



UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO
POSGRADO EN CIENCIA E INGENIERÍA DE LA COMPUTACIÓN

**CARACTERIZACIÓN DE TEXTURAS USANDO
MODELOS DE VISIÓN BIOLÓGICOS**

TESIS

**QUE PARA OPTAR POR EL GRADO DE:
DOCTOR EN CIENCIAS (COMPUTACIÓN)**

PRESENTA:

ALFONSO ESTUDILLO ROMERO

TUTOR:

DR. BORIS ESCALANTE RAMÍREZ
FACULTAD DE INGENIERÍA

MÉXICO, D. F., FEBRERO DE 2013



Universidad Nacional
Autónoma de México



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.

Agradecimientos

A todas las personas que influyeron positivamente en la elaboración de este proyecto de tesis a lo largo de todo este tiempo.

A mi tutor, el Dr. Boris Escalante, por su invaluable apoyo a lo largo de este proceso y, a su casi agotada ya, paciencia para que lograra terminar una etapa más de mi proyecto de vida.

A la Dra. María Elena Martínez, al Dr. Ernesto Bribiesca, al Dr. Fernando Arámbula, al Dr. Edgar Garduño y al Dr. Carlos Rivero por todos sus comentarios críticos, me han ayudado a fortalecer el análisis y las buenas maneras para conducirse en la investigación, forman parte sin duda alguna de ejemplos a seguir.

A mi familia, amigos y colegas.

Al CONACyT por la beca de estudios de Doctorado CVU. 173482

A la UNAM a través de los proyectos PAPIIT IN 113611 e IG 100814

Al Posgrado en Ciencia e Ingeniería de la Computación de la UNAM.

A los tres motores de mi vida: mi madre, mi abuelita y Marthita.

Resumen

El análisis de las texturas es uno de los temas que ha sido de gran importancia en el procesamiento digital de imágenes debido a que la información de la textura extraída de las imágenes ha sido utilizada en distintas aplicaciones como segmentación, clasificación, búsqueda de imágenes por contexto, percepción remota, análisis de documentos y sistemas biométricos entre otros. La descripción invariante a la rotación de las texturas a través de representaciones compactas se puede utilizar como información de bajo nivel para complementar a una metodología que caracterice no sólo a una textura sino que incluso pueda caracterizar a una imagen en su totalidad.

En este trabajo se utilizaron dos métodos basados en filtros para caracterizar a las texturas, los filtros de análisis de Hermite y los filtros Gabor. El diseño de estos filtros está inspirado biológicamente en los modelos de los sistemas de visión encontrados en algunos mamíferos. El objetivo principal de este trabajo es encontrar una representación compacta e invariante a la rotación de la textura usando modelos de visión inspirados biológicamente para el análisis de las imágenes. Para obtener representaciones compactas de las texturas se emplearon algunos métodos de reducción de dimensión. La metodología experimental incluyó la evaluación de diferentes tipos de características de textura. Cada tipo de característica de textura fue evaluado individualmente en términos de la precisión de la clasificación para conocer los tipos de características de textura que son los más adecuados. En este trabajo también fueron evaluadas distintas medidas de distancia. Las evaluaciones estadísticas de la tasa de clasificación correcta de varios experimentos fueron utilizadas para encontrar las combinaciones del método de reducción de dimensión, el tipo de característica de textura y las medidas de distancia que son las más adecuadas para la clasificación de texturas filtradas con los filtros de análisis de Hermite. Evaluaciones similares también sirvieron para encontrar las combinaciones del ancho de banda frecuencial y del número de orientaciones para el caso de las texturas filtradas con los bancos de filtros de análisis Gabor.

Los resultados de la clasificación de los métodos empleados para la reducción del tamaño del vector de características permitieron observar que la *tasa de varianza aumentada* (AVR) es la más competitiva respecto al análisis de la varianza de inter-clase (ICV) y que al análisis por componentes principales (PCA), en este orden. La ventaja de los métodos de reducción de dimensión AVR e ICV, además de las tasas de clasificación altas que se pudieron obtener con vectores de características pequeños, es que permitieron seleccionar los ordenes de los filtros de análisis a partir de un conjunto pequeño de texturas de entrenamiento y como consecuencia se obtuvo una disminución de la complejidad computacional durante la etapa de prueba o clasificación. Respecto a los distintos tipos de características utilizados, se observó en la mayoría de los experimentos que solo algunos de estos tipos propiciaron tasas de clasificación altas. Para el caso del análisis de texturas usando los filtros Gabor se

encontró que cierta cantidad de traslape, de los filtros de análisis en el dominio frecuencial, es útil para caracterizar mejor de las texturas.

Abstract

Texture analysis has been an important research area on digital image processing due to the fact that texture information extracted from images has been used on applications such as segmentation, classification, context-based image retrieval, remote sensing, document analysis and biometrics for security. A rotation invariant texture descriptor through compact representations can be used as a low level information provider for a methodology that describes not only textures but also whole images.

In this work two filter-based methods were used in order to form texture descriptors, such filters are known as the Hermite and Gabor analysis filters. These filter designs are biologically inspired on vision systems found on some mammals. The main objective of this work is to find compact and rotation-invariant representations of textures by using biologically inspired vision models. In order to obtain compact texture representations some dimensionality reduction methods were used. The evaluation of different types of texture features was included on the experimental agenda. Classification accuracy of every single type of texture feature was evaluated in order to realize which types of texture features are the most suitable. Different distance measurements were also evaluated in this work. Statistical analysis were performed on the computed correct classification rates of several experiments in order to find combinations of the dimensionality reduction method, type of texture feature and distance measure that best suit the texture classification by filtering with the Hermite analysis filters. Similar analysis were also performed for the filtered textures that were obtained with the Gabor filters banks in order to find combinations of frequency bandwidths and the number of orientations.

Comparison of the classification results between the different dimensionality reduction methods allowed to choose the augmented-variance ratio (AVR) as the most suitable method followed by the inter-class variance (ICV) and finally by the principal components analysis. Benefits of the AVR and ICV methods are high classification rates and small feature vectors. Moreover, the knowledge of the vector elements that performed better helped to select the analysis filters orders from a small set of training textures and, then the computational complexity was reduced during the testing or classification stage. High classification rates were obtained for some types of texture features. Statistical analysis on the classification results with Gabor filters yielded that some overlapping between the frequency filter responses is required in order to obtain high classification rates.

Índice general

Índice general	XI
Índice de figuras	XIII
Índice de tablas	XVII
1. Introducción	1
1.1. Antecedentes	1
1.2. Objetivo	4
1.2.1. Objetivos particulares	4
1.3. Contribución de la tesis	4
1.4. Estructura de la tesis	5
2. Metodología	7
2.1. Justificación	7
2.2. Materiales y resumen de los métodos empleados	7
2.3. Análisis de texturas con métodos de selección y reducción de dimensión	8
3. Análisis de texturas con métodos de filtrado	11
3.1. Características estadísticas de la textura	12
3.2. Vector de Características	14
3.3. Distancias	14
3.4. Invarianza a la rotación	16
3.5. Discusión	18
4. Procesamiento de texturas inspirado en sistemas de visión biológicos	19
4.1. Transformada de Hermite	20
4.1.1. Transformada de Hermite orientada	24
4.1.2. Filtros de análisis de Hermite en el dominio frecuencial 2D	25
4.2. Filtros Gabor	27
4.2.1. Filtros Gabor en el dominio frecuencial 2D	29
4.3. Análisis de texturas invariante a la rotación usando la Transformada Hermite orientada	30
4.4. Análisis de texturas invariante a la rotación usando filtros Gabor	32

5. Separación de clases, métodos de reducción de dimensión y clasificación	35
5.1. Criterio de Fisher	35
5.2. Análisis por componentes principales (PCA)	36
5.3. Tasa de varianza aumentada (AVR)	37
5.4. Varianza de inter-clase (ICV)	38
5.5. Clasificación	38
5.5.1. Análisis de la sensibilidad y especificidad del sistema	39
6. Resultados	43
6.1. Resultados usando la transformada de Hermite	47
6.1.1. Criterio de Fisher usando la transformada de Hermite	47
6.1.2. Reducción del número de elementos que conforman al vector de características	48
6.1.3. Interacción de los factores utilizando un vector de dimensión reducida	53
6.1.4. Discusión de resultados	60
6.1.5. Sensibilidad del sistema usando la transformada de Hermite	60
6.1.6. Determinación del orden promedio de los filtros que conforman los vectores de 20 elementos	65
6.2. Resultados usando los filtros Gabor	69
6.2.1. Diseño de bancos de filtros Gabor	69
6.2.2. Determinación experimental de la combinación del ancho de banda ϕ y del número de orientaciones K	71
6.2.3. Criterio de Fisher usando los filtros Gabor	74
6.2.4. Sensibilidad del sistema usando los filtros Gabor	79
7. Conclusiones y perspectivas	83
A. Tasas de clasificación correcta obtenidas con los métodos AVR, ICV y PCA a partir de características extraídas con la transformada de Hermite orientada	87
B. Polinomios de Hermite	91
C. Normalización L^1 y L^2 de la función Gaussiana	97
D. Obtención de las derivadas de Gaussianas como filtros de análisis de la transformada Hermite	99
E. Análisis de la complejidad computacional	101
E.1. Análisis de la complejidad sin métodos de selección de filtros	101
E.1.1. Filtrado con bancos de filtros Gabor	101
E.1.2. Filtrado con filtros de análisis de Hermite	101
E.2. Análisis de la complejidad con métodos de selección de filtros	102
Bibliografía	103

Índice de figuras

2.1. Metodología: análisis de texturas y reducción de dimensión	8
4.1. Filtros de análisis de Hermite	22
4.2. Coeficientes de la transformada de Hermite cartesiana y orientada	23
4.3. Funciones de análisis de Hermite en el espacio frecuencial	26
4.4. Filtros Gabor real e imaginario	28
4.5. Banco de filtros Gabor en el dominio frecuencial	29
4.6. Rotación de una textura y sus coeficientes 1D de Hermite orientados	31
4.7. Rotación de una textura y la suma de los distintos coeficientes direccionales de Gabor imaginario y real	34
5.1. Distribuciones en un problema típico de decisión binaria.	40
6.1. Comparación del criterio de Fisher entre las distintas combinaciones de tipos características y medidas de distancias para el conjunto de texturas Brodatz usando la transformada de Hermite orientada	48
6.2. Comparación del criterio de Fisher entre las distintas combinaciones de tipos características y medidas de distancias para el conjunto de texturas VisTex usando la transformada de Hermite orientada	49
6.3. Comparación del criterio de Fisher entre las distintas combinaciones de tipos características y medidas de distancias para el conjunto de texturas Outex usando la transformada de Hermite orientada	50
6.4. Interacción entre el método de reducción de dimensión y la medida de distancia e interacción entre el método de reducción de dimensión y el tipo de característica.	54
6.5. Interacción entre la base de datos y el método de reducción de dimensión e interacción entre la base de datos y la medida de distancia	55
6.6. Interacción entre los tipos de características y la base de datos e interacción entre los tipos de características y la medida de distancia	56
6.7. Interacción de los factores: distancia, tipos de características y base de datos	57
6.8. Interacción de los factores: distancia, tipos de características y método de reducción de dimensión	58
6.9. Interacción de los factores: distancia, método de reducción de dimensión y base de datos	58
6.10. Interacción de los factores: tipos de características, método de reducción de dimensión y base de datos	59

6.11. Comparación de curvas ROC usando la transformada de Hermite para la base de datos Brodatz	62
6.12. Comparación de curvas ROC usando la transformada de Hermite para la base de datos VisTex	63
6.13. Comparación de curvas ROC usando la transformada de Hermite para la base de datos Outex	64
6.14. Selección de los índices correspondientes a los 20 filtros más sobresalientes usando AVR para la base de datos Brodatz	66
6.15. Selección de los índices correspondientes a los 20 filtros más sobresalientes usando ICV para la base de datos Brodatz	66
6.16. Selección de los índices correspondientes a los 20 filtros más sobresalientes usando AVR para la base de datos VisTex	67
6.17. Selección de los índices correspondientes a los 20 filtros más sobresalientes usando ICV para la base de datos VisTex	67
6.18. Selección de los índices correspondientes a los 20 filtros más sobresalientes usando AVR para la base de datos Outex	68
6.19. Selección de los índices correspondientes a los 20 filtros más sobresalientes usando ICV para la base de datos Outex	68
6.20. Bancos de filtros Gabor utilizados en los experimentos	70
6.21. CCRs usando diversas combinaciones del ancho de banda frecuencial y del número de orientaciones para la base de datos Brodatz	72
6.22. CCRs usando diversas combinaciones del ancho de banda frecuencial y del número de orientaciones para la base de datos VisTex	74
6.23. CCRs usando diversas combinaciones del ancho de banda frecuencial y del número de orientaciones para la base de datos Outex	75
6.24. Comparación del criterio de Fisher entre las distintas combinaciones de tipos características y medidas de distancias para el conjunto de texturas Brodatz usando filtros Gabor	76
6.25. Comparación del criterio de Fisher entre las distintas combinaciones de tipos características y medidas de distancias para el conjunto de texturas VisTex usando filtros Gabor	77
6.26. Comparación del criterio de Fisher entre las distintas combinaciones de tipos características y medidas de distancias para el conjunto de texturas Outex usando filtros Gabor	78
6.27. Comparación de curvas ROC usando un banco de filtros Gabor para la base de datos Brodatz	80
6.28. Comparación de curvas ROC usando un banco de filtros Gabor para la base de datos VisTex	81
6.29. Comparación de curvas ROC usando un banco de filtros Gabor para la base de datos Outex	82
A.1. Comparación de CCRs usando la transformada de Hermite para la base de datos Brodatz	88

A.2. Comparación de CCRs usando la transformada de Hermite para la base de datos VisTex	89
A.3. Comparación de CCRs usando la transformada de Hermite para la base de datos Outex	90

Índice de tablas

4.1. Porcentajes de energía de los coeficientes de Hermite cartesianos	26
4.2. Porcentajes de energía de los coeficientes de Hermite orientados	26
4.3. Características de texturas invariantes a la rotación obtenidas de los coeficientes de la transformada de Hermite orientada	30
4.4. Características de texturas invariantes a la rotación obtenidas de los coeficientes Gabor real e imaginario	33
4.5. Características de texturas invariantes a la rotación obtenidas de la magnitud de la DFT del vector de características	33
6.1. Resumen de los tipos de características de textura utilizadas.	44
6.2. Efecto principal del número de elementos que conforman al vector de características r con cada método de reducción de dimensión: PCA, AVR, ICV.	51
6.3. Efecto principal de la medida de distancia.	52
6.4. Interacción de los factores ϕ y K para la elección de un banco de filtros Gabor	73

1. Introducción

1.1. Antecedentes

Durante las últimas décadas el análisis de texturas ha sido uno de los temas de gran interés y ha sido utilizado en algunas etapas del procesamiento de imágenes como en la segmentación [46, 55], clasificación [55, 66], indexado y recuperación de imágenes [52] (usadas en navegación por contexto en bases de datos multimedia); formando una parte esencial para el funcionamiento del proceso en su totalidad. El uso de la textura se puede encontrar en aplicaciones como percepción remota [2, 36], imagenología, análisis de documentos y sistemas biométricos [21]. En ambientes industriales, la inspección y la automatización de sistemas puede incorporar información de textura para evaluar la calidad de un proceso determinado [3, 31].

No existe una definición precisa de la textura, sin embargo, ésta tiene una gran cantidad de propiedades que generalmente es aceptada por muchos investigadores. La textura puede ser descrita sobre una vecindad espacial, el tamaño de la vecindad puede depender del tipo de textura y la escala o resolución a la cual se percibe la textura. Es posible relacionar a la textura con un conjunto de patrones repetitivos, algunas veces llamados “texturas primitivas”, que ocurren dentro de la vecindad. Los valores de nivel de gris de una textura también pueden formar una distribución y con ella se puede caracterizar a la textura. El hecho de que la percepción de la textura tenga tantas dimensiones es una razón importante por la cual no exista una sola técnica para representar a una gran variedad de texturas por lo cual el análisis de textura sigue siendo un tema de frontera.

En el trabajo de Tuceryan y Jain [77] se propone una categorización de los métodos de análisis de textura separando los métodos estadísticos, los métodos geométricos, los basados en modelos y los que están basados en el procesamiento de señales. Los métodos estadísticos, los basados en modelos y los métodos de procesamiento de señales son los más comunes que se pueden encontrar en la literatura del análisis de texturas. Las técnicas estadísticas más conocidas describen la textura en términos de la distribución espacial de los valores de nivel de gris de la textura usando características obtenidas de la matriz de co-ocurrencia y de la función de auto-correlación [69]. Las mejores técnicas basadas en modelos emplean los campos aleatorios de Markov (MRF) no solo para describir a la textura sino también para sintetizarla. Los MRF caracterizan las relaciones entre las intensidades de píxeles vecinos [79]. Finalmente, los métodos de procesamiento de señales, también conocidos como métodos basados en filtrado, han ganado importancia debido a la evidencia psicovisual de que el cerebro humano realiza un análisis por frecuencias de las imágenes [9]. Aunado a esto, existen modelos inspirados en el sistema de visión humano que simulan algunas de las etapas

primarias de visión y son capaces de extraer información visual de las imágenes [19, 57, 82].

A pesar de la falta de una definición precisa de la textura, frecuentemente es deseable que los métodos de análisis de textura cumplan con uno de los requisitos más importantes: que sean invariantes a la rotación o bien que se pueda determinar la direccionalidad de la textura. Los primeros métodos que consideraron la importancia de la *invarianza* a la rotación usaron “polarogramas” [22]. Los métodos basados en modelos propusieron un “modelo auto-regresivo de simetría circular” [43] y campos aleatorios de Markov Gaussianos (GMRF) [16] o combinaciones de filtros espejo en cuadratura en combinación con un modelo oculto de Markov (HMM) [12].

Los métodos recientes basados en la transformada *wavelet* proponen estrategias de *invarianza* a la rotación a través de etapas de pre-procesamiento sobre la imagen de textura; como una transformación polar [71], o bien una transformada Radón [40] de la textura. Sin embargo, una desventaja de esta estrategia es el aumento en el número de parámetros libres y como consecuencia es necesario realizar un análisis con mayor profundidad para determinar los parámetros óptimos para cada conjunto de datos. Por ejemplo, como ha sido indicado por los mismos autores [40], uno de los problemas más importantes al usar la transformada de Radón es encontrar el número óptimo de proyecciones que permita obtener el mínimo error de clasificación. Debido a que la tasa de muestro angular disminuye al incrementarse la distancia desde el origen en la dirección radial, las integrales de línea actúan como filtros pasobajas. Una posible solución es aumentar la tasa de muestreo angular. Sin embargo, la tasa de muestreo angular máxima deberá evitar incrementos en la orientación menores a pasos de un píxel. Por otro lado, se deben realizar muchos experimentos para determinar un número óptimo de proyecciones ya que éste depende en mayor parte al contenido de frecuencias de cada clase en particular de texturas.

Otros métodos han propuesto descriptores de textura invariantes a la rotación, a la escala e incluso a cambios de perspectiva [49, 79, 84]. El método presentado en [79] supone que un píxel es condicional solamente de su vecindario y que es independiente del resto de la imagen. Los descriptores del vecindario están basados en los MRF y éstos son agrupados usando el algoritmo de *k*-medias, y en consecuencia generan un conjunto representativo por cada clase de textura llamados “textones”. El centro del grupo o textón es almacenado en un diccionario para todas las muestras dentro del conjunto de datos. Para construir una representación de una textura, el diccionario es usado para etiquetar las imágenes. Los métodos presentados en [49, 84] usan características extraídas desde un conjunto disperso de puntos clave, regiones o “parches”, obtenidos al calcular las estructuras más sobresalientes de la imagen. Estas estructuras son calculadas usando detectores de regiones locales como los detectores Harris y Laplace, combinados con descriptores invariantes a la rotación como “spin” y “RIFT”. Las pruebas realizadas con los métodos descritos anteriormente muestran una gran precisión en experimentos de recuperación y clasificación empleando bases de datos de texturas complejas que constituyen un reto para las evaluaciones de los métodos de análisis de texturas, como la base de datos de imágenes UIUC [49]. Sin embargo, estos métodos demandan mucho tiempo computacional debido al cálculo de parámetros libres como el tamaño de los parches o regiones, el cálculo de histogramas (por ejemplo, histogramas de dimensión 32 o 100 por región en cada imagen, [49]) y el algoritmo de agrupamiento para generar una firma representada

por el número de grupos (“clusters”) y sus pesos.

Dentro de los modelos inspirados en el sistema de visión humano que simulan algunas de las etapas primarias de visión y que son capaces de extraer información visual de las imágenes podemos mencionar a las *wavelets* que emplean filtros Gabor [19] y a la Transformada de Hermite [57, 82].

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. Las técnicas propuestas para obtener características invariantes a la rotación pueden incluir un modelo espacio-frecuencia completo y el cálculo de micro y macro características [37]. Básicamente, la *invarianza* a la rotación se logra a partir del modelo de micro características y la auto-correlación de éstas; las micro características son obtenidas a partir de la transformada discreta de Fourier (DFT) de las texturas procesadas previamente con los filtros Gabor. Este método es complejo en términos computacionales. Otras técnicas involucran etapas de post-procesamiento como corrimientos circulares de los elementos que conforman al vector de características de acuerdo a una orientación dominante [63, 72] y simples sumas de los diferentes coeficientes direccionales [38] para obtener un solo coeficiente invariante a la rotación en cada escala de análisis.

La transformada de Hermite cartesiana es una técnica de descomposición local en la cual la imagen es localizada a través de una ventana Gaussiana y expandida en polinomios ortogonales con respecto a dicha ventana [57]. Una de las ventajas al usar la transformada de Hermite sobre otros métodos basados en *wavelets* es que las funciones de análisis son similares a las derivadas de Gaussianas (en semejanza con las wavelets Gabor). La evidencia psicovisual sugiere que las derivadas de Gaussianas se ajustan a los perfiles de los campos receptivos de los sistemas de visión de mamíferos [82]. De esta manera, las funciones de análisis de Hermite son capaces de extraer detalles visuales de las imágenes.

Algunos investigadores han encontrado que la transformada de Hermite orientada [58, 76, 78] es una manera eficiente de describir en forma compacta las características de una imagen con un número menor de coeficientes que la transformada de Hermite cartesiana. Con este enfoque, es posible obtener casi las mismas características de la imagen sin importar la orientación de ésta debido a que la *orientabilidad* puede hacerse dependiente de la dirección de máxima energía. A través de la transformada de Hermite orientada, los detalles visuales de la imagen pueden encontrarse a muchas orientaciones para ser localmente descritas como patrones unidimensionales. Esto representa una ventaja sobre otros diseños de bancos de filtros, los cuales requieren establecer un número fijo de orientaciones para el análisis [7].

Otros autores han utilizado el modelo de Hermite tipo Gabor [75] para extraer características de texturas y han sido probados en experimentos de eliminación de ruido e indexado [7, 25]. La transformada de Hermite orientada ha sido usada y se han obtenido resultados satisfactorios para la recuperación de imágenes [27]. Sin embargo, estos trabajos carecen de la evaluación experimental de los desempeños de la clasificación para distintos tipos de características, métodos de reducción de la dimensión del vector y de distancias que mejor clasifican las texturas.

1.2. Objetivo

Aumentar la discriminación entre texturas pertenecientes a clases distintas y minimizar la variabilidad entre las texturas que pertenecen a una misma clase a través de la fusión en un código compacto de patrones extraídos usando técnicas de procesamiento de imágenes inspiradas en sistemas de visión mamíferos.

1.2.1. Objetivos particulares

- Obtención de las características de la textura más significativas explotando las propiedades del análisis a distintas escalas y orientaciones que proporcionan las técnicas inspiradas en el sistema de visión de la transformada de Hermite y Gabor.
- Generación de un vector compacto que represente las características de la textura.
- Evaluación del desempeño de la clasificación de texturas con relación al traslape de las funciones de análisis en el dominio frecuencial.
- Selección de una métrica de discriminación acorde con el vector de características.

1.3. Contribución de la tesis

La metodología presentada contribuye al análisis de texturas a través de dos modelos con los cuales se obtiene una representación compacta de las características de la textura y que además es invariante a la rotación de dichas texturas.

Uno de los métodos presentados hace uso de la propiedad de la *orientabilidad* de los coeficientes obtenidos a partir de una descomposición de la imagen. Para el otro método es necesario definir con antelación el número de coeficientes orientados que se calcularán.

Dependiendo del uso que se le quiera dar o la aplicación, la metodología presentada usando el método de filtrado basado en la transformada de Hermite orientada permite explotar la detección de la orientación de la textura de forma más directa que utilizando un banco de filtros con orientaciones predefinidas.

Así mismo, la metodología usando la transformada de Hermite orientada permite seleccionar durante la etapa de entrenamiento a los filtros más apropiados que contribuirán a obtener buenas tasas de clasificación. Esta selección de filtros permite disminuir la carga computacional durante la etapa de prueba o verificación, en donde el número de imágenes a procesar es mayor.

Los experimentos y análisis presentados en esta tesis constituyen un indicio y una guía para seleccionar o descartar los tipos de características de texturas que durante mucho tiempo han sido utilizadas de forma discrecional. Los resultados presentados también forman un complemento de los hallazgos reportados en los trabajos [27, 28, 29, 30].

La textura y la metodología presentada en este trabajo es de utilidad para formar características de bajo nivel de la imagen. Esta información se puede conjuntar con otros tipos de características y metodologías para caracterizar las imágenes. Algunas aplicaciones que

involucran el análisis de textura y en donde es pertinente incursionar debido a los resultados presentados en esta tesis se enumeran a continuación: la búsqueda y recuperación de imágenes por contexto, clasificación de texturas en ambientes industriales como en la industria del tratamiento de aguas residuales, en la industria textil y en la industria automotriz para la detección de áreas con fallos, análisis de defectos en vías de comunicación como en carreteras o vías de ferrocarril.

1.4. Estructura de la tesis

En el Capítulo 2 se presenta la metodología que incluye la justificación, los materiales y el resumen de los métodos empleados así como la metodología propuesta para el análisis de texturas con métodos de selección y reducción de dimensión. Algunas de las técnicas que emplean filtros, una revisión de los tipos de características, distancias entre los vectores de características y de las estrategias para lograr invarianza a la rotación que han sido utilizadas con mayor frecuencia por otros autores se presentan en el Capítulo 3. El Capítulo 4 corresponde al tema del procesamiento de las imágenes de texturas utilizando filtros que han sido diseñados basándose en elementos de los sistemas de visión biológicos: transformada de Hermite y filtros Gabor. En el Capítulo 5 se aborda un método para medir la separación entre las clases de texturas y se presentan también los métodos de reducción de dimensión y de selección de características y clasificación. Los resultados de los experimentos se reportan en el Capítulo 6. Finalmente las conclusiones y perspectivas se presentan en el Capítulo 7.

2. Metodología

2.1. Justificación

El sistema de visión humano es capaz de detectar y discriminar detalles a distintas distancias de observación como bordes y texturas. Los modelos computacionales asociados a los sistemas de visión biológicos permiten obtener información visual de las imágenes digitales emulando las funciones primitivas de dichos sistemas. Esta información de bordes y texturas puede ayudar en la solución de problemas generales como la segmentación, clasificación, navegación y búsqueda de imágenes en bases de datos, reconocimiento de objetos o de personas empleando identificadores biométricos. La utilidad de poder discriminar texturas a través de algoritmos computacionales es amplia. La implementación real de un sistema de visión computacional que caracterice texturas de forma precisa puede ser determinante para su éxito. El conocimiento de las características, las métricas y la relación del traslape de la respuesta en frecuencia de las funciones de análisis es necesario para poder plantear una caracterización de texturas precisa.

2.2. Materiales y resumen de los métodos empleados

- Bases de datos de imágenes de textura Brodatz [8], VisTex [80], Outex [66, 67].
- Generación aleatoria de conjuntos disjuntos de entrenamiento y de clasificación y repetición de experimentos.
- Análisis de texturas empleando la transformada de Hermite y las wavelets Gabor.
- Uso y propuesta de métodos de reducción de dimensión y separación de clases.
- Determinación de la separación entre clases de texturas usando el criterio de Fisher.
- Determinación de los tipos de características más apropiados para el análisis de texturas, selección y reducción de dimensión del vector de características y selección de las medidas de distancia más apropiadas usando análisis estadísticos como ANOVAs y prueba de Tukey.
- Determinación de la sensibilidad usando curvas de característica de operación de receptor (ROC).

2.3. Análisis de texturas con métodos de selección y reducción de dimensión

El esquema de la propuesta general para el análisis de texturas se presenta en la Fig. 2.1. Debido a que el uso de la transformada de Hermite en el análisis de texturas ha sido menos difundido en la literatura, se ha optado por utilizar ésta en primer lugar y después realizar las comparaciones necesarias con el modelo de wavelets Gabor. Hay que destacar que en esta metodología la técnica de análisis a través de la transformada de Hermite puede ser reemplazada por cualquier otra.

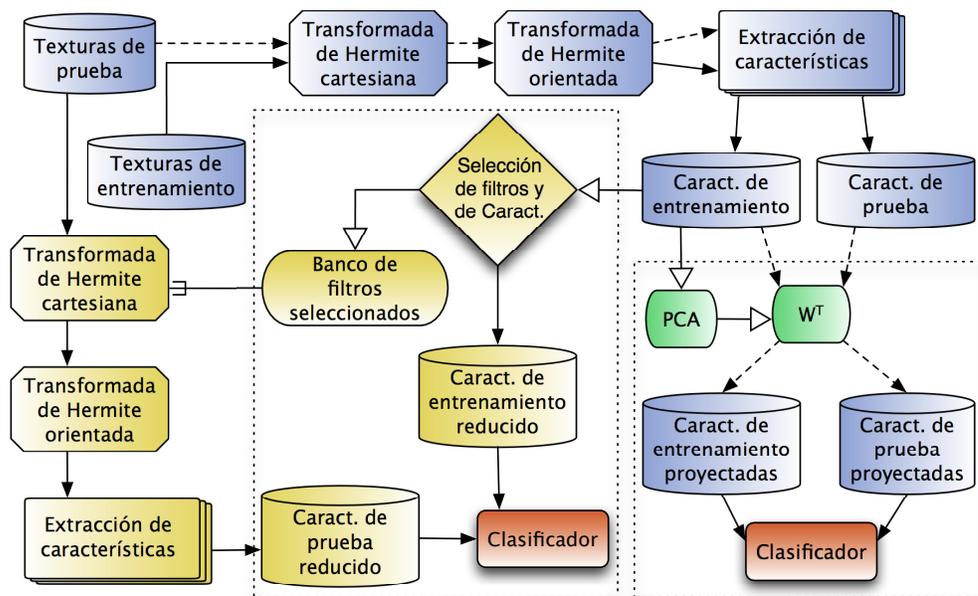


Figura 2.1.: Metodología del análisis de texturas usando la transformada de Hermite orientada y reducción de dimensión a través de la selección de características.

En este esquema se propone una primera etapa de selección de filtros y de las características que ofrecen mejor poder de discriminación. La selección de los filtros se realiza a través de los métodos de reducción de dimensión que se presentan en el Cap. 5.

Para efectos de comparación y validación de esta propuesta se incluye el método de análisis por componentes principales (PCA) para reducción de dimensión¹ (véase Cap. 5). La hipótesis es de que los resultados de la clasificación empleando otros métodos de reducción de dimensión deben ser por lo menos iguales a los obtenidos a través de PCA. Con estos métodos se busca aumentar la discriminación entre texturas pertenecientes a clases distintas y minimizar la variabilidad entre las texturas que pertenecen a una misma clase a la vez que se usa un

¹Una desventaja de PCA es que no se pueden determinar las fuentes originales que ayudan a la *ortogonalización* del espacio ya que cada componente principal es una combinación lineal de todas las fuentes originales, Véase Cap. 5

conjunto pequeño de coeficientes de la transformada de Hermite orientada o de cualquier otra técnica durante la etapa de prueba.

3. Análisis de texturas con métodos de filtrado

En este capítulo se presentan algunas de las técnicas que emplean filtros más comúnmente utilizadas. También se presenta una revisión del tipo de características, distancias y las estrategias para lograr invarianza a la rotación que han sido utilizadas por diversos autores.

El objetivo de los métodos de análisis de texturas basados en filtros o transformadas es representar una imagen en otro dominio el cual está de alguna manera relacionado con las características de la textura como selectividad en frecuencia y orientación. Dentro de los métodos basados en filtros se pueden enumerar aquellos que emplean a la transformada de Fourier y a las que utilizan algún tipo de transformada wavelet.

Los métodos de análisis de texturas utilizando filtros tienen su origen en los métodos que utilizan únicamente la transformada de Fourier para obtener características de la textura como la magnitud y la fase. La desventaja que se observa al utilizar éstas características es la falta de localización espacial y el inconveniente de obtener una fase indefinida cuando en alguna posición de la imagen su valor es cero [69].

La transformada wavelet provee información espacial y frecuencial de las texturas que es utilizada en las tareas de clasificación y segmentación. Sin embargo, las wavelets no son invariantes a la traslación y a la rotación. Diversas propuestas han sido publicadas para realizar análisis de texturas invariante a la rotación aplicando la transformada de Radón al inicio de la descomposición wavelet 1D [40] (También conocida como Ridgelet [11, 68]) y aproximadamente invariante a la traslación usando la transformada wavelet compleja de doble árbol (DT-CWT) [13]. Otros autores optan por una transformación log-polar para eliminar los efectos de la rotación y la escala [71]. Algunos autores utilizan la transformada wavelet discreta 2D (2D-DWT), árboles ocultos de Markov y corrimientos sobre el mapa o vector de características para lograr la invarianza a la rotación [53, 72] y a través de la rotación de los filtros 2D de una transformada wavelet discreta [44, 45]. Una transformada wavelet que ha sido diseñada con el propósito de obtener una representación invariante a la rotación es la llamada *shearlet*. Esta transformada es una versión multidimensional de la transformada wavelet tradicional y está específicamente diseñada para tratar la información anisotrópica y direccional a varias escalas de análisis utilizando sistemas afines con dilataciones *compuestas* [24, 47].

En cuanto a los métodos basados en filtros el más común que se encuentra en la literatura del análisis de texturas se basa en los filtros Gabor. Estos filtros tuvieron mucha popularidad a partir de los trabajos presentados simultáneamente en 1980 por Marcelja para una dimensión [56] y por Daugman para dos dimensiones [18]. Los autores que utilizan estos filtros proponen diversas metodologías para el cálculo de los mismos así como algunas combinaciones de

métodos basados en modelos. Cada autor propone una metodología para el desarrollo de sus filtros basado en alguna propiedad de éstos [5, 6, 15, 38] normalmente buscando cubrir la mayor parte del espectro frecuencial con combinaciones de frecuencia central, anchos de banda radiales y angulares. También se observan aquellos métodos que incorporan alguna etapa de procesamiento posterior sobre los coeficientes, entre los que podemos mencionar: obtención de la magnitud de las texturas filtradas con los filtros Gabor [6] y además con suavizado Gaussiano [6, 41], usando una función sigmoide [41] o bien usando únicamente la componente real la función Gabor [54].

3.1. Características estadísticas de la textura

La mayoría de los autores que presentan sus propuestas para el análisis de texturas empleando métodos basados en filtrado utilizan características globales de la texturas bajo análisis. Estas características globales son obtenidas a partir de las estadísticas de los píxeles de la textura filtrada. No existe un soporte más que el empírico o bien como una forma de comparación entre los distintos métodos existentes, que explique el origen o la razón por la cual usar uno u otro tipo de característica. Los tipos de características comúnmente reportadas en la literatura del análisis de texturas se enumeran a continuación.

Sea $f(h, w) \in \mathbb{R}$ una matriz que contiene a los valores de una imagen cualquiera después de haber sido filtrada de dimensiones H y W donde w y h representan los índices de cada píxel. Entonces, la característica de textura conocida como media [3, 26, 68] está definida como:

$$\mu = \frac{1}{H \times W} \sum_{h=1}^H \sum_{w=1}^W f(h, w) . \quad (3.1)$$

El tipo de característica de textura llamada desviación estándar (con sesgo) está definida como [3, 26, 44, 45, 64, 68, 71]:

$$\sigma = \sqrt{\frac{1}{(H \times W - 1)} \sum_{h=1}^H \sum_{w=1}^W [f(h, w) - \mu]^2} \quad (3.2)$$

y una variante de la desviación estándar comúnmente encontrada en la literatura del análisis de texturas [5, 38, 52, 53, 63] está definida como:

$$\sigma_b = \sqrt{\frac{1}{(H \times W - 1)} \sum_{h=1}^H \sum_{w=1}^W [|f(h, w)| - \mu]^2} . \quad (3.3)$$

Otros autores han propuesto el uso de energías que son calculadas a partir de la norma- p de la imagen. Estos tipos de características de texturas están definidas como el cuadrado de la

norma l_2 :

$$E_0 = \sum_{h=1}^H \sum_{w=1}^W [f(h, w)]^2 \quad (3.4)$$

y una variante de la ecuación (3.4) definida como [40, 64, 71]:

$$E_1 = \frac{E_0}{H \times W} \quad (3.5)$$

o bien la media absoluta calculada a partir de la norma l_1 [5, 38, 40, 44, 45, 52, 53, 54, 63, 64, 68, 71]:

$$E_2 = \frac{1}{H \times W} \sum_{h=1}^H \sum_{w=1}^W |f(h, w)| \quad (3.6)$$

mientras que algunos autores [40] definen a la norma $l_{\frac{1}{2}}$ como:

$$E_3 = \frac{1}{H \times W} \sum_{h=1}^H \sum_{w=1}^W \sqrt{|f(h, w)|} \quad (3.7)$$

Otros autores utilizan como características de las texturas otras formas de energía. El residuo promedio (varianza) lo definen como [64][71]:

$$E_4 = \sum_{h=1}^H \sum_{w=1}^W [f(h, w) - \mu]^2 \quad (3.8)$$

La forma de energía basada en la entropía, está definida a partir del cuadrado de las probabilidades de los valores absolutos de los píxeles de la imagen:

$$E_5 = - \sum_{i=0}^{B-1} |p(i)|^2 \log |p(i)|^2 \quad (3.9)$$

donde la probabilidad de encontrar píxeles con valores representados por la categoría i está definida como:

$$p(i) = \frac{h(i)}{H \times W} \quad ; \quad (3.10)$$

la idea es primero cuantificar los valores de los píxeles con B niveles de cuantificación y después, para cada nivel de cuantificación se obtiene el número de píxeles que caen dentro de la categoría i , de tal manera que $h(i)$ está definido como:

$$h(i) = \sum_{h=1}^H \sum_{w=1}^W \epsilon_i(h, w) \quad (3.11)$$

donde

$$\epsilon_i(h, w) = \begin{cases} 1 & \text{si } f(h, w) \in i \\ 0 & \text{en otro caso.} \end{cases} \quad (3.12)$$

En otras palabras, se genera un histograma h donde $h(i)$ representa el número de píxeles en cada nivel de cuantificación $0 \leq i \leq B - 1$ (las categorías se representan con el índice i).

3.2. Vector de Características

Un vector de características es un miembro de un espacio n -dimensional formado a partir de las características de una textura para representar a dicha textura de una forma más compacta. Lo más utilizado en la literatura del análisis de texturas para la generación del vector de características es la concatenación de estas características. En el caso del análisis de texturas donde se emplean los métodos basados en filtrado, se calculan los distintos tipos de características (como los que fueron presentados en la Sec. 3.1) a partir de la descomposición de la imagen original en un número de imágenes de textura filtradas (*sub-imágenes* calculadas a partir de la imagen original y un banco de filtros). Además, este número de imágenes de textura filtradas frecuentemente determinará el tamaño del vector de características.

Muchos autores [5, 38, 44, 45, 52, 53, 64, 68] concatenan distintos tipos de características en un solo vector. Por ejemplo:

$$\mathbf{v} = \{\mu_1, \sigma_1, \mu_2, \sigma_2 \dots, \mu_r, \sigma_r\} \quad (3.13)$$

donde r es el número de imágenes de textura filtradas o *sub-imágenes* a las cuales se le extraen las características μ y σ .

Al mezclar distintos tipos de características hay que tener en cuenta que no será posible conocer que tipo de característica es la que produce una mejor tasa de clasificación además de que la elección de una medida de distancia entre las clases será crucial para obtener buenas tasas de clasificación. Al carecer de un conocimiento de las características y distancias que producen buenas tasas de clasificación podría significar una desventaja de está práctica comúnmente encontrada en la literatura.

Existen también vectores de características y técnicas para lograr que esta representación compacta sea invariante a la rotación de la textura. Como se verá en la Sec. 3.4 es posible calcular las características de una textura a partir de los coeficientes direccionales de la textura original. Algunas técnicas de filtrado por si mismas están definidas como invariantes a la rotación de la textura mientras que con otras técnicas será necesario realizar corrimientos circulares sobre los elementos del vector de características para obtener una representación invariante a la rotación de la textura.

3.3. Distancias

La elección de la medida de distancia depende del método utilizado para representar las características de las texturas. En el caso de las texturas representadas como histogramas

la métrica más común para comparaciones entre estos histogramas es a través de la prueba estadística χ^2 [17, 48] aunque otros autores han usado la medida de similitud logarítmica [66]:

$$L(S, M) = \sum_{b=1}^B S_b \log M_b \quad (3.14)$$

donde S es la muestra de prueba, M es el modelo que maximiza la medida de similitud logarítmica, B es el número de frecuencias, S_b y M_b corresponden a la muestra y a la probabilidad del modelo a la frecuencia b respectivamente. La muestra S es asignada a la clase del modelo M que maximiza la medida de similitud logarítmica.

La distancia utilizada con mayor frecuencia en la literatura del análisis de texturas basado en los métodos de filtrado es la distancia Euclidiana [40, 45]. Sean $\mathbf{x} = \{x_1, x_2, \dots, x_r\}$ y $\mathbf{y} = \{y_1, y_2, \dots, y_r\}$ dos vectores de dimensión r , es decir el número total de características en ambos vectores es r , entonces la distancia Euclidiana está definida como:

$$d_E(\mathbf{x}, \mathbf{y}) = \sqrt{\sum_{j=1}^r (x_j - y_j)^2} . \quad (3.15)$$

Otra distancia menos utilizada pero también reportada en la literatura del análisis de texturas es la distancia Canberra [45], está definida como:

$$d_C(\mathbf{x}, \mathbf{y}) = \sum_{j=1}^r \frac{|x_j - y_j|}{|x_j| + |y_j|} . \quad (3.16)$$

De manera similar, otra forma de distancia utilizada en el análisis de texturas es la distancia de Mahalanobis [71]. Ésta asigna una etiqueta de clase a un vector muestra al seleccionar la distancia mínima de todas las distancias calculadas entre el vector medio o *centroide* de cada clase. Sea \mathbf{C} el conjunto de vectores de la clase \mathbf{C} tal que $\mathbf{C} = \{\mathbf{c}_1, \mathbf{c}_2, \dots, \mathbf{c}_q\}$ y sea $\bar{\mathbf{c}}$ el vector promedio o *centroide* de la clase \mathbf{C} . Entonces la distancia de Mahalanobis de un vector muestra \mathbf{x} a una clase \mathbf{C} está definida como:

$$d_M(\mathbf{x}, \mathbf{C}) = \sqrt{(\mathbf{x} - \bar{\mathbf{c}})\Sigma_{\mathbf{C}}^{-1}(\mathbf{x} - \bar{\mathbf{c}})^T} , \quad (3.17)$$

donde $\Sigma_{\mathbf{C}}^{-1}$ denota a la inversa de la matriz de covarianza de la clase \mathbf{C} . Si $\Sigma_{\mathbf{C}}^{-1}$ existe entonces es posible calcular la distancia de Mahalanobis.

Algunos autores han utilizado además una forma de distancia llamada distancia Euclidiana normalizada [38, 44, 45, 52, 63]. Consiste en normalizar la diferencia entre los vectores conocidos que pertenecen a una clase y el vector muestra. Sea \mathbf{x} un vector muestra con elementos $\mathbf{x} = \{x_1, x_2, \dots, x_r\}$. Sea \mathbf{C} el conjunto de vectores de la clase \mathbf{C} tal que $\mathbf{C} = \{\mathbf{c}_1, \mathbf{c}_2, \dots, \mathbf{c}_q\}$. Sea $\mathbf{c}_i \in \mathbf{C}$ el i -ésimo vector cuyos elementos son $\mathbf{c}_i = \{c_{i1}, c_{i2}, \dots, c_{ir}\}$. Entonces la distancia

Euclidiana normalizada entre el vector muestra \mathbf{x} y \mathbf{c}_i está definido como:

$$d_{E_N}(\mathbf{x}, \mathbf{c}_i) = \sum_{j=1}^r \left| \frac{x_j - c_{ij}}{\alpha(\mathbf{C})} \right|, \quad (3.18)$$

donde $\alpha(\mathbf{C})$ es la desviación estándar de la clase C .

3.4. Invarianza a la rotación

Uno de los problemas de gran importancia en el análisis de texturas es lograr obtener una descripción que sea invariante a la orientación en que se presenta la textura. Los primeros métodos que consideraron la importancia de la *invarianza* a la rotación usaron “polarogramas” [22]. Los métodos basados en modelos propusieron un “modelo auto-regresivo de simetría circular” [43] y campos aleatorios de Markov Gaussianos (GMRF) [16] o combinaciones de filtros espejo en cuadratura en combinación con un modelo oculto de Markov (HMM) [12]. Los métodos basados en filtrado han utilizado distintas técnicas para lograr la invarianza a la rotación, algunas de ellas se mencionan a continuación.

Sean $f(h, w)$ una imagen cualquiera y $g(h, w)_{k,s}$ un filtro direccional donde $k = 1, \dots, K$ y $s = 1, \dots, S$ son los índices de la orientación y de la escala del filtro respectivamente. Un método propuesto para obtener una representación invariante a la rotación se basa en sumar todos los filtros direccionales [38] en cada escala de análisis s de tal manera que

$$g(h, w)_s = \sum_{k=1}^K g(h, w)_{k,s} \quad (3.19)$$

es un filtro que extrae características a una escala específica cubriendo todas las orientaciones K . Posteriormente, la imagen original es filtrada a través de la convolución $f \star g$ para obtener la representación invariante a la orientación de la textura.

Otro de los métodos propuestos que logran *invarianza* a la orientación de la textura se basa en promediar todos los coeficientes direccionales¹ [64] para cada escala de análisis s . Sea $F(h, w)_{k,s}$ la imagen filtrada en la dirección k y a la escala s , calculada a través de la convolución $f \star g$:

$$F(h, w)_{k,s} = f(h, w) \star g(h, w)_{k,s} . \quad (3.20)$$

Entonces la representación invariante a la orientación de la textura queda definida a través de un coeficiente que resulta de promediar todos los coeficientes direccionales:

$$F(h, w)_s = \frac{1}{K} \sum_{k=1}^K F(h, w)_{k,s} \quad (3.21)$$

o bien a través del cálculo de la norma l_2 de los coeficientes direccionales [3] [26] en cada

¹El término coeficiente direccional se refiere a una imagen filtrada con un filtro a cierta orientación.

escala de análisis s :

$$F(h, w)_s = \left[\sum_{k=1}^K F(h, w)_{k,s}^2 \right]^{1/2} . \quad (3.22)$$

Otros autores proponen obtener descriptores de texturas invariantes a la rotación manipulando el vector de características de textura (como las presentadas en la Sec. 3.1). Sea $\mathbf{X} = \{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_K\}$ un conjunto de vectores de características obtenidos de los K coeficientes direccionales. El vector de características \mathbf{x}_k es generado a partir de la concatenación de todas las características extraídas del coeficientes direccional k . Entonces, el vector de características invariante a la rotación \mathbf{y} es calculado como:

$$\mathbf{y} = \sum_{k=1}^K \mathbf{x}_k . \quad (3.23)$$

Los autores [54] y [68] proponen calcular la magnitud de la DFT unidimensional de un vector generado a partir de las características obtenidas de los coeficientes direccionales. Se fundamenta en la propiedad de invarianza a la traslación de la DFT. Las distintas configuraciones posibles de posicionamiento de los elementos que conforman al vector de características corresponden a versiones trasladadas de un mismo vector de características. Las características trasladadas cíclicamente dentro del vector introducen únicamente un cambio de fase en la transformada de Fourier por lo que la magnitud de la DFT del vector de características permanece constante. Sea $\mathbf{x} = \{x_1, x_2, \dots, x_K\}$ un vector de características tal que una característica x_k se extrae del coeficiente direccional k , de tal manera que el vector \mathbf{x} se forma concatenando una sola característica a la vez extraída del coeficiente direccional k . Entonces una representación invariante a la rotación \mathbf{Y} del vector de características \mathbf{x} se obtiene a través de la magnitud de la transformada de Fourier del vector \mathbf{x} :

$$\mathbf{Y} = |\mathcal{F}\{\mathbf{x}\}| . \quad (3.24)$$

El vector \mathbf{Y} que resulta es invariante a la rotación y además tiene dimensión:

$$\hat{K} = \lfloor K/2 \rfloor + 1 . \quad (3.25)$$

Otros autores también han propuesto ordenar todas las características en orden descendiente y elegir los k primeros filtros direccionales más significativos [71] o bien realizando desplazamientos circulares de tal manera que los primeros elementos del vector de características para cada escala corresponda a la orientación dominante [63]. Otros métodos logran obtener representaciones invariantes al cambiar el espacio original de la imagen hacia otro espacio a través de una transformada polar [71] o bien una transformada de Radón [40]. Así también se pueden encontrar métodos que logran invarianza a la orientación de la textura a través de la rotación de los coeficientes de los filtros wavelets [45] o bien usando la transformada Ridgelet [68].

3.5. Discusión

Los tipos de características descritos en la Sec. 3.1 han sido utilizadas de manera arbitraria y se han llegado a comparar los resultados de las tasas de clasificación empleando además distintas metodologías para la obtención de dichas características. En este sentido, es posible que las comparaciones no hayan sido equitativas. Por esta razón, en este trabajo se propuso considerar un análisis de cada una de las características por separado y además formar el vector de características exclusivo para un solo tipo de características. La idea es obtener de forma experimental un escenario de tipos de características y escoger aquellos tipos de características con los cuales sería posible obtener una tasa de clasificación buena.

Respecto al tema de las distancias su uso también ha sido arbitrario. Para corroborar el desempeño en la clasificación de las distancias también se optó por realizar experimentos para cada uno de los tipos de distancia d_E , d_C , d_M y d_{E_N} definidas en las ecuaciones (3.15), (3.16), (3.17) y (3.18).

Lograr obtener representaciones de la textura invariantes a la orientación de ésta es aún un tema del estado del arte. En el caso de los métodos basados en filtros se han propuesto diversas formas de lograr representaciones invariantes como los que se presentaron en la Sec. 3.4. En este trabajo para el caso del análisis de las texturas utilizando los filtros Gabor se optó por obtener un vector bastante compacto utilizando el método presentado en [54, 68] y que se resume en la Ec. (3.24). En el caso del análisis de texturas utilizando la transformada de Hermite existe una versión de esta transformada que es invariante a la orientación y se aborda en la Sec. 4.1.1

4. Procesamiento de texturas inspirado en sistemas de visión biológicos

En este capítulo se presentan dos técnicas de procesamiento de imágenes que están inspiradas en sistemas de visión biológicos y que fueron utilizadas para el desarrollo de esta tesis.

Existen varios modelos de visión que han sido propuestos para el procesamiento de la información visual. Éstos 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 al nivel de la retina como de la 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 para una mejor comprensión de la visión humana [83].

Los campos receptivos son mapas de las localidades en el espacio y tiempo donde la intensidad de la luz puede afectar la respuesta 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 visual, 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 pre-procesamiento, 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 [83].

Algunos de los primeros experimentos realizados sobre células simples usaron estímulos unidimensionales como bordes extendidos o barras [39]. 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 [39] 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.

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 [20].

Otro modelo menos conocido y que puede ser utilizado en la discriminación de texturas se basa en la transformada de Hermite. La transformada de Hermite, originalmente introducida por Martens [57] 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 de 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 [82] 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].

4.1. Transformada de Hermite

Para el caso unidimensional, una transformada polinomial $L_n(x)$ es una descomposición local en la cual una señal de entrada $L(x)$ es localizada a través de una ventana $V(x)$ y después proyectada en polinomios ortogonales $G_n(x)$ en cada posición de la ventana [57]:

$$L_n(x_0) = \int_x L(x)G_n(x_0 - x)V_n^2(x_0 - x)dx . \quad (4.1)$$

La transformada de Hermite surge cuando G_n corresponde a los polinomios de Hermite $H_n(x)$ dados por la fórmula de Rodrigues:

$$H_n(x) = (-1)^n e^{x^2} \frac{d^n e^{-x^2}}{dx^n}, \quad n = 0, 1, 2, \dots \quad (4.2)$$

y la ventana ortogonal corresponde a una ventana similar a la función Gaussiana (excepto por un factor de escala, véase apéndices B, C y D):

$$V(x) = \frac{1}{\sqrt{\sqrt{\pi}\sigma}} \cdot e^{-x^2/2\sigma^2} . \quad (4.3)$$

Siguiendo la Eq. (4.1), la expansión de los coeficientes $L_n(x)$ se puede derivar de la convolución de la señal $L(x)$ con las funciones de análisis de Hermite $d_n(x)$, las cuales están dadas en

términos de la función ventana y los polinomios de Hermite como:

$$d_n(x) = \frac{(-1)^n}{\sqrt{2^n n!}} \cdot \frac{1}{\sigma \sqrt{\pi}} H_n \left(\frac{x}{\sigma} \right) e^{-x^2/\sigma^2} . \quad (4.4)$$

Las funciones de análisis de Hermite unidimensionales se pueden generalizar al caso bidimensional fácilmente, ya que éstas poseen la propiedad de ser tanto espacialmente separables como rotacionalmente simétricas. Podemos escribir entonces las funciones de análisis de Hermite en dos dimensiones como:

$$d_{n-m,m}(x, y) = d_{n-m}(x) d_m(y) \quad (4.5)$$

donde $n - m$ y m denotan el orden del análisis en las direcciones x y y respectivamente.

En la figura 4.1 se pueden observar a los filtros de análisis de Hermite desde orden $n = 1$ hasta orden $n = 8$ para una sola dirección de análisis x . Al incrementarse el orden del análisis n se aumenta la frecuencia de análisis del filtro.

La descomposición local en el punto (x_0, y_0) de una imagen $L(x, y)$ puede calcularse usando como base las funciones de análisis de Hermite en dos dimensiones presentadas en la Ec. (4.5) como:

$$L_{n-m,m}(x_0, y_0) = \int_x \int_y L(x, y) \cdot d_{n-m,m}(x_0 - x, y_0 - y) dx dy \quad (4.6)$$

para $n = 0, 1, \dots, \infty$ y $m = 0, \dots, n$. A partir de la Ec. (4.6) podemos generalizar la descomposición de la imagen en su totalidad $L_{n-m,m}(x, y)$ como una convolución¹ \star de la imagen original $L(x, y)$ con los filtros de análisis de Hermite $d_{n-m,m}(x, y)$:

$$L_{n-m,m}(x, y) = L(x, y) \star d_{n-m,m}(x, y) . \quad (4.7)$$

Los coeficientes $L_{n-m,m}(x, y)$ son llamados coeficientes cartesianos de la transformada de Hermite cartesiana porque el orden del análisis $n - m$ y m se realiza en las direcciones x y y respectivamente. El orden máximo de la descomposición de una imagen en sus coeficientes de Hermite cartesianos lo denotamos como:

$$N = n + m . \quad (4.8)$$

La figura 4.2(a) muestra una descomposición de orden máximo $N = 2$ de una imagen en sus coeficientes de Hermite cartesianos. Dado que el orden máximo de la descomposición puede ser definido, las combinaciones $n - m$ y m que cumplan con el orden máximo de la descomposición (en este caso $N = 2$) permite utilizar solo 6 filtros de análisis $d_{n-m,m}(x, y)$ y obtener como resultado 6 coeficientes de Hermite cartesianos.

¹La convolución puede calcularse eficientemente en el dominio frecuencial haciendo uso del teorema de la convolución. Sea $h(x) = (f \star g)(x) = \int_{-\infty}^{\infty} f(z)g(x-z)dz$ la convolución de f con g . El teorema $\mathcal{F}\{f \star g\} = \mathcal{F}\{f\} \cdot \mathcal{F}\{g\}$ permite calcular h de forma eficiente como $h(x) = \mathcal{F}^{-1}\{\mathcal{F}\{f\} \cdot \mathcal{F}\{g\}\}$, donde \mathcal{F} y \mathcal{F}^{-1} son los operadores de la transformada y transformada inversa de Fourier respectivamente.

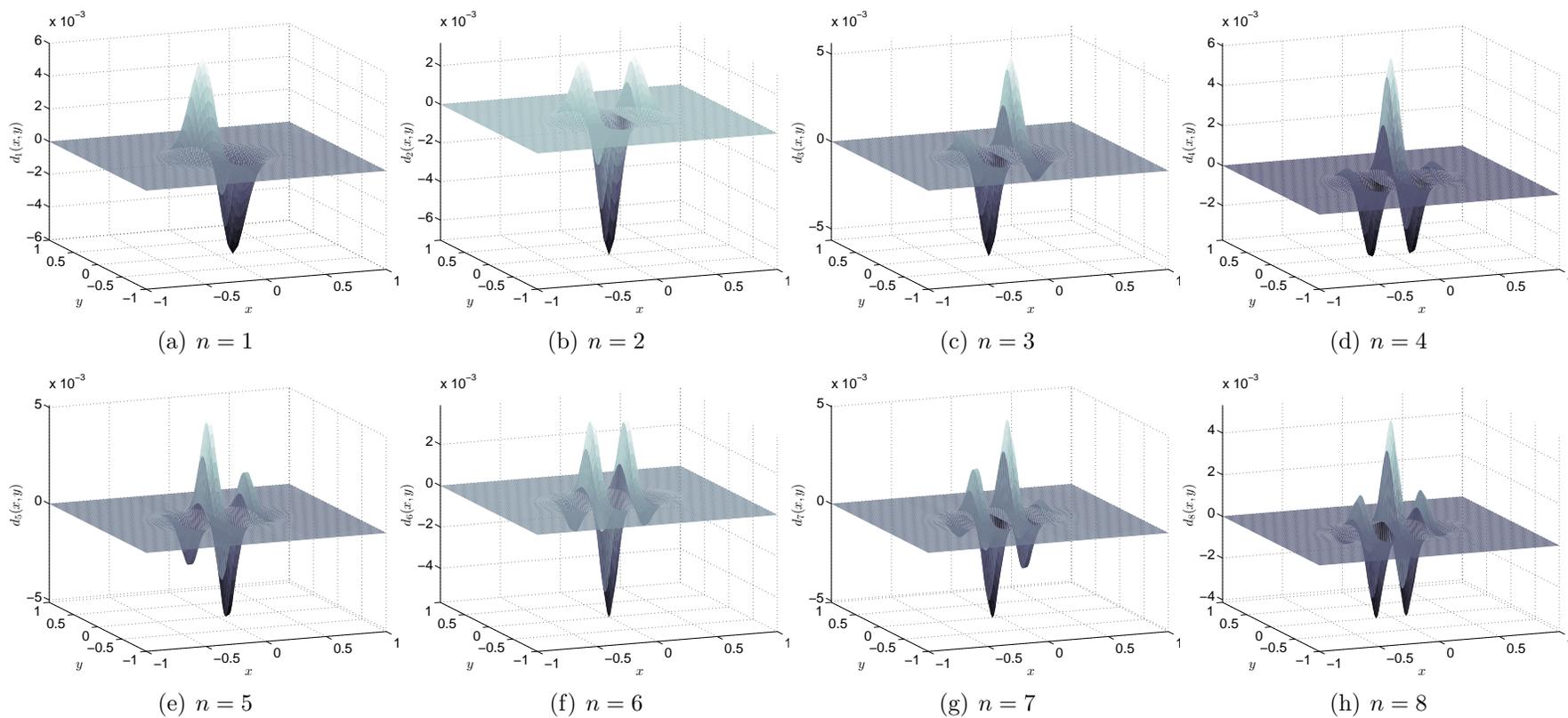


Figura 4.1.: Filtros de análisis de Hermite en el dominio espacial desde orden $n = 1$ hasta orden $n = 8$ en la dirección x . Al incrementarse el orden del análisis n se aumenta la frecuencia de análisis del filtro.

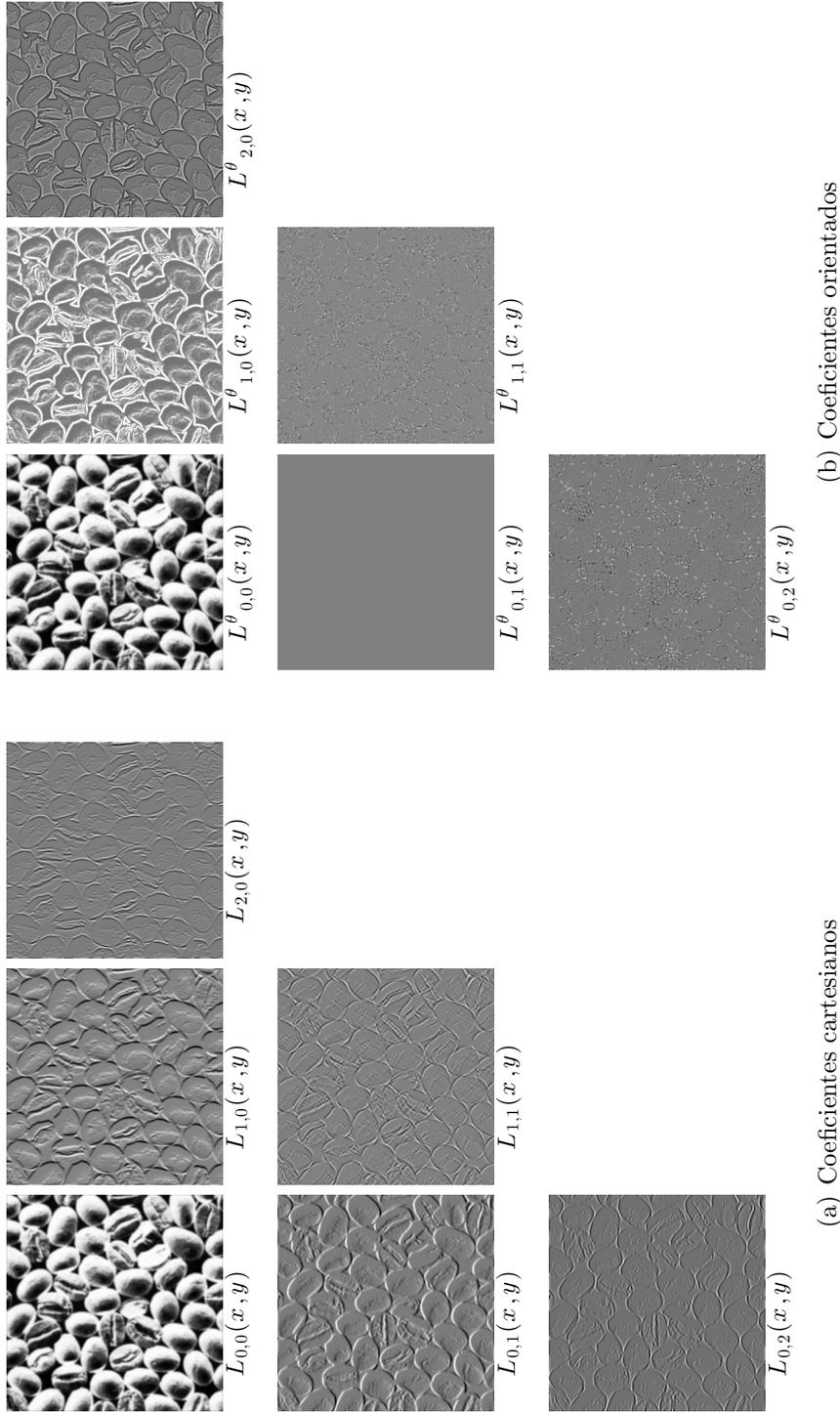


Figura 4.2.: Coeficientes de la transformada de Hermite cartesiana $L_{n-m,m}^\theta(x,y)$ (a) y orientada $L_{n-m,m}^\theta(x,y)$ (b). Los coeficientes $L_{0,0}(x,y)$ y $L_{0,0}^\theta(x,y)$ corresponden a los coeficientes de baja frecuencia (conocido también como promedio de la imagen original) de la descomposición. En ambos casos el orden máximo de la descomposición es $N = 2$. Obsérvese que la información visual distribuida en los coeficientes cartesianos (a) es compactada en los coeficientes orientados (b) $L_{1,0}^\theta(x,y)$ y $L_{2,0}^\theta(x,y)$. La textura corresponde a una región de la textura D74 del álbum de Brodatz [8].

4.1.1. Transformada de Hermite orientada

Un filtro *orientable* está descrito como una clase de filtros en la cual un filtro con una orientación arbitraria es sintetizado como una combinación lineal de un conjunto de *filtros base* [33]. En el caso de las funciones de análisis de Hermite, se pueden construir versiones rotadas de un filtro de orden n al tomar las combinaciones lineales de los filtros cartesianos de orden n . De esta manera, una expresión más general de los coeficientes cartesianos de Hermite $L_{n-m,m}(x, y)$ puede ser escrita en términos de la selección de orientación $\theta(x, y)$ [76] como:

$$L_{n-m,m}^\theta(x, y) = \sum_{k=0}^n L_{n-k,k}(x, y) \alpha_{n-k,k}(x, y) . \quad (4.9)$$

Esta descomposición ha sido llamada la transformada de Hermite *orientable* [78], orientada o rotada. Los términos $\alpha_{n-m,m}(x, y)$ son las funciones angulares cartesianas de orden n las cuales permiten agregar una función de peso de acuerdo a la orientación local $\theta(x, y)$ y del orden n ; están definidas como:

$$\alpha_{n-m,m}(x, y) = \sqrt{C_n^m} \cos^{n-m}(\theta(x, y)) \sin^m(\theta(x, y)) \quad (4.10)$$

donde C_n^m es una función binomial que corresponde a la versión discreta de una función Gaussiana [57]:

$$C_n^m = \frac{n!}{m!(n-m)!} . \quad (4.11)$$

Para el cálculo de $\theta(x, y)$ se han propuesto diversas estrategias [76, 78]. En este trabajo aproximaremos $\theta(x, y) = \arctan(L_{0,1}(x, y)/L_{1,0}(x, y))$.

Los coeficientes de Hermite orientados pueden interpretarse como el resultado de orientar los filtros de análisis de Hermite con un ángulo θ de forma *adaptiva* y posteriormente filtrar la imagen con estos filtros orientados. De manera tal que el análisis sobre una imagen se puede realizar sin necesidad de fijar un número de orientaciones de análisis.

En la figura 4.2(b) se muestra el resultado después de orientar los coeficientes cartesianos de Hermite calculados previamente en la figura 4.2(a)

Considerando una rotación ideal de una imagen dentro de una ventana circularmente simétrica, y asumiendo que no se introducen *artefactos* debido al método de rotación empleado, discretización y cambios en la dirección de la iluminación de la textura, entonces podemos asumir que no hay pérdida de la información original. Si este es el caso, la energía se preserva para cada imagen rotada, entonces podemos escribir la energía local en términos de los coeficientes de la transformada de Hermite orientada como:

$$E_N(x, y) = \sum_{n=0}^N \sum_{m=0}^n [L_{n-m,m}(x, y)]^2 = \sum_{n=0}^N \sum_{m=0}^n [L_{n-m,m}^\theta(x, y)]^2 \quad (4.12)$$

donde $N \geq 0$ es el orden máximo de la descomposición.

En imágenes naturales, muchos de los detalles que son de primer importancia, como bordes y líneas pueden ser localmente descritos como patrones unidimensionales (1D), es decir,

patrones que varían solo en una dirección (y son constantes en la dirección ortogonal). Se pueden distinguir términos de energía local 1D y 2D al separar estos términos a partir de la Ec. (4.12) como:

$$E_N(x, y) = [L_{0,0}(x, y)]^2 + E_N^{1D}(x, y) + E_N^{2D}(x, y) \quad (4.13)$$

donde $L_{0,0}(x, y)$ representa al coeficiente de Hermite de baja frecuencia (i.e., baja energía) o promedio. La energía de los patrones 1D se obtiene a partir de los coeficientes orientados $L_{n,0}^\theta(x, y)$ como:

$$E_N^{1D}(x, y) = \sum_{n=1}^N [L_{n,0}^\theta(x, y)]^2 \quad (4.14)$$

mientras que la energía de los patrones 2D se calcula como:

$$E_N^{2D}(x, y) = \sum_{n=1}^N \sum_{m=1}^n [L_{n-m,m}^\theta(x, y)]^2 . \quad (4.15)$$

Uno de los objetivos al orientar los coeficientes es maximizar la detección de patrones sobre una dirección local θ . De esta manera, [58], [76] y [78] proponen estrategias en las cuales θ es seleccionado de tal forma que E_N^{1D} es maximizada. Como consecuencia se puede lograr una compresión muy eficiente de la energía de los coeficientes de la transformada de Hermite cartesiana.

Las tablas 4.1 y 4.2 muestran los cálculos de los porcentajes de energía de cada uno de los coeficientes mostrados en las figuras 4.2(a) y 4.2(b) respectivamente. Los porcentajes de energía fueron calculados con respecto a la energía total de la descomposición (Ec. (4.12)). Como puede observarse en la tabla 4.1, la energía está distribuida entre todos los coeficientes cartesianos mientras que en la tabla 4.2 la energía ha sido compactada mayormente en los coeficientes orientados $L_{n,0}^\theta(x, y)$.

4.1.2. Filtros de análisis de Hermite en el dominio frecuencial 2D

Las funciones de análisis de la transformada de Hermite en el dominio de Fourier pueden ser escritas como:

$$\hat{d}_{n-m,m}(\omega_x, \omega_y) = \frac{(-j\omega_x\sigma)^{n-m}(-j\omega_y\sigma)^m}{\sqrt{2^n(n-m)!m!}} \cdot \hat{d}_{0,0}(\omega_x, \omega_y) \quad (4.16)$$

donde

$$\hat{d}_{0,0}(\omega_x, \omega_y) = e^{-\sigma^2(\omega_x^2 + \omega_y^2)/4} \quad (4.17)$$

corresponde a una función Gaussiana con parámetro de escala σ . La frecuencia de análisis y el ancho de banda frecuencial están relacionados al orden y la escala de análisis respectivamente. De la Ec. (4.16) podemos ver que el análisis de una imagen a diferentes frecuencias puede realizarse al incrementar el orden de análisis de la función de análisis de Hermite. También

$m \backslash n$	0	1	2
0	87.5856 %	5.1526 %	1.1512 %
1	4.4419 %	0.7404 %	-
2	0.9283 %	-	-

Tabla 4.1.: Porcentajes de energía de los coeficientes cartesianos de la Fig. 4.2(a). La energía está distribuida entre todos los coeficientes cartesianos $L_{n-m,m}(x,y)$. El cálculo de la energía para cada coeficiente individual se calcula con la Ec. (3.4)

$m \backslash n$	0	1	2
0	87.5856 %	9.5944 %	1.8884 %
1	0.0 %	0.4170 %	-
2	0.5145 %	-	-

Tabla 4.2.: Porcentajes de energía de los coeficientes orientados de la Fig. 4.2(b). La energía ha sido compactada mayormente en los coeficientes orientados $L_{n,0}^\theta(x,y)$. El cálculo de la energía para cada coeficiente individual se calcula con la Ec. (3.4)

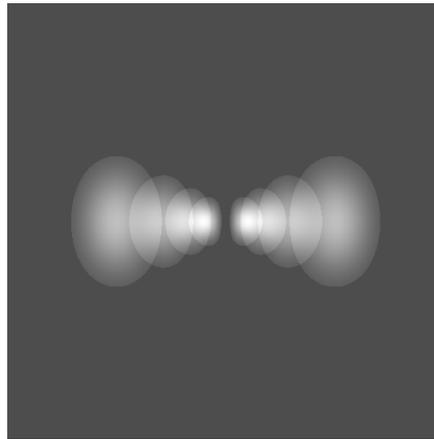


Figura 4.3.: Espacio frecuencial cubierto por un conjunto de funciones de análisis de Hermite en la dirección horizontal de orden $n = 2, 4, 6, 8$ a distintas escalas de análisis.

de la Ec. (4.17) observamos que al aumentar la desviación estándar en el dominio espacial se reduce el espacio cubierto de la función Gaussiana en el dominio frecuencial. De esta manera, el comportamiento de la descomposición de una imagen a través de un conjunto de funciones de análisis de Hermite de orden creciente a distintas escalas, se asemeja al análisis sistemático comúnmente usado en el esquema *wavelet*.

En la Fig. 4.3 se muestra el espacio cubierto en el dominio frecuencial de un conjunto de funciones de análisis de Hermite de orden $n = 2, 4, 6$ y 8 a escalas “diádicas”² de análisis en la dirección horizontal. Nótese que existe cierto grado de traslape entre las diferentes respuestas en frecuencia del banco de filtros de análisis.

²Cada escala es un múltiplo de 2 respecto a la escala que le precede.

4.2. Filtros Gabor

El modelo Gabor para campos receptivos visuales, basado en una descripción matemática de dichos campos como funciones Gabor, fue planteado por primera vez y de manera simultánea en 1980 para una dimensión por Marcelja [56] y para dos dimensiones por Daugman [18]. Básicamente, la función Gabor es el producto de una onda de tipo seno o coseno con una función Gaussiana.

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.

Daugman propuso funciones Gabor en dos dimensiones para modelar los perfiles de los campos receptivos (RFP) de células simples [19]. Las funciones Gabor 2D son filtros espaciales paso-banda que están perfectamente localizadas en los dominios espacial y frecuencial. Son funciones que logran alcanzar el límite teórico de la resolución conjunta de la información en ambos dominios [19].

Una función Gabor 2D consiste de una onda sinusoidal compleja de cierta frecuencia ω_0 y selectividad a la orientación θ , modulada por una función Gaussiana 2D:

$$g(x, y) = \frac{1}{2\pi\sigma_x\sigma_y} e^{-\pi\left(\frac{x_r^2}{\sigma_x^2} + \frac{y_r^2}{\sigma_y^2}\right)} \cdot e^{(j\omega_0 x_r)} . \quad (4.18)$$

La envolvente Gaussiana es usualmente elíptica, con una relación de aspecto σ_y/σ_x de 1.5-2.0. La onda sinusoidal compleja y el eje corto de la función Gaussiana elíptica están orientadas en la misma dirección. Cuando el plano de la onda rota también lo hace la función Gaussiana elíptica [6] [50]:

$$\begin{aligned} x_r &= x \cos \theta + y \sin \theta \\ y_r &= -x \sin \theta + y \cos \theta . \end{aligned} \quad (4.19)$$

Los planos frecuenciales en las direcciones x y y están dados por $\xi_0 = \omega_0 \cos \theta$ y $\nu_0 = \omega_0 \sin \theta$ respectivamente con $\theta = \arctan(\nu_0/\xi_0)$.

Una familia de funciones Gabor 2D que satisfacen la teoría wavelet y algunas restricciones neurofisiológicas fue propuesto por Lee [50]. Estas restricciones incluyen una función Gaussiana con una relación de aspecto elíptico σ_y/σ_x de 2:1, la mitad de la amplitud del ancho de banda frecuencial variando desde 1 a 1.5 octavas a lo largo del eje orientado y media igual a cero; produciendo la siguiente wavelet madre:

$$\psi(x, y) = \frac{\omega_0}{\sqrt{2\pi\kappa}} e^{-\frac{\omega_0^2}{8\kappa^2}(4x_r^2 + y_r^2)} \cdot \left[e^{(j\omega_0 x_r)} - e^{-\frac{\kappa^2}{2}} \right] . \quad (4.20)$$

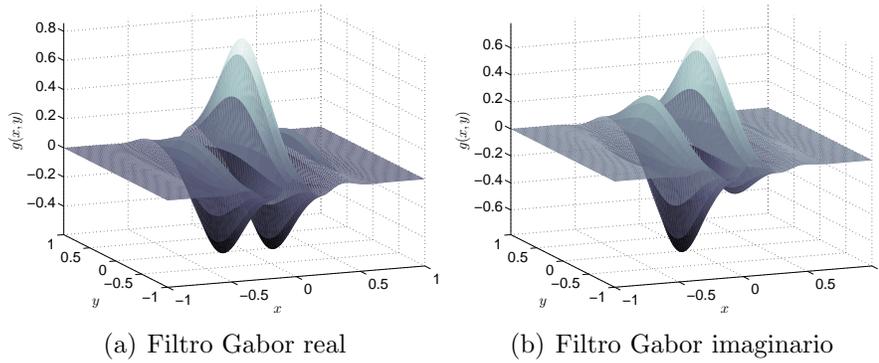


Figura 4.4.: Filtros Gabor real (a) e imaginario (b) en el dominio espacial con parámetros $\phi = 1$, $\theta = 0$ y frecuencia central de 0.078 cpp.

La relación entre σ y ω_0 es $\sigma = \kappa/\omega_0$ donde

$$\kappa = \sqrt{2 \ln 2} \left(\frac{2^\phi + 1}{2^\phi - 1} \right) \quad (4.21)$$

y ϕ es el ancho de banda frecuencial en octavas.

El parámetro de la fase en una wavelet Gabor 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 [70], 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 °). 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. En la Fig. 4.4 se muestran los filtros de Gabor 2D real e imaginario respectivamente, con parámetros $\phi = 1$, $\theta = 0$ y frecuencia central de 0.078 cpp³.

La wavelet madre en sus representaciones real e imaginaria puede escribirse como:

$$\begin{aligned} \psi(x, y)_{\text{Re}} &= \frac{\omega_0}{\sqrt{2\pi\kappa}} e^{-\frac{\omega_0^2}{8\kappa^2}(4x_r^2 + y_r^2)} \cdot \left[\cos(\omega_0 x_r) - e^{-\frac{\kappa^2}{2}} \right] \\ \psi(x, y)_{\text{Im}} &= \frac{\omega_0}{\sqrt{2\pi\kappa}} e^{-\frac{\omega_0^2}{8\kappa^2}(4x_r^2 + y_r^2)} \cdot [\sin(\omega_0 x_r)] \end{aligned} \quad (4.22)$$

³cpp: ciclos por píxel. En el espacio discreto $\omega_{cpp} = \frac{\text{Num. Ciclos}}{\text{Num. Píxeles}}$. Si se toma como unidad básica 1 ciclo, la frecuencia ω_{cpp} disminuye al aumentar el número de píxeles (muestras). En el espacio discreto la frecuencia máxima que se puede representar corresponde a cambios entre 2 píxeles consecutivos, es decir: $Max(\omega_{cpp}) = \frac{1\text{ciclo}}{2\text{píxeles}} = 0.5\text{cpp}$.

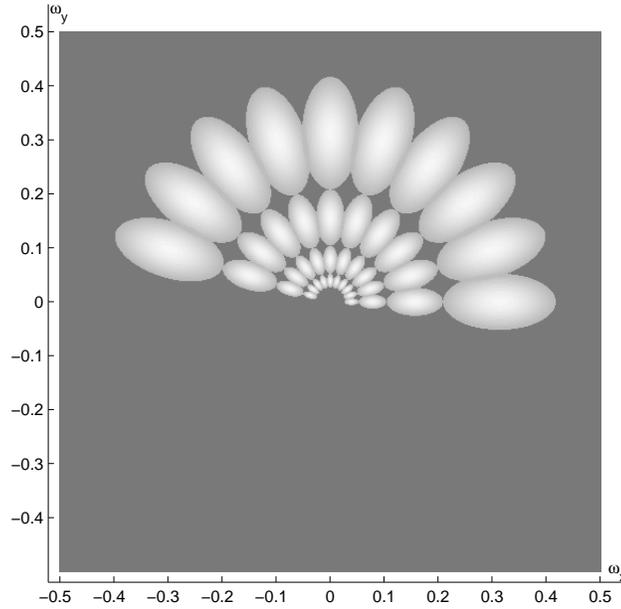


Figura 4.5.: Banco de filtros Gabor en el dominio frecuencial. $S = 4$ escalas de análisis y separación de una octava entre las frecuencias centrales. Ancho de banda frecuencial $\phi = 1$. El número de orientaciones K es igual 10. Las frecuencias centrales más baja y más alta son 0.039 y 0.312 cpp respectivamente.

y su ángulo de fase como:

$$\Phi(x, y) = \arctan \left[\frac{\psi(x, y)_{\text{Im}}}{\psi(x, y)_{\text{Re}}} \right]. \quad (4.23)$$

4.2.1. Filtros Gabor en el dominio frecuencial 2D

La transformada de Fourier de la wavelet madre en la Ec. (4.20) es:

$$\tilde{\psi}(\xi, \nu) = \frac{\sqrt{8\pi\kappa}}{\omega_0} \left[e^{-\frac{\kappa^2}{2\omega_0^2}(\Omega_{xr}^2 + 4\Omega_{yr}^2)} - e^{-\frac{\kappa^2}{2\omega_0^2}(\omega_{xr}^2 + 4\omega_{yr}^2 + \omega_0^2)} \right] \quad (4.24)$$

donde

$$\begin{aligned} \Omega_{xr} &= (\xi - \xi_0) \cos \theta + (\nu - \nu_0) \sin \theta, \\ \Omega_{yr} &= -(\xi - \xi_0) \sin \theta + (\nu - \nu_0) \cos \theta, \\ \omega_{xr} &= \xi \cos \theta + \nu \sin \theta, \\ \omega_{yr} &= -\xi \sin \theta + \nu \cos \theta. \end{aligned} \quad (4.25)$$

En la Fig. 4.5 se muestra un banco de filtros Gabor en el dominio frecuencial con $S = 4$ escalas de análisis y separación de una octava entre las frecuencias centrales. La mitad de

la amplitud del ancho de banda frecuencial ϕ es de una octava. El número de orientaciones K es igual 10, permitiendo que exista un filtro Gabor a cada 18° a partir del eje ω_x . Las frecuencias centrales más baja y más alta son 0.039 y 0.312 cpp respectivamente.

Algunos autores utilizan una transformación no lineal que evoca a una función de activación sigmoide utilizada en las redes neuronales artificiales [41], y suavizado de Gaussiano [15] para mejorar la precisión de la clasificación. En este trabajo utilizamos los filtros Gabor omitiendo esos pasos de post-procesamiento que podrían sesgar los resultados de la clasificación

4.3. Análisis de texturas invariante a la rotación usando la Transformada Hermite orientada

Consideremos una imagen de un mismo tipo de textura dispuesta a diferentes orientaciones. Asumiendo que no se introducen artefactos debido a la interpolación, discretización y cambios de iluminación podemos suponer que la información de la textura es aproximadamente la misma en todas las imágenes rotadas. Debido a que la transformada de Hermite aborda el problema de la dirección de máxima energía, como se vio en el la Sec. 4.1.1, ésta es idónea para analizar texturas a cualquier orientación arbitraria. Para el caso de un mismo tipo de textura dispuesta a distintas orientaciones es posible obtener aproximadamente las mismas características.

La Fig. 4.6 muestra los ejemplos de una textura a distintas orientaciones y sus respectivos coeficientes orientados de Hermite. Como se puede observar, el mismo contenido de información visual es obtenido en cada uno de los coeficientes sin importar la orientación de la textura.

Además de esta información visual, el cálculo de algunas características como la energía (3.4), la media (3.1) y la desviación estándar (3.2) que fueron obtenidos a partir de los coeficientes orientados de Hermite desde el orden $n = 1$ hasta $n = 2$ se muestran en la Tabla 4.3.

Tabla 4.3.: Características de texturas invariantes a la rotación extraídas de los coeficientes orientados de la transformada de Hermite de orden $n = 1$ y $n = 2$ que se muestran en la Fig. 4.6. Las características extraídas son la energía (3.4), la media (3.1) y la desviación estándar (3.2). Obsérvese que estas cantidades son cuantitativamente similares para todos los ángulos de rotación de la imagen original. Los valores de energía están dados $\times 10^3$ píxeles².

	$E_{L_1^\theta}$	$E_{L_2^\theta}$	$\mu_{L_1^\theta}$	$\mu_{L_2^\theta}$	$\sigma_{L_1^\theta}$	$\sigma_{L_2^\theta}$
0°	0.9197	0.1909	0.0582	-0.0012	0.0676	0.0406
17°	0.8932	0.1540	0.0576	-0.0010	0.0664	0.0365
45°	0.9489	0.1625	0.0596	-0.0011	0.0683	0.0375
65°	0.9201	0.1577	0.0589	-0.0010	0.0670	0.0369
120°	0.9277	0.1601	0.0589	-0.0015	0.0675	0.0372

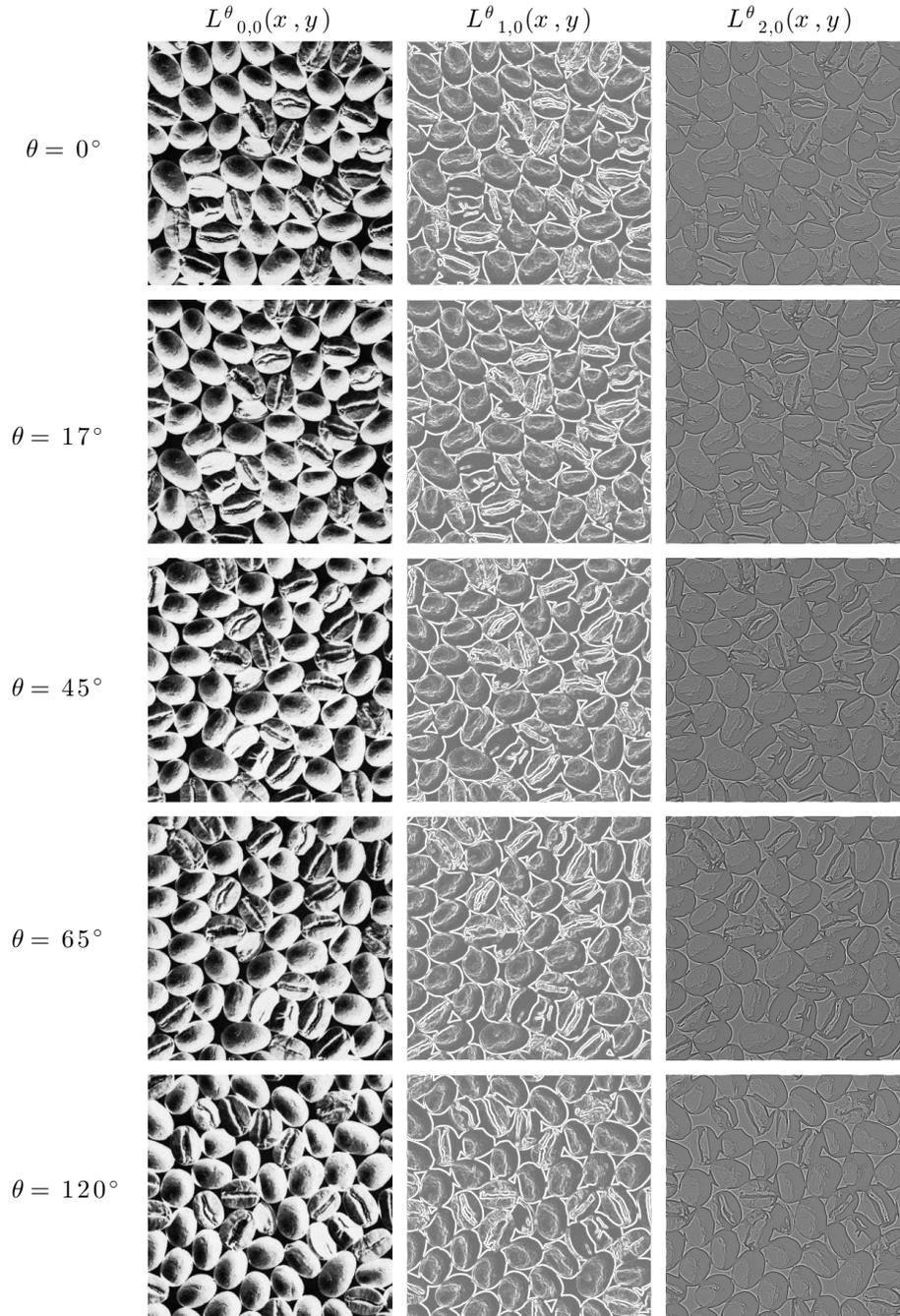


Figura 4.6.: Rotación de una textura y sus coeficientes 1D de Hermite orientados. De izquierda a derecha el orden de los filtros para en cada columna es $n = 0$, $n = 1$ y $n = 2$, respectivamente. La textura corresponde a una región de la textura de Brodatz D74 [8]. Nótese que cualitativamente la misma información visual se conserva sin importar la orientación de la textura. Desde arriba hacia abajo los ángulos de rotación en el sentido de las manecillas del reloj son 0° , 17° , 45° , 65° y 120° .

4.4. Análisis de texturas invariante a la rotación usando filtros Gabor

Para ejemplificar la invarianza a la rotación con los filtros Gabor las técnicas de análisis utilizadas involucran el diseño de un banco de filtros a una sola frecuencia de análisis a distintas orientaciones, siguiendo el modelo presentado en la Sec. 4.2. En este ejemplo se utilizó una frecuencia central de 0.1560 cpp, un ancho de banda radial de 1.5 octavas y 8 orientaciones de análisis.

Las imágenes de textura filtradas son obtenidas a través de la convolución⁴ \star de las funciones Gabor complejas $\psi(x, y)_{k,s}$ con la imágenes de textura $I(x, y)$. Cada componente real e imaginaria de las imágenes filtradas puede ser procesada de manera separada al reemplazar el término complejo de la Ec. (4.22) con su identidad de Euler $e^{j\omega_0 x_r} = \cos \omega_0 x_r + j \sin \omega_0 x_r$, en consecuencia se tiene:

$$F_{im}(x, y)_{k,s} = I(x, y) \star \text{Im}\{\psi(x, y)_{k,s}\} \quad (4.26)$$

$$F_{re}(x, y)_{k,s} = I(x, y) \star \text{Re}\{(\psi(x, y)_{k,s})\} \quad (4.27)$$

donde $k = 1, \dots, K$ y $s = 1, \dots, S$ son los índices de la orientación y la escala respectivamente. Algunos autores realizan la extracción de características únicamente de la componente real de la imagen filtrada mientras que otros emplean la magnitud de la forma compleja calculada como:

$$F(x, y)_{k,s} = \sqrt{F_{im}(x, y)_{k,s}^2 + F_{re}(x, y)_{k,s}^2} \quad (4.28)$$

Existen técnicas para lograr obtener representaciones de la textura invariantes a la rotación a partir de las texturas filtradas con los filtros Gabor como las descritas en la sección 3.4. A continuación se ejemplificarán los resultados de dos técnicas que crean representaciones invariantes a la rotación.

La primer técnica a ejemplificar es semejante a la técnica de la transformada de Hermite orientada. Para encontrar una representación invariante a la rotación se promedian las imágenes resultantes de filtrar con cada uno de los filtros Gabor dispuestos en las distintas orientaciones [64]. La ecuación (3.21) describe este procedimiento. Los resultados visuales se pueden observar en la Fig. 4.7 mientras que los resultados numéricos de los cálculos de algunos tipos de características como la energía (3.4), la media (3.1) y la desviación estándar (3.2) se detallan en la Tabla 4.4.

La segunda técnica a ejemplificar y que es utilizada en esta tesis para el análisis de texturas utilizando los filtros Gabor, consiste en calcular la magnitud de la DFT unidimensional de un vector generado a partir de las características obtenidas de los coeficientes direccionales, como fue propuesto en [54] y [68]. Véase la ecuación (3.24). En las Tabla 4.5 se detallan los cálculos para los tipos de características como la energía (3.4), la media (3.1) y la desviación estándar (3.2).

⁴La convolución puede calcularse eficientemente en el dominio frecuencial haciendo uso del teorema de la convolución. Sea $h(x) = (f \star g)(x) = \mathcal{F}^{-1}\{\mathcal{F}\{f\} \cdot \mathcal{F}\{g\}\}$, donde \mathcal{F} y \mathcal{F}^{-1} son los operadores de la transformada y transformada inversa de Fourier respectivamente.

Tabla 4.4.: Características de texturas invariantes a la rotación obtenidas de los coeficientes Gabor real F_{re}^θ e imaginario F_{im}^θ de la Fig. 4.7. Las características extraídas son: la energía (3.4), la media (3.1) y la desviación estándar (3.2). Estas cantidades son cuantitativamente similares para todos los ángulos de rotación de la imagen original. Los valores de energía están dados $\times 10^3$ píxeles².

	$E_{F_{im}^\theta}$	$E_{F_{re}^\theta}$	$\mu_{F_{im}^\theta}$	$\mu_{F_{re}^\theta}$	$\sigma_{F_{im}^\theta}$	$\sigma_{F_{re}^\theta}$
0°	58.7669	49.2780	-0.0008	0.0345	0.7130	0.65208
17°	59.5039	49.4515	-0.0006	0.0318	0.7175	0.6533
45°	63.3888	52.7766	0.0027	0.0341	0.7405	0.6748
65°	61.3269	52.0091	-0.0008	0.0345	0.7284	0.6699
120°	55.7314	50.6511	0.0002	0.0323	0.6943	0.661

Tabla 4.5.: Características de texturas invariantes a la rotación obtenidas a partir de la magnitud de la DFT del vector de características. Para cada coeficiente Gabor direccional real (F_{re}) e imaginario (F_{im}) se forma un vector de característica de: energía (3.4), media (3.1) y desviación estándar (3.2). Posteriormente se calcula la magnitud de la DFT del vector de características para obtener los elementos del vector invariante a la rotación \mathbf{y} : y_1, y_2 y y_3 (Ec. (3.24)). Estas cantidades son cuantitativamente similares para todos los ángulos de rotación de la imagen original.

	y_1	y_2	y_3	y_1	y_2	y_3
	$E_{F_{im}^\theta}$			$E_{F_{re}^\theta}$		
0°	7.1665	0.6347	0.6265	5.6359	0.4626	0.5651
17°	7.1713	0.5436	0.7860	5.6609	0.4354	0.7488
45°	7.6349	0.6427	0.4982	6.0186	0.4916	0.4887
65°	7.4932	0.7954	0.0732	5.9468	0.6284	0.0788
120°	7.3781	0.4803	0.5613	5.7621	0.3429	0.5112
	$\mu_{F_{im}^\theta}$			$\mu_{F_{re}^\theta}$		
0°	0.0032	0.0020	0.0042	0.1382	0.0043	0.0685
17°	0.0023	0.0012	0.0000	0.1272	0.0086	0.0651
45°	0.0108	0.0027	0.0030	0.1365	0.0013	0.0693
65°	0.0032	0.0035	0.0023	0.1380	0.0013	0.0713
120°	0.0007	0.0011	0.0003	0.1293	0.0013	0.0619
	$\sigma_{F_{im}^\theta}$			$\sigma_{F_{re}^\theta}$		
0°	4.9657	0.2165	0.2137	4.4007	0.1788	0.2175
17°	4.9676	0.1807	0.2680	4.4083	0.1617	0.2866
45°	5.1273	0.2222	0.1750	4.5489	0.1922	0.1892
65°	5.0772	0.2731	0.0390	4.5204	0.2427	0.0408
120°	5.0436	0.1653	0.1931	4.4545	0.1340	0.1973

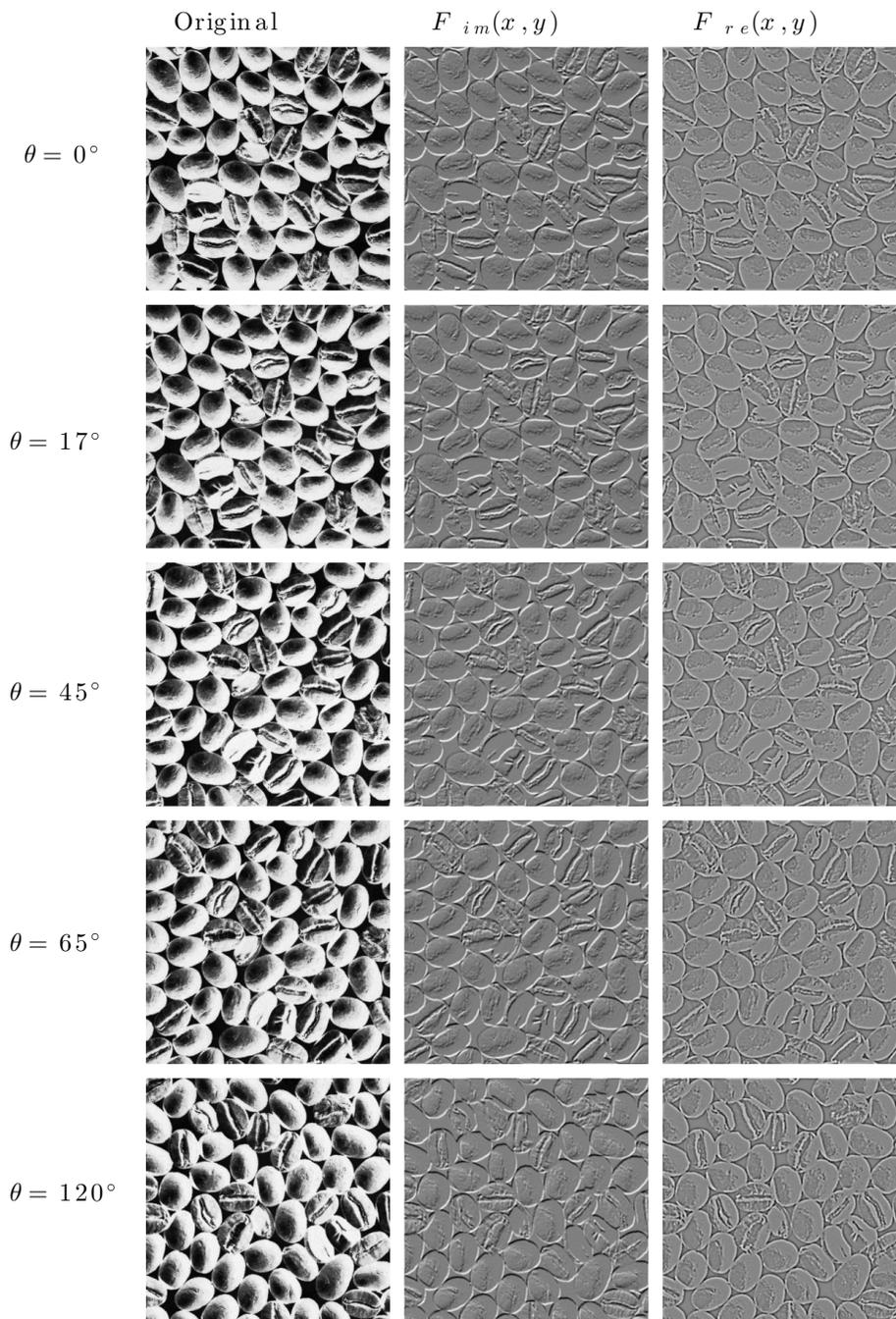


Figura 4.7.: Rotación de una textura y sus coeficientes Gabor imaginario y real. De izquierda a derecha se presenta la imagen original, la imagen filtrada con la parte imaginaria del filtro F_{im} y la filtrada con la parte real del filtro F_{re} , respectivamente. La textura corresponde a una región de la textura de Brodatz D74 [8]. Nótese que cualitativamente la misma información visual se conserva sin importar la orientación de la textura. Desde arriba hacia abajo los ángulos de rotación en el sentido de las manecillas del reloj son 0° , 17° , 45° , 65° y 120° .

5. Separación de clases, métodos de reducción de dimensión y clasificación

Los métodos de reducción de dimensión trabajan sobre el conjunto de características de entrenamiento o de prueba que han sido obtenidos después del análisis de texturas a través de algún método de extracción de características. La reducción de la dimensión del conjunto de características es un tema de importancia cuando se extraen características a través de bancos de filtros. Una desventaja de la estrategia presentada en [14] es que ésta es realizada después de la etapa de la clasificación; seleccionando después de muchas evaluaciones de clasificación que características permiten obtener las tasas más altas. Esta estrategia no permite conocer las características más relevantes previamente a la etapa de clasificación.

El método empleado en [10] para la reducción de la dimensión de los vectores de características está basado en el análisis por componentes principales (PCA). Sin embargo, como cada componente principal es obtenida a partir de una combinación lineal de todos los elementos del vector de características no hay una reducción de dimensión de los datos originales. Esto representa una desventaja, pues se desea reducir el número de operaciones de filtrado durante la clasificación o etapa de pruebas. Otra desventaja es que después de una transformación con PCA no es posible identificar que elementos en el espacio original permiten obtener una mejor discriminación de las texturas.

En este trabajo se utilizará PCA debido a que ha sido uno de los métodos más comunes utilizado para clasificación y reducción de dimensión en la literatura, véase por ejemplo: [4, 35, 42, 59, 73, 81]. Debido a esto, conviene utilizarse para comparar resultados.

5.1. Criterio de Fisher

Los vectores de características calculados a partir de las muestras de textura para cada clase no son idénticos. Estos vectores forman un cúmulo (grupo o *cluster*) en un espacio multidimensional. Entre más grande sea la distancia que separa a dos cúmulos que corresponden a dos tipos de texturas diferentes, mejor serán las propiedades de la discriminación del operador de texturas que produjo los vectores de características.

Sean $\mathbf{C}_1 = \{\mathbf{c}_{1,1}, \mathbf{c}_{1,2}, \dots, \mathbf{c}_{1,q}\}$ y $\mathbf{C}_2 = \{\mathbf{c}_{2,1}, \mathbf{c}_{2,2}, \dots, \mathbf{c}_{2,q}\}$ dos conjuntos de vectores pertenecientes a las clases o cúmulos C_1 y C_2 respectivamente. Para determinar la distancia entre dos cúmulos de vectores de características, se asume que es suficiente con observar sus proyecciones en un espacio 1D, es decir, una línea. La suposición es que dicha proyección maximiza la separación entre clusters en el espacio 1D. Fisher [32] introdujo una transformación lineal que bajo ciertas condiciones lleva a cabo dicha proyección y es llamada la *función*

discriminante lineal de Fisher. Tiene la siguiente forma:

$$x = \mathbf{x} \Sigma_{\mathbf{C}_1 \mathbf{C}_2}^{-1} (\bar{\mathbf{c}}_1 - \bar{\mathbf{c}}_2)^T \quad (5.1)$$

donde los vectores promedio o *centroides* de los cúmulos \mathbf{C}_1 y \mathbf{C}_2 son $\bar{\mathbf{c}}_1$ y $\bar{\mathbf{c}}_2$ respectivamente, $\Sigma_{\mathbf{C}_1 \mathbf{C}_2}^{-1}$ es la inversa de la matriz de covarianza entre ambos cúmulos, \mathbf{x} es cualquier vector de características y x es su proyección 1D.

Esta proyección de los vectores de características hacia el espacio 1D maximiza el criterio de Fisher, el cual mide la separación entre dos cúmulos en el espacio reducido:

$$f = \frac{|\eta_1 - \eta_2|}{\sqrt{\sigma_1^2 + \sigma_2^2}} \quad (5.2)$$

donde σ_1 y σ_2 son las desviaciones estándar de las distribuciones de los vectores de características proyectadas de los dos cúmulos y η_1 y η_2 son las proyecciones de las medias $\bar{\mathbf{c}}_1$ y $\bar{\mathbf{c}}_2$ respectivamente. El criterio de Fisher expresa en una sola cantidad la distancia entre dos cúmulos relativa al tamaño de los cúmulos. Entre más grande sea el valor del criterio de Fisher calculado para dos cúmulos, mayor es la separación entre ambos

5.2. Análisis por componentes principales (PCA)

La técnica de análisis por componentes principales (PCA), también conocida como la transformada Karhunen Loeve (KL) o la transformada “Hotelling”, está basada en técnicas de factorización de álgebra lineal [65]. PCA puede ser visto como una transformación del conjunto de datos original a un nuevo espacio vectorial generado por funciones base que descomponen los datos en componentes de energía de orden decreciente. Con el fin de reducir la dimensión o para comprimir los datos originales las componentes de menor energía se pueden eliminar. Muchos esquemas de clasificación trabajan en este espacio para lograr mejor separación de clases.

Sea \mathbf{V} un conjunto de q vectores o mediciones:

$$\mathbf{V} = \{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_q\} \quad (5.3)$$

tales que el i -ésimo vector \mathbf{v}_i tiene r elementos o características:

$$\mathbf{v}_i = \{v_{i,1}, v_{i,2}, \dots, v_{i,r}\} \quad (5.4)$$

Podemos agrupar todas las características en una matriz de características $\mathbf{F}_{\mathbf{V}}$ como:

$$\mathbf{F}_{\mathbf{V}} = \begin{bmatrix} v_{1,1} & v_{1,2} & \dots & v_{1,r} \\ v_{2,1} & v_{2,2} & \dots & v_{2,r} \\ \vdots & \vdots & \vdots & \vdots \\ v_{q,1} & v_{q,2} & \dots & v_{q,r} \end{bmatrix} \quad (5.5)$$

La idea detrás de PCA es resaltar los datos que producen las variaciones máximas calculadas por la matriz de covarianza definida como:

$$\Sigma_{\mathbf{V}} = \begin{bmatrix} cov_{1,1} & cov_{1,2} & \dots & cov_{1,r} \\ cov_{2,1} & cov_{2,2} & \dots & cov_{2,r} \\ \vdots & \vdots & \vdots & \vdots \\ cov_{r,1} & cov_{r,2} & \dots & cov_{r,r} \end{bmatrix} \quad (5.6)$$

siendo $cov_{i,j}$ la covarianza entre las características i y j . A partir de la matriz de covarianza se puede obtener información sobre los datos. Por ejemplo, si todos los valores de la matriz son cero excepto por su diagonal entonces significa que las características son independientes. El objetivo de PCA es transformar los datos de tal manera que la matriz de covarianza sea diagonal.

Para transformar cada vector de características definido en el conjunto \mathbf{V} en otro vector de características definido en el nuevo conjunto \mathbf{Y} se busca una transformación \mathbf{W} , tal que la matriz de covarianza de los elementos en \mathbf{Y} es diagonal. La transformación es lineal y está definida como:

$$\mathbf{F}_{\mathbf{Y}} = \mathbf{F}_{\mathbf{V}} \mathbf{W}^T \quad (5.7)$$

Podemos encontrar la matriz de transformación \mathbf{W} al resolver la siguiente Ec. [65]:

$$(\lambda_i \mathbf{I} - \Sigma_{\mathbf{V}}) \mathbf{w}_i = 0 \quad (5.8)$$

donde \mathbf{w}_i son los renglones de \mathbf{W} , λ_i y \mathbf{w}_i definen los eigen-valores y eigen-vectores respectivamente e \mathbf{I} es la matriz de identidad. La ecuación característica $\det(\lambda_i \mathbf{I} - \Sigma_{\mathbf{V}}) = 0$ es usada para resolver para los valores de λ_i y consecuentemente obtener \mathbf{w}_i de la Ec. (5.8).

5.3. Tasa de varianza aumentada (AVR)

La *tasa de varianza aumentada* (AVR) propuesta en [51] ha mostrado ser una base cuantitativa para separar las características que no aportan discriminación entre las clases. AVR está definida como:

$$AVR(F) = \frac{Var(S_F)}{\frac{1}{C} \sum_{i=1..C} \min_{i \neq j} (|mean_i(S_F) - mean_j(S_F)|)} \quad (5.9)$$

donde $Var(S_F)$ es la varianza inter-clase de la característica F , $Var_i(S_F)$ y $mean_i(S_F)$ son la varianza intra-clase y el promedio de la característica F para la clase i dentro de C clases distintas respectivamente. De manera similar al criterio de Fisher, AVR representa la tasa de la varianza intra-clase de la característica entre la varianza inter-clase pero incluye además un factor de penalización a las características inter-clase que tienen medias muy cercanas. El índice j indica el recorrido para calcular las separaciones entre la media de la clase i y las medias de las otras clases con $i \neq j$ y posteriormente escoger la separación inter-clase

mínima con la función $\min()$.

Al ordenar los elementos de los vectores de características con tasas AVR altas es posible saber los índices de los filtros que fueron utilizados para obtener dichas características. La selección de filtros que serán utilizados durante la etapa de prueba es un procedimiento sencillo a partir de este ordenamiento.

5.4. Varianza de inter-clase (ICV)

Este método se introduce en este trabajo como un método más inmediato para obtener de manera estimada que características son las más importantes. La idea es suponer que los valores altos de la función de auto-covarianza (diagonal de la matriz de covarianza en la ecuación (5.6)) de un conjunto de características agrupados en forma matricial, como en la ecuación (5.5), representan a aquellas características con alta variabilidad inter-clase:

$$\text{ICV} = \text{diag}(\Sigma_{\mathbf{V}}) , \quad (5.10)$$

donde $\Sigma_{\mathbf{V}}$ es la matriz de covarianza del conjunto de vectores de características \mathbf{V} .

Teniendo un método robusto de análisis de texturas invariante a la rotación, se espera que los vectores de características que pertenecen a una misma clase estén muy cercanos entre si de tal manera que las variaciones sean producidas debido a la diferencia entre las distintas clases de textura.

Al igual que sucede con AVR, al ordenar los elementos de los vectores de características con tasas ICV (auto-covarianza) altas es posible saber los índices de los filtros que fueron utilizados para obtener dichas características. La selección de filtros que serán utilizados durante la etapa de prueba es un procedimiento sencillo a partir de este ordenamiento.

5.5. Clasificación

Usualmente se tiene una descripción de una muestra de textura y se desea encontrar que elemento de la base de datos es lo más parecida a dicha muestra. En esencia esto es el corazón del reconocimiento de patrones.

Existen muchos métodos de clasificación en la literatura, sin embargo, en esta tesis se limitó el uso a uno de los clasificadores más sencillos y no se profundiza en investigar o experimentar con clasificadores más elaborados, pues el objetivo de la tesis es investigar la mejor forma de seleccionar subconjuntos de características (como los presentados en la Secciones 5.2-5.4) y los tipos de características, que podrían propiciar una mejor clasificación, considerando que cualquier otro método de clasificación que mejore las tasas de clasificación podría introducir un sesgo en una selección objetiva de los tipos de características.

La clasificación en el contexto del análisis de texturas se refiere a asociar la etiqueta de la clase apropiada (tipo de textura) con la muestra de prueba usando una medida que la describe. Si seleccionamos el punto en el espacio de características que está más cercano a la muestra bajo análisis entonces habremos seleccionado el vecino de la muestra más cercano.

Uno de los clasificadores más simples que se encuentran en la literatura es el de los k vecinos más cercanos; el cual consiste en asignar a la muestra bajo análisis la clase seleccionada por la *moda* de los k vecinos más cercanos de dicha muestra. La selección de más de un punto introduce una forma de suavizado del espacio de características y permite que la decisión de la clasificación no sea afectada por puntos atípicos considerados como ruido [65]. En las pruebas experimentales se decidió emplear al primer vecino más cercano $k = 1$ para evaluar de manera más rígida la clasificación de las texturas al tener elementos que introducen ruido en el resultado final de ésta.

5.5.1. Análisis de la sensibilidad y especificidad del sistema

El problema de reconocer que una textura dada pertenece a una clase en particular, 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 ($Z_1 \cap Z_2 = \emptyset$) 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 [60].

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 [60]:

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

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

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

$$P\{d_2|m_1\} = P\{z \in Z_2|m_1\} = \int_{Z_2} p(z|m_1)dz \quad (5.11)$$

$$P\{d_1|m_2\} = P\{z \in Z_1|m_2\} = \int_{Z_1} p(z|m_2)dz \quad . \quad (5.12)$$

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 5.1 se pueden apreciar estas regiones de decisión. En el contexto de la toma de una decisión para clasificar una textura, llamaremos a estos errores como error de falso negativo (FN) y

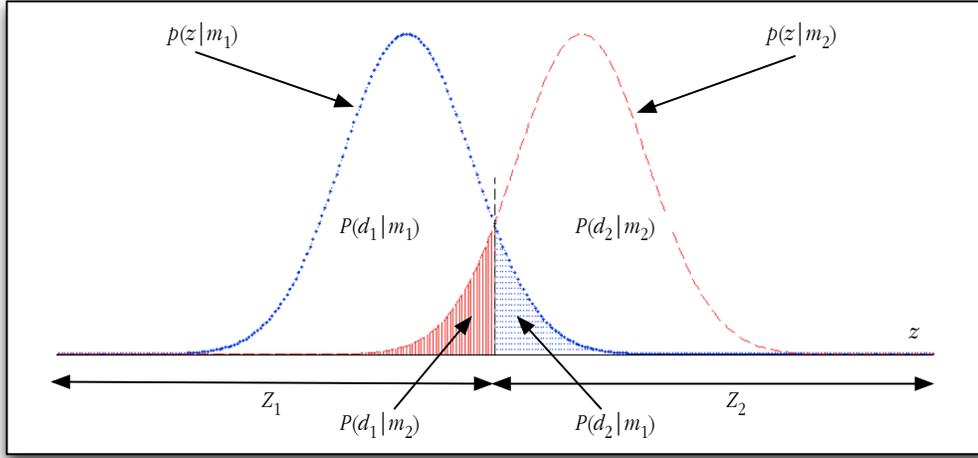


Figura 5.1.: Distribuciones en un problema típico de decisión binaria. La distribución $P\{d_1|m_1\}$ es un histograma de la separación entre elementos que pertenecen a una misma clase (intra-clase). La distribución $P\{d_2|m_2\}$ es un histograma de la separación entre elementos que pertenecen a clases distintas (inter-clase). Al variar z podemos obtener distintas tasas para los errores $P\{d_2|m_1\}$ y $P\{d_1|m_2\}$.

error de falso positivo (FP) 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 5.1). Estas decisiones correctas tienen las siguientes probabilidades asociadas [60]:

$P\{d_1|m_1\}$ = 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:

$$P\{d_1|m_1\} = P\{z \in Z_1|m_1\} = \int_{Z_1} p(z|m_1)dz \tag{5.13}$$

$$P\{d_2|m_2\} = P\{z \in Z_2|m_2\} = \int_{Z_2} p(z|m_2)dz . \tag{5.14}$$

Nuevamente en el contexto de clasificación de texturas, llamaremos a estas dos decisiones correctas como positivos verdaderos (TP) y negativos verdaderos (TN). 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:

$$P\{d_1|m_1\} + P\{d_2|m_1\} = 1 \tag{5.15}$$

$$P\{d_2|m_2\} + P\{d_1|m_2\} = 1 . \tag{5.16}$$

Este resultado se puede obtener al sustituir las ecuaciones (5.12) y (5.13) en la ecuación (5.15), es decir:

$$\begin{aligned} P\{d_1|m_1\} + P\{d_2|m_1\} &= \int_{z_1} p(z|m_1)dz + \int_{z_2} p(z|m_1)dz \\ &= \int_Z p(z|m_1)dz = 1 . \end{aligned}$$

A partir de las distribuciones intra-clase e inter-clase obtenidas con cada tipo de distancia (Ecuaciones (3.15), (3.16) y (3.18)) es posible obtener el desempeño de la clasificación (en términos de una curva ROC¹ de la sensibilidad S y especificidad E del sistema) para cada uno de los tipos de características, distancias y los métodos de reducción de dimensión presentados anteriormente. Podemos aproximar las probabilidades dadas por las ecuaciones (5.11), (5.12), (5.13) y (5.14) a partir de los TP , TN , FN y FP obtenidos experimentalmente como:

$$S = P\{d_1|m_1\} = \int_0^C P_{Au}(z)dz \quad (5.17)$$

$$FRR = P\{d_2|m_1\} = \int_C^1 P_{Au}(z)dz \quad (5.18)$$

$$FAR = P\{d_1|m_2\} = \int_0^C P_{Im}(z)dz \quad (5.19)$$

$$E = P\{d_2|m_2\} = \int_C^1 P_{Im}(z)dz , \quad (5.20)$$

donde

$$P_{Au}(z) = \frac{TP(z)}{TP(z) + FN(z)} \quad (5.21)$$

$$P_{Im}(z) = \frac{TN(z)}{TN(z) + FP(z)} , \quad (5.22)$$

S es la sensibilidad y E es la especificidad del sistema en el intervalo de 0 a C y de C a 1 respectivamente. FRR y FAR son las probabilidades condicionales de los errores que se interpretan como las tasas de error de los falsos negativos y de los falsos positivos del sistema en el intervalo de C a 1 y de 0 a C , respectivamente.

De las ecuaciones (5.15) y (5.16) se observa que S y FRR son complementarias [61] así como también los son E y FAR .

$$S = 1 - FRR \quad (5.23)$$

$$E = 1 - FAR . \quad (5.24)$$

La curva ROC se puede generar entonces al graficar S como función de FAR , ya que

¹ROC, por sus siglas en inglés "Receiver Operating Characteristic"

$FAR = 1 - E$. Una curva ROC de un sistema ideal se caracteriza por tener una $FAR = 0$ y una $FRR = 0$. Para reportar los valores de la sensibilidad se suele fijar una tasa de error FAR y a partir de ésta reportar S . Es una práctica común reportar el valor de S cuando las probabilidades condicionales de error son iguales, es decir $ERR = FAR = FRR$. La interpretación de esta convención es que el sistema asume un mismo costo para ambos errores. En los sistemas de identificación biométricos se busca minimizar FAR , ya que es de mayor costo permitir el acceso a un impostor que rechazar incorrectamente a un usuario previamente registrado en el sistema [21].

6. Resultados

En este capítulo se presentan los resultados de la clasificación y el análisis de los mismos usando la transformada de Hermite y los filtros Gabor. Para las pruebas experimentales se emplearon las bases de datos de imágenes de textura Brodatz [8], Outex [66, 67] y VisTex [80]. La metodología se basa en un análisis estadístico de 80 repeticiones o experimentos en los cuales para cada uno de ellos se seleccionaron aleatoriamente conjuntos disjuntos de entrenamiento y de pruebas. La agenda experimental incluyó una serie de evaluaciones como: el número de imágenes de entrenamiento, las tipos de características (Véase Tabla 6.1), tamaño del vector de características y configuraciones de los filtros de análisis que propician una mejor tasa de clasificación.

Bases de datos de imágenes de texturas

Para la base de datos de imágenes de texturas **Brodatz** [8] se extrajeron 25 texturas del álbum del mismo nombre de 512×512 píxeles como fue propuesto en [40] y [71]. Las clases de textura están codificadas en la base de datos de Brodatz como *DY* y son: $D\{1, 4, 6, 19, 20, 21, 22, 24, 28, 34, 52, 53, 56, 57, 66, 74, 76, 78, 82, 84, 102, 103, 105, 110, 111\}$.

Con respecto a la base de datos imágenes de texturas **VisTex** [80], ésta se formó a partir de 54 imágenes de texturas de 512×512 píxeles. Las clases de texturas utilizadas están codificadas en la base de datos VisTex como *Nombre.YYYY* y son: *Bark*.{0000, 0001, 0003, 0005, 0006, 0007, 0008, 0010}, *Brick*.{0000, 0002, 0004, 0005}, *Buildings*.{0003, 0009}, *Fabric*.{0000, 0002, 0004, 0008, 0011, 0014, 0015, 0017}, *Flowers*.{0000, 0003, 0005}, *Food*.{0001, 0003, 0005, 0008}, *Grass*.{0002}, *Leaves*.{0000, 0002, 0003, 0006, 0010, 0013, 0014, 0016}, *Metal*.{0000, 0002, 0004}, *Misc*.{0001, 0002}, *Sand*.{0002, 0005}, *Stone*.{0002, 0004}, *Terrain*.{0010}, *Tile*.{0001, 0007}, *Water*.{0000, 0004, 0006}, *Wood*.{0002}.

La base de datos **Outex** [66, 67] fue formada utilizando 24 de imágenes de textura de 538×746 píxeles codificadas en la base de datos Outex como *nombreYYY* y son: : *canvas*{001, 002, 003, 005, 006, 009, 011, 021, 022, 023, 025, 026, 031, 032, 033, 035, 038, 039}, *tile*{005, 006} y *carpet*{002, 004, 005, 009}. Las imágenes de textura fueron adquiridas con iluminación incandescente, denominada “inca” dentro de la base de datos, con una resolución espacial de 100 dpi.

Para ambas bases de datos **Brodatz** y **VisTex**, cada una de las clases de textura fueron rotadas artificialmente con 17 ángulos distintos, desde 0° hasta 170° en incrementos de 10° . Después se dividió esta imagen de textura rotada en 4 al tomar las regiones más externas (a partir de las esquinas de la imagen original) de 128×128 píxeles para minimizar el traslape entre dichas regiones[40, 71]. La base de datos Brodatz fue formada con $25 \times 17 \times 4 = 1,700$ texturas y la base de datos VisTex con $54 \times 17 \times 4 = 3,672$ texturas.

Tabla 6.1.: Resumen de los distintos tipos de características de textura utilizadas en la agenda experimental. Para cada tipo de característica se formó un vector de características al concatenar la característica calculada a partir de cada imagen filtrada.

Característica	Ecuación
μ	$\frac{1}{H \times W} \sum_{h=1}^H \sum_{w=1}^W f(h, w)$
σ	$\sqrt{\frac{1}{(H \times W - 1)} \sum_{h=1}^H \sum_{w=1}^W [f(h, w) - \mu]^2}$
σ_b	$\sqrt{\frac{1}{(H \times W - 1)} \sum_{h=1}^H \sum_{w=1}^W [f(h, w) - \mu]^2}$
E_0	$\sum_{h=1}^H \sum_{w=1}^W [f(h, w)]^2$
E_1	$\frac{E_0}{H \times W}$
E_2	$\frac{1}{H \times W} \sum_{h=1}^H \sum_{w=1}^W f(h, w) $
E_3	$\frac{1}{H \times W} \sum_{h=1}^H \sum_{w=1}^W \sqrt{ f(h, w) }$
E_4	$\sum_{h=1}^H \sum_{w=1}^W [f(h, w) - \mu]^2$
E_5	$-\sum_{i=0}^{B-1} p(i) ^2 \log p(i) ^2$

En el caso de las base de datos de **Outex** las texturas fueron capturadas al rotar éstas de forma natural con 9 ángulos de rotación distintos (0° , 5° , 10° , 15° , 30° , 45° , 60° , 75° y 90°). Las imágenes en RGB de 24 bits fueron posteriormente transformadas en imágenes con 8 bits de intensidad usando la ecuación [66]:

$$I = 0.299R + 0.587G + 0.114B \quad (6.1)$$

A partir de cada imagen transformada se extrajeron 20 regiones de texturas sin traslape de 128×128 píxeles, conformando la base de datos de texturas experimentales Outex en su totalidad con $24 \times 9 \times 20 = 4,320$ texturas [66].

Metodología experimental usando la transformada de Hermite

1. Filtrado de todas las imágenes para cada base de datos usando la transformada de Hermite orientada desde orden $n = 1$ hasta $n = 8$ con $S = 4$ escalas de análisis.
2. En un vector exclusivo para cada tipo de característica, concatenación de cada característica extraída a partir de las imágenes filtradas. La dimensión de vector resultante es de $n_{max} \times S$, donde $n_{max} = 8$ y $S = 4$ escalas de análisis (Véase Ec. (5.3)). Los tipos de características extraídos se resumen en la tabla 6.1.
3. Obtención de las distribuciones intra-clase e inter-clase para calcular el criterio de Fisher. El criterio de Fisher es calculado para cada tipo de característica por separado, utilizando el número máximo de imágenes de entrenamiento (20 imágenes) y todos

los elementos que conforman al vector de características así como también para las distancias Euclidiana, Canberra y Euclidiana normalizada (Ecuaciones (3.15), (3.16) y (3.18) respectivamente) descritas en las Sec. 3.3.

4. A través de una prueba ANOVA, determinar estadísticamente el número mínimo de los elementos que conforman al vector de características tal que la diferencia de las tasas de clasificación correctas (CCRs, %) es estadísticamente similar ($p < 0.05$) a la obtenida cuando se usan todos los elementos del vector de características. Comparar los resultados entre los métodos de reducción de dimensión presentados en el Cap. 5.
5. Cálculo de la sensibilidad usando el número mínimo de elementos que pueden conforman al vector de características tal que la diferencia de las tasas de clasificación correctas (CCRs, %) es estadísticamente similar ($p < 0.05$) a la obtenida cuando se usan todos los elementos del vector de características.
6. Determinación experimental de los órdenes de los filtros requeridos por los métodos de reducción de dimensión AVR e ICV. Esto se obtendrá a partir de un histograma conteniendo la frecuencia con la cual los índices de los filtros de análisis de Hermite fueron seleccionados en cada repetición del experimento.
7. Cálculos de CCR usando un número mínimo de elementos que conforman al vector de características y evaluación del comportamiento de las distancias incluyendo la distancia de Mahalanobis presentada en el Cap. 3.3. Comparar los resultados entre los métodos de reducción de dimensión presentados en el Cap. 5.

Metodología experimental usando filtros Gabor

1. Diseño de varios bancos de filtros Gabor con 3 anchos de banda frecuenciales distintos (variando ϕ en 1, 1.5 y 2 octavas) y 3 números de orientaciones distintas (variando K en 6, 8 y 10 orientaciones). Este diseño incluye $S = 4$ escalas de análisis.
2. Filtrado de todas las imágenes para cada base de datos usando los bancos de filtros Gabor.
3. Concatenación en un vector exclusivo para el mismo tipo de característica, de la característica extraída a partir de cada imagen filtrada. Los tipos de características extraídos se resumen en la tabla 6.1.
4. Cálculo de la magnitud de la DFT del vector de características originales para obtener un vector de características invariante a la rotación. La dimensión final del vector de características invariante a la rotación está dado por la Ec. (3.25).
5. Utilizar una prueba ANOVA para determinar estadísticamente la interacción de los factores y de esta manera proponer las posibles combinaciones del número de orientaciones y de los anchos de banda frecuenciales con los cuales se obtienen estadísticamente las mejores tasas de clasificación.

6. A partir de la combinación obtenida en el punto anterior, obtención de las distribuciones intra-clase e inter-clase para calcular el criterio de Fisher. El criterio de Fisher es calculado para cada tipo de característica por separado, utilizando el número máximo de imágenes de entrenamiento y todos los elementos que conforman al vector de características así como también para las distancias Euclidiana, Canberra y Euclidiana normalizada (Ecuaciones (3.15), (3.16) y (3.18) respectivamente) descritas en las Sec. 3.3.
7. Cálculo de la sensibilidad usando la combinación del número de orientaciones y de los anchos de banda frecuenciales encontrados en el punto 5.

6.1. Resultados usando la transformada de Hermite

6.1.1. Criterio de Fisher usando la transformada de Hermite

El criterio de Fisher es utilizado para medir la separación entre las distribuciones intra-clase e inter-clase. Las distribuciones intra-clase se generan al calcular para cada clase las distancias entre los vectores de características pertenecientes a una misma clase. Las distribuciones inter-clase se generan al calcular para cada clase las distancias entre los vectores de características pertenecientes a dicha clase con respecto a los vectores de características pertenecientes al resto de las clases.

El valor del criterio de Fisher que se presenta en los siguientes resultados, obtenidos al seleccionar aleatoriamente conjuntos disjuntos de entrenamiento y de prueba, es un indicativo de una separación promedio entre las clases. Los resultados que se presentan a continuación fueron obtenidos usando las distancias Euclidiana (d_E), Canberra (d_C) y Euclidiana normalizada (d_{E_N}) descritas en las Sec. 3.3 (Ecuaciones (3.15), (3.16) y (3.18) respectivamente) para cada una de las distintas características que fueron presentadas en la Sec. 3.1.

Las Figs. 6.1-6.3 corresponden a las distribuciones de los cálculos del valor del criterio de Fisher para 80 repeticiones del experimento de clasificación usando los conjuntos de texturas Brodatz [8], VisTex [80] y Outex [66, 67] respectivamente. El resultado del valor del criterio de Fisher al combinar un tipo de característica y una medida de distancia, empleada para comparar los vectores de características, se muestran por separado en cada diagrama de caja.

Para los tres conjuntos de texturas se observó que al comparar entre las combinaciones de los tipos de características y las medidas de distancias, el número de Fisher promedio fue mayor al utilizar la característica μ (Ec. (3.1)). La característica σ_b (Ec. (3.3)) presentó también un número de Fisher promedio mayor que el que se obtiene al utilizar la característica σ (Ec. (3.2)).

La distancia Euclidiana (Ec. (3.15)) es susceptible a presentar grandes variaciones en la separación de clases cuando las características son modificadas por un valor escalar. Obsérvese que las características E_0 y E_1 en las ecuaciones (Ec. (3.4)) y (Ec. (3.5)) respectivamente solo difieren por $\frac{1}{H \times W}$. Los resultados muestran que tanto la distancia Canberra (Ec. (3.16)) como la distancia Euclidiana normalizada (Ec. (3.18)) son robustas a los cambios escalares en los valores de las características.

Para la característica E_4 (Ec. (3.8)), que resulta de sustraer el valor medio al coeficiente y posteriormente calcular su energía de manera similar al cálculo de E_0 , las distancias Canberra y Euclidiana normalizada presentan un criterio de Fisher similar al obtenido con E_0 y E_1 a diferencia del criterio de Fisher obtenido al comparar los vectores usando la distancia Euclidiana. Sin embargo, se logra una mayor separación entre las clases usando las características E_0 y E_1 .

La característica E_3 (Ec. (3.7)) genera una separación ligeramente mejor entre clases que E_2 (Ec. (3.6)). Como se observó anteriormente, al utilizar la distancia Euclidiana se presenta una mayor variación entre ambas características que con las distancias Canberra y Euclidiana normalizada a pesar de que dichas características son obtenidas de manera muy similar. Por otro lado, al usar la característica E_5 (Ec. (3.9)) se observó que la separación entre

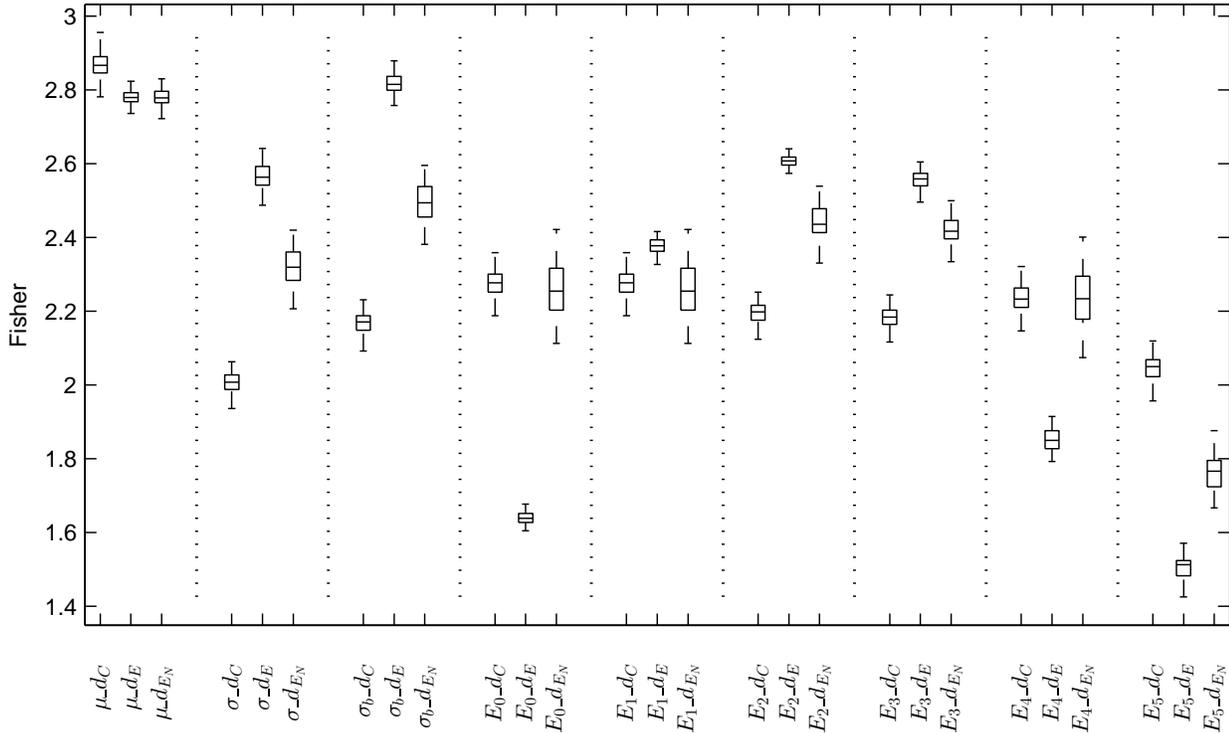


Figura 6.1.: Comparación del criterio de Fisher entre las distintas combinaciones de tipos características y medidas de distancias para el conjunto de texturas Brodatz usando la transformada de Hermite orientada [8] al obtener las distribuciones intra-clase e inter-clase con 80 repeticiones del experimento de clasificación. Cada diagrama de caja muestra el valor del criterio de Fisher que resulta al combinar un tipo de característica con una medida de distancia.

clases obtenida es muy baja respecto a la separación obtenida con el resto de los tipos de características sin importar que medida de distancia haya sido utilizada.

6.1.2. Reducción del número de elementos que conforman al vector de características

En esta sección se presentan los resultados de la evaluación estadística sobre el número mínimo de elementos que conforman al vector de características tales que la tasa de clasificación correcta (CCR) es estadísticamente similar a la obtenida cuando se usan todos los elementos del vector de características. Los resultados se presentan para cada una de las distancias utilizando los tres métodos de reducción de dimensión: PCA, AVR e ICV.

Efectos principales

Cada tipo de característica de textura fue evaluado de manera separada usando una prueba ANOVA de 2 vías con los siguientes factores: el número de elementos del vector de característi-

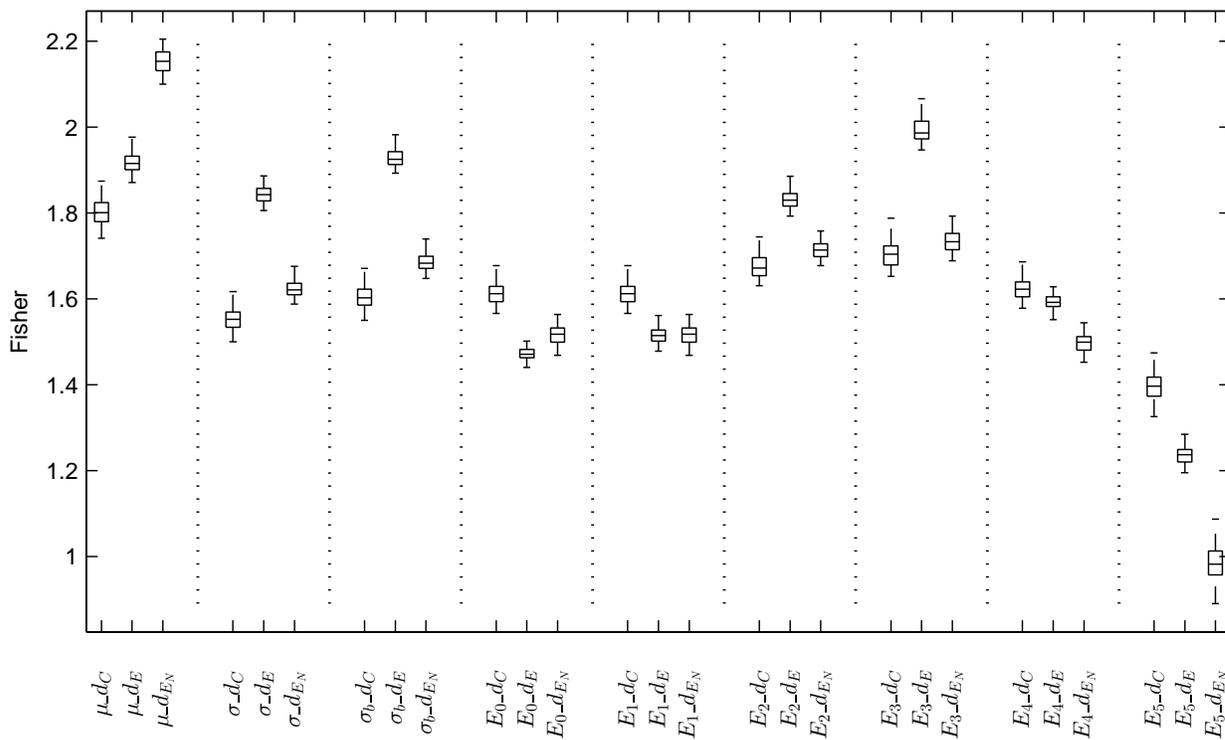


Figura 6.2.: Comparación del criterio de Fisher entre las distintas combinaciones de tipos características y medidas de distancias para el conjunto de texturas VisTex usando la transformada de Hermite orientada [80] al obtener las distribuciones intra-clase e inter-clase con 80 repeticiones del experimento de clasificación. Cada diagrama de caja muestra el valor del criterio de Fisher que resulta al combinar un tipo de característica con una medida de distancia.

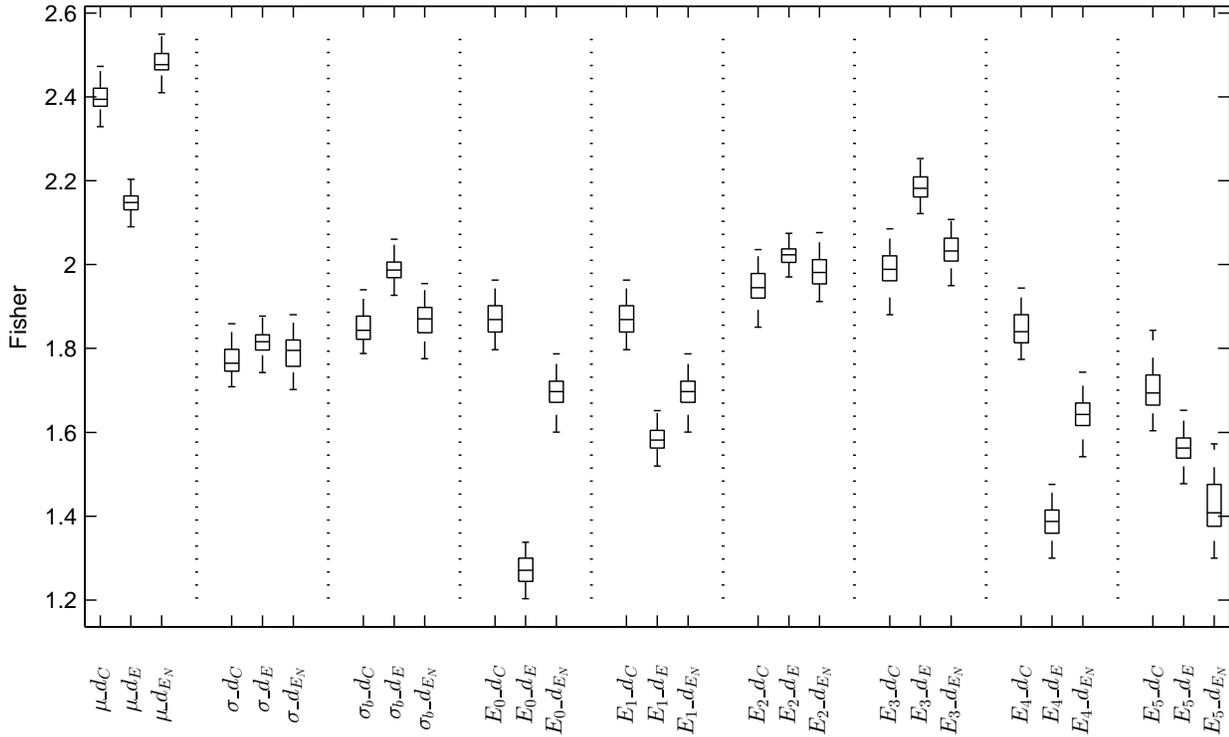


Figura 6.3.: Comparación del criterio de Fisher entre las distintas combinaciones de tipos características y medidas de distancias para el conjunto de texturas Outex usando la transformada de Hermite orientada [66, 67] al obtener las distribuciones intra-clase e inter-clase con 80 repeticiones del experimento de clasificación. Cada diagrama de caja muestra el valor del criterio de Fisher que resulta al combinar un tipo de característica con una medida de distancia.

cas r (desde 1 hasta 32 elementos) y el tipo de distancia utilizado, d_E , d_C y d_{EN} (Ecuaciones (3.15), (3.16) y (3.18) respectivamente) para cada una de las bases de datos de texturas. Se observó que los principales efectos de los dos factores fueron estadísticamente significativos ($p < 0.05$) para las tres bases de datos y para cada tipo de característica de textura.

El principal efecto del factor r que observamos en todos los casos (excepto cuando se usa PCA) es que la tasa de clasificación correcta $CCR(\%)$ no necesariamente se mejora conforme el número de elementos del vector r aumenta. Por el contrario, usando AVR e ICV se confirma que existe un número y una selección específica de ciertos elementos del vector de características que mejoran la $CCR(\%)$ con respecto a la obtenida con el método PCA.

En la Tabla 6.2 se muestran los resultados de esta evaluación. Al utilizar el método PCA es necesario considerar más elementos del vector de características que al utilizar AVR o ICV para obtener $CCRs(\%)$ altas. Esto tiene sentido ya que las componentes resultantes del método PCA tienen su origen en una combinación lineal de todas las características originales. El método AVR proporcionó una mayor reducción del número de elementos del vector de características comparado con PCA e ICV. El método ICV requiere de un número

Tabla 6.2.: Efecto principal del número de elementos que conforman al vector de características r con cada método de reducción de dimensión: PCA, AVR, ICV. Se reporta la CCR (%) obtenida con el número mínimo de elementos (el valor inmediato al símbolo @) que conforman al vector de características que fue encontrada estadísticamente similar a la mejor CCR (Tukey=0.05) al considerar un vector de 32 elementos. Se reporta además un número de elementos promedio por característica, base de datos, un promedio por método y un promedio global.

Método		Brodatz	VisTex	Outex	Promedio
PCA	μ	98.69@17	90.50@26	94.64@24	22.33
	σ	99.14@16	91.68@27	90.16@26	23.00
	σ_b	99.30@15	93.38@25	92.70@26	22.00
	E_0	98.21@22	90.96@17	84.82@24	21.00
	E_1	98.55@17	85.01@17	85.18@24	19.33
	E_2	99.33@18	92.99@18	92.87@@26	20.67
	E_3	98.92@17	94.21@26	94.08@26	23.00
	E_4	99.29@18	93.11@18	88.29@25	20.33
	E_5	87.95@22	73.71@19	72.71@20	20.33
	Promedio	18.00	21.44	24.56	21.33
AVR	μ	99.17@12	93.31@16	96.05@18	15.33
	σ	99.32@14	93.31@11	89.60@25	16.67
	σ_b	99.31@11	94.98@11	92.49@22	14.67
	E_0	98.37@12	92.44@22	85.82@14	16.00
	E_1	98.92@13	89.62@14	86.51@21	16.00
	E_2	99.34@11	93.99@13	92.29@19	14.33
	E_3	99.45@10	94.66@13	94.11@19	14.00
	E_4	99.21@12	93.50@23	87.64@16	17.00
	E_5	91.46@14	79.40@23	78.51@21	19.33
	Promedio	12.11	16.22	19.44	15.93
ICV	μ	99.41@18	93.50@19	96.51@18	18.33
	σ	99.59@25	93.95@24	90.22@29	26.00
	σ_b	99.72@24	95.05@17	93.25@27	22.67
	E_0	98.63@15	92.60@16	86.53@17	16.00
	E_1	99.13@20	89.39@21	87.07@28	23.00
	E_2	99.65@22	94.36@22	92.18@21	21.67
	E_3	99.66@12	95.18@18	94.05@24	18.00
	E_4	99.35@14	93.24@16	88.04@23	17.67
	E_5	92.86@18	79.67@18	78.85@23	22.33
	Promedio	18.67	19.00	23.33	20.33
Global	16.26	18.89	22.44	19.20	

Tabla 6.3.: Efecto principal de la medida de distancia. En la Tabla se muestran los resultados de la CCR (%) dependiendo de la medida de distancia utilizada. El reporte incluye la evaluación para cada tipo de característica de textura, para las tres bases de datos utilizadas y para cada método de reducción de dimensión. Los valores resaltados indican los mayores valores de CCRs para cada tipo de característica respecto a la medida de distancia empleada.

		μ	σ	σ_b	E_0	E_1	E_2	E_3	E_4	E_5
Brodatz										
PCA	d_C	94.97	94.18	95.10	94.72	94.83	94.78	93.21	96.01	73.15
	d_E	96.00	95.77	96.13	91.90	95.51	96.18	95.36	95.88	81.52
	d_{E_N}	95.99	95.70	96.07	95.27	95.67	96.11	95.20	96.71	80.07
AVR	d_C	96.75	95.90	96.25	96.54	96.60	96.60	96.84	96.58	86.86
	d_E	96.23	96.06	96.35	93.67	95.76	96.64	96.98	95.84	84.61
	d_{E_N}	96.84	95.96	96.27	96.42	96.51	96.61	96.86	96.59	85.34
ICV	d_C	96.89	94.80	96.67	96.35	96.43	96.97	96.92	96.62	82.96
	d_E	96.45	95.02	96.78	93.65	95.68	97.05	97.07	96.10	80.90
	d_{E_N}	96.89	94.88	96.65	96.25	96.32	97.01	96.94	96.65	81.44
VisTex										
PCA	d_C	81.00	76.65	82.16	86.48	77.58	79.38	81.36	86.42	58.98
	d_E	84.52	82.47	86.93	82.60	80.07	84.15	86.32	83.98	70.06
	d_{E_N}	84.61	81.81	87.04	87.78	80.26	83.80	86.43	87.88	67.65
AVR	d_C	88.58	88.77	90.72	89.91	87.51	90.30	90.22	88.94	71.76
	d_E	84.92	88.20	89.02	83.03	78.32	87.56	89.92	86.04	69.35
	d_{E_N}	89.26	88.59	90.41	88.56	86.60	89.88	90.00	88.72	68.28
ICV	d_C	86.78	86.55	88.38	88.90	84.58	87.78	89.05	88.85	69.86
	d_E	84.46	85.90	87.21	83.50	77.43	85.97	88.80	86.46	67.83
	d_{E_N}	87.29	86.13	87.81	88.53	83.56	87.14	88.75	88.65	66.40
Outex										
PCA	d_C	87.60	77.35	81.56	79.70	78.67	83.57	81.72	81.14	58.57
	d_E	89.37	81.53	85.53	72.76	79.08	85.83	86.06	77.02	67.10
	d_{E_N}	89.61	81.49	85.15	81.27	80.28	86.16	85.92	82.24	66.06
AVR	d_C	90.27	77.18	84.03	84.61	81.29	85.44	87.60	84.20	68.66
	d_E	88.15	78.63	84.63	75.10	77.37	85.32	88.09	78.04	65.70
	d_{E_N}	91.09	77.68	83.94	84.05	81.34	85.54	87.66	83.88	68.74
ICV	d_C	91.46	79.47	84.31	84.21	82.72	87.20	88.45	84.12	65.77
	d_E	89.49	81.11	84.83	75.52	79.25	87.54	89.28	79.59	63.58
	d_{E_N}	91.44	80.23	84.28	83.51	82.48	87.28	88.57	83.80	65.62

mayor de elementos del vector de características que AVR. La CCR (%) obtenida con ICV en la mayoría de los casos fue mejor que la obtenida con AVR mientras que las CCRs (%) obtenidas con PCA quedaron rezagadas si son comparadas con las obtenidas con AVR e ICV.

Utilizando todos los tipos de características y las tres bases de datos se encontró que en promedio el mínimo número de elementos del vector de características es 21.33, 15.93 y 20.33 para PCA, AVR e ICV respectivamente mientras que el promedio global fue de 19.20 características (Véase la Tabla 6.2). Es decir, en promedio se requiere que el vector de características esté formado por 20 elementos. También fue posible observar que para la base de datos Outex se requieren más elementos del vector de características seguido por la base de datos VisTex y por último la base de datos Brodatz es la que requiere menos elementos en el vector de características.

Con respecto al factor distancia d_x se puede observar a partir de la Tabla 6.3 que para las tres bases de datos las distancias más competitivas son d_C y d_{E_N} . Sin embargo, estos resultados son insuficientes aún y no es posible determinar que distancia es mejor debido a que se observa que hay cierta dependencia del tipo de característica utilizado. Los análisis de interacción de los factores, los cuales se reportan en la sección 6.1.3, permitieron responder a esta pregunta.

6.1.3. Interacción de los factores utilizando un vector de dimensión reducida

Utilizando vectores de características conformados por 20 elementos, los cuales fueron obtenidos al ser seleccionados a través de los métodos de reducción de dimensión (PCA, AVR e ICV), se evaluó la interacción estadística de los 4 factores: medidas de distancia, tipos de características, base de datos y del método de reducción de dimensión empleado. El primer resultado obtenido a partir de una prueba ANOVA de 4 vías indicó que los efectos principales de los 4 factores son estadísticamente significativos ($p < 0.05$) en términos del promedio de la tasa de clasificación correcta, CCR(%). Así mismo, se encontró que todas las posibles combinaciones de interacción entre estos factores también son estadísticamente significativas ($p < 0.05$). La interpretación gráfica de la interacciones de los promedios de las tasas de clasificación se presentan a continuación.

Interacción método-distancia

Las mejores tasas de clasificación promedio se obtuvieron al utilizar el método AVR, seguido por ICV y en último lugar PCA. La distancia d_C produjo las mejores tasas de clasificación con los métodos AVR e ICV pero no para PCA. La distancia d_{E_N} generó mejores tasas de clasificación que la distancia d_E en todos los casos. En la figura 6.4(a) se muestra la interacción entre estos factores. La interacción entre el método de reducción de dimensión AVR y la distancia Canberra (d_C) generó la mejor CCR promedio.

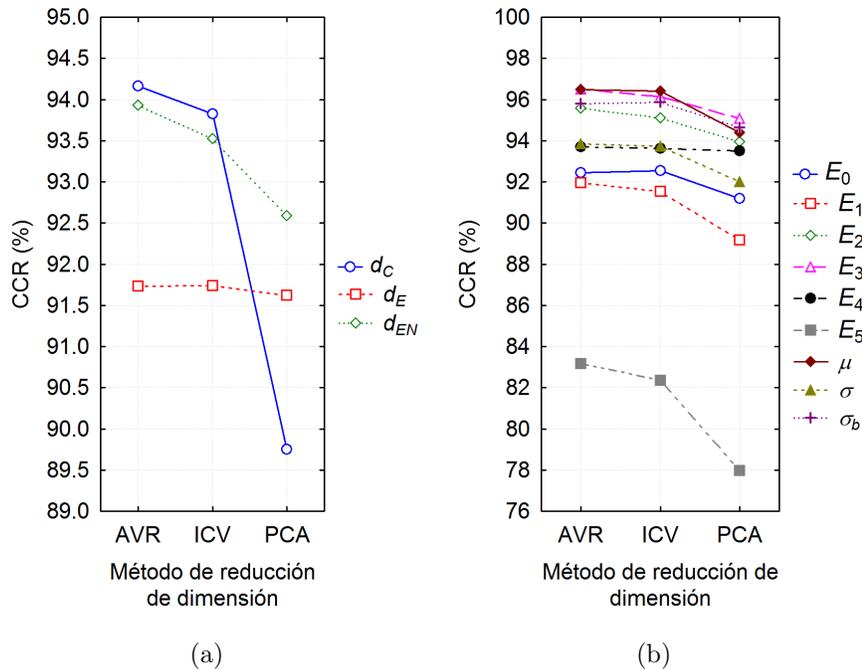


Figura 6.4.: Interacciones en términos de la tasa de clasificación correcta promedio, CCR(%). (a) Interacción entre el método de reducción de dimensión y la medida de distancia. (b) Interacción entre el método de reducción de dimensión y el tipo de característica. En ambos casos (a) y (b) las mejores tasas de clasificación se obtuvieron al utilizar el método AVR, seguido por ICV y en último lugar PCA. (a) La distancia d_C produjo las mejores tasas de clasificación con los métodos AVR e ICV pero no para PCA. La distancia d_{E_N} generó mejores tasas de clasificación que la distancia d_E en todos los casos. La interacción entre AVR y la distancia d_C generó la mejor CCR promedio. (b) Las mejores tasas de clasificación correcta fueron obtenidas con μ , E_3 , σ_b y E_2 . Las características E_1 y E_5 produjeron las tasas de clasificación más bajas. La interacción entre AVR y μ o bien entre AVR y E_3 generó la mejor CCR promedio.

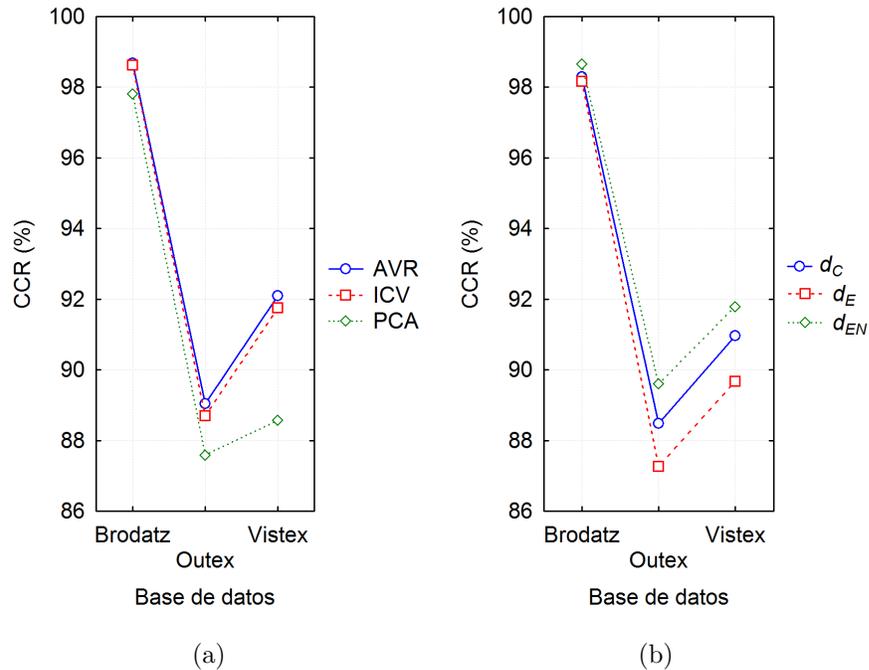


Figura 6.5.: Interacciones en términos de la tasa de clasificación correcta promedio, CCR(%). (a) Interacción entre la base de datos y el método de reducción de dimensión. (b) Interacción entre la base de datos y la medida de distancia. En ambos casos (a) y (b) se puede observar que la base de datos que presenta mayor complejidad para obtener CCRs altas fue Outex en contraste con la base de datos Brodatz, que es la de menor complejidad. (a) Las mejores tasas de clasificación se obtuvieron al utilizar el método AVR, seguido por ICV y en último lugar PCA. (b) Las mejores tasas de clasificación se obtuvieron al utilizar la distancia d_{EN} , seguido por d_C y en último lugar d_E .

Interacción método-característica

Las mejores tasas de clasificación promedio se obtuvieron al utilizar el método AVR, seguido por ICV y en último lugar PCA. Los 4 tipos de características que arrojaron las mejores tasas de clasificación (superiores a 93.94%) fueron μ , E_3 , σ_b y E_2 . Las siguientes 3 características con las cuales se obtuvo una CCR alta pero inferior a la obtenida con los tipos de características mencionados anteriormente fueron E_4 , σ y E_0 (con CCRs superiores a 91.19%), mientras que usando los tipos de características E_1 y E_5 las tasas de clasificación que se obtuvieron fueron las más bajas. La interacción entre el método de reducción de dimensión AVR y μ o bien entre AVR y E_3 generó en promedio la mejor CCR (aproximadamente 96.50%). En la figura 6.4(b) se muestran las interacciones entre estos factores.

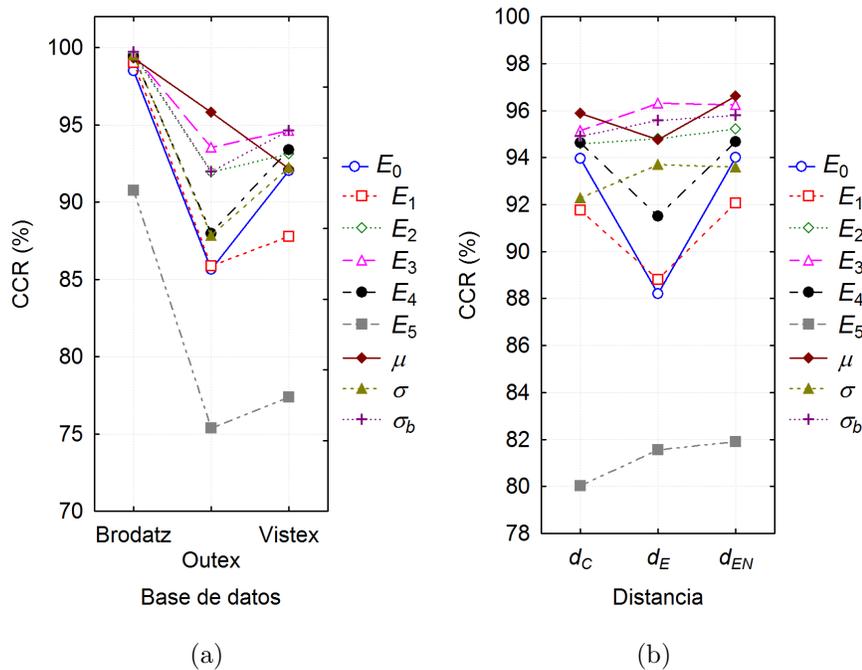


Figura 6.6.: Interacción entre los tipos de características y la base de datos (a) e interacción entre los tipos de características y la medida de distancia (b) en términos de la tasa de clasificación correcta promedio, CCR(%). (a) La interacción con la cual se observaron las mejores tasas de clasificación fue con la base de datos Brodatz y con todos los tipos de características excepto con E_5 . (b) La interacción con la cual se observó la mejor tasa de clasificación promedio fue con la característica μ y la distancia d_{EN} .

Interacción base de datos-método

Los tipos de texturas y la forma como éstas fueron adquiridas puede influir en la clasificación de las texturas. Para las tres bases de datos se encontró que al emplear el método AVR durante la selección de los elementos del vector de características se obtuvo el mejor desempeño en términos del promedio de la tasa de clasificación correcta, CCR(%). El siguiente método con el que se obtuvieron CCRs altas compitiendo con AVR fue ICV. Por último, el método PCA no logró ser más robusto que los otros dos. Se observó que la base de datos que presenta mayor complejidad para obtener CCRs altas fue Outex en contraste con la base de datos Brodatz, que es la de menor complejidad. En la figura 6.5(a) se muestran las interacciones entre la base de datos y el método de reducción de dimensión.

Interacción base de datos-distancia

Para las tres bases de datos se encontró que al emplear la distancia d_{EN} se obtiene el mejor desempeño en términos del promedio de la tasa de clasificación correcta, CCR(%). La siguiente distancia con la que se obtuvieron CCRs altas compitiendo con d_{EN} fue d_C . La

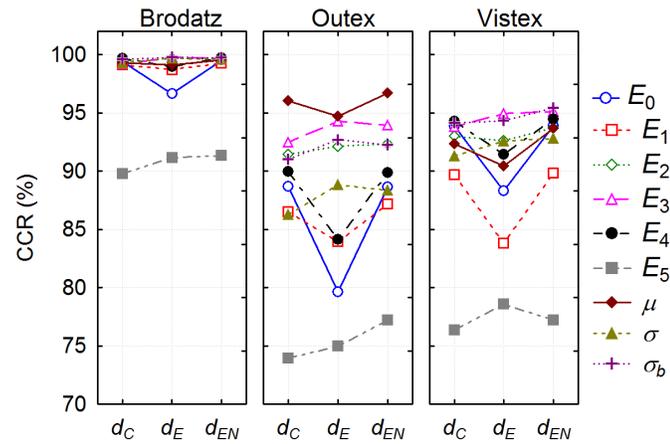


Figura 6.7.: Interacción entre la medida de distancia, los tipos de características y la base de datos de prueba en términos de la tasa promedio de clasificación correcta, CCR(%).

distancia que generó menor desempeño en la CCR fue la distancia d_E . En la figura 6.5(b) se muestran la interacciones entre la base de datos empleada y el tipo de distancia.

Interacción característica-base de datos

Para la base de datos Brodatz se obtuvieron buenas tasas de clasificación (superiores al 98.5 %) con cualquier tipo de característica de textura excepto al usar E_5 . Con E_5 se obtuvo la CCR más baja incluso de manera general al probar con cualquiera de las tres bases de datos. Al probar con la base de datos Outex, los 4 tipos de características que permitieron obtener las mejores CCRs (superiores a 91.98 %) fueron μ , E_3 , E_2 y σ_b . Las siguientes 4 características que produjeron CCRs por debajo del 88 % fueron E_4 , σ , E_1 y E_0 . Con respecto a la base de datos VisTex, las 4 características que produjeron las CCRs más altas (superiores a 93.11 %) fueron σ_b , E_3 , E_4 y E_2 mientras que μ , σ , E_0 y E_1 produjeron las tasas de clasificación más bajas, por debajo del 92.2 %. En la figura 6.6(a) se muestran la interacciones entre los tipos de característica y la base de datos empleada.

Interacción característica-distancia

En la mayoría de los casos, la distancia d_{EN} presentó mejor o similar desempeño que sus competidoras d_C y d_E . Al comparar las CCRs de la interacción de los factores se observó que existe una coincidencia para d_C y d_{EN} respecto a los mejores tipos de características (en orden descendiente de CCR) μ , E_3 , σ_b , E_2 , E_4 , E_0 , σ , E_1 y E_5 . Para la distancia d_E se observó la siguiente tendencia (en orden descendiente de CCR): E_3 , σ_b , E_2 , μ , σ , E_4 , E_1 , E_0 y E_5 . En la figura 6.6(b) se muestran la interacciones entre los tipos de características y la distancia empleada.

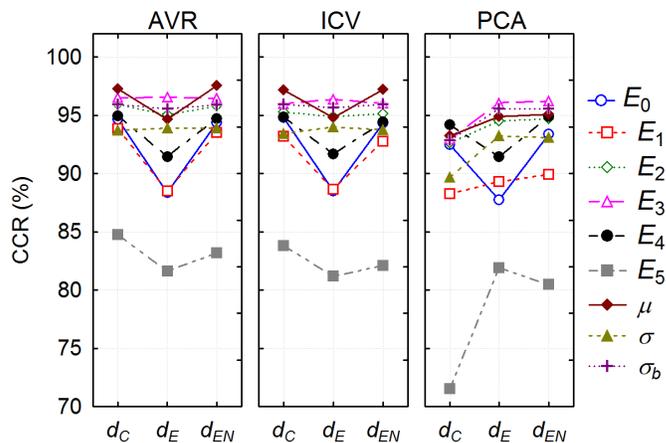


Figura 6.8.: Interacción entre la medida de distancia, los tipos de características y el método de reducción de dimensión en términos del promedio de la tasa de clasificación correcta, CCR(%).

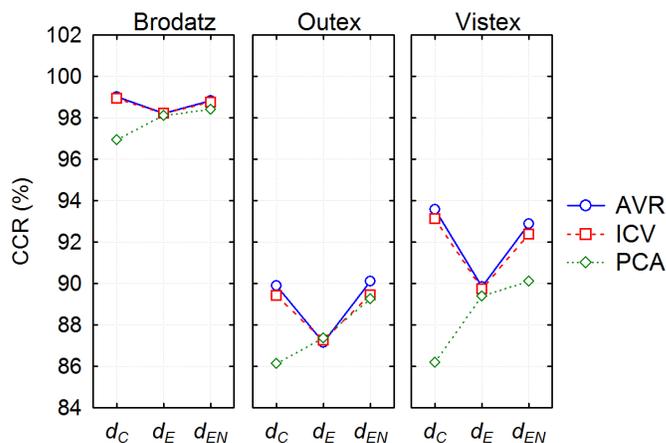


Figura 6.9.: Interacción entre la medida de distancia, el método de reducción de dimensión y la base de datos de prueba en términos del promedio de la tasa de clasificación correcta, CCR(%).

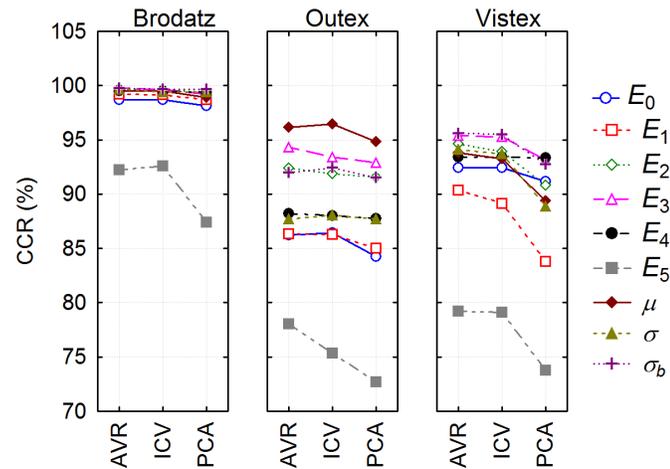


Figura 6.10.: Interacción entre los tipos de características, el método de reducción de dimensión y la base de datos de prueba en términos del promedio de la tasa de clasificación correcta, CCR(%).

Interacción distancia-característica-base de datos

La mayoría de las interacciones entre la distancia d_{E_N} y cualquiera de los tipos de características manifestaron una tendencia a producir mejores tasas de clasificación que al usar d_C para las tres bases de datos de pruebas. En algunos casos la interacción con d_E fue ligeramente mejor que con d_{E_N} . La interacción entre las distancias, la mayoría de los tipos de características y la base de datos Brodatz produjeron las mejores CCRs promedio. En la figura 6.7 se muestran estas interacciones.

Interacción distancia-característica-método

Las interacciones de los métodos de reducción de dimensión AVR e ICV con los tipos de características μ , E_3 , σ_b , E_2 y las distancias d_C y d_{E_N} produjeron en general buenas tasas de clasificación superiores al 95.15%. Esta tasa de clasificación fue similar a las interacciones entre PCA, μ , E_3 , σ_b , E_2 y las distancias d_E y d_{E_N} . La figura 6.8 muestra las interacciones entre la distancia empleada, los tipos de características y los métodos de reducción de dimensión.

Interacción distancia-método-base de datos

En general, las interacciones de los métodos de reducción de dimensión AVR e ICV con las distancias d_C y d_{E_N} y las tres bases de datos de prueba produjeron las mayores CCRs. La figura 6.9 muestra las interacciones entre la distancia empleada, los métodos de reducción de dimensión y la base de datos de prueba. Las interacciones de todas las distancias, el método de reducción de dimensión PCA y las tres bases de datos produjeron las menores CCRs excepto en la interacción entre la base de datos Outex, la distancia d_E y el método PCA.

Interacción característica-método-base de datos

En esta interacción de factores se muestra la robustez que presentan los métodos de reducción de dimensión AVR e ICV respecto a PCA, con cualquier combinación de característica y base de datos empleada, véase la figura 6.10. En el caso de las bases de datos Outex y VisTex las características que propician las mejores tasas de clasificación fueron μ , E_3 , σ_b y E_2 en combinación con cualquier método, AVR o bien ICV. En el caso de la base de datos VisTex las características σ y E_4 también produjeron buenas CCRs promedio. En la mayoría de los casos las características E_5 , E_1 y E_0 generaron las CCRs más bajas.

6.1.4. Discusión de resultados

Se efectuó un análisis de las interacciones entre los cuatro factores: método de reducción de dimensión, tipos de características, medidas de distancias y las bases de datos empleadas. Este análisis permitió observar que para las tres bases de datos los métodos AVR e ICV fueron más robustos y proporcionaron buenas tasas de clasificación en comparación con el método PCA; al utilizar la distancia d_{E_N} o bien d_C en conjunto con las características μ , E_3 , σ_b y E_2 .

El tipo de característica E_4 también presentó en diversas ocasiones tasas de clasificación altas, pero no las mejores, en particular cuando se combinó con las distancias d_C o bien con d_{E_N} y con cualquier método de reducción de dimensión.

En la mayoría de los casos los tipos de características y su interacción con cualquier medida de distancia y método de reducción de dimensión que proporcionaron las tasas de clasificación más bajas fueron E_5 , E_1 , E_0 y en algunas ocasiones σ . Sin embargo, σ llegó a presentar tasas de clasificación altas, pero no las mejores, en combinación con la distancia d_E y en combinación con cualquier método de reducción de dimensión y base de datos.

La base de datos Brodatz es una base de datos que presenta un nivel de dificultad baja durante la clasificación, pues las clases de texturas parecen estar bien delimitadas ya que en la mayoría de los casos con cualquier combinación del tipo de característica con el método de reducción de dimensión y medida de distancia se obtuvieron tasas de clasificación altas. Respecto a la base de datos VisTex podría catalogarse como una base de datos de complejidad media donde las clases de texturas están bien delimitadas pero existen algunas clases de texturas que probablemente son muy parecidas entre sí y provocan que la clasificación no sea la correcta. Por último, la base de datos Outex contiene clases de texturas con las cuales la dificultad es mayor para obtener tasas de clasificación altas.

6.1.5. Sensibilidad del sistema usando la transformada de Hermite

A partir de las distribuciones de distancia (d_x) intra-clase e inter-clase (Ecuaciones (5.17), (5.19)) es posible obtener el desempeño de las clasificación para las combinaciones de los tipos de características, distancias y los métodos de reducción de dimensión presentados anteriormente.

Usando los resultados presentados en la Sec. 6.1.2 se restringió el máximo número de

elementos que conforman al vector de características en 20. En la Fig. 6.11 se muestran las curvas ROC para la base datos Brodatz para todos los tipos de características, distancias y los métodos de reducción de dimensión. De manera similar las Figs. 6.12 y 6.13 muestran las curvas ROC para la base de datos VisTex y Outex respectivamente. En todos los casos se muestra además el valor de la sensibilidad S para una tasa de error igual $EER = FAR = FRR$, como se explicó en la sección 5.5.1, donde $FAR = 1 - E$ y E es la especificidad del modelo.

Para el caso de la sensibilidad al utilizar la base de datos de texturas Brodatz (Fig. 6.11), se puede observar que la sensibilidad S más alta (o bien la tasa de error igual (EER) más baja) se obtiene con los tipos de características μ , E_3 , σ_b y E_2 y el método de reducción de dimensión AVR, seguido por ICV y por último PCA.

La sensibilidad al utilizar la base de datos de texturas VisTex (Fig. 6.12) presentó un comportamiento similar al observado con la base de datos de texturas Brodatz. La sensibilidad S más alta (o bien la tasa de error igual (EER) más baja) se obtiene con los tipos de características μ , E_3 , σ_b y E_2 y el método de reducción de dimensión AVR, seguido por ICV y por último PCA.

Contrario a lo observado para las bases de datos de texturas Brodatz y VisTex, al utilizar la base de datos de texturas Outex (Fig. 6.13) se observó que se obtiene la sensibilidad S más alta (o bien la tasa de error igual (EER) más baja) con el método PCA seguido por AVR y por último ICV. Los tipos de características que permiten obtener una S alta coincidió con los resultados previos, siendo μ , E_3 , σ_b y E_2 dichos tipos de características

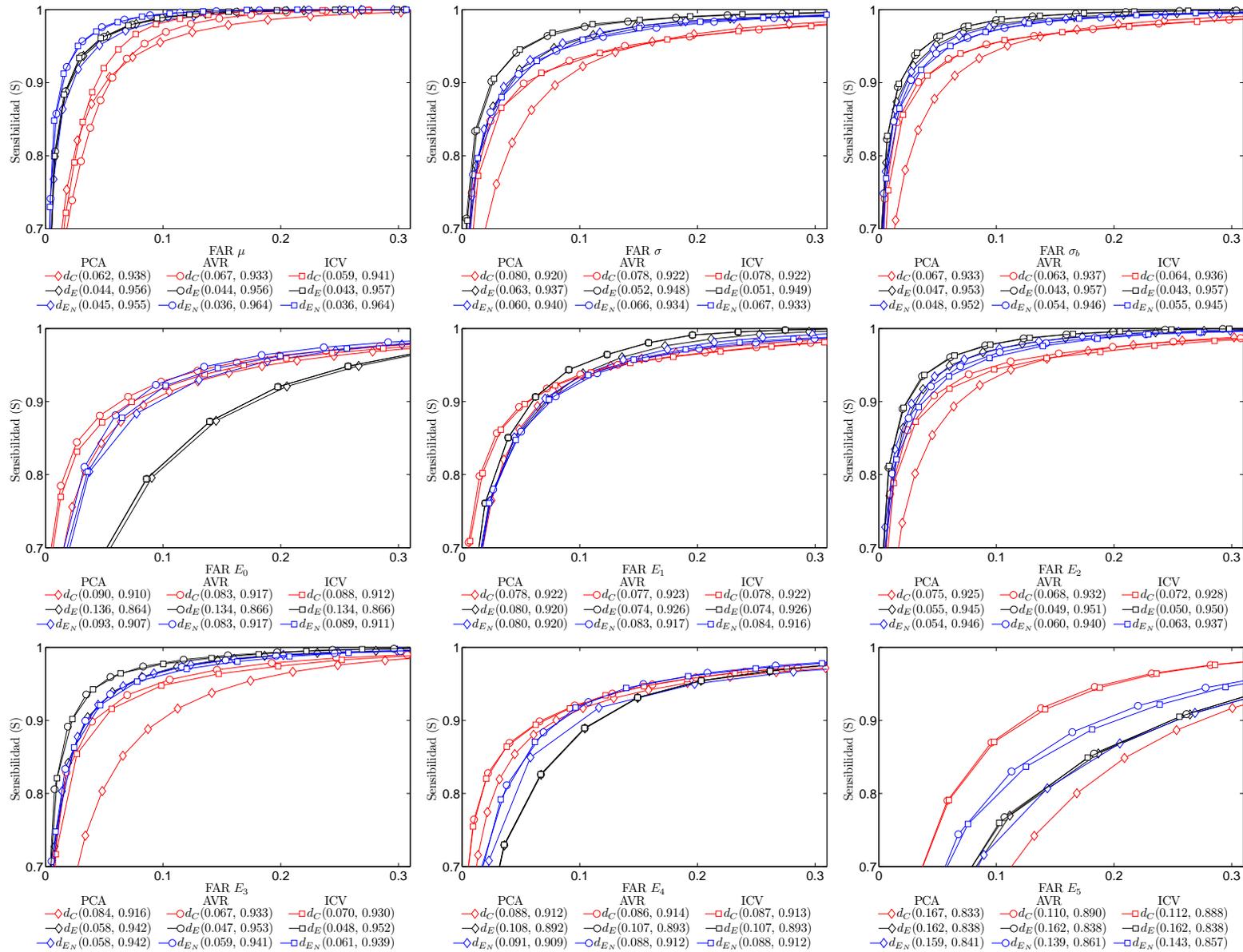


Figura 6.11.: Comparación de curvas ROC usando la transformada de Hermite para la base de datos Brodatz [8]. Cada figura representa la gráfica de la sensibilidad (S) con respecto al error FAR. Para cada método de reducción de dimensión se reporta S y FAR obtenido con las medidas de distancia.

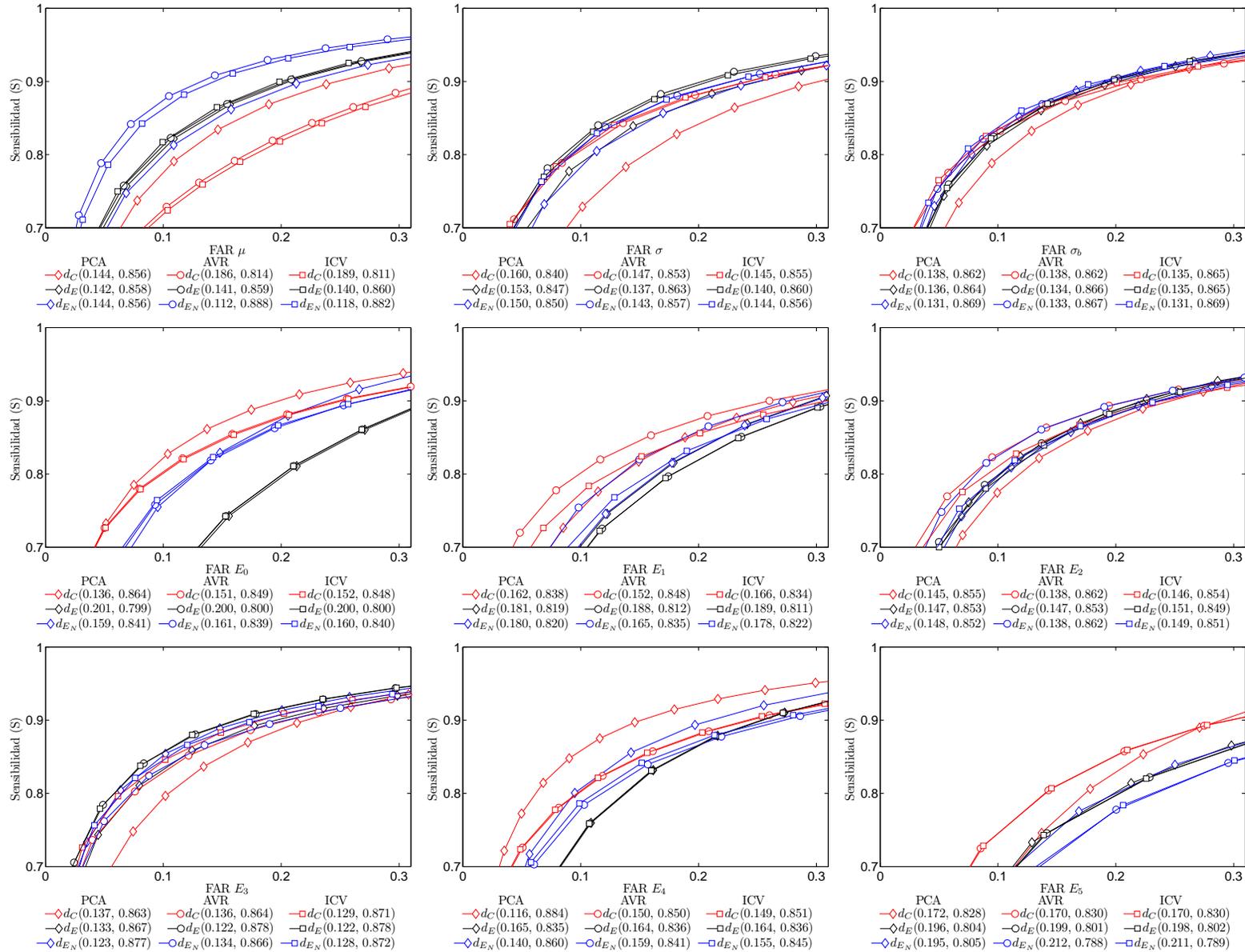


Figura 6.12.: Comparación de curvas ROC usando la transformada de Hermite para la base de datos VisTex [80]. Cada figura representa la gráfica de la sensibilidad (S) con respecto al error FAR. Para cada método de reducción de dimensión se reporta S y FAR obtenido con las medidas de distancia.

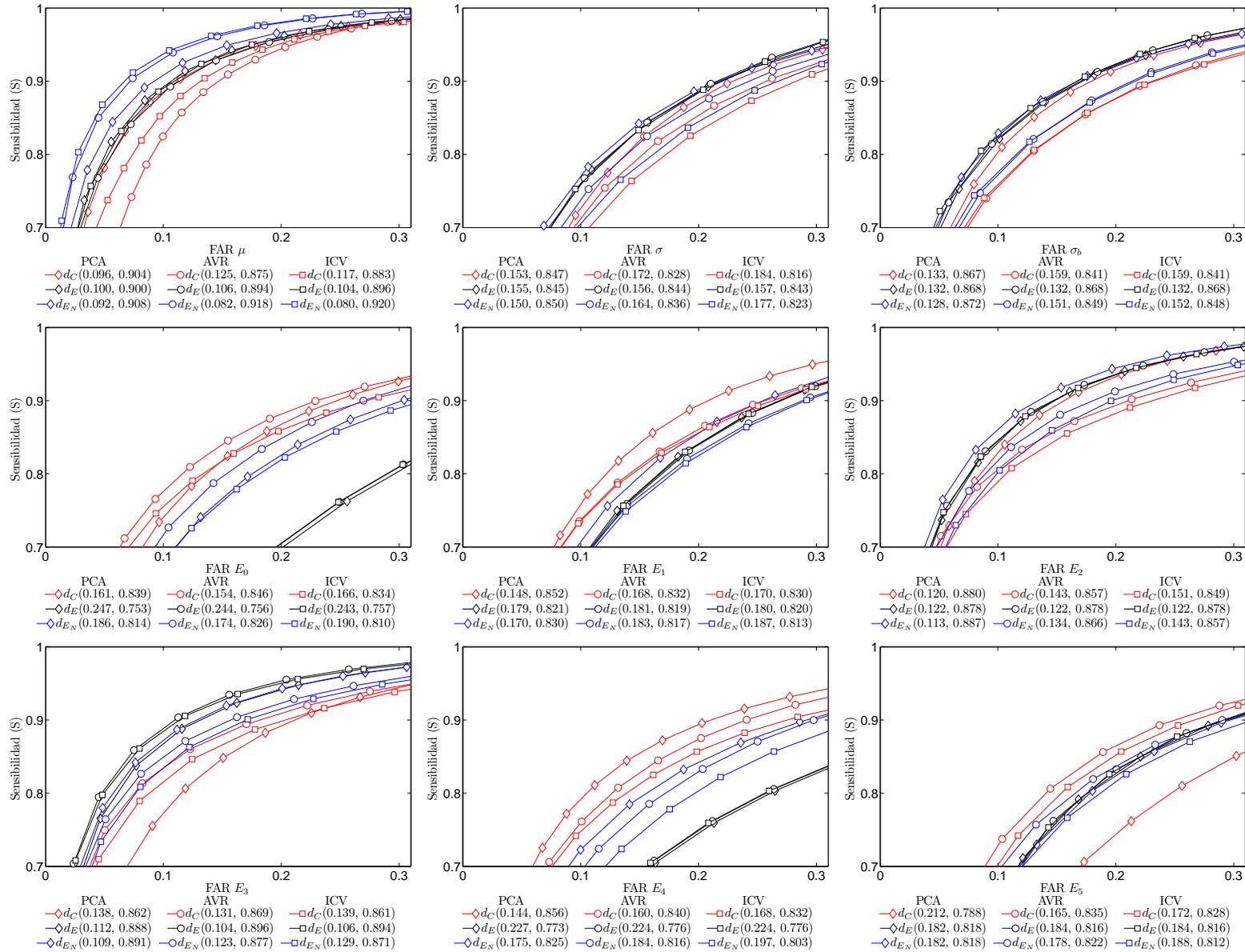


Figura 6.13.: Comparación de curvas ROC usando la transformada de Hermite para la base de datos Outex [66, 67]. Cada figura representa la gráfica de la sensibilidad (S) con respecto al error FAR. Para cada método de reducción de dimensión se reporta S y FAR obtenido con las medidas de distancia.

6.1.6. Determinación del orden promedio de los filtros que conforman los vectores de 20 elementos

A partir de las repeticiones experimentales se calculó la frecuencia con la cual se seleccionaban los índices de las características más importantes utilizando los métodos AVR e ICV. Estas frecuencias originaron los histogramas presentados en las siguientes figuras. Para cada tipo de característica se calculó su histograma de frecuencias. Al seleccionar las 20 frecuencias más altas y eliminar las 12 restantes (recordando que el vector de características es de 32 elementos) se seleccionaron los índices que en promedio para todos los experimentos realizados aportaron más de acuerdo a la selección generada por AVR o ICV. Existe una relación directa entre el índice que ocupan los elementos que conforman al vector de características y el orden de los filtros necesarios para su obtención. Cada elemento del vector de características fue colocado en el índice $i = (n - 1) + (s - 1)(n_{max})$, donde $n = 1, \dots, 8$ es el orden del filtro ($n_{max} = 8$) y $s = 1, \dots, S$ la escala de análisis con $S = 4$, que resulta en un vector de dimensión máxima de 32 elementos. De esta manera, el índice $i = 11$ del vector de características representa al filtro de análisis de Hermite de orden $n = 4$ y escala de análisis $s = 2$.

Las Figs. 6.14-6.15 representan a los índices utilizando los métodos de reducción de dimensión AVR e ICV respectivamente para la base de datos Brodatz. De manera similar las Figs. 6.16-6.17 para la base de datos VisTex mientras que las Figs. 6.18-6.19 para la base de datos Outex. En las gráficas se muestra únicamente los índices de las escalas de análisis; en cada escala de análisis existen $n_{max} = 8$ índices que representan los órdenes de los filtros más sobresalientes.

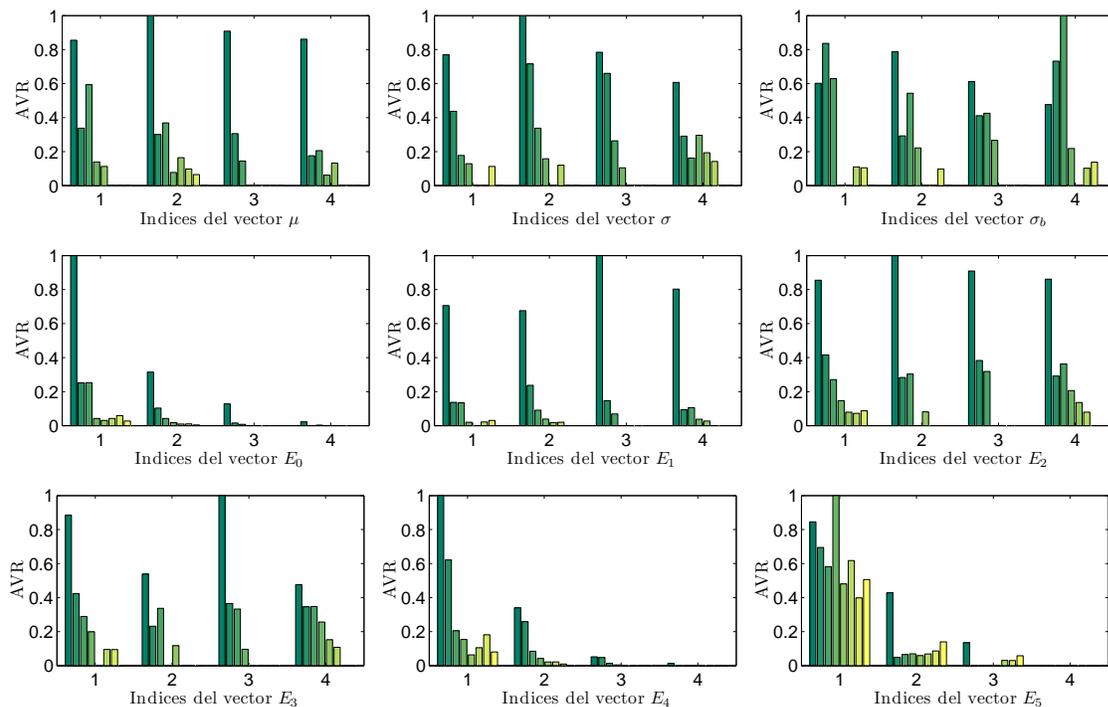


Figura 6.14.: Selección de los índices correspondientes a los 20 filtros más sobresalientes usando AVR para la base de datos Brodatz

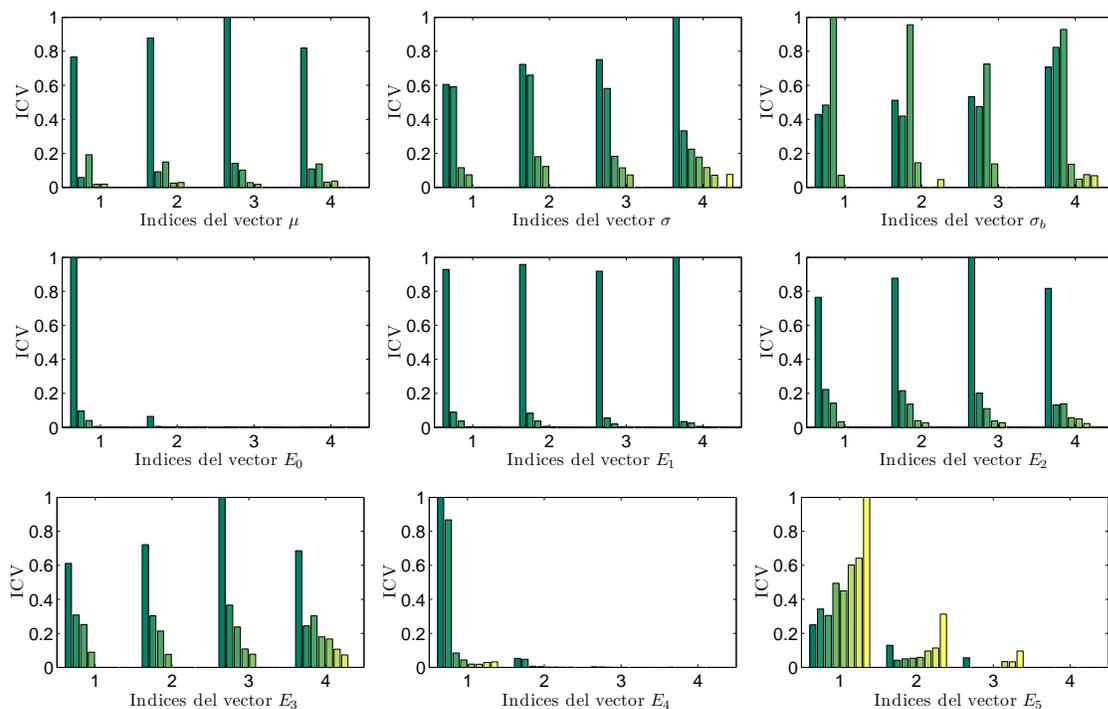


Figura 6.15.: Selección de los índices correspondientes a los 20 filtros más sobresalientes usando ICV para la base de datos Brodatz

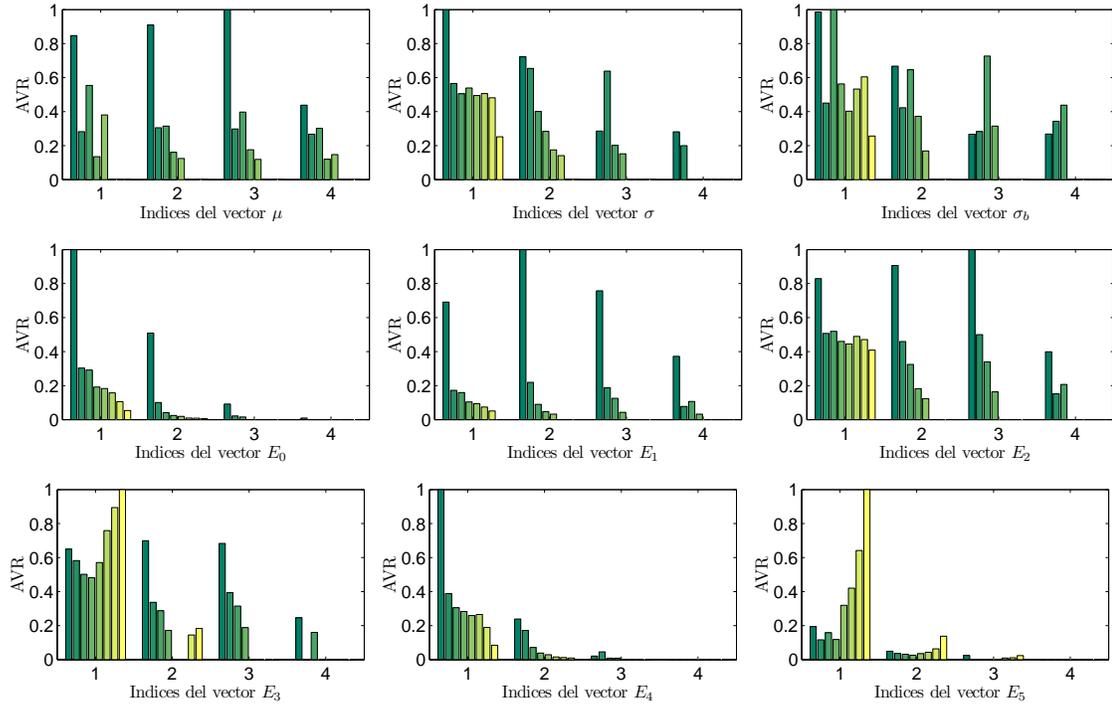


Figura 6.16.: Selección de los índices correspondientes a los 20 filtros más sobresalientes usando AVR para la base de datos VisTex

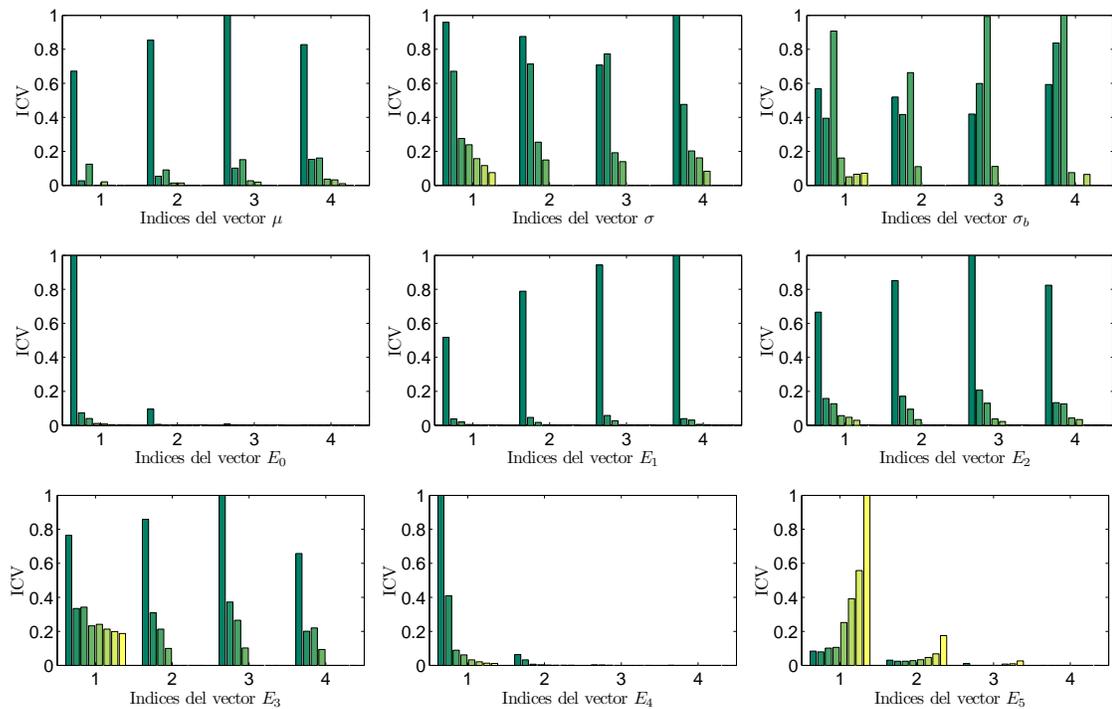


Figura 6.17.: Selección de los índices correspondientes a los 20 filtros más sobresalientes usando ICV para la base de datos VisTex

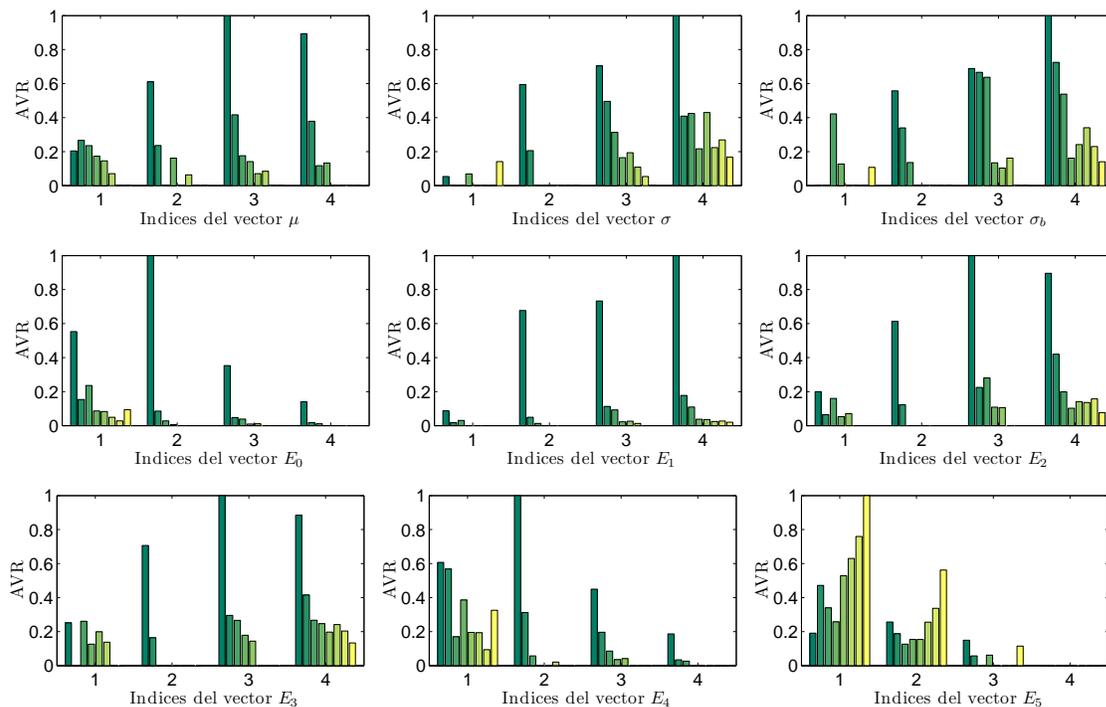


Figura 6.18.: Selección de los índices correspondientes a los 20 filtros más sobresalientes usando AVR para la base de datos Outex

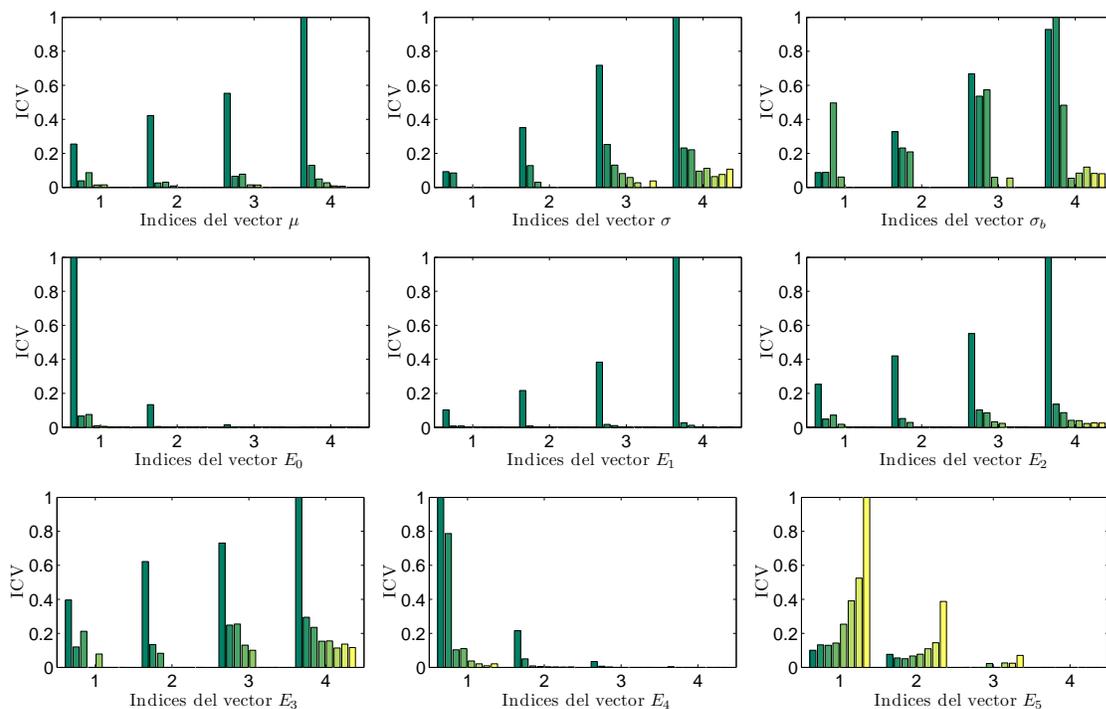


Figura 6.19.: Selección de los índices correspondientes a los 20 filtros más sobresalientes usando ICV para la base de datos Outex

6.2. Resultados usando los filtros Gabor

6.2.1. Diseño de bancos de filtros Gabor

En el diseño de los bancos de filtros Gabor se empleó la definición presentada en la Sec. 4.2. Los bancos de filtros Gabor propuestos han sido diseñados en el dominio frecuencial con $S = 4$ escalas de análisis empezando desde una frecuencia central de 0.039 y terminando con una frecuencia máxima de 0.312 cpp. La separación entre las frecuencias centrales es de una octava.

Debido a que el primer interés de este trabajo con respecto a los filtros Gabor es encontrar los efectos de los anchos de banda así como también del número de orientaciones sobre la clasificación, todas las frecuencias centrales se fijaron en el mismo valor para todos los bancos de filtros. Para investigar las distintas combinaciones de traslape radial y angular de la mitad de la amplitud del ancho de banda frecuencial se diseñaron bancos de filtros Gabor para tres anchos de banda $\phi = 1, 1.5$ y 2 octavas. Para cada ancho de banda se han propuesto tres configuraciones con respecto al número de orientaciones $K = 6, 8$ y 10. Estos bancos de filtros permiten tener nueve diseños diferentes como se muestra en la Fig. 6.20.

El resultado de este diseño, como se puede apreciar en la Fig. 6.20, es una serie de bancos de filtros Gabor que incluyen distintas configuraciones de traslape entre las respuestas en frecuencia de cada filtro individual.

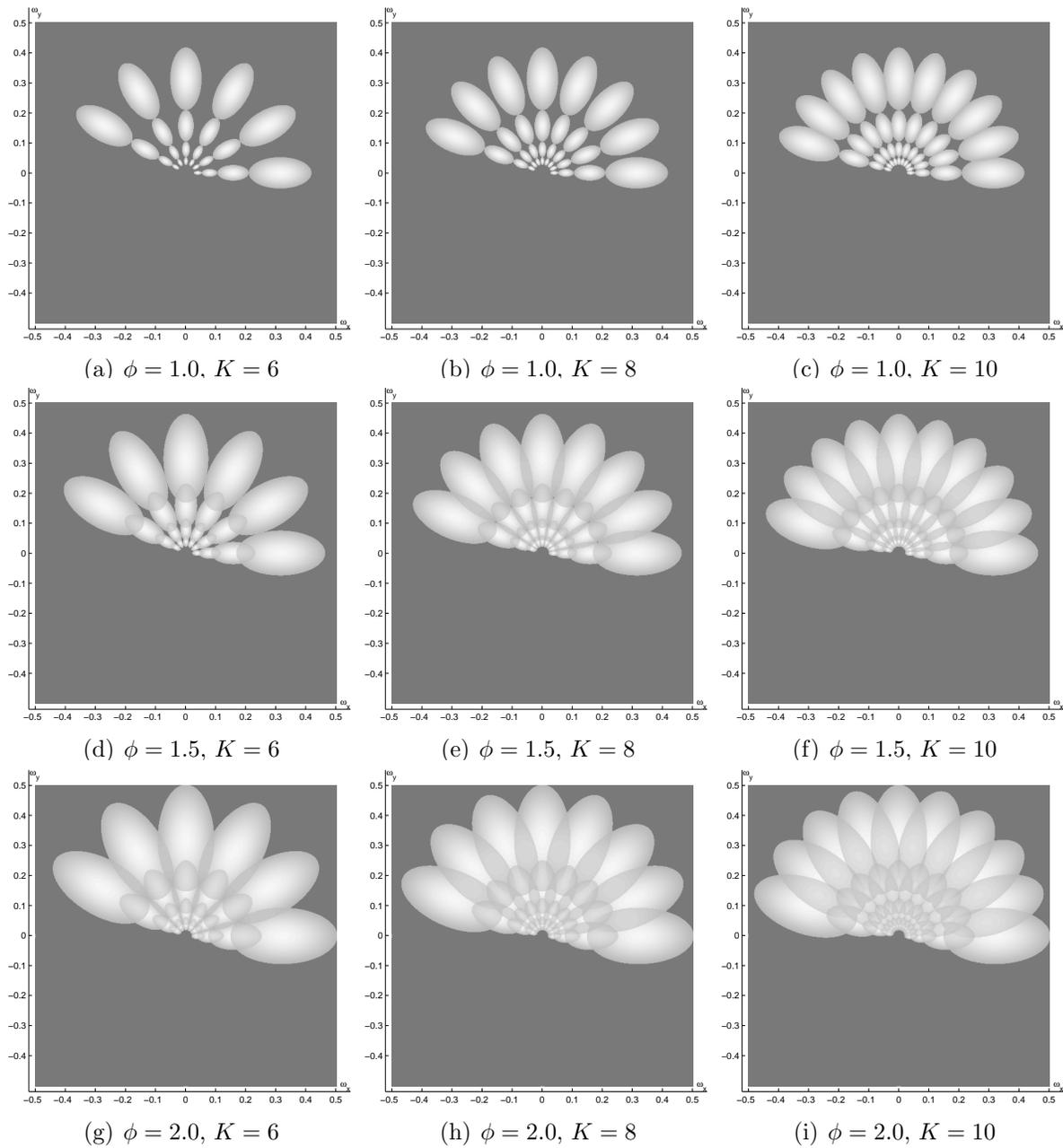


Figura 6.20.: Bancos de filtros Gabor utilizados en los experimentos. Cada banco de filtros se generó con $S = 4$ escalas de análisis. Las frecuencias centrales más baja y más alta es de 0.039 y 0.312 cpp respectivamente. La separación entre frecuencias centrales es de una octava. Cada banco de filtros fue generado al combinar 3 números de orientaciones ($K = 6, 8, 10$ orientaciones) y 3 anchos de banda ($\phi = 1.0, 1.5$ y 2.0).

6.2.2. Determinación experimental de la combinación del ancho de banda ϕ y del número de orientaciones K

Efectos principales

Cada tipo de característica de textura fue evaluado de manera separada usando una prueba ANOVA de 3 vías con los siguientes factores: el ancho de banda frecuencial radial ϕ , el número de orientaciones K y el tipo de distancia utilizado, d_E , d_C y d_{E_N} (Ecuaciones (3.15), (3.16) y (3.18) respectivamente). Se observó que los principales efectos de los tres factores fueron estadísticamente significativos ($p < 0.05$) para las tres bases de datos y para cada tipo de característica de textura.

El principal efecto del factor ϕ que se observó es que en promedio con un ancho de banda frecuencial radial de $\phi = 2.0$ las tasas de clasificación CCRs(%) son mejores con respecto a un ancho de banda frecuencial radial de $\phi = 1.0$. Usando $\phi = 1.5$ las CCRs también son las mejores con algunos tipos de características de texturas o bien estadísticamente similares a las obtenidas con $\phi = 2.0$.

Con respecto al factor K se observó que aumentar desde $K = 6$ hasta $K = 10$ orientaciones tiene un efecto positivo en el incremento de las tasas de clasificación. En algunos casos $K = 8$ es suficiente para obtener tasas de clasificación estadísticamente iguales o superiores que cuando se usa $K = 10$.

El principal efecto del factor d_x para todos los casos en las tres bases de datos utilizadas se encontró que la distancia d_{E_N} generó las mejores tasas estadísticamente significativas de clasificación excepto cuando se usó la característica E_5 en donde para las tres bases de datos de texturas se observó que la d_E se desempeñó mejor.

Interacción de los factores

La evaluación de la interacción entre los factores ϕ , K y d_x indicó que esta interacción es estadísticamente significativa ($p < 0.05$) para las tres bases de datos y para cada tipo de característica de textura.

En las Figs. 6.21-6.23 se muestra gráficamente la evaluación de las tasas de clasificación correcta (CCRs (%)) para cada uno de los factores y para las tres bases de datos de texturas. El análisis de la interacción de los tres factores (ancho de banda frecuencial ϕ , número de orientaciones K y medida de distancia d_x) permitió observar que para las tres bases de datos de texturas y para todas las características, excepto para E_5 , la distancia d_{E_N} produce las mejores tasas de clasificación con cualquier combinación de los factores $\phi = 1.5$, $\phi = 2.0$, $K = 8$ y $k = 10$ En la mayoría de los casos la tasa de clasificación correcta entre éstas combinaciones es estadísticamente similar. En el caso de E_5 la distancia d_E es la que mostró un mejor desempeño.

Ya que la distancia d_{E_N} produjo estadísticamente las mejores tasas de clasificación, entonces se restringió el análisis de la interacción a solo 2 factores: ϕ y K . El resultado de esta interacción se muestra en la Tabla 6.4. A partir del análisis de interacción de los factores ϕ y K , podemos concluir que en promedio se obtienen las mejores tasas de clasificación al utilizar $K = 10$ y un ancho de banda frecuencial de $\phi = 1.5$. En muchos casos al utilizar $K = 8$

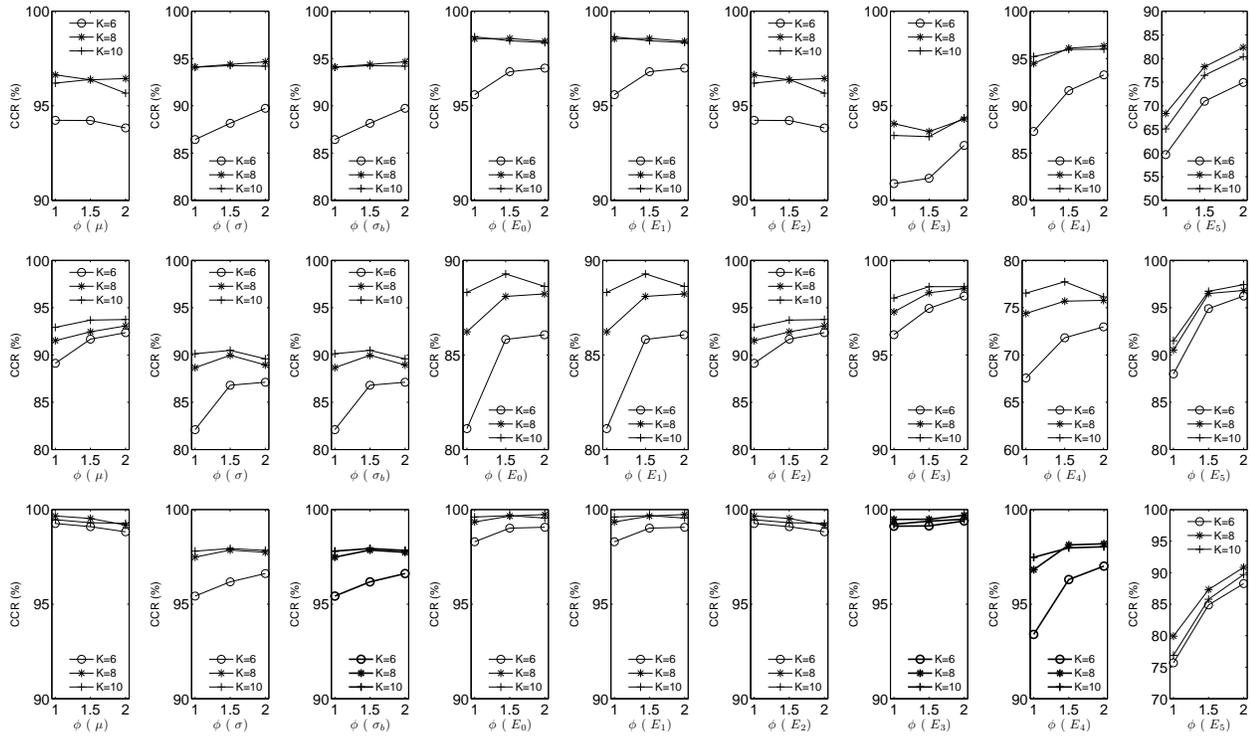


Figura 6.21.: Gráficas de CCRs (%) respecto al ancho de banda frecuencial ϕ usando diversos números de orientaciones K , para cada una de las distintas características presentadas en la Sec. 3.1. De arriba hacia abajo, la primer fila representa las CCRs para la base de datos Brodatz obtenidas con la distancia Canberra, la segunda con la distancia Euclidiana y la fila inferior con la distancia Euclidiana normalizada.

y $\phi = 2.0$ se obtienen tasas de clasificación estadísticamente similares al caso anterior. Esto permite suponer que los errores del análisis cuando se usa un número de orientaciones pequeño se pueden compensar al utilizar un ancho de banda frecuencial ϕ mayor; pues indirectamente se aumenta el ancho de banda angular por la relación intrínseca en la definición de la ecuación de los filtros Gabor (4.20).

invariante a la rotación es obtenido a través de la magnitud de la DFT del vector original,

Tabla 6.4.: Interacción de los factores ϕ y K usando la distancia d_{E_N} para la elección de un banco de filtros Gabor. Para cada tipo de características se reporta la CCR(%) en el primer renglón que fue encontrada estadísticamente similar a la mejor CCR (Tukey=0.05 en la mayoría de los casos) en el segundo renglón.

Característica	Brodatz CCR@(ϕ, K)	VisTex CCR@(ϕ, K)	Outex CCR@(ϕ, K)
μ	96.23@(2.0, 8)	83.46@(1.5, 8)	92.40@(1.5, 10)
	96.46@(1.5, 10)	83.74@(1.5, 10)	92.54@(2.0, 10)
σ	94.06@(1.5, 8)	77.11@(1.5, 8)	80.16@(1.0, 8)
	94.23@(1.5, 10)	77.13@(1.0, 10)	81.65@(1.0, 10)
σ_b	94.06@(1.5, 8)	77.11@(1.5, 8)	80.16@(1.0, 8)
	94.23@(1.5, 10)	77.13@(1.0, 10)	81.65@(1.0, 10)
E_0	95.46@(2.0, 8)	78.64@(1.5, 10)	84.89@(2.0, 10)
	96.52@(1.0, 10)	78.69@(2.0, 10)	84.95@(1.5, 10)
E_1	95.46@(2.0, 8)	78.64@(1.5, 10)	84.89@(2.0, 10)
	96.52@(1.0, 10)	78.69@(2.0, 10)	84.95@(1.5, 10)
E_2	96.23@(2.0, 8)	83.46@(1.5, 8)	92.40@(1.5, 10)
	96.46@(1.5, 10)	83.74@(1.5, 10)	92.54@(2.0, 10)
E_3	97.49@(2.0, 10)	86.51@(1.5, 10)	94.50@(1.5, 10)
	97.50@(2.0, 8)	86.89@(2.0, 8)	95.03@(2.0, 10)
E_4	90.10@(2.0, 8)	71.33@(1.5, 8)	70.01@(1.0, 8)
	90.56@(1.5, 10)	71.53@(2.0, 10)	70.43@(1.5, 10)
E_5	89.18@(2.0, 8)	74.80@(2.0, 8)	68.80@(1.5, 8)
	90.00@(2.0, 10)	74.99@(1.5, 8)	69.12@(2.0, 8)

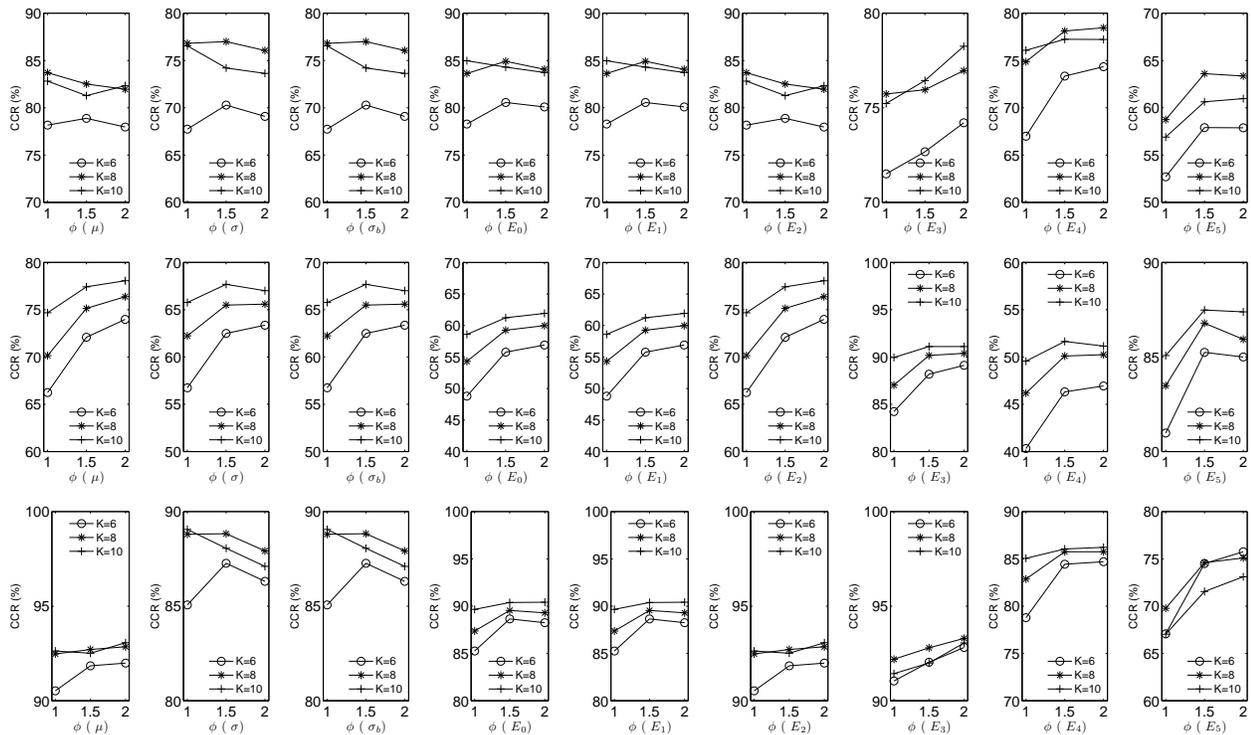


Figura 6.22.: Gráficas de CCRs (%) respecto al ancho de banda frecuencial ϕ usando diversos números de orientaciones K , para cada una de las distintas características presentadas en la Sec. 3.1. De arriba hacia abajo, la primer fila representa las CCRs para la base de datos VisTex obtenidas con la distancia Canberra, la segunda con la distancia Euclidiana y la fila inferior con la distancia Euclidiana normalizada.

6.2.3. Criterio de Fisher usando los filtros Gabor

El criterio de Fisher es utilizado para medir la separación entre las distribuciones intra-clase e inter-clase. Las distribuciones intra-clase se generan al calcular para cada clase las distancias entre los vectores de características pertenecientes a una misma clase. Las distribuciones inter-clase se generan al calcular para cada clase las distancias entre los vectores de características pertenecientes a dicha clase con respecto a los vectores de características pertenecientes al resto de las clases.

El valor del criterio de Fisher que se presenta en los siguientes resultados, obtenidos al seleccionar aleatoriamente conjuntos disjuntos de entrenamiento y prueba, es un indicativo de una separación promedio entre las clases. Los resultados que se presentan a continuación fueron obtenidos usando las distancias Euclidiana, Canberra y Euclidiana normalizada descritas en las Sec. 3.3 para cada una de las distintas características que fueron presentadas en la Sec. 3.1.

Para calcular el criterio de Fisher y posteriormente poder comparar los resultados con los obtenidos en la sección de resultados usando la transformada de Hermite y considerando además los resultados presentados en la Sec. 6.2.2 se utilizó un banco de filtros de análisis

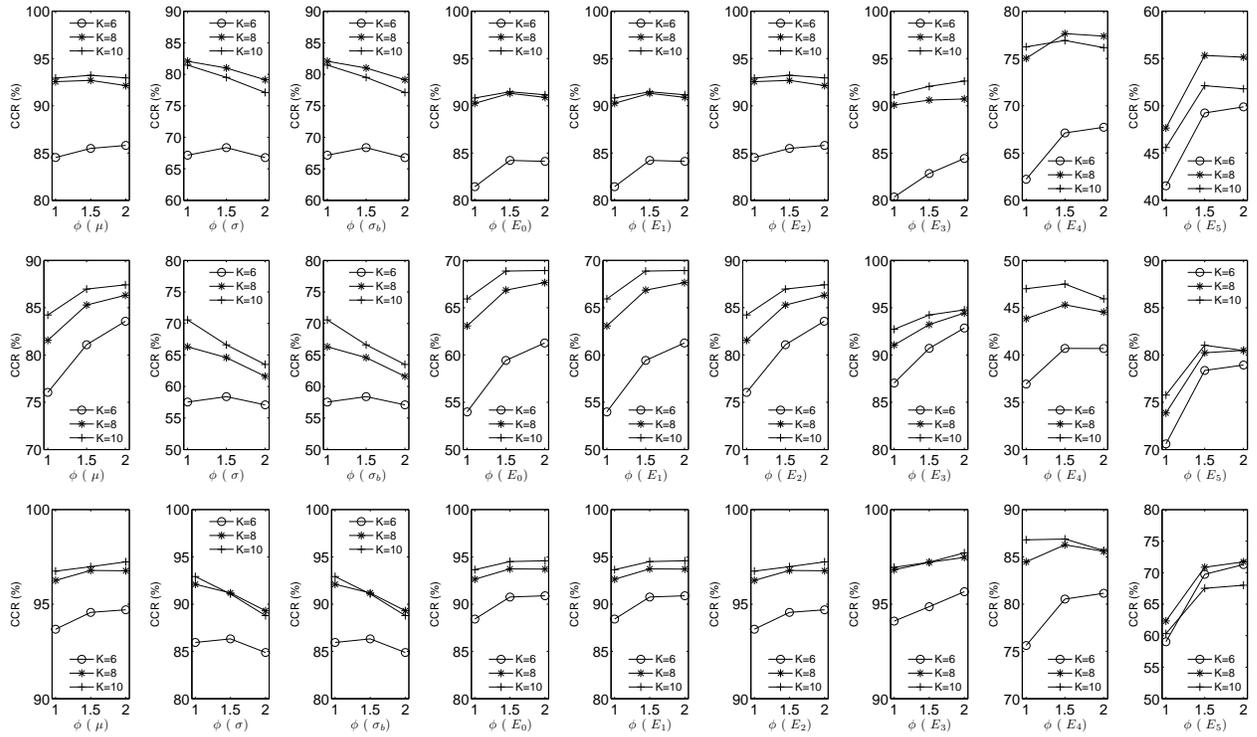


Figura 6.23.: Gráficas de CCRs (%) respecto al ancho de banda frecuencial ϕ usando diversos números de orientaciones K , para cada una de las distintas características presentadas en la Sec. 3.1. De arriba hacia abajo, la primer fila representa las CCRs para la base de datos Outex obtenidas con la distancia Canberra, la segunda con la distancia Euclidiana y la fila inferior con la distancia Euclidiana normalizada.

Gabor con $K = 10$ orientaciones de análisis, $\phi = 1.5$, $S = 4$ escalas de análisis y separación de una octava entre las frecuencias centrales. Las frecuencias centrales más baja y más alta son 0.039 y 0.312 cpp respectivamente.

Las Figs. 6.24-6.26 corresponden a las distribuciones de los cálculos del criterio de Fisher para las 80 repeticiones del experimento usando los conjuntos de texturas Brodatz [8], VisTex [80] y Outex [66, 67] respectivamente. El resultado de cada característica y la distancia empleada para comparar los vectores de características se muestran por separado.

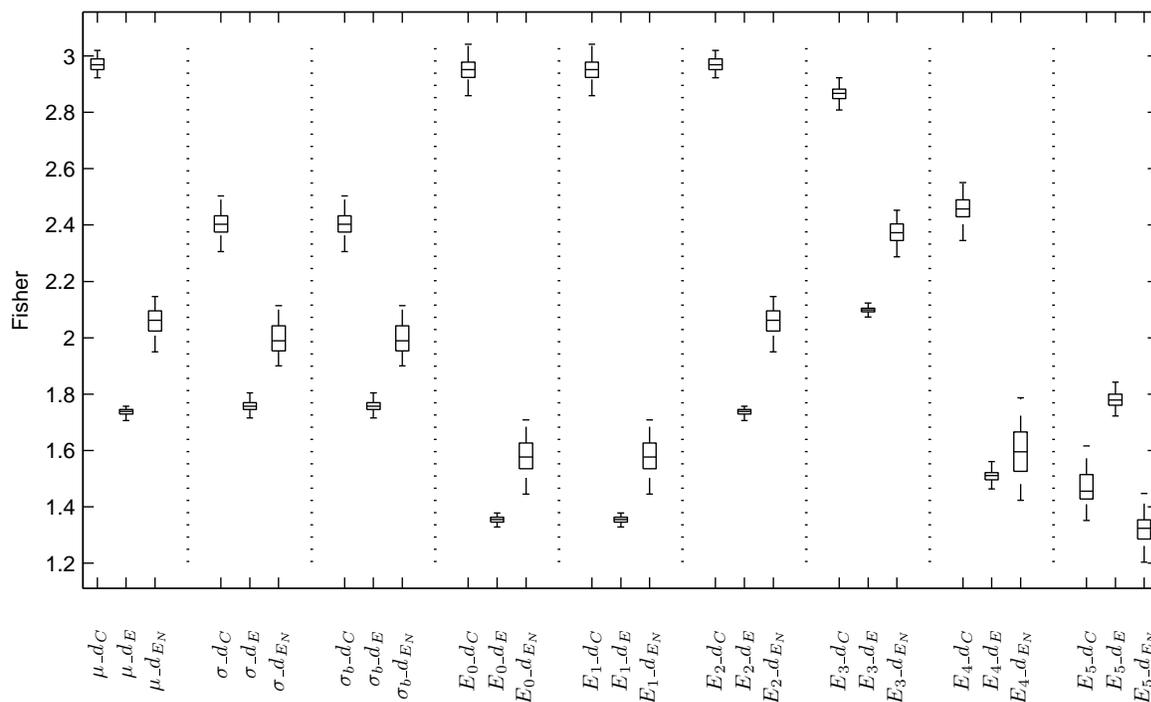


Figura 6.24.: Comparación del criterio de Fisher entre las distintas combinaciones de tipos características y medidas de distancias para el conjunto de texturas Brodatz usando filtros Gabor [8] al obtener las distribuciones intra-clase e inter-clase con 80 repeticiones del experimento de clasificación. Cada diagrama de caja muestra el valor del criterio de Fisher que resulta al combinar un tipo de característica con una medida de distancia.

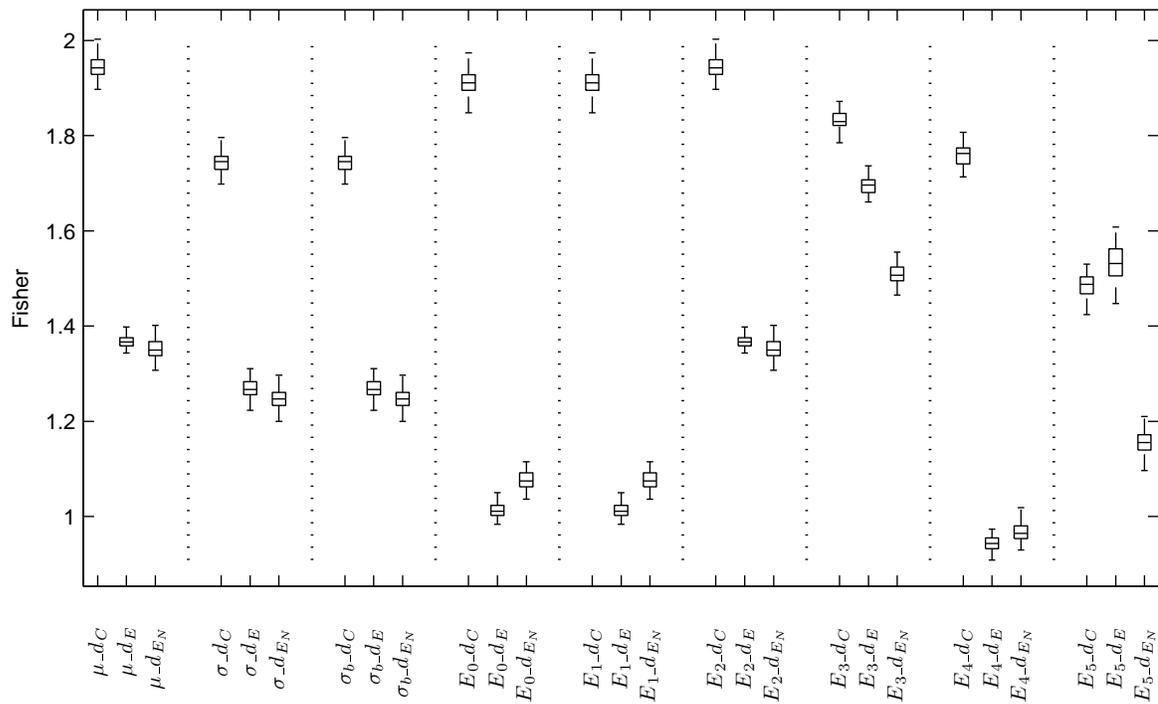


Figura 6.25.: Comparación del criterio de Fisher entre las distintas combinaciones de tipos características y medidas de distancias para el conjunto de texturas VisTex usando filtros Gabor [80] al obtener las distribuciones intra-clase e inter-clase con 80 repeticiones del experimento de clasificación. Cada diagrama de caja muestra el valor del criterio de Fisher que resulta al combinar un tipo de característica con una medida de distancia.

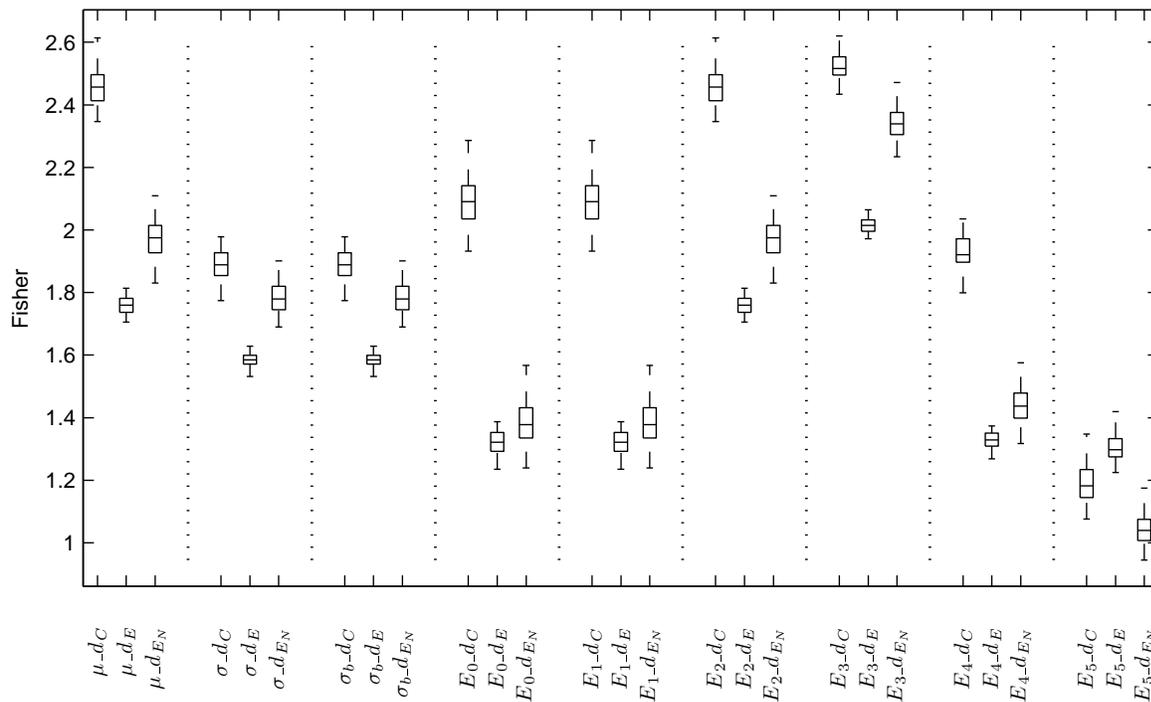


Figura 6.26.: Comparación del criterio de Fisher entre las distintas combinaciones de tipos características y medidas de distancias para el conjunto de texturas Outex usando filtros Gabor [66, 67] al obtener las distribuciones intra-clase e inter-clase con 80 repeticiones del experimento de clasificación. Cada diagrama de caja muestra el valor del criterio de Fisher que resulta al combinar un tipo de característica con una medida de distancia.

6.2.4. Sensibilidad del sistema usando los filtros Gabor

A partir de las distribuciones de distancia (d_x) intra-clase e inter-clase (Ecuaciones (5.17), (5.19)) es posible obtener el desempeño de la clasificación para las combinaciones de los tipos de características y medidas de distancias usando un banco de filtros Gabor diseñado con un ancho de banda $\phi = 1.5$ y $K = 10$.

En la Fig. 6.27 se muestran las curvas ROC de la sensibilidad S como función del error FAR para la base datos Brodatz para todos los tipos de características y medidas de distancias. De manera similar las Figs. 6.28 y 6.29 muestran las curvas ROC para la base de datos VisTex y Outex respectivamente. En todos los casos se muestra además el valor de la sensibilidad S para una tasa de error igual $EER = FAR = FRR$, como se explicó en la sección 5.5.1, donde $FAR = 1 - E$ y E es la especificidad del modelo.

Para todas las bases de datos de texturas se pudo observar que en la mayoría de los casos se obtiene la sensibilidad S más alta (o bien la tasa de error igual (EER) más baja) al utilizar la distancia Canberra (d_C) seguida de la distancia d_{E_N} y por último de la distancia d_E , con cualquier tipo de característica, excepto cuando se utilizan las características E_3 y E_5 . Los tipos de características que permiten obtener una S alta son μ , E_3 , σ_b y E_2 .

Para todas las bases de datos se observó que la sensibilidad S es la misma para el par de características σ y σ_b , así como para E_0 y E_1 . Esto se debe a la transformación que sufre el vector de características original a través de la DFT para obtener el vector de características invariante a la rotación (véase ecuación (3.24)).

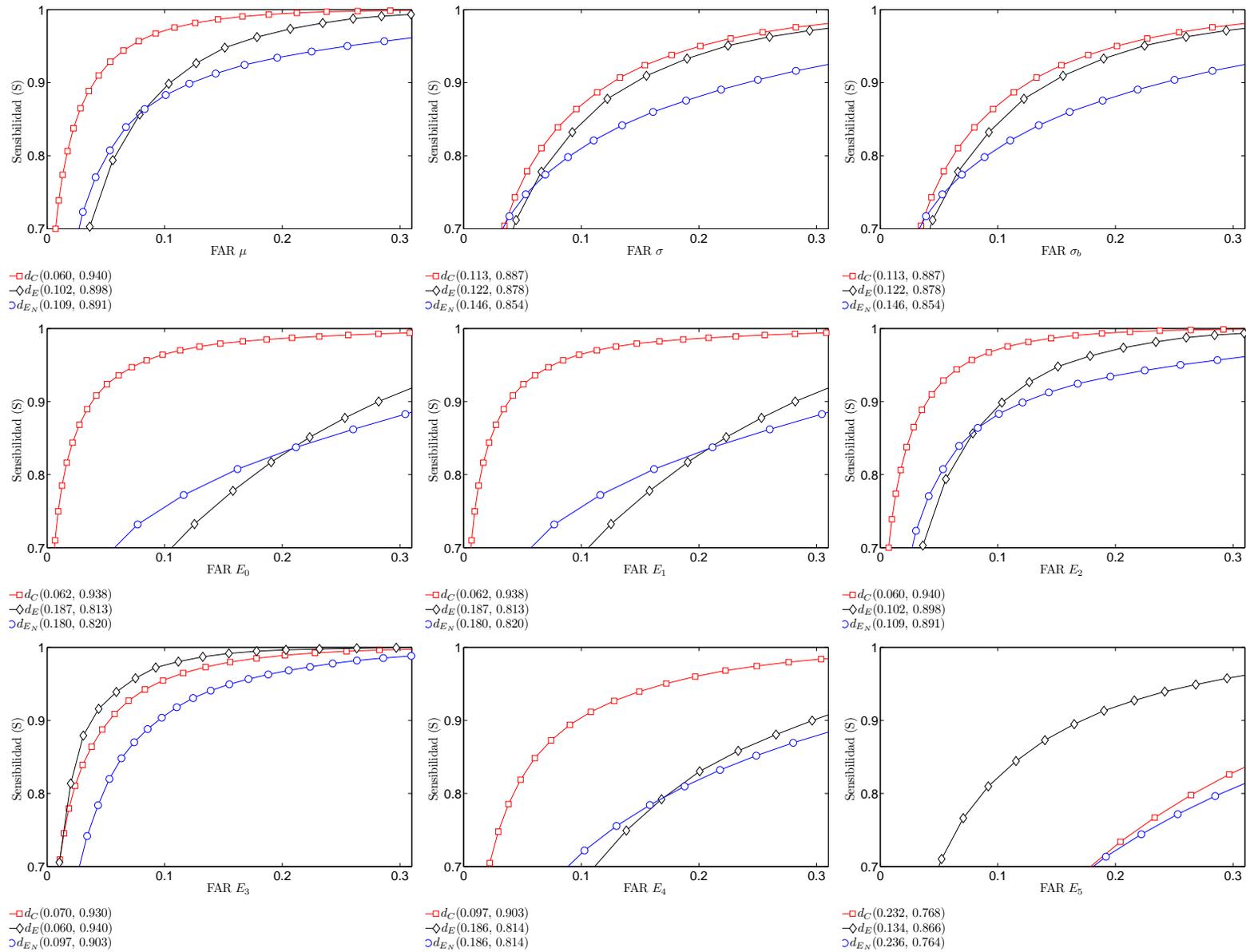


Figura 6.27.: Comparación de curvas ROC usando un banco de filtros Gabor para la base de datos Brodatz [8]. Cada figura representa la gráfica de la sensibilidad (S) con respecto al error FAR. Para cada método de reducción de dimensión se reporta S y FAR obtenido con las medidas de distancia.

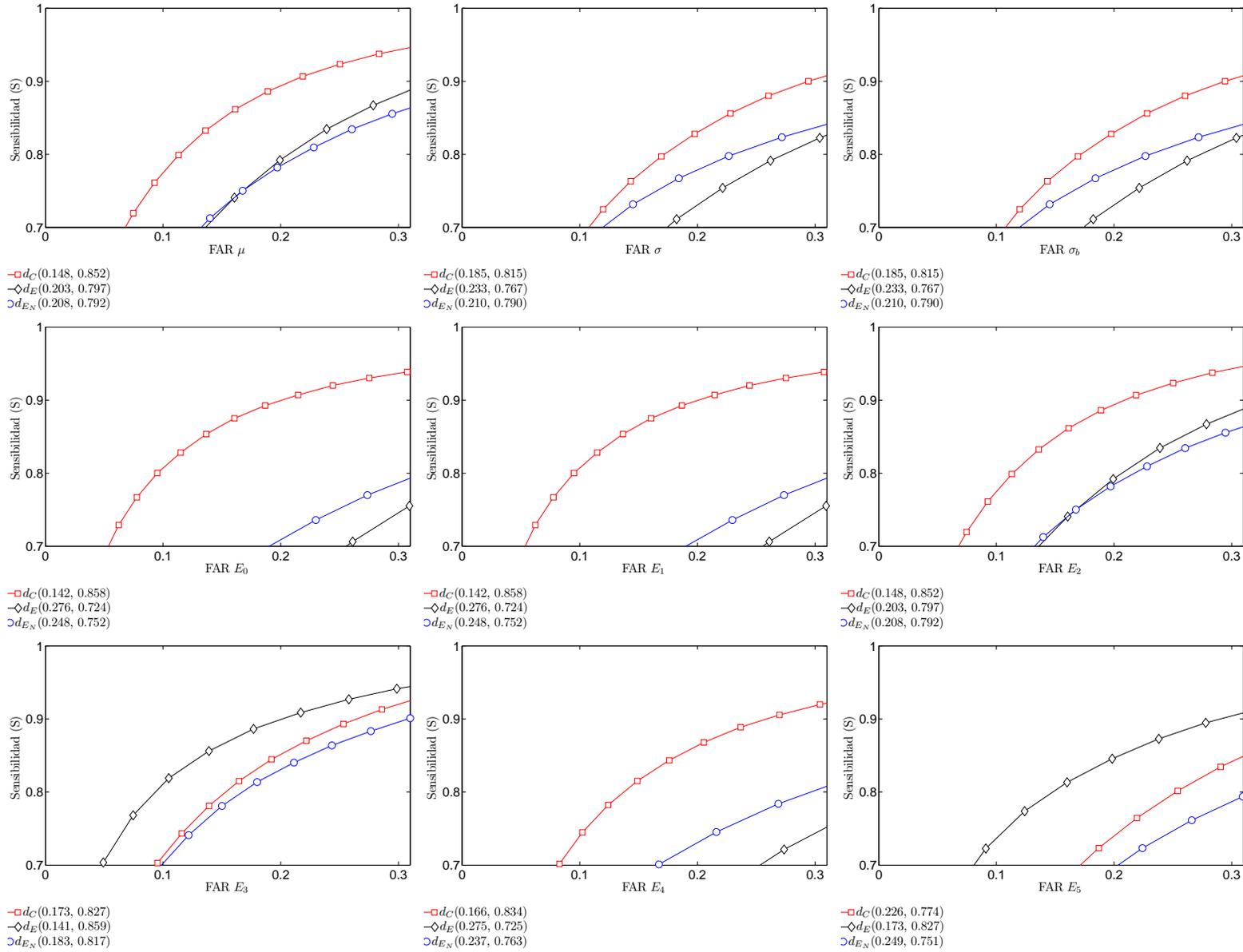


Figura 6.28.: Comparación de curvas ROC usando un banco de filtros Gabor para la base de datos VisTex [80]. Cada figura representa la gráfica de la sensibilidad (S) con respecto al error FAR. Para cada método de reducción de dimensión se reporta S y FAR obtenido con las medidas de distancia.

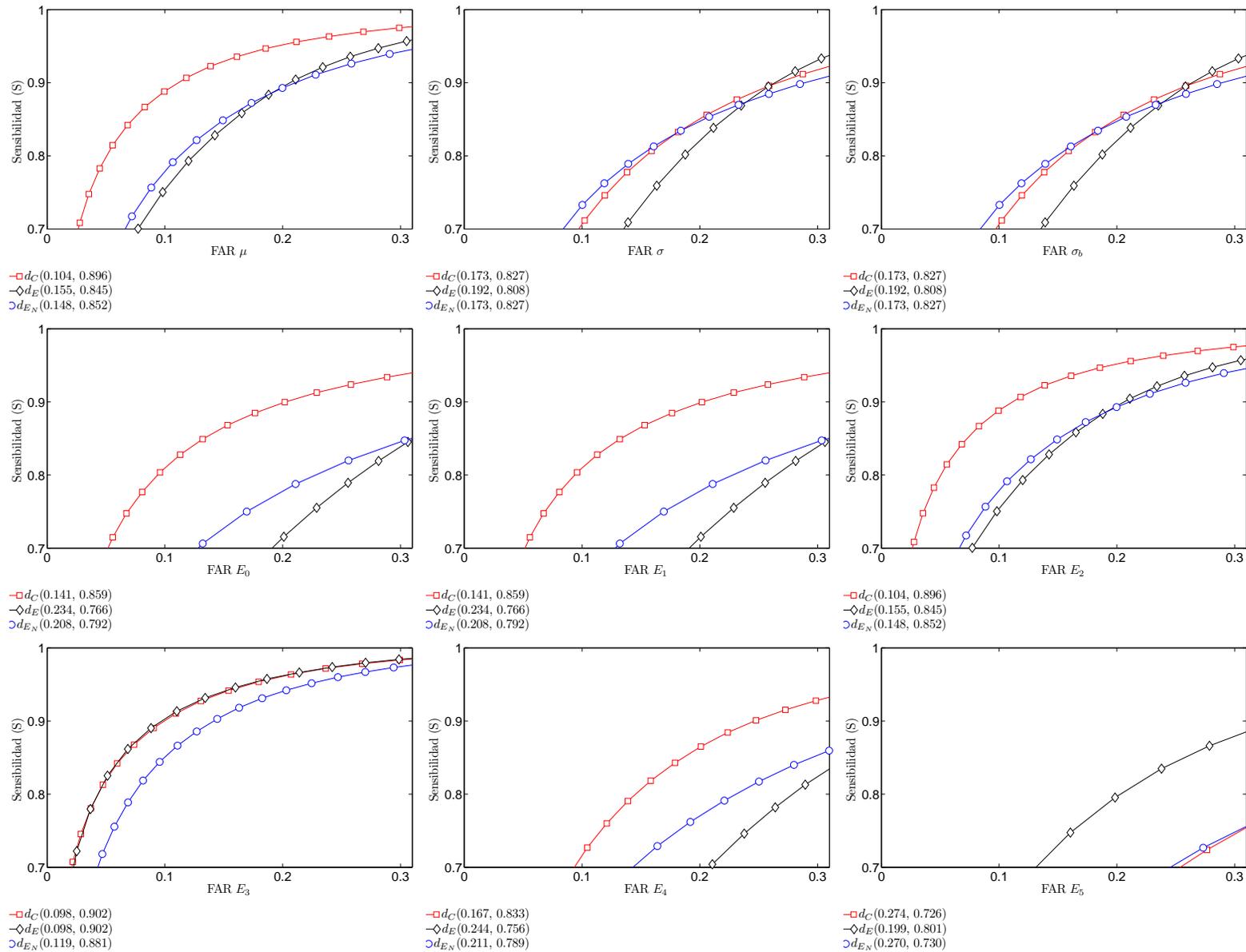


Figura 6.29.: Comparación de curvas ROC usando un banco de filtros Gabor para la base de datos Outex [66, 67]. Cada figura representa la gráfica de la sensibilidad (S) con respecto al error FAR. Para cada método de reducción de dimensión se reporta S y FAR obtenido con las medidas de distancia.

7. Conclusiones y perspectivas

En esta tesis se presentaron metodologías para aumentar las tasas de discriminación entre texturas que pertenecen a diferentes clases y disminuir la variabilidad entre texturas pertenecientes a una misma clase. Existen distintos factores que pueden alterar las tasas de clasificación los cuales se engloban principalmente en el modelo utilizado para el procesamiento de las imágenes de texturas y sus parámetros, los tipos de características y las medidas de distancia utilizadas en el espacio de las características.

Aunque existen muchos métodos para el análisis de las imágenes de texturas se optó y se limitó el trabajo al utilizar los modelos computacionales inspirados en los sistemas de visión biológicos por sus propiedades matemáticas y biológicas que estos representan. Existen muchos resultados positivos reportados en la literatura del análisis de texturas que refieren principalmente al modelo utilizando filtros Gabor y un uso a discreción de las medidas de distancia y de los tipos de características extraídos a partir de las imágenes filtradas. En esta tesis se utilizó además la transformada de Hermite como un modelo de visión alternativo, con menor difusión en la literatura pero que posee sus propias ventajas sobre otros.

La metodología principal presentada en esta tesis partió de que no existe un conocimiento previo sobre el tipo de texturas que serán analizadas y tampoco de las características y los tipos de características que son más representativas de la clase de textura. Por esta razón se planteó realizar un análisis exhaustivo para determinar desde un inicio, a partir de las imágenes de entrenamiento, cuales son los filtros idóneos que con mayor certeza propician una tasa de clasificación alta y seleccionarlos para configurar un banco de filtros durante la etapa de pruebas o clasificación, en la cual el conjunto de imágenes es mayor, y reducir los costos computacionales.

Se analizaron de forma individual los tipos de características y de distancias más comunes en la literatura. A partir de los resultados de la clasificación se determinó experimentalmente cuales son los que pueden generar las mejores tasas. Para el análisis de resultados se utilizaron el criterio de Fisher, el análisis de varianzas sobre las tasas de clasificación (ANOVA) y curvas ROC.

Para el caso de la transformada de Hermite la idea fue utilizar un conjunto pequeño de imágenes de entrenamiento las cuales fueron filtradas exhaustivamente usando las funciones de análisis de la transformada de Hermite. Las imágenes filtradas fueron procesadas para obtener versiones invariantes a la rotación y a partir de estas versiones invariantes a la rotación se extrajeron distintos tipos de características, las más reportadas en la literatura. Posteriormente un método de reducción de dimensión (AVR o ICV) y selección de los índices de los filtros más representativos fue utilizado para obtener la información necesaria para configurar el conjunto de los filtros de análisis durante la etapa de pruebas o clasificación. Para comprobar si los métodos de reducción de dimensión y selección de los índices de los

filtros más representativos propiciaban que se obtuviera un desempeño alto en las tasas de clasificación, los resultados de la separación de clases y de la clasificación durante la etapa de pruebas fueron comparados con los obtenidos utilizando el método de gran difusión en la literatura PCA.

Una contribución fue encontrar experimentalmente que PCA no es una herramienta que ayude a mejorar las tasas de clasificación. Además, los métodos AVR e ICV si permitieron conocer desde el espacio de características que índices de los filtros son los más representativos e idóneos para generar tasas de clasificación altas a partir del conjunto de imágenes de entrenamiento y reducir operaciones computacionales durante la etapa de pruebas. Un análisis de la complejidad usando esta estrategia se presenta en el Apéndice E.

El método de selección de características y de reducción del tamaño del vector en el que se agrupan éstas llamado AVR fue el más robusto. El método propuesto ICV, demostró ser competitivo con respecto a PCA pero inferior a AVR. Las ventajas de estos métodos poco convencionales se resaltan en las tasas de clasificación altas que se pudieron obtener con una reducción del tamaño del vector de características y además el conocimiento de los elementos que conforman al vector de características y como consecuencia de los índices de los filtros.

Al determinar las características más importantes a través de AVR e ICV como consecuencia se logro conocer que índices de los filtros de análisis de Hermite son los que ayudan a obtener una mejor tasa de clasificación. Es frecuente que para cada escala de análisis (de las 4 en su totalidad en estos experimentos) los filtros más importantes sean los de primer y segundo orden. Sin embargo no es una regla ya que depende en gran medida del tipo de característica utilizado pero sobre todo de las texturas bajo análisis. Mientras que para alguna base de datos de texturas la primer escala de análisis es fundamental, para otra base de datos puede ser la última escala de análisis quien posea este valor y de la misma manera sucede con los órdenes de los filtros de análisis. Con esta técnica no es posible generalizar a un solo criterio la cantidad de traslape entre los filtros de análisis para lograr obtener tasas de clasificación altas, por lo que se puede considerar como una técnica adaptiva a las propiedades de las texturas bajo análisis.

Respecto a los tipos de características usando la transformada de Hermite podemos concluir que de manera general usando las características E_3 , σ_b , y μ es posible obtener con mayor certeza una tasa de clasificación correcta alta. Por otro lado, la tasa de clasificación disminuye al usar las características E_4 , E_0 y E_5 . En términos de obtener una tasa de clasificación correcta “moderada” las características E_2 , σ y E_1 son suficientes.

La distancia de Mahalanobis d_M mostró ser superior a las otras distancias en varios casos pero solo cuando el vector se formaba con un número muy pequeño de elementos de características. Las distancias más competitivas son d_C y d_{E_N} . No fue posible determinar que distancia es mejor debido a que se observó que hay cierta dependencia del tipo de característica. La distancia d_E en la mayoría de los casos presentó un desempeño bajo excepto para algunas características como σ , E_2 y E_5 , aunque de manera global estas características no representen las mejores CCRs.

Respecto al análisis usando los filtros Gabor la idea fue diseñar distintos bancos de filtros Gabor para encontrar experimentalmente que combinación de parámetros relacionados con la cantidad de traslape en el dominio frecuencial de las funciones de análisis permitía obtener

las mejores tasas de clasificación. Una contribución fue encontrar que, a diferencia de lo reportado en muchos trabajos en la literatura, la redundancia que genera el traslape es necesaria para obtener tasas de clasificación altas. Sin embargo, una cantidad en exceso de traslape no mejora la tasa de clasificación con respecto a una cantidad de traslape moderada.

Experimentalmente se encontró que la combinación $\phi = 1.5$ y $K = 10$ permite obtener las mejores tasas de clasificación correctas (CCRs, %). Este resultado coincide con los resultados obtenidos previamente con otros autores en el sentido de que para poder describir una textura las funciones o filtros de análisis utilizadas no son necesariamente ortogonales, por el contrario son funciones que en el dominio de la frecuencia presentan cierto grado de traslape entre éstas. Por otro lado, los bancos de filtros con grandes cantidades de traslape generalmente no mejorarán las tasas de clasificación en términos de diferencias estadísticamente significativas, por ejemplo usando $\phi = 2.0$ y $K = 10$.

El análisis a través del criterio de Fisher de la medida de distancia al utilizar los filtros Gabor permitió obtener un resultado interesante: en promedio para todos los tipos de características, excepto para E_5 , se obtuvo una mejor separación entre las distribuciones utilizando la distancia Canberra d_C e incluso esta separación es mejor comparada con la obtenida con los filtros de Hermite utilizando la misma medida de distancia. Sin embargo, esta separación no necesariamente significa que se obtuvieran las mejores tasas de clasificación; el análisis estadístico indica que para el caso de las texturas procesadas con los filtros Gabor la distancia d_{E_N} es apta para obtener las mejores CCRs.

Las características invariantes a la rotación, calculadas con la magnitud de la DFT sobre un vector de características obtenido previamente a partir de las texturas filtradas con los filtros Gabor, simplificaron el análisis puesto que al usar cualquier característica del par (σ, σ_b) se obtuvieron las mismas CCRs. De la misma manera sucedió con las parejas (E_0, E_1) y (μ, E_2) . De esta manera, se encontró que en general usando las características $E_3, (\mu, E_2)$ es posible obtener con mayor certeza una tasa de clasificación correcta alta. Las características que por el contrario propician a que se obtengan con mayor certeza una tasa de clasificación correcta baja son E_5 y E_4 . En términos de obtener una tasa de clasificación correcta “moderada” las características (E_0, E_1) generan mejores CCRs que las obtenidas con (σ, σ_b) .

Uno de los trabajos que se sugieren para mejorar la robustez de la clasificación es proponer una metodología para la fusión de los vectores de característica que fueron obtenidos de forma independiente para cada tipo de característica. Las evaluaciones presentadas en esta tesis son un indicio para la selección de los tipos de características con las que se puede llegar a obtener tasas de clasificación competitivas. La combinación de clasificadores o bien de estos vectores sugiere que en teoría las tasas de clasificación pueden mejorar aún más.

Es necesario comparar los resultados de la clasificación a través de otras técnicas de obtención de características invariantes a la rotación basadas en los filtros Gabor utilizando los parámetros de los filtros relacionados al traslape que mejoraron las tasas de clasificación según los experimentos presentados.

A. Tasas de clasificación correcta obtenidas con los métodos AVR, ICV y PCA a partir de características extraídas con la transformada de Hermite orientada

Usando los resultados presentados en la Sec. 6.1.2 se restringió el máximo número de elementos del vector de características a 20, seleccionados a partir de los métodos de reducción de dimensión AVR, ICV y PCA para evaluar la CCRs. En las Figs. A.1-A.3 se muestran los resultados para cada base de datos Brodatz, VisTex y Outex respectivamente.

Para las tres bases de datos se observa que los métodos AVR e ICV proporcionaron las mejores tasas de clasificación con respecto a las obtenidas con PCA, excepto con algunos tipos de características. En algunos casos se obtienen mejores resultados con ICV que con AVR. Sin embargo, el comportamiento del crecimiento en la CCR(%) conforme se aumenta el número de elementos del vector de características es más suave usando AVR a diferencia de ICV que presenta cambios abruptos en su crecimiento.

Con respecto a las distancias se observó de manera general que las distancias más competitivas son d_C y d_{E_N} . No es posible determinar que distancia es mejor debido a que se observa que hay cierta dependencia del tipo de característica. Sin embargo, la distancia d_E presentó en la mayoría de los casos un desempeño bajo excepto para algunas características como σ , E_2 y E_5 , aunque en comparación con otras características σ y E_5 no representen las mejores CCRs. Las condiciones que se observaron para que al utilizar la distancia de Mahalanobis d_M se obtuviera un mejor desempeño fue que el vector de características debía conformarse con un número pequeño de elementos. Para el caso de la CCR(%) obtenida con la distancia d_M se determinó que en promedio se requieren 10 elementos del vector de características para obtener una CCR(%) estadísticamente similar a la mejores.

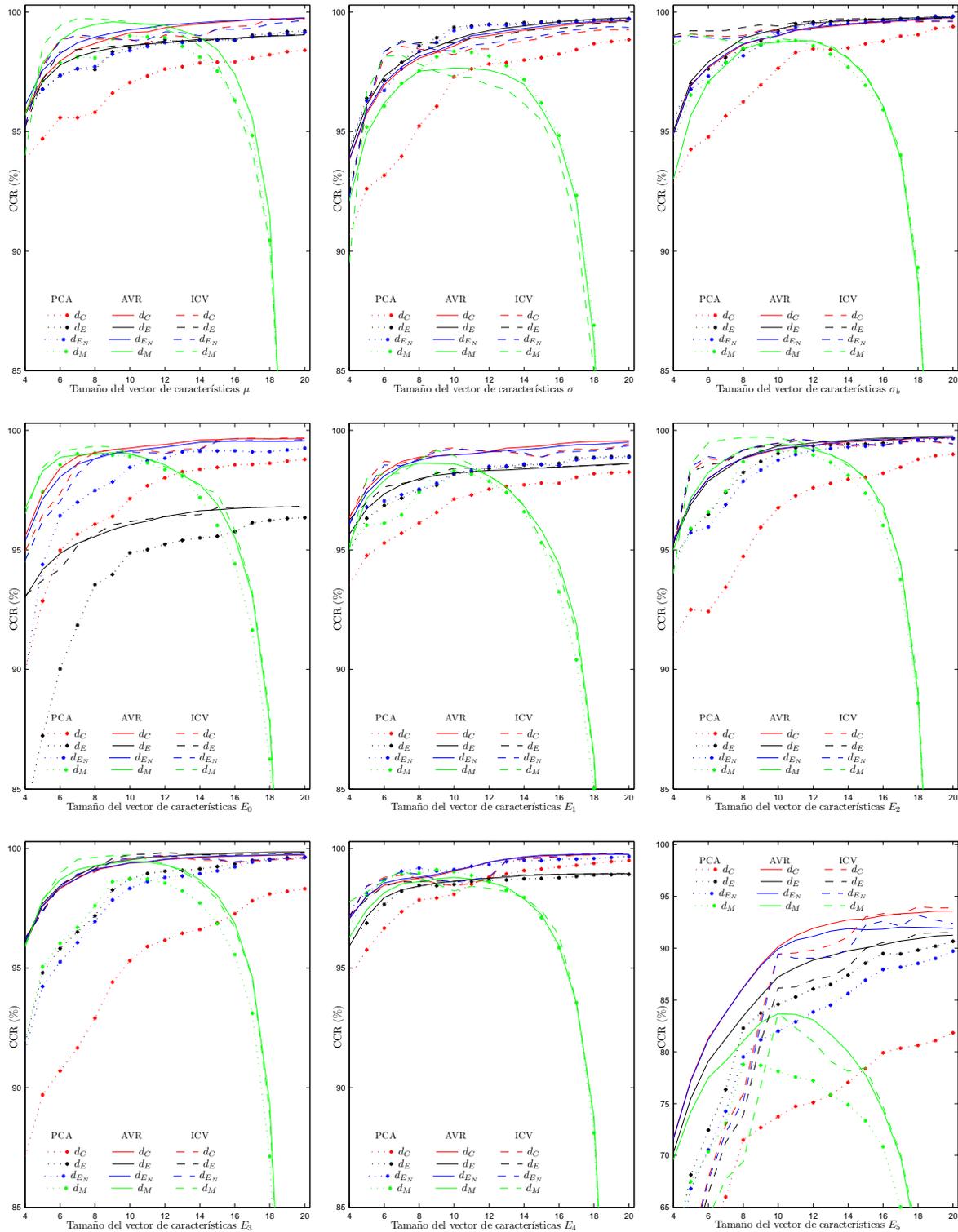


Figura A.1.: Comparación de CCRs usando la transformada de Hermite para la base de datos Brodatz [8].

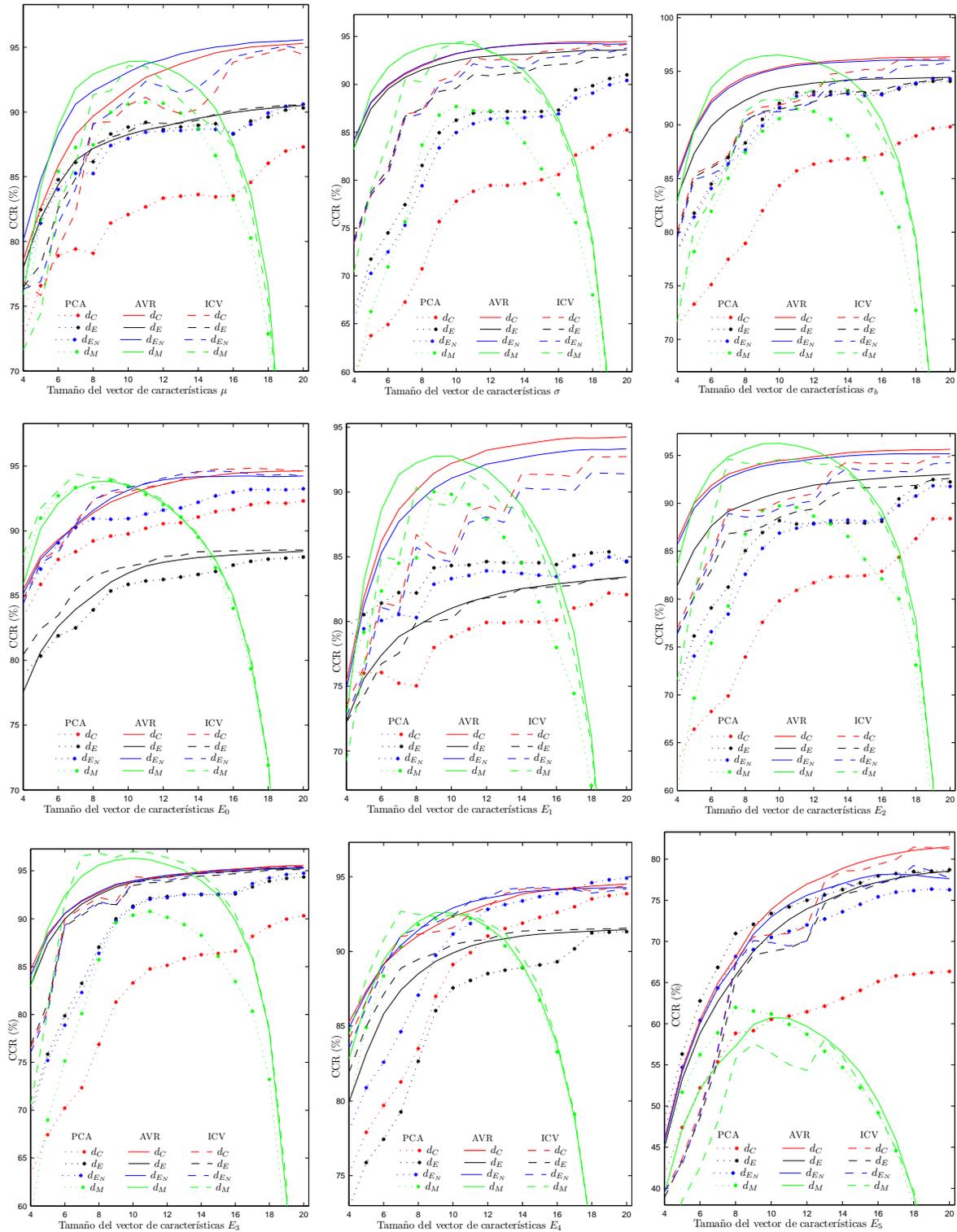


Figura A.2.: Comparación de CCRs la transformada de Hermite para la base de datos VisTex [80].

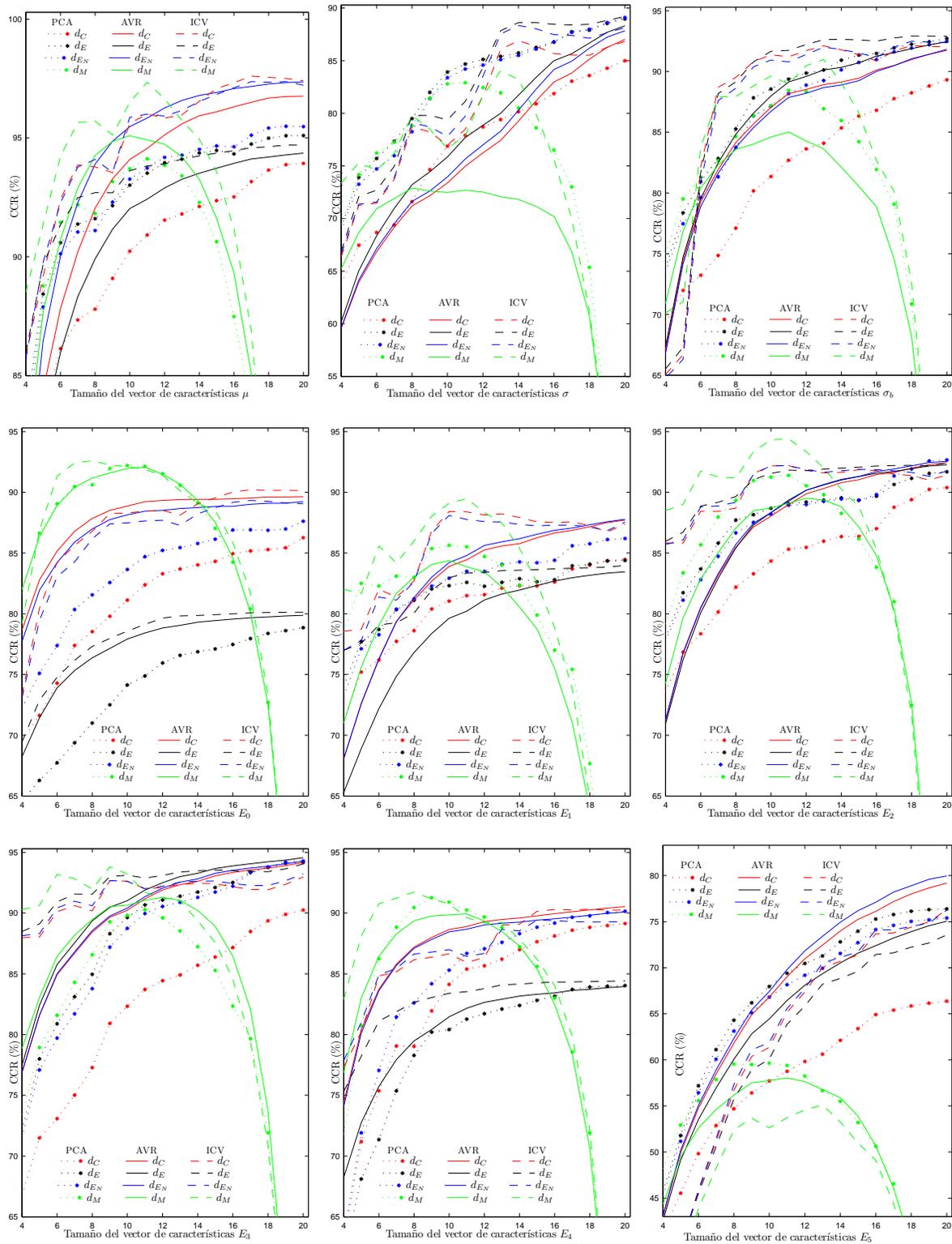


Figura A.3.: Comparación de CCRs usando la transformada de Hermite para la base de datos Outex [66, 67].

B. Polinomios de Hermite

La ecuación de Hermite tiene la forma siguiente [1, 74]:

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

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 (B.1) está dada por:

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

La demostración de que (B.2) es solución a (B.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:

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

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 (B.3) y dividiendo por e^{-x^2} nos queda:

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

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

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 órdenes de los polinomios de Hermite es:

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

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:

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

donde es posible observar que

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

Ortogonalidad

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

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

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:

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

multiplicamos esta ecuación por H_m :

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

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

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

restando estas últimas dos ecuaciones, (B.7) y (B.8) nos queda:

$$\begin{aligned} H_m(x) \frac{d}{dx} \left(e^{-x^2} H_n'(x) \right) - H_n(x) \frac{d}{dx} \left(e^{-x^2} H_m'(x) \right) = \\ -H_m(x) 2ne^{-x^2} H_n(x) + H_n(x) 2me^{-x^2} H_m(x) , \end{aligned} \quad (\text{B.9})$$

integrando sobre x de $-\infty$ a ∞ ; los términos de la izquierda se integran por partes

($\int u dv = uv - \int v du$). Para el primer término de la izquierda tenemos:

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

$$\begin{aligned} & \int_{-\infty}^{\infty} H_m(x) \frac{d}{dx} \left(e^{-x^2} H_n'(x) \right) dx = \\ & H_m(x) e^{-x^2} \frac{d}{dx} H_n(x) \Big|_{-\infty}^{\infty} - \int_{-\infty}^{\infty} \frac{d}{dx} H_m(x) e^{-x^2} \frac{d}{dx} H_n(x) dx \end{aligned} \quad (\text{B.10})$$

y para el segundo término tenemos:

$$\begin{aligned} & - \int_{-\infty}^{\infty} H_n(x) \frac{d}{dx} \left(e^{-x^2} H_m'(x) \right) dx = \\ & - H_n(x) e^{-x^2} \frac{d}{dx} H_m(x) \Big|_{-\infty}^{\infty} + \int_{-\infty}^{\infty} \frac{d}{dx} H_n(x) e^{-x^2} \frac{d}{dx} H_m(x) dx \end{aligned} \quad (\text{B.11})$$

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 (B.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 (B.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 (B.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 δ_{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 $dx = \frac{1}{\sigma} du$.

C. Normalización L^1 y L^2 de la función Gaussiana

Una Gaussiana tiene la siguiente forma general [62]:

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

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 σ 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 [62] como la definida en la ecuación (C.1) es:

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

por lo que podemos calcular L^2 :

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

y de la misma manera L^1 :

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

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 (C.3) $a = (2\sigma^2)^{-1}$:

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

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

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

D. Obtención de las derivadas de Gaussianas como filtros de análisis de la transformada Hermite

La función Gaussiana G_0 definida en la ecuación (C.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 σ 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:

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

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

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

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 :

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

las funciones filtro en la transformada Hermite están definidas como:

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

sustituyendo la ecuación (D.3) en (D.4):

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

pero

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

sustituyendo esta última ecuación en (D.5) obtenemos:

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

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 (D.2). Entonces podemos escribir a $D_n(x)$ como [62]:

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

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 (D.1) evaluadas en $x = \frac{x}{\sigma}$, pero multiplicadas por una constante de normalización L^2 [57]:

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

E. Análisis de la complejidad computacional

En lo referente a la complejidad computacional de los métodos propuestos la mayor carga computacional reside en las operaciones de procesamiento de las imágenes. En este trabajo se han filtrado las imágenes en el dominio frecuencial. Este proceso involucra el cálculo de la transformada discreta de Fourier (DFT) de la imagen de entrada y del filtro. Para obtener la DFT se utilizó una biblioteca optimizada para su cálculo a través del algoritmo de la transformada rápida de Fourier (FFT). Esta biblioteca escrita originalmente en lenguaje C es ampliamente utilizada y se llama FFTW [34]. Una DFT a través de este algoritmo requiere $O(t \log t)$ para su cálculo, donde t es el tamaño de la muestra, para cualquier tamaño t y no impone restricciones a algún subconjunto respecto al tamaño de la muestra [34].

Suponiendo una imagen de entrada $I(x, y)$ de tamaño $t = H \times W$ y un filtro de análisis de tamaño a lo más $t = H \times W$, entonces podemos escribir que filtrar una imagen requiere:

$$O(2 \times t \log t) = O(t \log t) . \quad (\text{E.1})$$

E.1. Análisis de la complejidad sin métodos de selección de filtros

E.1.1. Filtrado con bancos de filtros Gabor

A partir de (E.1) es posible determinar que filtrar una imagen con un banco de filtros Gabor calculado con K orientaciones y S escalas de análisis requiere:

$$O(S \times K \times t \log t) = \Omega(t \log t) . \quad (\text{E.2})$$

E.1.2. Filtrado con filtros de análisis de Hermite

El cálculo de toda la descomposición con la transformada de Hermite requiere realizar cada vez más filtrados conforme el grado de la descomposición aumenta. Suponiendo que el grado de la descomposición de una transformada de Hermite es n , entonces el número de operaciones de filtrado requeridas para obtener toda la descomposición es

$$N = (n + 1) + (n - 1) + (n - 2) + \dots + (1) , \quad (\text{E.3})$$

donde N es la suma aritmética de la serie, es decir:

$$N = \frac{n^2 + 3n + 2}{2} . \quad (\text{E.4})$$

Entonces obtener toda la descomposición de una imagen con la transformada de Hermite a una sola escala de análisis requiere:

$$O(N \times t \log t) . \quad (\text{E.5})$$

mientras que para S escalas de análisis se requiere:

$$O(S \times N \times t \log t) . \quad (\text{E.6})$$

E.2. Análisis de la complejidad con métodos de selección de filtros

En el caso de que exista algún método que seleccione filtros específicos previamente a las etapas de prueba o de clasificación, entonces es de esperarse que la complejidad disminuya o al menos esa es la hipótesis. En este trabajo se limitó la metodología de selección de filtros únicamente para el caso de la descomposición con la transformada de Hermite. Como en cada prueba realizada los órdenes de los filtros de análisis de Hermite que fueron seleccionados variaban, es necesario realizar las evaluaciones experimentales para encontrar alguna cota de referencia. Sin embargo, se debe señalar que el cálculo de una imagen filtrada, con algún filtro de orden n en particular, no requiere que se filtre la imagen con los filtros antecesores a n de orden $n - 1, n - 2, \dots, 0$. Es decir, dependiendo del orden del filtro seleccionado, las operaciones requeridas están acotadas por:

$$O((n + 1) \times t \log t) \quad (\text{E.7})$$

Las operaciones más costosas se presentaran cuando el grado del filtro seleccionado es grande. Sea

$$\{d_{n_1}, d_{n_2}, \dots, d_{n_M}\} \quad (\text{E.8})$$

el conjunto de todos los filtros seleccionados y el máximo grado de todos los filtros es n_{\max} :

$$n_{\max} = \max\{n_1, n_2, \dots, n_M\} \quad (\text{E.9})$$

entonces es posible establecer la cota del tiempo de procesamiento de una imagen con el método de selección de filtros como:

$$\Theta((n_{\max} + 1) \times t \log t) . \quad (\text{E.10})$$

Bibliografía

- [1] ABRAMOWITZ, M., AND STEGUN, I. *Handbook of mathematical functions*. Dover, 1965.
- [2] ARIVAZHAGAN, S., AND GANESAN, L. Texture segmentation using wavelet transform. *Pattern Recognition Letters* 24, 16 (2003), 3197–3203.
- [3] ARIVAZHAGAN, S., GANESAN, L., AND BAMA, S. Fault segmentation in fabric images using Gabor wavelet transform. *Machine Vision and Applications* 16, 6 (2006), 356–363.
- [4] BHARATI, M. H., LIU, J., AND MACGREGOR, J. F. Image texture analysis: methods and comparisons. *Chemometrics and Intelligent Laboratory Systems* 72, 1 (2004), 57–71.
- [5] BIANCONI, F., AND FERNANDEZ, A. Evaluation of the effects of Gabor filter parameters on texture classification. *Pattern Recognition* 40, 12 (2007), 3325–3335.
- [6] 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.
- [7] BRES, S., EGLIN, V., AND RIVERO, C. Handwriting documents denoising and indexing using Hermite transform. En *Pattern Recognition and Data Mining* (2005), vol. 3686 de *LNCS*, pp. 664–673.
- [8] BRODATZ, P. Texture: a photographic album for artists and designers. <http://www.ux.uis.no/~tranden/brodatz.html>, 1966. último acceso 09-09-2013.
- [9] CAMPBELL, F. W., AND ROBSON, J. G. Application of fourier analysis to the visibility of gratings. *The Journal of Physiology* 197, 3 (1968), 551–566.
- [10] CELIK, T., AND TJAHJADI, T. Bayesian texture classification and retrieval based on multiscale feature vector. *Pattern Recognition Letters* 32, 2 (2011), 159–167.
- [11] CHEN, G., BUI, T., AND KRZYŻAK, A. Rotation invariant pattern recognition using ridgelets, wavelet cycle-spinning and Fourier features. *Pattern Recognition* 38, 12 (2005), 2314–2322.
- [12] CHEN, J. L., AND KUNDU, A. Rotation and gray scale transform invariant texture identification using wavelet decomposition and hidden Markov model. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 16, 2 (1994), 208–214.
- [13] CHEN, S., SHANG, Y., MAO, B., AND LIAN, Q. Rotation invariant texture classification algorithm based on DT-CWT and SVM. En *Advances in Neural Networks – ISNN* (2007), vol. 4493 de *LNCS*, pp. 454–460.

-
- [14] CHU, X., AND CHAN, K. Rotation and scale invariant texture analysis with tunable Gabor filter banks. En *Advances in Image and Video Technology* (2009), vol. 5414 de *LNCS*, pp. 83–93.
- [15] CLAUSI, D. A., AND ED JERNIGAN, M. Designing Gabor filters for optimal texture separability. *Pattern Recognition* 33, 11 (2000), 1835–1849.
- [16] COHEN, F., FAN, Z., AND PATEL, M. Classification of rotated and scaled textured images using Gaussian Markov random field models. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 13, 2 (1991), 192–202.
- [17] CROSIER, M., AND GRIFFIN, L. Using basic image features for texture classification. *International Journal of Computer Vision* 88, 3 (2010), 447–460.
- [18] DAUGMAN, J. G. Two-dimensional spectral analysis of cortical receptive field profiles. *Vision Research* 20, 10 (1980), 847–856.
- [19] DAUGMAN, J. G. Uncertainty relation for resolution in space, spatial frequency, and orientation optimized by two-dimensional visual cortical filters. *Journal of the Optical Society of America A* 2, 7 (1985), 1160–1169.
- [20] 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*, [véase también *IEEE Transactions on Signal Processing*] 36, 7 (1988), 1169–1179.
- [21] 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.
- [22] DAVIS, L. S. Polarograms: A new tool for image texture analysis. *Pattern Recognition* 13, 3 (1981), 219–223.
- [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] EASLEY, G., LABATE, D., AND LIM, W.-Q. Sparse directional image representations using the discrete shearlet transform. *Applied and Computational Harmonic Analysis* 25 (2008), 25–46.
- [25] EGLIN, V., BRES, S., AND RIVERO, C. Hermite and Gabor transforms for noise reduction and handwriting classification in ancient manuscripts. *International Journal on Document Analysis and Recognition* 9 (2007), 101–122.
- [26] 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.
- [27] ESTUDILLO-ROMERO, A., AND ESCALANTE-RAMIREZ, B. Advances in rotation-invariant texture analysis. En *Progress in Pattern Recognition, Image Analysis, Computer Vision, and Applications* (2009), vol. 5856 de *LNCS*, pp. 145–152.

- [28] ESTUDILLO-ROMERO, A., AND ESCALANTE-RAMIREZ, B. Classification of low level visual texture features based on the Hermite transform. En *Signal-Image Technology and Internet-Based Systems (SITIS)* (2011), pp. 461–467.
- [29] ESTUDILLO-ROMERO, A., AND ESCALANTE-RAMIREZ, B. Rotation-invariant texture features from the steered Hermite transform. *Pattern Recognition Letters* 32, 16 (2011), 2150–2162.
- [30] ESTUDILLO-ROMERO, A., ESCALANTE-RAMIREZ, B., AND SAVAGE-CARMONA, J. Texture analysis based on the Hermite transform for image classification and segmentation. En *Optics, Photonics, and Digital Technologies for Multimedia Applications* (2012), vol. 8436 de *Proceedings of the SPIE*, pp. 843619–843619–12.
- [31] FERREIRA, M., SANTOS, C., AND MONTEIRO, J. Cork parquet quality control vision system based on texture segmentation and fuzzy grammar. *IEEE Transactions on Industrial Electronics* 56, 3 (2009), 756–765.
- [32] FISHER, R. A. The use of multiple measurements in taxonomic problems. *Annals of Human Genetics*, 7 (1936), 179–188.
- [33] FREEMAN, W., AND ADELSON, E. The design and use of steerable filters. *IEEE 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] GORDO, A., PERRONNIN, F., AND VALVENY, E. Large-scale document image retrieval and classification with runlength histograms and binary embeddings. *Pattern Recognition* 46, 7 (2013), 1898–1905.
- [36] GUO, D., ATLURI, V., AND ADAM, N. Texture-based remote-sensing image segmentation. En *IEEE International Conference on Multimedia and Expo, ICME* (2005), pp. 1472–1475.
- [37] HALEY, G., AND MANJUNATH, B. Rotation-invariant texture classification using a complete space-frequency model. *IEEE Transactions on Image Processing* 8, 2 (1999), 255–269.
- [38] HAN, J., AND MA, K.-K. Rotation-invariant and scale-invariant Gabor features for texture image retrieval. *Image and Vision Computing* 25, 9 (2007), 1474–1481.
- [39] HUBEL, D. H., AND WIESEL, T. N. Receptive fields, binocular interaction, and functional architecture in the cat’s visual cortex. *Journal of Physiology* 1, 160 (1962), 106–154.
- [40] JAFARI-KHOUSANI, K., AND SOLTANIAN-ZADEH, H. Rotation-invariant multiresolution texture analysis using Radon and wavelet transforms. *IEEE Transactions on Image Processing* 14, 6 (2005), 783–795.
- [41] JAIN, A. K., AND FARROKHNI, F. Unsupervised texture segmentation using Gabor filters. *IEEE International Conference on Systems, Man and Cybernetics* (1990), 14–19.

- [42] JENSSEN, R., AND ELTOFT, T. Independent component analysis for texture segmentation. *Pattern Recognition* 36, 10 (2003), 2301–2315.
- [43] KASHYAP, R. L., AND KHOTANZAD, A. A model-based method for rotation invariant texture classification. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 8, 4 (1986), 472–481.
- [44] KOKARE, M., BISWAS, P. K., AND CHATTERJI, B. N. Rotation-invariant texture image retrieval using rotated complex wavelet filters. *IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics* 36, 6 (2006), 1273–1282.
- [45] KOKARE, M., BISWAS, P. K., AND CHATTERJI, B. N. Texture image retrieval using rotated wavelet filters. *Pattern Recognition Letters* 28, 10 (2007), 1240–1249.
- [46] KRUIZINGA, P. Grating cell operator features for oriented texture segmentation. En *Proc. Fourteenth International Conference on Pattern Recognition* (1998), vol. 2, pp. 1010–1014.
- [47] LABATE, D., LIM, W.-Q., KUTYNIOK, G., AND WEISS, G. Sparse multidimensional representation using shearlets. En *Wavelets XI* (2005), vol. 5914 de *Proceedings of the SPIE*, pp. 59140U–59140U–9.
- [48] LAHAJNAR, F., AND KOVACIC, S. Rotation-invariant texture classification. *Pattern Recognition Letters* 24, 9-10 (2003), 1151–1161.
- [49] LAZEBNIK, S., SCHMID, C., AND PONCE, J. A sparse texture representation using local affine regions. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 27, 8 (2005), 1265–1278.
- [50] LEE, T. S. Image representation using 2D Gabor wavelets. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 18, 10 (1996), 959–971.
- [51] LIU, Y., TEVEROVSKIY, L., CARMICHAEL, O., KIKINIS, R., SHENTON, M., CARTER, C. S., STENGER, V. A., DAVIS, S., AIZENSTEIN, H., BECKER, J. T., LOPEZ, O. L., AND MELTZER, C. C. Discriminative MR image feature analysis for automatic schizophrenia and Alzheimer’s disease classification. En *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2004* (2004), vol. 3216 de *LNCS*, pp. 393–401.
- [52] 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.
- [53] MANTHALKAR, R., BISWAS, P. K., AND CHATTERJI, B. N. Rotation and scale invariant texture features using discrete wavelet packet transform. *Pattern Recognition Letters* 24, 14 (2003), 2455–2462.
- [54] MANTHALKAR, R., BISWAS, P. K., AND CHATTERJI, B. N. Rotation invariant texture classification using even symmetric Gabor filters. *Pattern Recognition Letters* 24, 12 (2003), 2061–2068.

- [55] MAO, J., AND JAIN, A. K. Texture classification and segmentation using multiresolution simultaneous autoregressive models. *Pattern Recognition* 25, 2 (1992), 173–188.
- [56] MARCELJA, S. Mathematical description of the responses of simple cortical cells. *Journal of the Optical Society of America* 70 (1980), 1297–1300.
- [57] MARTENS, J.-B. The Hermite transform-theory. *IEEE Transactions on Acoustics, Speech and Signal Processing* 38, 9 (1990), 1595–1606.
- [58] MARTENS, J.-B. The Hermite transform: a survey. *EURASIP Journal of Applied Signal Processing* 26145 (2006), 1110–8657.
- [59] MARTINEZ, A. M., AND KAK, A. PCA versus LDA. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 23, 2 (2001), 228–233.
- [60] MELSA, J. L. *Decision and estimation theory*. McGraw-Hill, 1978.
- [61] METZ, C. E. Receiver operating characteristic analysis: a tool for the quantitative evaluation of observer performance and imaging systems. *Journal of the American College of Radiology* 3, 6 (2006), 413–422.
- [62] 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, Alemania, 1994.
- [63] MONTROYA-ZEGARRA, J. A., PAPA, J. P., LEITE, N. J., TORRES, R. D. S., AND FALCÃO, A. X. Learning how to extract rotation-invariant and scale-invariant features from texture images. *EURASIP Journal on Advances in Signal Processing* 691924 (2008), 1–15.
- [64] MUNESWARAN, K., GANESAN, L., ARUMUGAM, S., AND SOUNDAR, K. R. Texture classification with combined rotation and scale invariant wavelet features. *Pattern Recognition* 38, 10 (2005), 1495–1506.
- [65] NIXON, M. S., AND AGUADO, A. S. *Feature Extraction and Image Processing*, 2 ed. Academic Press, 2008.
- [66] OJALA, T., PIETIKÄINEN, M., AND MÄENPÄÄ, T. Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 24, 7 (2002), 971–987.
- [67] OUTEX. Base de datos de texturas, University of Oulu, Finland. <http://www.outex.oulu.fi/>, 2002. Último acceso 09-09-2013.
- [68] PAN, W., BUI, T. D., AND SUEN, C. Y. Rotation invariant texture classification by ridgelet transform and frequency-orientation space decomposition. *Signal Processing* 88, 1 (2008), 189–199.
- [69] PETROU, M., AND SEVILLA, P. G. *Image processing: dealing with texture*. John Wiley & Sons, Ltd, 2006.

- [70] POLLEN, D. A., AND RONNER, S. F. Phase relationships between adjacent simple cells in the visual cortex. *Science* 212, 4501 (1981), 1409–1411.
- [71] PUN, C.-M., AND LEE, M.-C. Log-polar wavelet energy signatures for rotation and scale invariant texture classification. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 25, 5 (2003), 590–603.
- [72] RALLABANDI, V. R., AND RALLABANDI, V. S. Rotation-invariant texture retrieval using wavelet-based hidden Markov trees. *Signal Processing* 88, 10 (2008), 2593–2598.
- [73] REIS, M. S., AND BAUER, A. Wavelet texture analysis of on-line acquired images for paper formation assessment and monitoring. *Chemometrics and Intelligent Laboratory Systems* 95, 2 (2009), 129–137.
- [74] RILEY, K. F., HOBSON, M. P., AND BENCE, S. J. *Mathematical methods for physics and engineering*. Cambridge University Press, 2006.
- [75] RIVERO-MORENO, C. J., AND BRES, S. Conditions of similarity between Hermite and Gabor filters as models of the human visual system. En *Computer Analysis of Images and Patterns* (2003), vol. 2756 de LNCS, pp. 762–769.
- [76] SILVAN-CARDENAS, J., AND ESCALANTE-RAMIREZ, B. The multiscale Hermite transform for local orientation analysis. *IEEE Transactions on Image Processing* 15, 5 (2006), 1236–1253.
- [77] TUCERYAN, M., AND JAIN, A. K. Texture analysis. *Handbook of pattern recognition & computer vision* (1993), 235–276.
- [78] VAN DIJK, A. M., AND MARTENS, J.-B. Image representation and compression with steered Hermite transforms. *Signal Processing* 56, 1 (1997), 1–16.
- [79] VARMA, M., AND ZISSERMAN, A. Texture classification: are filter banks necessary? En *Proc. of the IEEE conference on Computer Vision and Pattern Recognition* (2003), vol. 2, pp. 691–698.
- [80] VISTEX. Base de datos de texturas, MIT Media Lab Vision Textures. <http://vismod.media.mit.edu/pub/VisTex/>, 2002. Último acceso 09-09-2013.
- [81] YANG, X., AND TIAN, Y. Texture representations using subspace embeddings. *Pattern Recognition Letters* 34, 10 (2013), 1130–1137.
- [82] YOUNG, R. A. Orthogonal basis functions for form vision derived from eigenvector analysis. En *ARVO Abstracts* (Sarasota, FL, 1978), Association for Research in Vision and Ophthalmology, p. 22.
- [83] YOUNG, R. A., AND LESPERANCE, R. M. The Gaussian derivative model for spatial-temporal vision: II. Cortical data. *Spatial Vision* (2001).
- [84] ZHANG, J., MARSZALEK, M., LAZEBNIK, S., AND SCHMID, C. Local features and kernels for classification of texture and object categories: a comprehensive study. *International Journal of Computer Vision* 73 (2007), 213–238.