también orienta hacia la posible etiología. En general, las lesiones ocupantes benignas se asocian con una menor densidad que las lesiones ocupantes malignas. Las lesiones ocupantes malignas generalmente se asocian con una mayor densidad radiológica que las lesiones ocupantes benignas o el tejido mamario normal circundante. Sin embargo, la densidad de una lesión no siempre es un indicador confiable de benignidad o malignidad [Ja91].

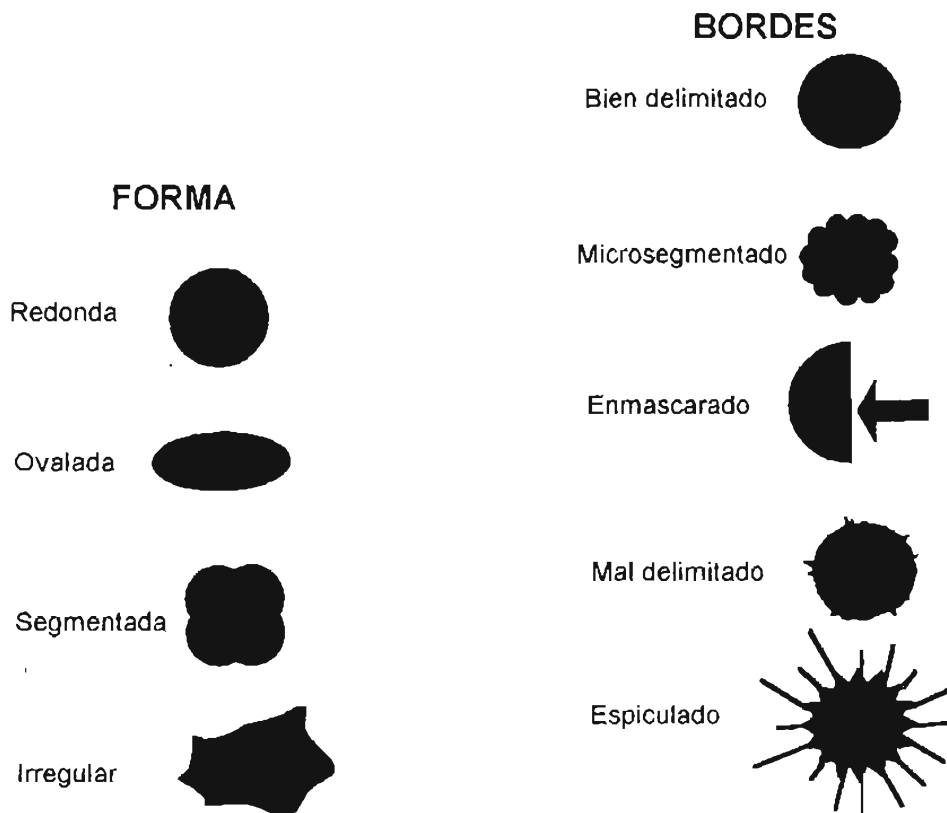


Figura A.4. Terminología estandarizada para la configuración de las lesiones.

Figura A.5. Terminología estandarizada para los bordes de las lesiones.

Existen varios hallazgos asociados con lesiones ocupantes que pueden orientar hacia la etiología de la lesión. Por ejemplo, la presencia de calcificaciones en una lesión ocupante a

menudo brinda información definitiva. Las calcificaciones de los fibroadenomas generalmente son grandes y densas; estas calcificaciones pueden estar localizadas en la parte central o en el borde del fibroadenoma. Es posible detectar calcificaciones granulares finas en lesiones benignas o malignas. Otros hallazgos que pueden asociarse con una lesión ocupante consisten en retracción cutánea, retracción del pezón y engrosamiento de la piel.

A.5.B. Calcificaciones

El calcio es un elemento metálico bivalente de color blanco que pertenece al grupo alcalino. La calcificación es el depósito de sales de calcio en los tejidos. En las mamas las calcificaciones generalmente se producen en la forma de hidroxapatita cálcica o fosfato tricálcico [La86]. En la terminología estandarizada las calcificaciones se dividen en tres grupos generales: las típicamente benignas, las indeterminadas y las asociadas con mayor probabilidad de malignidad. Las calcificaciones *típicamente benignas* generalmente pueden identificarse por sus rasgos mamográficos y son cutáneas, vasculares, groseras, bastoniformes, redondeadas, en cáscara de huevo y en forma de "leche de calcio" (ver figura A.6). Las calcificaciones *indeterminadas* y las calcificaciones *asociadas con mayor probabilidad de malignidad* pueden ser descritas como amorfas o indistintas: pleomorfas o heterogéneas o finas, lineales y ramificadas ("en molde") (ver figura A.7).

Las calcificaciones también se dividen según su distribución: las calcificaciones *agrupadas o en racimo* abarcan más de 5 calcificaciones en un área pequeña (menos de 2 cm²) y pueden ser benignas o malignas. Las calcificaciones *lineales* se distribuyen en una línea que puede presentar pequeñas ramificaciones. Las calcificaciones *segmentarias* siguen la distribución de un conducto y sus ramificaciones. Las calcificaciones *regionales* abarcan un mayor volumen de glándula mamaria y no presentan necesariamente una distribución ductal. Las calcificaciones *difusas/diseminadas* se distribuyen en forma aleatoria en toda la mama y casi siempre son benignas.

A.6. EJEMPLOS DE LESIONES BENIGNAS Y MALIGNAS

Generalmente, varias regiones sospechosas que se encuentran en la mama no son necesariamente malignas. Kopans [Ko98] y Wolberg [Wo00] identifican las siguientes lesiones benignas y sus características asociadas.

Fibroadenoma: Estas son las lesiones benignas usualmente más observadas en mujeres menores de 30 años de edad. Éstas son redondas y bastante móviles. Aún siendo lesiones benignas, tienen un factor de riesgo a largo plazo para la aparición de cáncer de mama.

Lesión fibroquística: Es un término ambiguo que incluye muchos tipos de lesiones mamarias benignas. Las autopsias muestran que la mitad de todas las mujeres tienen cambios microscópicos en la mama consistentes con lesiones fibroquísticas. La lesión fibroquística no es un precursor de malignidad.

Típicamente benignas



Figura A.6. Morfología de las calcificaciones benignas típicas.

Indeterminada



Figura A.7. Morfología de calcificaciones indeterminadas y calcificaciones asociadas con alta probabilidad de malignidad.

Hiperplasia atípica de la mama: Esta condición benigna está caracterizada por una acentuada proliferación y atipia del epitelio (ductal o lobular). Se encuentra en un 3% de todas las biopsias con hallazgos benignos. Esta condición se asocia con un 13% de subsiguiente desarrollo de cáncer de mama.

Tumor phyllodes: Esta condición benigna es un tumor fibroepitelial y aproximadamente un 10% de estos casos produce metástasis a partir de una lesión histológicamente maligna o benigna.

Mastitis periductal: En una de sus formas, esta condición es la causa de la mayor parte de las inflamaciones de la mama de origen no lacto. Bacterias aeróbicas o anaeróbicas pueden complicar el caso o ser estériles. La ocurrencia crónica es la causa más frecuente de la descarga del pezón en mujeres premenopáusicas.

Papilomas: Estas masas son típicamente más pequeñas que 1 cm y usualmente se observan como crecimientos intraductales. Las pacientes frecuentemente tienen descarga por el pezón y los papilomas pueden ser únicos o múltiples. Las pacientes con papilomas frecuentemente desarrollan cáncer.

La detección de lesiones ocupantes malignas requiere que el médico radiólogo considere las características de cada región. Todas las masas diagnosticadas como malignas presentan diferencias, algunas consisten en una gran variedad de tipos de células diferentes con diversas apariencias y diferentes características de comportamiento. Las tasas de crecimiento de los tumores varían con cada paciente. Usualmente el cáncer de mama promedio toma aproximadamente 9 años para alcanzar 1 cm. El cáncer de mama causa la muerte al esparcirse a otras regiones del cuerpo (metástasis) y tal propagación puede ocurrir en cualquier momento durante el crecimiento del tumor. Típicamente se recurre a la mastectomía, cuando la metástasis ha ocurrido y existe como tejido canceroso que es muy pequeño para ser detectado por cualquier estudio clínico. Ciertos cánceres producen metástasis con más frecuencia que otros.

Carcinoma intraductal in situ (CIDS) es una manifestación común de cáncer de mama. Lesiones ocupantes malignas aparecen a partir de las células epiteliales del revestimiento de los ductos y lóbulos. Los cánceres que no penetran fuera del confinamiento de los ductos de los lóbulos donde aparecen, raramente producen metástasis hacia los nódulos linfáticos o hacia otras regiones distantes en el cuerpo. Tales cánceres no infiltrantes se denominan cáncer de mama *in situ*. Usualmente existen 3 condiciones para que ocurra la metástasis:

1. Penetración de los vasos sanguíneos o del sistema linfático.
2. Las células cancerígenas deben permanecer intactas durante el tránsito.
3. Las células deben alojarse y proliferar en una región distante.

El cáncer de mama típicamente prolifera a partir de la mama a los nódulos linfáticos y de aquí a sitios distantes. Un cáncer que se propaga a los nódulos linfáticos tiene una gran posibilidad de poseer otras características malignas. Wolberg [Wo00] identifica 2 tipos de cáncer de mama no infiltrante.

Carcinoma lobulillar in situ: Este tipo de cáncer no produce metástasis pero proporciona de un 20% a 30% de riesgo de carcinoma invasor en la misma mama o en la mama contralateral. Este tipo de cáncer se trata típicamente con mastectomía bilateral.

Carcinoma intraductal: Similarmente este tipo de cáncer generalmente no produce metástasis y el riesgo subsecuente de cáncer invasivo se confina a la misma mama. Esta forma de cáncer se clasifica de acuerdo con su morfología como cribiforme, micropapilar, o comedo. La forma comedo, usualmente, puede viajar ampliamente sin detectarse a través de la mama. De esta forma, existe un gran riesgo de desarrollar un cáncer invasivo después de que éste se ha removido. Radioterapia o mastectomía son las formas adicionales más comunes de tratamiento.

Otro cáncer de mama típico es el cáncer de mama invasivo. Estos son tipos de cáncer que penetran fuera del confinamiento de los ductos de los lóbulos donde éstos aparecen y frecuentemente producen metástasis en los nódulos linfáticos o en regiones distantes del cuerpo. El incremento de la edad es el factor de riesgo principal para desarrollar esta forma de cáncer. Éste puede asociarse con engrosamiento de la piel o retracción del pezón. Egan [Eg88] identifica los siguientes tipos comunes de carcinomas invasivos; ductal, lobular, medular, comedocarcinoma, papilar, cirroso y tubular. Al localizar una masa en una región sospechosa, el médico radiólogo analizará sus características claves. Kopans [Ko98]

subraya las siguientes características: 1) Región: el cáncer de mama puede ocurrir en cualquier parte de la mama o donde se encuentre tejido mamario. Estadísticamente la localización más común para el cáncer de mama es en el cuadrante superoexterno de la mama; 2) Tamaño: el tamaño del cáncer es importante para el pronóstico. El pronóstico es mucho mejor para tumores de tamaño pequeño. Resultados de los estudios de escrutinio típicos indican que más del 50% de los cánceres invasivos detectados son menores que 1 cm de diámetro; 3) Forma: las masas pueden dividirse en 4 formas, redonda, oval, lobular e irregular, siendo que la probabilidad de malignidad aumenta conforme la forma se vuelve más irregular; 4) Bordes: las masas que poseen márgenes que forman una interfaz definida y abrupta con el tejido circundante son casi siempre hallazgos benignos. El cáncer de mama clásico posee márgenes espiculados. La mayoría de los cánceres de mama tienen límites irregulares mientras invaden a los tejidos circundantes, produciendo así bordes mal definidos.

Para una discusión visual de las diversas formas y signos del cáncer de mama, referirse a Kopans [Ko98] y Gamagami [Ga96].

ANEXO B

ASPECTOS TÉCNICOS EN EL ANÁLISIS DE IMÁGENES DIGITALES

ANEXO B

ASPECTOS TÉCNICOS EN EL ANÁLISIS DE IMÁGENES DIGITALES

B.1. CONSIDERACIONES REFERENTES AL ANÁLISIS DE LAS MAMOGRAFÍAS DIGITALES

El programa empleado para escribir las rutinas de análisis (llamadas scripts) fue Matlab versión 7.0 ("Release" 14). El formato empleado para el análisis de las mamografías digitales es el DICOM, modalidad MG (mamografía digital), que utiliza el Senographe 2000D para guardar las imágenes obtenidas. Sin embargo, existen algunos comentarios pertinentes referentes a estas elecciones que a continuación se describen.

B.1.A. Matlab v7

Debido a las bondades de este paquete de programas comercial en lo que concierne al procesamiento de imágenes, se seleccionó como plataforma de desarrollo para generar las rutinas necesarias para analizar las mamografías digitales de acuerdo con los parámetros anteriormente presentados.

En la figura B.1 se presenta el organigrama empleado de procesamiento y análisis de las imágenes.

En un principio se trabajó con una versión anterior de Matlab (6.5R13), sin embargo se cambió por la versión 7 pues se detectó que la anterior tenía algunas carencias para analizar las imágenes tipo DICOM (.dcm). Estas carencias de la versión 6.5 (v6.5) nos habían llevado a pasar las mamografías digitales de DICOM a un formato más reconocible (TIFF) por Matlab v6.5 usando para tal efecto el programa ImageJ v1.33 (del dominio público). Lamentablemente, nos dimos cuenta que ImageJ efectuaba cambios que modificaban la información de la imagen al pasarla de DICOM a TIFF. Al comparar los resultados obtenidos al usar Matlab v6.5, para las imágenes TIFF seleccionadas en este trabajo de tesis, con los de [Bu04], para el histograma correspondiente a los niveles de gris obtenidos con un equipo Senographe similar al nuestro, resultó que existían grandes variaciones.

Una vez que se utilizó Matlab v7 para analizar las mamografías digitales en formato DICOM, se obtuvo un buen acuerdo entre los histogramas de ambos trabajos. En el capítulo 5, sección 5.1, se presentaron los resultados de esta comparación.

B.1.B. Mamografías digitales en formato DICOM

El formato empleado para la transferencia de imágenes de uso clínico es el DICOM (Digital Imaging and Communications in Medicine). Este formato, acordado por el Colegio Estadounidense de Radiología (ACR) y la Asociación Nacional de Fabricantes Eléctricos (NEMA) en 1983 [DI03], pretende estandarizar el formato de las imágenes médicas de uso clínico para su transferencia entre diferentes equipos e instituciones (versión 1.0 en 1985,

versión 2.0 en enero 1988 y versión 3.0 en 1989). Sin embargo, para cada aplicación médica que use una imagen digital este formato cambia, ya que debe incluir información sobre el paciente y sobre los datos utilizados inherentes a cada aplicación. Esta información, junto con los datos propios de la imagen, se registra en diferentes campos definidos para cada formato DICOM de aplicaciones médicas diferentes (Tomografía axial, mamografía digital, radiografía digital, entre otros). Esto obliga a analizar los encabezados (primera parte de la imagen donde se registran los campos de información), para saber como analizar correctamente las imágenes. En el caso de las mamografías digitales obtenidas del equipo Senographe 2000D (modalidad MG), éstas son imágenes de 16 bits (2294 renglones y 1914 columnas), con un encabezado de 153 renglones en los cuales se guardan datos de la técnica radiológica de exposición (tensión, carga, combinación blanco/filtro, entre otros), información de la paciente (nombre, número de estudio, edad, entre otros), datos del hospital (nombre del hospital, dirección, entre otros), datos de la técnica radióloga e información de lectura y guardado de la imagen (tipo de lectura, en nuestro caso la lectura se realiza del bit menos representativo al más representativo, clase de archivo definido por [DI03, GE00], entre otros)

Los médicos radiólogos están acostumbrados a evaluar las imágenes mamográficas en película radiográfica, donde el tejido con mayor coeficiente de atenuación (tejido radiopaco) que los tejidos circundantes (tejidos radiolúcidos), se visualiza más claro que los demás tejidos que lo rodean. Sin embargo en la radiografía digital sucede justamente lo opuesto, pues en las zonas donde hay menor atenuación de los rayos X, la imagen muestra niveles de gris más claros que en las zonas con mayor atenuación (como en el caso del archivo RAW, ver más abajo). Debido a esta razón los archivos de imagen DICOM (como en el caso de los archivos de imágenes procesadas, ver más abajo), muestran una inversión en la escala de niveles de gris de tal forma que la imagen visualizada (y guardada) sea igual a la de una imagen en película radiográfica, i.e. niveles más bajos de gris corresponden a tonalidades claras y niveles altos a tonalidades oscuras, tomando en cuenta una transformación logarítmica acorde a la visión humana. De esta forma se debe tener cuidado en el tratamiento de las imágenes con los programas convencionales de imágenes donde los niveles más bajos de gris corresponden a tonalidades oscuras y niveles altos a tonalidades claras.

El equipo Senographe 2000D ofrece 2 tipos de imágenes (ver sección 2.3.B):

- **RAW:** afectada por una corrección de “aplanado”, para obtener uniformidad del brillo, alta frecuencia de modulación, uniformidad en el cociente señal a ruido y la corrección de píxeles. Además se realiza una detección del borde de la mama y todo lo que esté fuera de ésta es cubierto por una máscara negra para facilitar la evaluación visual de la imagen.
- **Procesada:** ver sección 2.3.B. Este tipo de imagen muestra un aspecto visual de la mama, tal como los médicos radiólogos están habituados y entrenados para evaluar una imagen mamográfica en película radiográfica.

Por estas razones, para cualquier tipo de análisis que se desee llevar a cabo, éste debe hacerse utilizando las imágenes RAW, aunque sería deseable analizar los datos sin ningún procesamiento de la imagen, a fin de estudiar el sistema de adquisición y tal vez proponer algoritmos alternativos de procesamiento.

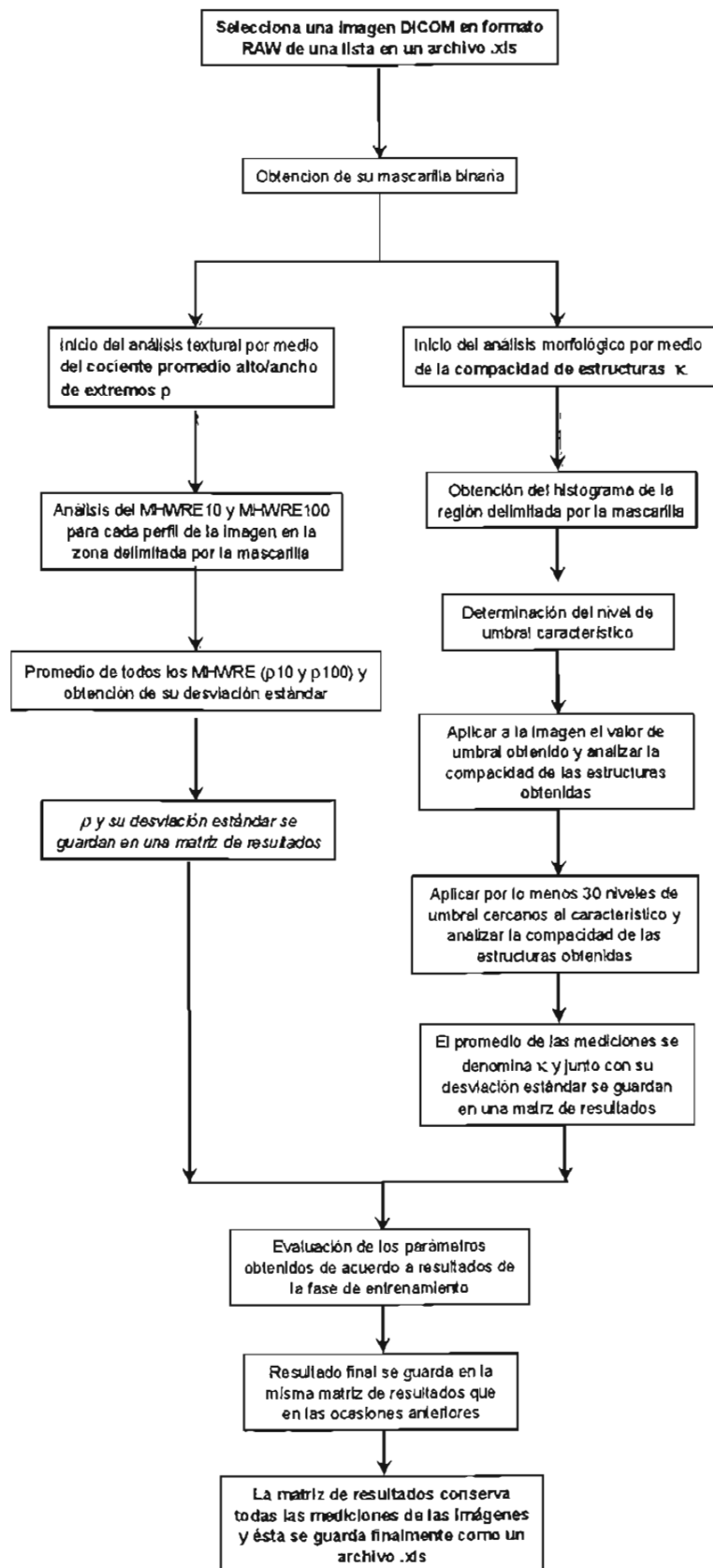


Figura B.1. Organigrama del análisis de las mamografías digitales por medio de las rutinas desarrolladas en Matlab v7.

ANEXO C

OBTENCIÓN DE LA MASCARILLA BINARIA

ANEXO C

OBTENCIÓN DE LA MASCARILLA BINARIA

C.1. PROCESO EMPLEADO PARA LA OBTENCIÓN DE LA MASCARILLA BINARIA

Para analizar únicamente el área interna de la mama, para cada imagen, se propone el uso de una mascarilla binaria. De esta forma, los procesos de análisis de la textura y la morfología de la mama se realizan en la región delimitada por la mascarilla empleada. Para la obtención de la mascarilla se utiliza un algoritmo de proceso binario de cerradura morfológica seguido por un proceso de apertura morfológica de la imagen, al utilizar un elemento estructurante circular de 10 píxeles de diámetro. En nuestro caso la mascarilla obtenida fue una región con nivel de gris cero. Ambos procesos se realizan utilizando las rutinas programadas de la plataforma de desarrollo empleada (Matlab v7 "Release" 14). A continuación se ofrece la definición y teoría de los procesos empleados por Matlab para aislar la región de interés (ROI) deseada. Primero se revisa la dilatación y erosión morfológicas, y finalmente su operación en la cerradura y apertura morfológicas.

C.1.A. Dilatación (\oplus) y erosión (\ominus) morfológicas

El elemento estructurante aplicado a la imagen de entrada (I_{in}), es una parte esencial en las operaciones de dilatación y erosión morfológicas. Elementos estructurantes consisten de una matriz de 0's y 1's, típicamente más pequeños que la imagen que se está procesando. El pixel central del elemento estructurante, llamado el origen, identifica al pixel de interés (el pixel de la imagen que será procesado). Los píxeles en el elemento estructurante contienen 1's, los cuales definen su vecindad. Son estos píxeles los que usualmente se consideran en un proceso de dilatación y erosión morfológicas. En la figura C.1 se muestra un ejemplo de elemento estructurante en forma de diamante con tamaño de 7 x 7 píxeles. Usualmente se elige un elemento estructurante del mismo tamaño y forma de los objetos que se desean procesar en la imagen de entrada, de esta forma si se desean encontrar líneas en una imagen, se crea un elemento estructurante lineal.

0	0	0	1	0	0	0
0	0	1	1	1	0	0
0	1	1	1	1	1	0
1	1	1	1	1	1	1
0	1	1	1	1	1	0
0	0	1	1	1	0	0
0	0	0	1	0	0	0

Figura C.1. Ejemplo de un elemento estructurante con forma de diamante y tamaño 7 x 7 píxeles.

En las operaciones de dilatación y erosión morfológicas, el estado de cualquier pixel en la imagen de salida (I_{out}) se determina al aplicar una regla de correspondencia al pixel y a sus vecinos en la imagen de entrada. La regla usada para el procesamiento del pixel, define

la operación empleada como una dilatación o una erosión morfológicas. La tabla C.1 muestra las reglas empleadas para ambos procesos.

Tabla C.1. Reglas de dilatación y erosión morfológicas.

Operación	Regla
Dilatación \oplus	El valor del pixel en la imagen de salida es el valor máximo de todos sus pixeles vecinos en la imagen de entrada. En una imagen binaria, si cualquiera de los pixeles vecinos tiene un valor de 1, el valor del pixel de salida será 1.
Erosión \ominus	El valor del pixel en la imagen de salida es el valor mínimo de todos sus pixeles vecinos en la imagen de entrada. En una imagen binaria, si cualquiera de los pixeles vecinos tiene un valor de 0, el valor del pixel de salida será 0.

En la figura C.2 se muestra un ejemplo de dilatación morfológica aplicada a una imagen binaria (I_{in}) de 5 x 5 pixeles con un elemento estructurante horizontal (ES) de 1 x 3 pixeles: $I_{out} = I_{in} \oplus ES$. Cabe remarcar que las operaciones de dilatación y erosión no son operaciones inversas, i.e., por ejemplo, si se realiza una erosión a una imagen, seguida de una dilatación, no se recuperará la imagen original.

<table border="1" style="display: inline-table; margin-right: 20px;"> <tr> <td style="padding: 2px 10px;">1</td> <td style="border: 2px solid black; padding: 2px 10px;">1</td> <td style="padding: 2px 10px;">1</td> <td colspan="2"></td> </tr> </table>					1	1	1		
1	1	1							
Elemento estructurante									
1	0	0	0	0	1	1	0	0	0
1	0	1	0	0	1	1	1	1	1
0	0	1	1	0	0	0	0	0	0
0	0	0	1	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0
Imagen de entrada I_{in}					Imagen de salida I_{out}				

Figura C.2. Ejemplo de dilatación morfológica en un objeto binario de tamaño 5 x 5 pixeles usando un elemento estructurante de 1 x 3 pixeles.: $I_{out} = I_{in} \oplus ES$.

Funciones morfológicas colocan el origen del elemento estructurante (su elemento central), sobre el pixel de interés en la imagen de entrada. Sin embargo, para los pixeles que se encuentran en el borde de la imagen, los pixeles vecinos fuera del borde, se definen por las partes del elemento estructurante que se extienden más allá del borde de la imagen. Para el procesamiento de los pixeles del borde, las funciones morfológicas asignan un valor a estos pixeles indefinidos (más allá del borde original de la imagen de entrada), como si la función hubiera pegado columnas y renglones adicionales a la imagen de entrada. El valor de estos pixeles añadidos varía para las operaciones de dilatación y erosión. En la tabla C.2 se muestra un ejemplo que describe las reglas de adición para la dilatación y la erosión.

C.1.B. Apertura (\circ) y cerradura (\bullet) morfológicas

La dilatación y la erosión morfológicas se combinan usualmente para mejorar las operaciones de procesamiento de imágenes. Por ejemplo, la definición de **apertura morfológica** de una imagen es una operación de erosión seguida por una de dilatación,

aplicando el mismo elemento estructurante para ambos procesos. La operación de **cerradura morfológica** es el proceso inverso, al aplicar una operación de dilatación seguida por uno de erosión.

Tabla C.2. Reglas para la adición de píxeles a los bordes de la imagen I_{in} .

Operación	Regla
Dilatación	A los píxeles más allá del borde de la imagen se les asigna el valor mínimo permitible para el tipo de imagen procesada. Para imágenes binarias, el valor que se asigna a estos píxeles es 0.
Erosión	A los píxeles más allá del borde de la imagen se les asigna el valor máximo permitible para el tipo de imagen procesada. Para imágenes binarias, el valor que se asigna a estos píxeles es 1.

En la figura C.3 se muestra el ejemplo de una apertura morfológica. En este ejemplo se tiene una imagen I_{in} (272 x 280 píxeles), de una tarjeta de circuitos que contiene microchips rectangulares y líneas de circuitos. El tamaño del elemento estructurante (ES) debe ser suficientemente grande para eliminar las líneas de circuitos, pero preservar las de los rectángulos. Se utiliza un ES de 40 x 30 píxeles para eliminar las líneas de circuitos y apreciar la forma rectangular de los microchips: $I_{out} = I_{in} \circ ES$. En la imagen de la izquierda se muestra la imagen original, al centro la imagen erosionada y a la derecha la imagen resultante de la erosión seguida de la dilatación.

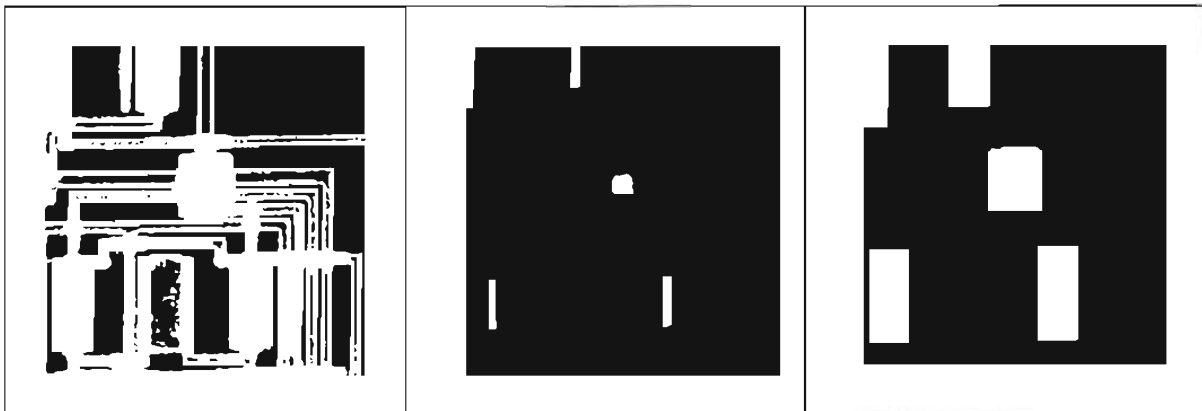


Figura C.3. Ejemplo de un proceso de apertura aplicado a una imagen de líneas de circuitos y microchips, para eliminar las líneas de circuitos y observar solamente las formas rectangulares de los microchips: $I_{out} = I_{in} \circ ES$. Izquierda: imagen original. Centro: imagen original erosionada. Derecha: dilatación de la imagen central.

En nuestro caso, se realiza primero una cerradura para asegurar que no se pierde información en el borde de la mama (pequeños bordes o imperfecciones de la piel), y luego se procede a una apertura morfológica para eliminar todos los objetos pequeños que pudiera haber en el exterior y sólo quedarnos con el interior de la mama (rigurosamente hablando este proceso elimina del borde mamario, un área mínima de 5 x 5 píxeles, aproximadamente).

