



UNIVERSIDAD NACIONAL  
AUTÓNOMA DE  
MÉXICO

**UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO**

POSGRADO EN CIENCIA E INGENIERÍA DE LA COMPUTACIÓN

RECONOCIMIENTO DE IRIS USANDO  
LA TRANSFORMADA HERMITE MULTIESCALA

**T E S I S**

QUE PARA OBTENER EL GRADO DE:

**MAESTRO EN INGENIERÍA  
(COMPUTACIÓN)**

**P R E S E N T A:**

**ALFONSO ESTUDILLO ROMERO**

**DIRECTOR DE TESIS: BORIS ESCALANTE RAMÍREZ**

México, D.F.

2008



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.

# Dedicatoria y agradecimientos

A mis padres, hermanos, abuelita y a toda mi familia, a quienes les agradezco el apoyo que siempre me han brindado.

A Nadia, mi fuente de inspiración, le agradezco su invaluable compañía, paciencia y apoyo incondicional en todo momento. Gracias también a sus padres por las noches de desvelo que algunas veces les ocasioné.

Amalia, Diana, Lulú, muchas gracias por los jalones de oreja, necesarios para la culminación satisfactoria de nuestros estudios sin dejar de mencionar la valiosa amistad recibida.

Ernesto, Cira y Oscar, gracias por permitir tener tardes amenas en el laboratorio.

# Agradecimientos especiales

Dr. Yann Frauel por la atención prestada y comentarios realizados durante la revisión de este documento.

Dra. Elena Martínez Pérez por sus valiosos comentarios y sugerencias que se plasmaron en este documento para permitir una mejor comprensión del mismo.

Dr. Carlos J. Rivero Moreno por su valiosa ayuda y aclaraciones realizadas durante el desarrollo de este proyecto.

Dra. Lucía Medina Gómez por presentarme el fantástico mundo de las wavelets y su dedicación en la revisión de este documento.

Dr. Boris Escalante Ramírez por haberme otorgado la oportunidad de integrarme al equipo de trabajo, orientarme y acercarme hacía el camino de la investigación.

Al Posgrado en Ciencia e Ingeniería de la Computación de la UNAM, planta docente y servicios que me permitieron recibir estudios de excelente calidad.

Al Consejo Nacional de Ciencia y Tecnología (CONACYT) por haber otorgado la beca para la realización de los estudios de maestría con el número de registro 200564.

Al Programa de Apoyo a Proyectos de Investigación e Innovación Tecnológica (PAPIIT) de la UNAM ya que una parte de este proyecto fue financiada con la clave IN105505.

# Resumen

Diversas tecnologías biométricas han sido utilizadas a lo largo de los últimos años y aún continúan en fase de desarrollo. Cada una de estas tecnologías posee sus ventajas y desventajas. Uno de los sistemas biométricos que presenta una mejor confianza en aplicaciones que demandan alta seguridad es el reconocimiento de iris.

Un sistema típico de reconocimiento de iris requiere de un modelo de representación de imágenes para extraer las características de la textura del iris. Una vez que las características han sido extraídas, éstas son almacenadas en una plantilla biométrica (código) asociada a cada individuo. En el paso siguiente, para poder efectuar el reconocimiento entre personas registradas en uno de estos sistemas, las plantillas biométricas registradas son comparadas con las que presenta el individuo que pretende ser reconocido por el sistema.

Frecuentemente, los sistemas de reconocimiento de iris han utilizado a las wavelets (ondeletas) durante el proceso de extracción de características de texturas de iris, particularmente la familia de wavelets de tipo Gabor. Esta familia utiliza propiedades importantes encontradas en algunos sistemas de visión de mamíferos, ajustándose en forma a los perfiles de los campos receptivos de ciertas células ubicadas en la retina y a nivel de la corteza visual, para extraer la información más importante de las imágenes. Las wavelets son construidas a partir de las funciones Gabor elementales, las cuales son utilizadas para analizar una señal de entrada dada.

En este trabajo se desarrolló un sistema de reconocimiento de iris, en el cual se utilizó a la transformada Hermite como una herramienta alternativa de extracción de características de texturas de iris. La transformada Hermite puede simular algunas propiedades de los sistemas de visión de mamíferos de manera similar a las wavelets Gabor. La transformada Hermite es una técnica de descomposición local, en la cual una imagen es localizada a través de una ventana Gaussiana y luego proyectada sobre polinomios ortogonales respecto a dicha ventana, cuyas funciones de Hermite son soluciones a la ecuación del oscilador armónico. La conexión entre la transformada Hermite y los sistemas de visión de mamíferos reside en que los filtros de análisis usados son similares a las derivadas de Gaussiana, las cuales también son capaces de modelar y ajustarse a los perfiles de los campos receptivos de células simples.

Para efectos de comparación, en este proyecto también se utilizaron a las wavelets Gabor como un modelo extractor de características de texturas de iris y se reprodujeron algunas de las etapas del sistema de reconocimiento de iris presentado por Daugman, quien es pionero en esta aplicación usando wavelets Gabor. Los resultados obtenidos con el sistema de reconocimiento de iris propuesto, usando la transformada Hermite como una herramienta alternativa de extracción de características de texturas de iris, fueron satisfactorios y se encontró que esta transformada puede ser utilizada como un modelo para caracterizar texturas de iris.

# Abstract

In the last years, diverse biometric technologies have been used and its development still continues. Each technology has advantages and disadvantages. One of the most confident biometric systems in applications demanding high security levels is iris recognition.

A typical iris recognition system requires an image representation model to extract the iris texture features. Once the features have been extracted, these are stored on a biometric template (code) associated to each person. The next stage, in order to recognize across the people registered in one of such systems, the registered biometric templates are matched against the ones presented by the user, who claims to be recognized by the system.

Frequently, iris recognition systems have used wavelets during the texture feature extraction process, especially the Gabor wavelet family. This family uses important properties found on some mammalian visual systems, fitting in shape to the receptive field profiles of cells at retinal and visual cortical levels, to extract the most important information from the images. Wavelets are constructed from the Gabor elementary functions, which are used to analyze a given input signal.

In this work an iris recognition system was developed in which the Hermite transform was used as an alternative tool for iris texture feature extraction. The Hermite transform can simulate some properties of the mammalian visual systems in a similar way as the Gabor model does. The Hermite transform is a local decomposition technique in which an input image is localized through a Gaussian window and then expanded into orthogonal polynomials with respect to such window, whose Hermite functions are solutions to the harmonic oscillator equation. The connection between the Hermite transform and the mammalian visual systems comes from the fact that the elementary functions used in the Hermite transform are similar to the Gaussian derivatives, which are able to model and fit the receptive field profiles of simple cells.

For comparison purposes, in this project Gabor wavelets were also used as an iris texture feature extractor model and some stages of the iris recognition system presented by Daugman, who has pioneered this application using the Gabor wavelet, were reproduced. Results obtained with the proposed iris recognition system, using the Hermite transform as an alternative iris texture feature extraction tool, were successful and it was found that this transform can be used to characterize iris textures.

# Índice general

Dedicatoria y agradecimientos	I
Agradecimientos especiales	III
Resumen	V
Abstract	VII
Índice de figuras	XIII
Lista de tablas	XV
Índice de algoritmos	XVII
Organización de la tesis	XIX
<b>1. Introducción</b>	<b>1</b>
1.1. Sistemas biométricos . . . . .	1
1.1.1. Reconocimiento de huellas dactilares . . . . .	3
1.1.2. Reconocimiento de rostro . . . . .	4
1.1.3. Reconocimiento de iris . . . . .	6
1.2. El sistema de visión humano . . . . .	8
1.2.1. Anatomía del sistema de visión . . . . .	8
1.2.2. La retina como una parte del sistema nervioso central . . . . .	10
1.2.3. Proyección de los campos receptivos a la corteza visual . . . . .	12
<b>2. Procesamiento digital de imágenes basado en sistemas de visión biológicos y el análisis de texturas</b>	<b>15</b>
2.1. Análisis de textura inspirado en modelos de visión biológicos . . . . .	15
2.2. Modelos de representación de imágenes inspirados en sistemas de visión biológicos . . . . .	17
2.2.1. Modelo detector de característica . . . . .	18
2.2.2. Modelo diferencia de Gaussiana (DOG) . . . . .	18
2.2.3. Combinación de Modelos DOG . . . . .	18
2.2.4. Modelo derivada de Gaussiana (DG) . . . . .	19
2.2.5. Modelo de Gabor . . . . .	19
2.2.6. Modelo de la función Hermite . . . . .	19
2.3. Wavelets Gabor . . . . .	20
2.3.1. Funciones Gabor 2D elementales y sus propiedades . . . . .	20
2.3.2. Diseño de las wavelets Gabor 2D . . . . .	22
Muestreo . . . . .	24
Selección de los coeficientes de los filtros Gabor . . . . .	25

2.4.	Transformada Hermite . . . . .	26
2.4.1.	Transformada polinomial . . . . .	27
2.4.2.	Transformada Hermite . . . . .	28
2.4.3.	Polinomios de Krawtchouk . . . . .	31
2.4.4.	Transformada Hermite 2D . . . . .	32
2.4.5.	Transformada Hermite 2D orientada . . . . .	32
<b>3.</b>	<b>Reconocimiento de Iris</b>	<b>35</b>
3.1.	El Iris del ojo humano: identificador biométrico robusto . . . . .	35
3.2.	Descripción de las diferentes etapas de un sistema típico de reconocimiento de iris . . . . .	36
3.2.1.	Adquisición de imágenes . . . . .	37
3.2.2.	Extracción de características y codificación . . . . .	38
3.2.2.1.	Localización del iris . . . . .	38
3.2.2.2.	Normalización o alineación . . . . .	41
3.2.2.3.	Análisis, extracción de características y codificación . . . . .	42
3.2.3.	Comparación entre códigos de iris ( <i>matching</i> ) . . . . .	45
3.3.	Sistema de reconocimiento de iris propuesto . . . . .	46
3.3.1.	Implementación del módulo de adquisición de imágenes de iris . . . . .	46
3.3.2.	Implementación del módulo de localización de iris . . . . .	47
	Obtención del mapa de bordes: operador Canny . . . . .	47
	Parametrización: transformada Hough . . . . .	50
3.3.3.	Implementación del módulo de normalización y eliminación de artefactos . . . . .	53
3.3.4.	Implementación del módulo extractor de características Hermite . . . . .	55
3.3.5.	Implementación del algoritmo de comparación ( <i>matching</i> ) . . . . .	57
<b>4.</b>	<b>Análisis de resultados</b>	<b>59</b>
4.1.	Descripción de las bases de datos de imágenes de iris utilizadas . . . . .	59
4.1.1.	Base de datos de imágenes de iris de nuestra propiedad . . . . .	59
4.1.2.	Base de datos de imágenes de iris CASIA . . . . .	59
4.2.	Teoría de la decisión estadística . . . . .	60
4.2.1.	Criterio Neyman-Pearson . . . . .	60
4.2.2.	Aplicación de la teoría de la decisión en el reconocimiento de iris . . . . .	62
4.3.	Resultados usando la base de datos de imágenes de iris de nuestra propiedad . . . . .	64
	Extractor basado en wavelets Gabor . . . . .	64
	Extractor basado en la transformada Hermite . . . . .	66
	Extractor basado en la transformada Hermite orientada . . . . .	69
	Comparación de desempeño entre ambos extractores de características . . . . .	70
4.4.	Resultados usando la base de datos de imágenes de iris CASIA . . . . .	72
	Extractor basado en wavelets Gabor . . . . .	72
	Extractor basado en la transformada Hermite . . . . .	73
	Extractor basado en la transformada Hermite orientada . . . . .	76
	Comparación de desempeño entre ambos extractores de características . . . . .	76
4.5.	Discusión de resultados . . . . .	78
<b>5.</b>	<b>Conclusiones y perspectivas</b>	<b>81</b>
<b>A.</b>	<b>Polinomios de Hermite</b>	<b>83</b>



B. Normalización $L^1$ y $L^2$ de la función Gaussiana	87
C. Obtención de las derivadas de Gaussianas como filtros de análisis de la transformada Hermite	89
D. Obtención de las funciones patrón como filtros de síntesis de la transformada Hermite	91
E. Algoritmos de la transformada Hough	93
Bibliografía	95

# Índice de figuras

1.1. Minucias de una huella dactilar . . . . .	4
1.2. Eigenfaces: vectores de características usando PCA . . . . .	5
1.3. Ejemplo de seis clases usando LDA . . . . .	6
1.4. Mapa de grafos elásticos . . . . .	6
1.5. Imagen de iris humano. . . . .	7
1.6. Corte del ojo derecho visto desde la parte superior . . . . .	9
1.7. Organización de la retina . . . . .	10
1.8. Distribución de los bastones y conos en la retina humana . . . . .	12
1.9. Organización de campo receptivo de las células ganglionares de la retina . . . . .	13
1.10. Vía visual principal vista desde la base del cerebro . . . . .	14
2.1. Componentes de un filtro Gabor 2D complejo. (a) Componente par (real). (b) Componente impar (imaginaria). . . . .	22
2.2. Familia de wavelets Gabor 2D . . . . .	24
2.3. Gráfica de la razón ( $r_s$ ) de la tasa de la frecuencia de muestreo mínima a dos veces la frecuencia central del filtro Gabor contra el ancho de banda de frecuencia radial $B$ . . . . .	25
3.1. Etapas de un sistema de reconocimiento de iris. . . . .	36
3.2. Adquisición de imágenes de iris . . . . .	37
3.3. Imagen de iris capturada. Se observan elementos no importantes como la esclerótica, la pupila y los párpados. . . . .	39
3.4. Localización de iris utilizando máscaras predefinidas. . . . .	40
3.5. Proceso de demodulación de la fase usado para codificar los patrones del iris. . . . .	44
3.6. Sistema de reconocimiento de iris propuesto. . . . .	46
3.7. Filtro paso banda en la región del NIR usado en el módulo de adquisición de imágenes del sistema de reconocimiento de iris propuesto. . . . .	47
3.8. Obtención del mapa de bordes normalizado usando el operador de Canny para la parametrización de la pupila y del iris. . . . .	48
3.9. Obtención de los mapas de bordes binarios a través de la umbralización y binarización por histéresis del operador de Canny . . . . .	50
3.10. Transformada Hough para delimitar las fronteras del iris . . . . .	51
3.11. El iris segmentado: resultado de la parametrización de la pupila y de la esclerótica como dos círculos no concéntricos. . . . .	53
3.12. Esquema del proceso de normalización . . . . .	54
3.13. Normalización y eliminación de párpados y pestañas . . . . .	55
3.14. Filtros de análisis usando la transformada Hermite 2D . . . . .	56
3.15. Descomposición Hermite $N = 6$ , $D_{max} = 2$ . . . . .	56
3.16. Código de iris usando la transformada Hermite de primer y segundo orden . . . . .	57
3.17. Proyección de la transformada Hermite cartesiana 2D a la transformada Hermite orientada 1D. . . . .	57

Índice de figuras

4.1. Distribuciones en un problema típico de decisión binaria. . . . .	61
4.2. Histogramas obtenidos al comparar la distancia de Hamming usando la transformada Hermite . . . . .	63
4.3. Ejemplo de tasas de error $FRR$ , $FAR$ y curva de característica de operación de receptor (ROC) usando la transformada Hermite para la base de datos de imágenes de iris de nuestra propiedad . . . . .	64
4.4. Evaluación del desempeño entre los extractores de características wavelet Gabor y transformada Hermite para la base de datos de imágenes de iris de nuestra propiedad.	71
4.5. Evaluación del desempeño entre los extractores de características wavelet Gabor y transformada Hermite para la base de datos de imágenes de iris CASIA. . . . .	77

# Lista de tablas

4.1. Tabla comparativa de los resultados obtenidos al variar el tamaño del código o plantilla biométrica y el número de bandas o regiones de análisis. Para un ancho de banda angular $B_\theta = \pi/4$ usando wavelets Gabor, el mejor resultado se observó con un código de 2048 bits y 8 bandas de análisis. . . . .	65
4.2. Tabla comparativa de los resultados obtenidos al variar el tamaño del código o plantilla biométrica y el número de bandas o regiones de análisis. Para un ancho de banda angular $B_\theta = \pi/6$ usando wavelets Gabor, el mejor resultado se observó con un código de 2048 bits y 8 bandas de análisis. . . . .	66
4.3. Tabla comparativa de los resultados obtenidos al variar el tamaño del código o plantilla biométrica y el número de bandas o regiones de análisis. Para un ancho de banda angular $N = 4$ usando la transformada Hermite, el mejor resultado se observó con un código de 2048 bits y 8 bandas de análisis. . . . .	67
4.4. Tabla comparativa de los resultados obtenidos al variar el tamaño del código o plantilla biométrica y el número de bandas o regiones de análisis. Para un ancho de banda angular $N = 6$ usando la transformada Hermite, el mejor resultado se observó con un código de 2048 bits y 8 bandas de análisis. . . . .	67
4.5. Tabla comparativa de los resultados obtenidos al variar el tamaño del código o plantilla biométrica y el número de bandas o regiones de análisis. Para un ancho de banda angular $N = 8$ usando la transformada Hermite, el mejor resultado se observó con un código de 2048 bits y 8 bandas de análisis. . . . .	68
4.6. Tabla comparativa de los resultados obtenidos al variar el tamaño del código o plantilla biométrica y el número de bandas o regiones de análisis. Para un ancho de banda angular $N = 10$ usando la transformada Hermite, el mejor resultado se observó con un código de 2048 bits y 8 bandas de análisis. . . . .	69
4.7. Tabla comparativa de los resultados obtenidos al variar el tamaño del código o plantilla biométrica y el número de bandas o regiones de análisis. Para un ancho de banda angular $N = 6$ usando la transformada Hermite orientada, el mejor resultado se observó con un código de 2048 bits y 8 bandas de análisis. . . . .	69
4.8. Tabla comparativa de los resultados obtenidos al variar el tamaño del código o plantilla biométrica y el número de bandas o regiones de análisis. Para un ancho de banda angular $N = 8$ usando la transformada Hermite orientada, el mejor resultado se observó con un código de 2048 bits y 8 bandas de análisis. . . . .	70
4.9. Tabla comparativa de los resultados obtenidos al variar el tamaño del código o plantilla biométrica y el número de bandas o regiones de análisis. Para un ancho de banda angular $B_\theta = \pi/4$ usando wavelets Gabor, el mejor resultado se observó con un código de 2048 bits y 8 bandas de análisis. . . . .	72
4.10. Tabla comparativa de los resultados obtenidos al variar el tamaño del código o plantilla biométrica y el número de bandas o regiones de análisis. Para un ancho de banda angular $B_\theta = \pi/6$ usando wavelets Gabor, el mejor resultado se observó con un código de 2048 bits y 8 bandas de análisis. . . . .	72

4.11. Tabla comparativa de los resultados obtenidos al variar el tamaño del código o plantilla biométrica y el número de bandas o regiones de análisis. Para un ancho de banda angular  $N = 4$  usando la transformada Hermite orientada, el mejor resultado se observó con un código de 2048 bits y 8 bandas de análisis. . . . . 73

4.12. Tabla comparativa de los resultados obtenidos al variar el tamaño del código o plantilla biométrica y el número de bandas o regiones de análisis. Para un ancho de banda angular  $N = 6$  usando la transformada Hermite orientada, el mejor resultado se observó con un código de 2048 bits y 8 bandas de análisis. . . . . 74

4.13. Tabla comparativa de los resultados obtenidos al variar el tamaño del código o plantilla biométrica y el número de bandas o regiones de análisis. Para un ancho de banda angular  $N = 8$  usando la transformada Hermite orientada, el mejor resultado se observó con un código de 2048 bits y 8 bandas de análisis. . . . . 75

4.14. Tabla comparativa de los resultados obtenidos al variar el tamaño del código o plantilla biométrica y el número de bandas o regiones de análisis. Para un ancho de banda angular  $N = 10$  usando la transformada Hermite orientada, el mejor resultado se observó con un código de 2048 bits y 8 bandas de análisis. . . . . 75

4.15. Tabla comparativa de los resultados obtenidos al variar el tamaño del código o plantilla biométrica y el número de bandas o regiones de análisis. Para un ancho de banda angular  $N = 6$  usando la transformada Hermite orientada, el mejor resultado se observó con un código de 2048 bits y 8 bandas de análisis. . . . . 76

4.16. Comparación de los resultados obtenidos con los reportados por otros autores . . . . . 79

# Índice de algoritmos

E.1. TransformadaHough parte 1. El algoritmo encuentra y acumula los posibles centros de los círculos. . . . .	93
E.2. TransformadaHough parte 2. El algoritmo selecciona las coordenadas del centro $(x_0, y_0)$ que recibió la mayor cantidad de votos. . . . .	93
E.3. TransformadaHough parte 3. El algoritmo acumula los radios posibles del círculo. . . . .	94
E.4. TransformadaHough parte 4. El algoritmo selecciona el radio del círculo $R_0$ que recibió la mayor cantidad de votos. . . . .	94

# 1 Introducción

Antes de comenzar con la descripción del sistema de reconocimiento de iris propuesto, es conveniente presentar brevemente algunos sistemas de reconocimiento que utilizan otros identificadores como el rostro o las huellas dactilares. En la sección 1.1 de este capítulo se introduce el término biometría, su estudio y las diferentes tecnologías existentes. Así mismo en dicha sección se abordan los tres sistemas de reconocimiento basados en imágenes que han sido utilizados con mayor frecuencia como son: el reconocimiento de huellas dactilares, de rostro y de iris. Muchas de las tecnologías que sirven para analizar texturas hacen uso de modelos matemáticos basados en sistemas de visión biológicos. Por esta razón, en la sección 1.2 se abordan los conceptos y algunos de los procesos que el sistema de visión de algunos mamíferos lleva a cabo durante el procesamiento de la información visual.

## 1.1. Sistemas biométricos

El término “biometría” se deriva de las palabras griegas “bios” (vida) y “metron” (medición) [64]. La biometría es un campo de investigación relativamente nuevo y estudia los métodos para el reconocimiento único de cada individuo a través de alguna característica intrínseca del individuo mismo. Surge del principio de que cada individuo tiene características únicas, a las cuales se les llama identificadores biométricos. La elección de un identificador biométrico dependerá de los siguientes factores: universal, es decir, toda persona debe poseer esa característica; único, porque dos o más personas no deben poseer los mismos rasgos; permanente o estable, ya que los rasgos no deben cambiar o ser alterados con el paso del tiempo; por último que sea obtenible, es decir, debe ser posible medir y cuantificar la característica con algún tipo de sensor.

La biometría aglutina diferentes áreas del conocimiento como las matemáticas, el procesamiento digital de señales e imágenes, la computación y la electrónica. Los sistemas biométricos automáticos han llegado a ser realizables en los últimos décadas gracias a los grandes avances en la computación. Entre las principales tecnologías biométricas se encuentran el reconocimiento de huellas dactilares, reconocimiento de la palma de la mano, el reconocimiento de rostro, el reconocimiento de voz o del hablante y el reconocimiento del iris.

Muchas de estas nuevas técnicas automáticas están basadas en las ideas que fueron originalmente concebidas hace cientos e incluso miles de años atrás. Una de las características básicas y de las más antiguas utilizada para el reconocimiento por los humanos ha sido el rostro. Desde el comienzo de las civilizaciones los seres humanos han utilizado al rostro para distinguir entre individuos conocidos e individuos desconocidos. El concepto de reconocimiento humano-humano es también visto en aspectos biométricos predominantes en el comportamiento de los individuos; como lo es el reconocimiento del habla y de la forma de caminar. Los individuos utilizan estas características día a día de forma casi inconsciente para reconocer individuos conocidos. Otras características han sido utilizadas a través de la historia de la civilización para el reconocimiento como las que se muestran a continuación [58]:

- En una caverna, con una antigüedad de aproximadamente 31,000 años, fueron descubiertas pinturas en las paredes de dicha caverna que fueron hechas por hombres prehistóricos. Alrededor de estas pinturas existen varias huellas de manos que posiblemente funcionaron como una firma de los creadores.

## 1 Introducción

- Existe también evidencia de que las huellas dactilares fueron utilizadas como marca de personas hace aproximadamente 500 años A.C. Las transacciones e intercambios de los babilonios están registrados en tablas de arcilla en las que se incluían huellas dactilares.
- El explorador João de Barros, en su libro “Décadas da Ásia” publicado en 1563, describe la manera en que los mercaderes chinos identificaban a los niños tomando las huellas de sus pies y manos con tinta sobre papel [36]. Cabe señalar que estos papeles eran contratos de venta de los menores.

A mediados del siglo XIX, con el rápido crecimiento de las ciudades debido a la revolución industrial, aumentó la necesidad de identificar a las personas. Los comerciantes y autoridades ya no podían confiar en sus propias experiencias y conocimiento local con poblaciones cada vez mayores y con habitantes que se trasladaban de una ciudad a otra con mayor facilidad. Los sistemas de justicia trataban de castigar a los infractores recurrentes de manera más severa que aquellos que cometían alguna infracción por primera vez. Esto creó la necesidad de formar un sistema que registrara las infracciones a través de ciertas características de los infractores. El primero de dos sistemas propuestos fue el sistema Bertillon en Francia y consistía en escribir sobre tarjetas las medidas de los infractores, las cuales podrían ser ordenadas por altura, longitud de brazo o cualquier otro parámetro. El segundo sistema fue el uso formal de las huellas dactilares por parte de la policía. Este proceso emergió en el sur de América, Asia, y Europa. A finales del siglo se desarrolló un método para indexar las huellas dactilares y era capaz de mostrar los registros como lo hacía el método Bertillon pero basado en el uso de una métrica más individualizada como los patrones de las huellas dactilares. El primer sistema robusto para indexar huellas dactilares fue el sistema Henry.

Los sistemas biométricos comenzaron a emerger en la última mitad del siglo XX coincidiendo con los avances de los sistemas de computo. A partir de 1990 hubo una gran actividad en el desarrollo de estos sistemas y desde principios del año 2000 y hasta nuestras fechas se les puede ver en diversas aplicaciones.

De manera general, un sistema biométrico automático está compuesto de cinco componentes integrados [58]:

- un sensor utilizado para recolectar los datos y convertir la información en un formato digital.
- algoritmos de procesamiento digital de señales encargados de generar plantillas biométricas.
- una base de datos que sirve para almacenar la información con la cual nuevas plantillas biométricas serán comparadas.
- un algoritmo de comparación (*matching*) que se encarga de comparar una nueva plantilla biométrica contra una o más plantillas biométricas que se encuentran dentro de la base de datos.
- finalmente un proceso de decisión, que puede ser automático o asistido por personal humano, utiliza los resultados del algoritmo de comparación para tomar una decisión.

Un sistema de reconocimiento automático puede ser usado para verificación o bien para identificación. El reconocimiento es un término genérico, el cual no implica necesariamente verificación o identificación. Todos los sistemas biométricos realizan reconocimiento de algún individuo registrado en el sistema con anterioridad. Las diferencias entre estos dos tipos de reconocimiento se describen a continuación [58]:



- Verificación, también conocida como autenticación, es la tarea donde el sistema biométrico trata de confirmar que la identidad que un individuo aclama pertenece realmente a dicho individuo; esto se hace comparando una muestra capturada por el sistema con una o más plantillas biométricas previamente registradas.
- Identificación es la tarea donde el sistema biométrico trata de determinar cual es la identidad de un individuo. Un identificador biométrico del individuo es recolectado y posteriormente comparado con todas las plantillas biométricas en la base de datos. La identificación puede determinar si el individuo pertenece a la base de datos, o bien que si está en la base de datos y además decir de que individuo se trata.

Cada tecnología biométrica, reconocimiento de voz, reconocimiento de huellas dactilares, reconocimiento de rostro, reconocimiento de iris, etc., tiene sus ventajas y limitaciones. Hasta el momento no hay alguna que sea mejor que las demás ya que muchos factores deben ser tomados en cuenta cuando se implementa alguna de éstas, por ejemplo: riesgos de seguridad, tarea (verificación o identificación), número de usuarios esperado, datos disponibles, circunstancias de lugar o bien del usuario, etc. Es posible también combinar varias tecnologías en un solo sistema y aumentar su confiabilidad. Sin embargo, se sabe que de todas las tecnologías mencionadas anteriormente, el reconocimiento de iris ha presentado mejores expectativas y resultados favorables en aplicaciones de alta seguridad.

A continuación se describen algunas de las tecnologías biométricas en las que se utilizan técnicas de procesamiento digital de imágenes para la obtención de las características y generación de las plantillas biométricas.

### 1.1.1. Reconocimiento de huellas dactilares

El reconocimiento de huellas dactilares es uno de los mejores ejemplos de tecnología biométrica conocida. Debido a que las huellas dactilares son únicas y consistentes a lo largo del tiempo (siempre y cuando éstas no sean alteradas físicamente en forma irreversible accidental o intencionalmente [71]), han sido utilizadas para identificación desde hace poco más de un siglo cuando Sir Francis Galton definió algunos de los puntos o características, con los cuales las huellas pueden ser identificadas, llamados “puntos de Galton”. El reconocimiento automático de huellas dactilares comenzó a finales de 1960, a la par del surgimiento de nuevas tecnologías de computo, cuando el FBI<sup>1</sup> tuvo la necesidad de desarrollar un sistema para automatizar el proceso de identificación de huellas dactilares, ya que éste había sido un proceso manual que requería muchas horas y personal para realizarlo. Posteriormente, en conjunto con el NIST<sup>2</sup>, se desarrollaron escáneres de huellas dactilares que extraían y clasificaban de manera automática las huellas dactilares.

Durante las siguientes décadas, el NIST se centró en el desarrollo de métodos para la digitalización automática de huellas dactilares impresas en papel, así como de los efectos de la compresión de imágenes en la calidad de éstas, en los procesos de clasificación, extracción y comparación. Esto produjo como resultado el desarrollo del algoritmo M40 como uno de los primeros que el FBI colocó en operación. La tecnología de huellas dactilares fue mejorando y posteriormente se extendió a otros estados así como a otros países, los cuales también desarrollaron sus propios sistemas, dando lugar a que algunas huellas que habían sido recolectadas por un sistema fueran incompatibles con otro sistema, por lo que hubo necesidad de proponer estándares para esta tecnología. Actualmente se cuentan con varios productos comerciales de verificación [58].

<sup>1</sup>FBI, por sus siglas en inglés “Federal Bureau of Investigation”, organismo de investigación del Departamento de Justicia de los Estados Unidos de América.

<sup>2</sup>NIST, por sus siglas en inglés “National Institute of Standards and Technology”, agencia de la Administración de Tecnología del Departamento de Comercio de los Estados Unidos de América.



Figura 1.1: Minucias de una huella dactilar [58].

Los patrones de una huella dactilar usualmente aparecen como una serie de líneas oscuras llamadas crestas en combinación de líneas claras llamadas valles. Estas estructuras proveen la información esencial para el reconocimiento. Convencionalmente se han utilizado una serie de terminaciones de crestas o de bifurcaciones, conocidas como “minucias”, como las características de los patrones de huellas digitales. En la figura 1.1 se muestran este tipo de características. Las minucias proveen información de discriminación muy compacta, sin embargo, los algoritmos basados en este tipo de características poseen varias desventajas. En primer lugar no es fácil obtener los puntos de minucia automáticamente y con precisión. En segundo lugar el número de puntos de minucia puede ser insuficiente, especialmente en sistemas que utilizan sensores que capturan huellas digitales con un tamaño pequeño. Para superar estos problemas, algunos esquemas proponen el uso de bancos de filtros direccionales cuyo fin es representar una huella dactilar en términos de energías direccionales [60]. Otros esquemas también proponen el uso de *wavelets*<sup>3</sup> Gabor [44]. De igual manera, existen sistemas de reconocimiento de huellas dactilares que pueden estar compuestos de una o más de técnicas por lo que se pueden formar plantillas con una combinación de distintas características [58].

### 1.1.2. Reconocimiento de rostro

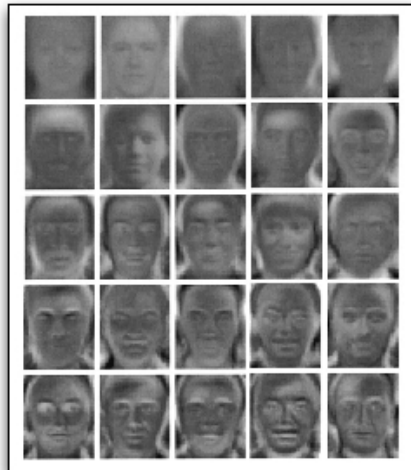
En los seres humanos se utilizan casi siempre los rostros para reconocerse entre si. Los nuevos avances en computo han permitido que en la actualidad reconocimientos similares puedan hacerse de manera automática. En un principio los algoritmos utilizaban simples modelos geométricos, pero en la actualidad estos procesos de reconocimiento utilizan representaciones matemáticas sofisticadas. El reconocimiento de rostro puede ser utilizado tanto para verificación como para identificación.

Existen dos técnicas que predominan en el problema de reconocimiento de rostro: geométrico (basado en características) y fotométrico (basado en el aspecto). Muchos algoritmos han sido desarrollados y reportados en la literatura entre ellos se puede mencionar [58]: el análisis de componentes principales (PCA, Principal Component Analysis), el análisis discriminante lineal (LDA, Linear Dis-

<sup>3</sup>En la teoría de análisis multiresolución introducida por S. Mallat [52], la wavelet es una técnica de descomposición de señales (imágenes en nuestro contexto). En dicha técnica, la señal se descompone en un conjunto completo de funciones, las cuales forman una base, usualmente una base ortogonal. Estas funciones son construidas trasladando y dilatando una función, llamada wavelet madre, la cual es localizada en ambos dominios frecuencial y espacial. En el idioma español se le conoce como ondeleta u onduleta. En el presente trabajo, se adopta el nombre en inglés *wavelet* para nombrar a este modelo.

criminant Analysis) y comparación basado en mapas de grafos elásticos (EBGM, Elastic Bunch Graph Matching).

- El esquema PCA permite reducir la dimensión de los datos a través de elementos básicos de compresión y revela la estructura más efectiva con una dimensión baja de los patrones del rostro. Esta reducción en dimensión elimina información que no es útil [9], descomponiendo la estructura del rostro en componentes ortogonales o eigenvectores (llamados “eigenfaces”) después de ser aplicado dicho esquema sobre un conjunto de imágenes. El objetivo de este esquema es encontrar el mejor conjunto de imágenes base, de tal manera que en esta base nueva los coeficientes PCA no se puedan predecir uno a partir de otro en forma lineal [5]. Cada una de estas imágenes de rostro puede entonces ser descrita en términos de la proyección de sus coeficientes en los eigenvectores generados como una suma ponderada (vector de característica) de eigenfaces, los cuales son almacenados en un arreglo unidimensional (Véase figura 1.2). Una imagen de prueba puede ser comparada contra una galería de imágenes midiendo la distancia que existe entre sus respectivos vectores característica. Este esquema normalmente necesita que se presente la parte frontal del rostro completa.



**Figura 1.2:** Eigenfaces: vectores de características usando PCA [5].

- El LDA es un esquema estadístico para clasificar muestras de clases desconocidas basado en muestras de entrenamiento de clases conocidas [9]. El objetivo es formar un subespacio que es linealmente separable entre clases. LDA encuentra un conjunto de vectores que maximiza el criterio discriminante de Fisher al tiempo que maximiza la varianza entre clases (i.e. entre los usuarios) y minimiza la varianza dentro de cada clase (i.e. de cada usuario) [59]. Mientras que PCA puede ser agrupado dentro de las técnicas de aprendizaje no supervisado, LDA es una técnica de aprendizaje supervisada ya que necesita información de la clase para cada imagen en el proceso de entrenamiento. Este método supera algunas de las limitaciones del método de eigenfaces al aplicar el criterio discriminante lineal de Fisher. Los problemas que se presentan con esta técnica se originan cuando existen pocas muestras de entrenamiento comparado con la dimensión del espacio total de muestras. En la figura 1.3 se muestran ejemplos de seis clases utilizando esta técnica.



Figura 1.3: Ejemplo de seis clases usando LDA [50].

- EBGM se basa en el concepto de que las imágenes de rostro reales poseen muchas características que son no lineales y que no pueden ser analizados usando los métodos lineales discutidos anteriormente, como variaciones de iluminación, posición del usuario al momento de capturar las imágenes o distintas expresiones en el rostro del individuo. En este esquema, una wavelet Gabor crea una arquitectura de liga dinámica que proyecta el rostro a una malla elástica [77]. Un nodo sobre la malla elástica describe el comportamiento de la imagen alrededor de un píxel dado, como se muestra en la figura 1.4. Este nodo es un vector característica formado a partir de las respuestas de los filtros Gabor (a distintas escalas y direcciones de análisis) sobre un mismo punto<sup>4</sup>, dichas respuestas son obtenidas a través de la convolución de la imagen con los filtros. Este nodo es usado para detectar formas y extraer características y el reconocimiento se basa en la medida de similitud de las respuestas del filtro Gabor en cada nodo Gabor.

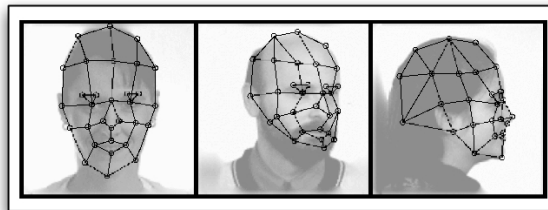


Figura 1.4: Mapa de grafos elásticos [77].

### 1.1.3. Reconocimiento de iris

El reconocimiento de iris es el proceso de reconocer a una persona analizando el patrón aleatorio del iris. Los sistemas automáticos de reconocimiento de iris son relativamente nuevos. El iris es un músculo que tiene la función de regular el tamaño de la pupila controlando la cantidad de luz que entra al ojo, cuyo color depende de la cantidad de melanina que pigmenta a dicho músculo. A pesar de que la coloración del iris está generalmente ligado a factores genéticos, los detalles de los patrones se deben a factores aleatorios.

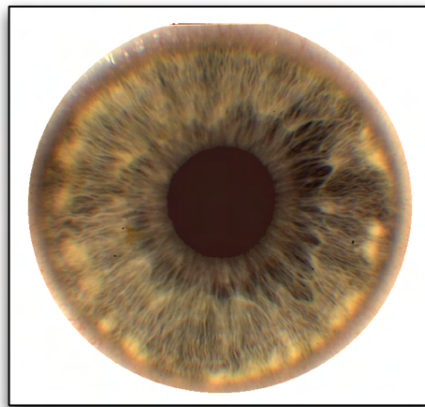
La posibilidad de que el iris de un ojo pueda ser usado como un tipo de huella digital para identificación personal fue sugerido primeramente por oftalmólogos [2][32], quienes notaron a partir de la

<sup>4</sup>Al vector característica formado a partir de las salidas de los filtros a distintas escalas y direcciones de análisis sobre un mismo punto, se le ha llamado en inglés “jet”.

experiencia clínica que cada iris tenía una textura altamente detallada y única que permanecía sin cambios en fotografías médicas tomadas durante décadas (contrario a los diagnósticos ocultos que postula la “iridología”).

En 1993 la agencia de defensa nuclear, nombrada a la fecha como DTRA<sup>5</sup>, comenzó a trabajar y a desarrollar una unidad prototipo, la cual fue completada satisfactoriamente gracias al esfuerzo en conjunto de Flom, Safir y Daugman. En 1994 Daugman patentó sus algoritmos automáticos de reconocimiento de iris y en 1995 los primeros productos comerciales surgieron. En 2005 la patente que cubría el concepto básico de reconocimiento de iris expiró, dando oportunidades de mercado a otras compañías que habían desarrollado sus propios algoritmos para reconocimiento de iris. Sin embargo, la patente de la implementación de los códigos de iris (IrisCodes) desarrollada por Daugman expirará hasta el 2011 [42].

Además de las características visibles en un iris, como las que pueden verse en la figura 1.5, existe una red trabecular compuesta de tejido conectivo, fibras colágenas, cuerpo ciliar, surcos, anillos, corona, coloración y manchas. La red trabecular estriada de células de fibroblastos y vesículas contráctiles de pigmentación crea la textura predominante bajo la luz visible, sin embargo, todas las fuentes de variaciones radial y angular tomadas en conjunto constituyen una “huella digital” distintiva que puede ser capturada a cierta distancia de la persona [19]. Otras propiedades del iris que refuerzan su uso en identificación automática son: 1) siendo éste un órgano interno del ojo, ubicado detrás de la cornea y del humor acuoso, presenta un aislamiento y protección propia al ambiente externo; 2) la imposibilidad de modificarlo a través de cirugías sin riesgos secundarios para la visión y 3) su respuesta fisiológica a la luz, provee de una prueba natural contra algún artificio.



**Figura 1.5:** Imagen de iris humano.

En el sistema propuesto por Daugman [19], el iris debe ser localizado y aislado con precisión del resto de la imagen. La localización del iris es sumamente importante ya que pueden resultar objetos extraños considerados como ruido de realizarse incorrectamente. Ejemplo de estos objetos son: pestañas, reflejos, pupila y párpados. Después de la localización y extracción del iris es posible obtener sus características. Daugman [19] propone el uso de las wavelets Gabor 2D como modelo de análisis de la textura del iris para la extracción de características; filtrando y mapeando los segmentos de iris en fasores que incluyen información de orientación, frecuencia espacial y posición de estas áreas en la imagen. Esta información es usada para generar los códigos de iris y de esta manera, los patrones

<sup>5</sup>DTRA, por sus siglas en inglés “Defense Threat Reduction Agency”, organismo del Departamento de Defensa de los Estados Unidos de América.

del iris quedan descritos usando información de la fase obtenida a partir de los fasores. Se utiliza a la fase como característica ya que ésta no es afectada por el contraste, la ganancia de la cámara utilizada o por los niveles de iluminación.

El reconocimiento de iris se lleva a cabo comparando dos códigos de iris. La diferencia entre los dos códigos de iris es usada como una prueba de independencia estadística. En el caso del algoritmo de Daugman [21], los códigos son de 256 bytes y se utiliza la distancia de Hamming como medida de similitud. Si la distancia de Hamming indica que menos de una cierta cantidad pequeña de bytes en el código de iris son diferentes, entonces el código de iris no pasa la prueba de *significancia* estadística, indicando que ambos códigos de iris pertenecen a una misma persona.

En la capítulo 3 se describirán con mayor detalle los procesos de un sistema típico de reconocimiento de iris y de los algoritmos que han sido utilizados comúnmente. Así mismo se detallarán los algoritmos utilizados para la realización de este proyecto usando la transformada Hermite para el análisis de la textura del iris.

## 1.2. El sistema de visión humano

Diversos algoritmos de procesamiento digital de imágenes han sido propuestos y se han basado en los modelos de visión biológicos de primates y de seres humanos. Es conveniente realizar una revisión de manera general de los órganos y procesos involucrados en el procesamiento de la visión para poder entender algunos de los modelos utilizados a la fecha.

El procesamiento de la visión comienza en el ojo cuando la luz, que viaja a través de éste hasta la retina, es disminuida, refractada, absorbida y convertida por varias estructuras en el ojo. La luz al acercarse al ojo entra en contacto primeramente con la cornea. La cornea refracta la luz provocando que la imagen converja en su camino hacia el iris y la pupila. Dependiendo de la intensidad de la luz, el iris se contrae o expande para ajustar el tamaño de la pupila. En situaciones de baja intensidad luminosa la pupila será más grande, permitiendo el paso de la luz suficiente para formar una imagen discernible. En condiciones opuestas la pupila será más pequeña ya que el exceso de luz genera imágenes pobres.

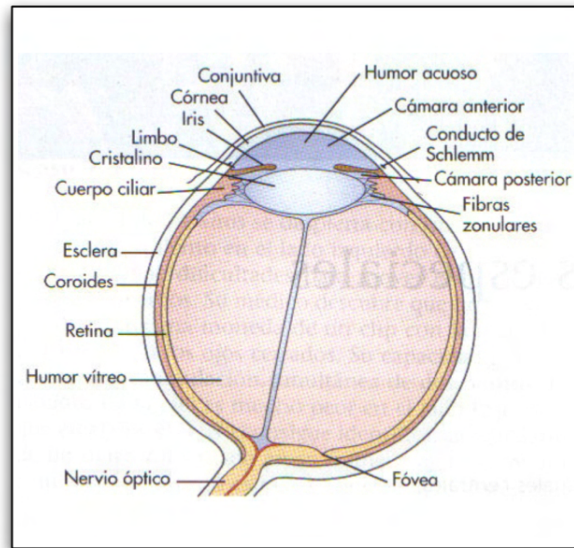
Una vez que la luz atraviesa la pupila, la luz es recibida por la lente o cristalino. Con la ayuda de músculos auxiliares, la lente puede cambiar su forma y dependiendo de la forma que ésta tome, los objetos pueden ser enfocados adecuadamente a varias distancias. La lente mejora la imagen tomada por la cornea y la proyecta sobre la retina. La retina captura la luz a través de sus fotorreceptores y células epiteliales pigmentarias. Las moléculas de fotopigmento del fotorreceptor absorben la luz provocando un cambio de potencial en la membrana del fotorreceptor. Con esto se inicia una serie de señales que viajan a través de las neuronas de la retina hacia el nervio óptico para llegar al cerebro. La señal es entonces recibida y procesada por el cerebro como una imagen interpretable.

### 1.2.1. Anatomía del sistema de visión

#### La cornea

Como se mencionó anteriormente, la cornea es el primer cuerpo con el cual la luz choca en su camino hacia adentro del ojo, como se ve en la figura 1.6. La cornea está compuesta de un cuerpo transparente que consiste de un epitelio, una lámina elástica homogénea y una sola capa de células de endotelio. La cornea es el contribuidor primario que enfoca la luz hacia la retina. Siguiendo las leyes principales de la refracción, cuando la luz incide sobre un medio que posee un índice de refracción mayor que el del aire, la velocidad de propagación disminuye y la luz cambia de dirección. La cornea podría ser un ejemplo de ese medio, ya que cuando la luz alcanza la superficie de la cornea, esta última disminuye

la velocidad de propagación de la luz y la hace converger hacia el centro del ojo, de tal manera que la imagen proyectada desde afuera es reducida en tamaño.



**Figura 1.6:** Corte del ojo derecho visto desde la parte superior [6].

### La pupila y el iris

La luz debe viajar a través del humor acuoso, el cual es un cuerpo líquido que llena el espacio entre la cornea y la lente o cristalino, para alcanzar al siguiente grupo de estructuras: el iris y la pupila, véase figura 1.6. Estas dos estructuras trabajan como reguladores de la cantidad de luz que entra al sistema. El iris es un diafragma pigmentado que se encuentra en frente de la lente, compuesto de fibras musculares lisas de orientación radial y circular que constituyen los músculos dilatador y esfínter de la pupila, que tienen la función de contraerse o dilatarse para regular la apertura del ojo, llamada pupila. En caso de que exista demasiada luz, el iris disminuye la apertura de la pupila con la ayuda del músculo esfínter, ya que demasiada luz podría provocar opacidad al momento de la formación de la imagen. Lo opuesto ocurre cuando la luz es pobre; la pupila es dilatada para tratar de capturar tantos fotones como sea posible para la formación adecuada de la imagen.

### La lente

Una vez que la cantidad correcta de luz ha entrado al ojo a través de la pupila, se encuentra la lente o cristalino como se puede apreciar en la figura 1.6. La lente está compuesta de una capa de epitelio cubriendo una masa de fibras. Esta capa está primeramente compuesta de proteínas llamadas cristalinas con las que se refina la imagen que llega desde la cornea. Las moléculas de la lente, al igual que las de la cornea, se encuentran densamente y uniformemente espaciadas. Esto es necesario para lograr su transparencia. La lente tiene un índice de refracción mayor que el de la cornea y que el del humor acuoso. Además del índice de refracción alto que la lente posee, también tiene la habilidad de cambiar su grado de refracción con la ayuda de los músculos ciliares [49]. La lente se sostiene y se mantiene en posición gracias a las fibras zonulares o ligamentos de suspensión. Estos ligamentos o fibras están unidos al área de los músculos ciliares que rodea al ojo llamado cuerpo ciliar. El cuerpo

## 1 Introducción

ciliar y las fibras zonulares trabajan en conjunto para alterar el punto focal del ojo. Se le llama acomodación al proceso por el cual la contracción del músculo ciliar hace que la lente adquiera una forma más redondeada. Como resultado de la acomodación las imágenes son enfocadas y enviadas a la retina.

### 1.2.2. La retina como una parte del sistema nervioso central

La luz procedente de la lente es enviada a la retina. La retina puede considerarse como un centro de procesamiento de la luz, donde las señales de luz son transformadas en señales neuronales que pueden ser percibidas y procesadas por el cerebro. La parte funcional de la retina cubre toda la porción posterior del ojo, excepto la mancha ciega, que es una salida del nervio óptico llamada papila o disco óptico. La agudeza visual es máxima en la parte central de la retina llamada la mancha amarilla o mácula lútea, sobre todo en la fovea, una depresión con forma de fosa en el centro de la mácula donde se enfocan los objetos de mayor interés visual; es el centro de fijación o de enfoque cuando los rayos de luz se dirigen hacia el objetivo visual correspondiente. La fovea contiene la densidad más alta de fotorreceptores [38].

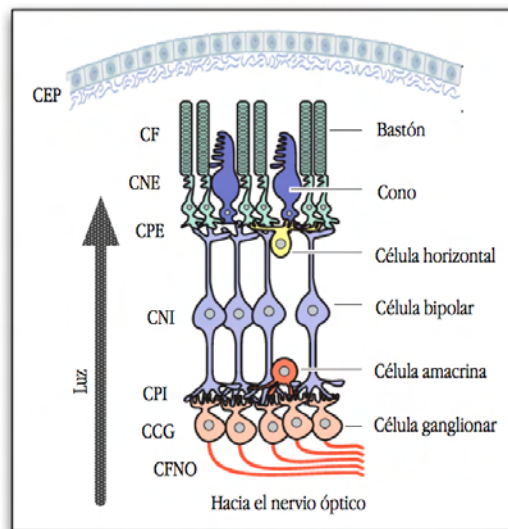


Figura 1.7: Organización de la retina [37, 38].

### Organización de la retina

La retina puede ser dividida en varias capas externas e internas. La capa que se encuentra hasta el fondo de la retina es la capa de epitelio pigmentario (CEP), la cual ayuda a evitar que la luz se disperse por todo el globo ocular permitiendo una visión nítida [37] (Véase figura 1.7). Esta capa además otorga el soporte a las células fotorreceptoras que se encuentran arriba de esta capa. La capa de los fotorreceptores (CF) consiste de los segmentos interno y externo de los conos y bastones, los cuales reciben y transforman los fotones de la luz. El núcleo de estos fotorreceptores se encuentra en la capa nuclear externa (CNE) mientras que sus axones y terminaciones celulares se encuentran en la capa plexiforme externa (CPE). La capa nuclear interna (CNI) contiene el núcleo de las células bipolares, a las células horizontales y a la mayor parte de las células amacrinas. La capa plexiforme interna (CPI) es el lugar donde se lleva a cabo la comunicación vertical entre las células bipolares y ganglionares,



haciendo un segundo contacto sináptico. En la última capa, la capa de células ganglionares (CCG), se encuentran los cuerpos de las células ganglionares. Las dendritas de estas células se extienden hacia la CPI, mientras que sus axones lo hacen en la dirección opuesta, hacia la capa de fibras del nervio óptico (CFNO) [49].

Ahora que han sido vistas las diferentes capas en que se divide la retina, es conveniente revisar las células que intervienen en el procesamiento de la luz, desde la captación de ésta hasta su transformación y envío al cerebro a través del nervio óptico.

### Conos y bastones

Las células fotorreceptoras son las responsables del inicio de una serie de eventos en cascada para convertir una imagen proyectada sobre el tejido de la retina, de fotones a señales electroquímicas que pueden ser procesadas por el cerebro. A esta conversión se le denomina fototransducción. Los dos tipos de células involucradas en este proceso son las células fotorreceptoras llamadas bastones y conos.

De los 130 millones de fotorreceptores que poseen las retinas de los seres humanos, aproximadamente cerca de 120 millones son estructuras largas y cilíndricas conocidas como bastones [49]. Los bastones son altamente sensibles a la luz y por ello están especializados en la detección de luz de baja intensidad (visión escotópica), siendo homogéneos en su sensibilidad de longitud de onda. Los conos son usualmente más cortos que los bastones y están especializados en la detección de intensidades altas que registran el color y detalles finos; individualmente son más sensibles a longitudes de onda largas (conos L), medias (conos M) y cortas (conos C) [38] (visión fotópica).

El elemento más crítico en este proceso de fototransducción es el fotopigmento que contienen los conos y bastones; una proteína sensible a la luz llamada opsina. En los bastones, cuando un solo fotón es absorbido por la opsina, se genera un cambio de potencial que genera a su vez un impulso eléctrico. Los conos poseen tres distintos tipos de opsina que son los responsables de los tres colores primarios (rojo, azul, amarillo). Como se mencionó anteriormente, los conos son menos sensibles por lo que necesitan una longitud de onda muy específica para iniciar un impulso eléctrico.

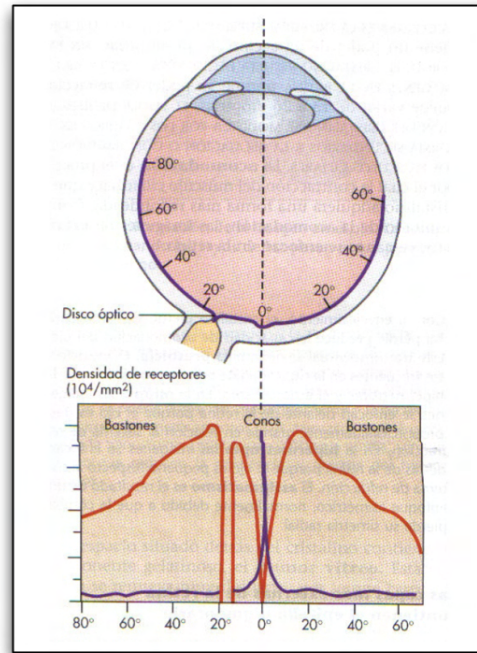
La fovea, la región más sensible y de mayor agudeza de la retina, contiene exclusivamente conos. La densidad de los conos decrece al aumentar la distancia a la fovea. Los bastones por su parte, no están presentes en la fovea y se encuentran mayormente concentrados en la parafovea [6]. En la zona del disco óptico, que es el lugar donde se juntan los axones de las células ganglionares para abandonar el ojo y formar el nervio óptico, no existen fotorreceptores y constituye un punto ciego. La distribución y proporción de los conos y bastones varía entre las especies de animales. En la figura 1.8 se muestra un gráfico de la densidad de conos y bastones en diferentes partes de la retina.

### Células bipolares

El siguiente conjunto de neuronas son las llamadas células bipolares, cuya función es propagar en la dirección vertical la comunicación. El cuerpo de estas células reside en la CNI de la retina (Véase figura 1.7) mientras que sus dendritas reciben las señales de los fotorreceptores en la primera unión sináptica. En la terminación opuesta del cuerpo de la célula, la señal viaja a través del axón de la célula bipolar para realizar una segunda sinapsis con la siguiente neurona vertical, la célula ganglionar.

### Neuronas laterales: células horizontales y amacrinas

Las señales eléctricas que fluyen a través de las neuronas verticales no son completamente independientes entre sí, ya que muchas están ligadas por neuronas laterales. Un tipo de estas neuronas



**Figura 1.8:** Distribución de los bastones y conos en la retina humana [6].

laterales son las células horizontales que se encuentran en la CNI de la retina (Véase figura 1.7). Comúnmente estas células están ligadas a más de un fotorreceptor, lo que significa que las células bipolares subsecuentes reciben señales de más de un fotorreceptor. Estas células hacen conexiones inhibitorias con las células bipolares cuando hay luz en su alrededor; en otras palabras, estas células tienen como función la variación del contraste.

Por su parte, las células amacrinas forman conexiones entre neuronas en la dirección vertical en la capa interna y algunas veces con las células ganglionares. La función de estas células no es clara aún pero se cree que contribuyen al efecto de contraste.

### Células ganglionares

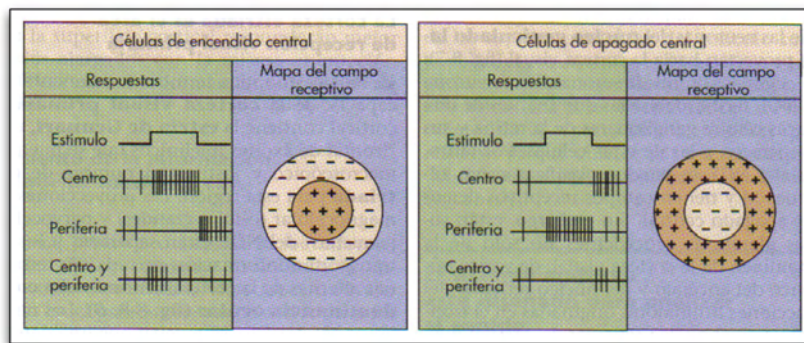
Las últimas neuronas de la red que reciben las señales son las células ganglionares de la retina. Cuando se activan debido a la llegada de una señal, la célula ganglionar produce un potencial de acción que comienza su travesía hacia los axones de la célula. Los axones de las células ganglionares de la retina convergen, formando el nervio óptico. El nervio óptico, compuesto de un manojo de fibras ópticas, representa el camino de las señales que se dirigen hacia el cerebro.

### 1.2.3. Proyección de los campos receptivos a la corteza visual

#### Campos receptivos

En cada una de las estructuras descritas anteriormente, una célula individual responde a imágenes en una pequeña parte del campo de visión y sólo responde fuertemente a patrones particulares de la imagen. La parte del campo de visión a la cual una célula responde se le llama el campo receptivo de la célula. A la relación entre los patrones de la imagen en el campo receptivo y la actividad de la célula es referida como las propiedades del campo receptivo de la célula [38].

El campo receptivo de un fotorreceptor suele ser una pequeña área circular que coincide con el área de la retina ocupada por éste. Las células bipolares y ganglionares de muchos mamíferos tienen una organización de campo receptivo antagonista concéntrica formado por dos áreas llamadas centro y periferia respectivamente. Las células bipolares y ganglionares pueden ser de dos tipos: de encendido central (también conocidas como células ON) o de apagado central (o células OFF). Estas células son excitadas por un estímulo en el centro de dicho campo receptivo e inhibido por otro estímulo en el área que rodea al centro llamada periferia (Véase figura 1.9). Una célula de tipo ON se despolariza cuando la luz incide en el centro de su campo receptivo y se hiperpolariza cuando lo hace en un anillo alrededor del centro de dicho campo. Un célula de tipo OFF se comporta de la manera contraria. En otras palabras, una célula de tipo ON responde a un contraste luz-obscuridad entre el centro y la periferia de su campo receptivo mientras que una célula de tipo OFF lo hace cuando existe un contraste oscuridad-luz entre el centro y la periferia del campo receptivo.



**Figura 1.9:** Organización de campo receptivo de las células ganglionares de la retina. Lado izquierdo: célula de encendido central (ON). Lado derecho: célula de apagado central (OFF) [6].

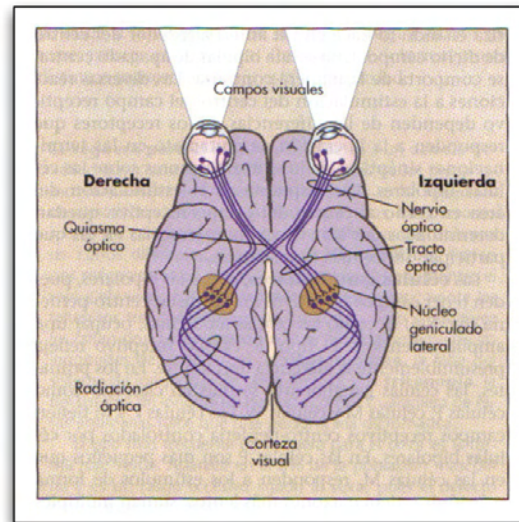
En las células bipolares, las respuestas al estímulo del área en torno al centro del campo receptivo quedan determinadas por las vías de interneuronas en las que participan las células horizontales [6]. Las salidas de estas células horizontales siempre son inhibitorias, lo que provoca que se mejore el contraste [37].

Existen tres tipos de células ganglionares las cuales se clasifican de acuerdo a las funciones y extensión de sus campos receptivos [37]. Las células ganglionares de tipo W constituyen el 40% del total de las células ganglionares. La excitación que estas células reciben proviene mayormente de las células bipolares conectadas sinápticamente con los bastones. Las células ganglionares de tipo W tienen un campo receptivo amplio ya que sus dendritas se extienden por toda la capa plexiforme interna y parecen estar especializadas en la detección de dirección de movimiento en el campo de visión. El grupo de células ganglionares de tipo X es el más numeroso y representa un 55% del total de las células ganglionares. Estas células tienen un campo receptivo pequeño debido a que sus dendritas no se extienden ampliamente en la retina, sino que solamente representan áreas discretas de la retina, transmitiendo los detalles finos de la imagen. Debido a que cada célula X recibe una señal de por lo menos un cono, la transmisión de una célula X es probablemente la responsable de toda la visión a color. El último tipo de células ganglionares y el menos numeroso, representando el 5% del total de células ganglionares, es el tipo Y. Estas células poseen campos receptivos extensos y responden a cambios rápidos en la imagen visual y movimientos rápidos o cambios rápidos en la intensidad de la luz.

## Corteza visual

Los nervios ópticos procedentes de ambos ojos convergen en el quiasma óptico. Algunas de las fibras de este nervio se cruzan en el quiasma y se unen al tracto óptico mientras que otras continúan su camino en sentido posterior en el tracto homolateral al ojo de origen. Como resultado de esta organización, cada tracto óptico contiene fibras cruzadas y no cruzadas. Las fibras de los tractos ópticos establecen sinapsis en el núcleo geniculado lateral (NGL) [6]. La mayoría de las neuronas del NGL proyectan la información visual hacia la corteza visual; sin embargo, algunas son interneuronas. En los primates, casi en un 90% de estas proyecciones se llevan a cabo sobre el NGL, donde los axones de las células ganglionares son agrupados en capas dependiendo del ojo de origen y otras propiedades (Véase figura 1.10). El NGL sirve de relevo y envía la información hacia la corteza visual primaria (V1). Las neuronas que hacen la sinapsis en el NGL, llamadas neuronas de relevo, poseen las mismas propiedades que las ganglionares W, X y Y respectivamente [38]. Los axones de las neuronas de relevo, que representan cada ojo son agrupados en una forma alternada, dando lugar a una banda llamada radiación óptica que viaja hacia la V1. Antes de llegar a la V1, los axones de las neuronas de relevo hacen contactos sinápticos con diferentes capas córticas.

En la corteza visual, existen neuronas en algunas capas que tienen propiedades de campo receptivo muy diferentes a las que se encuentran en el NGL. Estas células responden mejor a bordes o barras a una orientación en particular; estos bordes orientados son una característica importante para el sistema nervioso ya que frecuentemente los bordes definen los contornos de los objetos. V1 tiene dos tipos de neuronas sensibles a la orientación: células simples y células complejas [41]. Los campos receptivos de células simples tienen regiones separadas que responden a variaciones de luz, respondiendo entonces a barras o bordes en una posición en particular en el espacio. Las células complejas en contraste, responden a la presencia de una barra localizada en cualquier lugar dentro de su campo receptivo y no tienen regiones específicas que puedan ser estimuladas por cambios de luz. Estas evidencias de Hubel y Wiesel provienen de estudios anatómicos y fisiológicos realizados en la corteza visual primaria de gato.



**Figura 1.10:** Vía visual principal vista desde la base del cerebro [6].

## 2 Procesamiento digital de imágenes basado en sistemas de visión biológicos y el análisis de texturas

Un modelo de representación de imágenes, inspirado en el sistema de visión de algunos mamíferos y que ha sido mayormente utilizado en distintas tareas de procesamiento digital de imágenes, es el basado en las funciones Gabor. A partir de estas funciones se construyen bancos de filtros a distintas orientaciones y frecuencias de análisis con los cuales se filtra la imagen [17].

Otro modelo poco conocido y que puede ser utilizado en la discriminación de texturas se basa en la transformada Hermite. La transformada Hermite, originalmente introducida por Martens [55] al área de procesamiento digital de imágenes, es una técnica de descomposición local, en la cual una imagen es localizada a través de una ventana Gaussiana y proyectada sobre bases ortonormales respecto a dicha ventana de análisis. La conexión entre la transformada Hermite y los modelos de visión de mamíferos reside en que los filtros o funciones elementales usadas en dicha transformada son similares a las derivadas de Gaussiana (excepto por un factor de escala de la Gaussiana). Young [81] fue el primero en descubrir la similitud con respecto a la forma entre las funciones unidimensionales de las derivadas de Gaussianas y la forma de los campos receptivos de células simples en primates, basándose en el análisis de un número grande de espectros unidimensionales en el dominio frecuencial obtenidos experimentalmente [23].

En este capítulo se introducen conceptos de textura y algunos problemas relacionados con la discriminación de ésta (sección 2.1). Así mismo, se revisan algunos de los modelos inspirados en sistemas de visión biológicos que han sido propuestos, con especial énfasis en los modelos: wavelets Gabor en la sección 2.3 y transformada Hermite en la sección 2.4. Con respecto al diseño de los filtros Gabor para una implementación discreta, en la sección 2.3.2, se menciona una posible aproximación realizada por Bovik et al., [10], basada en la determinación de densidades de submuestreo mínimas de la forma continua de dichos filtros. En la sección 2.4.3 se presentan los filtros de Krawtchouk, los cuales son utilizados en la implementación discreta de las derivadas de Gaussiana que corresponden a los filtros de análisis en la transformada Hermite. En la sección 2.4.4 se aborda la transformada Hermite 2D y en la sección 2.4.5 a la transformada Hermite 2D orientada.

### 2.1. Análisis de textura inspirado en modelos de visión biológicos

Una textura puede ser definida como un arreglo local de patrones de irradiancia proyectados desde una área de una superficie, los cuales son distinguibles entre sí por concentraciones de frecuencia espacial y orientación de los patrones [10]. El análisis de textura en el procesamiento digital de imágenes puede ser utilizado de distintas maneras para cumplir con una tarea en particular, comúnmente en procesos automatizados basados en visión por computadora, a saber de manera general: en la segmentación de escenas en distintos objetos y regiones, en la discriminación, clasificación y reconocimiento de superficies de materiales, etc. En este análisis es indispensable la identificación de aquellos atributos que caractericen o representen mejor a la textura [10].

La segmentación de imágenes basada en la textura consiste en dividir de manera precisa una imagen en diferentes regiones que comparten una misma característica o textura. Requiere el uso simultáneo de ciertas mediciones tanto en el dominio espacial como en el dominio frecuencial [73]. Dentro de los métodos utilizados para la segmentación de texturas se pueden mencionar aquellos basados en primitivas locales geométricas, características locales estadísticas, modelos de campos aleatorios, y fractales. Por otro lado, los modelos basados en algunas propiedades de los sistemas de visión biológicos, como las wavelets Gabor, han alcanzado un auge mayor.

El problema de la discriminación de texturas puede ser confundido con el problema de la segmentación basado en texturas y aunque a veces un problema de segmentación se reduce a un problema de discriminación [10], en el caso de reconocimiento de iris y en otras tareas son problemas distintos. En el reconocimiento de iris por ejemplo, la etapa de segmentación sirve para localizar la región de interés y segregarla de aquellas áreas donde no existe información útil (pupila, esclerótica, párpados, etc); mientras que, para lograr discriminar a un iris perteneciente a un individuo en forma confiable de un iris que pertenece a un segundo individuo, es indispensable obtener cierta información local y global del comportamiento de la textura que conforma al iris, sin que esto implique realizar segmentaciones, sino únicamente la extracción de características mediante algún modelo ya sea puramente matemático o un modelo matemático basado en elementos de psicovisión.

En general, tanto la discriminación como la segmentación de texturas pueden realizarse a través de métodos supervisados o no supervisados. Los métodos supervisados involucran un conocimiento previo del tipo de textura a analizar generalmente obtenido en una etapa de preprocesamiento de la imagen [10]; por ejemplo: usando filtros Gabor se diseña uno o muy pocos de estos filtros para una aplicación en particular seleccionando aquellos filtros de acuerdo a un conocimiento previo de la textura. Los métodos no supervisados, en el caso del uso del filtros Gabor, involucra el diseño de un banco de filtros a distintas frecuencias y a distintas orientaciones, con el cual las características de la textura son extraídas. Uno de los problemas encontrado en el uso de los bancos de filtros Gabor es la selección de los parámetros libres, ya que una selección incorrecta de dichos parámetros podría resultar en un conjunto muy grande de filtros Gabor [73].

En el caso particular del reconocimiento de iris, se sabe *a priori* que los patrones de una textura de iris corresponden a cuerpos ciliares y tejido trabecular extendiéndose radialmente hacia adentro del ojo [21]. Es por ello que se sugiere el uso de un filtro, que represente mejor las características de la textura en dicha dirección. Dicho filtro puede seleccionarse a partir de un banco de filtros al observar la capacidad de discriminación que cada filtro entregue.

Una vez que una textura ha sido segmentada y/o se han discriminado ciertos patrones relacionados a sus componentes esenciales, es común que a través de algún tipo de aprendizaje, usando técnicas estadísticas o bien técnicas de inteligencia artificial, sea posible en una etapa posterior o puesta en marcha del sistema, reconocer los tipos de texturas u objetos que se le presenten para su manejo adecuado o clasificación. Para este propósito, normalmente se generan códigos de características. Una gran cantidad de información puede obtenerse a partir de los coeficientes obtenidos de filtrar una imagen con un filtro o un banco de filtros. Los códigos de características pueden ser formados de diversas maneras: tomando la media y desviación estándar de la magnitud de los coeficientes [53], usando la magnitud de los coeficientes [10], usando la magnitud de la componente real [43], creando momentos basados en el plano frecuencial [7], usando una medida de varianza local de un filtro para caracterizar la complejidad de la textura [15], etc.

El uso de las wavelets Gabor ha sido común en el análisis de texturas. La habilidad de las funciones Gabor en la discriminación de texturas parece estar relacionada al conjunto óptimo de la resolución espacio-frecuencia y a su capacidad de modelar la respuesta de células córticas (células simples) observadas durante el procesamiento de señales visuales. La descomposición wavelet de las funciones

elementales Gabor ha sido utilizada en problemas de inspección industrial; como en la detección de fallas en los procesos de fabricación de papelería, madera y materiales textiles [4][30]. Así mismo ha sido utilizado mayor frecuencia en tareas de reconocimiento basadas en identificadores biométricos como: reconocimiento de rostro [68][77], de huellas dactilares [85][44] y de iris [19][47][67] (así como también su variante log-Gabor [3]). En particular, uno de los primeros sistemas de reconocimiento de iris que apareció en la literatura, ha utilizado las wavelets Gabor en la etapa de extracción de características de la textura del iris [19].

Por otro lado, la transformada Hermite ha sido utilizada en aplicaciones de reducción de ruido [28] codificación y estimación de movimiento para el caso de secuencias de imágenes [29], así como también en fusión de imágenes [27]. Sin embargo, la transformada Hermite ha sido utilizada recientemente también en la caracterización de texturas e indexación [66] y ha arrojado buenos resultados.

A continuación se abordan con mayor profundidad los modelos de representación de imágenes inspirados en algunas propiedades importantes de los sistemas de visión biológicos, como las wavelets Gabor y la transformada Hermite, que pueden ser utilizados en la discriminación de texturas de iris.

## 2.2. Modelos de representación de imágenes inspirados en sistemas de visión biológicos

Existen varios modelos de visión que han sido propuestos para el procesamiento de la información visual, los cuales se basan en ciertos mecanismos de los sistemas de visión de algunos mamíferos; como los gatos y primates. Estos mecanismos se encuentran en algunas etapas del procesamiento, tanto a nivel retinal como de corteza visual, específicamente en los campos receptivos de unas células llamadas células simples. Estos modelos tratan de describir de mejor manera, cumpliendo además con ciertos requisitos matemáticos, los perfiles de los campos receptivos a nivel de la corteza visual. Debido a que muchos modelos pueden describir dichos campos receptivos, la respuesta a la pregunta de qué modelo es el mejor está gobernada por: 1) un análisis cuantitativo de la utilidad y consistencia de los modelos para describir y predecir los datos observados y 2) la utilidad relativa de los modelos en relación a una comprensión a mayor profundidad de la visión [84].

Como se vio en la sección 1.2, los campos receptivos son mapas de las localidades en el espacio y tiempo donde la intensidad de la luz puede afectar la salida eléctrica de una célula. Todas las células responsables de la visión en el sistema visual, desde las células receptoras en el ojo hasta las células de visión a un alto nivel en la corteza, tienen campos receptivos. La luz entra por el ojo y excita a las células receptoras en la retina, las cuales convierten a los fotones en energía eléctrica. Las señales son procesadas en la retina y posteriormente enviadas a través del tracto óptico hacia el cerebro para un análisis posterior. Después de un preprocesamiento, en un nivel anterior a la corteza visual, las señales alcanzan a ésta en la parte trasera del cerebro. Las células simples son las primeras células en la corteza visual en recibir la información entrante y se ha visto que también son las primeras células responsables de la sensibilidad a la dirección del movimiento en estudios de cerebros con primates [84].

Algunos de los primeros experimentos realizados sobre células simples usaron estímulos unidimensionales como bordes extendidos o barras [41]. En dichos experimentos, se presenta un estímulo visual y el experimentador mapea los campos receptivos de las células en el sistema de visión, de un animal por ejemplo, al cual se le presentan los estímulos. Las observaciones cualitativas, de las propiedades espaciales estáticas de los campos receptivos de las células simples, condujeron a la noción de las respuestas a bordes y barras alineados a diferentes orientaciones espaciales por parte de los campos receptivos. Hubel y Wiesel [41] enumeraron tres categorías de neuronas encontradas en la corteza

estriada como células de tipo simples, complejas e hiper complejas. De estas tres categorías de neuronas encontradas, sólo las células simples son consideradas generalmente como integradores lineales de luminiscencia dentro de sus campos receptivos por sus funciones de peso.

De acuerdo con Young [84], los modelos en general pueden dividirse en funcionales y no funcionales. Si un modelo funcional es descriptivo de los datos de los campos receptivos, conduce a una implicación del porque el campo posee una forma o perfil dado. Esto es, un modelo funcional puede indicar qué ventajas para el análisis visual posee dicho campo. A continuación se enumeran y se describen algunos de los modelos de visión mencionados en la literatura.

### 2.2.1. Modelo detector de característica

El modelo detector de característica de los campos receptivos de células simples es uno de los primeros modelos espaciales propuesto [41]. Es un modelo funcional ya que los campos de dos regiones opuestas por medio de un signo fueron denominados detectores de bordes y aquellos con tres regiones como detectores de barras. Sin embargo, el modelo detector de característica es puramente cualitativo por lo que se dificulta probarlo de una manera cuantitativa.

### 2.2.2. Modelo diferencia de Gaussiana (DOG)

A niveles de la retina, el modelo DOG<sup>1</sup> describe los campos receptivos en células ganglionares de gatos como una diferencia de Gaussianas concéntricas de diferentes tamaños [84]. En fisiología también es un modelo funcional cualitativo en el sentido de que trata de relacionar las operaciones de diferencias con la extracción de contraste. En visión computacional, el modelo DOG ha tenido un papel cuantitativo ya que ha sido mostrado que es apropiado en la formación de los campos para el modelo Laplaciano de Gaussiana.

Sin embargo, de acuerdo a Young este modelo tiene varias desventajas:

- No describe apropiadamente las células ganglionares en la región de la fovea en los monos.
- Está matemáticamente “mal formulado”, ya que las Gaussianas que representan el centro y la periferia de un número de diferentes anchos y amplitudes pueden combinarse para formar un modelo de campo DOG idéntico. De ahí que el modelo pueda tener muchas soluciones cuando se ajusta a un campo receptivo.
- Requiere de campos concéntricos (Gaussianas traslapadas con el mismo centro pero a diferentes escalas).

### 2.2.3. Combinación de Modelos DOG

Varias combinaciones lineales del modelo DOG han sido utilizados para modelar las formas de los campos receptivos de células simples. Por un lado se proponen sumas alineadas de campos del modelo DOG en la explicación de datos psicofísicos y por otro lado se proponen diferencias alineadas de campos del modelo DOG. Las desventajas encontradas con este modelo son su complejidad, múltiples soluciones para los mismos datos y la necesidad de crear campos receptivos con un mínimo de cuatro regiones cuando los datos fisiológicos muestran que muchos de los campos receptivos tienen menos de cuatro regiones [83].

---

<sup>1</sup>DOG, por sus siglas en inglés “difference-of-Gaussian”.



#### 2.2.4. Modelo derivada de Gaussiana (DG)

La semejanza en forma de las funciones unidimensionales DG y las formas de los campos receptivos en células simples de primates fue descubierta en un principio por Young [81], basado en un análisis de un número grande de espectros de frecuencia unidimensionales. La variación del ancho de banda frecuencial observada, en una muestra grande de campos receptivos de células simples en la corteza visual de monos, fue de 0.4 a un poco más de 4 octavas [23]. Esta variación del ancho de banda está directa y claramente representada por el modelo espacial DG, ya que cada campo del modelo de derivada de Gaussiana, de un número en particular de derivada, difiere de las otras en términos del ancho de banda. El modelo espacial de DG ha sido extendido de una a dos y tres dimensiones [84] y han mostrado ser consistentes con los perfiles de los campos receptivos espaciales de la retina.

Un mecanismo de implementación eficiente para el modelo de campos de DG son las Diferencias de Offset de Gaussianas (DOOG) [84], el cual de hecho es matemáticamente la forma discreta del modelo continuo de DG. A esta forma discreta se le conoce como los polinomios de Krawtchouk, de los cuales se hablará en la sección 2.4.3.

#### 2.2.5. Modelo de Gabor

El modelo de Gabor para campos receptivos visuales, basado en una descripción matemática de dichos campos como funciones Gabor, fue planteado por primera vez en 1980 [54]. Básicamente, la función Gabor es el producto de una onda de tipo seno o coseno con una función Gaussiana; y no de un polinomio de Hermite con una función Gaussiana como es el caso del modelo DG.

El modelo Gabor concibe los campos de células simples como “filtros locales del espacio frecuencial” [23], que analizan la escena visual en términos de áreas de frecuencia localizadas. Es un modelo funcional cuantitativo ya que agrega un significado funcional capaz de describir las formas de los campos receptivos.

Los campos receptivos de células simples son generalmente de naturaleza paso banda en el dominio frecuencial, con un ancho de banda en promedio de 1.4 octavas en las regiones de la fóvea en monos. Esta evidencia ha sido utilizada para asociar dichos campos con “canales” de frecuencia local.

El modelo de Gabor ha sido comparado cuantitativamente con el modelo espacial DG en ajustes unidimensionales de las respuestas a los estímulos para dichos campos receptivos. En general, ambos modelos producen buenas aproximaciones en términos de la minimización del error residual entre los datos observados y los predichos. Sin embargo, el modelo espacial de DG unidimensional se ajusta mejor a las porciones de baja y alta frecuencia de las respuestas a los estímulos frecuenciales que el modelo Gabor [84]. En la sección 2.3 se abordará con mayor profundidad este modelo.

#### 2.2.6. Modelo de la función Hermite

El modelo de la función Hermite es un modelo íntimamente relacionado con el modelo de DG. Sus propiedades matemáticas están bien establecidas en la mecánica cuántica, donde es conocida como la función del oscilador armónico lineal. Usualmente esta función es confundida con las DG debido a su forma analítica muy similar. La diferencia entre ambas es un factor de escala en la definición de las DG como se ve en las ecuaciones 2.1 y 2.2:

$$P_n(x) = NH_n(x)e^{-\frac{x^2}{2}} \quad (2.1)$$

$$G_n(x) = NH_n(x)e^{-x^2} \quad (2.2)$$

donde  $H_n$  denota el polinomio de Hermite de  $n$ -ésimo orden,  $N$  es un factor de normalización.  $P_n$  y  $G_n$  son la función de Hermite y la derivada de Gaussiana de  $n$ -ésimo orden respectivamente. La conexión

entre las funciones de Hermite y las DG es la siguiente: las funciones filtro (análisis) son las DG mientras que las funciones patrón (síntesis o reconstrucción) son las funciones Hermite. Esta asimetría es necesaria ya que las DG no son ortogonales y como consecuencia la transformada Hermite puede verse como la transformada de una Gaussiana “ventaneada” [57]. Las DG son entonces ortogonales con respecto a la función ventana Gaussiana. En la sección 2.4 trataremos con mayor detalle la transformada Hermite.

## 2.3. Wavelets Gabor

Las investigaciones teóricas [16], sobre los perfiles de los campos receptivos de las células simples en la corteza visual, han sugerido la utilización de las funciones Gabor 2D como modelo de dichos perfiles y que un grupo de células simples es mejor modelado con una familia de wavelets Gabor 2D muestreando el dominio frecuencial en forma logarítmica.

De las tres categorías clásicas de neuronas encontradas en la corteza estriada, como fueron enumeradas por Hubel y Wiesel [41], llamadas células simples, complejas e hipercomplejas; sólo las células simples son generalmente consideradas como integradores lineales de luminiscencia de sus campos receptivos por sus funciones de peso; de tal manera que cuando el umbral de descarga de estas células es excedido, las células modifican linealmente la distribución espacial de luminiscencia local de acuerdo a su perfil de campo receptivo bidimensional, que consiste de influencias excitatorias e inhibitorias a cierta frecuencia. Bajo esta condición de linealidad, la respuesta de una neurona dada a una imagen arbitraria es la integral 2D del producto de su perfil de campo receptivo con la distribución 2D de luminiscencia de la imagen, y entonces la neurona puede ser tratada como un filtro espacial 2D [16].

### 2.3.1. Funciones Gabor 2D elementales y sus propiedades

Espacialmente un filtro Gabor 2D es una sinusoidal modulada por una Gaussiana. Los campos receptivos de una familia de filtros Gabor 2D poseen ocho grados de libertad: dos coordenadas  $(x_0, y_0)$ , que especifican la posición del filtro 2D en el espacio visual; dos coordenadas de modulación  $(\xi_0, \nu_0)$ , que especifican la posición del filtro en el espacio frecuencial 2D, que se interpretan como la preferencia de orientación y de frecuencia espacial; la fase de la componente de modulación, la cual determina la combinación de simetría-asimetría del filtro representadas por una parte real y una parte imaginaria; el ancho y longitud de las envolventes elípticas Gaussianas 2D  $\sigma_x$  y  $\sigma_y$  respectivamente, las cuales son recíprocas en los dos dominios 2D; y por último, el ángulo relativo que existe entre los ejes de la Gaussiana 2D y la orientación del vector de la onda de modulación en el plano frecuencial. De estos ocho parámetros, los primeros cuatro  $(x_0, y_0, \xi_0, \nu_0)$  son las variables independientes que forman los ejes del hiper espacio de información que deben ser extendidos para cubrir el espacio.

Utilizamos  $G_{gab}$  para denotar a un filtro Gabor y marcar la diferencia entre  $G_n$ , que se utiliza para denotar a una función derivada de Gaussiana de grado  $n$ . Sea la ecuación 2.3 la definición espacial de un filtro Gabor 2D complejo centrado en  $(x_0, y_0)$  [10]:

$$G_{gab}(x, y) = g(x' - x_0, y' - y_0) \cdot \exp [j2\pi(\xi_0 x + \nu_0 y)] \quad (2.3)$$

donde  $g(x, y)$  denota a la función Gaussiana moduladora, cuyo ancho y longitud están dados por las desviaciones estándar  $\sigma_x$  y  $\sigma_y$  respectivamente y cuya razón ancho/longitud es  $\sigma_x/\sigma_y$  :

$$g(x, y) = \frac{1}{2\pi\sigma_x\sigma_y} \exp \left[ -\frac{1}{2} \left( \frac{x^2}{\sigma_x^2} + \frac{y^2}{\sigma_y^2} \right) \right] \quad (2.4)$$

La función Gaussiana está orientada (con respecto al eje coordenado  $x$  en la dirección opuesta a las manecillas del reloj) a un ángulo  $\theta$  de forma tal que las variables  $x'$  y  $y'$  indican la posición de la Gaussiana en el plano espacial 2D:

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} \cos\theta & \text{sen}\theta \\ -\text{sen}\theta & \cos\theta \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} \quad (2.5)$$

mientras que el vector de onda formado por la exponencial compleja, i.e. la parte real e imaginaria del filtro Gabor 2D, tiene una frecuencia radial central igual a  $F_0 = \sqrt{\xi_0^2 + \nu_0^2}$  y un ángulo entre sus componentes dado por  $\alpha = \text{arctg}(\frac{\nu_0}{\xi_0})$ .

Sin embargo, es conveniente considerar aquellos filtros Gabor 2D cuya función Gaussiana moduladora y el vector de onda comparten la misma orientación, es decir  $\theta = \alpha$ . De esta forma, cuando el plano del vector de onda gira, la Gaussiana lo hace de la misma manera. Esto implica que la frecuencia central del filtro  $(\xi_0, \nu_0)$  está relacionada con el ángulo de rotación de la Gaussiana por  $\xi_0 = F_0 \cos\theta$  y  $\nu_0 = F_0 \text{sen}\theta$  [45]. Con estas consideraciones y colocando en el espacio la posición del centro del filtro en  $x_0 = 0$  y  $y_0 = 0$ , la ecuación 2.3 se reduce entonces a:

$$G_{gab}(x, y) = g(x', y') \cdot \exp [j2\pi F_0 x'] \quad (2.6)$$

y podemos obtener que la respuesta en el dominio frecuencial está dada por:

$$H_{gab}(\xi, \nu) = \exp \left[ -\frac{1}{2} \left( \frac{(\xi - F_0)^2}{\sigma_\xi^2} + \frac{\nu^2}{\sigma_\nu^2} \right) \right] \quad (2.7)$$

donde  $\sigma_\xi = 1/(2\pi\sigma_x)$  y  $\sigma_\nu = 1/(2\pi\sigma_y)$  [15][53].

El parámetro de la fase, en una wavelet Gabor estándar, está dado por las partes reales e imaginarias de la función compleja que la representa. La utilización de la fase tiene su fundamento en los resultados experimentales de Pollen y Ronner [61], donde se argumenta la existencia de células simples con una fase en cuadratura. Pollen y Ronner descubrieron que pares de células simples adyacentes compartían una misma preferencia de selectividad frecuencial y orientación pero a través de una relación de fase en cuadratura (un desfaseamiento de  $90^\circ$ ). De tal manera que cada campo receptivo, con una simetría par y una simetría impar en ese par de células simples, es posible representarlo con un filtro Gabor 2D complejo como se muestra en la figura 2.1. A partir de un filtro Gabor 2D complejo se pueden obtener las partes reales e imaginarias al substituir en la ecuación 2.6 la exponencial por sus componentes de acuerdo a la identidad de Euler:

$$\text{Re}\{G_{gab}(x, y)\} = g(x', y') \cdot \cos(2\pi F_0 x') \quad (2.8)$$

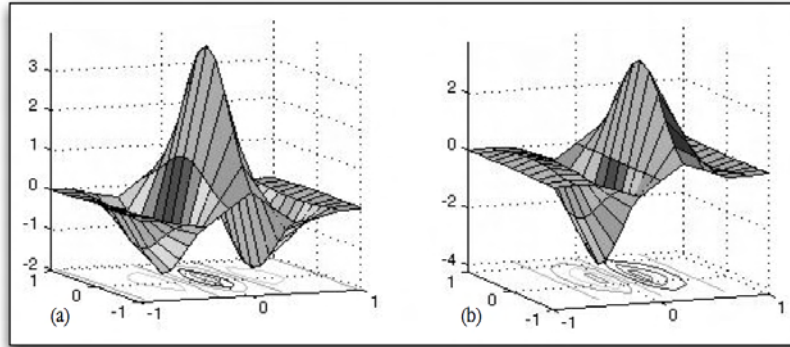
$$\text{Im}\{G_{gab}(x, y)\} = g(x', y') \cdot \text{sen}(2\pi F_0 x') \quad (2.9)$$

y la fase puede entonces obtenerse como:

$$\phi = \text{arctg}(\text{Im}\{G_{gab}(x, y)\}/\text{Re}\{G_{gab}(x, y)\}) \quad (2.10)$$

Los parámetros que definen los filtros Gabor 2D vistos hasta ahora son  $(F_0, \theta, \sigma_x, \sigma_y)$ . A continuación se determinan con base a pruebas experimentales los parámetros de los anchos de banda frecuencial  $B_F$  y angular  $B_\theta$  en el espacio frecuencial. Se sabe que las neuronas en la corteza visual tienen un amplio rango de dimensiones de los campos receptivos, de anchos de banda angulares y anchos de banda espacio-frecuencial [16][23]. Estos parámetros empíricos son correlacionados estadísticamente variando los grados de uno con respecto al otro y con otras variables, como la excentricidad

(la distancia en el espacio visual desde la fovea) y la frecuencia preferida de las células. Cuando las correlaciones obtenidas son altas, éstas implican restricciones significativas de los grados de libertad de la familia de los filtros 2D y pueden revelar la lógica del muestreo con el fin de cubrir la información en el hiper-espacio. Los valores típicos en promedio del ancho de banda frecuencial y del ancho de banda angular, encontrados en células simples de gato reportados por Daugman [16], son de 1.32 octavas y  $17.5^\circ$  respectivamente, mientras que DeValois [23] reporta anchos de banda similares (ancho de banda frecuencial promedio 1.4 octavas y ancho de banda de angular  $20^\circ$ ) encontrados en células simples de monos.



**Figura 2.1:** Componentes de un filtro Gabor 2D complejo. (a) Componente par (real). (b) Componente impar (imaginaria).

Los valores de  $B_F$  y  $B_\theta$  pueden establecerse con valores constantes que se ajusten con los datos psicovisuales como los vistos anteriormente. La frecuencia  $F_0$  y la orientación  $\theta$  definen la posición del centro del filtro en el espacio frecuencial. Para determinar las variables desconocidas  $\sigma_x$  y  $\sigma_y$ , i.e.  $\sigma_\xi$  y  $\sigma_\nu$ , las siguientes ecuaciones son establecidas; la variable  $\sigma_x$  es determinada colocando la frecuencia de corte a -6 db [15]:

$$\sigma_x = \frac{\sqrt{\ln 2}(2^{B_F} + 1)}{\sqrt{2\pi}F_0(2^{B_F} - 1)} \quad (2.11)$$

y de manera similar, colocando la frecuencia de corte en la dirección angular a -6 db se determina  $\sigma_y$ :

$$\sigma_y = \frac{\sqrt{\ln 2}}{\sqrt{2\pi} \tan(B_\theta/2)} \quad (2.12)$$

Si se deseara que la Gaussiana fuese circular, la Gaussiana ocuparía una área espacial constante en todas las direcciones. En dicho caso  $\sigma_x = \sigma_y = \sigma$  y  $B_\theta$  tiene que ser calculado. Los parámetros de posición  $(F, \theta)$  y anchos de banda deben ser establecidos cuidadosamente para capturar apropiadamente la información de la textura. Las frecuencias centrales de los filtros deben escogerse lo más cercanas a las frecuencias de la textura que la caracterizan o de lo contrario las respuestas de los filtros decaen rápidamente. Los filtros Gabor han sido implementados para análisis de textura usando únicamente su parte real (ecuación 2.8) [43] o bien usando tanto la parte real como la parte imaginaria (ecuaciones 2.8 y 2.9) [53][19].

### 2.3.2. Diseño de las wavelets Gabor 2D

Las funciones elementales de los filtros Gabor 2D pueden formar una base completa pero no ortogonal. Descomponiendo una señal en términos de esta base provee una descripción de localización en

frecuencia mientras, que la orientación del filtro una selectividad deseada. A partir de la ecuación 2.6 podemos escribir una función genérica Gabor 2D elemental como:

$$G_{gab}(x, y) = g(x, y) \cdot \exp [j2\pi F_m x] \quad (2.13)$$

donde  $m$  indica el índice de la frecuencia de análisis en pasos de octavas. Definimos una wavelet madre arbitraria como [19]:

$$\Psi(x, y) = G_{gab}(x, y) \quad (2.14)$$

Entonces podemos generar a partir de este miembro una familia completa de wavelets Gabor 2D, similares entre sí, a través de la función generadora (o wavelet hija):

$$\Psi_{m,n}(x, y) = a^{-m} \Psi(x', y'); \quad a > 1; \quad m, n \in \mathbb{N} \quad (2.15)$$

donde las variables  $(x', y')$  incorporan dilataciones logarítmicas en el tamaño de la wavelet por  $2^{-m}$  y traslaciones en una forma polar dadas por el ángulo de orientación  $\theta$  (Véase ecuación de la matriz de rotación 2.5):

$$x' = a^{-m} [x \cos \theta + y \sin \theta] \quad (2.16)$$

$$y' = a^{-m} [-x \sin \theta + y \cos \theta] \quad (2.17)$$

siendo  $\theta = n\pi/K$  y  $K$  el número total de orientaciones. El factor de escala  $a^{-m}$  es necesario para asegurar que la energía sea independiente de  $m$  [53].

Las wavelets admisibles son aquellas funciones que tienen un promedio de cero o que integran a cero (respuesta de DC), i.e. que su norma es finita. En un filtro Gabor 2D complejo la componente impar (ecuación 2.9) tiene un promedio de cero; sin embargo, la componente par (ecuación 2.8) no integra a cero. La componente de DC puede aproximarse a partir de la transformada de Fourier de la función, en cada escala de análisis (ecuación 2.7), evaluada en  $\xi = 0$  y  $\nu = 0$  [45].

$$DC_m = H(\xi = 0, \nu = 0) = \exp \left[ -\frac{1}{2} \left( \frac{F_m}{\sigma_\xi} \right)^2 \right] \quad (2.18)$$

Ahora estamos en condiciones de proponer una aproximación de wavelet Gabor 2D admisible, al substrair de la wavelet hija (ecuación 2.15) la componente de DC a dicha octava de análisis  $m$  y obtener una familia de wavelets Gabor 2D:

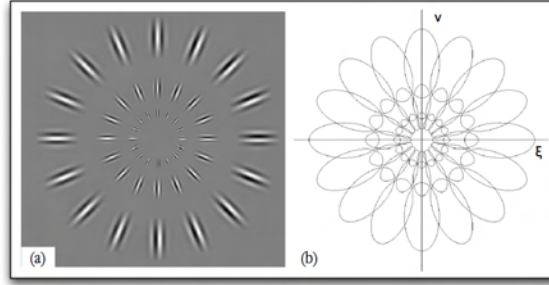
$$\Psi_{m,n}(x, y) = a^{-m} \Psi(x', y') - DC_m; \quad a > 1; \quad m, n \in \mathbb{N} \quad (2.19)$$

Un ejemplo de una familia de wavelets Gabor 2D se puede ver en la figura 2.2. Cada miembro de esta familia de wavelets Gabor 2D modela la estructura del campo receptivo espacial de una célula simple. La respuesta de una célula simple es la proyección de una imagen con una wavelet Gabor, la cual se puede calcular a partir del producto interno de la imagen  $f$  con el campo receptivo centrado en  $x_0, y_0$ :

$$R_s = \langle \Psi_{m,n}(x_0, y_0), f \rangle = \int_x \int_y \Psi_{m,n}(x_0 - x, y_0 - y) f(x, y) dx dy \quad (2.20)$$

donde  $\langle, \rangle$  denota el producto interno.

La carencia de ortogonalidad en una wavelet 2D implica que existe información redundante en las imágenes filtradas. Las wavelets que forman una base ortonormal tienen la propiedad de poder



**Figura 2.2:** Familia de wavelets Gabor 2D [45]. (a) Comportamiento en el dominio espacial. (b) Contornos del área cubierta en el dominio frecuencial.

reconstruir completamente una señal  $f$  a través de una superposición lineal de las proyecciones de la señal (bases ponderadas) con los coeficientes de la wavelet:

$$f = \sum_{m,n} \langle f, \Psi_{m,n} \rangle \Psi_{m,n} \quad (2.21)$$

Sin embargo, tanto las wavelets discretas Gabor 1D al igual que las Gabor 2D no son ortogonales; como consecuencia no forman una base ortonormal y se les agrupa dentro de las wavelets llamadas wavelets no ortogonales. Debido a que realizar reconstrucción está fuera del propósito en este proyecto, no se aborda este tema y únicamente se deja al lector abundar en la literatura relacionada con algunos paradigmas que pueden adoptarse para llegar a caracterizar a una función completamente y reconstruirla en forma “numéricamente estable” tales como las redes neuronales [18][24] y el modelo conocido en inglés como frames [45].

## Implementación de filtros Gabor 2D

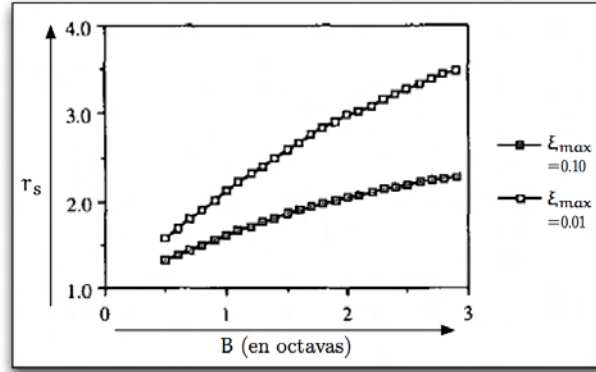
### Muestreo

Hasta ahora se ha considerado únicamente la forma continua de los filtros Gabor 2D. Como el objetivo es aplicar estas técnicas al análisis de texturas con sistemas de procesamiento digitales, es conveniente considerar el efecto de la discretización de estos filtros. Para lograr una representación apropiada de los filtros en una forma discreta, se deben usar densidades de cuantización y de muestreo adecuadas. En este aspecto, Bovik et al., [10] analizan las densidades de muestreo mínimas necesarias para asegurar que los efectos de “aliasing” no lleguen a ser excesivos.

Mantener una densidad de muestreo adecuada, sobre todo al momento de realizar convoluciones discretas con los filtros Gabor 2D, es sumamente importante ya que diferentes elecciones de los parámetros de los filtros pueden conducir a diferentes necesidades de muestreo.

En Bovik et al., [10] se derivó una tasa de muestreo mínima para evitar el “aliasing”, pero puede variar la forma de aplicarla porque depende de la flexibilidad con la cual las imágenes pueden ser obtenidas. Para evitar tener un “aliasing” severo, se define un valor máximo de porcentaje de energía de traslape  $\xi_{max}$  permitido en términos de un filtro con un ancho de banda  $B$  (en octavas). De acuerdo al teorema de muestreo, la tasa de muestreo mínima permitida (también llamada tasa de Nyquist) está dada por  $f_s = 2F$ , donde  $F$  es la frecuencia del filtro. Para evitar el “aliasing”,  $f_s$  debe ser seleccionada mayor a  $2F$ ; es decir, podemos escribir:

$$f_s = r_s(2F) \quad (2.22)$$



**Figura 2.3:** Gráfica de la razón ( $r_s$ ) de la tasa de la frecuencia de muestreo mínima a dos veces la frecuencia central del filtro Gabor contra el ancho de banda de frecuencia radial  $B$  [10].

donde  $r_s \geq 1$  es la razón que asegura cumplir con el teorema de muestreo. Despejando a  $r_s$  de la ecuación 2.22 obtenemos la razón [10]:

$$r_s = \frac{f_s}{2F} \quad (2.23)$$

A partir de los umbrales de energía de traslape y anchos de banda de frecuencia radial mostrados en las gráficas de la figura 2.3, se pueden seleccionar los valores apropiados para  $r_s$ . Para mantener el umbral de energía de traslape extremo de  $\xi_{max} = 10\%$  para anchos de banda de frecuencia radial que varían sobre el intervalo  $1/2 \leq B \leq 3$ , los valores mínimos de  $r_s$  obtenidos varían desde 1.3 hasta 2.3. Así mismo, para mantener un umbral de energía de traslape más razonable de  $\xi_{max} = 1\%$ , los valores mínimos requeridos para  $r_s$  varían desde 1.6 hasta 3.5 sobre los mismos anchos de banda de frecuencia radial. De esta manera, se puede observar que las necesidades de la tasa de muestreo mínima ( $f_s$ ) de los filtros Gabor son mayores para el segundo caso pero asegura mantener una energía de traslape a lo más de 1% [10].

Si se cuenta con imágenes con una densidad de muestreo (resolución espacial) preestablecida, entonces puede ser más eficiente la implementación de filtros Gabor con frecuencias centrales bajas aplicados a versiones submuestreadas de las imágenes de textura. La utilización de frecuencias centrales altas podría necesitar que los filtros Gabor fuesen muestreados más densamente que la rejilla disponible, a menos que el ancho de banda del filtro Gabor sea especificado lo suficientemente estrecho. Desafortunadamente, esto podría requerir el uso de un conjunto más grande de filtros si se desea cubrir todo el plano frecuencial; además el uso de filtros extremadamente estrechos puede resultar en una pérdida de la capacidad de resolución espacial entre las respuestas de los canales.

Por otro lado, si es posible definir las densidades de submuestreo para cada subimagen individualmente, implica que una descomposición jerárquica de dichas imágenes, a resoluciones variables, puede obtenerse a través de filtrar con varios conjuntos de filtros Gabor con diferentes densidades de muestreo y anchos de banda razonables, donde cada conjunto de filtros Gabor tendría una frecuencia radial en común y únicamente variaciones en la sensibilidad de la orientación.

### Selección de los coeficientes de los filtros Gabor

Existen varios métodos para la selección óptima de los coeficientes de los filtros Gabor 2D. Muchos de estos métodos están basados en el uso de redes neuronales, en donde normalmente se entrena a

la red neuronal con texturas previamente conocidas para caracterizar sus respuestas; comparando dichas respuestas con las respuestas obtenidas con otras texturas también conocidas y poder entonces efectuar la discriminación. Otros métodos emplean un banco de filtros Gabor 2D para caracterizar las texturas sin necesidad de redes de aprendizaje con texturas conocidas. Es posible agrupar estos métodos en métodos supervisados y métodos no supervisados [15].

- Métodos supervisados

Bovik et al., [10] mencionan tres aproximaciones para seleccionar las posiciones de los filtros utilizando información empírica basada en la potencia del espectro característica de cada textura individual. Para texturas predominantemente orientadas, el pico espectral más sobresaliente sobre la dirección de la orientación predominante es usado como una posición del filtro. Las texturas periódicas se pueden identificar eligiendo la frecuencia fundamental más baja. Para texturas no orientadas se recomienda usar la frecuencia central de los dos máximos. Debido a que la identificación frecuentemente requiere de muchos picos para cualquier textura genérica, estas aproximaciones son difíciles de implementar.

Dunn y Higgins [26] desarrollaron un método para seleccionar los parámetros óptimos de los filtros basado en muestras conocidas de texturas. El objetivo es identificar texturas fronterizas usando únicamente un número mínimo de filtros. Sólo un filtro en particular que separa de forma óptima las dos clases de textura es seleccionado para particionar la imagen. Este filtro óptimo puede reflejar altamente la textura característica de una clase pero puede expresar una carencia de información de textura para otra clase. Utilizar únicamente un filtro para discriminar cada par de texturas puede ser difícil, ya que cada textura puede estar sujeta a cambios espaciales y puede contener muchos componentes dominantes.

- Métodos no supervisados

Obviamente los métodos automáticos que identifican texturas características son más atractivos. En lugar de intentar identificar picos únicos que pertenecen a texturas únicas, otra aproximación es utilizar un banco de filtros, cuyos filtros se extiendan por todo el campo del dominio frecuencial. En este caso, la información de textura en una imagen puede ser capturada con una descomposición de esta imagen a través de un banco de filtros de Gabor 2D a diferentes escalas y orientaciones [15][53](Véase figura 2.2).

## 2.4. Transformada Hermite

En muchas aplicaciones tales como codificación de imágenes y visión computacional, así como en el caso del sistema de percepción visual humano, es necesario que los datos de una imagen digital, los cuales están dados como un arreglo de valores de intensidad, sean interpretados como patrones visualmente significativos a través de algún procesamiento local espacio-temporal de los datos originales. Este tipo de procesamiento local se lleva a cabo multiplicando la imagen original por una ventana de análisis cuyo tamaño establece el conjunto de elementos que contribuyen en ese paso de procesamiento en particular, mientras que la forma de la ventana establece el peso relativo con el que cada elemento contribuye. Para describir la imagen en su totalidad es necesario que este procesamiento local sea repetido para un número suficiente de posiciones de la ventana sobre la imagen original [55].

La teoría de wavelets muestra como las señales pueden ser expandidas en una familia de funciones, las cuales son dilataciones y translaciones de una única función o wavelet madre (ventana) [55]. Sin embargo, las funciones ventana que satisfacen todas las condiciones necesarias de ortogonalidad tienen dos desventajas. Primero, usualmente son mucho más grandes que el espaciamiento entre las



ventanas. Segundo, frecuentemente son menos suaves que las funciones ventana deducidas a partir de los perfiles de los campos receptivos (RFP) en la percepción visual humana. Para poder obtener funciones ventana suaves es necesario desechar la condición de ortogonalidad entre funciones base de ventanas adyacentes como sucede por ejemplo en las wavelets Gabor mencionadas en la sección 2.3.

Un parámetro importante de una ventana de análisis es su tamaño o escala. La selección de una ventana de análisis lo suficientemente grande permite obtener una reducción de datos considerablemente alta, sin embargo, la complejidad del análisis dentro de cada ventana se incrementa al aumentar el tamaño de dicha ventana. Las medidas que se pueden tomar son: a) fijar el tamaño de la ventana de análisis de tal manera que todos los patrones visuales de interés sean incluidos en el análisis o bien b) limitar la complejidad del análisis en cada ventana y consecuentemente determinar el tamaño de la ventana necesario para describir localmente con suficiente precisión la imagen. Como consecuencia, en lugar de restringir el procesamiento a una escala, se puede repetir el mismo procesamiento a múltiples escalas y consecuentemente usar las salidas en estas etapas de procesamiento para seleccionar la escala óptima en cada posición [55]. Se ha visto que este tipo de procesamiento es similar al utilizado por el sistema de visión humano.

Frecuentemente, cuando se usan funciones ventana de diferentes tamaños, el espaciado entre las funciones ventana es tomado proporcionalmente a esos tamaños. En la mayoría de las estructuras de descomposición piramidal de señales, el tamaño y el espaciado entre ventanas se incrementan en pasos de dos<sup>2</sup>. Este incremento exponencial en el tamaño y espaciado entre ventanas es preferible tanto desde el punto de vista de la información como de la eficiencia computacional.

### 2.4.1. Transformada polinomial

La transformada polinomial es una técnica de descomposición de señales en la cual una señal es aproximada localmente a través de polinomios. El análisis por una transformada polinomial consiste de dos pasos. Primero la señal original  $f(x)$  es localizada al multiplicarla por una función ventana  $V(x)$ . Una descripción completa de la señal necesitaría que la localización, a través de la ventana  $V(x)$ , se llevara a cabo en un número suficiente de posiciones de la ventana sobre la señal. Podemos definir una función de peso a partir de la función ventana  $V(x)$  como una repetición periódica [55]:

$$W(x) = \sum_k V(x - kT) \quad (2.24)$$

La función de peso  $W$  es periódica con periodo  $T$ . Asegurando que  $W(x)$  es diferente de cero para toda  $x$ , obtenemos:

$$f(x) = \frac{1}{W(x)} \sum_k f(x) \cdot V(x - kT) \quad (2.25)$$

El segundo paso consiste en aproximar la región de la señal dentro de la ventana  $V(x - kT)$  por un polinomio  $Q_n(x)$ , donde el subíndice en  $Q$  denota el grado del polinomio que en este caso es  $n$ . Como funciones base para la expansión polinomial, se toman los polinomios  $Q_n(x)$  que son ortonormales con respecto a la ventana de análisis  $V^2(x)$ :

$$\int_{-\infty}^{\infty} V^2(x) Q_m(x) Q_n(x) dx = \delta_{mn} \quad (2.26)$$

En una transformada polinomial, los coeficientes de una señal  $f(x)$  (proyecciones) pueden ser obtenidos como:

$$f_n(kT) = \int_{-\infty}^{\infty} f(x) \cdot Q_n(x - kT) V^2(x - kT) dx$$

<sup>2</sup>Al incremento en pasos de dos se le conoce como octava.

seguido de un muestreo en múltiplos de  $T$ . Es posible notar que la señal original  $f(x)$  se convoluciona con las funciones filtro denotadas por:

$$D_n(x) = Q_n(-x)V^2(-x) \quad (2.27)$$

es decir:

$$f_n(kT) = \int_{-\infty}^{\infty} f(x) \cdot D_n(x - kT) dx \quad (2.28)$$

La aproximación de la señal original  $f(x)$ , a partir de los coeficientes  $f_n(kT)$  está dada por:

$$f(x) = \sum_{n=0}^{\infty} \sum_k f_n(kT) \cdot P_n(x - kT) \quad (2.29)$$

donde las funciones patrón están definidas como:

$$P_n(x) = Q_n(x)V(x)/W(x) \quad (2.30)$$

A la expansión de una señal original  $f(x)$  en sus coeficientes  $f_n(kT)$  (ecuación 2.28) se le llama transformada polinomial. Por otro lado, a la reconstrucción de la señal a partir de los coeficientes (ecuación 2.29) se le llama transformada polinomial inversa, la cual consiste en un proceso de interpolación de los filtros  $f_n(kT)$  con una función patrón  $P_n(x)$  y suma sobre todos los grados  $n$ .

### 2.4.2. Transformada Hermite

La transformada Hermite, formalmente introducida por Martens [55] al área de procesamiento digital de imágenes, es un caso particular de una transformada más general denominada transformada polinomial. Debido a que se ha prestado a confundir a las derivadas de Gaussianas con las funciones Hermite [70][78], es conveniente explicar las diferencias que existen entre ambas antes de plantear a la transformada Hermite. De igual manera comenzamos el estudio en una dimensión ya que el caso bidimensional es muy similar.

Los polinomios de Hermite son las soluciones a la ecuación diferencial [65]:

$$y'' - 2xy' + 2ny = 0 \quad (2.31)$$

donde  $x$  es una variable espacial y  $y$  es una función de  $x$ . La ecuación anterior es conocida como la ecuación de onda del oscilador armónico (encontrada cuando se resuelve la ecuación de Schrödinger). La forma matemática de los polinomios de Hermite dada por la fórmula de Rodrigues [1] es:

$$H_n(x) = (-1)^n e^{x^2} \frac{d^n e^{-x^2}}{dx^n}, \quad n = 0, 1, 2, \dots, \quad (2.32)$$

donde el operador  $\frac{d^n}{dx^n}$  se refiere a la  $n$ -ésima derivada de la función. La derivación de esta fórmula puede verse en el apéndice A. Esta serie es la que se ha usado en visión computacional después de multiplicar la ecuación 2.32 por una función ventana. Las derivadas de Gaussianas pueden ser definidas como:

$$G_n(x) = \frac{d^n e^{-x^2}}{dx^n}, \quad n = 0, 1, 2, \dots, \quad (2.33)$$

Por simplicidad, adoptamos por un momento la convención de que la desviación estándar de la Gaussiana es  $\sigma = 1$  y omitimos el factor de normalización. Como puede verse, la ecuación 2.33 corresponde a los polinomios de Hermite multiplicados por una función Gaussiana, de tal manera que

podemos reescribir a las derivadas de Gaussianas en función de dichos polinomios y de la función Gaussiana:

$$G_n(x) = (-1)^n \cdot H_n(x)e^{(-x^2)} \quad (2.34)$$

Cuando los polinomios de Hermite son multiplicados por una Gaussiana diferente, entonces una nueva función es obtenida. Esta función es conocida en la literatura como una función Hermite ortogonal, función cilindro parabólica, función Weber-Hermite, función Hermite-Gaussian y en nuestro caso la llamaremos función patrón o de síntesis. La función Hermite, obtenida a partir de multiplicar el polinomio de Hermite por la ventana Gaussiana ( $e^{-x^2/2}$ ), puede escribirse entonces como:

$$P_n(x) = H_n(x)e^{-\frac{x^2}{2}} \quad (2.35)$$

donde es posible notar que la desviación estándar de la ventana Gaussiana en esta última ecuación (2.35) es mayor que en la ecuación de las derivadas de Gaussiana (2.34) por un factor de  $\sqrt{2}$ . La expresión para este caso en particular puede ser reescrita como

$$P_n(x) = (-1)^n e^{\frac{x^2}{2}} \frac{d^n e^{-x^2}}{dx^n}, \quad n = 0, 1, 2, \dots \quad (2.36)$$

Una vez que se ha aclarado la diferencia entre un polinomio de Hermite  $H_n$ , una derivada de Gaussiana  $G_n(x)$  y una función de Hermite  $P_n(x)$  (conocida también como función patrón) es conveniente realizar la conexión entre la definiciones de transformada polinomial y estas últimas para definir a la transformada Hermite.

### Funciones de análisis (funciones filtro)

La transformada Hermite es una transformada “ventaneada” en la que es posible distinguir a la función ventana así como a la funciones de análisis y de reconstrucción. Ahora nos centramos en el caso particular de que la función ventana es una función Gaussiana [55]:

$$V(x) = \frac{1}{\sqrt{\sqrt{\pi}\sigma}} e^{-\frac{x^2}{2\sigma^2}} \quad (2.37)$$

donde el factor de normalización  $L^2$  es tal que  $V^2(x)$  tiene energía unitaria y  $\sigma$  es la desviación estándar de la Gaussiana (Véase apéndice B).

Los polinomios ortogonales que están asociados a  $V^2(x)$  son conocidos como los polinomios de Hermite (ecuación 2.32), de ahí que el nombre que se le otorgue a esta técnica de descomposición local sea la transformada Hermite.

Las funciones filtro o de análisis determinan que información está hecha explícita en los coeficientes de la transformada Hermite. Las propiedades principales de la transformada Hermite están entonces determinadas por dichas funciones. Denotaremos como  $D_n(x)$  a las funciones filtro utilizadas en la transformación Hermite para evitar confundirlas con las derivadas de Gaussiana<sup>3</sup>  $G_n$ , ya que las primeras incluyen un factor de normalización  $L^2$ . Entonces, es posible sustituir en la definición de la función filtro 2.27 de una transformada polinomial, la función ventana  $V(x)$  y el polinomio  $Q_n(x)$  por la ventana Gaussiana de la ecuación 2.37 y los polinomios de Hermite de la ecuación 2.32 normalizados por el factor  $1/\sqrt{2^n n!}$  y evaluados en  $x = x/\sigma$  respectivamente (consultar apéndice A):

$$D_n(x) = \frac{(-1)^n}{\sqrt{2^n n!}} \cdot \frac{1}{\sigma\sqrt{\pi}} H_n\left(\frac{x}{\sigma}\right) e^{-\frac{x^2}{\sigma^2}} \quad (2.38)$$

<sup>3</sup>Es posible demostrar que  $D_n(x) = kG_n^2(x)$ , donde el exponente introduce únicamente un factor de escala en la definición de derivada de Gaussiana (Véase apéndice C).

que podemos reescribirla en términos de la definición de derivadas de Gaussiana (ecuación 2.33) evaluada en  $x = (x/\sigma)$  (Véase apéndice C) como [55]:

$$D_n(x) = \frac{1}{\sqrt{2^n n!}} \cdot \frac{1}{\sigma \sqrt{\pi}} \cdot G_n\left(\frac{x}{\sigma}\right) \quad (2.39)$$

Es aquí donde se produce la conexión entre los perfiles de los campos receptivos del sistema de visión humano modelados por derivadas de Gaussiana (diferentes a éstas por un factor de escala) y los polinomios de Hermite ortonormales respecto a la ventana de análisis.

La transformada de Fourier de las funciones filtro tiene una expresión muy simple [55]:

$$d_n(\omega) = \frac{1}{\sqrt{2^n n!}} \cdot (j\omega\sigma)^n e^{-(\omega\sigma)^2/4} \quad (2.40)$$

teniendo un valor extremo para  $(\omega\sigma)^2 = 2n$ . Incrementado el grado de los filtros conduce a un análisis sucesivo de frecuencias cada vez más altas en la señal. Sin embargo, es común encontrar que para grados grandes, los picos de frecuencia se mueven muy cercanos entre sí, de tal manera que los filtros sucesivos solo obtienen muy poca información.

Las investigaciones sobre modelos matemáticos de visión humana que simulen algunas etapas y procesos a nivel retinal o corteza visual han estado concentrados mayormente sobre funciones Gabor. El principal argumento al usar las expansiones de Gabor es que los perfiles de los campos receptivos pueden ser modelados con la funciones Gabor. Sin embargo, los perfiles de campos receptivos no son conocidos con precisión aún, de tal manera que, como ha sido propuesto por Young [82], también es posible modelarlos con derivadas de Gaussianas. Éstas tienen la ventaja de necesitar menos parámetros que las funciones Gabor.

La preferencia en el uso de funciones Gabor puede deberse al hecho de que existe una teoría matemática para la expansión de señales en funciones Gabor. Sin embargo, como se vio en la sección 2.3.2, una desventaja en el uso de funciones Gabor es que no son ortogonales entre sí y como consecuencia los coeficientes correctos no pueden ser obtenidos por la regla usual del producto interno.

La transformada Hermite es entonces una técnica de expansión de señales alternativa, en la cual las funciones de análisis son incluidas modelando los perfiles de los campos receptivos.

## Función de peso

Debido a que la función de peso es periódica con periodo  $T$ , ésta puede expandirse en una serie de Fourier, es decir:

$$W(x) = \frac{\sqrt{2\sqrt{\pi}\sigma}}{T} w(x) \quad (2.41)$$

donde

$$w(x) = 1 + 2 \sum_{k=1}^{\infty} \exp\left[-\frac{1}{2} \left(k \frac{2\pi\sigma}{T}\right)^2\right] \cdot \cos\left(k \frac{2\pi x}{T}\right) \quad (2.42)$$

El contraste de esta función de peso es determinada por el parámetro de muestreo  $\tau = T/\sigma$ . Debido a que usualmente se quiere limitar el número de descomposiciones locales, especialmente en algoritmos de codificación, es preferible seleccionar a  $\tau$  tan grande como sea posible. Por otro lado, de acuerdo a la ecuación 2.25, se puede argumentar que  $W(x)$  debe ser aproximadamente constante. De otra manera, especialmente en aplicaciones digitales, la división por  $W(x)$  podría introducir una sensibilidad a la variación de la escala.

### Funciones de síntesis (funciones patrón)

La síntesis o reconstrucción de la señal requiere del uso de las funciones patrón para reconstruir a la señal original a partir de los coeficientes de la transformada Hermite. A partir de la ecuación 2.30 podemos escribir a las funciones de síntesis como (Véase desarrollo en apéndice D):

$$\begin{aligned} P_n(x) &= H_n\left(\frac{x}{\sigma}\right) \cdot V(x)/W(x) \\ &= \frac{T}{\sqrt{2^n n!}} \cdot \frac{1}{\sigma\sqrt{2\pi}} H_n\left(\frac{x}{\sigma}\right) \cdot e^{-\frac{x^2}{2\sigma^2}}/w(x) \end{aligned} \quad (2.43)$$

donde  $w(x)$  es la función de peso definida en 2.42. Si  $w(x) = 1$  (para valores de muestreo del parámetro  $\tau < 2$ ), la función patrón es igual a la función Hermite de grado  $n$ . La función Hermite tiene la propiedad de ser isomorfa a su transformada de Fourier:

$$p_n(\omega) = \frac{T}{\sqrt{2^n n!}} \cdot (-j)^n H_n(\omega\sigma) e^{-\frac{(\omega\sigma)^2}{2}} \quad (2.44)$$

#### 2.4.3. Polinomios de Krawtchouk

A diferencia de los filtros Gabor, la transformada Hermite cuenta con una aproximación discreta basada en los polinomios de Krawtchouk. Ya que los filtros de análisis de la transformada de Hermite son similares a las derivadas de Gaussiana excepto por un factor de escala, es posible utilizar la forma discreta de las derivadas de Gaussiana dada por los coeficientes binomiales:

$$C_N^x = \frac{N!}{x!(N-x)!} \quad (2.45)$$

donde  $N$  es la longitud de la ventana binomial. Los polinomios ortonormales de Krawtchouk son obtenidos a partir del producto de los polinomios de Krawtchouk por una ventana binomial definida por:

$$v^2(x) = C_N^x / 2^N \quad (2.46)$$

de tal manera que los polinomios ortonormales de Krawtchouk están definidos como:

$$K_n(x) = \frac{1}{\sqrt{C_N^n}} \sum_{\tau=0}^n (-1)^{n-\tau} C_{N-x}^{n-\tau} C_x^\tau \quad (2.47)$$

para  $x = 0, \dots, N$  y  $n = 0, \dots, D_{max}$  donde  $D_{max}$  es el máximo orden de la derivada y  $D_{max} \leq N$ . La transformada discreta de Hermite de longitud  $N$  aproxima a la transformada Hermite continua por la siguiente relación  $\sigma = \sqrt{M/2}$  [55]. Ejemplos de los polinomios de Krawtchouk se presentan a continuación para  $D = 1$ ,  $D = 2$  y  $D = 4$  con  $D = N$ :

$$\begin{aligned} K_1 &= \frac{1}{\sqrt{2^1}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \\ K_2 &= \frac{1}{\sqrt{2^2}} \begin{bmatrix} 1 & 1 & 1 \\ 2 & 0 & -2 \\ 1 & -1 & 1 \end{bmatrix} \\ K_4 &= \frac{1}{\sqrt{2^4}} \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 4 & 2 & 0 & -2 & -4 \\ 6 & 0 & -2 & 0 & 6 \\ 4 & -2 & 0 & 2 & -4 \\ 1 & -1 & 1 & -1 & 1 \end{bmatrix} \end{aligned}$$

#### 2.4.4. Transformada Hermite 2D

Un caso especialmente interesante en las transformadas polinomiales bidimensionales resulta cuando la función ventana de localización es separable:

$$V(x, y) = V(x)V(y) \quad (2.48)$$

y la rejilla de muestreo es cuadrada. Las funciones de análisis y las funciones de síntesis son por ende separables y pueden ser implementadas de manera muy eficiente [55]. La transformada Hermite se obtiene cuando las funciones ventana son Gaussianas como las que se obtienen de la ecuación 2.37. Como las Gaussianas son isotrópicas podemos escribir  $\sigma = \sigma_x = \sigma_y$ :

$$V(x, y) = \frac{1}{\sigma\sqrt{\pi}} e^{-\frac{x^2+y^2}{2\sigma^2}} \quad (2.49)$$

Los polinomios de Hermite ortonormales con respecto a  $V^2(x, y)$  pueden escribirse entonces como:

$$Q_{n-m,m}(x, y) = \frac{1}{\sqrt{2^n(n-m)!m!}} H_{n-m}\left(\frac{x}{\sigma}\right) H_m\left(\frac{y}{\sigma}\right) \quad (2.50)$$

Los coeficientes de la señal localizada  $f_{n-m,m}(x, y)$  se obtienen al filtrar la imagen con las funciones filtro separables dada por:

$$D_{n-m,m}(x, y) = V^2(-x, -y)Q_{n-m,m}(-x, -y) \quad (2.51)$$

#### Transformada Hermite multiescala 2D

Una descomposición multirresolución usando la transformada Hermite puede ser obtenida a través de un esquema piramidal. En una descomposición piramidal la imagen es descompuesta en un número de subimágenes paso banda o paso baja, las cuales son posteriormente submuestreadas en proporción a su resolución espacial. En cada capa se transforman los coeficientes de orden cero para obtener, en una capa inferior, una versión escalada de la anterior.

#### 2.4.5. Transformada Hermite 2D orientada

Freeman y Adelson [33] describen el problema de encontrar la respuesta de un filtro a varias orientaciones. El término de filtro orientable describe una clase de filtros, en la cual un filtro a una orientación arbitraria es sintetizado como una combinación lineal de un conjunto de filtros bases.

Con la propiedad de separabilidad, es posible incorporar una rotación de los filtros que puede ser preestablecida durante el diseño de los filtros [66] o adaptativa [69][74] de acuerdo a una función de selectividad de orientación de los patrones en la señal 2D. Los filtros de Hermite 2D  $n + 1$  forman una base de los filtros de orden  $n$  sensibles a una orientación  $\theta$  en particular. De manera más específica, los filtros rotados de orden  $n$  pueden ser construidos a través de combinaciones lineales de los filtros de orden  $n$ .

La transformada de Fourier de las funciones de análisis bidimensionales  $D_{m,n-m}(x, y)$  es:

$$d_{m,n-m}(\omega_x, \omega_y) = \alpha_{m,n-m}(\theta)d_n(\omega) \quad (2.52)$$

donde  $\omega_x = \omega\cos\theta$  y  $\omega_y = \omega\sen\theta$ . La transformada de Fourier de cada función de análisis puede escribirse entonces como un producto que separadamente expresa una preferencia de orientación y preferencia de frecuencia radial. La frecuencia radial de las funciones de análisis 2D de orden  $n$  iguala

a la derivada de Gaussiana de orden  $n$  y están dadas por la ecuación 2.40, mientras que la preferencia de orientación está dada por la siguiente función angular:

$$\alpha_{n-m,m}(\theta) = \sqrt{C_n^m} \cos^{n-m} \theta \sin^m \theta \quad (2.53)$$

Denotamos a  $f_{n-m,m}^\theta(x, y)$  como una versión rotada o proyectada de los coeficientes  $f_{n-m,m}(x, y)$  con respecto al ángulo de preferencia  $\theta$ . Es decir, podemos escribir a la transformada Hermite orientada de la siguiente manera:

$$f_{n-m,m}^\theta(x, y) = \sum_{m=0}^n f_{n-m,m} \cdot \alpha_{n-m,m}(\theta) \quad (2.54)$$

con la cual se tiene la ventaja de reducir el número de coeficientes usados para representar a la imagen.

## 3 Reconocimiento de Iris

El desempeño de un sistema de reconocimiento de iris está ligado a diversos factores que deben tomarse en cuenta en cada etapa de desarrollo. Por ejemplo, es importante contar con imágenes de iris con calidad razonablemente buena, que proporcionen la información necesaria para poder realizar un buen reconocimiento; los algoritmos de localización del iris dentro de una imagen deben ser precisos, rápidos y robustos a ciertas interferencias que se pudieran encontrar durante la adquisición de las imágenes; así mismo, los modelos utilizados para el análisis de las texturas del iris deben extraer las características que permitan discriminar, con un margen de error sumamente bajo, dos texturas de iris que pertenecen a ojos distintos y proporcionar una correlación alta entre dos texturas de iris que pertenecen a un mismo ojo. Otro de los grandes retos, que no se aborda en este proyecto, consiste en desarrollar algoritmos de reconocimiento veloces y capaces de trabajar en condiciones adversas [63].

Cabe señalar que un sistema de reconocimiento de iris puede ser utilizado como autenticador o como verificador de la identidad de una persona en una base de datos. Cuando un sistema de reconocimiento de iris actúa como autenticador, una persona se presenta ante el sistema y reclama una identidad dada, en tanto que el sistema realiza las mediciones de similitud entre la plantilla de la identidad reclamada por la persona y la plantilla de la persona obtenida en ese momento, para otorgar una aceptación o un rechazo a través de un criterio de decisión. Por otro lado, cuando el sistema actúa como verificador de la identidad de una persona, una persona se presenta ante el sistema y éste compara la plantilla de la persona obtenida en ese momento contra todas las demás plantillas en la base de datos y, a través de un criterio de decisión, indica a la persona si ésta pertenece o no a la base de datos y de quien se trata.

En este capítulo se describen a profundidad las diferentes etapas que conforman un sistema típico de reconocimiento de iris así como también de los algoritmos que han sido utilizados por diversos autores en cada una de ellas. De igual manera, se describe el sistema de reconocimiento de iris propuesto. El capítulo se divide en tres secciones. En la sección 3.1, se señalan los motivos por los cuales una textura de iris puede ser utilizada como un identificador biométrico robusto. En la sección 3.2 se abordan las etapas de un sistema de reconocimiento de iris y los diferentes algoritmos que han sido utilizados con anterioridad. En la sección siguiente, sección 3.3, se describen los algoritmos implementados en el sistema de reconocimiento de iris propuesto.

### 3.1. El Iris del ojo humano: identificador biométrico robusto

Como fue mostrado en la sección 1.2 del sistema de visión humano, el iris es un diafragma circular que está ubicado entre la cornea y la lente del ojo humano. En la figura 1.6 de la sección 1.2.1, se muestra un corte lateral y la ubicación del iris dentro del ojo. El diámetro en promedio del iris es de 12 mm y el tamaño de la pupila puede variar desde un 10% hasta un 80% del diámetro del iris dependiendo de la cantidad de luz a la que es expuesta [19].

El iris tiene cuatro capas: la capa anterior o epitelio, la capa estromal, el músculo dilatador de la pupila y la capa posterior que es la superficie visible [46]. La capa anterior o epitelio consiste en su mayor parte de una colección densa de fibroblastos y melanocitos. En algunas áreas este tejido es bastante delgado o con poca pigmentación, lo que provoca que se formen líneas oscuras

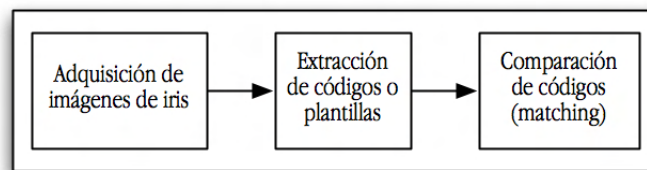


mientras que en otras áreas la pigmentación es abundante y como consecuencia se forman una serie de objetos pigmentados en forma de burbujas. El estroma yace sobre el epitelio; es una capa de tejido conectado de manera relajada que contiene colágeno, melanocitos, mastocitos, macrófagos y a los músculos elásticos. Muchos de los melanocitos y macrófagos contienen pigmento. El músculo dilatador de la pupila, como su nombre lo indica, es el responsable de la dilatación de la misma. Esta dilatación ocurre en condiciones de baja iluminación u oscuridad. Sus células están ligeramente pigmentadas. La capa posterior o el epitelio pigmentado posterior, es de apariencia oscura hablando macroscópicamente, su función es evitar que la luz se disperse a través del iris. Finalmente, el músculo que rodea a la pupila es concéntrico, yace cerca del borde de la pupila y es el responsable de la contracción de ésta. Todas estas capas juntas determinan el color del iris debido una combinación de efectos de dirección y pigmentación aleatorios. Una “huella digital” puede constituirse tomando todas las variaciones radiales y angulares que proporcionan en conjunto tanto la red trabecular como los efectos de pigmentación de los distintos tejidos y músculos que conforman el iris [19].

Usando técnicas de análisis de texturas como las descritas en el capítulo anterior, es posible capturar todos estos patrones aleatoriamente formados para generar una firma única para cada persona. La robustez de este identificador biométrico proviene de sus características naturales, pues es un órgano que se encuentra bien protegido, no cambia con el tiempo, es relativamente fácil de capturar y es posible obtener una alta variabilidad entre imágenes de iris que pertenecen a distintas personas y una baja variabilidad entre imágenes de iris de una misma persona cuando se analizan sus patrones característicos [19].

## 3.2. Descripción de las diferentes etapas de un sistema típico de reconocimiento de iris

Un sistema de reconocimiento de iris puede ser dividido en tres etapas principales como se muestra en la figura 3.1. En cada una de estas etapas se implementan diferentes algoritmos para procesar las imágenes de iris de manera eficiente y lograr codificar en una plantilla las características únicas para cada persona.



**Figura 3.1:** Etapas de un sistema de reconocimiento de iris.

Durante la etapa de adquisición, las imágenes de iris son capturadas a través de una cámara digital y se consideran aspectos importantes como la iluminación y el contraste de las imágenes obtenidas. Puesto en marcha un sistema de reconocimiento de iris, la adquisición de imágenes se realiza en tiempo real en conjunto con las demás etapas.

En la segunda etapa del sistema de reconocimiento de iris, cada imagen es procesada para extraer las características importantes del iris y codificada en una plantilla de tamaño universal para todas las personas. Estas plantillas son almacenadas en una base de datos de plantillas o códigos de iris.

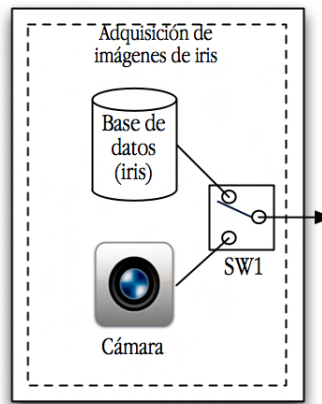
Durante la última etapa, después del establecimiento del criterio de decisión de aceptación o de rechazo, a través del estudio estadístico de las distribuciones de la medida de similitud intra-clase e

inter-clase, se presenta un sujeto al sistema de reconocimiento de iris, al cual se le extrae su código de iris y se compara éste contra otro código de iris, previamente registrado en la base de datos, usando una medida de similitud. Con el criterio de decisión establecido, el sistema puede reconocer o desconocer a la persona en cuestión.

A continuación se detallan las etapas que pueden ser encontradas comúnmente en un sistema de reconocimiento de iris y algunas de las propuestas realizadas por otros autores.

#### 3.2.1. Adquisición de imágenes

Uno de los principales retos de un sistema automático de reconocimiento de iris es capturar imágenes del iris de alta calidad. Este proceso debe realizarse de manera no invasiva y además debe ser robusto a las condiciones naturales del ambiente en el cual se encuentra el sistema (por ejemplo una sala de aeropuerto), de tal manera que este proceso no interfiera demasiado con las actividades normales de los usuarios y sea cómodo para éstos (cooperación mínima de los usuarios). Varios puntos deben ser tomados en cuenta por el sistema de captura de imágenes implementado para efectuar un buen reconocimiento. Primero, es deseable adquirir imágenes con suficiente resolución y con buen contraste que sirva tanto para localizar al iris dentro de la imagen como para poder capturar todos los detalles de los patrones dentro del iris. Para cumplir con este propósito, es importante no irritar al usuario con las fuentes de iluminación utilizadas, así como también controlar aquellas fuentes de iluminación que pudieran crear artefactos como sombras, reflejos, o aberraciones ópticas [75].



**Figura 3.2:** Adquisición de imágenes de iris. El interruptor SW1 selecciona si se trabaja en tiempo real o con una base de datos de imágenes de iris previamente capturadas.

De acuerdo con Daugman [21], un sistema de captura de imágenes de iris debe ser capaz de resolver un mínimo de 70 píxeles como radio de iris. Los sistemas típicos de adquisición utilizados han sido cámaras monocromáticas CCD<sup>1</sup> (480 x 640) en combinación con iluminación en la banda de los 700nm a 900nm del espectro electromagnético mejor conocida como NIR<sup>2</sup>. Este tipo de luz de baja intensidad,

<sup>1</sup>CCD, por sus siglas en inglés “Charge-Coupled Device” (dispositivo de cargas (eléctricas) interconectadas), es un circuito integrado que contiene un número determinado de condensadores enlazados o acoplados. Bajo el control de un circuito interno, cada condensador puede transferir su carga eléctrica a uno o a varios de los condensadores que estén a su lado en el circuito impreso.

<sup>2</sup>NIR, por sus siglas en inglés “Near InfraRed” (infrarrojo cercano) denota un intervalo de longitud de onda que se encuentra por encima del rojo visible del espectro electromagnético.

no invasiva e invisible para los humanos, es utilizada para evitar dilataciones y/o contracciones severas de la pupila al momento de capturar las imágenes, otorgar una mayor comodidad para el sujeto y obtener imágenes de buena calidad. Las imágenes de iris pertenecientes a la base de datos de libre acceso CASIAv1 [14] fueron capturadas usando las condiciones de iluminación mencionadas anteriormente; sin embargo, otras bases de datos de imágenes de iris, como UBIRIS [63], han sido capturadas usando una iluminación visible por lo que las imágenes obtenidas incluyen el color.

Una vez puesto en marcha un sistema de reconocimiento de iris, la etapa de adquisición debe ser capaz de determinar qué imágenes sirven y cuáles no. Esta etapa se implementa en tiempo real, pero para fines de desarrollo e investigación, las imágenes pueden capturarse y guardarse en una base de datos para su posterior procesamiento, véase figura 3.2. El sistema desarrollado por Daugman [21] incluye un mecanismo de enfoque en tiempo real (*autofocus*) usando una cámara de vídeo monocromática CCD. Primero se mide la potencia total de alta frecuencia en el espectro de Fourier en dos dimensiones en cada cuadro (*frame*) de captura. Posteriormente el mecanismo de la lente se mueve hacia afuera o hacia adentro y/o se le informa al sujeto a través de una grabación que se acerque o se aleje de la cámara para poder capturar la imagen apropiadamente. Aquellas imágenes que pasan un criterio de enfoque mínimo son seleccionadas para las etapas posteriores del sistema de reconocimiento de iris.

Las base de datos de libre acceso CASIAv1 y UBIRIS han sido creadas con ayuda de un operador, por lo que las imágenes obtenidas siempre muestran un buen contraste en el interior del iris. Sin embargo, la base de datos UBIRIS incluye un conjunto de imágenes tomadas bajo condiciones adversas con el propósito de evaluar la robustez de los algoritmos empleados.

#### 3.2.2. Extracción de características y codificación

La etapa de extracción de características y codificación puede dividirse en una serie de pasos previos. Primero, es necesario localizar al iris dentro de la imagen adquirida durante la etapa de adquisición. Frecuentemente se hace uso de círculos como modelos de las fronteras que delimitan al iris. Una vez que el iris ha sido detectado, es conveniente someterlo a un proceso de normalización, o bien de registro dependiendo del sistema, para lograr una consistencia entre todas las imágenes de iris adquiridas, códigos, etc. Posteriormente se extraen las características utilizando alguna herramienta matemática que permita representar a los patrones del iris de manera compacta y robusta, los cuales son almacenadas en forma de códigos o plantillas.

##### 3.2.2.1. Localización del iris

La adquisición de una imagen de iris no produce una imagen que contiene solamente al iris. Varios elementos en la vecindad de la región del iris y del ojo en general son capturados a la vez: pupila, esclerótica, párpados, pestañas, cejas, etc. Antes de hacer cualquier extracción de características es importante localizar la porción de la imagen adquirida que corresponde al iris. La región de interés en el ojo que corresponde al iris está delimitada por el limbus (el borde entre la esclerótica y el iris) y la pupila (Véase la figura 3.3). Si las pestañas y párpados obstruyen la región de interés, entonces únicamente la región sin obstrucciones debe tomarse en cuenta para el análisis. Se han reportado diversos métodos para localizar y delimitar la región de interés como los que se describen a continuación: operador integro-diferencial, transformada Hough, localización a través de máscaras predefinidas.

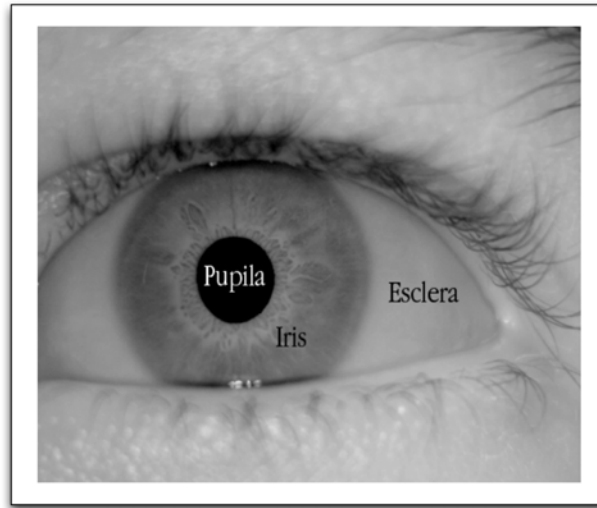
**Operador integro-diferencial.** Daugman [19] y Wildes [75] hacen uso de las primeras derivadas de los niveles de intensidad de la imagen para localizar los bordes que demarcan las fronteras del

### 3.2 Descripción de las diferentes etapas de un sistema típico de reconocimiento de iris

iris. Ambos sistemas modelan la región de interés usando modelos geométricos simples. Por ejemplo, la pupila y el iris son modelados como contornos circulares. Daugman por su parte, resume su algoritmo para localizar el iris a través de un operador integro-diferencial como se muestra en la siguiente expresión matemática:

$$\max_{(r,x_0,y_0)} |G_\sigma(r) * \frac{\partial}{\partial r} \oint_{r,x_0,y_0} \frac{I(x,y)}{2\pi r} ds| \quad (3.1)$$

donde  $*$  denota convolución y  $G_\sigma(r)$  es una función Gaussiana de suavizado con escala  $\sigma$ . Dicho operador integro-diferencial busca sobre el dominio de la imagen  $(x,y)$  el máximo en una derivada parcial suavizada, con respecto al radio variable  $r$ , de la integral de contorno normalizada de  $I(x,y)$  sobre un arco circular  $ds$  de radio  $r$  y coordenadas de centro  $(x_0, y_0)$ . El operador completo funciona como un detector de bordes circular; que busca iterativamente a escalas sucesivamente más finas de análisis a través de los tres espacios de parámetros  $(x_0, y_0, r)$ , variando el radio, por una derivada de integral de contorno máxima.



**Figura 3.3:** Imagen de iris capturada. Se observan elementos no importantes como la esclerótica, la pupila y los párpados.

El operador de la ecuación 3.1 es utilizado para encontrar las fronteras de la pupila y del limbus. Una vez que la búsqueda de las fronteras para ambas regiones alcanzan precisión de sólo un píxel, entonces se utiliza una aproximación similar para detectar bordes curvilíneos y localizar las fronteras de los párpados.

**Transformada Hough.** La transformada Hough es usada en la segmentación del iris por Wildes [76] y Ma et al., [51]. Esta transformada es un algoritmo de visión computacional que puede ser utilizado para determinar los parámetros de objetos geométricos simples, presentes en una imagen, como líneas o círculos. En términos prácticos, el método utilizado por Daugman [19] puede verse como una transformada Hough para localizar círculos. Hough (1962) patentó un método para detectar conjuntos colineales de puntos en una imagen [39]. Posteriormente se utilizó para la parametrización de líneas y recientemente en la parametrización de círculos y elipses.

En general, una curva parametrizada por  $n$  parámetros necesita un acumulador de  $n$  dimensiones para la transformada Hough [12]. Esta transformada para detección de círculos utiliza un mapa de

bordes de una imagen y puede describirse a través de una serie de pasos previos para lograr obtener las coordenadas de centro y el radio de dicho círculo. Un círculo dentro de la imagen puede ser descrito por la ecuación:

$$(x - x_0)^2 + (y - y_0)^2 = R^2 \quad (3.2)$$

donde  $(x_0, y_0)$  y  $R$  son los tres parámetros a buscar que representan las coordenadas del centro y al radio del círculo respectivamente. El uso de acumuladores para guardar y actualizar los parámetros es necesario. En el espacio de parámetros  $A$ , cada elemento  $A(x, y, r)$  tiene un valor que satisface la ecuación 3.2. Suponiendo que ya se ha calculado previamente un mapa de bordes binario  $M$ , con la ayuda de algún operador para localizar bordes, entonces para cada píxel  $P_{x,y} \in M$  se aproxima un posible  $A(x_0, y_0, R)$  calculado como:

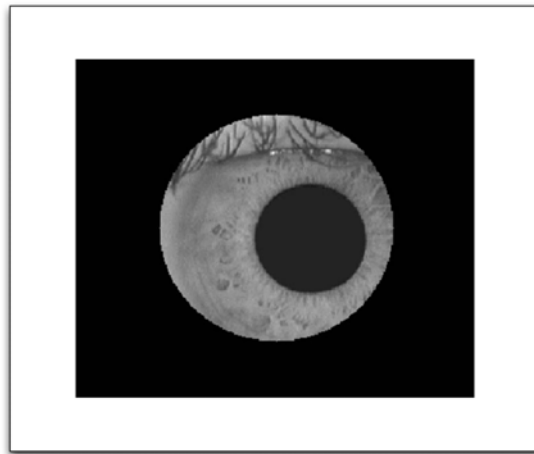
$$A(x, y, r) = A(x, y, r) + 1 \quad (3.3)$$

Después de visitar cada borde y de manera implícita votar por un posible círculo con centro  $(x_0, y_0)$  y radio  $R$ , el acumulador de parámetros  $A$  indica las posiciones en las cuales existen posibles círculos con sus tres parámetros. Basado en el histograma arrojado por el acumulador, se elige de acuerdo a alguna regla, frecuentemente la que maximiza el número de votos que recibe cada parámetro en el acumulador  $A(x, y, r)$ , que parámetros aproximan mejor al círculo en cuestión:

$$A(x_0, y_0, R) = \max(\cup A(x, y, r)) \quad (3.4)$$

y el resultado es el círculo con centro  $(x_0, y_0)$  y radio  $R$ .

Existen varios problemas en el uso de la transformada Hough. Primero, la transformada necesita valores umbrales que deben elegirse con el conocimiento previo del comportamiento de los parámetros a buscar; por ejemplo, para un círculo debe conocerse previamente un intervalo posible de valores de radio  $r \in [r_{min}, r_{max}]$ . Segundo, la transformada Hough es computacionalmente intensa y requiere mucho tiempo de procesamiento, aunque es posible con el conocimiento previo del problema obtener versiones optimizadas de esta transformada.



**Figura 3.4:** Localización de iris utilizando máscaras predefinidas.

**Máscaras predefinidas.** Otros métodos sugieren el uso de máscaras predefinidas en forma de círculos que son colocadas sobre la imagen para localizar al iris [35]. La posición del iris se obtiene

cuando se logra encerrar al iris completamente, previo movimientos de la máscara sobre la imagen como se muestra en la figura 3.4. Sin embargo, esto puede resultar ineficiente debido a que el tamaño de los ojos varía de un individuo a otro y dicho método requiere que se defina un diámetro fijo de la máscara en forma de círculo para cualquier imagen en la cual se trate de localizar un iris.

### 3.2.2.2. Normalización o alineación

Una vez que la región del iris ha sido segmentada de una imagen de ojo de manera satisfactoria, la siguiente etapa es transformar dicha región de tal manera que ésta posea ciertas dimensiones establecidas que permitan realizar las comparaciones. Las inconsistencias dimensionales que se podrían encontrar entre las imágenes de iris están relacionadas en su mayoría con el estrechamiento del iris; al dilatar o contraer la pupila debido a las variaciones de los niveles de intensidad de la luz.

Las representaciones robustas en el reconocimiento de patrones deben ser invariantes a cambios en el tamaño, posición y orientación de los patrones. En el caso de reconocimiento de iris, esto significa que se debe crear una representación que sea invariante al tamaño óptico del iris en la imagen (que puede depender de la distancia a la cual el ojo es capturado y del factor de amplificación de la cámara); el tamaño de la pupila dentro del ojo (que introduce una deformación de los patrones); la localización del iris dentro de la imagen y la orientación del iris, la cual puede depender de algún movimiento o inclinación de la cabeza al momento de capturar la imagen, movimiento del ojo dentro de su cavidad ocular, ángulo de la cámara, que en combinación producen un factor de rotación de la imagen [21].

El proceso de normalización producirá regiones de iris, las cuales tendrán dimensiones constantes, de tal manera que dos imágenes de un mismo iris tendrán sus características típicas en la misma posición espacial aproximadamente y bastará con realizar pequeñas rotaciones cíclicas para alinearlas en los algoritmos de comparación posteriores. A continuación se describen algunos de los algoritmos utilizados comúnmente en el proceso de normalización o alineación.

**Modelo de desdoblamiento elástico.** El modelo de desdoblamiento elástico de Daugman [21] remapea cada punto dentro de la región del iris a un par de coordenadas polares  $(r, \theta)$ , donde  $r$  varía en el intervalo  $[0, 1]$  y  $\theta$  es el ángulo que varía en el intervalo  $[0, 2\pi]$ . El remapeo a este sistema de coordenadas pseudopolares no es necesariamente concéntrico, ya que en muchas de las imágenes de iris capturadas es posible encontrar a la pupila en una posición no concéntrica con respecto al iris. La variable radial  $r$  varía desde la frontera de la pupila hasta el limbus (la frontera entre el iris y la esclerótica). El remapeo de la región del iris de coordenadas cartesianas al sistema normalizado de coordenadas polares no concéntricas es modelado como:

$$I(x(r, \theta), y(r, \theta)) \rightarrow I(r, \theta) \quad (3.5)$$

donde  $x(r, \theta)$  y  $y(r, \theta)$  son definidos como combinaciones lineales tanto de los puntos de la frontera de la pupila  $(x_p(\theta), y_p(\theta))$  como del conjunto de puntos del limbus sobre el perímetro externo del iris  $(x_i(\theta), y_i(\theta))$  que bordea la esclerótica, los cuales son previamente detectados por el operador de localización del iris utilizado, y están definidos como:

$$x(r, \theta) = (1 - r)x_p(\theta) + rx_i(\theta) \quad (3.6)$$

$$y(r, \theta) = (1 - r)y_p(\theta) + ry_i(\theta) \quad (3.7)$$

Debido a que la coordenada radial varía desde el perímetro interno del iris hasta el perímetro externo en un intervalo unitario, esta transformación implícitamente corrige algunas deformaciones elásticas de los patrones en el iris cuando la pupila cambia su tamaño, produciendo de esta manera una representación constante en dimensión.

Este sistema de coordenadas pseudopolares cumple con cierta invariancia al tamaño y posición del iris así como también con los cambios de tamaño que sufre la pupila; sin embargo, esta representación no es invariante a la orientación del iris dentro del plano de la imagen. En el sistema de Daugman [21], la forma más eficiente para cumplir con la invariancia a la orientación es no rotar la imagen misma usando por ejemplo la matriz de Euler; sino por el contrario, una vez que se obtuvieron los códigos de iris en forma de plantilla (una representación compacta) durante la etapa de comparación de plantillas o *matching*, éstas son rotadas a varias orientaciones discretas en la dirección angular  $\theta$  de manera cíclica hasta que las dos plantillas de iris son alineadas.

**Registro de imagen.** El sistema de reconocimiento de iris propuesto por Wildes et al., [76], utiliza una técnica de registro de imagen con el objetivo de compensar el escalamiento y rotación de la imagen. Este esquema deforma geoméricamente una imagen nueva  $I_a(x, y)$  para alinearla con una imagen seleccionada de una base de datos  $I_d(x, y)$  de acuerdo a una función de mapeo  $(u(x, y), v(x, y))$ , de tal forma que para toda posición  $(x, y)$ , el valor de la intensidad de la imagen en ese punto  $(x, y)$  en  $I_a$  es cercano a aquel valor en el mismo punto en  $I_d$ . Es decir, la función de mapeo  $(u, v)$  debe ser elegida de tal manera que se minimice:

$$\int_x \int_y (I_d(x, y) - I_a(x - u, y - v))^2 dx dy \quad (3.8)$$

mientras es limitada a capturar una transformación de similitud de coordenadas de imagen  $(x, y)$  a  $(x', y')$ , esto es:

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} x \\ y \end{pmatrix} - sR(\phi) \begin{pmatrix} x \\ y \end{pmatrix} \quad (3.9)$$

siendo  $s$  un factor de escalamiento y  $R(\phi)$  una matriz representando la rotación por un ángulo  $\phi$ . En la práctica, dado un par de imágenes de iris  $I_a$  e  $I_d$ , los parámetros  $s$  y  $\phi$  son obtenidos a través de un algoritmo iterativo de minimización [75].

**Círculos virtuales.** En el sistema de Boles [8], las imágenes de iris son primeramente escaladas para tener un diámetro constante de tal manera que cuando dos imágenes son comparadas, una imagen es considerada como una imagen de referencia. Esta técnica es diferente a las demás ya que la normalización no es realizada sino hasta el momento de comparar dos regiones de iris. Una vez que los dos iris tienen las mismas dimensiones, las características son extraídas de la región de iris almacenando los valores de intensidad al recorrer círculos concéntricos virtuales sobre la región, siendo el origen el centro de la pupila. Una resolución de normalización es seleccionada, de tal manera que el número de datos extraídos en cada iris sea el mismo para todos los demás. Esto es esencialmente similar a la técnica descrita por Daugman [19]; sin embargo, el escalamiento es hecho al momento de la comparación y es relativo a la región de iris a comparar.

### 3.2.2.3. Análisis, extracción de características y codificación

Las características espaciales discriminantes del iris humano se manifiestan a distintas escalas. Por ejemplo, las estructuras discriminantes del iris pueden variar desde la forma misma del iris hasta la distribución de los pequeños surcos y las texturas tan detalladas que lo conforman [75]. Para poder reconocer individuos de manera eficiente, la información más discriminante presente en un patrón de iris debe ser extraída. Estas características más significativas de un iris deben ser codificadas para poder efectuar las comparaciones entre las plantillas.

**Filtros Gabor.** El sistema de Daugman [19] hace uso de una descomposición derivada de la aplicación de filtros Gabor 2D a la imagen del iris normalizada en un sistema de coordenadas polares. Cada patrón de iris aislado es entonces demodulado para extraer la información de la fase usando las wavelets Gabor 2D en cuadratura. Como se mencionó en la sección 2.3, un par en cuadratura de filtros Gabor es construido modulando una señal coseno (parte real) y una señal seno (parte imaginaria) por una función Gaussiana. Los filtros reales e imaginarios también son conocidos como las componentes simétricas par e impar respectivamente.

Los filtros Gabor 2D utilizados por Daugman [19] están definidos en un sistema de coordenadas polares  $(r, \theta)$  como:

$$G(r, \theta) = e^{-j\omega(\theta-\theta_0)} e^{-(r-r_0)^2/\alpha^2} e^{-(\theta-\theta_0)^2/\beta^2} \quad (3.10)$$

Los parámetros  $\alpha$  y  $\beta$ , que son parámetros de tamaño de la wavelet multiescala, covarían en proporción inversa a  $\omega$  (la frecuencia de la wavelet) para generar una familia wavelet de filtros en cuadratura multiescala, de frecuencia selectiva y con un ancho de banda logarítmico, cuyas posiciones especificadas por  $\theta_0$  y  $r_0$  varían sobre la zonas de análisis del iris. La parte real de los filtros Gabor 2D son ligeramente ajustados, truncándolos para darles un volumen de cero y como consecuencia eliminar la respuesta de DC; de esta manera se reduce la dependencia a factores de iluminación del código de iris generado. La parte imaginaria de los filtros no tienen respuesta de DC debido a su naturaleza simétrica impar, como se vio en la sección 2.3.2.

Cada bit  $h$ , en un código de iris, es calculado al evaluar, a una escala de análisis dada y en una región local de la imagen del iris  $I(\rho, \phi)$ , el signo de ambas partes real e imaginaria que resultan de proyectar la imagen del iris con un filtro complejo Gabor 2D en particular, como se muestra en las siguientes ecuaciones [19]:

$$h_{re} = 1 \text{ si } Re \int_{\rho} \int_{\phi} I(\rho, \phi) e^{-j\omega(\theta_0-\phi)} e^{-(r_0-\rho)^2/\alpha^2} e^{-(\theta_0-\phi)^2/\beta^2} \rho d\rho d\phi \geq 0 \quad (3.11)$$

$$h_{re} = 0 \text{ si } Re \int_{\rho} \int_{\phi} I(\rho, \phi) e^{-j\omega(\theta_0-\phi)} e^{-(r_0-\rho)^2/\alpha^2} e^{-(\theta_0-\phi)^2/\beta^2} \rho d\rho d\phi < 0 \quad (3.12)$$

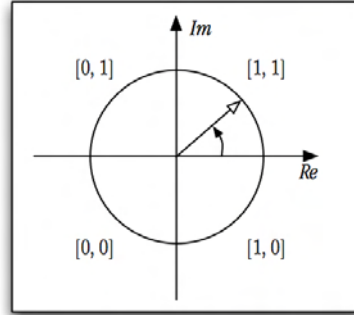
$$h_{im} = 1 \text{ si } Im \int_{\rho} \int_{\phi} I(\rho, \phi) e^{-j\omega(\theta_0-\phi)} e^{-(r_0-\rho)^2/\alpha^2} e^{-(\theta_0-\phi)^2/\beta^2} \rho d\rho d\phi \geq 0 \quad (3.13)$$

$$h_{im} = 0 \text{ si } Im \int_{\rho} \int_{\phi} I(\rho, \phi) e^{-j\omega(\theta_0-\phi)} e^{-(r_0-\rho)^2/\alpha^2} e^{-(\theta_0-\phi)^2/\beta^2} \rho d\rho d\phi < 0 \quad (3.14)$$

De tal manera que, un solo filtro Gabor 2D, teniendo un conjunto en particular de parámetros de posición y tamaño  $(r_0, \theta_0; \alpha, \beta, \omega)$  en el sistema de coordenadas polares del iris  $(r, \theta)$  realiza una cuantización de la fase de la señal de textura local aproximándola como un par de bits  $(h_{re}, h_{im})$  cuyo valor depende de los condicionales en las ecuaciones 3.11-3.14. La figura 3.5 muestra gráficamente el proceso de la cuantización de la fase.

La longitud del código generado por el algoritmo de Daugman [19] es de 2048 bits, de los cuales 1024 bits corresponden a las salidas arrojadas al proyectar la imagen normalizada del iris con un filtro real y 1024 bits corresponden a la proyección de la misma imagen con un filtro imaginario; generando una plantilla biométrica compacta de 256 bytes que permite almacenarse y realizar comparaciones de forma eficiente. De acuerdo al algoritmo general presentado por Daugman [21], la porción del iris que se analizará es mapeada (normalización) y subdividida en bandas de análisis. Estas bandas de análisis se definen en un sistema de coordenadas polares, cuya coordenada radial se puede distorsionar ligeramente si, como frecuentemente ocurre, las circunferencias que delimitan al iris no son concéntricas (ecuación 3.5). De acuerdo a Daugman [21], la imagen normalizada de iris debe ser dividida en a lo más 8 bandas de análisis. Cada banda de análisis es procesada por separado, al proyectar cada una





**Figura 3.5:** Proceso de demodulación de la fase usado para codificar los patrones del iris.

de éstas sobre los filtros complejos Gabor 2D, y posteriormente codificada al cuantizar la fase como se vio anteriormente. Al dividir la imagen del iris en 8 bandas, se obtiene que cada banda de análisis es codificada con solamente 256 bits; 128 bits de la parte real y 128 bits de la parte imaginaria.

Además de generar el código que representa en forma compacta las características principales de la textura del iris, se calcula una máscara de bits del mismo tamaño del código generado, que contiene información sobre la región del iris que es obstruida por pestañas, párpados o algunos reflejos.

**Filtros log-Gabor.** De acuerdo a Field [31], los filtros Gabor tienen la desventaja de presentar componente de DC cuando el ancho de banda es mayor a 1 octava. Sin embargo, es posible obtener componente de DC igual a cero usando un filtro Gabor que es gaussiano en una escala logarítmica. A estos filtros se les llama log-Gabor. La respuesta en frecuencia de un filtro log-Gabor está dada por:

$$G(f) = \exp\left(\frac{-(\log(f/f_0))^2}{2(\log(\sigma/f_0))^2}\right) \quad (3.15)$$

donde  $f_0$  representa la frecuencia central y  $\sigma$  determina el ancho de banda del filtro en la dirección radial. Algunos de los sistemas de reconocimiento de iris que han hecho uso de estos filtros se pueden encontrar en [3][40][80][86].

**Filtros laplacianos de Gaussianas.** El sistema descrito por Wildes et al., [75][76] utiliza una descomposición isotrópica paso banda derivada de la aplicación de filtros Laplacianos de Gaussianas a la imagen del iris. Estos filtros pueden ser especificados como:

$$\frac{1}{\pi\sigma^4} \left(1 - \frac{\rho^2}{2\sigma^2}\right) e^{-\rho^2/2\sigma^2} \quad (3.16)$$

donde  $\sigma$  es la desviación estándar de la Gaussianas y  $\rho$  es la distancia radial de un punto hacia el centro del filtro. En la práctica, la imagen filtrada es obtenida a partir de una descomposición piramidal Laplaciana, la cual en este caso es construida con cuatro niveles diferentes de resolución. A diferencia del esquema presentado por Daugman [19], las plantillas biométricas generadas con este sistema tienen un tamaño del orden del número de bytes en la región del iris original, reteniendo una mayor cantidad de información del iris.

**Wavelet Haar.** Lim et al., [48] por su parte, hace uso de la transformada wavelet Haar para extraer las características de la región del iris. A partir de un filtrado multirresolución generan un vector código de 87 dimensiones. Debido a que cada valor del vector resultante es real y varía entre -1.0 y 1.0, de manera semejante al algoritmo de Daugman [19] cuantizan cada valor real convirtiendo aquellos valores positivos como 1 y los valores negativos como 0. Finalmente, se genera una plantilla biométrica de 87 bits.

### 3.2.3. Comparación entre códigos de iris (*matching*)

**Distancia de Hamming.** En el esquema presentado por Daugman [20], la distancia de Hamming (HD) normalizada es una medida de similitud que indica el número de bits que son diferentes entre dos plantillas biométricas:

$$HD = \frac{\sum_{j=1}^N (codeA_j \otimes codeB_j) \cap maskA_j \cap maskB_j}{\sum_{i=1}^N maskA_i \cap maskB_i} \quad (3.17)$$

donde  $N$  el número de bits que conforman la plantilla biométrica. El operador XOR  $\otimes$  detecta los bits que difieren entre dos patrones de iris, enmascarados ( $\cap$ ) por ambas máscaras de bits que contienen artefactos ajenos al iris para prevenir que éstos influyan en las comparaciones de iris. El denominador es el número total de bits que son tomados en cuenta durante la comparación, después de haber descartado los bits pertenecientes a artefactos usando el operador  $\cap$ .

La distancia de Hamming es, de esta forma, una medida de similitud en la que un valor de 0, resultante al comparar dos plantillas biométricas, representaría que dichas plantillas son exactamente iguales. Si dos plantillas de iris son completamente independientes, como las generadas a partir de dos iris diferentes, la HD entre las dos plantillas deberá ser igual a 0.5. Esto ocurre porque la independencia estadística de dos códigos aleatorios implica que, cualquier bit puede ser 1 o 0, dando lugar a una probabilidad de 0.5 de encontrar un bit igual a 1 o igual a 0 [20].

**Distancia euclidiana ponderada.** La distancia euclidiana ponderada puede ser utilizada para comparar dos plantillas, especialmente si las plantillas están formadas con valores enteros [87]. La distancia euclidiana ponderada entrega una medida de que tan similar es una colección de valores entre dos plantillas.

$$WED(k) = \sum_{i=1}^N \frac{(f_i - f_i^{(k)})^2}{(\delta_i^{(k)})^2} \quad (3.18)$$

donde  $f_i$  es la  $i$ -ésima característica de un iris desconocido,  $f_i^{(k)}$  es la  $i$ -ésima característica de una plantilla  $k$  y  $\delta_i^{(k)}$  es la desviación estándar de la característica  $i$ -ésima en la plantilla  $k$ . La plantilla del iris desconocido se encuentra que es semejante a la plantilla  $k$  cuando  $WED$  es mínima para dicha plantilla  $k$ .

**Correlación normalizada.** El sistema de Wildes et al., [76] emplea un procedimiento para cuantificar la similitud entre dos plantillas. Éste está basado en la correlación normalizada entre  $p_1$  y  $p_2$  definida como:

$$\frac{\sum_{i=1}^n \sum_{j=1}^m (p_1[i, j] - \mu_1)(p_2[i, j] - \mu_2)}{nm\sigma_1\sigma_2} \quad (3.19)$$

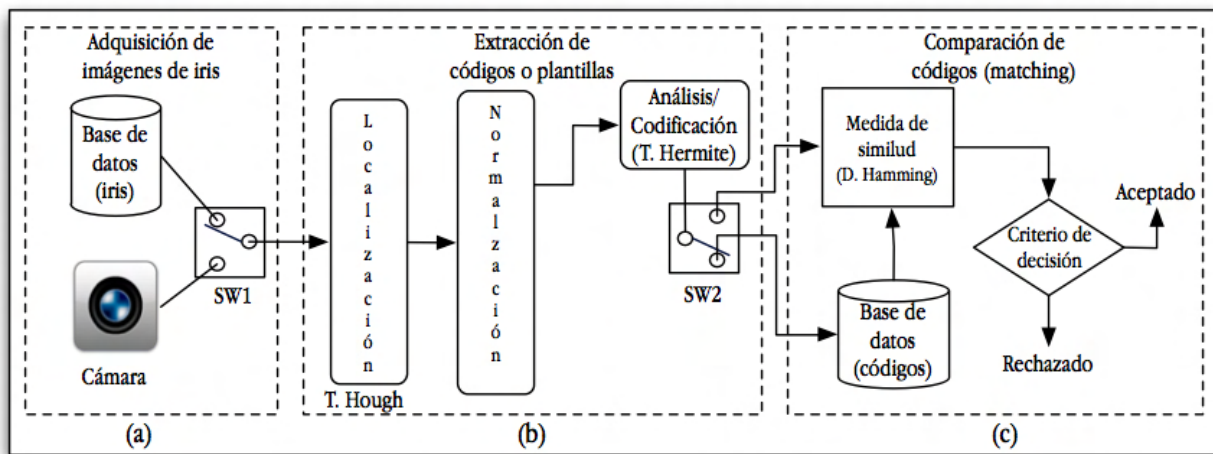
donde  $p_1[i, j]$  y  $p_2[i, j]$  son las dos imágenes de tamaño  $n \times m$ ,  $\mu_1$  y  $\sigma_1$  son la media y desviación respectivamente de  $p_1$ , mientras que  $\mu_2$  y  $\sigma_2$  la media y desviación de  $p_2$ . La correlación normalizada

captura el mismo tipo de información que la correlación estándar; sin embargo, también toma en cuenta las variaciones locales en la intensidad de la imagen que provoca errores en la correlación estándar.

### 3.3. Sistema de reconocimiento de iris propuesto

El objetivo principal en el sistema de reconocimiento de iris propuesto es utilizar el modelo de la transformada Hermite como una alternativa a los modelos de análisis de textura y extractores de características que han sido utilizados con anterioridad. Para efectos de comparación, en la etapa de análisis de textura y extracción de características de iris, además del algoritmo extractor de características usando la transformada Hermite, también fue implementado una versión de esta etapa utilizando wavelets Gabor.

El sistema de reconocimiento de iris fue implementado en su totalidad en lenguaje C y se utilizaron las bibliotecas de GTK [72] como herramienta de manipulación de entrada/salida de imágenes y de interfaz del sistema. De igual manera, el filtrado de una imagen con un filtro se implementó en el dominio de la frecuencia con la ayuda de las bibliotecas de la FFTW [34].



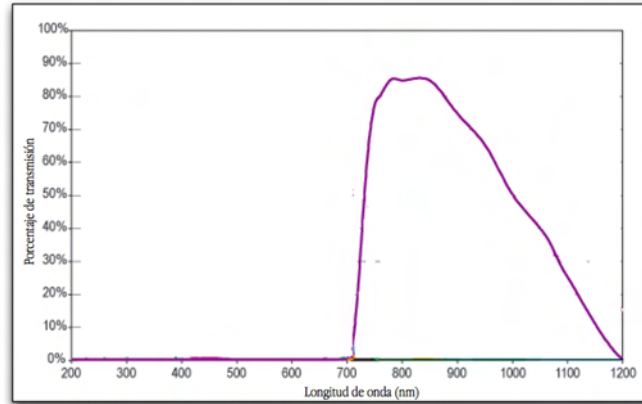
**Figura 3.6:** Sistema de reconocimiento de iris propuesto. (a) Etapa de adquisición de imágenes. (b) Etapa de extracción de códigos usando la transformada Hough en la segmentación y la transformada Hermite durante la codificación. (c) Etapa de comparación; distancia de Hamming como medida de similitud.

En la figura 3.6 se pueden observar los distintos módulos del sistema de reconocimiento de iris propuesto. A continuación se describe la implementación de cada uno de los módulos.

#### 3.3.1. Implementación del módulo de adquisición de imágenes de iris

El sistema de reconocimiento de iris trabaja en su versión fuera de línea, es decir, aunque se implementó el módulo de adquisición de imágenes de iris, este módulo no trabaja en tiempo real en conjunto con los otros módulos; por lo que las imágenes obtenidas son almacenadas en una base de datos para ser utilizadas posteriormente.

El módulo de adquisición de imágenes se montó usando una cámara de vídeo (tecnología CMOS<sup>3</sup>) con interfaz Firewire, equipada con lentes con una longitud focal de 35mm y filtros paso banda en la región del NIR. La iluminación NIR se logró con la ayuda de LEDs<sup>4</sup> infrarrojos comúnmente encontrados en el mercado electrónico. El comportamiento de los filtros utilizados para la adquisición de imágenes de iris se puede observar en la figura 3.7.



**Figura 3.7:** Filtro paso banda en la región del NIR usado en el módulo de adquisición de imágenes del sistema de reconocimiento de iris propuesto.

### 3.3.2. Implementación del módulo de localización de iris

La implementación del algoritmo de localización o segmentación se llevó a cabo utilizando una versión de la transformada Hough para detección de círculos. Como se vio en la sección 3.2.2.1, es necesario obtener un mapa de bordes binarios mediante alguna técnica de detección de bordes antes de encontrar los parámetros de centro  $x_0, y_0$  y radio  $R$  de un círculo dado. El operador utilizado para obtener el mapa de bordes binarios está basado en el operador de Canny.

#### Obtención del mapa de bordes: operador de Canny

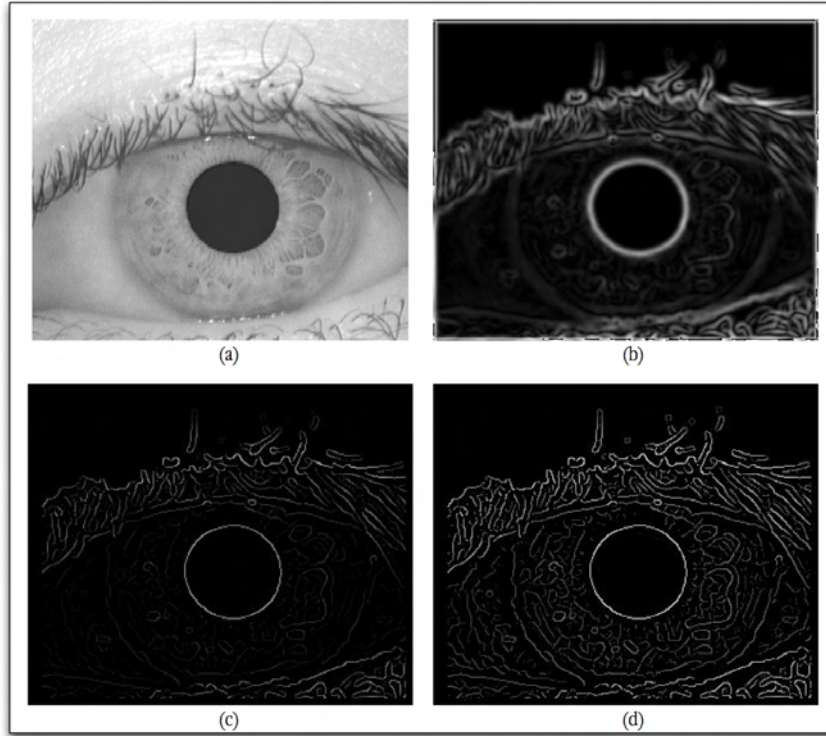
El operador de Canny [11] puede dividirse en varios procesos: reducción de ruido, obtención del gradiente y su dirección, eliminación de valores no máximos, normalización de valores, umbralización y binarización.

A continuación se describen dichos procesos implementados en el sistema de reconocimiento de iris para la obtención de los mapas de bordes binarios. Estos mapas de bordes binarios son utilizados para parametrizar la zona de la pupila y la zona del iris [25][79].

1. Reducción de ruido. En esta etapa la imagen es suavizada utilizando un filtro paso bajas de tipo gaussiano para reducir tanto el ruido como los detalles finos de la imagen. En la implementación se utilizó un filtro de tamaño 13x13.

<sup>3</sup>CMOS, por sus siglas en inglés “Complementary Metal Oxide Semiconductor” (Semiconductor Complementario de Óxido Metálico), es una de las familias lógicas empleadas en la fabricación de circuitos integrados. Su característica principal consiste en la utilización de transistores de tipo pMOS y de tipo nMOS configurados de tal forma que, en estado de reposo, el consumo de energía es únicamente el debido a las corrientes parásitas.

<sup>4</sup>LED, por sus siglas en inglés “Light-Emitting Diode” (Diodo Emisor de Luz), es un dispositivo semiconductor que emite luz, con un espectro muy angosto, cuando se polariza de forma directa y es atravesado por una corriente eléctrica. El color (longitud de onda), depende del material semiconductor empleado en la construcción del diodo.



**Figura 3.8:** Obtención del mapa de bordes normalizado usando el operador de Canny para la parametrización de la pupila y de la frontera entre el iris y la esclerótica. (a) Imagen del ojo original. (b) Gradiente de la imagen. (c) Eliminación de valores no máximos. (d) Normalización de los niveles de gris.

2. Obtención de la magnitud del gradiente de la imagen y su orientación. Una vez que la imagen ha sido suavizada con el filtro gaussiano, con la ayuda de los filtros detectores de borde, se calcula la magnitud del gradiente de la imagen y la dirección de dicho gradiente. El gradiente de una imagen  $f(x, y)$  es un vector renglón que puede escribirse como:

$$\nabla f = [G_x, G_y] = \left[ \frac{\partial f}{\partial x}, \frac{\partial f}{\partial y} \right] \quad (3.20)$$

donde  $\frac{\partial f}{\partial x}$  y  $\frac{\partial f}{\partial y}$  son las dos primeras derivadas de  $f$  con respecto a las dos coordenadas  $x$  y  $y$ . La magnitud del gradiente  $G$  está dada por:

$$G = \text{mag}(\nabla f) \quad (3.21)$$

$$= \sqrt{G_x^2 + G_y^2} \quad (3.22)$$

y su dirección como:

$$G_\theta = \text{arctg} \left( \frac{G_y}{G_x} \right) \quad (3.23)$$

En el operador de Canny propuesto se utilizaron los filtros detectores de borde Sobel en la

dirección  $x$  y Sobel en la dirección  $y$  :

$$G_x = \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix} \times [ 1 \ 0 \ -1 ] = \begin{bmatrix} 1 & 0 & -1 \\ 2 & 0 & -2 \\ 1 & 0 & -1 \end{bmatrix} \quad (3.24)$$

$$G_y = \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix} \times [ 1 \ 2 \ 1 ] = \begin{bmatrix} 1 & 2 & 1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{bmatrix} \quad (3.25)$$

En la figura 3.8.b se observa el gradiente de la imagen de ojo previamente capturada.

3. Eliminación de valores no máximos. Debido a los efectos de “aliasing” de un borde sobre sus píxeles vecinos después del cálculo del gradiente, es necesario eliminar los bordes falsos que pudieran provocar que la posición exacta del borde sea extremadamente errónea. Los píxeles que pertenecen al borde son aquellos puntos cuya magnitud local del gradiente es máxima en la dirección de dicho gradiente; de forma tal que, un píxel es retenido como un punto de borde potencial sólo si tiene una magnitud de gradiente mayor que la que tienen sus vecinos que se encuentran más cercanos en la dirección del gradiente y que se originaron por efectos del suavizado. La aproximación adoptada fue seleccionar aquellos vecinos que se encontraran en la dirección normal a la dirección del gradiente, para ser comparados con el borde potencial en cuestión. En la figura 3.8.c se observa el resultado después de eliminar los valores no máximos.
4. Normalización de los valores. Para poder aplicar una umbralización, es necesario que los valores máximo y mínimo del gradiente sean comunes para cualquier imagen procesada. Después del proceso de normalización, los valores de la magnitud del gradiente varían de 0 a 1. En la figura 3.8.d se puede observar el resultado después de la normalización de los niveles de gris.
5. Umbralización y binarización por histéresis. En esta última etapa del operador Canny se seleccionan y se marcan con un valor de 1, a través de un mecanismo de umbralización, los bordes que servirán a la transformada Hough. Para ello, se visita cada uno de los píxeles candidatos a borde que fueron obtenidos en el paso anterior (paso 4) y usando dos umbrales  $U_H$  y  $U_L$  tal que  $U_H > U_L$  se seleccionan o desechan de acuerdo a la siguiente regla:

Sea  $P_{x,y}$  el valor del píxel candidato a borde que se visita y sea  $P_{i,j}$  el valor del píxel vecino a  $P_{x,y}$ . Entonces  $P_{x,y}$  es seleccionado como borde y marcado como 1 si y sólo si:

$$P_{x,y} > U_H \quad (3.26)$$

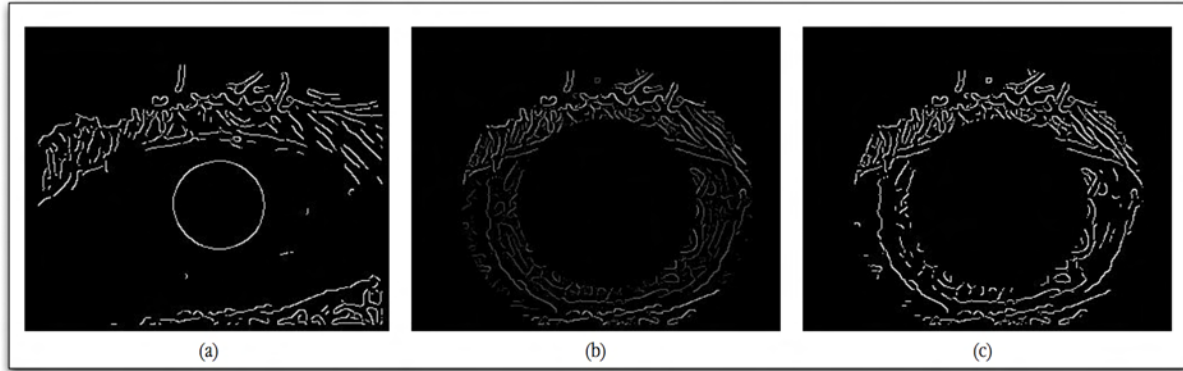
y en caso contrario se desecha el candidato marcándolo con un 0. Para cada uno de los vecinos de  $P_{x,y}$  (usando una vecindad 8-conexo),  $P_{i,j}$  es seleccionado como borde si  $P_{x,y}$  satisface la ecuación 3.26 y además el vecino  $P_{i,j}$  satisface:

$$P_{i,j} > U_L \quad (3.27)$$

donde  $i = x - 1, x + 1$  y  $j = y - 1, y + 1$ . De esta manera, el mapa de bordes binario es obtenido después de visitar y seleccionar o desechar todos los puntos candidatos a borde.

Distintos mapas de bordes binarios pueden ser obtenidos, pues dependiendo de los umbrales elegidos es posible seleccionar aquellos bordes que están presentes con una mayor intensidad que otros. En la figura 3.9a se puede observar el mapa de bordes binario obtenido a través de la umbralización y

binarización por histéresis cuando se pretende localizar a la pupila, así como también el mapa de bordes binario obtenido a través de la umbralización y binarización por histéresis cuando se pretende localizar a la frontera entre el iris y la esclerótica (figura 3.9.c). El mapa de bordes en la figura 3.9.b corresponde al mismo mapa calculado en el paso 3 (figura 3.8.c) pero cuya parte central, una vez que la pupila ha sido parametrizada, fue eliminada para reducir el mapa de bordes binarios de la esclerótica. A continuación describiremos los pasos a seguir para la parametrización de la región de interés.

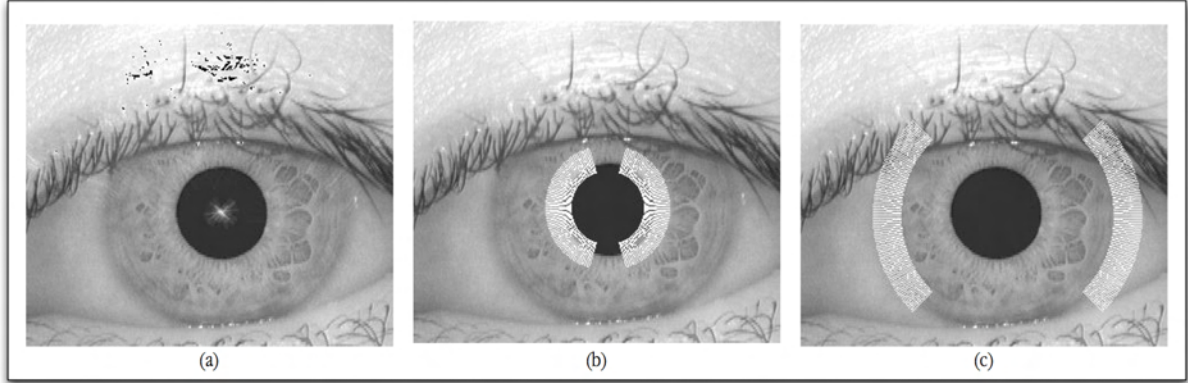


**Figura 3.9:** Obtención de los mapas de bordes binarios. (a) Mapa de bordes binario de la pupila ( $M_P$ ) obtenido a través de la umbralización y binarización por histéresis del operador de Canny. (b) Mapa de bordes de la esclerótica sin umbralizar. Observe que los bordes pertenecientes a la pupila han sido eliminados después de que ésta fue parametrizada. (c) Mapa de bordes binario de la esclerótica ( $M_S$ ) obtenido después de la umbralización y binarización por histéresis usando el operador de Canny.

### Parametrización: transformada Hough

El algoritmo propuesto para la detección de los parámetros de las fronteras entre la pupila y el iris  $A_P(x_{0p}, y_{0p}, R_P)$  y entre el iris y la esclerótica  $A_S(x_{0s}, y_{0s}, R_S)$ , comienza buscando desde el interior hacia el exterior, es decir, primero se localiza la frontera entre la pupila y el iris  $A_P$  y posteriormente se localiza la frontera entre el iris y la esclerótica (limbus)  $A_S$ . La descripción general de un algoritmo de parametrización de círculos a través de la transformada Hough fue introducida en la sección 3.2.2.1. Sin embargo, es conveniente mencionar algunos aspectos de optimización y reducción del espacio de búsqueda de los parámetros en nuestra implementación.

La primera optimización se logra durante la obtención de los mapas de bordes binarios a través del operador de Canny. Para localizar  $A_P$  es necesario obtener el mapa de bordes binario de la pupila  $M_P$ . Ésto se consigue aplicando la serie de pasos del operador de Canny sobre la imagen original del iris (figura 3.8.a), hasta completar el paso 5, que es la umbralización por histéresis (Véase figura 3.9.a). De manera similar, para la localización de  $A_S$  es necesario obtener el mapa de bordes binario de la esclerótica  $M_S$ . Se realiza otra umbralización para obtener el mapa de bordes  $M_S$ , pero esta segunda umbralización se lleva a cabo después de normalizar los mismos valores obtenidos durante la eliminación de valores no máximos (figura 3.8.c) en el paso 3, de tal forma que los pasos 1-3 del operador de Canny son ejecutados solamente una vez para cada imagen en la cual se pretende delimitar la región de interés que es el iris (figuras 3.9.b y 3.9.c).



**Figura 3.10:** Transformada Hough. (a) Búsqueda de las coordenadas del centro de la pupila  $(x_{0p}, y_{0p})$ . (b) Búsqueda del radio de la pupila  $R_P$  en un espacio reducido  $(M_P)$  para delimitar su frontera  $A_P$ . (c) Búsqueda del radio del limbus  $R_S$  en un espacio reducido  $(M_S)$  para delimitar esta frontera  $A_S$ .

La reducción del espacio de búsqueda de parámetros permite optimizar la transformada Hough. Las optimizaciones concernientes a dichas reducciones de los espacios de búsqueda se describen a continuación:

- Primero, para localizar la frontera de la pupila  $A_P$  restringimos la búsqueda sobre aquellos puntos  $P_{x,y} \in M_P$  cuya dirección de gradiente  $G_\theta$  se encuentre dentro de los intervalos  $[-2\pi/5, 2\pi/5]$ ,  $[4\pi/5, \pi]$  y  $[-4\pi/5, -\pi]$ , dejando dos espacios en forma de cono, centrados en  $\pi/2$  y  $-\pi/2$  respectivamente, cuyos puntos  $P_{x,y} \in M_P$  no son tomados en cuenta para la obtención de los parámetros  $A_P(x_{0p}, y_{0p}, R_P)$  debido a que estas regiones pueden estar obstruidas por párpados o pestañas, véase la figura 3.10.b.
- La segunda optimización se logra durante la localización de la frontera entre el iris y la esclerótica  $A_S$ . Una vez que los parámetros que delimitan la región de la pupila  $A_P(x_{0p}, y_{0p}, R_P)$  han sido encontrados a través de la transformada Hough, se elimina de la máscara de valores no máximos (calculada una sola vez en el paso 3 del operador de Canny, figura 3.8.c) el área del círculo correspondiente a la pupila parametrizada (aumentada por un factor, figura 3.9.b). El espacio de búsqueda de los parámetros de la esclerótica a través de la transformada Hough se reduce, porque en el mapa de bordes de la esclerótica  $M_S$  la información de los bordes binarios de la pupila ya no interfiere en dicho proceso de búsqueda, como se ve en la figura 3.9.c.
- El espacio de búsqueda de los parámetros de la frontera entre el iris y la esclerótica  $A_S$ , a través de la transformada Hough, se reduce aún más al restringir que esta búsqueda se realice sobre aquellos puntos  $P_{x,y} \in M_S$  cuya dirección de gradiente  $G_\theta$  se encuentre dentro de los intervalos  $[-\pi/4, \pi/4]$ ,  $[3\pi/4, \pi]$  y  $[-3\pi/4, -\pi]$ , dejando nuevamente dos espacios en forma de cono, centrados en  $\pi/2$  y  $-\pi/2$ , cuyos puntos  $P_{x,y} \in M_S$  no se consideran en la obtención del segundo grupo de parámetros  $A_S(x_{0s}, y_{0s}, R_S)$ . Estos puntos no son considerados ya que en dichas áreas también se presentan con frecuencia párpados o pestañas que pueden afectar a la parametrización, véase figura 3.10.c.

El algoritmo de la transformada Hough implementado primeramente acumula los centros de los posibles círculos y luego selecciona un centro. Posteriormente con esa información, se acumulan los

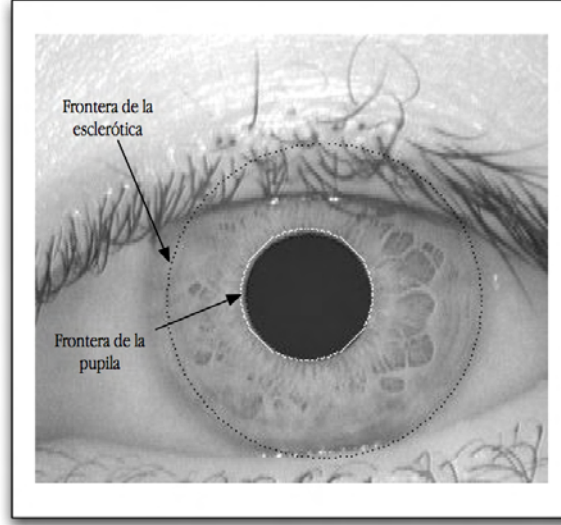


radios posibles que pueda tener el círculo. Finalmente se selecciona un radio y la parametrización del círculo se completa. El algoritmo de la transformada Hough se ejecuta dos veces: una vez utilizando el mapa de bordes de la pupila  $M_P$  para parametrizar a la misma  $A_P(x_{0p}, y_{0p}, R_P)$  y la segunda vez utilizando el mapa de bordes de la esclerótica  $M_S$  para la parametrización de la frontera entre el iris y la esclerótica  $A_S(x_{0s}, y_{0s}, R_S)$ . En las dos parametrizaciones, la transformada Hough utiliza los puntos  $P_{x,y} \in M_P$  y  $P_{x,y} \in M_S$ , cuya dirección de gradiente  $G_\theta$  está dentro de los intervalos vistos anteriormente, para calcular primeramente los parámetros que delimitan a la pupila y posteriormente los que delimitan al iris respectivamente (recordando que el área del círculo correspondiente a la pupila parametrizada es eliminada durante la parametrización de la frontera entre el iris y la esclerótica). En esta versión de la transformada Hough, además de restringir el espacio de búsqueda, la búsqueda de parámetros es dividida; de tal forma que se usan dos acumuladores: un acumulador bidimensional, cuya función es almacenar los votos en las posiciones posibles de centros de círculos, y un acumulador unidimensional, que tiene como función guardar los votos de los posibles radios del círculo.

El algoritmo se detalla en el apéndice E y se describe a continuación para un mapa de bordes cualquiera  $M$ :

1. Acumulación de centros de círculos. Lo primero que se puede notar a partir de la dirección del gradiente  $G_\theta$  en la frontera de un círculo, es que esta dirección siempre está apuntando hacia el posible centro del círculo [12]. De hecho, este paso puede verse como una transformada Radón inversa, en la cual la imagen a reconstruir a partir de las proyecciones (en este caso los mapas de bordes binarios) es un punto. Como se mencionó anteriormente, se utilizó un acumulador bidimensional  $C_A[x, y]$  cuya función es almacenar en cada posición  $(x, y)$ , en la dirección del gradiente, los votos de los posibles centros, véase figura 3.10.a. La variación del radio  $r_{min} < r < r_{max}$  se obtiene con el conocimiento previo de las propiedades de la pupila y esclerótica, y que son particulares para cada base de datos de imágenes de iris utilizada (Véase el algoritmo E.1 en el apéndice E).
2. Selección del centro del círculo. Una vez que se han visitado todos los puntos  $P_{x,y} \in M$  y se han proyectado y acumulado los posibles centros en la dirección del gradiente variando el radio  $r_{min} < r < r_{max}$ , es necesario elegir que centro en  $C_A[x, y]$  representa al círculo. La aproximación adoptada fue seleccionar aquellas coordenadas  $x_0$  y  $y_0$  cuyo número de votos acumulados en  $C_A[x, y]$  fue el mayor durante la acumulación de votos (Véase el algoritmo E.2 en el apéndice E).
3. Acumulación de radios del círculo. Ahora que ya se han encontrado las coordenadas  $(x_0, y_0)$  del círculo, es posible proyectar ese punto, variando el radio  $r_{min} < r < r_{max}$  y variando el ángulo de proyección en un intervalo apropiado sobre el mapa de bordes  $M$  y acumular en  $R_A[r]$  cada vez que en dicha proyección exista un borde (Véase algoritmo E.3 en el apéndice E). En las figuras 3.10.b y 3.10.c se pueden ver impresas sobre una imagen de iris las proyecciones de búsqueda del radio, en ese caso para la pupila usando un mapa de bordes  $M_P$  y para el iris un mapa de bordes  $M_S$  respectivamente.
4. Selección del radio del círculo. Acumulados los votos para los posibles radios del círculo, se aproxima el radio de dicho círculo al seleccionar aquel radio  $R_0$  en el acumulador  $R_A[r]$  que obtuvo el mayor número de votos (Véase algoritmo E.4 en el apéndice E).

El resultado de la parametrización Hough para localizar las fronteras entre la pupila y el iris y entre el iris y la esclerótica se muestra en la figura 3.11.



**Figura 3.11:** El iris segmentado: resultado de la parametrización de la pupila y de la esclerótica como dos círculos no concéntricos.

### 3.3.3. Implementación del módulo de normalización y eliminación de artefactos

Una vez que se localizaron los parámetros de centro y radio,  $A_P(x_{0p}, y_{0p}, R_p)$  y  $A_S(x_{0s}, y_{0s}, R_s)$  para la pupila y para la esclerótica respectivamente, el paso siguiente consistió en normalizar la región del iris usando el esquema de desdoblamiento elástico presentado por Daugman [19] y que se revisó en la sección 3.2.2.2.

El centro de la pupila fue considerado como el punto de referencia (o pivote) desde el cual se calculan vectores radiales que atraviesan la región del iris. Se selecciona un número de puntos que contienen datos dentro de la región del iris sobre cada línea radial. La cantidad de puntos elegidos en esa dirección indica la resolución radial. El número de líneas radiales que se toman sobre la zona del iris está determinada por la resolución angular escogida, véase figura 3.12.

Debido a que la pupila y el iris pueden no ser concéntricos, es necesario el uso de una función de remapeo para escalar los puntos dependiendo del ángulo alrededor del círculo. El desplazamiento del centro de la pupila  $(x_{0p}, y_{0p})$  es relativo al centro de la esclerótica  $(x_{0s}, y_{0s})$  y está dado por  $O_x$  y  $O_y$ :

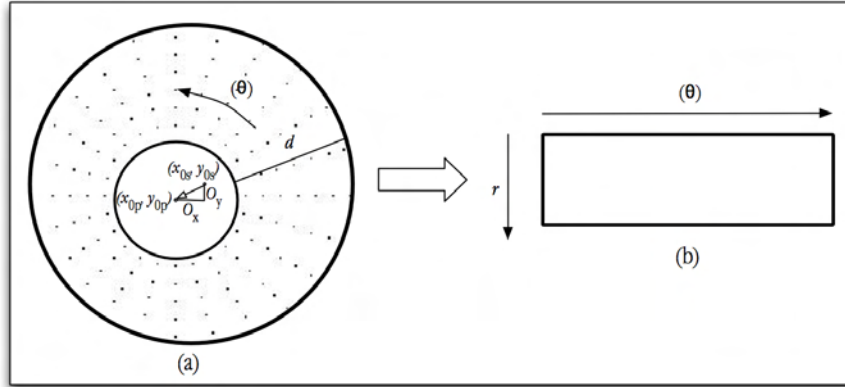
$$O_x = x_{0p} - x_{0s} \quad (3.28)$$

$$O_y = y_{0p} - y_{0s} \quad (3.29)$$

La distancia entre el borde de la pupila y el borde de la esclerótica  $d$  está en función del ángulo  $\theta$  sobre la región del iris; de tal manera que la ecuación 3.30 calcula para cada ángulo alrededor del iris la separación que hay entre la esclerótica y la pupila relativa al desplazamiento de un centro con respecto al otro:

$$d = \sqrt{\alpha}\beta \pm \sqrt{\alpha\beta^2 - \alpha - r_s^2} \quad (3.30)$$

donde  $\alpha = O_x^2 + O_y^2$  y  $\beta = \cos(\pi - \arctg(O_y/O_x) - \theta)$ . La resolución radial  $r$  determina un número constante de puntos, que son remapeados a lo largo de cada línea radial hacia el sistema normalizado de coordenadas polares no concéntricas (ecuación 3.5), independientemente de que tan estrecha o amplia sea la distancia entre la esclerótica y la pupila en un ángulo en particular, véase figura 3.12.a.



**Figura 3.12:** Esquema del proceso de normalización. (a) Los puntos seleccionados en la dirección radial son calculados en forma adaptativa en función del ángulo  $\theta$ . (b) La imagen normalizada resultante con una resolución radial  $r$  y una resolución angular  $\theta$

La normalización produce un arreglo bidimensional (figura 3.12.b), cuya dimensión horizontal representa a la resolución angular  $\theta$  y a la dimensión vertical la resolución radial  $r$ . Para asegurar no remapear regiones no pertenecientes al iris que se encuentren justo en la frontera de la pupila, el radio de la pupila es aumentado en 2% y el radio de la esclerótica disminuido en 3% aproximadamente, dependiendo de la base de datos utilizada.

Hasta este momento el iris ha quedado aislado de la pupila y de la esclerótica, y se ha mapeado la región de interés a un sistema normalizado de coordenadas polares no concéntricas. Sin embargo, existen artefactos que deben ser removidos de la región de interés como son los párpados y pestañas como se puede ver en la figura 3.13.a. La aproximación propuesta fue modelar las regiones del párpado y de las pestañas como una función elíptica, pues frecuentemente las imágenes después de haber sido normalizadas presentaban estos artefactos en forma de media elipse.

Para encontrar los parámetros de la elipse  $(a, b)$  se utilizó una versión menos sofisticada de la transformada Hough para detección de elipses. La ecuación de una elipse está dada por:

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1 \quad (3.31)$$

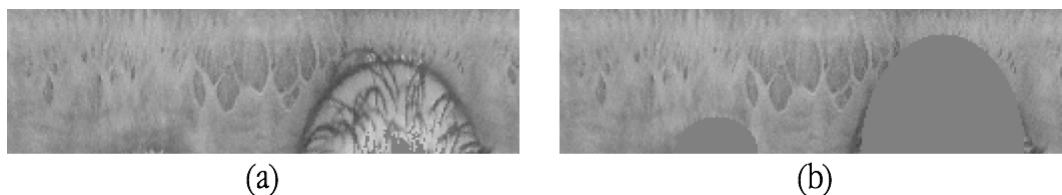
Las imágenes normalizadas con presencia de párpados y pestañas contienen estos artefactos la mayoría de las veces centrados en la primera y en la cuarta parte de ésta, es decir, en  $(\frac{\theta}{4}, r)$  y  $(\frac{3\theta}{4}, r)$ , donde  $\theta$  es la resolución angular y  $r$  la resolución radial. La imagen normalizada es dividida a la mitad y en cada una de ellas se parametriza por separado una elipse. De tal forma que pueden existir dos, una o en su defecto ninguna elipse dentro de la imagen normalizada. Los parámetros a encontrar son entonces:  $a_1, b_1$  para la primera elipse, y  $a_2, b_2$  para la segunda.

Para cada una de las mitades de la imagen normalizada se calcula un mapa de bordes binario, obtenido a través del detector de Canny sobre dichas mitades. Para reducir el espacio de búsqueda, formulamos la hipótesis de que el centro de las elipses se encuentra en  $(\frac{\theta}{4}, r)$  y  $(\frac{3\theta}{4}, r)$ . A partir de los supuestos centros, se realizan proyecciones (búsquedas) hacia los extremos, primero para parametrizar  $a_1$  y posteriormente a  $a_2$ . Para lograrlo, se propuso usar dos acumuladores en cada parametrización. En cada mitad del mapa de bordes binario de la imagen normalizada, se varía el ángulo de búsqueda formado entre el eje de la resolución angular  $\theta$  y la proyección. Se escogieron los intervalos de búsqueda

$0^\circ < \alpha < 13^\circ$  y  $167^\circ < \alpha < 180^\circ$  y durante dicha búsqueda, un acumulador es usado en cada intervalo. Cada vez que la proyección encuentra en su dirección un punto en el mapa de bordes, el acumulador asociado a dicho intervalo incrementa su valor en una unidad. De esta manera, cuando el algoritmo se ejecuta sobre la primera mitad del mapa de bordes binarios de la imagen normalizada, se parametriza  $a_1$  a partir del promedio de los valores máximos obtenidos del par de acumuladores. En forma similar, ejecutando el algoritmo sobre la segunda mitad del mapa de bordes binarios de la imagen normalizada, el promedio de los valores máximos obtenidos del par de acumuladores parametriza a  $a_2$ .

La localización sobre el eje angular  $\theta$  de los centros de las elipses queda confirmada, corregida o descartada (si no existe elipse alguna), a través de los parámetros  $a_1$  y  $a_2$ . La parametrización de  $b_1$  y  $b_2$  se lleva a cabo a través de proyecciones a partir de los respectivos centros  $a_1$  y  $a_2$ , variando el ángulo de búsqueda sobre el intervalo  $65^\circ < \alpha < 115^\circ$  para cada elipse. En este caso únicamente se utiliza un acumulador durante cada búsqueda, el cual incrementa su valor cada vez que la proyección encuentra en su dirección un punto en el mapa de bordes. Tanto  $b_1$  como  $b_2$  se parametrizan tomando el índice en el acumulador que obtiene el valor mayor de votos para cada caso.

Los resultados de esta aproximación para eliminación de párpados y pestañas se pueden observar en la figura 3.13.



**Figura 3.13:** Normalización y eliminación de párpados y pestañas. (a) Imagen normalizada con artefactos: párpados y pestañas. (b) La misma imagen normalizada; los artefactos han sido removidos.

#### 3.3.4. Implementación del módulo extractor de características Hermite

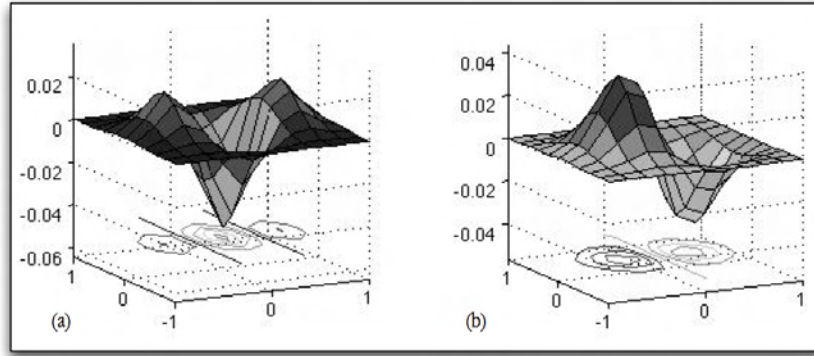
Para extraer las características de las texturas normalizadas del iris se utilizó la herramienta alternativa: la transformada Hermite. De manera similar al algoritmo de Daugman [21], se utilizó un par de filtros par e impar obtenidos al combinar dos ordenes de filtro. De acuerdo a [57], es posible formar un par óptimo de filtros de análisis Hermite tomando:

$$F = D_{2nx} \mp jk\sigma^{\pm 1} D_{(2n\pm 1)x} \quad (3.32)$$

Los parámetros de optimización del par de filtros no fueron considerados en este proyecto ya que es un proceso complejo que se encontraba fuera del alcance de esta tesis, por lo que se aproximó a una forma compleja de la transformada Hermite tomando únicamente una combinación del primer y segundo orden de los filtros de análisis de Hermite:

$$F \approx D_{2,0} + jD_{1,0} \quad (3.33)$$

Se eligieron los filtros de segundo y primer orden como los filtros par e impar respectivamente como se muestra en la figura 3.14. La fase fue cuantificada en forma similar al algoritmo propuesto por Daugman [19] (ecuaciones 3.11-3.14) en la sección 3.2.2.3, sustituyendo las coordenadas polares por



**Figura 3.14:** Filtros de análisis usando la transformada Hermite 2D. (a) Filtro par  $D_{20}$ . (b) Filtro impar  $D_{10}$ .

coordenadas cartesianas y tomando las funciones filtro de Hermite como los filtros de análisis de la textura del iris. De esta manera, el código se generó a partir de las ecuaciones siguientes:

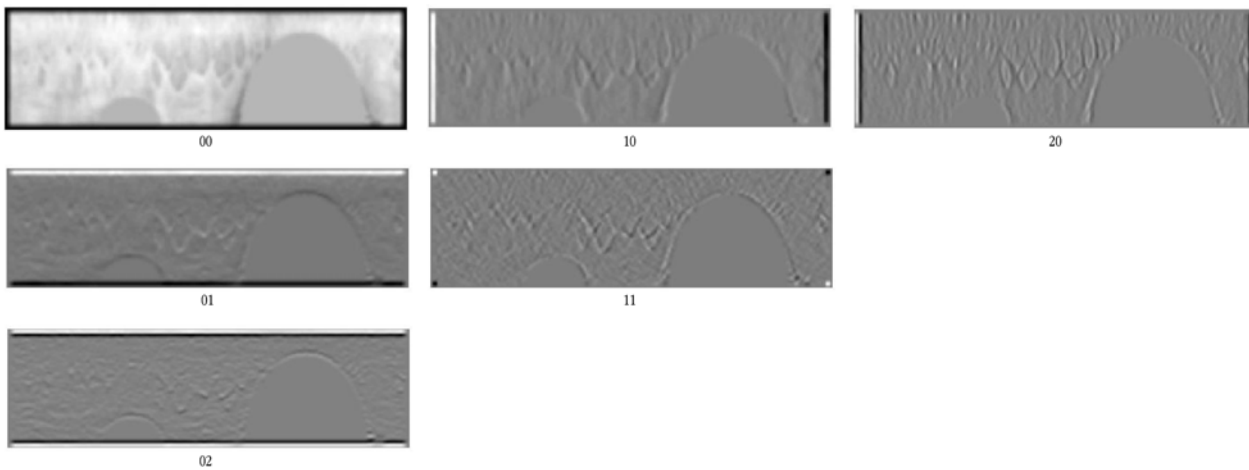
$$f_{re} = f_{2,0} = 1 \text{ si } \int_x \int_y f(x, y) D_{2,0}(x_0 - x, y_0 - y) \geq 0 \quad (3.34)$$

$$f_{re} = f_{2,0} = 0 \text{ si } \int_x \int_y f(x, y) D_{2,0}(x_0 - x, y_0 - y) < 0 \quad (3.35)$$

$$f_{im} = f_{1,0} = 1 \text{ si } \int_x \int_y f(x, y) D_{1,0}(x_0 - x, y_0 - y) \geq 0 \quad (3.36)$$

$$f_{im} = f_{1,0} = 0 \text{ si } \int_x \int_y f(x, y) D_{1,0}(x_0 - x, y_0 - y) < 0 \quad (3.37)$$

A diferencia de los filtros basados en las funciones Gabor, los filtros de análisis de la transformada Hermite no necesitan corrección de DC y el proceso de obtención de los coeficientes se realiza directamente usando los polinomios de Krawtchouk (ecuación 2.47) como se vio en la sección 2.4.3.

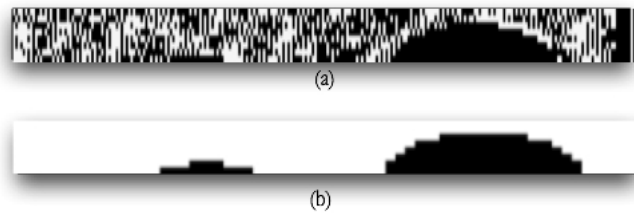


**Figura 3.15:** Descomposición Hermite  $N = 6$ ,  $D_{max} = 2$ .

Una descomposición de hasta segundo orden de la textura normalizada de un iris, a través de la

transformada Hermite, puede obtenerse como se muestra en la figura 3.15. Los coeficientes numerados como “10” y “20” representan a la textura de iris localizada-filtrada con los filtros de análisis Hermite impar y par respectivamente en la dirección  $x$ . Mientras que los coeficientes numerados como “01” y “02” representan a la textura de iris localizada-filtrada con los filtros de análisis Hermite impar y par respectivamente en la dirección  $y$ . El coeficiente “00” representa la respuesta al filtro 2D de análisis Hermite de orden cero, es decir, la componente de baja frecuencia de una textura de iris. Por último, el coeficiente “11” representa a la textura de iris localizada-filtrada con un filtro de orden 1 tanto en la dirección  $x$  como en la dirección  $y$ .

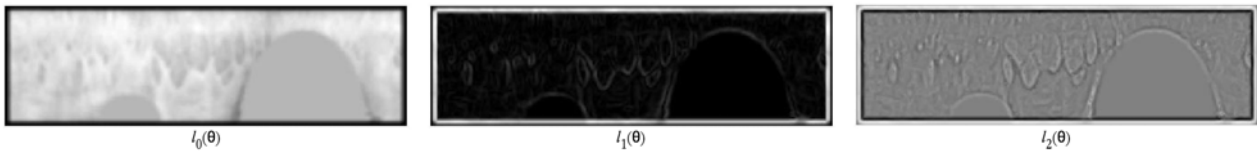
El primer extractor de características empleado fue implementado al combinar los coeficientes “10” y “20” en un código binario. Los coeficientes obtenidos fueron agrupados en una forma similar al algoritmo de Daugman [21]. Es decir, a partir de los coeficientes “10” y “20” se codificó el signo de la fase aproximada en un código de 256 bytes de acuerdo a las ecuaciones 3.34-3.37. Un ejemplo de este código obtenido (amplificado para fines de visualización) se muestra en la figura 3.16.



**Figura 3.16:** Código de iris usando la transformada Hermite de primer y segundo orden. (a) Código. (b) Máscara de artefactos, puede contener información de la región del iris que es obstruida por pestañas, párpados o algunos reflejos.

El segundo extractor de características empleado fue el obtenido a partir de una proyección de los coeficientes 2D a 1D usando la ecuación 2.54 que se vio en la sección 2.4.5, como se muestra en la figura 3.17. En este caso, únicamente se utilizó el coeficiente proyectado  $l_2(\theta)$  para generar el código del iris de manera similar al algoritmo de Daugman [21], es decir:

$$l_2(\theta) = \begin{cases} 1 & \text{si } l_2(\theta) \geq 0 \\ 0 & \text{si } l_2(\theta) < 0 \end{cases} \quad (3.38)$$



**Figura 3.17:** Proyección de la transformada Hermite cartesiana 2D a la transformada Hermite orientada 1D.

### 3.3.5. Implementación del algoritmo de comparación (*matching*)

El algoritmo de comparación de plantillas o códigos fue implementado usando la distancia de Hamming presentada en la ecuación 3.17 en la sección 3.2.2.3. Para lograr que sea invariante a la rotación,

### 3 Reconocimiento de Iris

durante la etapa de *matching* entre dos códigos, se rota de forma cíclica uno de ellos mientras el otro permanece fijo. Para cada ángulo discreto rotado, se mide la distancia de Hamming entre ambos códigos. La distancia de Hamming mínima, encontrada después de rotar un número discreto de ángulos, es considerada como la distancia de Hamming entre ambos códigos.

## 4 Análisis de resultados

### 4.1. Descripción de las bases de datos de imágenes de iris utilizadas

A continuación se describen las bases de datos de imágenes de iris utilizadas. Cabe señalar que el número total de imágenes de prueba, a las cuales se les extrae satisfactoriamente su código de iris, puede variar y ser menor que el número total de imágenes en las bases de datos, a las cuales se les pretende extraer su código de iris o plantilla biométrica correspondiente. Ésto se debe a que durante el proceso de localización o segmentación algunas imágenes son descartadas automáticamente por el algoritmo, ya que esta etapa aún es vulnerable a localizar erróneamente las regiones de la pupila y del iris. Cuando la circunferencia que delimita a la pupila está demasiado cerca o bien ésta se encuentra fuera de la circunferencia que delimita al iris, entonces la imagen es descartada para el procesamiento subsecuente, ya que es posible que el algoritmo haya localizado incorrectamente alguna de estas regiones, como consecuencia de la presencia de artefactos (como pestañas y párpados) en la imagen del iris. Aún así, el número de comparaciones que se logran obtener puede considerarse suficiente para realizar los análisis estadísticos. Posteriormente se indicará el número de comparaciones (intra-clase e inter-clase) realizadas para cada base de datos de imágenes de iris.

#### 4.1.1. Base de datos de imágenes de iris de nuestra propiedad

La base de datos de imágenes de iris de nuestra propiedad consiste de 120 imágenes de  $1280 \times 1024$  píxeles. Esta base de datos cuenta con imágenes de iris provenientes de 15 usuarios. Cada uno de los usuarios proporcionó alrededor de 8 imágenes de iris de un mismo ojo, por lo que sólo se cuenta con un conjunto. El formato de las imágenes es bmp y éstas se encuentran en escala de grises, pues fueron adquiridas usando iluminación infrarroja a través de nuestro propio sistema de adquisición. Para acelerar el proceso de localización del iris, las imágenes son escaladas por el mismo sistema. Las características de textura son extraídas sobre la imagen sin escalar.

#### 4.1.2. Base de datos de imágenes de iris CASIA

La base de datos de imágenes de iris CASIA [14] consiste de 756 imágenes de  $320 \times 280$  píxeles. Esta base de datos cuenta con imágenes de iris provenientes de 108 usuarios. Cada uno de los usuarios proporcionó alrededor de 7 imágenes de iris en dos conjuntos; un primer conjunto con 3 imágenes de un mismo ojo y el segundo conjunto con 4 imágenes del otro ojo. El formato de las imágenes es bmp y éstas se encuentran en escala de grises, pues fueron adquiridas usando iluminación infrarroja. En el desarrollo de este proyecto se utilizó el segundo conjunto de imágenes, es decir 432 imágenes. Dentro de este conjunto existen alrededor de 20 imágenes cuya información de textura de iris es nula debido a que el iris está obstruido casi totalmente por los párpados y pestañas. Dichas imágenes fueron descartadas en forma manual por simple inspección. El total de imágenes utilizadas fue de 402.



## 4.2. Teoría de la decisión estadística

El problema de reconocer que la firma de un iris dado le pertenece a un individuo en particular, ya sea después de una búsqueda exhaustiva dentro de una base de datos grande (identificación) o simplemente por la comparación con una sola plantilla de autenticación (verificación), puede ser formulado bajo el esquema de la teoría de la decisión estadística. Este tipo de problemas puede agruparse en la teoría de la decisión dentro de los problemas de decisión binaria.

Por una decisión binaria se entiende que hay dos tipos de mensajes  $m_1$  y  $m_2$  en el espacio de mensajes  $M$  y que el espacio de decisión  $D$  tiene también sólo dos elementos  $d_1$  y  $d_2$ . El mensaje  $m_1$  es enviado si  $d_1$  es correcta. El problema es seleccionar la regla de decisión  $d(z)$  que mapea, de alguna manera óptima, el espacio de observaciones  $Z$  a un espacio de decisiones binarias. Debido a que sólo hay dos decisiones posibles, es equivalente a dividir  $Z$  en dos regiones de decisión  $Z_1$  y  $Z_2$  de tal manera que  $d(z) = d_1$  si  $z \in Z_1$  y  $d(z) = d_2$  si  $z \in Z_2$ . Las regiones  $Z_1$  y  $Z_2$  deben ser disjuntas ( $Z_1 Z_2 = \emptyset$ ) para que cada punto en  $Z$  tenga una decisión única. Las regiones  $Z_1$  y  $Z_2$  deben cubrir  $Z$  ( $Z_1 \cup Z_2 = Z$ ) para que cada punto tenga una decisión asociada en  $Z$  [56].

Existen varios criterios en el diseño de una regla de decisión: criterio de máxima similitud, criterio de Neyman-Pearson, criterio de probabilidad de error, criterio de riesgo de Bayes, criterio min-max. El criterio de máxima similitud es una de las técnicas más simples que se pueden encontrar; sin embargo, representar problemas reales adecuadamente usando este criterio puede resultar poco confiable por lo que es necesario el uso de otra técnica [56]. En este trabajo se analizaron los resultados que se obtuvieron al comparar las plantillas de iris usando el criterio de Neyman-Pearson, por lo que únicamente se describe en que consiste dicho criterio.

### 4.2.1. Criterio Neyman-Pearson

Antes de presentar el criterio de Neyman-Pearson es conveniente introducir algunas definiciones y la notación utilizada. En un problema de decisión binaria existen dos tipos de errores que se pueden cometer. Primero, es posible decidir  $d_2$  cuando  $m_1$  es verdadero. Segundo, podemos decidir  $d_1$  cuando  $m_2$  es verdadero. Cada uno de estos errores tiene asociada una probabilidad de error, la cual depende de la regla de decisión y de las funciones de densidad de probabilidad condicionales [56]:

$$P\{d_2|m_1\} = \text{Probabilidad de tomar la decisión } d_2 \text{ cuando } m_1 \text{ es verdadero}$$

$$P\{d_1|m_2\} = \text{Probabilidad de tomar la decisión } d_1 \text{ cuando } m_2 \text{ es verdadero}$$

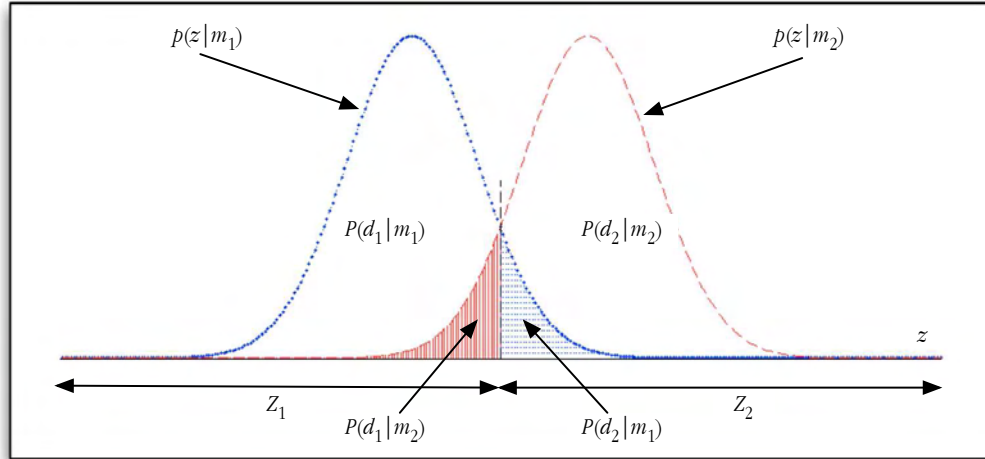
las cuales pueden expresarse en términos de las regiones de decisión como:

$$P\{d_2|m_1\} = P\{z \in Z_2|m_1\} = \int_{Z_2} p(z|m_1)dz \quad (4.1)$$

$$P\{d_1|m_2\} = P\{z \in Z_1|m_2\} = \int_{Z_1} p(z|m_2)dz \quad (4.2)$$

Al error de decidir  $d_2$  cuando  $m_1$  es verdadero se le llama frecuentemente error de tipo uno mientras que decidir  $d_1$  cuando  $m_2$  es verdadero se le llama error de tipo dos. En la figura 4.1 se pueden apreciar estas regiones de decisión. En el contexto de la toma de una decisión biométrica, llamaremos a estos errores como error de falso rechazo (FR) y error de falso acceso (FA) respectivamente.

Además de los dos errores, existen también dos decisiones correctas que se pueden efectuar. Podemos decidir  $d_1$  cuando  $m_1$  es verdadero y también podemos decidir  $d_2$  cuando  $m_2$  es verdadero (Véase



**Figura 4.1:** Distribuciones en un problema típico de decisión binaria.

figura 4.1). Estas decisiones correctas tienen las siguientes probabilidades asociadas [56]:

$P\{d_1|m_1\}$  = Probabilidad de tomar la decisión  $d_1$  cuando  $m_1$  es verdadero

$P\{d_2|m_2\}$  = Probabilidad de tomar la decisión  $d_2$  cuando  $m_2$  es verdadero

que también pueden ser expresadas en términos de las regiones de decisión:

$$P\{d_1|m_1\} = P\{z \in Z_1|m_1\} = \int_{Z_1} p(z|m_1)dz \quad (4.3)$$

$$P\{d_2|m_2\} = P\{z \in Z_2|m_2\} = \int_{Z_2} p(z|m_2)dz \quad (4.4)$$

Nuevamente en el contexto biométrico, llamaremos a estas dos decisiones correctas como correcto acceso (CA) y correcto rechazo (CR). Debido a que sólo se pueden tomar las decisiones  $d_1$  o  $d_2$ , las probabilidades de error más la probabilidad de una decisión correcta deben sumar uno para ambos mensajes, de tal manera que tenemos:

$$P\{d_1|m_1\} + P\{d_2|m_1\} = 1 \quad (4.5)$$

$$P\{d_2|m_2\} + P\{d_1|m_2\} = 1 \quad (4.6)$$

Este resultado se puede obtener al sustituir las ecuaciones 4.2 y 4.3 en la ecuación 4.5:

$$\begin{aligned} P\{d_1|m_1\} + P\{d_2|m_1\} &= \int_{Z_1} p(z|m_1)dz + \int_{Z_2} p(z|m_1)dz \\ &= \int_Z p(z|m_1)dz = 1 \end{aligned}$$

En nuestro contexto, el *nivel de significancia*  $P(d_1|m_2)$  representa la probabilidad de incorrectamente anunciar que una persona es el individuo quien dice ser. La *potencia de la prueba*  $P(d_1|m_1)$  describe la probabilidad de anunciar que un individuo es en efecto quien reclama ser. Con el objetivo de lograr una buena decisión, es necesario maximizar tanto como sea posible la potencia de la prueba  $P(d_1|m_1)$  y hacer  $P(d_1|m_2)$  lo más pequeña posible. Sin embargo, no es posible simultáneamente

maximizar  $P(d_1|m_1)$  y minimizar  $P(d_1|m_2)$ . Cuando se aumenta  $P(d_1|m_1)$  también se tiende a incrementar  $P(d_1|m_2)$  y viceversa si se disminuye  $P(d_1|m_2)$  también se decreta  $P(d_1|m_1)$ . El criterio de Neyman-Pearson provee de una solución posible a esta paradoja.

**Criterio de decisión Neyman-Pearson:** Fijar el valor de  $P(d_1|m_2)$  en un valor preseleccionado  $\alpha_0$  y luego maximizar  $P(d_1|m_1)$  [56].

El criterio de Neyman-Pearson dice que se debe maximizar la potencia de la prueba para un nivel dado de *significancia*. En otras palabras, afuera de todas las regiones de decisión  $Z_1$  para el cual  $P(d_1|m_2) = \alpha_0$  se tiene que seleccionar aquella para la cual  $P(d_1|m_1)$  es máxima. En este criterio la regla de decisión esta especificada completamente por el valor de  $\alpha_0$ .

El uso de una curva de característica de operación de receptor (ROC<sup>1</sup>) provee una guía en la selección del valor  $\alpha_0$ . La ROC es una gráfica de la probabilidad de detección (potencia)  $P(d_1|m_1)$  para la regla de decisión de Neyman-Pearson contra la probabilidad del nivel de *significancia*  $P(d_1|m_2)$  como una función de uno o más parámetros. La ROC puede ser obtenida determinando  $P(d_1|m_1)$  como una función de  $\alpha_0 = P(d_1|m_2)$ , variando  $\alpha_0$  desde cero hasta uno.

#### 4.2.2. Aplicación de la teoría de la decisión en el reconocimiento de iris

Una distribución de la distancia de similitud se puede obtener al realizar las comparaciones de un código perteneciente a una persona, el cual es obtenido durante el proceso de extracción de características, con respecto a todos los códigos generados. Entre más cercana sea a cero esta similitud, significa que la semejanza entre los códigos es mayor. Realizando estas comparaciones para todos los individuos registrados en la base de datos de imágenes de iris, comparando para cada uno de ellos los códigos pertenecientes a las mismas personas, nos produce una distribución intra-clase. Por otro lado, realizar comparaciones entre códigos que pertenecen a diferentes personas, para todos los individuos registrados en la base de datos de imágenes de iris, produce una distribución inter-clase. Un ejemplo de estas distribuciones lo podemos observar en la figura 4.2. En dicha figura se muestran las dos distribuciones obtenidas experimentalmente, usando el extractor de características basado en la transformada Hermite, para la base de datos de imágenes de iris de nuestra propiedad. El número de comparaciones intra-clase fue  $n_1 = 378$ , mientras que el número de comparaciones inter-clase fue  $n_2 = 4182$ .

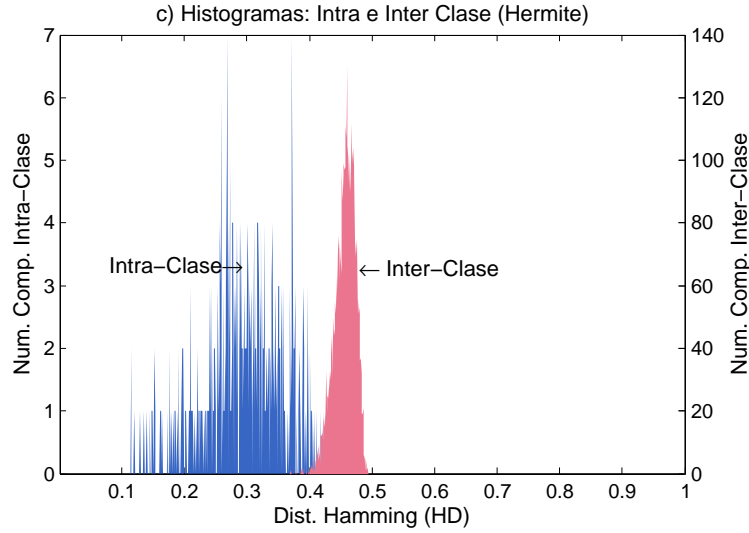
El desempeño de un sistema, el cual debe tomar decisiones del tipo si/no, también está determinado por la cantidad de traslape que existe entre las dos distribuciones. Una medida de separación entre las dos distribuciones intra-clase e inter-clase puede escribirse como:

$$d' = \frac{|\mu_1 - \mu_2|}{\sqrt{\frac{1}{2}(\sigma_1^2 + \sigma_2^2)}} \quad (4.7)$$

donde  $\mu_1$  y  $\sigma_1$  son la media y la desviación estándar de la distribución intra-clase respectivamente y  $\mu_2$  y  $\sigma_2$  la media y la desviación estándar de la distribución inter-clase respectivamente. En el ejemplo de la figura 4.2, la distancia de separación entre las dos distribuciones, usando el extractor de características basado en la transformada Hermite, fue  $d' = 3.4405$ .

Sean las dos distribuciones,  $P_{Au}(x)$  y  $P_{Im}(x)$  las que denotan las densidades de probabilidad, obtenidas experimentalmente a partir de la medida de similitud (distancia de Hamming), de dos fuentes biométricas distintas: auténticos e impostores (intra-clase e inter-clase) respectivamente. Entonces podemos aproximar las probabilidades dadas por las ecuaciones 4.1, 4.2, 4.3 y 4.4 a partir de las tasas

<sup>1</sup>ROC, por sus siglas en inglés "Receiver Operating Characteristic".



**Figura 4.2:** Histogramas obtenidos al comparar la distancia de Hamming usando la transformada Hermite. Lado izquierdo comparaciones intra-clase; media  $\mu_1 = 0.2932$ , desviación estándar  $\sigma_1 = 0.0648$ . Lado derecho comparaciones inter-clase; media  $\mu_2 = 0.4558$ , desviación estándar  $\sigma_2 = 0.0164$ . Separación  $d' = 3.4405$ .

obtenidas experimentalmente como:

$$P\{d_1|m_1\} = CAR = \int_0^C P_{Au}(x)dx \quad (4.8)$$

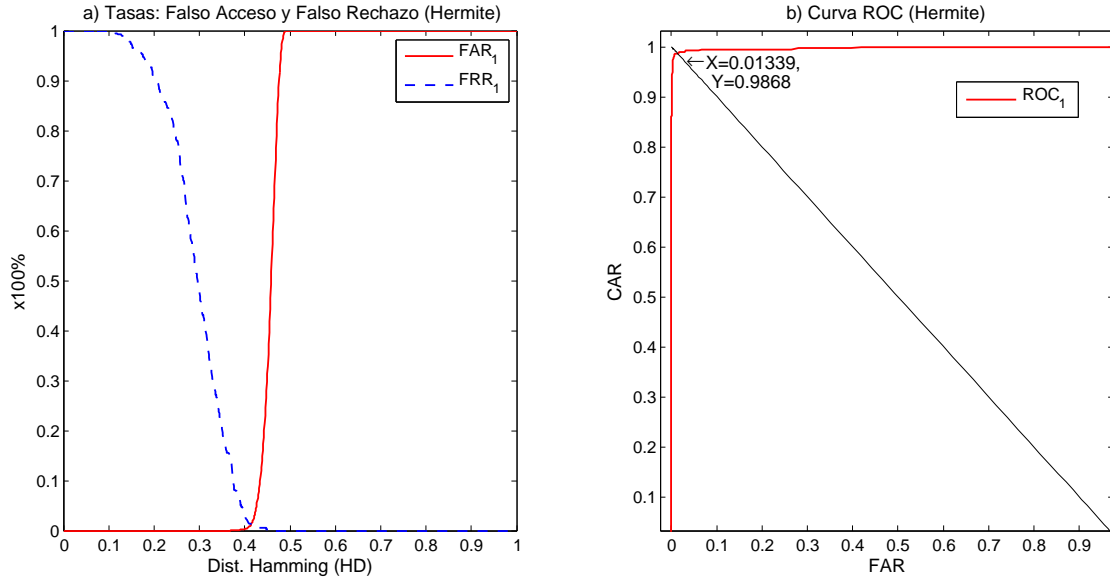
$$P\{d_2|m_1\} = FRR = \int_C^1 P_{Au}(x)dx \quad (4.9)$$

$$P\{d_1|m_2\} = FAR = \int_0^C P_{Im}(x)dx \quad (4.10)$$

$$P\{d_2|m_2\} = CRR = \int_C^1 P_{Im}(x)dx \quad (4.11)$$

donde  $m_1$  es el mensaje de acceso que indica que un usuario ha sido reconocido (recordando que el reconocimiento puede ser verificación o bien identificación, véase sección 1.1), dentro de la base de datos de las imágenes de iris. El mensaje  $m_2$  es de rechazo, es decir, que el usuario no fue reconocido, dentro de la base de datos de las imágenes de iris.  $d_1$  es la decisión de otorgar acceso y  $d_2$  es la decisión de rechazar el acceso a un individuo que el sistema puede tomar.  $CAR$  es la tasa de otorgar accesos correctamente,  $FRR$  es la tasa de error al realizar incorrectamente rechazos (falso negativo),  $FAR$  es la tasa de error al otorgar incorrectamente accesos (falso positivo) y  $CRR$  es la tasa de realizar rechazos de forma correcta y son iguales a las áreas bajo las dos distribuciones de probabilidad en ambos lados del criterio de decisión  $C$  elegido como lo indican las ecuaciones 4.8-4.11.

Cometer un falso rechazo puede ser menos costoso que cometer un falso acceso dependiendo de los niveles de seguridad donde se requiera implantar el sistema. Las curvas presentadas en la figura 4.3 permiten realizar una interpretación del nivel de confianza del sistema. En el ejemplo de la gráfica (a) en la figura 4.3, usando la transformada Hermite se observa una tasa de error igual ( $EER$ ) cercano al 1.34% ( $FAR = FRR = 1.34\%$ ). De la gráfica (b) en la figura 4.3, la tasa de accesos



**Figura 4.3:** Ejemplo de tasas de error  $FRR$ ,  $FAR$  y curva de característica de operación de receptor (ROC) usando la transformada Hermite para la base de datos de imágenes de iris de nuestra propiedad. (a) Tasa de error de falso acceso  $FAR$  (línea continua) y tasa de error de falso rechazo  $FRR$  (línea discontinua). (b) Curva ROC.

correctos ( $CAR$ ) para una tasa de error de  $FAR = 0\%$  fue cercano al 86% mientras que en la misma gráfica se observa un desempeño general del sistema cercano al 98.66% para la tasa de error igual ( $FAR = FRR = 1.34\%$ ).

Con estas tasas es posible establecer un umbral de la distancia de Hamming aproximadamente igual a 0.36, bajo el cual se permite de forma correcta rechazar a todos los individuos que son impostores y reconocer al 86% de los individuos registrados en el sistema, pero con la desventaja de rechazar equivocadamente al 14% de individuos que no son impostores. Este fue uno de los resultados obtenidos usando la transformada Hermite como modelo de extracción de características para la base de datos de nuestra propiedad obtenida en el laboratorio.

A continuación se muestran las pruebas experimentales que se realizaron con la base de datos de imágenes de iris de nuestra propiedad y con la base de datos de imágenes de iris CASIAv1, usando ambos modelos de extracción de características: wavelets Gabor y la transformada Hermite. Por último se seleccionan los mejores resultados con cada uno de los modelos para efectos de comparación del desempeño obtenido.

### 4.3. Resultados usando la base de datos de imágenes de iris de nuestra propiedad

#### Extractor basado en wavelets Gabor

Las tablas 4.1-4.2 muestran los resultados obtenidos para la base de datos de imágenes de iris de nuestra propiedad usando la wavelet Gabor como extractor de características. Las pruebas consistieron

en variar algunos parámetros del extractor de características y de la formación del código.

Se seleccionaron dos anchos de banda angular distintos  $B_\theta = \pi/4$  y  $B_\theta = \pi/6$ . Para cada  $B_\theta$  se generaron códigos de iris, con los cuales se obtuvieron dos distribuciones de la distancia de Hamming al comparar códigos intra-clase (la distribución obtenida cuando se comparan los códigos de un mismo individuo, para cada individuo) y códigos inter-clase (la distribución obtenida cuando se comparan códigos de un individuo con respecto a los códigos de los demás individuos, para cada individuo). El número de comparaciones intra-clase e inter-clase para todos los casos fue de  $n_1 = 378$  y  $n_2 = 4182$  respectivamente. Las distribuciones arrojaron las estadísticas observadas en dichas tablas. Así mismo, se eligió el número de bandas o regiones de análisis de 8 y 4 bandas. Este número de regiones de análisis impuso un límite con respecto al tamaño del código de iris. Los tamaños de los códigos de iris utilizados fueron de 2048 bits y 1024 bits como se observa en las tablas 4.1-4.2.

**Tabla 4.1:** Tabla comparativa de los resultados obtenidos al variar el tamaño del código o plantilla biométrica y el número de bandas o regiones de análisis. Para un ancho de banda angular  $B_\theta = \pi/4$  usando wavelets Gabor, el mejor resultado se observó con un código de 2048 bits y 8 bandas de análisis.

$B_\theta = \pi/4$	1024 bits / 4 bandas	1024 bits / 8 bandas	2048 bits / 8 bandas
$n_1$	378	378	378
$\mu_1$	0.2953	0.3004	<b>0.2971</b>
$\sigma_1$	0.0671	0.0704	<b>0.0689</b>
$n_2$	4182	4182	4182
$\mu_2$	0.4403	0.4477	<b>0.4469</b>
$\sigma_2$	0.0216	0.0176	<b>0.0176</b>
$d'$	2.9098	2.8708	<b>2.9815</b>
$EER$	4.18%	5.1%	<b>4.16%</b>
Desempeño	95.82%	94.9%	<b>95.84%</b>
$CAR(FAR = 0\%)$	47.88%	69.05%	<b>83.86%</b>
$FRR(FAR = 0\%)$	52.12%	30.95%	<b>16.14%</b>

A partir de la tabla 4.1 es posible observar, para un filtro que posee un ancho de banda de  $B_\theta = \pi/4$ , que al dividir la región de interés en 8 bandas de análisis y generar un código de 2048 bits, se obtuvieron los mejores resultados para dicho ancho de banda. De esta manera, la distancia de separación entre clases fue de  $d' = 2.9815$  y la tasa de error igual ( $EER = FAR = FRR$ ) fue cercana al 4.16%. Para el caso en que las tasas de error son iguales, se obtuvo un desempeño general del sistema de 95.84%. Sin embargo, para poder evaluar el desempeño del sistema de reconocimiento de iris en condiciones de alta seguridad, fijamos un criterio de decisión estricto, bajo el cual a ningún impostor se le es permitido ingresar al sistema. En dicho caso, la tasa de error de otorgar incorrectamente accesos a impostores debe ajustarse a cero ( $FAR = 0\%$ ), asegurando que ningún impostor pueda ganar acceso al sistema. Como consecuencia, algunos usuarios son rechazados o bien se les solicita que registren la imagen del iris para intentar nuevamente acceder al sistema. Esta tasa de error, al realizar incorrectamente rechazos de usuarios registrados en el sistema, fue de  $FRR = 16.14\%$ , es decir, al 83.86% de los usuarios se les permitirá el acceso de manera correcta ( $CAR$ ).

El análisis para la tabla 4.2 es muy similar al caso anterior. En esta tabla es posible observar que usando un código de 2048 bits y dividiendo la región de interés en 8 bandas de análisis, cuando el filtro posee un ancho de banda de  $B_\theta = \pi/6$ , fue el que arrojó mejores resultados. La separación entre las clases fue de  $d' = 3.0352$ . La tasa de error igual ( $EER = FAR = FRR$ ) obtenida fue cercana

**Tabla 4.2:** Tabla comparativa de los resultados obtenidos al variar el tamaño del código o plantilla biométrica y el número de bandas o regiones de análisis. Para un ancho de banda angular  $B_\theta = \pi/6$  usando wavelets Gabor, el mejor resultado se observó con un código de 2048 bits y 8 bandas de análisis.

$B_\theta = \pi/6$	1024 bits / 4 bandas	1024 bits / 8 bandas	2048 bits / 8 bandas
$n_1$	378	378	378
$\mu_1$	0.2901	0.2963	<b>0.2940</b>
$\sigma_1$	0.0669	0.0703	<b>0.0688</b>
$\mu_2$	0.4396	0.4470	<b>0.4464</b>
$n_2$	4182	4182	4182
$\sigma_2$	0.0221	0.0177	<b>0.0177</b>
$d'$	3.0008	2.9396	<b>3.0352</b>
$EER$	4.3%	4.6%	<b>3.73%</b>
Desempeño	95.7%	95.4%	<b>96.27%</b>
$CAR(FAR = 0\%)$	46.03%	73.55%	<b>84.66%</b>
$FRR(FAR = 0\%)$	53.97%	26.45%	<b>15.34%</b>

al 3.73% y el desempeño general del sistema obtenido para el caso en que las tasas de error son iguales fue de 96.27%. Evaluando el desempeño del sistema de reconocimiento de iris en condiciones de alta seguridad, fijamos un criterio de decisión estricto, bajo el cual la tasa de error de otorgar incorrectamente accesos a impostores se ajusta a cero ( $FAR = 0\%$ ), asegurando que ningún impostor pueda obtener acceso al sistema. Como consecuencia, algunos usuarios son rechazados o bien se les solicita nuevamente registren la imagen del iris para intentar acceder al sistema. La tasa de error al rechazar incorrectamente a usuarios registrados en el sistema fue de  $FRR = 15.34\%$ , es decir, al 84.66% de los usuarios se les permitirá el acceso de manera correcta ( $CAR$ ).

Con estos resultados es posible observar para la tabla 4.2, al igual que ocurrió para la tabla 4.1, que con un código de 2048 bits formado a partir de 8 bandas de análisis se obtuvieron los mejores resultados. Al comparar el mejor resultado obtenido en cada una de las tablas (tablas 4.1 y 4.2), es posible seleccionar los parámetros del código, número de bandas de análisis y características del filtro, que permitirán obtener una mayor separación entre las clases intra-clase e inter-clase. De esta manera, el mejor desempeño y la mayor distancia de separación entre clases fue obtenida con un filtro de ancho de banda angular  $B_\theta = \pi/6$  sobre 8 regiones o bandas de análisis y codificando los coeficientes de las proyecciones en un código de 2048 bits. La separación entre clases proporcionada por dichos parámetros fue de  $d' = 3.0352$  y el desempeño de accesos correctos ( $CAR$ ) fue cercano al 84.66% para una tasa de falsos accesos  $FAR$  igual a 0%. El desempeño general del sistema fue de 96.27%.

### Extractor basado en la transformada Hermite

Pruebas similares a las anteriores fueron realizadas para el caso en que el extractor de características fue la transformada Hermite. Las tablas 4.3-4.6 muestran estos resultados para la base de datos de imágenes de iris de nuestra propiedad.

Los parámetros de prueba consistieron en elegir distintas escalas de análisis de la transformada Hermite, variando  $N = 4, 6, 8$  y 10. Se generaron códigos de iris para cada  $N$ , con los cuales dos distribuciones de la distancia de Hamming, obtenidas a partir de las comparaciones intra-clase y de las comparaciones inter-clase, arrojaron las estadísticas observadas en dichas tablas. El número de

### 4.3 Resultados usando la base de datos de imágenes de iris de nuestra propiedad

comparaciones intra-clase e inter-clase en todos los casos fue de  $n_1 = 378$  y  $n_2 = 4182$  respectivamente. El número de bandas o regiones de análisis también se varió de 8 a 4 bandas. El número de regiones impuso un límite con respecto al tamaño del código de iris. Estos códigos de iris fueron de 2048 bits y 1024 bits.

**Tabla 4.3:** Tabla comparativa de los resultados obtenidos al variar el tamaño del código o plantilla biométrica y el número de bandas o regiones de análisis. Para un ancho de banda angular  $N = 4$  usando la transformada Hermite, el mejor resultado se observó con un código de 2048 bits y 8 bandas de análisis.

$N = 4$	1024 bits / 4 bandas	1024 bits / 8 bandas	2048 bits / 8 bandas
$n_1$	378	378	378
$\mu_1$	0.3141	0.3185	<b>0.3176</b>
$\sigma_1$	0.0677	0.0681	<b>0.0659</b>
$n_2$	4182	4182	4182
$\mu_2$	0.4559	0.4549	<b>0.4636</b>
$\sigma_2$	0.0183	0.0145	<b>0.0141</b>
$d'$	2.8599	2.7704	<b>3.0634</b>
$EER$	3.4%	4.0%	<b>1.93%</b>
Desempeño	96.6%	96.0%	<b>98.07%</b>
$CAR(FAR = 0\%)$	54.5%	72.49%	<b>84.13%</b>
$FRR(FAR = 0\%)$	45.5%	27.51%	<b>15.87%</b>

En la tabla 4.3 se observa que al usar un código de 2048 bits y dividir la región de interés en 8 bandas de análisis, cuando el tamaño de la ventana del filtro es  $N = 4$ , se obtienen mejores resultados.

**Tabla 4.4:** Tabla comparativa de los resultados obtenidos al variar el tamaño del código o plantilla biométrica y el número de bandas o regiones de análisis. Para un ancho de banda angular  $N = 6$  usando la transformada Hermite, el mejor resultado se observó con un código de 2048 bits y 8 bandas de análisis.

$N = 6$	1024 bits / 4 bandas	1024 bits / 8 bandas	2048 bits / 8 bandas
$n_1$	378	378	378
$\mu_1$	0.2978	0.3070	<b>0.3023</b>
$\sigma_1$	0.0677	0.0663	<b>0.0657</b>
$n_2$	4182	4182	4182
$\mu_2$	0.4525	0.4532	<b>0.4601</b>
$\sigma_2$	0.0195	0.0152	<b>0.0154</b>
$d'$	3.1067	3.0387	<b>3.3043</b>
$EER$	2.87%	1.84%	<b>1.91%</b>
Desempeño	97.13%	98.16%	<b>98.09%</b>
$CAR(FAR = 0\%)$	60.85%	85.71%	<b>90.74%</b>
$FRR(FAR = 0\%)$	39.15%	14.29%	<b>9.26%</b>

La distancia de separación entre las clases fue de  $d' = 3.0634$  y la tasa de error igual ( $EER = FAR = FRR$ ) fue cercana al 1.93%. Cuando las tasas de error son iguales el desempeño general del sistema obtenido fue de 98.07%. Evaluando el desempeño del sistema de reconocimiento de iris en



condiciones de alta seguridad, fijamos un criterio de decisión estricto, bajo el cual a ningún impostor se le es permitido ingresar al sistema. En dicho caso, la tasa de error de otorgar incorrectamente accesos a impostores debe ajustarse a cero ( $FAR = 0\%$ ), asegurando que ningún impostor pueda ganar acceso al sistema. El porcentaje de usuarios que incorrectamente son rechazados por el sistema, o bien que se les solicita registren la imagen del iris para intentar nuevamente acceder al sistema fue de  $FRR = 15.87\%$ , es decir, a un 84.13% de los usuarios se les permitirá el acceso de manera correcta ( $CAR$ ).

**Tabla 4.5:** Tabla comparativa de los resultados obtenidos al variar el tamaño del código o plantilla biométrica y el número de bandas o regiones de análisis. Para un ancho de banda angular  $N = 8$  usando la transformada Hermite, el mejor resultado se observó con un código de 2048 bits y 8 bandas de análisis.

$N = 8$	1024 bits / 4 bandas	1024 bits / 8 bandas	2048 bits / 8 bandas
$n_1$	378	378	378
$\mu_1$	0.2850	0.2957	<b>0.2932</b>
$\sigma_1$	0.0669	0.0661	<b>0.0648</b>
$n_2$	4182	4182	4182
$\mu_2$	0.4486	0.4517	<b>0.4558</b>
$\sigma_2$	0.0212	0.0157	<b>0.0164</b>
$d'$	3.2972	3.2492	<b>3.4405</b>
$EER$	2.82%	2.9%	<b>1.34%</b>
Desempeño	97.18%	97.10%	<b>98.66%</b>
$CAR(FAR = 0\%)$	73.55%	85.45%	<b>85.98%</b>
$FRR(FAR = 0\%)$	26.45%	14.55%	<b>14.02%</b>

Como se mencionó anteriormente, este tipo de pruebas fueron realizadas para distintos tamaños de ventana y los resultados pueden observarse en las tablas 4.4-4.6 para  $N = 6, 8$  y  $N = 10$  respectivamente. En todas las tablas es común observar que la mejor distancia de separación entre clases (intra-clase e inter-clase), así como también el mejor desempeño al ajustar la tasa de error de otorgar incorrectamente accesos a impostores ( $FAR$ ) igual a 0%, se logra cuando la región de interés se divide en 8 bandas de análisis y se generan códigos de 2048 bits.

A partir de los mejores resultados obtenidos para cada tamaño de ventana ( $N = 4, 6, 8$  y  $10$ ), es posible determinar los parámetros del código, número de bandas de análisis y características del filtro, que permiten obtener una mayor distancia de separación entre las clases intra-clase e inter-clase, así como también un mejor desempeño. Comparando el mejor resultado obtenido en cada una de las tablas (4.3 - 4.6), se encontró que al usar los filtros de análisis Hermite (real e impar) que compartían un tamaño de ventana de  $N = 6$  (tabla 4.4), proporcionaban la mayor distancia de separación entre clases ( $d' = 3.3043$ ) y un mejor desempeño de accesos correctos ( $CAR = 90.74\%$ ), bajo un criterio de decisión en el cual a ningún impostor se le es permitido ingresar al sistema ( $FAR = 0\%$ ), con respecto a los otros filtros. De esta manera, se obtuvo una tasa de error de usuarios incorrectamente rechazados de  $FRR = 9.26\%$  y una tasa de error igual ( $EER$ ) cercana al 1.91%, proporcionando un desempeño general del sistema del 98.09%. Como fue mencionado anteriormente, estos resultados fueron obtenidos al dividir la región de interés en 8 bandas de análisis y se generaron códigos o plantillas biométricas de 2048 bits.

**Tabla 4.6:** Tabla comparativa de los resultados obtenidos al variar el tamaño del código o plantilla biométrica y el número de bandas o regiones de análisis. Para un ancho de banda angular  $N = 10$  usando la transformada Hermite, el mejor resultado se observó con un código de 2048 bits y 8 bandas de análisis.

$N = 10$	1024 bits / 4 bandas	1024 bits / 8 bandas	2048 bits / 8 bandas
$n_1$	378	378	378
$\mu_1$	0.2783	0.2869	<b>0.2881</b>
$\sigma_1$	0.0677	0.0660	<b>0.0642</b>
$n_2$	4182	4182	4182
$\mu_2$	0.4458	0.4470	<b>0.4523</b>
$\sigma_2$	0.0215	0.0168	<b>0.0175</b>
$d'$	3.3343	3.3262	<b>3.4868</b>
$EER$	2.94%	1.82%	<b>1.7%</b>
Desempeño	97.06%	98.18%	<b>98.3%</b>
$CAR(FAR = 0\%)$	68.25%	78.84%	<b>86.51%</b>
$FRR(FAR = 0\%)$	31.75%	21.16%	<b>13.49%</b>

### Extractor basado en la transformada Hermite orientada

Para el caso de la transformada Hermite orientada no fue posible obtener mejores resultados. Las tablas 4.7 y 4.8 muestran estos resultados obtenidos para tamaños de ventana  $N = 6$  y  $N = 8$  respectivamente.

En la utilización de la transformada Hermite orientada solamente se consideró el coeficiente  $l_2(\theta)$  para su codificación en un esquema semejante al descrito en la sección 3.2.2.3 (ecuación 3.38 en la sección 3.3.4). Debido a que los valores obtenidos con el coeficiente  $l_1(\theta)$  no permiten aplicar este mismo esquema de codificación, dicho coeficiente no fue considerado.

**Tabla 4.7:** Tabla comparativa de los resultados obtenidos al variar el tamaño del código o plantilla biométrica y el número de bandas o regiones de análisis. Para un ancho de banda angular  $N = 6$  usando la transformada Hermite orientada, el mejor resultado se observó con un código de 2048 bits y 8 bandas de análisis.

$N = 6$	1024 bits / 4 bandas	1024 bits / 8 bandas	2048 bits / 8 bandas
$n_1$	378	378	378
$\mu_1$	0.3782	0.3813	0.3783
$\sigma_1$	0.0526	0.0531	0.0533
$\mu_2$	0.4465	0.4490	0.4519
$n_2$	4182	4182	4182
$\sigma_2$	0.0189	0.0164	0.0160
$d'$	1.7284	1.7227	1.8709
$EER$	17.56%	18.30%	15.83%
Desempeño	82.44%	81.70%	84.17%
$CAR(FAR = 0\%)$	16.67%	24.86%	44.81%
$FRR(FAR = 0\%)$	83.33%	75.14%	55.19%

**Tabla 4.8:** Tabla comparativa de los resultados obtenidos al variar el tamaño del código o plantilla biométrica y el número de bandas o regiones de análisis. Para un ancho de banda angular  $N = 8$  usando la transformada Hermite orientada, el mejor resultado se observó con un código de 2048 bits y 8 bandas de análisis.

$N = 8$	1024 bits / 4 bandas	1024 bits / 8 bandas	2048 bits / 8 bandas
$n_1$	378	378	378
$\mu_1$	0.3706	0.3729	0.3706
$\sigma_1$	0.0554	0.0556	0.0551
$n_2$	4182	4182	4182
$\mu_2$	0.4441	0.4444	0.4471
$\sigma_2$	0.0203	0.0189	0.0186
$d'$	1.7607	1.7207	1.8602
$EER$	16.35%	17.72%	15.77%
Desempeño	83.65%	82.28%	84.23%
$CAR(FAR = 0\%)$	16.12%	24.86%	28.43%
$FRR(FAR = 0\%)$	83.88%	75.14%	71.57%

La razón por la cual se obtiene un desempeño tan bajo, probablemente se debe a la forma de codificar este coeficiente. Es necesario proponer una forma diferente de codificar en la que además se pueda incluir a los otros coeficientes obtenidos después de la orientación o bien incluir la información del ángulo bajo el cual se orientaron los coeficientes de la transformada.

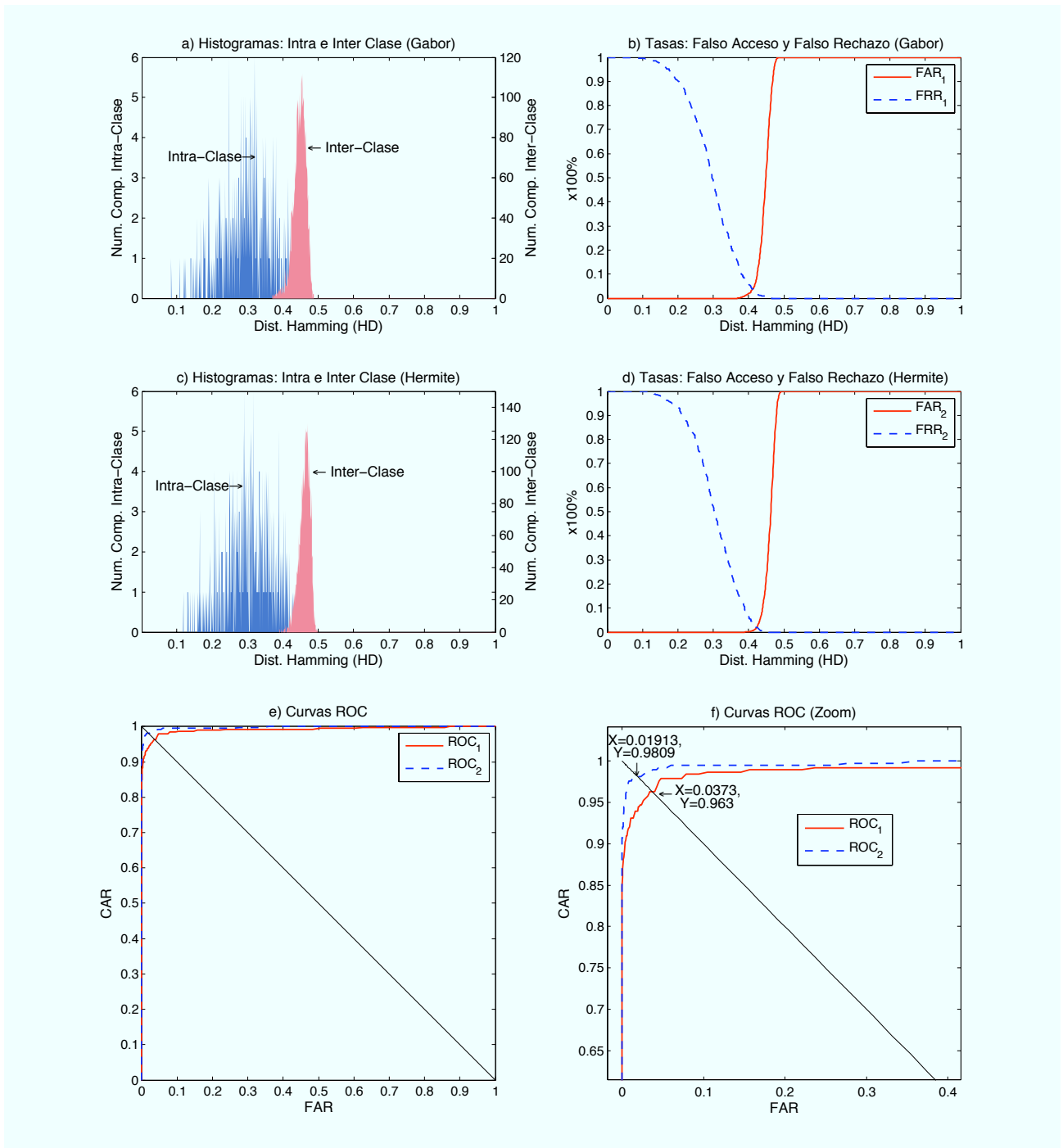
### Comparación de desempeño entre ambos extractores de características

A continuación evaluamos el desempeño obtenido por cada uno de los modelos utilizados (wavelet Gabor y transformada Hermite), con respecto a la base de datos de imágenes de iris de nuestra propiedad. Seleccionamos los parámetros del código y del filtro que otorgaron mejores desempeños en cada uno de los modelos.

Para ambos modelos se encontró que el mejor desempeño se obtenía dividiendo la región de interés en 8 bandas de análisis y formando códigos o plantillas biométricas de 2048 bits. Con respecto al modelo de la wavelet Gabor, los mejores resultados se obtuvieron con filtros cuyo ancho de banda fue de  $B_\theta = \pi/6$  (tabla 4.2). Para el modelo de la transformada Hermite, los mejores resultados se obtuvieron con filtros cuyo tamaño de ventana fue de  $N = 6$  (tabla 4.4).

A partir de las gráficas en la figura 4.4 y comparando las tablas 4.2 y 4.4, es posible observar que la distancia de separación entre clases (intra-clase e inter-clase) fue mayor usando la transformada Hermite ( $d' = 3.3043$ , tabla 4.4) con respecto a la wavelet Gabor ( $d' = 3.0352$ , tabla 4.2). Así mismo, el desempeño de accesos correctos de la transformada Hermite ( $CAR = 90.74\%$ ) fue mejor con respecto al de la wavelet Gabor ( $CAR = 84.66\%$ ) cuando se restringe a 0% la tasa de acceso incorrecto a impostores ( $FAR$ ). El desempeño general del sistema observado en la gráfica (f) de la figura 4.4, usando la wavelet Gabor, fue de aproximadamente 96.3%, mientras que el desempeño general del sistema usando la transformada Hermite fue de 98.09%. El área bajo la curva  $ROC_1$  fue de 0.99200, mientras que el área bajo la curva  $ROC_2$  fue de 0.995126, lo cual coincide con los resultados de mejor desempeño obtenidos para la transformada Hermite.

### 4.3 Resultados usando la base de datos de imágenes de iris de nuestra propiedad



**Figura 4.4:** Evaluación del desempeño entre ambos extractores de características para la base de datos de imágenes de iris de nuestra propiedad. Wavelet Gabor: (a) distribuciones intra-clase e inter-clase y (b) tasas de falso acceso ( $FAR$ ) y falso rechazo ( $FRR$ ). Transformada Hermite: (c) Distribuciones intra-clase e inter-clase y (d) tasas de falso acceso ( $FAR$ ) y falso rechazo ( $FRR$ ). (e) Comparación de curvas ROC:  $ROC_1$  usando wavelet Gabor,  $ROC_2$  usando la transformada Hermite. (f) Versión de (e) con acercamiento.

## 4.4. Resultados usando la base de datos de imágenes de iris CASIA

### Extractor basado en wavelets Gabor

Las pruebas realizadas sobre la base de datos de imágenes de iris CASIAv1 [14], usando la wavelet Gabor se muestran a continuación. Al igual que para la base de datos de imágenes de iris de nuestra propiedad, dichas pruebas consistieron en elegir y variar algunos parámetros del extractor de características y de la formación del código.

**Tabla 4.9:** Tabla comparativa de los resultados obtenidos al variar el tamaño del código o plantilla biométrica y el número de bandas o regiones de análisis. Para un ancho de banda angular  $B_\theta = \pi/4$  usando wavelets Gabor, el mejor resultado se observó con un código de 2048 bits y 8 bandas de análisis.

$B_\theta = \pi/4$	1024 bits / 4 bandas	1024 bits / 8 bandas	2048 bits / 8 bandas
$n_1$	547	547	547
$\mu_1$	0.3382	0.3541	<b>0.3449</b>
$\sigma_1$	0.0536	0.0554	<b>0.0524</b>
$n_2$	4278	4278	4278
$\mu_2$	0.4422	0.4511	<b>0.4503</b>
$\sigma_2$	0.0173	0.0147	<b>0.0146</b>
$d'$	2.6137	2.3957	<b>2.7382</b>
$EER$	6.75%	7.78%	<b>4.72%</b>
Desempeño	93.25%	92.22%	<b>95.28%</b>
$CAR(FAR = 0\%)$	60.23%	63.44%	<b>69.84%</b>
$FRR(FAR = 0\%)$	39.77%	36.56%	<b>30.16%</b>

**Tabla 4.10:** Tabla comparativa de los resultados obtenidos al variar el tamaño del código o plantilla biométrica y el número de bandas o regiones de análisis. Para un ancho de banda angular  $B_\theta = \pi/6$  usando wavelets Gabor, el mejor resultado se observó con un código de 2048 bits y 8 bandas de análisis.

$B_\theta = \pi/6$	1024 bits / 4 bandas	1024 bits / 8 bandas	2048 bits / 8 bandas
$n_1$	547	547	547
$\mu_1$	0.3319	0.3522	<b>0.3423</b>
$\sigma_1$	0.0542	0.0559	<b>0.0533</b>
$n_2$	4278	4278	4278
$\mu_2$	0.4414	0.4508	<b>0.4498</b>
$\sigma_2$	0.0176	0.0148	<b>0.0147</b>
$d'$	2.7163	2.4132	<b>2.7520</b>
$EER$	5.91%	7.41%	<b>4.6%</b>
Desempeño	94.09%	92.59%	<b>95.4%</b>
$CAR(FAR = 0\%)$	58.68%	64.35%	<b>69.65%</b>
$FRR(FAR = 0\%)$	41.32%	35.65%	<b>30.35%</b>

Se seleccionaron dos anchos de banda angular distintos  $B_\theta = \pi/4$  y  $B_\theta = \pi/6$ . Para cada  $B_\theta$  se generaron códigos de iris, con los cuales se obtuvieron dos distribuciones de la distancia de Hamming

al comparar códigos intra-clase (la distribución obtenida cuando se comparan los códigos de un mismo individuo, para cada individuo) y códigos inter-clase (la distribución obtenida cuando se comparan códigos de un individuo con respecto a los códigos de los demás individuos, para cada individuo). El número de comparaciones intra-clase e inter-clase fue de  $n_1 = 547$  y  $n_2 = 4278$  respectivamente.

En las tablas 4.9-4.10 se muestran las estadísticas observadas. Así mismo, se eligieron el número de bandas o regiones de análisis, de 8 a 4 bandas. El número de bandas de análisis impuso un límite con respecto al tamaño del código o plantilla biométrica del iris. Los tamaños de los códigos de iris utilizados fueron de 2048 bits y 1024 bits.

Para ambas tablas (4.9 y 4.10), es posible notar que los mejores resultados se obtienen usando códigos de 2048 bits y al dividir la imagen en 8 bandas de análisis. Para un ancho de banda angular  $B_\theta = \pi/4$  el desempeño general del sistema fue de 95.28% y la tasa de error igual ( $EER = FAR = FRR$ ) fue cercana al 4.72%, mientras que para un ancho de banda angular  $B_\theta = \pi/6$  el desempeño general del sistema fue de 95.4% y la tasa de error igual fue de  $EER = 4.6\%$ . Determinar que ancho de banda angular otorga mejores resultados es un poco más complicado, ya que para ambos anchos de banda angular los resultados son muy similares. Sin embargo, a pesar de que para un ancho de banda angular  $B_\theta = \pi/6$ , bajo un criterio de decisión estricto en el cual a ningún impostor se le permite obtener acceso del sistema ( $FAR = 0\%$ ), el desempeño de accesos correctos fue menor ( $CAR = 69.65\%$ ) que para un ancho de banda angular de  $B_\theta = \pi/4$  ( $CAR = 69.84\%$ ), se pueden seleccionar los parámetros encontrados para  $B_\theta = \pi/6$  como los mejores, pues tomando en cuenta que el desempeño general del sistema y que la distancia de separación entre las dos distribuciones ( $d'_{\pi/6} = 2.7520$ ,  $d'_{\pi/4} = 2.7382$ ) fueron mayores para éste que para  $B_\theta = \pi/4$ , entonces podría resultar más seguro elegir un ancho de banda angular igual a  $B_\theta = \pi/6$ .

### Extractor basado en la transformada Hermite

Para el caso en que el extractor de características fue la transformada Hermite también se realizaron pruebas similares a las anteriores. Las tablas (tablas 4.11 - 4.14) muestran estos resultados para la base de datos de imágenes de iris CASIAv1 [14].

**Tabla 4.11:** Tabla comparativa de los resultados obtenidos al variar el tamaño del código o plantilla biométrica y el número de bandas o regiones de análisis. Para un ancho de banda angular  $N = 4$  usando la transformada Hermite orientada, el mejor resultado se observó con un código de 2048 bits y 8 bandas de análisis.

$N = 4$	1024 bits / 4 bandas	1024 bits / 8 bandas	2048 bits / 8 bandas
$n_1$	547	547	547
$\mu_1$	0.3560	0.3716	<b>0.3643</b>
$\sigma_1$	0.0494	0.0471	<b>0.0454</b>
$n_2$	4278	4278	4278
$\mu_2$	0.4539	0.4542	<b>0.4618</b>
$\sigma_2$	0.0159	0.0140	<b>0.0135</b>
$d'$	2.6668	2.3804	<b>2.9070</b>
$EER$	6.19%	8.69%	<b>4.44%</b>
Desempeño	93.81%	91.31%	<b>95.56%</b>
$CAR(FAR = 0\%)$	70.75%	48.08%	<b>79.52%</b>
$FRR(FAR = 0\%)$	29.25%	51.92%	<b>20.48%</b>

#### 4 Análisis de resultados

Se eligieron distintas escalas de análisis de la transformada Hermite, variando  $N = 4, 6, 8$  y  $10$ . Para cada  $N$  se generaron códigos de iris, con los cuales dos distribuciones de la distancia de Hamming (comparaciones intra-clase y comparaciones inter-clase) arrojaron las estadísticas observadas en dichas tablas. El número de bandas o regiones de análisis también se varió de 8 a 4 bandas. El número de regiones impuso un límite con respecto al tamaño del código de iris. Estos códigos de iris fueron de 2048 bits y 1024 bits.

En todas las tablas, al igual que sucedió con la base de datos de imágenes de iris de nuestra propiedad, es común observar que la mejor separación entre clases, así como también el mejor desempeño ajustando  $FAR = 0\%$ , se logra usando un código de 2048 bits, formado a partir de 8 bandas de análisis. A partir de los mejores resultados obtenidos para cada tamaño de ventana, es posible determinar los parámetros del código, número de bandas de análisis y características del filtro, que permiten obtener una mayor distancia de separación entre las clases intra-clase e inter-clase, así como también un mejor desempeño.

Al comparar las tablas 4.11-4.14, se encontró que un código de 2048 bits, formado a partir de 8 bandas de análisis, con los filtros de análisis Hermite (real e impar) compartiendo un tamaño de ventana de  $N = 8$  (Véase tabla 4.13), y bajo un criterio de decisión estricto en el cual a ningún impostor se le permite ingresar al sistema, es decir  $FAR = 0\%$ , otorgó mejores resultados que para otros tamaños de ventana. El desempeño de accesos correctos de usuarios para  $N = 8$  fue de  $CAR = 79.52\%$  contra un  $20.48\%$  de usuarios incorrectamente rechazados ( $FRR$ ). La distancia de separación entre las clases intra-clase e inter-clase para este caso fue de  $d' = 3.0567$ . La tasa de error igual ( $EER = 3.99\%$ ) proporcionó un desempeño general del sistema del  $96.01\%$ .

**Tabla 4.12:** Tabla comparativa de los resultados obtenidos al variar el tamaño del código o plantilla biométrica y el número de bandas o regiones de análisis. Para un ancho de banda angular  $N = 6$  usando la transformada Hermite orientada, el mejor resultado se observó con un código de 2048 bits y 8 bandas de análisis.

$N = 6$	1024 bits / 4 bandas	1024 bits / 8 bandas	2048 bits / 8 bandas
$n_1$	547	547	547
$\mu_1$	0.3370	0.3590	<b>0.3499</b>
$\sigma_1$	0.0522	0.0494	<b>0.0487</b>
$n_2$	4278	4278	4278
$\mu_2$	0.4501	0.4526	<b>0.4569</b>
$\sigma_2$	0.0175	0.0146	<b>0.0155</b>
$d'$	2.9045	2.5704	<b>2.9655</b>
$EER$	5.82%	6.35%	<b>4.62%</b>
Desempeño	94.18%	93.65%	<b>95.38%</b>
$CAR(FAR = 0\%)$	65.27%	60.69%	<b>66.91%</b>
$FRR(FAR = 0\%)$	34.73%	39.31%	<b>33.09%</b>

**Tabla 4.13:** Tabla comparativa de los resultados obtenidos al variar el tamaño del código o plantilla biométrica y el número de bandas o regiones de análisis. Para un ancho de banda angular  $N = 8$  usando la transformada Hermite orientada, el mejor resultado se observó con un código de 2048 bits y 8 bandas de análisis.

$N = 8$	1024 bits / 4 bandas	1024 bits / 8 bandas	2048 bits / 8 bandas
$n_1$	547	547	547
$\mu_1$	0.3235	0.3470	<b>0.3396</b>
$\sigma_1$	0.0545	0.0502	<b>0.0494</b>
$n_2$	4278	4278	4278
$\mu_2$	0.4452	0.4489	<b>0.4523</b>
$\sigma_2$	0.0196	0.0164	<b>0.0168</b>
$d'$	2.9709	2.7310	<b>3.0567</b>
$EER$	5.26%	6.28%	<b>3.99%</b>
Desempeño	94.74%	93.72%	<b>96.01%</b>
$CAR(FAR = 0\%)$	73.67%	56.67%	<b>79.52%</b>
$FRR(FAR = 0\%)$	26.33%	43.33%	<b>20.48%</b>

**Tabla 4.14:** Tabla comparativa de los resultados obtenidos al variar el tamaño del código o plantilla biométrica y el número de bandas o regiones de análisis. Para un ancho de banda angular  $N = 10$  usando la transformada Hermite orientada, el mejor resultado se observó con un código de 2048 bits y 8 bandas de análisis.

$N = 10$	1024 bits / 4 bandas	1024 bits / 8 bandas	2048 bits / 8 bandas
$n_1$	547	547	547
$\mu_1$	0.3156	0.3380	<b>0.3338</b>
$\sigma_1$	0.0552	0.0505	<b>0.0492</b>
$n_2$	4278	4278	4278
$\mu_2$	0.4411	0.4435	<b>0.4478</b>
$\sigma_2$	0.0207	0.0178	<b>0.0179</b>
$d'$	3.0089	2.7841	<b>3.0771</b>
$EER$	5.54%	6.21%	<b>5.4%</b>
Desempeño	94.46%	93.79%	<b>94.6%</b>
$CAR(FAR = 0\%)$	68.01%	52.47%	<b>69.65%</b>
$FRR(FAR = 0\%)$	31.99%	47.53%	<b>30.35%</b>



### Extractor basado en la transformada Hermite orientada

Para este caso, al igual que sucedió al utilizar la base de datos de nuestra propiedad, usando la transformada Hermite orientada no fue posible obtener mejores resultados. Únicamente se reporta una tabla, obtenida para una ventana de tamaño  $N = 6$ , ya que los resultados para ventanas de tamaño diferente fueron muy similares

**Tabla 4.15:** Tabla comparativa de los resultados obtenidos al variar el tamaño del código o plantilla biométrica y el número de bandas o regiones de análisis. Para un ancho de banda angular  $N = 6$  usando la transformada Hermite orientada, el mejor resultado se observó con un código de 2048 bits y 8 bandas de análisis.

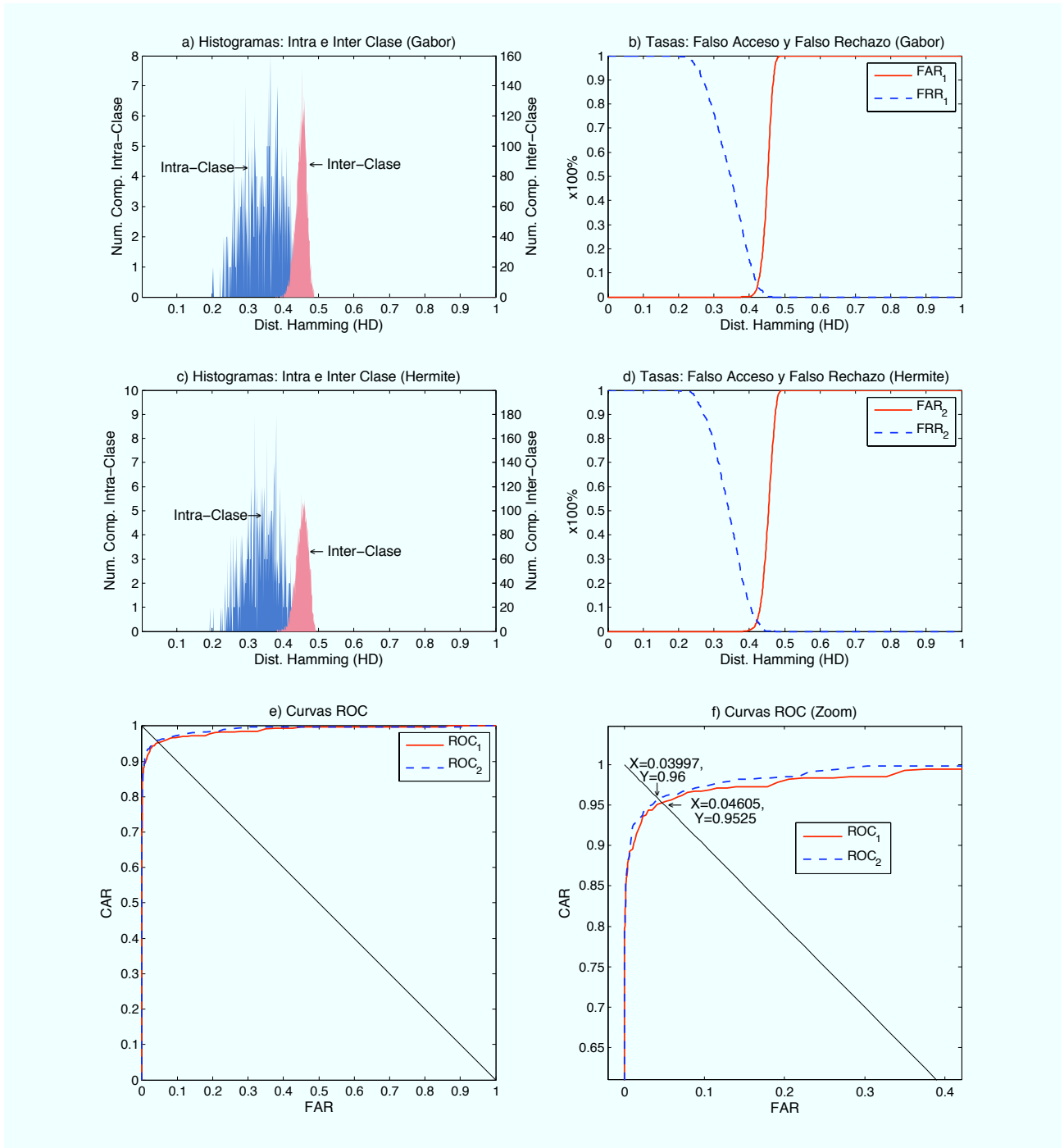
$N = 6$	1024 bits / 4 bandas	1024 bits / 8 bandas	2048 bits / 8 bandas
$n_1$	547	547	547
$\mu_1$	0.4109	0.4181	0.4149
$\sigma_1$	0.0346	0.0319	0.0337
$n_2$	4278	4278	4278
$\mu_2$	0.4480	0.4508	0.4547
$\sigma_2$	0.0164	0.0139	0.0133
$d'$	1.3711	1.3316	1.5559
$EEER$	24.24%	24.47%	20.52%
Desempeño	75.76%	75.53%	79.48%
$CAR(FAR = 0\%)$	17.00%	15.00%	21.21%
$FRR(FAR = 0\%)$	83.00%	85.00%	78.79%

### Comparación de desempeño entre ambos extractores de características

A continuación evaluamos el desempeño obtenido por cada uno de los modelos utilizados (wavelet Gabor y transformada Hermite), con respecto a la base de datos de imágenes de iris CASIAv1 [14]. Para cada uno de los modelos seleccionamos los parámetros del código y del filtro que otorgaron los mejores desempeños.

Para ambos modelos se encontró que el mejor desempeño se obtenía dividiendo la región de interés en 8 bandas de análisis y formando códigos o plantillas biométricas de 2048 bits. Con respecto al modelo de la wavelet Gabor, los mejores resultados se obtuvieron con filtros cuyo ancho de banda fue de  $B_\theta = \pi/6$  (tabla 4.10). Para el modelo de la transformada Hermite, los mejores resultados se obtuvieron con filtros cuyo tamaño de ventana fue de  $N = 8$  (tabla 4.13).

A partir de las gráficas en la figura 4.5 y comparando las tablas 4.10 y 4.13, es posible observar que la distancia de separación entre clases (intra-clase e inter-clase) fue mayor usando la transformada Hermite ( $d' = 3.0567$ , tabla 4.13) con respecto a la wavelet Gabor ( $d' = 2.7520$ , tabla 4.10). Así mismo, el desempeño de accesos correctos de la transformada Hermite ( $CAR = 79.52\%$ ) fue mejor con respecto al de la wavelet Gabor ( $CAR = 69.65\%$ ) cuando se restringe a 0% la tasa de acceso incorrecto a impostores ( $FAR$ ). El desempeño general del sistema observado en la gráfica (f) de la figura 4.5, usando la wavelet Gabor, fue de aproximadamente 95.4%, mientras que el desempeño general del sistema usando la transformada Hermite fue de 96.01%. El área bajo la curva  $ROC_1$  fue de 0.985809, mientras que el área bajo la curva  $ROC_2$  fue de 0.990085, lo cual coincide con los resultados de mejor desempeño obtenidos para la transformada Hermite.



**Figura 4.5:** Evaluación del desempeño entre ambos extractores de características para la base de datos de imágenes de iris CASIA. Wavelet Gabor: (a) distribuciones intra-clase e inter-clase y (b) tasas de falso acceso ( $FAR$ ) y falso rechazo ( $FRR$ ). Transformada Hermite: (c) Distribuciones intra-clase e inter-clase y (d) tasas de falso acceso ( $FAR$ ) y falso rechazo ( $FRR$ ). (e) Comparación de curvas ROC:  $ROC_1$  usando wavelet Gabor,  $ROC_2$  usando la transformada Hermite. (f) Versión de (e) con acercamiento.

## 4.5. Discusión de resultados

Para el caso en que se utilizó nuestra propia base de datos de imágenes de iris, se logró observar un mejor desempeño del sistema de reconocimiento de iris usando la transformada Hermite con respecto a los resultados obtenidos con la wavelet Gabor. De igual manera, aunque con menor diferencia de desempeño entre ambos extractores de características, se pudo observar un mejor desempeño usando la transformada Hermite para el caso en que la base de datos de imágenes de iris es la proveniente de CASIA [14].

A partir de los resultados obtenidos usando la transformada Hermite sobre nuestra propia base de datos de imágenes de iris, se obtuvo una tasa de accesos correctos  $CAR = 90.74\%$  y una tasa de error de falsos rechazos  $FRR = 9.26\%$  cuando se restringe la tasa de error de falso acceso a impostores a  $FAR = 0\%$ . La tasa de error igual obtenida fue  $EEER = 1.91\%$ . La distancia de separación entre las clases intra-clase e inter-clase obtenida en este caso fue de  $d' = 3.3043$ . Cuando se utilizó la base de datos de imágenes de iris CASIA, se obtuvo una tasa de accesos correctos  $CAR = 79.52\%$  de usuarios correctamente aceptados contra una tasa de usuarios incorrectamente rechazados de  $FRR = 20.48\%$  bajo un criterio de decisión estricto en el cual se restringe la tasa de acceso a impostores  $FAR = 0\%$ . La tasa de error igual fue  $EEER = 3.99\%$  y la distancia de separación entre clases intra-clase e inter-clase fue de  $d' = 3.0567$ .

Uno de los factores por el cual las tasas de error de falsos rechazos  $FRR$  fueron diferentes entre las bases de datos usadas, y que para el caso de la base de datos de imágenes de iris CASIA fue mucho mayor que el obtenido para la base de datos imágenes de iris de nuestra propiedad, se debe a la calidad de las imágenes de iris. Mientras que para nuestra base de datos de imágenes de iris, éstas fueron adquiridas con alta resolución, la base de datos de imágenes de iris CASIA fueron capturadas con baja resolución.

Otro factor que desafortunadamente ayuda a aumentar las tasas de error de falsos rechazos  $FRR$  se debe a una localización del iris y eliminación incorrectas de párpados y pestañas dentro de la imagen. Las imágenes de iris provenientes de la base de datos CASIA están altamente obstruidas por párpados y pestañas, lo que dificulta la tarea de localización del iris. Cuando una imagen está demasiado obstruida por estos artefactos, la localización del iris resulta incorrecta e induce el incremento de la tasa de error de falso rechazo ( $FRR$ ). En algunos sistemas de reconocimiento de iris, después de la etapa de localización se seleccionan (comúnmente en forma manual) únicamente aquellas imágenes de iris que fueron correctamente localizadas, permitiendo que la tasa de  $FRR$  disminuya considerablemente. Sin embargo, este no fue el caso para el sistema propuesto, aunque el algoritmo es capaz de detectar y descartar algunas imágenes que localizan incorrectamente la región del iris, no se inspeccionan en forma manual para comprobarlo.

Podemos comparar los resultados obtenidos usando ambas bases de datos de imágenes de iris con algunos de los métodos de extracción de características reportados en la literatura. En la tabla 4.16 se condensan algunos resultados obtenidos por otros autores.

Usando filtros log-Gabor, Zhang et al., [86] reporta para una base de datos de imágenes de iris de su propiedad, en el mejor de los casos colocando la tasa de error de falso acceso en  $FAR = 0\%$ , una tasa de error de falso rechazo de  $FRR = 10\%$ .

Shinyoung-Lim et al., [48] obtienen para una base de datos de imágenes de iris de su propiedad, con selección previa de aquellas imágenes cuyo iris fue localizado correctamente, para una tasa de error de falso acceso  $FAR = 0\%$  una tasa de error de falso rechazo  $FRR = 17.75\%$  usando una descomposición con la wavelet Haar.

Por otro lado, Chen et al., [13] reporta para la base de datos de imágenes de iris CASIA, usando la modulación delta de la dirección del gradiente, en el mejor de los casos una tasa de error igual

**Tabla 4.16:** Comparación de los resultados obtenidos con los reportados por otros autores.

Autor	Extractor	$FAR$ (%)	$FRR$ (%)	$EER$ (%)	$d'$	Base de datos
Propuesto	Hermite	0.0	9.26	1.91	3.3043	Propietario
Propuesto	Hermite	0.0	20.48	3.99	3.0567	CASIA
Zhang et al., [86]	log-Gabor	0.0	10.0	–	–	Propietario
Shinyoung-Lim et al., [48]	Haar	0.0	17.75	2.9	–	Propietario
Chen et al., [13]	Modulación delta de la dirección del gradiente	0.0	8.13	3.79	–	CASIA
Poursaberi et al., [62]	Daubechies 2	0.001	2.98	0.2687	–	CASIA
Sanchez-Avila et al., [67]	Gabor	0.0	4.0	3.3	–	Propietario
Daugman [22]	Gabor	0.001	0.12	0.11	–	NIST (ICE-1)

cercana a  $EER = 3.79\%$  y una tasa de falso rechazo de  $FRR = 8.13\%$  para una tasa de error de falso acceso  $FAR = 0\%$ . En este caso, durante las pruebas experimentales seleccionan de manera subjetiva la precisión de las fronteras del iris en la etapa de localización.

En Poursaberi et al., [62] usando la wavelet Daubechies 2, reporta una tasa de error igual cercana al  $EER = 0.2687\%$  y una tasa de falso rechazo de  $FRR = 2.98\%$  para una tasa de error de falso acceso  $FAR = 0.001\%$ , usando únicamente la mitad de la la región del iris con imágenes provenientes de la base de datos CASIA.

Usando una base de datos de imágenes de iris de su propiedad, Sanchez-Avila et al., [67] obtuvieron una tasa de error de falso rechazo cercana al  $FRR = 4.0\%$  cuando la tasa de error de falso acceso fue fijada en  $FAR = 0\%$ . El sistema en general otorgó una tasa de error igual de  $EER = 3.3\%$ .

El algoritmo presentado por Daugman [19] es de los pocos algoritmos que posee bajas tasas de error ( $FRR$  y  $FAR$ ) y un desempeño alto en términos de alta seguridad. Recientemente, Daugman [22] ha reportado para la base de datos de imágenes de iris (ICE-1) perteneciente al NIST <sup>2</sup>, una tasa de error igual cercana a  $EER = 0.11\%$  y una tasa de falso rechazo de  $FRR = 0.12\%$  para una tasa de error de falso acceso  $FAR = 0.001\%$ . En otros artículos del mismo autor [20][21], también ha reportado una distancia de separación entre las distribuciones intra-clase e inter-clase de  $d' = 7.3$ , pero no menciona nada respecto a las tasas de error.

Comparando nuestros resultados con los obtenidos por Daugman [22], quien ha utilizado bases de datos de imágenes de iris grandes (alrededor de 600,000 códigos de iris diferentes [22]), y aunque para el caso en que se utilizó la wavelet Gabor se obtenga un menor desempeño con nuestra interpretación de los algoritmos reportados por Daugman, no es posible decir aún que la transformada Hermite es

<sup>2</sup>NIST, por sus siglas en inglés “National Institute of Standards and Technology”, agencia de la Administración de Tecnología del Departamento de Comercio de los Estados Unidos de América.

#### 4 Análisis de resultados

mejor que la wavelet Gabor en esta tarea en particular, así como tampoco es posible decir lo contrario, pues existen etapas a lo largo de todo el proceso de extracción de características (como la localización del iris), cuyos errores influyen finalmente en las tasas de falso rechazo ( $FRR$ ), degradando a la vez las tasas de error igual  $EER$ . Sin embargo, si podemos decir que nuestros resultados son competitivos con los resultados de otros autores, quienes reportan en la literatura el uso de diversos extractores de características. Esto se debe a que la forma de generación de los códigos de iris no es clara todavía, pues este proceso se encuentra protegido aún por una patente propiedad de Iridian Technologies, Inc [42]. Por esta razón, es necesario realizar nuevas propuestas de generación de códigos característicos que permitan obtener una mayor separación entre las texturas de iris que pertenecen a una misma clase y una baja variabilidad para aquellas que pertenecen a clases distintas.

## 5 Conclusiones y perspectivas

La viabilidad del uso de la transformada Hermite, como una herramienta alternativa a las wavelets Gabor para la extracción de características de textura, fue demostrada por los resultados obtenidos en este proyecto de investigación. El tipo de texturas que fueron analizadas corresponden a texturas de imágenes de iris. Para ello, se desarrolló satisfactoriamente un sistema de reconocimiento de iris. La capacidad de extracción de patrones significativos de texturas de iris a través de la transformada Hermite, se ve reflejada en los niveles de confianza que se lograron obtener.

Se desarrollaron todas las etapas comúnmente encontradas en los sistemas de reconocimiento de iris. Se observó también que todas las etapas que conforman un sistema de reconocimiento de iris son sumamente importantes, ya que la capacidad de discriminación inter-clase y la disminución de variabilidad intra-clase, se puede ver afectada por cualquiera de ellas que no realice adecuadamente su tarea. En el caso del sistema desarrollado en este trabajo, cada etapa o módulo puede ser mejorada por separado sin afectar a las demás, creando oportunidades de futuras implementaciones que mejoren el desempeño del sistema en cada una de dichas etapas.

Con respecto a la etapa de adquisición de imágenes de iris, a pesar de que ésta no funciona en tiempo real, para propósitos de investigación las imágenes capturadas son aceptables y forman parte de una base de datos de imágenes de iris, la cual deberá incrementar el número de usuarios registrados gradualmente. La automatización de esta etapa para su funcionamiento en tiempo real, requiere de lentes controlados por hardware y software para lograr enfocar de forma adecuada las imágenes de iris (*autofocus*).

Se utilizaron dos bases de datos de imágenes de iris, una obtenida a través del módulo de adquisición y otra proveniente de un grupo de investigación en China, llamada CASIA. Al comparar los resultados obtenidos con ambas bases de datos de imágenes de iris, se pudo observar que la variabilidad intra-clase se ve afectada por la resolución de las imágenes usada. Las imágenes capturadas con una resolución alta permitieron un mejor desempeño del sistema de reconocimiento de iris.

El sistema de reconocimiento de iris demostró un desempeño aceptable y resultados competitivos comparado con reportes encontrados en la literatura, a pesar de no obtener los resultados que han sido reportados por Daugman [21], quien es uno de los pioneros de los mejores sistemas de reconocimiento de iris que existen. Cabe señalar que el proceso de generación de los códigos de iris usados por Daugman no es clara aún ya que se encuentra protegido por una patente, por lo que esta desventaja hace difícil comparar únicamente los dos modelos de representación y extracción de características. Es por eso que surge la necesidad de realizar nuevas propuestas de generación de códigos característicos que permitan obtener una mayor separación entre las texturas de iris que pertenecen a una misma clase y aquellas que pertenecen a distintas clases.

Con estos resultados, también se confirmó la capacidad que puede ofrecer un órgano interno y bien protegido, como lo es el iris, para su utilización como un identificador biométrico robusto, en situaciones que demanden alta confiabilidad como el control de acceso a áreas restringidas o identificación de criminales por citar algunos ejemplos.

En trabajos futuros se prevé mejorar la etapa de localización del iris para lograr obtener una mayor precisión en la definición de las fronteras que delimitan al iris del resto de la zona del ojo, así como también de la rapidez con la cual se lleva a cabo esta localización. Para el caso de la etapa de

## *5 Conclusiones y perspectivas*

codificación, es necesario realizar una mayor investigación y proponer una nueva forma de agrupar los atributos característicos de la textura del iris, cumpliendo con que sea invariante a la rotación, de forma tal que sea posible mejorar la separación o discriminación entre las distribuciones intra-clase e inter-clase.

# A Polinomios de Hermite

La ecuación de Hermite tiene la forma siguiente:

$$y'' - 2xy' + 2ny = 0 \quad (\text{A.1})$$

donde  $n$  puede ser cualquier número real aunque casi siempre toma valores enteros en aplicaciones físicas. La ecuación de Hermite aparece en la descripción de la ecuación de onda del oscilador armónico cuántico. Las soluciones a la ecuación anterior son los llamados polinomios de Hermite. Los polinomios de Hermite y las funciones derivadas de ellos son importantes en el análisis del comportamiento de algunos sistemas físicos en la mecánica cuántica.

## Fórmula de Rodrigues

La fórmula de Rodrigues para los polinomios de Hermite asociados a la solución de la ecuación A.1 está dada por:

$$H_n(x) = (-1)^n e^{x^2} \frac{d^n e^{-x^2}}{dx^n} \quad (\text{A.2})$$

La demostración de que A.2 es solución a A.1 se da a continuación.

**Prueba:** sea  $u = e^{-x^2}$  y sea  $u^n = \frac{d^n e^{-x^2}}{dx^n}$  la derivada  $n$ -ésima de  $u$  con respecto a  $x$ . Obteniendo la primera derivada de  $u$  con respecto a  $x$  tenemos:

$$\begin{aligned} u' &= -2xe^{-x^2} \\ \Rightarrow u' + 2xe^{-x^2} &= 0 \end{aligned}$$

Si diferenciamos esta ecuación  $n + 1$  veces usando el teorema de Leibnitz entonces resulta:

$$u^{n+2} + 2xu^{n+1} + 2(n+1)u^n = 0$$

Realizando un cambio de variable con  $v = (-1)^n u^n$  donde  $u^n = \frac{d^n e^{-x^2}}{dx^n}$  obtenemos:

$$v'' + 2xv' + 2(n+1)v = 0 \quad (\text{A.3})$$

Ahora bien, sea  $y = e^{x^2}v$ . Podemos escribir a las derivadas de  $v$  con respecto de  $x$  como:

$$\begin{aligned} v &= e^{-x^2}y, \\ v' &= e^{-x^2}(y' - 2xy), \\ v'' &= e^{-x^2}(y'' - 4xy' + 4x^2y - 2y). \end{aligned}$$

Sustituyendo estas últimas expresiones en la ecuación A.3 y dividiendo por  $e^{-x^2}$  nos queda:

$$y'' - 2xy' + 2ny = 0 \quad \blacksquare$$

lo que indica que  $y = (-1)^n e^{x^2} \frac{d^n e^{-x^2}}{dx^n}$ .



## A Polinomios de Hermite

Los polinomios de Hermite tienen la siguiente propiedad:

$$H_n(-x) = (-1)^n H_n(x)$$

lo que implica que el polinomio  $H_n$  es par si  $n$  es par y viceversa.

La forma explícita de los primeros ordenes de los polinomios de Hermite es:

$$\begin{aligned}H_0(x) &= 1 \\H_1(x) &= 2x \\H_2(x) &= 4x^2 - 2 = 2^2 x^2 - 2 \\H_3(x) &= 8x^3 - 12x = 2^3 x^3 - 12x \\H_4(x) &= 16x^4 - 48x^2 + 12 = 2^4 x^4 - 48x^2 + 12 \\H_5(x) &= 32x^5 - 160x^3 + 120x = 2^5 x^5 - 160x^3 + 120x\end{aligned}$$

calculando las derivadas  $n = 0, 1, 2, \dots, 5$  para cada uno de los polinomios anteriores, podemos ver que los términos, cuyo orden es menor al orden  $n$ , no contribuyen en el resultado final del cálculo de dicha derivada:

$$\begin{aligned}D^0 H_0(x) &= D^0 1 = 1 \\D^1 H_1(x) &= D^1(2x) = 2 \\D^2 H_2(x) &= D^2(2^2 x^2 - 2) = 2^2 D^1(2x) = 2^2 \cdot 2 \\D^3 H_3(x) &= D^3(2^3 x^3 - 12x) = 2^3 \cdot 3 \cdot D^2(x^2) = 2^3 \cdot 3 \cdot 2 \cdot D^1(x) = 2^3 \cdot 6 \\D^4 H_4(x) &= D^4(2^4 x^4 - 48x^2 + 12) = 2^4 \cdot 4 \cdot D^3(x^3) = 2^4 \cdot 4 \cdot 3 \cdot D^2(x^2) = 2^4 \cdot 24 \\D^5 H_5(x) &= D^5(2^5 x^5 - 160x^3 + 120x) = 2^5 \cdot 5 \cdot D^4(x^4) = 2^5 \cdot 5 \cdot 4 \cdot D^3(x^3) = 2^5 \cdot 120\end{aligned}$$

donde es posible observar que

$$D^n H_n(x) = 2^n n! \tag{A.4}$$

## Ortogonalidad

Los polinomios de Hermite están definidos como polinomios ortogonales con respecto a un producto de peso escalar  $e^{-x^2}$  en el intervalo  $-\infty < x < \infty$ :

$$\int_{-\infty}^{\infty} H_m(x) H_n(x) e^{-x^2} dx = \begin{cases} 0, & m \neq n \\ 2^n n! \sqrt{\pi}, & m = n \end{cases} \tag{A.5}$$

Podemos demostrar el resultado anterior de la siguiente manera. Sean  $H_n$  y  $H_m$  dos polinomios de Hermite de grado  $n$  y  $m$  respectivamente.

**Prueba:** empezamos con la ecuación de Hermite, renombrando a  $y$  por  $H_n$  ya que son equivalentes como se demostró en el desarrollo de la fórmula de Rodrigues:

$$H_n''(x) - 2xH_n'(x) + 2nH_n(x) = 0$$

introducimos el factor  $e^{-x^2}$ :

$$H_n''(x)e^{-x^2} - 2xe^{-x^2}H_n'(x) = -2ne^{-x^2}H_n(x)$$

cuyo lado izquierdo, siguiendo la regla de la cadena de la derivada de un producto de dos funciones, podemos escribir como:

$$\begin{aligned}\frac{d}{dx} \left( e^{-x^2} H'_n(x) \right) &= -2ne^{-x^2} H_n(x) \\ \frac{d}{dx} \left( e^{-x^2} \frac{d}{dx} H_n(x) \right) &= -2ne^{-x^2} H_n(x)\end{aligned}\quad (\text{A.6})$$

multiplicamos esta ecuación por  $H_m$ :

$$H_m(x) \frac{d}{dx} \left( e^{-x^2} H'_n(x) \right) = -H_m(x) 2ne^{-x^2} H_n(x) \quad (\text{A.7})$$

De la misma manera que fue generada esta ecuación para un polinomio de grado  $n$ , generamos otra a partir de la ecuación A.6 para un polinomio de grado  $m$  y multiplicamos a ambos lados por  $H_n$ :

$$H_n(x) \frac{d}{dx} \left( e^{-x^2} H'_m(x) \right) = -H_n(x) 2me^{-x^2} H_m(x) \quad (\text{A.8})$$

restando estas últimas dos ecuaciones, A.7 y A.8 nos queda:

$$H_m(x) \frac{d}{dx} \left( e^{-x^2} H'_n(x) \right) - H_n(x) \frac{d}{dx} \left( e^{-x^2} H'_m(x) \right) = -H_m(x) 2ne^{-x^2} H_n(x) + H_n(x) 2me^{-x^2} H_m(x)$$

integrando sobre  $x$  de  $-\infty$  a  $\infty$ ; los términos de la izquierda se integran por partes ( $\int u dv = uv - \int v du$ ). Para el primer término de la izquierda tenemos:

$$\begin{aligned}u_m &= H_m(x); & du_m &= \frac{d}{dx} H_m(x) \\ dv_m &= \frac{d}{dx} \left( e^{-x^2} H'_n(x) \right); & v_m &= e^{-x^2} \frac{d}{dx} H_n(x)\end{aligned}$$

$$\int_{-\infty}^{\infty} H_m(x) \frac{d}{dx} \left( e^{-x^2} H'_n(x) \right) dx = H_m(x) e^{-x^2} \frac{d}{dx} H_n(x) \Big|_{-\infty}^{\infty} - \int_{-\infty}^{\infty} \frac{d}{dx} H_m(x) e^{-x^2} \frac{d}{dx} H_n(x) dx$$

y para el segundo término tenemos:

$$- \int_{-\infty}^{\infty} H_n(x) \frac{d}{dx} \left( e^{-x^2} H'_m(x) \right) dx = -H_n(x) e^{-x^2} \frac{d}{dx} H_m(x) \Big|_{-\infty}^{\infty} + \int_{-\infty}^{\infty} \frac{d}{dx} H_n(x) e^{-x^2} \frac{d}{dx} H_m(x) dx$$

Los términos de las dos integrales son simétricos y se cancelan al sumar. Los términos restantes son simétricos entonces desaparecen a ambos límites si  $m \neq n$ . La integral de la izquierda resulta cero. Integrando la parte derecha

$$\begin{aligned}0 &= 2(m-n) \int_{-\infty}^{\infty} H_m(x) H_n(x) e^{-x^2} dx \\ &\Rightarrow \int_{-\infty}^{\infty} H_m(x) H_n(x) e^{-x^2} dx = 0 \quad \blacksquare\end{aligned}$$

es la constante obtenida al integrar la parte izquierda.

Ahora deseamos probar el caso en que  $n = m$ :

$$\int_{-\infty}^{\infty} H_n(x) H_n(x) e^{-x^2} dx = 2^n n! \sqrt{\pi}$$

**Prueba:** Sustituyendo la definición de la fórmula de Rodrigues (ecuación A.2) tenemos:

$$I_{nn} = \int_{-\infty}^{\infty} H_n(x)H_n(x)e^{-x^2} dx = (-1)^n \int_{-\infty}^{\infty} H_n(x) \frac{d^n e^{-x^2}}{dx^n} dx$$

integrando por partes ( $\int u dv = uv - \int v du$ )  $n$  veces con  $u = H_n(x)$ ,  $d^n u = d^n H_n(x)$ ,  $d^n v = d^n e^{-x^2} dx$ ,  $v = e^{-x^2}$  y tomando en cuenta que los términos frontera desaparecen tenemos:

$$I_{nn} = \int_{-\infty}^{\infty} \frac{d^n H_n(x)}{dx^n} e^{-x^2} dx$$

sustituyendo la ecuación A.4 en la ecuación anterior nos queda:

$$I_{nn} = 2^n n! \int_{-\infty}^{\infty} e^{-x^2} dx = 2^n n! \sqrt{\pi} \quad \blacksquare$$

Usando el factor de normalización  $I_{nn} = 2^n n! \sqrt{\pi}$  podemos reescribir la ecuación A.5 para cualquier  $n$  y  $m$  como:

$$\int_{-\infty}^{\infty} H_n(x)H_m(x)e^{-x^2} dx = \delta_{nm} 2^n n! \sqrt{\pi}$$

donde  $\delta_{nm}$  es la delta de Kronecker ( $\delta_{nm} = 0$  si  $n \neq m$  y  $\delta_{nm} = 1$  si  $n = m$ ). Reescribiendo la ecuación anterior tenemos:

$$\int_{-\infty}^{\infty} \frac{e^{-x^2}}{\sqrt{\pi}} \frac{H_n(x)}{\sqrt{2^n n!}} \cdot \frac{H_m(x)}{\sqrt{2^m m!}} dx = \delta_{nm}$$

evaluando la ecuación anterior en  $x = \frac{u}{\sigma}$ :

$$\int_{-\infty}^{\infty} \frac{1}{\sigma \sqrt{\pi}} e^{-\frac{u^2}{\sigma^2}} \frac{H_n\left(\frac{u}{\sigma}\right)}{\sqrt{2^n n!}} \cdot \frac{H_m\left(\frac{u}{\sigma}\right)}{\sqrt{2^m m!}} du = \delta_{nm}$$

donde  $dx = \frac{1}{\sigma} du$ .

## B Normalización $L^1$ y $L^2$ de la función Gaussiana

Una Gaussiana tiene la siguiente forma general [57]:

$$G_0 = N(a)e^{-ax^2} \quad (\text{B.1})$$

$N$  es una constante de normalización que depende de  $a$  pero no de  $x$ . Hay muchas posibilidades para el coeficiente  $a$  en el exponente: (1) La forma estándar es  $a = (2\sigma^2)^{-1}$  donde  $\sigma$  es la desviación estándar. (2) En la transformada Hermite la ventana cuadrada Gaussiana al igual que los polinomios de Hermite  $a = (\sigma^2)^{-1}$ . (3) En la teoría espacio-escala el parámetro natural de escala  $s$  está dado por  $a = (4s)^{-1}$ .

La constante de normalización puede ser tal que  $L^1(G_0) = 1(N_1)$  o bien  $L^2(G_0) = 1(N_2)$ .  $L^2$  es interesante especialmente si se necesita realizar reconstrucción como es el caso de la transformada Hermite.  $L^1$  también es considerada ya que a veces es preferible cuando se desea la interpretación de las respuestas.

La integral de una Gaussiana [57] como la definida en la ecuación B.1 es:

$$\int_0^\infty e^{-ax^2} dx = \frac{\sqrt{\pi}}{2\sqrt{a}} \quad (\text{B.2})$$

por lo que podemos calcular  $L^2$ :

$$\begin{aligned} (L^2(G_0))^2 &= 2 \int_0^\infty (N_2)^2 e^{-2ax^2} dx \\ &= (N_2)^2 \frac{\sqrt{\pi}}{\sqrt{2a}} \\ \Rightarrow N_2 &= (2a)^{1/4} (\pi)^{-1/4} \end{aligned} \quad (\text{B.3})$$

y de la misma manera  $L^1$ :

$$\begin{aligned} L^1(G_0) &= 2 \int_0^\infty N_1 e^{-ax^2} dx \\ &= N_1 \frac{\sqrt{\pi}}{\sqrt{a}} \\ \Rightarrow N_1 &= \frac{\sqrt{a}}{\sqrt{\pi}} \end{aligned} \quad (\text{B.4})$$

La función ventana usada en la obtención de las funciones filtro en la transformada Hermite está normalizada usando  $L^2$  y el coeficiente es  $a = (2\sigma^2)^{-1}$ . Entonces sustituimos en B.3  $a = (2\sigma^2)^{-1}$ :

$$N_2 = \sigma^{-1/2} \pi^{-1/4} = \frac{1}{\sqrt{\sqrt{\pi}\sigma}}$$

*B Normalización  $L^1$  y  $L^2$  de la función Gaussiana*

Por lo que podemos escribir dicha función ventana sustituyendo  $N = N_2$  y  $a = (2\sigma^2)^{-1}$  en la ecuación general B.1:

$$\begin{aligned} G_0 &= N_2 e^{-\frac{x^2}{2\sigma^2}} \\ &= \frac{1}{\sqrt{\sqrt{\pi}\sigma}} e^{-\frac{x^2}{2\sigma^2}} \end{aligned} \tag{B.5}$$

# C Obtención de las derivadas de Gaussianas como filtros de análisis de la transformada Hermite

La función Gaussiana  $G_0$  definida en la ecuación B.5 bajo la norma  $L^2$  es:

$$G_0(x)_{L^2} = \frac{1}{\sqrt{\sqrt{\pi}\sigma}} e^{-\frac{x^2}{2\sigma^2}}$$

donde  $\sigma$  es la desviación estándar de la función Gaussiana.

La definición de derivadas de Gaussianas sin considerar la constante de normalización y una desviación estándar  $\sigma = 1$  es:

$$G_n(x) = \frac{d^n e^{-x^2}}{dx^n} \quad (C.1)$$

Introduciendo la constante de normalización  $L^2$ , podemos escribir a las  $G_n$  en términos de  $G_0$  como:

$$G_n(x)_{L^2} = \frac{d^n G_0(x)_{L^2}}{dx^n} = \frac{1}{\sqrt{\sqrt{\pi}\sigma}} \frac{d^n e^{-\frac{x^2}{2\sigma^2}}}{dx^n} \quad (C.2)$$

La función ventana  $V$  utilizada en la obtención de los filtros de análisis (funciones filtro) de la transformada Hermite son iguales a la función Gaussiana  $G_0$  en  $L^2$ :

$$V(x) = G_0(x)_{L^2} = \frac{1}{\sqrt{\sqrt{\pi}\sigma}} e^{-\frac{x^2}{2\sigma^2}} \quad (C.3)$$

las funciones filtro en la transformada Hermite están definidas como:

$$D_n(x) = (-1)^n \cdot H\left(\frac{x}{\sigma}\right) \cdot V^2(x) \quad (C.4)$$

sustituyendo la ecuación C.3 en C.4:

$$\begin{aligned} D_n(x) &= (-1)^n \cdot H\left(\frac{x}{\sigma}\right) \cdot \left[ \frac{1}{\sqrt{\sqrt{\pi}\sigma}} e^{-\frac{x^2}{2\sigma^2}} \right]^2 \\ &= (-1)^n \cdot H\left(\frac{x}{\sigma}\right) \cdot \frac{1}{\sqrt{\pi}\sigma} e^{-\frac{x^2}{\sigma^2}} \end{aligned} \quad (C.5)$$

pero

$$H\left(\frac{x}{\sigma}\right) = \frac{(-1)^n}{\sqrt{2^n n!}} \cdot e^{\frac{x^2}{\sigma^2}} \cdot \frac{d^n e^{-\frac{x^2}{\sigma^2}}}{dx^n} \quad (C.6)$$

sustituyendo esta última ecuación en C.5 obtenemos:

$$D_n(x) = \frac{(-1)^{2n}}{\sqrt{2^n n!}} \cdot e^{\frac{x^2}{\sigma^2}} \cdot \frac{d^n e^{-\frac{x^2}{\sigma^2}}}{dx^n} \cdot \frac{1}{\sqrt{\pi\sigma}} e^{-\frac{x^2}{\sigma^2}}$$

$$D_n(x) = \frac{1}{\sqrt{2^n n!}} \cdot \frac{1}{\sqrt{\pi\sigma}} \cdot \frac{d^n e^{-\frac{x^2}{\sigma^2}}}{dx^n} \quad (\text{C.7})$$

Donde es posible notar que, excepto por el factor  $(1/\sqrt{2^n n!})$ , la ecuación anterior corresponde al cuadrado de la ecuación de derivadas de Gaussiana con  $L^2$  definidas en la ecuación C.2. Entonces podemos escribir a  $D_n(x)$  como [57]:

$$D_n(x) = \frac{1}{\sqrt{2^n n!}} \cdot [G_n(x)_{L^2}]^2 \quad (\text{C.8})$$

Con este resultado se puede observar también que las funciones de análisis o funciones filtro se pueden obtener a partir de derivadas de Gaussiana no normalizadas como las de la ecuación C.1 evaluadas en  $x = \frac{x}{\sigma}$ , pero multiplicadas por una constante de normalización  $L^2$  [55]:

$$D_n(x) = \frac{1}{\sqrt{2^n n!}} \cdot \frac{1}{\sqrt{\pi\sigma}} G_n\left(\frac{x}{\sigma}\right) \quad (\text{C.9})$$

## D Obtención de las funciones patrón como filtros de síntesis de la transformada Hermite

Las funciones patrón están definidas como:

$$P_n(x) = H_n\left(\frac{x}{\sigma}\right) V(x)/W(x) \quad (\text{D.1})$$

donde  $V(x)$  es la ventana de la función Gaussiana definida en C.3 usando  $L^2$ :

$$V(x) = \frac{1}{\sqrt{\sqrt{\pi}\sigma}} e^{-\frac{x^2}{2\sigma^2}} \quad (\text{D.2})$$

La función de peso  $W(x)$  está dada por:

$$W(x) = \frac{\sqrt{2\sqrt{\pi}\sigma}}{T} w(x) \quad (\text{D.3})$$

donde

$$w(x) = 1 + 2 \sum_{k=1}^{\infty} \exp\left[-\frac{1}{2} \left(k \frac{2\pi\sigma}{T}\right)^2\right] \cdot \cos\left(k \frac{2\pi x}{T}\right)$$

Entonces podemos escribir las funciones patrón como en función de D.2 y de D.3 como:

$$\begin{aligned} P_n(x) &= H_n\left(\frac{x}{\sigma}\right) \cdot \frac{1}{\sqrt{\sqrt{\pi}\sigma}} e^{-\frac{x^2}{2\sigma^2}} \cdot \frac{T}{\sqrt{2\sqrt{\pi}\sigma}} / w(x) \\ &= \frac{T}{\sigma\sqrt{2\pi}} \cdot H_n\left(\frac{x}{\sigma}\right) e^{-\frac{x^2}{2\sigma^2}} / w(x) \\ &= \frac{T}{\sigma\sqrt{2\pi}} \cdot \frac{(-1)^n}{\sqrt{2^n n!}} \cdot e^{\frac{x^2}{2\sigma^2}} \frac{d^n e^{-\frac{x^2}{\sigma^2}}}{dx^n} / w(x) \end{aligned}$$



## E Algoritmos de la transformada Hough

---

**Algoritmo E.1** TransformadaHough parte 1. El algoritmo encuentra y acumula los posibles centros de los círculos.

---

```
1:  $C_A[x, y] \leftarrow 0$ 
2: for  $h = 0$  to  $ImHeight$  do
3:   for  $w = 0$  to  $ImWidth$  do
4:      $Value \leftarrow M[w, h]$  { $M$  es el mapa de bordes binario}
5:     if  $Value > 0$  then
6:        $Angle \leftarrow GradientOr[w, h]$ 
7:       if ((( $Angle \geq -ConeWidth$ ) and ( $Angle \leq ConeWidth$ )) or (( $Angle \geq \pi - ConeWidth$ )
or ( $Angle \leq -\pi + ConeWidth$ ))) then
8:          $Px \leftarrow -\cos(Angle)$ ;  $Py \leftarrow -\sin(Angle)$ 
9:         for  $Radio = r_{min}$  to  $r_{max}$  do
10:           $Pxx \leftarrow \text{round}(Radio \times Px)$ ;  $Pyy \leftarrow \text{round}(Radio \times Py)$ 
11:          {Verificamos que las coordenadas esten dentro de la imagen}
12:          if  $InsideImage(w + Pxx, h + Pyy)$  then
13:             $Pxx \leftarrow w + Pxx$ ;  $Pyy \leftarrow h + Pyy$ 
14:             $C_A[Pxx, Pyy] \leftarrow C_A[Pxx, Pyy] + 1$ 
15:          end if
16:        end for
17:      end if
18:    end if
19:  end for
20: end for
```

---

**Algoritmo E.2** TransformadaHough parte 2. El algoritmo selecciona las coordenadas del centro  $(x_0, y_0)$  que recibió la mayor cantidad de votos.

---

```
1:  $Max \leftarrow 0$ 
2: for  $h = 0$  to  $ImHeight$  do
3:   for  $w = 0$  to  $ImWidth$  do
4:     if  $C_A[w, h] > Max$  then
5:        $Max \leftarrow C_A[w, h]$ 
6:        $x_0 \leftarrow w$ ;  $y_0 \leftarrow h$ 
7:     end if
8:   end for
9: end for
```

---

---

**Algoritmo E.3** TransformadaHough parte 3. El algoritmo acumula los radios posibles del círculo.

---

```
1:  $R_A[r] \leftarrow 0$ 
2:  $RadIndex \leftarrow 0$ 
3:  $Max \leftarrow 0$ 
4: for  $r = r_{min}$  to  $r_{max}$  do
5:   for  $Angle = -ConeWidth$  to  $ConeWidth$  do
6:      $Pxx \leftarrow \text{round}(r \times \cos(Angle)); Pyy \leftarrow \text{round}(r \times \sin(Angle))$ 
7:     {Verificamos que las coordenadas esten dentro de la imagen}
8:     if  $\text{InsideImage}(x_0 + Pxx, y_0 + Pyy)$  then
9:        $Value \leftarrow M[x_0 + Pxx, y_0 + Pyy]$ 
10:      if  $Value > 0$  then
11:         $R_A[RadIndex] \leftarrow R_A[RadIndex] + 1$ 
12:      end if
13:    end if
14:     $Pxx \leftarrow -\text{round}(r \times \cos(Angle))$ 
15:    {Verificamos que las coordenadas esten dentro de la imagen}
16:    if  $\text{InsideImage}(x_0 + Pxx, y_0 + Pyy)$  then
17:       $Value \leftarrow M[x_0 + Pxx, y_0 + Pyy]$ 
18:      if  $Value > 0$  then
19:         $R_A[RadIndex] \leftarrow R_A[RadIndex] + 1$ 
20:      end if
21:    end if
22:     $RadIndex \leftarrow RadIndex + 1$ 
23:  end for
24: end for
```

---

---

**Algoritmo E.4** TransformadaHough parte 4. El algoritmo selecciona el radio del círculo  $R_0$  que recibió la mayor cantidad de votos.

---

```
1:  $Radii = r_{max} - r_{min}$ 
2:  $Max \leftarrow 0$ 
3: for  $r = 0$  to  $Radii$  do
4:   if  $R_A[r] > Max$  then
5:      $Max \leftarrow R_A[r]$ 
6:      $R_0 \leftarrow r + r_{min}$ 
7:   end if
8: end for
```

---

# Bibliografía

- [1] ABRAMOWITZ, M., AND STEGUN, I. *Handbook of Mathematical Functions*. Dover, 1965.
- [2] ADLER, F. H. *Physiology of the Eye: Clinical Application*. London: The C.V. Mosby Company, 1965.
- [3] AL-ZUBI, R., AND ABU-AL-NADI, D. Automated personal identification system based on human iris analysis. *Pattern Analysis & Applications* 10, 2 (2007), 147–164.
- [4] ARIVAZHAGAN, S., GANESAN, L., AND BAMA, S. Fault segmentation in fabric images using Gabor wavelet transform. *Machine Vision and Applications* 16, 6 (January 2006), 356–363.
- [5] BARTLETT, M. S., MOVELLAN, J. R., AND SEJNOWSKI, T. J. Face recognition by independent component analysis. *IEEE Transactions on Neural Networks* 13, 6 (2002), 1450–1464.
- [6] BERNE, R. M., LEVY, M. N., STANTON, B. A., AND KOEPPEN, B. M. *Berne y Levy Fisiología*. Elsevier, 2006. Traducción de: Principles of physiology. 4th ed.
- [7] BIGUN, J., AND DU BUF, J. M. H. N-folded symmetries by complex moments in Gabor space and their application to unsupervised texture segmentation. *Transactions on Pattern Analysis and Machine Intelligence* 16, 1 (1994), 80–87.
- [8] BOLES, W. W., AND BOASHASH, B. A human identification technique using images of the iris and wavelet transform. *IEEE Transactions on Signal Processing*, [see also *IEEE Transactions on Acoustics, Speech, and Signal Processing*] 46, 4 (1998), 1185–1188.
- [9] BOLME, D. S., BEVERIDGE, J. R., TEIXEIRA, M., AND DRAPER, B. A. The CSU face identification evaluation system: Its purpose, features, and structure. In *Computer Vision Systems: Third International Conference, ICVS 2003* (Graz, Austria, April 2003), vol. 2626, Springer Berlin / Heidelberg, pp. 304–313.
- [10] BOVIK, A. C., CLARK, M., AND GEISLER, W. S. Multichannel texture analysis using localized spatial filters. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 12, 1 (1990), 55–73.
- [11] CANNY, F. J. A computational approach to edge detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 8 (1986), 679–698.
- [12] CHAN, R., AND SIU, W. C. A new approach for efficient Hough transform for circles. In *IEEE Pacific Rim Conference on Communications, Computers and Signal Processing* (1989), pp. 99–102.
- [13] CHEN, W.-S., CHIH, K.-H., SHIH, S.-W., AND HSIEH, C.-M. Personal identification technique based on human iris recognition with wavelet transform. In *IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP-05)* (2005), vol. 2, pp. 949–952.

- [14] CHINESE ACADEMY OF SCIENCES' INSTITUTE OF AUTOMATION. CASIAv1.0 Iris Image Database. <http://www.cbsr.ia.ac.cn/english/Iris%20Databases.asp>, (último acceso, Febrero 2008).
- [15] CLAUSI, D. A., AND ED JERNIGAN, M. Designing Gabor filters for optimal texture separability. *Pattern Recognition* 33, 11 (2000), 1835–1849.
- [16] DAUGMAN, J. G. Uncertainty relation for resolution in space, spatialfrequency, and orientation optimized by two-dimensional visual cortical filters. *J. Opt. Soc. Am. A* 2, 7 (1985), 1160–1169.
- [17] DAUGMAN, J. G. Complete discrete 2-D Gabor transforms by neural networks for image analysis and compression. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, [see also *IEEE Transactions on Signal Processing*] 36, 7 (1988), 1169–1179.
- [18] DAUGMAN, J. G. Relaxation neural network for nonorthogonal image transforms. In *IEEE International Conference on Neural Networks* (1988), vol. 1, pp. 547–560.
- [19] DAUGMAN, J. G. High confidence visual recognition of persons by a test of statistical independence. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 15, 11 (1993), 1148–1161.
- [20] DAUGMAN, J. G. The importance of being random: statistical principles of iris recognition. *Pattern Recognition* 36, 2 (2003), 279–291.
- [21] DAUGMAN, J. G. How iris recognition works. *IEEE Transactions on Circuits and Systems for Video Technology* 14, 1 (2004), 21–30.
- [22] DAUGMAN, J. G. New methods in iris recognition. *IEEE Transactions on Systems, Man and Cybernetics, Part B* 37, 5 (2007), 1167–1175.
- [23] DE VALOIS, R. L., ALBRECHT, D. G., AND THORELL, L. G. Spatial frequency selectivity of cells in macaque visual cortex. *Vision Research* 22, 5 (1982), 545–559.
- [24] DECO, G., AND SCHÜRMAN, B. Predictive coding in the visual cortex by a recurrent network with Gabor receptive fields. *Neural Processing Letters* 14, 2 (2001), 107–114.
- [25] DING, L., AND GOSHTASBY, A. On the Canny edge detector. *Pattern Recognition* 34, 3 (2001), 721–725.
- [26] DUNN, D. F., AND HIGGINS, W. E. Optimal Gabor-filter design for texture segmentation. In *IEEE International Conference on Acoustics, Speech, and Signal Processing, (ICASSP-93)* (1993), vol. 5, pp. 37–40.
- [27] ESCALANTE-RAMIREZ, B., AND LOPEZ-CALOCA, A. A. Image fusion with the multiscale Hermite transform. In *Applications of Digital Image Processing XXIX* (San Diego, CA, USA, 2006), A. G. Tescher, Ed., vol. 6312, Proceedings of SPIE.
- [28] ESCALANTE-RAMÍREZ, B., AND MARTENS, J. B. Noise reduction in computerized tomography images by means of polynomial transforms. *Journal of Visual Communication and Image Representation* 3 (1992), 272–285.

- [29] ESCALANTE-RAMIREZ, B., SILVAN-CARDENAS, J. L., AND YUEN-ZHUO, H. Optic flow estimation using the Hermite transform. In *Applications of Digital Image Processing XXVII* (Denver, CO, USA, 2004), A. G. Tescher, Ed., vol. 5558, Proceedings of SPIE, pp. 632–643.
- [30] ESCOFET, J., NAVARRO, R., MILLAN, M. S., AND PLADELLORENS, J. Detection of local defects in textile webs using Gabor filters. *Optical Engineering* 37, 8 (1998), 2297–2307.
- [31] FIELD, D. J. Relations between the statistics of natural images and the response properties of cortical cells. *J. Opt. Soc. Am. A* 4 (1987), 2379–2394.
- [32] FLOM, L., AND SAFIR, A. U.S. Patent no. 4641349. U.S. Government Printing Office, 1987.
- [33] FREEMAN, W. T., AND ADELSON, E. H. The design and use of steerable filters. *Transactions on Pattern Analysis and Machine Intelligence* 13, 9 (1991), 891–906.
- [34] FRIGO, M., AND JOHNSON, S. G. The design and implementation of FFTW3. *Proceedings of the IEEE* 93, 2 (2005), 216–231.
- [35] GANESHAN, B., THECKEDATH, D., YOUNG, R., AND CHATWIN, C. Biometric iris recognition system using a fast and robust iris localization and alignment procedure. *Optics and Lasers in Engineering* 44, 1 (2006), 1–24.
- [36] GARFINKEL, S. *Database Nation: The Death of Privacy in the 21st Century*, 1 ed. O’Reilly and Associates, January 2000.
- [37] GUYTON, A. C., AND HALL, J. E. *Textbook of Medical Physiology*. Elsevier saunders, 2006.
- [38] HOOSER, S. D. V., AND NELSON, S. B. Visual system. In *Encyclopedia of Life Sciences*. John Wiley & Sons, Ltd, 2005.
- [39] HOUGH, P. V. C. Method and means for recognizing complex patterns, U.S. Patent no. 3069654, 1962.
- [40] HUANG, J., WANG, Y., TAN, T., AND CUI, J. A new iris segmentation method for recognition. *Proceedings of the 17th International Conference on Pattern Recognition, (ICPR 2004)* 3 (2004), 554–557.
- [41] HUBEL, D. H., AND WIESEL, T. N. Receptive fields, binocular interaction, and functional architecture in the cat’s visual cortex. *J. Physiol.* 1, 160 (January 1962), 106–154.
- [42] IRIDIAN TECHNOLOGIES, INC. Iridian technologies. <http://www.iridianttechnologies.com>, (último acceso, Febrero 2008).
- [43] JAIN, A. K., AND FARROKHANIA, F. Unsupervised texture segmentation using Gabor filters. *IEEE International Conference on Systems, Man and Cybernetics* (1990), 14–19.
- [44] LEE, C.-J., WANG, S.-D., AND WU, K.-P. Fingerprint recognition using principal Gabor basis function. In *Proceedings of 2001 International Symposium on Intelligent Multimedia, Video and Speech Processing* (2001), pp. 393–396.
- [45] LEE, T. S. Image representation using 2D Gabor wavelets. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 18, 10 (1996), 959–971.

- [46] LEFOHN, A., BUDGE, B., SHIRLEY, P., CARUSO, R., AND REINHARD, E. An ocularist's approach to human iris synthesis. *IEEE Comput. Graph. Appl.* 23, 6 (2003), 70–75.
- [47] LI, X., AND XIE, M. A novel algorithm of human iris recognition. In *IEEE International Symposium on Communications and Information Technology, (ISCIT 2005)* (2005), vol. 2, pp. 1231–1234.
- [48] LIM, S., LEE, K., BYEON, O., AND KIM, T. Efficient iris recognition through improvement of feature vector and classifier. *ETRI Journal* 23, 2 (June 2001), 61–70.
- [49] LITZINGER, T. C., AND RIO-TSONIS, K. D. Eye anatomy. In *Encyclopedia of Life Sciences*. John Wiley & Sons, Ltd, 2002.
- [50] LU, J., PLATANIOTIS, K. N., AND VENETSANOPOULOS, A. N. Regularized discriminant analysis for the small sample size problem in face recognition. *Pattern Recognition Letters* 24, 16 (2003), 3079–3087.
- [51] MA, L., WANG, Y., AND TAN, T. Iris recognition using circular symmetric filters. In *16th International Conference on Pattern Recognition* (2002), vol. 2, pp. 414–417.
- [52] MALLAT, S. G. A theory for multiresolution signal decomposition: the wavelet representation. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 11, 7 (1989), 674–693.
- [53] MANJUNATH, B. S., AND MA, W. Y. Texture features for browsing and retrieval of image data. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 18, 8 (1996), 837–842.
- [54] MARCELJA, S. Mathematical description of the responses of simple cortical cells. *J. Opt. Soc. Am.* 70 (1980), 1297–1300.
- [55] MARTENS, J. B. The Hermite transform-theory. *IEEE Transactions on Acoustics, Speech, and Signal Processing [see also IEEE Transactions on Signal Processing]* 38, 9 (1990), 1595–1606.
- [56] MELSA, J. L. *Decision and estimation theory*. McGraw-Hill, 1978.
- [57] MICHAELIS, M., AND SOMMER, G. Basic functions for early vision. Bericht (Reporte técnico) 9413, Institut für Informatik und Praktische Mathematik Christian-Albrechts-Universität zu Kiel, August 1994.
- [58] NATIONAL SCIENCE AND TECHNOLOGY COUNCIL. Biometrics foundation documents. <http://www.biometrics.gov>, (último acceso, Agosto 2007).
- [59] NHAT, V. D. M., LEE, S. Y., AND YOUN, H. Y. Whitened LDA for face recognition. In *Proceedings of the 6th ACM international conference on Image and video retrieval (CIVR '07)* (New York, NY, USA, 2007), ACM Press, pp. 234–241.
- [60] PARK, C.-H., LEE, J.-J., SMITH, M. J. T., IL PARK, S., AND PARK, K.-H. Directional filter bank-based fingerprint feature extraction and matching. *IEEE Transactions on Circuits and Systems for Video Technology* 14, 1 (2004), 74–85.
- [61] POLLEN, D. A., AND RONNER, S. F. Phase relationships between adjacent simple cells in the visual cortex. *Science* 212, 4501 (1981), 1409–1411.

- [62] POURSABERI, A., AND ARAABI, B. N. A half-eye wavelet based method for iris recognition. In *Proceedings of the 5th International Conference on Intelligent Systems Design and Applications, (ISDA '05)* (2005), pp. 262–267.
- [63] PROENÇA, H., AND ALEXANDRE, L. A. UBIRIS: A noisy iris image database. In *Proceedings of the International Conference on Image Analysis and Processing, (ICIAP 2005)* (2005), vol. 1, pp. 970–977.
- [64] REAL ACADEMIA ESPAÑOLA. *Diccionario de la lengua española*, 22 ed. Real Academia Española (versión electrónica, <http://www.rae.es>), 2001.
- [65] RILEY, K. F., HOBSON, M. P., AND BENCE, S. J. *Mathematical Methods for Physics and Engineering*. Cambridge University Press, 2006.
- [66] RIVERO-MORENO, C. J., AND BRES, S. Texture feature extraction and indexing by Hermite filters. In *Proceedings of the 17th International Conference on Pattern Recognition, (ICPR 2004)* (2004), vol. 1, pp. 684–687.
- [67] SANCHEZ-AVILA, C., AND SANCHEZ-REILLO, R. Two different approaches for iris recognition using Gabor filters and multiscale zero-crossing representation. *Pattern Recognition* 38, 2 (2005), 231–240.
- [68] SHEN, L., BAI, L., AND FAIRHURST, M. Gabor wavelets and general discriminant analysis for face identification and verification. *Image and Vision Computing* 25, 5 (2007), 553–563.
- [69] SILVAN-CARDENAS, J. L., AND ESCALANTE-RAMIREZ, B. The multiscale Hermite transform for local orientation analysis. *IEEE Transactions on Image Processing* 15, 5 (2006), 1236–1253.
- [70] STORK, D. G., AND WILSON, H. R. Do Gabor functions provide appropriate descriptions of visual cortical receptive fields? *J. Opt. Soc. Am. A* 7, 8 (August 1990), 1362–1373.
- [71] SUBIRANA, M., LUCENA, J., PLANCHAT, L., CUQUERELLA, A., FERREIRO, L., AND VÁZQUEZ, M. La huella perdida: Identificación personal utilizando un dedo parcialmente amputado encontrado en la escena de un robo. *Cuadernos de Medicina Forense [online]*, 40 (2005), 139–144.
- [72] THE GIMP TOOLKIT. GTK+-2.0. <http://www.gtk.org>, (último acceso, Noviembre 2007).
- [73] TSAI, D.-M., WU, S.-K., AND CHEN, M.-C. Optimal Gabor filter design for texture segmentation using stochastic optimization. *Image and Vision Computing* 19, 5 (2001), 299–316.
- [74] VAN DIJK, A. M., AND MARTENS, J. B. Feature-based image compression with steered Hermite transforms. In *International Conference on Image Processing* (1996), vol. 1, pp. 205–208.
- [75] WILDES, R. P. Iris recognition: an emerging biometric technology. *Proceedings of the IEEE* 85, 9 (1997), 1348–1363.
- [76] WILDES, R. P., ASMUTH, J. C., GREEN, G. L., HSU, S. C., KOLCZYNSKI, R. J., MATEY, J. R., AND MCBRIDE, S. E. A system for automated iris recognition. In *Proceedings of the Second IEEE Workshop on Applications of Computer Vision* (1994), pp. 121–128.
- [77] WISKOTT, L., FELLOUS, J. M., KUIGER, N., AND VON DER MALSBERG, C. Face recognition by elastic bunch graph matching. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 19, 7 (1997), 775–779.

- [78] YANG, J. Do Gabor functions provide appropriate descriptions of visual cortical receptive fields, comment. *J. Opt. Soc. Am. A* 9, 2 (February 1992), 334–336.
- [79] YANGXING, L., SATOSHI, G., AND TAKESHI, I. An accurate and low complexity approach of detecting circular shape objects in still color images. In *IEEE International Conference on Image Processing (ICIP 2005)* (2005), vol. 1, pp. 333–336.
- [80] YAO, P., LI, J., YE, X., ZHUANG, Z., AND LI, B. Iris recognition algorithm using modified log-Gabor filters. In *18th International Conference on Pattern Recognition, (ICPR 2006)* (2006), vol. 4, pp. 461–464.
- [81] YOUNG, R. A. Orthogonal basis functions for form vision derived from eigenvector analysis. In *ARVO Abstracts* (Sarasota, FL, 1978), Association for Research in Vision and Ophthalmology, Abstract, p. 22.
- [82] YOUNG, R. A. The Gaussian derivative theory of spatial vision: Analysis of cortical cell receptive field line-weighting profiles. Tech. Rep. GRM-4920, General Motors Research, 1985.
- [83] YOUNG, R. A., AND LESPERANCE, R. M. The Gaussian derivative model for spatial-temporal vision: II. Cortical data. *Spatial Vision* 14, 3,4 (March 2001), 321–389.
- [84] YOUNG, R. A., LESPERANCE, R. M., AND MEYER, W. W. The Gaussian derivative model for spatial-temporal vision: I. Cortical model. *Spatial Vision* 14, 3,4 (March 2001), 261–319.
- [85] ZHANG, D. D. Texture feature using 2-D Gabor filters. In *Palmprint Authentication*, D. D. Zhang and A. K. Jain, Eds., vol. 3 of *Kluwer International Series on Biometrics*. Springer US, 2004, ch. III, pp. 85–98.
- [86] ZHANG, P.-F., LI, D.-S., AND WANG, Q. A novel iris recognition method based on feature fusion. In *Proceedings of 2004 International Conference on Machine Learning and Cybernetics* (2004), vol. 6, pp. 3661–3665.
- [87] ZHU, Y., TAN, T., AND WANG, Y. Biometric personal identification based on iris patterns. In *Proceedings of the 15th International Conference on Pattern Recognition* (2000), vol. 2, pp. 801–804.