



UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO
PROGRAMA DE MAESTRÍA Y DOCTORADO EN INGENIERÍA
INGENIERÍA ELÉCTRICA - PROCESAMIENTO DIGITAL DE SEÑALES

CARACTERIZACIÓN DE TEXELES INVARIANTE A ROTACIÓN MEDIANTE
LA TRANSFORMADA HERMITE EN MULTIESCALA

TESIS
QUE PARA OPTAR POR EL GRADO DE:
MAESTRO EN INGENIERÍA

PRESENTA:
ING. ERIK YIDELL CARBAJAL DEGANTE

TUTOR:
DR. BORIS ESCALANTE RAMÍREZ
FACULTAD DE INGENIERÍA

CIUDAD DE MÉXICO, DICIEMBRE 2016

JURADO ASIGNADO:

Presidente: Dr. Pablo Roberto Pérez Alcazar

Secretario: Dr. Fernando Arámbula Cosío

Primer vocal: Dr. Boris Escalante Ramírez

Segundo vocal: Dr. Jesús Savage Carmona

Tercer vocal: Dr. Edgar Garduño Ángeles

Lugar donde se realizó la tesis:

Laboratorio Avanzado de Procesamiento de Imágenes *LaPI*. Sección de Estudios de Posgrado, UNAM. Ciudad de México.

TUTOR DE TESIS:

Dr. Boris Escalante Ramírez

Agradecimientos y dedicatoria

Por y para Dios.

A mi madre, mi hermana y desde luego al resto de mi familia.

A Lilly y a los que considero también como parte de mi familia.

A mis amigos y compañeros de toda la vida. Grandes amistades de generación y al mejor grupo del posgrado, LaPI.

A mis profesores. A la valiosa ayuda de mi tutor y a los miembros del jurado.

A la prestigiosa UNAM y desde luego al Conacyt. He aquí un esfuerzo más.

Resumen

Las superficies texturizadas juegan un papel importante en el procesamiento de imágenes, especialmente para tareas de discriminación donde las escenas proveen de información relevante sobre los objetos capturados. Es común encontrar texturas con un comportamiento oscilatorio de píxeles que identifican distintas regiones de interés. Por esta razón, polinomios que asemejan la forma regular de las texturas son utilizados para resaltar el grado de variación entre los patrones. En consecuencia, el orden, dominio, e incluso el valor de la escala de un polinomio están ligados al tipo de textura.

El objetivo de este estudio es generar una base consistente para estimar la ventana óptima de análisis llamada *texel*, útil para detectar y capturar el elemento fundamental de textura y evitar redundancia. La herramienta se vuelve invariante a rotación cuando el llamado texel tiende a mantener su dimensión independientemente de la orientación. Dado que la formación de tales elementos se efectúa sobre cierta dirección, encontrar el patrón elemental que genera las texturas regulares implica analizar sobre un ángulo arbitrario.

La primera parte de este trabajo se llevó a cabo buscando una dirección dominante mediante la transformada de Radón. Subsecuentemente se calcularon los estadísticos de segundo orden los cuales han mostrado gran habilidad para caracterizar texturas, esto se realizó mediante la suma y diferencia de histogramas (SDH) que opera en función de la probabilidad de ocurrencia entre pares de píxeles. Aplicar conjuntamente estas técnicas nos permite abordar el modelo como un problema de optimización, computacionalmente resoluble. Para demostrar la efectividad de la propuesta, una base de texturas sintéticas fué evaluado bajo ciertas condiciones de degradación con el fin de extraer el texel correspondiente.

La sección final establece que la versión multiescala de la transformada de Hermite logra enfatizar las estructuras mas importantes al modelar el sistema de visión humano con alto nivel de precisión. A fin de validar nuestro método, realizamos algunos experimentos de clasificación utilizando un conjunto selecto de texturas naturales de dos bases de datos. Este enfoque fué comparado con metodologías tradicionales de ventanas fijas encontradas en la literatura. Los resultados son mostrados en base al número de subimágenes extraídas por clase que han sido clasificadas correctamente. Finalmente, los resultados indican que el uso de texeles tienden a ser un buen descriptor cuando se requiere la cantidad mínima de información espacial necesaria para definir imágenes completas.

Abstract

Textured surfaces play an important role in image processing, specially for discrimination tasks where scenes provide relevant information about the objects captured. It is common to find textures with an oscillatory behavior that identify different regions of interest. For this reason, polynomials that resemble regular shape of textures are used to highlight the variation degree among all patterns. Therefore, the order of polynomial, domain, and the scale are often linked to the texture type.

The aim of this study is to generate a consistent basis for estimating the optimal window of analysis called *texel*, useful to detect and capture the fundamental texture element and avoid redundancy. A rotation-invariant tool is obtained when the so-called *texel* tends to keep its dimension regardless its orientation. Since the formation of such elements is accomplished for a certain direction, finding the elemental pattern that generates regular textures involves analyzing an arbitrary angle.

The first part of this work was carried out by searching a dominant direction through the Radon transform. Subsequently, the second order statistics have shown high ability to characterize textures. The sum and difference of histograms (SDH) operate in function of the probability of occurrence between pairs of pixels. The joint application of these techniques allows to address the model as an optimization problem, solvable computationally. In order to demonstrate the robustness and effectiveness of the proposal, a set of synthetic textures was evaluated under certain conditions of degradations.

The last section states that the multiscale version of the Hermite transform achieves to emphasizes the most important structures by modeling the human visual system with high level of accuracy. In order to validate our method, we performed some classification

experiments using a set of natural textures selected from two well-known databases. This approach was compared to traditional fixed-size window methodologies found in the literature. The results are shown according to the number of subimages extracted by each class that have been correctly classified. Finally, the use of texels indicates to be a good descriptor when the minimum amount of spatial information is required and needed to define complete images.

Índice general

Resumen	I
Abstract	III
Índice de figuras	IX
1 Introducción	1
1.1 Antecedentes en la estimación del texel	2
1.2 Uso de bases ortogonales para caracterización	4
1.3 Estructura de la Tesis	6
2 Caracterización de texturas	9
2.1 Fundamentos de percepción visual en texturas	9
2.1.1 Superficies de textura	11
2.2 Métodos de análisis de texturas	11
2.2.1 Métodos estadísticos	12
2.2.2 Métodos estructurales	12
2.2.3 Métodos espectrales	12
2.2.4 Métodos estocásticos	13
2.3 Características de primer orden	13
2.4 Características de segundo orden	16
2.4.1 Matrices de co-ocurrencia	16
2.4.2 Suma y diferencia de histogramas	20
2.5 Tareas de clasificación	22
2.5.1 Clasificación no supervisada	23
2.5.2 Clasificación supervisada	23

2.5.3	Clasificador k-vecinos cercanos	23
3	Transformada de Hermite	27
3.1	Descomposición local de señales	27
3.2	Transformada de Hermite cartesiana	29
3.3	Transformada de Hermite rotada	30
3.4	Transformación en multiescala	32
3.4.1	Esquema piramidal	34
3.4.2	Caso multidireccional	35
3.5	Características estadísticas de textura	36
4	Metodología	39
4.1	Detección de la orientación primaria	39
4.1.1	Transformada de Radón aplicada a texturas	41
4.1.2	Base experimental 1. Orientación	44
4.2	Estimación del Texel	48
4.2.1	Valor de contraste	48
4.2.2	Picos y picos prominentes	52
4.2.3	Identificación de picos basado en múltiples escalas	53
4.2.4	Base experimental 2. Ventanas óptimas	55
5	Clasificación de estructuras	59
5.1	Extracción de características	59
5.1.1	Vector de características con la Transformada Hermite cartesiana en multiescala	60
5.1.2	Vector de características con la Transformada Hermite rotada en multiescala	61
5.2	Proceso de clasificación	62
5.2.1	Holdout validation	63
5.2.2	Random-fold cross-validation (RFCV)	63
5.3	Evaluación de la predicción en clasificación	65
5.4	Medidas de rendimiento	66

<i>Índice general</i>	VII
6 Conclusiones	73
A Matrices de confusión de valores AP	75
Bibliografía	81

Índice de figuras

2.1	Geometría para mediciones de GLCM a 4 distancias y 4 ángulos.	19
2.2	Diagrama del proceso de clasificación.	22
2.3	Algunos métodos de clasificación.	24
3.1	Funciones de Hermite hasta cuarto orden. Polinomios de Hermite H_n en (a) y funciones de análisis D_n en (b).	30
3.2	Funciones de análisis bidimensionales $D_{n-m,m}$, en el dominio espacial (a) y en el dominio frecuencial (b).	31
3.3	Transformada Hermite de tercer orden aplicada a una imagen de textura D1 del álbum [Brodatz, 1966]. Coeficientes cartesianos en (a) y coeficientes rotados en (b). Nótese la concentración de energía sobre los coeficientes $L_{n,0}$ de la versión rotada.	32
3.4	Esquema piramidal de la transformación multiescala para los procesos de análisis y síntesis	34
4.1	Diagrama a bloques de la metodología para la estimación del texel y extracción de características con SHT multiescala.	40
4.2	Inclinación dada la orientación. La flecha indica la inclinación para texturas con elementos ubicados en intervalos regulares en (a) e irregulares en (b).	41
4.3	Transformada de Radón.	42
4.4	Funciones. (a) Señal senoidal base. (b) Imagen sintética generada. (c) Sinograma de la imagen sintética desde 0° a 179°	45
4.5	Varianza en (a) y su segunda derivada como detector de picos prominentes en (b). Nótese el valor mínimo de la segunda derivada es en $\theta = \alpha = 0$	46

4.6	Transformada de Radón para distintos ángulos θ dónde $R(r, \theta = 0, 5, 10, 20, 45)$. Notar que $R(r, \theta = 0)$ corresponde a la misma forma de onda de la función base utilizada para generar la textura sintética.	46
4.7	Rotación y degradación. En (a) rotación sin degradación para $\theta^r = \{0, 20, 45, 80\}$ grados respectivamente. En (b) la textura rotada a $\theta^r = 30$ en sentido opuesto al avance de las manecillas del reloj, presenta degradación Gaussiana de $\sigma^2 = \{0.03, 0.1, 0.8, 1.6\}$. Las figuras en (c) presentan ruido impulsivo para $\theta^r = 60$ de $Oc = \{10\%, 30\%, 60\%, 90\%\}$	47
4.8	Error MAE en la detección de la orientación. Con ruido Gaussiano en (a) y con ruido impulsivo en (b).	48
4.9	Toeplitz de 8 niveles de cuantización para la métrica de contraste.	51
4.10	Curvas de contraste. Se muestran tres texturas en (a), periódica sintética, regular e irregular de dimensiones 512×512 pixeles. En (b) sus respectivas curvas de contraste.	52
4.11	Estimación de texel en tablero. Tablero de ajedrez de 512×512 pixeles en (a). (b) Curvas de contraste para la estimación de texel sin etapa previa de DO, arroja un texel de 128×128 pixeles. (c) Estimación de texel considerando etapa previa de DO con un texel resultante de 90×90 pixeles.	56
4.12	Estimación de Texel para el tablero con ruido. En (a) se presenta el tablero rotado a $\theta^r = 20$ en sentido opuesto al avance de las manecillas del reloj, con ruido Gaussiano $\sigma^2 = \{0.06, 0.2, 1, 1.9\}$ respectivamente. Sobre (b) existe ruido impulsivo de niveles $Oc = \{20\%, 40\%, 60\%, 80\%\}$ para la imagen de rotación a $\theta^r = 70$. Gráficas del error MAE para la estimación sobre ruido Gaussiano e Impulsivo (c).	58
5.1	HT-multiescala aplicada a la imagen Kufu09e del álbum [Klette, 2010]. Coeficientes en (a) para $n = 4, s = 1$, los coeficientes en (b) son resultado de aplicar $n = 4, s = 2$. Todos los coeficientes de las escalas son procesados para generar V_{HT}	60

- 5.2 SHT-multiescala aplicada a la imagen D101 del álbum [Brodatz, 1966]. Coeficientes en (a) para $n = 4, s = 1$, los coeficientes en (b) son resultado de aplicar $n = 4, s = 2$. La concentración de energía sobre los primeros ordenes de análisis disminuye la longitud del vector V_{SHT} sustancialmente. 62
- 5.3 Validación Holdout en (a), los elementos de entrenamiento y prueba son separados en grupos determinados. La validación RFCV (b) elige aleatoriamente el conjunto de datos para entrenamiento y prueba, se efectúan K -iteraciones. 64
- 5.4 Matrices de Confusión de AP para HT-multiescala. En (a), (b) y (c) se muestra el valor AP por clase de ventanas fijas de 128×128 , fijas de 64×64 y ventanas texel para validación *Holdout*. En (d), (e) y (f) se muestra las matrices del proceso *RFCV* para ventanas fijas de 128×128 , ventanas fijas de 64×64 y ventanas texel, respectivamente. 66
- 5.5 Matrices de Confusión de AP para SHT-multiescala. En (a), (b) y (c) se muestra el valor AP por clase de ventanas fijas de 128×128 , fijas de 64×64 y ventanas texel para validación *Holdout*. En (d), (e) y (f) se muestra las matrices del proceso *RFCV* para ventanas fijas de 128×128 , fijas de 64×64 y ventanas texel, respectivamente. 67
- 5.6 Niveles de clasificación CCR. Se presentan los resultados de la herramienta HT-multiescala para la validación Holdout en (a) y RFCV en (b). La herramienta SHT-multiescala arroja resultados de Holdout en (c) y RFCV en (d). Las gráficas muestran los niveles CCR por ventana de análisis y parámetro de distancia para cierto porcentaje de vectores de entrenamiento. 70
- 5.7 Promedio de pixeles para cada clase ocupando ventanas texel y ventanas de tamaño fijo. 71
- 5.8 Ventanas de análisis para las imágenes D20 y D64 en (a) y (b) del álbum [Brodatz, 1966] y Knoxville del álbum [Klette, 2010] en (c). Los tamaños de ventana por columna de imagen son de 128×128 , 64×64 y texel respectivamente. 72

A.1	Matriz de confusión de valores AP para la herramienta HT-multiescala con validación Holdout. (a), (b) y (c) con ventanas 128×128 , 64×64 y texel, respectivamente.	77
A.2	Matriz de confusión de valores AP para la herramienta HT-multiescala con validación RFCV. (a), (b) y (c) con ventanas 128×128 , 64×64 y texel, respectivamente.	78
A.3	Matriz de confusión de valores AP para la herramienta SHT-multiescala con validación Holdout. (a), (b) y (c) con ventanas 128×128 , 64×64 y texel, respectivamente.	79
A.4	Matriz de confusión de valores AP para la herramienta SHT-multiescala con validación RFCV. (a), (b) y (c) con ventanas 128×128 , 64×64 y texel, respectivamente.	80

Capítulo 1

Introducción

El análisis de texturas ha sido de gran importancia en muchas áreas de procesamiento de imágenes y de visión. Se podrían mencionar tareas como inspección y orientación de superficies, clasificación de escenas, estimación de formas entre las más utilizadas.

Desde sus inicios, la ciencia en el análisis de texturas ha ido en busca de un modelo que pueda brindar una descripción matemática para caracterizar la textura, sin embargo, debido al hecho de que la textura puede representarse de diversas formas, no existe un modelo preciso ni un concepto que pueda aplicarse de manera global. Tal cual, las definiciones que se han realizado sobre las imágenes de textura son tomadas como conceptos meramente generales, las cuales involucran reglas caracterizadas por una distribución espacial de niveles de intensidad en una vecindad. Es aceptado que dichas reglas, en muchas ocasiones, muestran que las imágenes poseen un comportamiento específico de manera local, así como propiedades estadísticas que incitan a pensar que las texturas están caracterizadas por variaciones espaciales con algún tipo de periodicidad. Esta propiedad está directamente relacionada al material, aspereza, densidad, tamaño o forma de la superficie capturada, por lo que provee de información esencial cuando se analizan a fondo las imágenes.

Se deben tener en cuenta diversos factores que influyen directa o indirectamente la percepción de una imagen. Cambios de iluminación actúan modificando ciertos niveles de intensidad para distintas zonas. La resolución con la que una imagen es capturada determina la escala en que la textura es observada, es decir, una distancia focal corta

implica que el mayor número de detalles o patrones finos se vuelven notables y viceversa.

Debido a una extensa variedad de formas que dan lugar a los patrones que definen las texturas, una clasificación enfocada a la formación y estadística de sus elementos brindará una perspectiva matemática en grupos definidos que puede ir desde un límite a otro. Por un lado las **texturas estocásticas** son imágenes que carecen de ciertos niveles de homogeneidad, son generadas con base en una regla aleatoria y por lo tanto un patrón espacial repetitivo está ausente. En el otro extremo, aquellas texturas que contienen un elemento fundamental percibido en forma periódica a lo largo de la imagen son conocidas como **texturas regulares**.

Con base en esto, una superficie también es considerada como textura si posee un número contable de elementos fundamentales, llamados *texeles*. Este concepto hace notable el hecho de que se pueden representar ciertas clases de textura mediante una regla estadística de correspondencia que asigne al texel en algún lugar a lo largo del espacio de la imagen y en determinada posición. Dicha regla espacial puede ser regular o aleatoria, con una dirección prominente, variaciones en el tamaño o rotación del texel e incluso con un nivel de traslape.

Dada la posibilidad de identificar visualmente un patrón de textura elemental, vuelve factible el hecho de relacionar todas las posibles propiedades de un texel con la imagen para intentar describirle de la misma manera que si se describiera a la imagen por completo. La identificación y extracción del texel se convierte en una tarea necesaria para caracterizar texturas y evitar información redundante que pudieran generar los patrones repetidos. Los métodos que realizan esta estimación deben ser robustos a la mayor cantidad de variaciones posibles.

1.1. Antecedentes en la estimación del texel

Los primeros trabajos que aproximan la estimación de la periodicidad sobre una textura [Terzopoulos, 1985, Parkkinen et al., 1990] proponen valores estadísticos y realizan una comparación computacional con ayuda de matrices de co-ocurrencia GLCM (del inglés Gray-Level Co-occurrence Matrix), definidas en [Haralick, 1979], dicha herramienta aplica a un enfoque estadístico de análisis a través de métricas por lo que

se obtiene un conjunto muy amplio de características. De igual forma, la complejidad computacional de GLCM aumenta en proporción a los parámetros definidos como distancia, escala y orientación, por lo que [Starovoitov et al., 1998] demuestra la eficiencia de los cálculos sobre la extracción de características y detección de patrones por medio de binarización o cuantización de niveles de intensidad. Estos avances fueron considerados por [Jan and Hsueh, 1998] al introducir la medida de covarianza como estimador de periodicidad. Un modelo eficiente de detección propuesto por [Oh et al., 1999] reduce de forma logarítmica los cálculos realizados en GLCM al plasmarlos en un vector de coincidencias de distancias y a su vez muestra la utilidad de los estadísticos de segundo orden.

Los métodos clásicos que extraen patrones repetitivos utilizan versiones modificadas de la función de autocorrelación; mientras que [Lin et al., 1997] propone la búsqueda de picos prominentes relacionadas a las oscilaciones de textura, [Leu, 2001] relaciona operadores basados en gradiente que facilitan la identificación de un periodo fundamental. Ambos métodos muestran mejoras en la búsqueda de picos que las aplicaciones tradicionales basadas en Fourier como en [Matsuyama et al., 1983], al ser menos sensibles a irregularidades.

Por otra parte, enfoques basados en transformaciones también han sido aplicados en tareas de extracción. La periodicidad y orientación de los patrones como parte de los descriptores de textura en tareas de adquisición fueron hechas por [Wu et al., 2000] mediante el uso de funciones de Gabor. De manera similar, la descomposición piramidal con Wavelets permite un análisis en múltiples resoluciones, la localización de contornos prominentes que dan lugar a la textura pueden detectarse en conjunto con la transformada de Hough como en [Lee and Chen, 2002]. De manera similar lo mostrado por [Asha et al., 2013], en el cual, dada una función de coincidencias, detecta la periodicidad de las curvas utilizando las funciones de Haar.

Una solución a la identificación de texeles es suponer que dichos patrones se presentan en formas cuadradas, por lo que una rejilla de análisis y un modelo estructural, como los utilizados por [Hays et al., 2006, Han et al., 2008], que permiten aproximar el tamaño de rejilla al comportamiento de las oscilaciones de los píxeles. Una versión modificada de rejilla presentada por [Gui et al., 2011] utiliza métodos de triangulación para realizar un cálculo de distancias entre elementos.

Las medidas de entropía presentadas por [Grigorescu and Petkov, 2003] sugieren que la cantidad de incertidumbre en la predicción de un evento probabilístico se reduce de manera proporcional a la entropía. De igual forma el concepto de homogeneidad mostrado por [Lizarraga-Morales et al., 2013] como indicador de similitud entre vecindades de pixeles a cierta distancia, en este caso el uso de GLCM mediante suma y diferencia de histogramas (SDH) propuesto por [Unser, 1986], brinda un soporte computacional eficiente. Finalmente, la métrica de correlación, aunada al análisis Cepstral, es empleada en [Carbajal-Degante et al., 2016] para la estimación de ventanas óptimas y su aplicación en tareas de clasificación, donde se logra reducir el espacio de características y evitar la inestabilidad en el uso de polinomios ortogonales de orden superior.

1.2. Uso de bases ortogonales para caracterización

El continuo desarrollo de técnicas en análisis de texturas ha permitido el incremento de información que se extrae de una región de interés o de manera global. Específicamente, las bases ortogonales han mostrado ser muy útiles para describir formas mediante la proyección de las imágenes sobre un conjunto de funciones que simulan el comportamiento oscilatorio de las formas y de los patrones. En sus comienzos, [Hu, 1962] introdujo los momentos invariantes basados en métodos algebraicos, posteriormente [Teh and Chin, 1988, Belkasim et al., 1991] realizaron la aplicación en imágenes de algunos momentos ortogonales como son Legendre, Zernike y pseudo-Zernike midiendo la sensibilidad al ruido, la redundancia y la capacidad de representación. Sin embargo, un problema común es el error de discretización que se acumula por el uso de momentos de orden alto, por lo cual, [Liao and Pawlak, 1996] propusieron un método para mantener el error bajo cierto nivel. Sin embargo, dichos momentos requieren de una transformación del espacio de la imagen a un sistema de coordenadas circulares donde los polinomios están definidos.

Las limitaciones de los momentos continuos consideradas por [Mukundan et al., 2001], lo llevan a presentar los momentos discretos de Tchebichef (DTM), cuyo proceso de ortonormalización elimina gran parte de la inestabilidad numérica. Un estudio similar es propuesto por [Yap et al., 2003] pero utilizando los polinomios de Krawtchouk y discute los aspectos computacionales para reducir el tiempo de cálculo de los

momentos. Una contribución reciente, hecha por [Marcos and Cristóbal, 2013], brinda el soporte matemático para generar un vector de características utilizado en tareas de clasificación. Los resultados de tal estudio muestran un buen desempeño de la técnica cuando las ventanas son relativamente grandes y utilizando un porcentaje alto de traslape, lo que implica manejar abundante información, ya que los momentos discretos son de igual dimensión que el tamaño de la ventana.

Metodologías alternativas, que también utilizan como base un kernel de polinomios, contribuyen al avance significativo de la representación de las funciones en espacio-escala y en procesamiento de imágenes ha ganado mucha atención. En sus inicios, [Bovik et al., 1990] realizaron un análisis de las texturas con un enfoque computacional donde definen un modelo de Gabor; posteriormente, [Manjunath and Ma, 1996] se enfocan en la construcción de un vector de características para evaluar la herramienta de Gabor en tareas de adquisición, similar a lo presentado por [Nava et al., 2012] con el uso de filtros logarítmicos de Gabor. La salida de dicho banco de filtros no es ortogonal, lo que resulta en información redundante que es proporcional a la cantidad de traslape entre cada filtro [Jain and Farrokhnia, 1991], por lo tanto, su diseño debe considerar un ajuste adecuado de parámetros para todas sus escalas y orientaciones.

La transformada Wavelet provee un marco de referencia preciso para la caracterización y análisis a diferentes escalas. La versión piramidal de Wavelet (PWT) [Mallat, 1989], utilizada en sus inicios en estudios de textura fué comparada por [Chang and Kuo, 1993] con la versión ramificada (TWT), demostrando la importancia de identificar el nivel de concentración de energía para determinadas sub-bandas y regiones sobre el dominio espacio-escala. Posteriormente [Unser, 1995] y [Arivazhagan et al., 2006] desarrollaron técnicas para extraer algunas características y clasificar estructuras. Sin embargo, una limitación importante es la variación que puede hallarse cuando las imágenes son rotadas, debido a que las texturas exhiben distintas respuestas frecuenciales a lo largo de direcciones específicas, por lo que también se han presentado en la literatura algunas maneras de realizar una clasificación invariante a rotación con Wavelets [Charalampidis and Kasparis, 2002, Manthalkar et al., 2003], los cuales requieren de una etapa de pre procesamiento o un número elevado de cálculos.

Las funciones derivadas de Gaussianas presentan la propiedad de ortogonalidad y miden con una alta precisión las señales a nivel de los campos receptivos [Young,

1991]. Estos también se definen como operadores naturales para la representación en espacio-escala [Witkin, 1983]; así que existe una relación directa entre el ancho de banda frecuencial de las funciones y los perfiles visuales encontrados en los sistemas biológicos de visión [Valois et al., 1982, Van Rullen and Thorpe, 2001]. Dichas funciones son utilizadas en [Martens, 1990] como un caso especial de la transformada polinomial cuando se utiliza una ventana Gaussiana de análisis, la cual resulta en la transformada de Hermite. Una versión mejorada de esta herramienta considera el uso de filtros rotables [Freeman and Adelson, 1991], como una manera compacta de describir las imágenes al detectar la orientación de la máxima energía local, y obtener descriptores invariantes a rotación, como se presenta en [van Dijk and Martens, 1997] para tareas de compresión de imágenes. Posteriormente [Silvan-Cardenas and Escalante-Ramirez, 2006] aprovechan las propiedades multiescala de las funciones base para construir un medio de representación más completo a través de un enfoque piramidal. La aplicación en análisis de texturas se realizó por [Estudillo-Romero and Escalante-Ramirez, 2011], mostrando a través de una reducción de dimensión en los vectores de características que en los primeros ordenes de los filtros de Hermite se concentran grandes niveles de energía útiles para discriminar texturas y estudiando los elementos estadísticos de mayor influencia para la descripción de una forma particular

1.3. Estructura de la Tesis

En el Capítulo 2 se presenta una descripción general del análisis de texturas y sus métodos. Se profundiza y se resaltan las características de primer y segundo orden que son ampliamente utilizadas en la literatura de las texturas, haciendo énfasis en la relación entre estadísticos de segundo orden y el sistema de visión humano. En el Capítulo 3 se presenta la transformada Hermite, una herramienta matemática extensa y de gran utilidad en imágenes; con ésta se resaltan los elementos más importantes de las texturas para su descripción y clasificación. Posteriormente, en el Capítulo 4, se aborda un método para estimar la orientación prominente de una textura y detectar el elemento fundamental, con esto se logra extraer el texel que adapte su forma invariante a la rotación de las imágenes. Finalmente, el Capítulo 5 muestra el esquema de clasificación utilizado para evaluar el desempeño de la metodología y los resultados

obtenidos. El Capítulo 6 concluye este trabajo y propone una serie de tareas en el área de procesamiento de señales para aplicar esta herramienta.

Capítulo 2

Caracterización de texturas

The analysis of texture is another topic that lies in a somewhat unsatisfactory state. The mathematics is easy, but the psychophysics is not, nor is it at all obvious to what extent the vagaries of the natural world allow the visual system to make use of the possible mathematical relations.

- David Marr, 1982.

2.1. Fundamentos de percepción visual en texturas

La percepción visual en texturas ha sido estudiada de manera amplia mediante experimentos físicos, evidencia psicovisual, incluso visión por computadora y gráficos. En esta medida el avance y las contribuciones han ayudado a mejorar la comprensión y brindarle nuevos enfoques de análisis.

Las primeras ideas sobre la importancia de las superficies texturizadas en imágenes, [Gibson, 1950], establecen que la textura produce un estímulo matemático y psicológico en la percepción humana; de tal forma que una textura provee de información suficiente para suponer un valor de distancia focal sobre la superficie y especificar una orientación local. Estas medidas intuitivas son resultado de un proceso de visión efectuado ya que la percepción como tal es una compilación del entorno de una persona y la interacción con el medio. Dado lo anterior, se sabe que las escenas visuales en su mayoría poseen características de textura útiles para describirlas, de manera que distintas perspectivas

de la escena proveen una variación significativa. El proceso de percepción recopilará esta información para crear un arreglo óptico en relación a los cambios de densidad de textura con el que simula el llamado flujo óptico.

Una de las primeras contribuciones notables sobre la percepción humana de texturas fué propuesta por [Julesz, 1962], en la que se llevó a cabo una investigación empírica sobre el significado perceptual de los estadísticos en patrones de textura, con el objetivo de determinar como el sistema de visión humano responde a variaciones particulares según el orden del estadístico. En primera instancia, el resultado suponía que las texturas sólo son distinguibles con combinaciones de estadísticos menores al de tercer orden. Dicho experimento consistió en incrustar una textura dentro de otra y determinar si el parche incrustado podía percibirse visualmente entre las dos texturas estadísticamente diferentes. Mediante pruebas de opinión, se lograron diferenciar fácilmente texturas con estadísticos de primer orden similares, pero con estadísticos de segundo orden distintos. Por otro lado, los experimentos mostraron que texturas con estadísticos de primer y segundo orden similares, pero de tercer orden distintos no podían ser detectados.

En trabajos subsecuentes, se demostró que la primera conjetura de Julesz era errónea al distinguir en [Julesz et al., 1978], texturas con estadísticos de segundo orden similares. Esto reveló que el mecanismo del sistema de visión humano (SVH) no necesariamente utiliza estadísticos de tercer orden o superior para percibir diferentes tipos de textura.

Los estudios de percepción del sistema de visión humano y la psicofísica se han enfocado en el desarrollo de modelos psicológicamente plausibles para discriminación de texturas. Éstos modelos determinan que mediciones en las variaciones de textura son más sensibles a detectarse por el SVH. Algunos estudios han mostrado evidencia que el cerebro humano realiza un análisis en tiempo y frecuencia de las imágenes, por lo que un modelo diseñado con bancos de filtros es factible, [Malik and Perona, 1990, Dunn et al., 1994]. Se ha observado que los filtros realizados con Gabor proveen la suficiente cantidad de estadísticos para su fácil percepción, debido a que son modificables en escala y orientación, [Manjunath and Ma, 1996]. Una opinión relevante de la neurobiología menciona que la corteza visual efectúa un método de aprendizaje llamado *sparse coding* donde una imagen es representada ocupando simultáneamente un número limitado de

neuronas activas [Olshausen and Field, 1997].

Un modelo general que es útil en el procesamiento de imágenes fué propuesto por [Young, 1991]. El ancho de banda en el dominio espacio-frecuencia observado en un conjunto de muestras de células simples de campos receptivos de los monos, variaban entre 0.4 hasta 4 octavas. Esta variación está directamente relacionada con modelos basados en derivadas de Gaussianas ya que cada función difiere de otras en anchos de bandas similares. Esto demuestra que las derivadas de Gaussianas miden el campo receptivo con mayor precisión que las funciones de Gabor, [Young and Lesperance, 2001], con la ventaja adicional de que las funciones presentan la propiedad de ortogonalidad.

2.1.1. Superficies de textura

La literatura define dos formas principales que describen una textura con respecto al observador. Se puede identificar una distancia local entre sus elementos fundamentales o especificar su orientación relativa, [Marr, 1982]. A su vez, la orientación de una superficie es naturalmente dividida en dos componentes, el ángulo en el cual la superficie simula un efecto de distancia y la dirección hacia donde toma lugar. [Stevens, 1979] establece una interrogante sobre cual de estas cantidades de distancia, ángulo de inclinación y dirección puede ser extraída directamente de medir las variaciones de textura. Dicho estudio se concluye de la siguiente manera:

- Dirección de inclinación es probablemente extraída explícitamente.
- Distancia es probablemente extraída explícitamente.
- Angulo de inclinación se infiere mediante la diferenciación de las estimaciones a distancias escaladas.

2.2. Métodos de análisis de texturas

El hecho de que la textura tenga un concepto no muy bien definido por su extensa área de estudio permite diferentes formas de analizarla, en tal caso, el análisis tiende a basarse mas en la aplicación que en sus fundamentos. La literatura ([Jain et al.,

1995, Petrou and Garcia-Sevilla, 2006, Mirmehdi et al., 2009]) establece varios posibles enfoques para abordar esta situación, los cuales pueden describirse como en las siguientes clases:

2.2.1. Métodos estadísticos

Un conjunto de características estadísticas se utiliza para representar la textura. Es particularmente útil cuando los patrones de textura son pequeños, resultando en *microtexturas*. Generalmente no es posible reconstruir las imágenes desde el espacio de características, por tal motivo esta clase de métodos son usualmente aplicados a tareas de clasificación. Las matrices de coocurrencias y la función de autocorrelación son ejemplos de estos métodos.

2.2.2. Métodos estructurales

Algunos tipos de textura pueden ser vistos como patrones bidimensionales que constituyen un conjunto de primitivas los cuales obedecen ciertas reglas de posicionamiento. En un principio, es necesario determinar la forma y las propiedades básicas que usualmente dibujan las *macrottexturas* y por ende, determinar las reglas de posición que las gobiernan.

2.2.3. Métodos espectrales

En ocasiones resulta conveniente representar las imágenes en otro dominio de forma que facilite la extracción de ciertas características que pueden ser de mayor importancia. Las técnicas espectrales son métodos en los que un conjunto de filtros actúan en respuesta a una señal de entrada, por lo que suelen ser eficientes y de buena precisión detectando canales de frecuencia u orientaciones del patrón de textura. Representar la señal en el dominio de la frecuencia mediante la transformada de Fourier, filtros de Gabor, polinomios de Hermite, son técnicas espectrales aplicadas a imágenes.

2.2.4. Métodos estocásticos

También conocidos como métodos basados en modelos o métodos paramétricos. Se asume que la textura es una realización de un proceso estocástico que es regida por ciertos parámetros, por lo tanto, los valores determinan las propiedades de la textura. El reto es estimar tal conjunto de parámetros para generar una textura sintética, que visualmente será similar a la textura analizada. Los campos aleatorios de Markov el nivel de intensidad de los pixeles es modelado como una combinación lineal de las intensidades de vecinos, una serie de parámetros es estimado y comparado con aquellas clases conocidas para determinar una clase en particular.

2.3. Características de primer orden

Asumir una imagen como una función $I = f(x, y)$, con dos variables espaciales $x = 0, 1, \dots, N - 1$ e $y = 0, 1, \dots, M - 1$. La función $f(x, y)$ puede tomar valores discretos $g = 0, 1, \dots, G - 1$, donde G es el número total de niveles de intensidad. El histograma de niveles de intensidades es una función que muestra el número de pixeles en la imagen completa,

$$h(g) = \sum_{x=0}^{N-1} \sum_{y=0}^{M-1} \delta(f(x, y), g), \quad (2.1)$$

dónde $\delta(I, g)$ es la función delta de Kronecker. El histograma de niveles de intensidad es una suma de la información estadística contenida en la imagen. El cálculo de histogramas de niveles de gris envuelve pixeles simples. Por lo que el histograma contiene información estadística de primer orden. Es común dividir el valor de $h(g)$ por el número total de pixeles, esto aproxima la densidad de probabilidad de ocurrencia de los niveles de gris.

$$H(g) = \frac{h(g)}{NM}. \quad (2.2)$$

De igual manera, suponiendo una imagen a color, comúnmente con representación en el espacio RGB. La generalización para histogramas de color viene dada por:

$$\chi(\bar{c}) = \frac{h(\bar{c})}{NM}; \bar{c} \in \Omega, \quad (2.3)$$

dónde Ω denota al espacio tridimensional $G_1 \times G_2 \times G_3$ dónde G_k representa el número de niveles de intensidad por canal y $h(\bar{c})$ es el número de pixeles de valor \bar{c} en la imagen.

Como se puede ver, H y χ son aproximaciones a las funciones de probabilidad de los valores de pixeles, por lo tanto se pueden caracterizar sus propiedades con un conjunto de parámetros estadísticos. A continuación se enlistan los parámetros estadísticos de primer orden, donde P_n representa el n-ésimo parámetro de H y Ψ_n denota al n-ésimo parámetro de χ . Es importante hacer notar que en algunos casos Ψ_n puede ser un escalar o vector de Ω . Tales características pueden calcularse como se muestra a continuación,

- Media:

$$\mu = \sum_{g=0}^{G-1} gH(g) \quad \mu_c = \sum_{\bar{c} \in \Omega} \bar{c}\chi(\bar{c}) \quad (2.4)$$

- Desviación estándar:

$$\sigma = \sqrt{\sum_{g=0}^{G-1} (g - \mu)^2 H(g)} \quad \sigma_c = \sqrt{\sum_{\bar{c} \in \Omega} (\bar{c} - \mu_c)^2 \chi(\bar{c})} \quad (2.5)$$

- Skew:

$$P_3 = \sigma^{-3} \sum_{g=0}^{G-1} (g - \mu)^3 H(g) \quad \Psi_3 = \sigma_c^{-3} \sum_{\bar{c} \in \Omega} (\bar{c} - \mu_c)^3 \chi(\bar{c}) \quad (2.6)$$

- Kurtosis:

$$P_4 = \sigma^{-4} \sum_{g=0}^{G-1} (g - \mu)^4 H(g) - 3 \quad \Psi_4 = \sigma_c^{-4} \sum_{\bar{c} \in \Omega} (\bar{c} - \mu_c)^4 \chi(\bar{c}) - 3 \quad (2.7)$$

- Energía:

$$P_5 = \sum_{g=0}^{G-1} H^2(g) \quad \Psi_5 = \sum_{\bar{c} \in \Omega} \chi^2(\bar{c}) \quad (2.8)$$

- Entropía:

$$P_6 = - \sum_{g=0}^{G-1} H(g) \log_2(H(g)) \quad \Psi_6 = - \sum_{\bar{c} \in \Omega} \chi(\bar{c}) \log_2(\chi(\bar{c})) \quad (2.9)$$

- Segundo momento:

$$P_7 = \sum_{g=0}^{G-1} (g - \mu)^2 H(g) \quad \Psi_7 = \sum_{\bar{c} \in \Omega} (\bar{c} - \mu_c)^2 \chi(\bar{c}) \quad (2.10)$$

- Tercer momento:

$$P_8 = \sum_{g=0}^{G-1} (g - \mu)^3 H(g) \quad \Psi_8 = \sum_{\bar{c} \in \Omega} (\bar{c} - \mu_c)^3 \chi(\bar{c}) \quad (2.11)$$

- Cuarto momento:

$$P_9 = \sum_{g=0}^{G-1} (g - \mu)^4 H(g) \quad \Psi_9 = \sum_{\bar{c} \in \Omega} (\bar{c} - \mu_c)^4 \chi(\bar{c}) \quad (2.12)$$

La media representa el nivel promedio de las intensidades en la textura, donde la desviación estándar describe las variaciones de valores alrededor de la media, el valor *skewness* es cero si el histograma es simétrico alrededor de la media, en otro caso puede adquirir valores positivos o negativos dependiendo de la tendencia por encima o debajo de la media, por lo tanto es un indicador de simetría. La *kurtosis* es una medida de uniformidad del histograma, el valor 3 en Eq.(2.7) normaliza la *kurtosis* a cero para histogramas en forma de Gaussiana. La entropía es una medida de aleatoriedad.

En el caso visual de imágenes, la media y desviación estándar no portan información acerca de la textura, simplemente representan un parámetro en el proceso de adquisición de la imagen, como el promedio de condiciones de luz, la ganancia o ampliación. El uso de imágenes normalizadas provee de mejor precisión en discriminación de texturas que los parámetros sin normalizar.

2.4. Características de segundo orden

Los estadísticos de segundo orden operan en función de la probabilidad de ocurrencia entre pares de píxeles de acuerdo a algún valor de desplazamiento d y orientación específica θ en la imagen. La función de probabilidad en este caso es llamada matriz de coocurrencias (Gray-Level Cooccurrence Matrix, GLCM) propuesta en [Haralick et al., 1973, Haralick, 1979] como una forma de extraer los estadísticos de segundo orden particularmente en texturas.

2.4.1. Matrices de co-ocurrencia

En una matriz GLCM el número de filas y columnas es igual al número de niveles de intensidad G . El elemento $C(i, j|\Delta x, \Delta y)$ indica la frecuencia relativa con la cual dos píxeles están separados una distancia en píxeles $(\Delta x, \Delta y)$ dentro de una vecindad, posiblemente del tamaño de la imagen $N \times M$, uno con nivel de intensidad i y otro con j . Representado de otra forma, la matriz de elementos $C(i, j|d, \theta)$ contiene el valor de probabilidad de los estadísticos de segundo orden

$$C(i, j|\Delta x, \Delta y) = \frac{Q(i, j|\Delta x, \Delta y)}{(N - \Delta x)(M - \Delta y)}, \quad (2.13)$$

donde:

$$Q(i, j|\Delta x, \Delta y) = \sum_{n=1}^{N-\Delta x} \sum_{m=1}^{M-\Delta y} A,$$

cuando:

$$A = \begin{cases} 1 & \text{Si } f(n, m) = f(n + \Delta x, m + \Delta y) \\ 0 & \text{En otro caso.} \end{cases}$$

Métricas de Haralick

Un conjunto de características pueden ser extraídas desde las matrices de coocurrencia a través de las siguientes métricas:

- Energía (Segundo momento angular). Es una medida de homogeneidad global en la imagen. Una escena homogénea contiene poca diferencia de niveles de intensidad, produciendo matrices con altos valores relativos en $C(i, j)$, en este caso la suma de los cuadrados será alta.

$$F_1 = \sum_{i=0}^{G-1} \sum_{j=0}^{G-1} C^2(i, j). \quad (2.14)$$

- Contraste (Inercia). Esta métrica mide la cantidad de cambios locales. Refleja la sensibilidad de la textura en relación a cambios de intensidad. De esta manera, largas transiciones de patrones son enfatizados.

$$F_2 = \sum_{i=0}^{G-1} \sum_{j=0}^{G-1} (i - j)^2 C(i, j). \quad (2.15)$$

- Dissimilarity. Es una medida de primer orden del contraste.

$$F_3 = \sum_{i=0}^{G-1} \sum_{j=0}^{G-1} |i - j| C(i, j). \quad (2.16)$$

- Homogeneidad (Momento de diferencia inverso). Esta métrica es influenciada por la homogeneidad local de la imagen. Debido al factor de diferencia inverso, el valor final contribuye a contrarrestar áreas no homogéneas.

$$F_4 = \sum_{i=0}^{G-1} \sum_{j=0}^{G-1} \frac{C(i, j)}{1 + (i - j)^2}. \quad (2.17)$$

- Entropía. Cuantifica el nivel de aleatoriedad en la región y muestra el valor más

alto cuando los valores de $C(i, j)$ son iguales.

$$F_5 = - \sum_{i=0}^{G-1} \sum_{j=0}^{G-1} C(i, j) \log(C(i, j)). \quad (2.18)$$

- Distribuciones marginales. Se definen como probabilidades marginales de $C(i, j)$:

$$C_x = \sum_{j=0}^{G-1} C(i, j) \quad C_y = \sum_{i=0}^{G-1} C(i, j), \quad (2.19)$$

por lo tanto, los estadísticos marginales:

$$\begin{aligned} \mu_x &= \sum_{i=0}^{G-1} i \sum_{j=0}^{G-1} C(i, j) = \sum_{i=0}^{G-1} i C(i) \\ \mu_y &= \sum_{j=0}^{G-1} j \sum_{i=0}^{G-1} C(i, j) = \sum_{j=0}^{G-1} j C(j) \end{aligned} \quad (2.20)$$

$$\begin{aligned} \sigma_x^2 &= \sum_{i=0}^{G-1} (i - \mu_x)^2 \sum_{j=0}^{G-1} C(i, j) = \sum_{i=0}^{G-1} (C(i) - \mu_x(i))^2 \\ \sigma_y^2 &= \sum_{j=0}^{G-1} (j - \mu_y)^2 \sum_{i=0}^{G-1} C(i, j) = \sum_{j=0}^{G-1} (C(j) - \mu_y(j))^2 \end{aligned} \quad (2.21)$$

- Varianza (Suma de cuadrados). Esta característica coloca pesos relativamente altos sobre elementos que difieren del valor promedio de $C(i, j)$

$$F_6 = \sum_{i=0}^{G-1} \sum_{j=0}^{G-1} (i - \mu)^2 C(i, j). \quad (2.22)$$

- Correlación. Mide la dependencia lineal de los niveles de intensidad en una vecindad de píxeles a posición específica. A medida de que los valores son similares, la correlación es alta.

$$F_7 = \sum_{i=0}^{G-1} \sum_{j=0}^{G-1} \frac{(i - \mu_x)(j - \mu_y)(C(i, j))}{\sigma_x \sigma_y}. \quad (2.23)$$

El uso de un amplio número de intensidades G implica un almacenamiento de datos masivo, ya que el tamaño de C va en relación a $G \times G$ para cada distancia d y orientación. En tales casos se tiene la situación de que las matrices para cada característica de textura son de mayor contenido que la imagen original. Por esta razón se tiene claro que la dimensionalidad de GLCM es proporcional al tamaño de la imagen, los niveles de intensidad y el número de orientaciones analizadas.

Usualmente GLCM es aplicado después de un proceso de cuantización de niveles para reducir el tamaño de las matrices, regularmente 8 ó 16 niveles, lo cuál ha mostrado que las oscilaciones de textura se mantienen visibles para su fácil discriminación. Este pre-procesamiento es equivalente a percibir la textura desde una escala ligeramente más áspera.

Dado que la matriz de dimensiones $G \times G$ puede ser calculada para cada imagen, subimagen o incluso ventana, y para cada conjunto de parámetros (d, θ) , es útil restringir o limitar dichos valores en caso de ser requerido. La Figura 2.1 muestra la relación geométrica de las mediciones hechas por GLCM para cuatro distancias $d = \max\{|\Delta x|, |\Delta y|\}$ y los cuatro posibles ángulos de cálculo $\theta = \{0, \pi/4, \pi/2, 3\pi/4\}$ radianes.

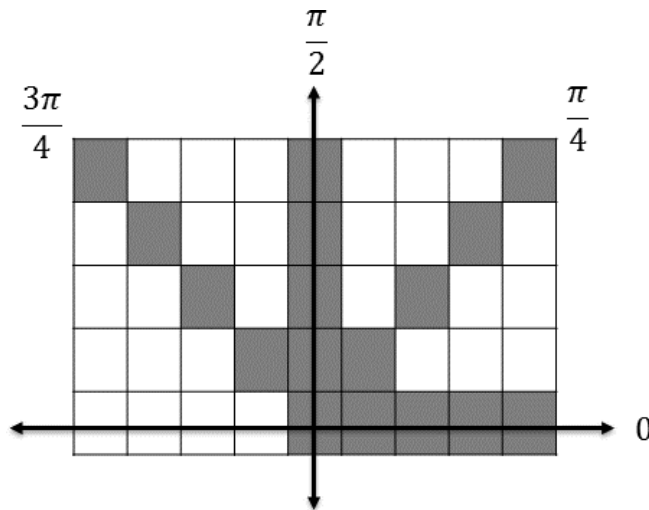


Figura 2.1: Geometría para mediciones de GLCM a 4 distancias y 4 ángulos.

Se asume simetría angular en GLCM al notar que un análisis en cierta dirección ϕ es similar a realizarlo en dirección contraria $\phi + \pi$. Se establece que $C(d, \phi) = C^T(d, \phi + \pi)$ donde C^T denota la transpuesta de C . Finalmente, incluir el cálculo

de ángulos suplementarios agrega información redundante.

Para obtener una alta precisión de la distribución de probabilidad conjunta sobre una imagen, las matrices deben contener un número elevado de nivel de ocupación. Esto puede realizarse mediante altos niveles de cuantización o ventanas de análisis grandes, lo cuál ha mostrado un aumento de dimensionalidad en proporción a los parámetros G y d respectivamente. A su vez, si se requiere evitar dependencia en cierta dirección es posible calcular un promedio de GLCM para sus cuatro direcciones (análisis isotrópico). [Conners et al., 1984] sugiere el uso de la media angular $M_F(d)$ y el rango $R_F(d)$ en la métrica de Haralick F correspondiente (Eq.(2.14) - Eq.(2.23)).

$$M_F(d) = \frac{1}{N_\theta} \sum_{\theta} F(d, \theta), \quad (2.24)$$

$$R_F(d) = \max_{\theta} \{F(d, \theta)\} - \min_{\theta} \{F(d, \theta)\}, \quad (2.25)$$

dónde la suma se realiza sobre las mediciones angulares y N_θ representa el número de cada medida. Entonces, la varianza angular independiente es definida como:

$$V_F^2(d) = \frac{1}{N_\theta} \sum [F(d, \theta) - M_F(d)]^2. \quad (2.26)$$

Mientras más grande sea el número de características de textura disponible, mayor será la correlación existente con las demás. Un post procesamiento puede ser aplicado para definir un conjunto o una combinación lineal de las características que mayor influyen en tareas de discriminación.

2.4.2. Suma y diferencia de histogramas

Se ha demostrado en [Unser, 1986], que la suma y diferencia de histogramas definen el eje principal de las funciones de probabilidad de segundo orden como un proceso estacionario. Lo cual sugiere que la descripción que establece GLCM sobre las imágenes sea reemplazada por la suma y diferencia de histogramas estimados directamente sobre la imagen.

Dada una imagen o vecindad $I = f(x, y)$ de dos variables espaciales $x = 0, 1, \dots, N-1$ e $y = 0, 1, \dots, M-1$. La función $f(x, y)$ puede tomar valores discretos $g = 0, 1, \dots, G-1$, donde G es el número total de niveles de intensidad contenidos en I . Los histogramas de Unser consideran un vector de desplazamiento $v = \{v_x, v_y\}$ como parámetro esencial que relaciona dos pixeles $\{f(x, y), f(x, y) + v\} \in I$. La suma y diferencia de pixeles para todo $\{I(x, y), I(x, y) + v\} \in I$:

$$\text{sum}(x, y|v) = I(x, y) + I(x + v_x, y + v_y), \quad (2.27)$$

$$\text{dif}(x, y|v) = I(x, y) - I(x + v_x, y + v_y), \quad (2.28)$$

por lo tanto, los histogramas de suma y diferencia son:

$$h_s(i|v) = \text{Card}\{(x, y) | \text{sum}(x, y|v) = i\} \quad , \quad i \in \{0, 1, \dots, 2G - 2\}, \quad (2.29)$$

$$h_d(j|v) = \text{Card}\{(x, y) | \text{dif}(x, y|v) = j\} \quad , \quad j \in \{-G + 1, -G + 2, \dots, G - 1\}, \quad (2.30)$$

donde el operador **Card** denota Cardinalidad. El número total de valores se mantiene para ambos histogramas

$$W = \sum_i h_s(i|v) = \sum_j h_d(j|v), \quad (2.31)$$

finalmente, los histogramas de suma y diferencia normalizados son:

$$H_s = \frac{h_s(i|v)}{W} \quad (2.32)$$

y

$$H_d = \frac{h_d(j|v)}{W}. \quad (2.33)$$

En el Cuadro 2.1 se muestra la equivalencia entre las características más utilizadas de GLCM y las asociadas con suma y diferencia de histogramas.

Parámetro	Expresión
Media	$\frac{1}{2} \sum_i i H_s(i)$
Varianza	$\frac{1}{2} (\sum_i (i - 2\mu)^2 H_s(i) + \sum_j H_d(j))$
Energía	$\sum_i \sum_j H_s(i)^2 H_d(j)^2$
Correlación	$\frac{1}{2} (\sum_i (i - 2\mu)^2 H_s(i) - \sum_j j^2 H_d(j))$
Entropía	$-\sum_i H_s(i) \log H_s(i) - \sum_j H_d(j) \log H_d(j)$
Contraste	$\sum_j j^2 H_d(j)$
Homogeneidad	$\sum_j (1 + j^2)^{-1} H_d(j)$

Cuadro 2.1: Métricas de Unser.

2.5. Tareas de clasificación

Esta sección describe la metodología de clasificación, la cuál consiste en generar un vector representativo de la imagen mediante la extracción de características para posteriormente definir una clase a través de una regla de correspondencia. Un enfoque común obedece las etapas descritas en la Figura 2.2

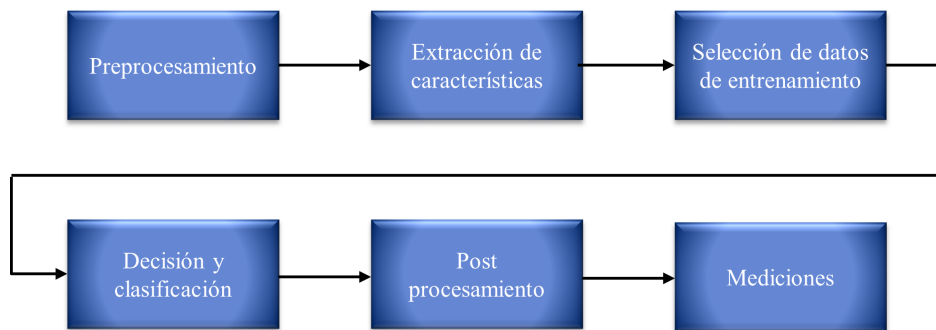


Figura 2.2: Diagrama del proceso de clasificación.

Tal proceso consiste básicamente en los siguientes pasos.

- **Preprocesamiento** es una fase de acoplamiento de las señales para su uso du-

rante el proceso de clasificación. La estimación del texel invariante a rotación es considerada durante esta fase.

- Se procede a la **detección y extracción de características** para describir los objetos y encontrar una regla de pertenencia.
- **Entrenamiento** es la fase que selecciona los atributos que mejor ayuden a relacionar el patrón correspondiente.
- **Clasificación** es el paso final que realiza la separación de los objetos en distintas clases, con base en estos resultados se realizan una serie de cálculos para medir el desempeño de los métodos de pre-procesamiento, extracción e incluso del clasificador.

2.5.1. Clasificación no supervisada

La clasificación no supervisada se realiza utilizando algoritmos de clasificación automática, no se requiere identificar alguna clase en particular pero un parámetro indispensable es definir el número total de clases que se buscan. Los algoritmos más comunes dividen el espacio de variables en una serie de regiones de manera que se propone una regla que agrupa las variables por cada clase, algunos métodos se describen en la Figura 2.3.

2.5.2. Clasificación supervisada

La clasificación supervisada se basa en la disponibilidad de clases que describen un conjunto de entrenamiento. Dicha disponibilidad incluye un número de muestras conocidas de cada clase y se intenta que estas sean lo más homogéneas posibles con respecto a sus características para facilitar su discriminación, ver Figura 2.3.

2.5.3. Clasificador k-vecinos cercanos

Es uno de los métodos más importantes y destacados por su gran funcionalidad, k -NN (k -Nearest Neighbor) asume que los datos se encuentran en un espacio de carac-

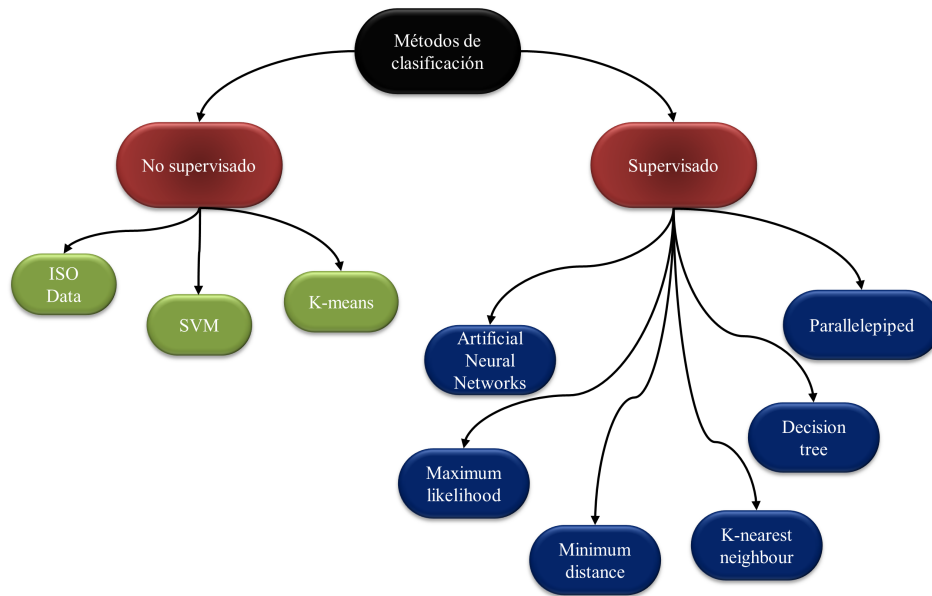


Figura 2.3: Algunos métodos de clasificación.

terísticas, tal que pueden tomar valores escalares o formar vectores simulando espacios multidimensionales, ya que los datos se encuentran en el mismo espacio se puede establecer una distancia entre ellos, regularmente euclideana.

Los datos de entrenamiento son un conjunto de valores que tienen asociada una clase, sin embargo k -NN funciona incluso con un número arbitrario de clases. El parámetro k decide cuantos vecinos deben considerarse para la clasificación. La manera más simple es considerar $k = 1$, con lo que el método se denomina el vecino más cercano (Nearest neighbour, NN).

Un enfoque para la estimación de densidad se menciona en [Bishop, 1996], utilizado en la construcción de clasificadores basados en el Teorema de Bayes. Se genera un modelo de densidades condicionales para cada clase y combinadas con probabilidades *a priori* para establecer modelos de probabilidad *a posteriori* y finalmente proponer un criterio de clasificación.

Supóngase un conjunto de \tilde{N}_k datos pertenecientes a la clase C_k y \tilde{N} datos en total. Si se dibuja una hiperesfera de volumen V con centro en \tilde{x} que contenga K_k datos de la clase C_k , es posible estimar la probabilidad condicional a partir del resultado $P(\tilde{x}) = K/\tilde{N}V$ que simboliza la probabilidad de que un patrón pertenezca al volumen V . Por

lo que la probabilidad condicional para cada clase se forma por $P(\tilde{x}|C_k) = K_k/\tilde{N}_kV$ mientras que las probabilidades a priori por clase son $P(C_k) = \tilde{N}_k/\tilde{N}$. Utilizando el teorema de Bayes:

$$P(C_k|\tilde{x}) = \frac{P(\tilde{x}|C_k)P(C_k)}{P(\tilde{x})} = \frac{K_k}{K}, \quad (2.34)$$

es decir, la probabilidad a posteriori de que el patrón \tilde{x} pertenezca a la clase C_k es igual al número de patrones K_k que pertenecen a la clase C_k dentro de un volumen que contiene un total de k elementos. Entonces, para minimizar la probabilidad de error en la clasificación de un nuevo elemento \tilde{x} , éste debe ser asignado a la clase C_k para la cual, la relación K_k/K sea mayor.

Parámetros

La elección del valor k es crítico en el proceso de clasificación. Un valor pequeño puede ocasionar errores por presencia del ruido, es decir, datos de entrenamiento que son difíciles de discriminar. Por otro lado, un valor elevado de k hace al algoritmo computacionalmente complejo. Una regla general es elegir $k = \sqrt{n}$ donde n es el total de clases.

Otro parámetro importante utilizado como criterio de decisión es la distancia. Sean dos conjuntos de puntos $x = \{x_1, x_2, \dots, x_m\}$ y $y = \{y_1, y_2, \dots, y_m\}$, la distancia en el espacio de variables se define de acuerdo a las expresiones mostradas en el Cuadro 2.2.

Distancia	Expresión
Euclideana	$d^2 = \sum_s (x_s - y_s)^2$
Minkowski	$d^p = \sum_s x_s - y_s ^p$
Mahalanobis	$d^2 = (x_s - y_s)Cov^{-1}(x_s - y_s)$
Cityblock	$d = \sum_s x_s - y_s $
Bhattacharyya	$d = \sum_s \sqrt{x_s y_s}$
Canberra	$d = \sum_s \frac{ x_s - y_s }{ x_s + y_s }$
Coseno	$d = \frac{\sum_s x_s y_s}{\sqrt{\sum_s x_s} \sqrt{\sum_s y_s}}$

Cuadro 2.2: Distancias utilizadas comunmente.

Capítulo 3

Transformada de Hermite

La gran similitud que las derivadas de Gaussianas poseen con las células de los campos receptivos y sus respectivas ventajas biológicas, como se describió en el Capítulo 2, proveen la mínima incertidumbre en el dominio tiempo-frecuencia, así como la relación de recurrencia que mantienen los polinomios de Hermite con respecto a funciones derivadas de Gaussianas, son algunos motivos para el uso de estas útiles funciones de análisis.

Los polinomios de Hermite introducidos originalmente por [Martens, 1990], y estudiados a fondo en [van Dijk and Martens, 1997, Silvan-Cardenas and Escalante-Ramirez, 2006], permiten realizar un análisis amplio de las imágenes al distinguir estructuras fundamentales tales como bordes, orientaciones o estimación de flujo en diferentes dominios como espacio, frecuencia y escala, semejante a la manera de percibir por parte del sistema de visión humano.

3.1. Descomposición local de señales

El aproximar una función a $f(x)$ se puede realizar mediante la representación de una combinación lineal de polinomios de orden n que van desde 0 a infinito. De esta manera, en sentido del error cuadrático medio, se intenta minimizar el error cuadrático ponderado ϵ definido por:

$$\epsilon = \int_{\mathbb{R}} w(x) \left(f(x) - \sum_{n=0}^{\infty} f_n \phi_n(x) \right)^2 dx, \quad (3.1)$$

dónde f_n definen los componentes polinomiales los cuales minimizan ϵ , tal que $w(x)$ es una función de peso. Los valores de f_n indican las proyecciones de la función $f(x)$ sobre la base funcional ϕ para cada n .

$$f_n = \frac{a_n}{c_n} = \frac{\langle f, \phi \rangle}{\langle \phi, \phi \rangle}. \quad (3.2)$$

Supóngase que el conjunto de funciones ϕ sirve como base generadora para crear un espacio funcional normalizado, con respecto a una ventana de análisis $v(x - k\tau)$ que localiza la señal alrededor del corrimiento $k\tau$ y a su vez contiene suficiente información de cada posición. Para tal caso se debe cumplir la condición de ortogonalidad:

$$\int_{\mathbb{R}} v^2(x) \phi_m \phi_n dx = \delta_{mn}, \quad (3.3)$$

dónde δ_{mn} representa la delta de Kronecker. Dado que el espacio funcional es de norma unitaria, los coeficientes f_n vienen dados únicamente por el valor de a_n , Eq.(3.2), que indica el producto interno entre la señal de origen y las funciones base limitadas por una ventana de análisis de posición $k\tau$ sobre x .

$$f_n(k\tau) = \int_{\mathbb{R}} f(x) \phi_n(x - k\tau) v^2(x - k\tau) dx. \quad (3.4)$$

Si los desplazamientos τ son versiones equidistantes de posiciones de ventana en periodos de T , es posible notar que la señal original se convoluciona con funciones de análisis $D_n(x) = \phi(-x)v^2(-x)$. La Eq.(3.4) toma forma en Eq.(3.5) conocida como transformada directa polinomial.

$$f_n(kT) = \int_{\mathbb{R}} f(x) D_n(x - kT) dx \quad (3.5)$$

3.2. Transformada de Hermite cartesiana

Se realiza una expansión cuando los polinomios de Hermite H_n forman una base funcional que son generados a partir de la fórmula de Rodrigues Eq.(3.6) y una ventana Gaussiana de análisis Eq.(3.7). Algunos polinomios se muestran en la Figura 3.1a.

$$H_n(x) = (-1)^n e^{x^2} \frac{d^n}{dx^n} e^{-x^2} \quad (3.6)$$

y

$$g(x) = \frac{1}{(\sqrt{\pi}\sigma)^{1/2}} e^{-x^2/2\sigma^2}. \quad (3.7)$$

Los coeficientes de Hermite $L_n(x)$ son obtenidos convolucionando la señal original $f(x)$ con las funciones de análisis de Hermite, llámese $D_n(x) = H_n(x)g^2(x)$ similar a la Eq.(3.5) con lo que se obtiene la Transformada de Hermite unidimensional.

$$L_n(x) = \int_x f(x) H_n(x - kT) g^2(x - kT) dx = \int_x f(x) D_n(x) dx. \quad (3.8)$$

Dichas funciones de análisis o también conocidos como filtros, brindan información explícita de la señal original sobre los coeficientes de Hermite. También pueden ser escritos en términos de la ventana en conjunto con los polinomios de Hermite.

$$D_n(x) = \frac{(-1)^n}{\sqrt{2^n n!}} \frac{1}{\sigma \sqrt{\pi}} H_n\left(\frac{x}{\sigma}\right) e^{-x^2/\sigma^2}. \quad (3.9)$$

Las funciones de análisis de Hermite unidimensionales, Figura 3.1b, se pueden generalizar al caso bidimensional debido a la propiedad de separabilidad espacial y simetría rotacional. La Eq.(3.9) puede reescribirse como:

$$D_{n-m,m}(x, y) = D_{n-m}(x) D_m(y), \quad (3.10)$$

dónde $n - m$ y m denotan el orden del análisis en las direcciones x e y respectivamente, un ejemplo de dichos filtros se muestran en la Figura 3.2. Como resultado, una imagen $f(x, y)$ se puede descomponer utilizando como base las funciones anteriores de manera que:

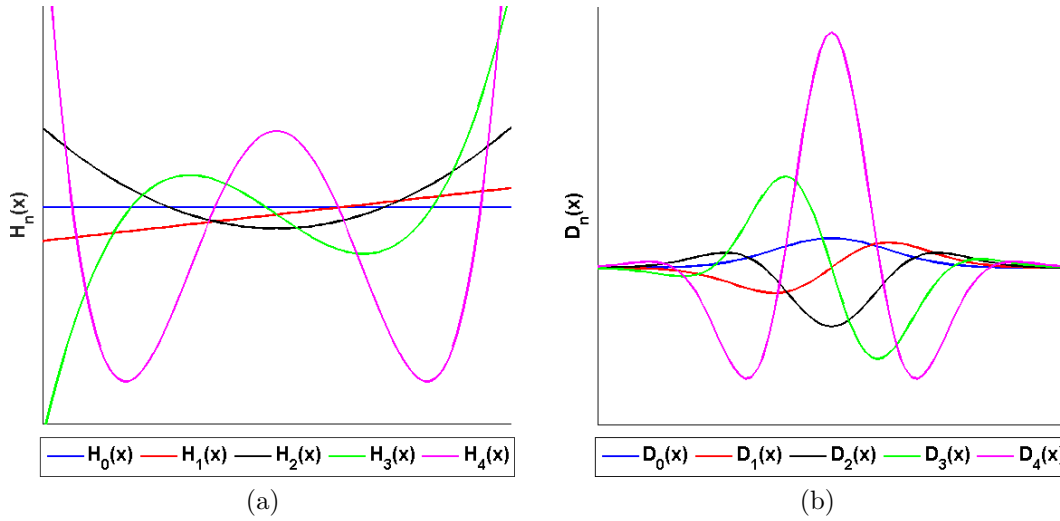


Figura 3.1: Funciones de Hermite hasta cuarto orden. Polinomios de Hermite H_n en (a) y funciones de análisis D_n en (b).

$$L_{n-m,m}(x_0, y_0) = \int_x \int_y f(x, y) D_{n-m,m}(x - x_0, y - y_0) dx dy, \quad (3.11)$$

para $n = 0, 1, \dots, \infty$ y $m = 0, 1, \dots, n$.

3.3. Transformada de Hermite rotada

Es posible construir una serie de filtros con una orientación arbitraria mediante la combinación lineal de un conjunto de filtros base, [Freeman and Adelson, 1991]. Debido a las propiedades de las funciones de análisis de Hermite, estos permiten construir versiones rotadas de un filtro de orden n al tomar combinaciones lineales de los filtros cartesianos de orden n . De esta manera, los coeficientes cartesianos de Hermite de la Eq.(3.11) son descritos a través de términos con respecto a una orientación θ , la cual es conocida como transformada de Hermite rotada por sus siglas SHT (Steered Hermite Transform) [van Dijk and Martens, 1997],

$$L_{n-m,m}^\theta(x_0, y_0, \theta) = \sum_{k=0}^n L_{n-k,k}(x_0, y_0) \alpha_{n-k,k}(\theta), \quad (3.12)$$

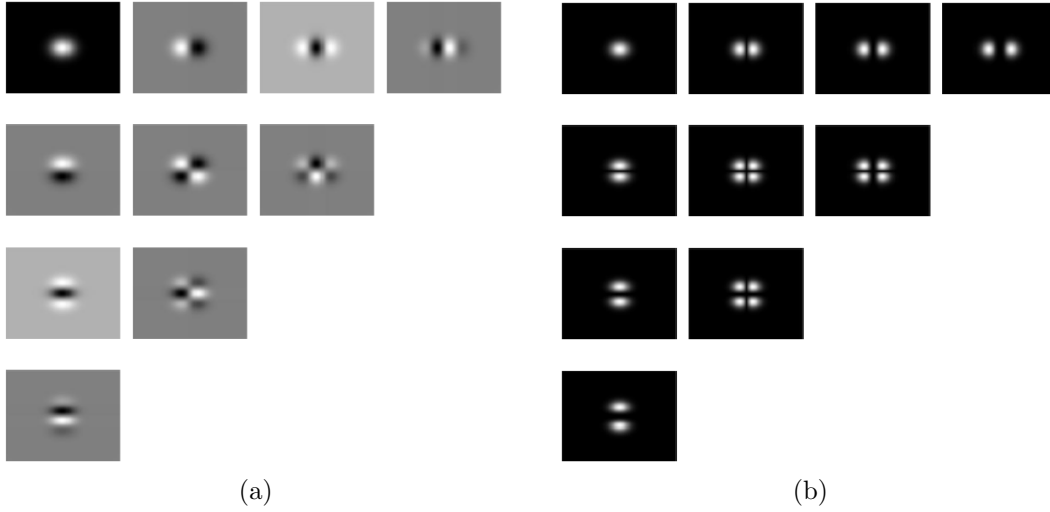


Figura 3.2: Funciones de análisis bidimensionales $D_{n-m,m}$, en el dominio espacial (a) y en el dominio frecuencial (b).

donde los términos $\alpha_{n-m,m}(\theta)$ son las funciones angulares cartesianas de orden n las cuales permiten la selección de orientación y donde C_n^m es la versión binomial,

$$\alpha_{n-m,m}(\theta) = \sqrt{C_n^m} \cos^{n-m}(\theta) \sin^m(\theta). \quad (3.13)$$

La rotación de los filtros de Hermite en cada posición sobre la imagen es una adaptación del contenido de orientación local. Primero, una transformación de Hermite se aplica para que los coeficientes sean rotados hacia un estimado de orientación local siguiendo un criterio de máxima energía orientada para cada posición de ventana, de tal manera que se puede realizar un análisis rotacional sin necesidad de fijar un número de orientaciones, Figura 3.3.

En la teoría se asume que la metodología de rotación de la Transformada de Hermite no introduce errores debido a la misma rotación, discretización, deformaciones o cambios de iluminación, el contenido de energía puede ser expresado mediante la relación de Parseval, y limitado hasta un orden E_N para $N \geq 0$.

$$E_\infty = \sum_{n=0}^{\infty} \sum_{m=0}^n [L_{n-m,m}^\theta]^2. \quad (3.14)$$

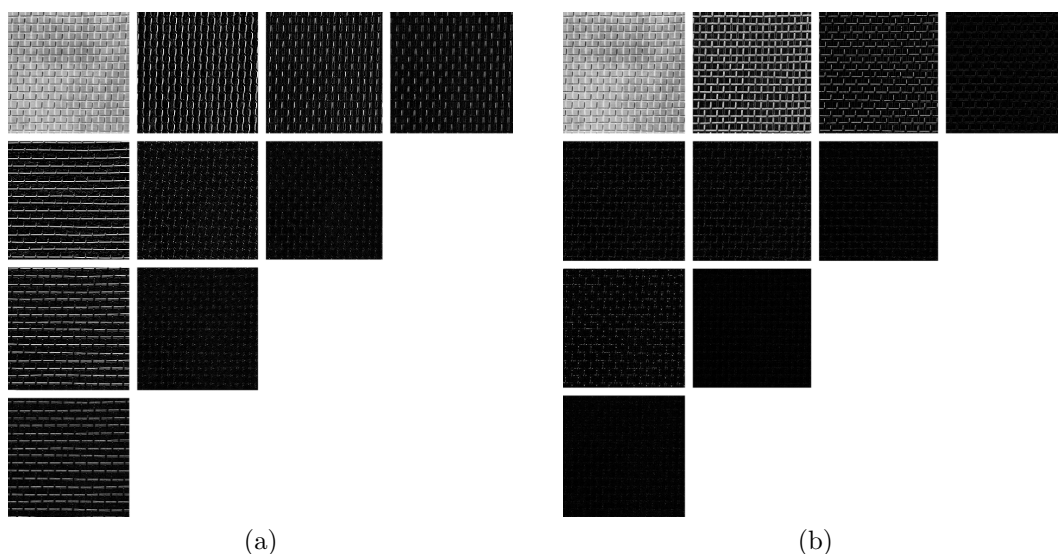


Figura 3.3: Transformada Hermite de tercer orden aplicada a una imagen de textura D1 del álbum [Brodatz, 1966]. Coeficientes cartesianos en (a) y coeficientes rotados en (b). Nótese la concentración de energía sobre los coeficientes $L_{n,0}$ de la versión rotada.

En la teoría de reconocimiento de patrones, las líneas y bordes son considerados como patrones unidimensionales por sus variaciones simples en una dirección y constantes en dirección ortogonal. Por lo tanto, se pueden distinguir patrones hasta de dos dimensiones y expresar la energía local mediante estos utilizando las siguientes ecuaciones,

$$E_N^{1D} = \sum_{n=1}^N [L_{n,0}^\theta]^2 \quad E_N^{2D} = \sum_{n=1}^N \sum_{m=1}^n [L_{n-m,m}^\theta]^2. \quad (3.15)$$

3.4. Transformación en multiescala

En muchas ocasiones, las imágenes son representaciones de una escena física, afectada por cantidades que pueden medirse, estimarse e incluso simularse. El efecto de la distancia focal se puede simbolizar como un escalamiento de la distribución de intensidades, lo que significa que en la realidad, al alejarnos de un objeto se incrementa la escala de representación y por ende los detalles finos se vuelven imperceptibles, por otro lado al acercarnos al objeto la escala disminuye y algunos patrones que requieren ma-

yor contenido espacial no son distinguidos, [Silvan-Cardenas and Escalante-Ramirez, 2006]. Entonces, la escala se encuentra directamente relacionada con los patrones que se requieren percibir y realizar un análisis a diferentes escalas se vuelve una necesidad para considerar una cantidad adecuada de perspectivas.

La representación espacio-escala emplea versiones escaladas de la ventana convolucionada con la señal de origen, $f(x, s) = f(x) * g(x, s)$, donde $g(x, s)$ es una función Gaussiana de escala $s > 0$. Éstas escalas son relaciones directas para la percepción de estructuras sobre una imagen y vinculadas con el sistema de visión humano. Los canales del campo visual a nivel retina exhiben una respuesta aproximada a una diferencia de Gaussiana (DoG) cuya respuesta al impulso se denota como:

$$DoG(x, s_1, s_2) = g(x, s_1) - g(x, s_2) \quad s_2 > s_1. \quad (3.16)$$

Por lo tanto, la respuesta de un filtro DoG contiene información de las escalas s_1 y s_2 espaciadas por un parámetro, llámese τ , que es constante para todo entero y cuya evidencia biológica muestra una relación de aproximadamente una octava, $\tau = 0.75$.

Finalmente, las derivadas de $f(x, s)$ se pueden obtener convolucionando la función $f(x)$ con los operadores de derivadas de Gaussiana $g_n(x, s) = \partial_x^n g(x, s)$, así que cada escala de análisis genera un conjunto de coeficientes de Hermite representativo de cada perspectiva al ser proyectado sobre las funciones de análisis de Hermite.

$$L_n^{(k)}(\psi) = \langle f(x), G_n^{(k)}(x, \psi) \rangle, \quad (3.17)$$

donde

$$G_n^{(k)}(x, \psi) = \frac{1}{\sqrt{2s_k}} g_n^* \left(\frac{x}{\sqrt{2s_k}} - \psi \right), \quad (3.18)$$

para $n = 1, 2, \dots$ y $s_k = (1 - \tau)^{-k} s_0, \forall k \in \mathbb{Z}$. Las cuales se conocen como transformada multiescala en Eq.(3.17) y sus funciones de análisis en Eq.(3.18).

3.4.1. Esquema piramidal

Un esquema de procesamiento con el cuál, dada una función se descompone en un conjunto de subimágenes submuestreadas en proporción a su resolución en pasobajas o pasobandas se conoce como esquema piramidal. Cada nivel de pirámide guarda una relación lineal con los niveles adyacentes y dicha relación es invariante a escala.

La relación piramidal de la transformada en multiescala se deriva del hecho de que las funciones de análisis de un orden dado forman una pirámide Gaussiana debido también a una relación lineal, dónde los coeficientes P_n de la pirámide representan en este caso las derivadas de Gaussiana a escala $\sigma = \tau T^2/4$ que se mantiene constante para todos los niveles. El término T indica la proporción de resolución entre un nivel dado y el anterior, donde $T = 1/\sqrt{1 - \tau}$.

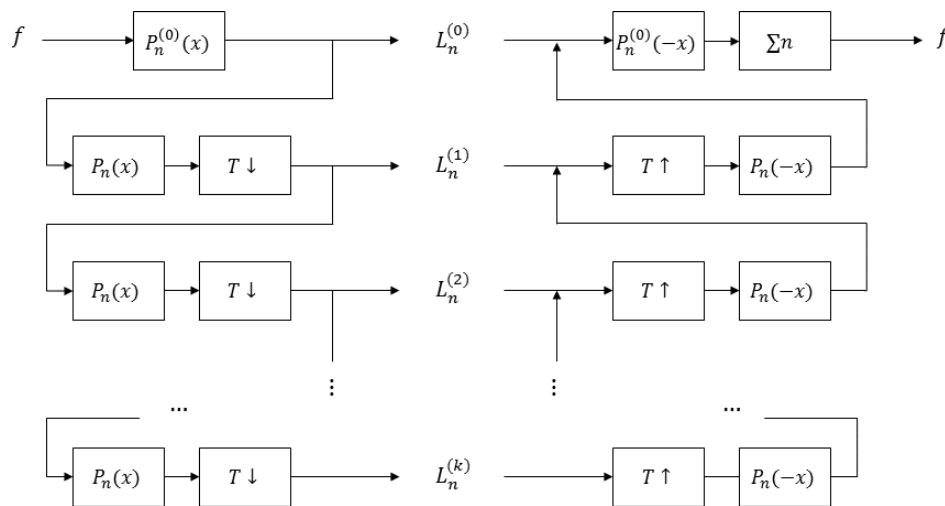


Figura 3.4: Esquema piramidal de la transformación multiescala para los procesos de análisis y síntesis

Como se observa en el diagrama de la Figura 3.4, el esquema piramidal tiene la característica de que la reconstrucción se efectúa por etapas agregando sucesivamente las contribuciones de cada nivel de la pirámide comenzando con el residuo pasobajas.

3.4.2. Caso multidireccional

El caso bidimensional de la transformación multiescala se efectúa filtrando la señal en cascada a lo largo de cada coordenada, de la misma forma para señales multidimensionales ya que las funciones Gaussianas pueden expresarse mediante productos de Gaussianas unidimensionales. Similar a las funciones de análisis de Hermite de la Eq.(3.10), considerando la escala s de resolución, los operadores bidimensionales de derivadas escaladas se definen como:

$$G_{m,n-m} = G_m(x, s)G_{n-m}(y, s), \quad (3.19)$$

para $m = 0, 1, \dots, n$ y $n = 0, 1, \dots$ donde n representa el orden de derivación total, de esta manera la descomposición de señales bidimensionales es análoga al caso unidimensional. Los operadores Gaussianos de la Eq.(3.19) tienen la propiedad de direccionalidad como en las Eq.(3.12) y Eq.(3.13) de la Transformada de Hermite rotada, lo que convierte a estas funciones sensibles a responder en determinada orientación $G_{m,n-m}(x, y, s, \theta) = G_{m,n-m}(x \cos \theta + y \sin \theta, -x \sin \theta + y \cos \theta, s)$, donde θ es el ángulo de rotación, conocidas como derivadas direccionales de Gaussianas (DDG)

Cuando el análisis requiere evaluar las DDG en multiescala, el parámetro s es factor para decidir el escalamiento que tomará la función en la resolución siguiente, como se determina en la Eq.(3.18) que establece la relación entre orden y escala la cual se expresa en la Eq.(3.20).

$$G_n^{(k,j)}(x, y, \xi, \eta) = \frac{1}{2s_k} G_{n,0}^* \left(\frac{x \cos \theta_j + y \sin \theta_j}{\sqrt{2s_k}} - \xi, \frac{-x \sin \theta_j + y \cos \theta_j}{\sqrt{2s_k}} - \eta \right) \quad (3.20)$$

dónde $\theta_j = \theta_0 + j/(n+1)$ para $j = 0, 1, \dots, n$; $n = 1, 2, \dots$ y $s_k = (1-\tau)^{-k} s_0, \forall k \in \mathbb{Z}$, lo cual se conoce como funciones de análisis de la transformación multiescala-multidireccional que emplea las DDG a múltiples escalas y orientaciones para el proceso de transformación de Hermite en multiescala-multidireccional (THMM):

$$L_n^{(k,j)}(\xi, \eta) = \langle f(x, y), G_n^{(k,j)}(x, y, \xi, \eta) \rangle \quad (3.21)$$

3.5. Características estadísticas de textura

Las características globales de textura son obtenidas a partir de los estadísticos de las imágenes. El análisis de Hermite provee un soporte amplio para representar las imágenes mediante distintas orientaciones y escalas que se definen mediante un conjunto de filtros, con lo que cada respuesta de una imagen a cada filtro produce un conjunto particular de coeficientes que describen el comportamiento de los patrones de la imagen.

Por su naturaleza, existen imágenes cuya respuesta a cada filtro es más pronunciada que otros, el nivel respuesta va en relación a la cantidad de variación de los niveles de intensidad y el tipo de estructura que se perciben, con lo que se definen también imágenes cuyo contenido frecuencial pueda ser delimitado en alguna banda. Los filtros que actúan extrayendo información de una banda distinta a la que define a la imagen, pueden ser considerados despreciables puesto que la cantidad de aportación para la caracterización es diminuta. Como lo es en el caso de la versión rotada de la Transformada de Hermite, cuando se utiliza un criterio para identificar la máxima energía a cada orientación los coeficientes de mayor valor quedan plasmados sobre los primeros ordenes de los filtros, Figura 3.3b

Lo anterior es fundamentalmente considerado por [Estudillo-Romero and Escalante-Ramirez, 2011] cuando define una serie de estadísticos que se pueden extraer sobre las imágenes filtradas de mayor contenido energético de la transformada rotada de Hermite, esto es sobre los coeficientes $L_{n,0}^\theta$.

- Media.

$$\mu = \frac{1}{H \times W} \sum_{w=1}^W \sum_{h=1}^H L_{n,0}^\theta. \quad (3.22)$$

- Desviación estándar.

$$\sigma = \sqrt{\frac{1}{H \times W - 1} \sum_{w=1}^W \sum_{h=1}^H (L_{n,0}^\theta - \mu)^2}. \quad (3.23)$$

- Energías.

$$E_0 = \sum_{w=1}^W \sum_{h=1}^H H(L_{n,0}^\theta)^2, \quad (3.24)$$

$$E_1 = \frac{E_0}{H \times W}, \quad (3.25)$$

$$E_2 = \frac{1}{H \times W} \sum_{w=1}^W \sum_{h=1}^H |L_{n,0}^\theta|, \quad (3.26)$$

y

$$E_3 = \frac{1}{H \times W} \sum_{w=1}^W \sum_{h=1}^H \sqrt{|L_{n,0}^\theta|}. \quad (3.27)$$

Capítulo 4

Metodología

El objetivo de este estudio es proponer una metodología para encontrar el tamaño adecuado de la ventana de análisis, y resolver el problema de invarianza a rotación, esto es, lograr identificar y capturar el patrón fundamental de oscilación de textura para cualquier orientación, conocido como texel, con el propósito de demostrar que el texel es capaz de describir a la imagen completa de textura mediante un proceso de clasificación.

En el esquema mostrado en la Figura 4.1 se proponen dos etapas de análisis. La primera etapa describe la obtención del texel invariante a rotación a través de un algoritmo para la detección de la orientación de una textura y subsecuentemente para la estimación de la ventana adecuada. Se mostrará que la herramienta matemática de Transformada de Radón en conjunto con SDH revela un resultado apropiado cuando se consideran aquellas características que ayudan a resaltar la identificación de patrones elementales.

4.1. Detección de la orientación primaria

La detección de la orientación en una textura es una tarea confusa debido a que los patrones fundamentales poseen variaciones de niveles de intensidad sobre una región delimitada de la imagen, dichos patrones dan lugar a contornos y formas que no necesariamente brindan una descripción completa de la orientación de la imagen, por lo tanto

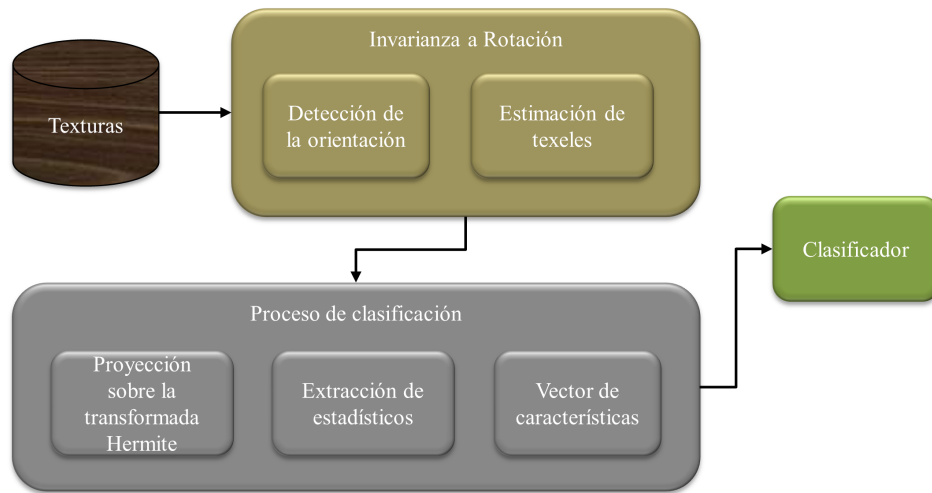


Figura 4.1: Diagrama a bloques de la metodología para la estimación del texel y extracción de características con SHT multiescala.

metodologías de estimación local como filtros direccionales sobre los ejes cartesianos [Rao and Jain, 1992] o filtros en base a gradientes no constituyen un medio para percibir si los patrones están alineados. En métodos alternativos como filtros de Gabor [Wu et al., 2000] la precisión en la detección es proporcional a la complejidad en el diseño de los filtros puesto que se establece una relación con el número de orientaciones, escalas o tamaños. Por lo tanto, el proceso completo de detección de la orientación debe verse como un problema global que involucra al conjunto de patrones de la imagen completa.

Las superficies de textura generan una respuesta en el observador de tal manera que uno de los parámetros que más influyen en la percepción es la inclinación, [Stevens, 1979]. La inclinación de una textura vista como superficie es la dirección en la cual los patrones se alejan del observador. Si la superficie mantiene una textura uniforme, la proyección del eje de inclinación en la imagen indica la dirección en la cual la densidad local de textura varía en mayor cantidad, esto es, un ángulo perpendicular a la dirección en la cual los elementos de textura están más uniformemente distribuidos, ver Figura 4.2a, esta dirección dominante de uniformidad simula un comportamiento constante de la textura sobre una orientación dominante. Por otro lado, si la textura presenta cierto nivel de variación con respecto a su inclinación, una metodología eficiente para estimar la orientación de la superficie es hallar una línea recta sobre los patrones de textura de tal manera que intersece a sus elementos en intervalos similares, ver Figura 4.2b.

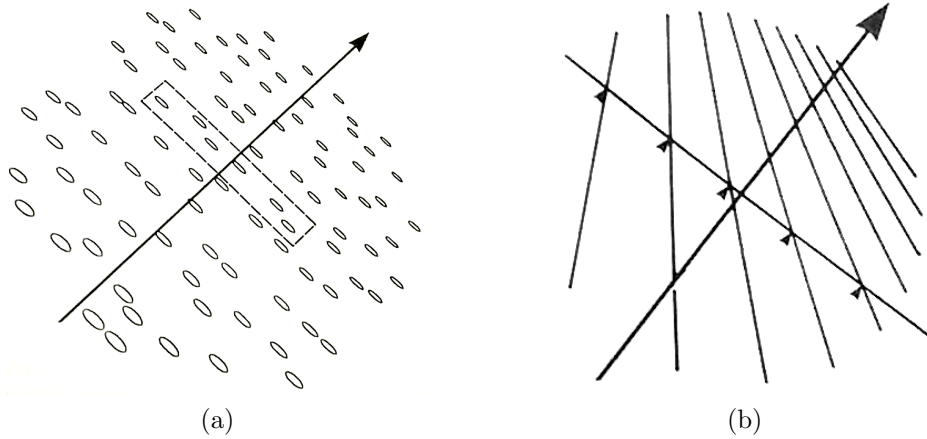


Figura 4.2: Inclínación dada la orientación. La flecha indica la inclinación para texturas con elementos ubicados en intervalos regulares en (a) e irregulares en (b).

4.1.1. Transformada de Radón aplicada a texturas

La transformada de Radón bidimensional Eq.(4.1), es básicamente la proyección radial de una función $f(x, y)$ o imagen, orientada a un ángulo específico.

$$R(r, \theta) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) \delta(r - x \cos \theta - y \sin \theta) dx dy, \quad (4.1)$$

donde la función de Dirac δ convierte la integral bidimensional en una integral de línea a lo largo de $x \cos \theta + y \sin \theta = r$. De acuerdo a un ángulo dado θ , la distancia perpendicular desde el origen es r (Figura 4.3). La función transformada $R(r, \theta)$ es también conocida como *sinograma*.

Las texturas regulares presentan patrones periódicos, donde el color, intensidad o forma de todos los elementos de la textura están repetidos en intervalos similares sobre el espacio de la imagen [Mirmehdi et al., 2009]. Algunas texturas se definen como estacionarias cuando la imagen contiene un sólo tipo de patrón, así que las propiedades estadísticas locales tienden a ser iguales.

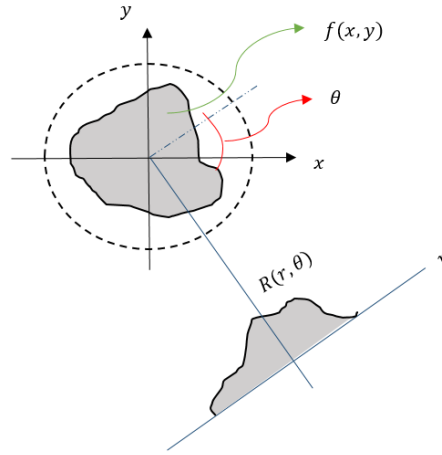


Figura 4.3: Transformada de Radón.

Propiedad de rotación

Considere una función bidimensional en coordenadas polares $f(l, \phi)$ y su transformada de Radón

$$\begin{aligned} R(r, \theta) &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(l, \phi) \delta(r - l \cos \phi \cos \theta - l \sin \phi \sin \theta) |l| dl d\phi \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(l, \phi) \delta(r - l \cos(\phi - \theta)) |l| dl d\phi, \end{aligned}$$

suponiendo un cambio de ángulo ϕ_0 de forma que $f(l, \phi - \phi_0)$ entonces:

$$R(r, \theta) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(l, \phi - \phi_0) \delta(r - l \cos(\phi - \theta)) |l| dl d\phi,$$

al introducir un cambio de variable $\phi' = \phi - \phi_0$:

$$R(r, \theta) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(l, \phi') \delta(r - l \cos(\phi' + \phi_0 - \theta)) |l| dl d\phi' = R(r, \phi_0 - \theta). \quad (4.2)$$

La ecuación anterior indica que un valor de rotación en el espacio de coordenadas de una imagen es proporcional a un desplazamiento en el dominio de Radón, manteniendo los valores del sinograma constantes.

Existe una relación entre la orientación de una textura con respecto a la Transformada de Radón. El cálculo se realiza sobre líneas rectas, partiendo del hecho de que los patrones repetitivos de textura tienden a formarse sobre líneas rectas a lo largo de la imagen con base en una orientación específica [Jafari-Khouzani and Soltanian-Zadeh, 2005]. Estimar la formación lineal de dichos patrones indicará la orientación prominente de oscilación.

El uso de texturas sintéticas provee un medio para evaluar la efectividad de los algoritmos. En [Starovoitov et al., 1998, Jan and Hsueh, 1998, Grigorescu and Petkov, 2003, Lizarraga-Morales et al., 2013], utilizan texturas sintéticas que son mapas de patrones repetitivos para estimar el grado de periodicidad, formar ventanas óptimas y medir el nivel de precisión en los cálculos. Una de las formas más simples de representar una textura sintética es mediante las variaciones periódicas de niveles de intensidad, sea binario o en varios niveles de gris. Las funciones senoidales por sus características de oscilación pueden ser utilizadas para crear una textura sintética elemental que simule los cambios periódicos de los niveles de intensidad, similar a [Leu, 2001], en el que se alternan bandas horizontales de distintos niveles de gris.

Sea $f(x) = A \sin(\omega x)$, una función senoidal con amplitud A y frecuencia angular ω , y sea $g(y) = u(y)$, donde $u(y)$ representa la función escalón de Heaviside de forma que $u(y) = \{0, y < 0; 1, y \geq 0\}$. Una función $f(x, y)$ se puede obtener mediante el producto de las funciones $g(y)f(x)$, limitado para valores de $x \in [0, M - 1]$ y $y \in [0, N - 1]$ cuyo rango estará dentro de $[M, N]$.

Dicha función simboliza variaciones senoidales de amplitud a lo largo del eje x , con lo que se distinguen bandas continuas verticales. Es decir, se alcanzan los valores originales de la función $f(x)$ cuando se observa a $f(x, y)$ en un valor fijo de y , de manera que $f(x, y = c)$. Considerando el ángulo $\theta = 0$. La transformada de Radón de $f(x, y)$ en dirección θ es:

$$R(r, \theta) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} A \sin(\omega x) \delta(r - x \cos \theta - y \sin \theta) dx dy ,$$

$$R(r, 0) = A \int_{-\infty}^{\infty} \sin(\omega x) \delta(r - x) dx .$$

Dadas las propiedades de integración de la función delta. Se obtiene:

$$R(r, 0) = A \sin(\omega r) . \quad (4.3)$$

La Eq.(4.3) indica que la forma de la función original $f(x)$ se mantiene en el dominio de la transformada de Radón para el ángulo en el cual los patrones muestran regularidad. Por lo tanto, el valor esperado es $E[f(x)] = E[R(r, 0)]$ y la varianza es $var[f(x)] = var[R(r, 0)]$. Así que el valor máximo de varianza en $R(r, \theta)$ sobre θ , se alcanza cuando existe una función de oscilación $f(x)$ orientada hacia $\theta = 0^\circ$. Un criterio utilizado para estimar el máximo prominente es a través de la segunda derivada, por lo que la detección de la orientación vendrá dada por:

$$\alpha = \arg \left(\min \frac{d^2 var[R(r, \theta)]}{d\theta^2} \right) \quad (4.4)$$

4.1.2. Base experimental 1. Orientación

Experimento 1. Con base en lo anterior se generó una textura sintética utilizando una función senoidal de parámetros $\omega = 2\pi/50$ y $M = N = 512$, con período fundamental de 50 pixeles, ver Figura 4.4. Su Transformada de Radón se aprecia en Figura 4.4c para $\theta = 0, 1, \dots, 179$. La varianza de $R(r, \theta) \forall \theta$ y su segunda derivada se muestran en la Figura 4.5. De acuerdo a la Eq.(4.4), el máximo prominente de la varianza corresponde a un mínimo global de su segunda derivada.

La Eq.(4.3) demuestra que los valores de la Transformada de Radón a la orientación estimada serán iguales a la función original, esto se aprecia en la Figura 4.6, donde se percibe que las proyecciones de la imagen sobre Radón se deforman al alejarse de la orientación prominente de la textura.

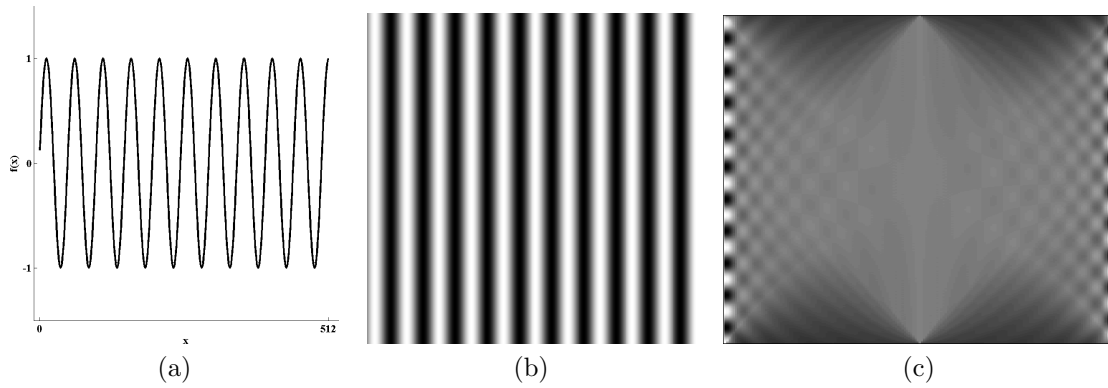


Figura 4.4: Funciones. (a) Señal senoidal base. (b) Imagen sintética generada. (c) Sinograma de la imagen sintética desde 0° a 179° .

Experimento 2. Consiste en evaluar la precisión en la estimación de la orientación de una textura bajo ciertos niveles de degradación. Utilizando el mismo esquema de generación de la textura sintética del experimento anterior, ver Figura 4.4, se hace rotar la imagen con algoritmos de interpolación bilineal de forma que el parámetro de rotación, llámese θ^r , toma valores en grados de $\theta^r = 0, 1, 2, \dots, 179$.

Para cada nivel de degradación se calcula el error absoluto como una medida que indica la cantidad de ángulos que puede variar la detección de la orientación en presencia de ciertas cantidades de ruido (Mean Absolute Error, MAE), donde θ_d es el ángulo detectado. Cuando no existe degradación sobre las imágenes rotadas, Figura 4.7a, el algoritmo arroja un valor MAE en la estimación de 0.01.

$$MAE = \frac{1}{n} \sum_{i=1}^n |\theta_i^r - \theta_i^d| \quad (4.5)$$

- La principal fuente de ruido Gaussiano que afecta las imágenes se encuentra en el proceso de adquisición causado por variaciones de iluminación, temperatura e incluso al transmitir la información mediante un canal defectuoso. Este ruido es modelable mediante la función de densidad de probabilidad con forma de Gaussiana. La Figura 4.7b muestra los efectos del ruido sobre la imagen sintética rotada. Los parámetros utilizados en este experimento son $\mu = 0$ y $\sigma^2 \in [0, 2]$, véase la Figura 4.8a.

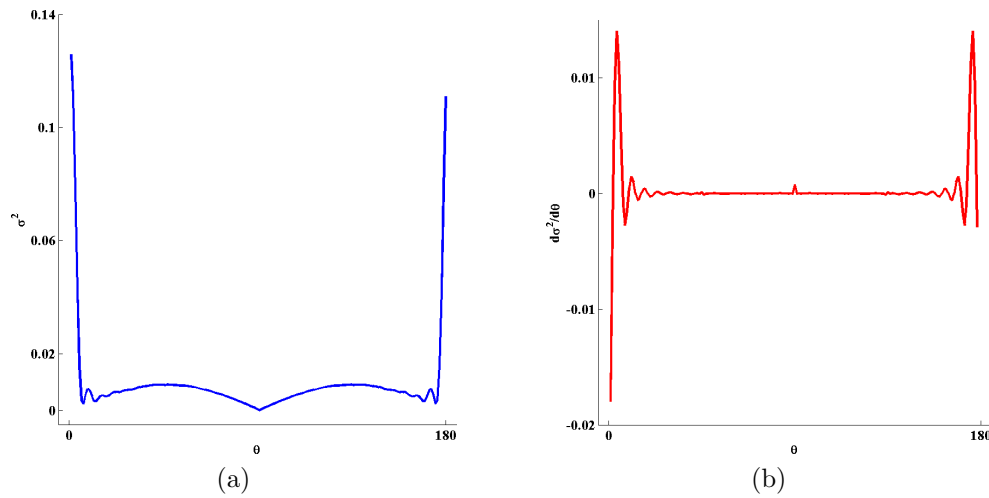


Figura 4.5: Varianza en (a) y su segunda derivada como detector de picos prominentes en (b). Nótese el valor mínimo de la segunda derivada es en $\theta = \alpha = 0$.

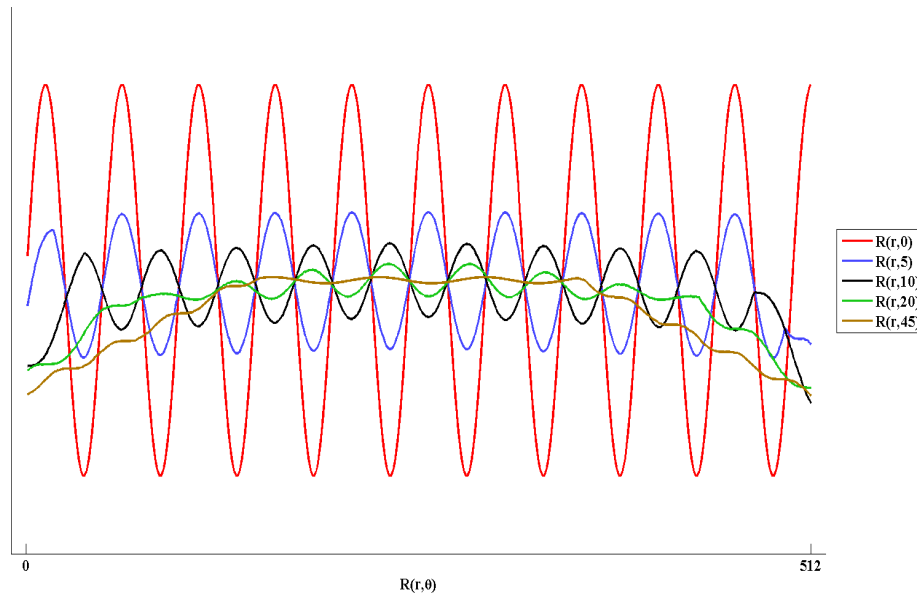


Figura 4.6: Transformada de Radón para distintos ángulos θ dónde $R(r, \theta = 0, 5, 10, 20, 45)$. Notar que $R(r, \theta = 0)$ corresponde a la misma forma de onda de la función base utilizada para generar la textura sintética.

- El ruido impulsivo es un tipo de degradación que simula errores de conversión analógico digital o el mal funcionamiento de algunos sensores en las cámaras, cuyo modelo se representa como una función delta de Kronecker al reproducir cambios bruscos en los niveles de intensidad de la imagen. La Figura 4.7c muestra

los efectos del ruido impulsivo en la imagen sintética rotada. El parámetro Oc denota el nivel de ocupación del ruido impulsivo sobre la imagen, cuando $Oc = 1$ es equivalente a obtener un 100 % de alteración. El MAE se gráfica en la Figura 4.8b.

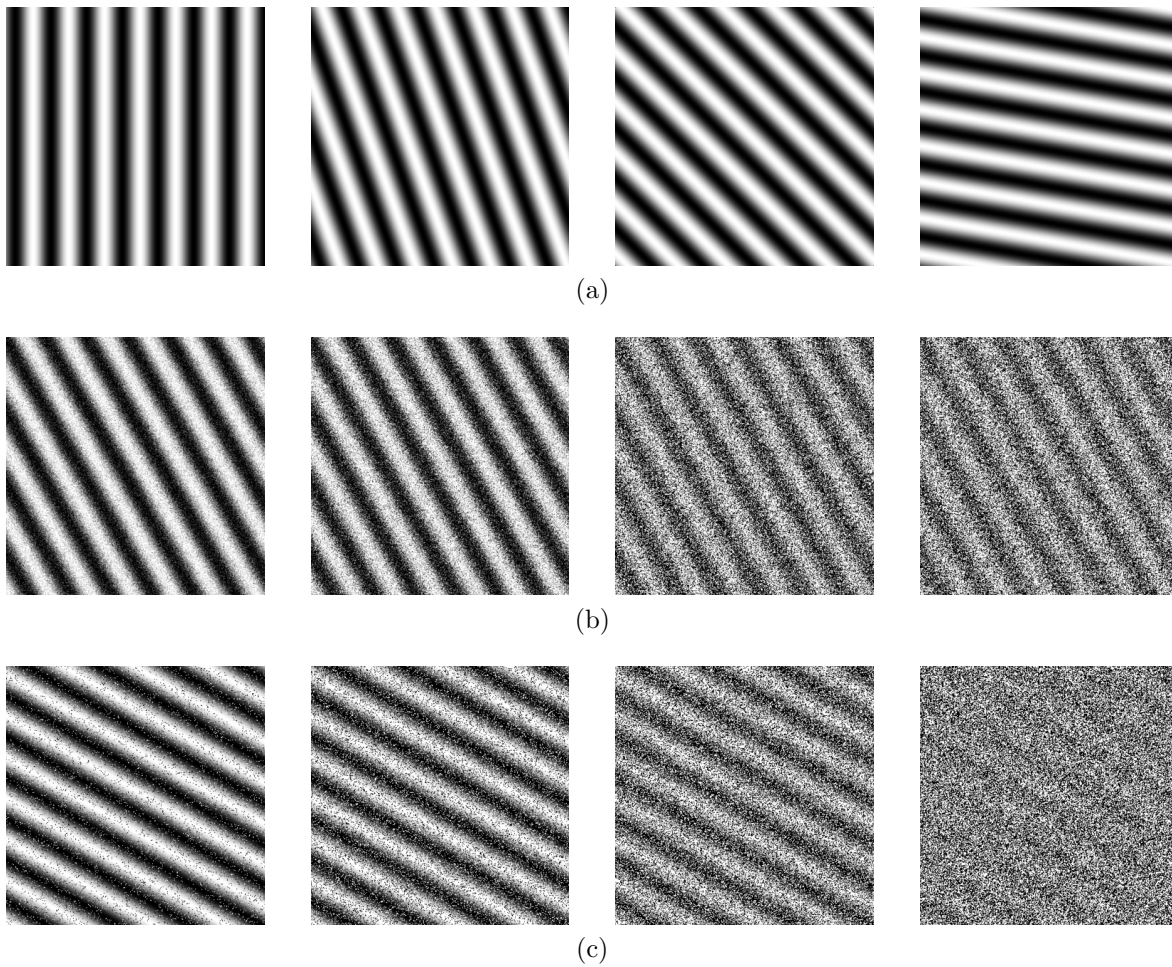


Figura 4.7: Rotación y degradación. En (a) rotación sin degradación para $\theta^r = \{0, 20, 45, 80\}$ grados respectivamente. En (b) la textura rotada a $\theta^r = 30$ en sentido opuesto al avance de las manecillas del reloj, presenta degradación Gaussiana de $\sigma^2 = \{0.03, 0.1, 0.8, 1.6\}$. Las figuras en (c) presentan ruido impulsivo para $\theta^r = 60$ de $Oc = \{10\%, 30\%, 60\%, 90\%\}$

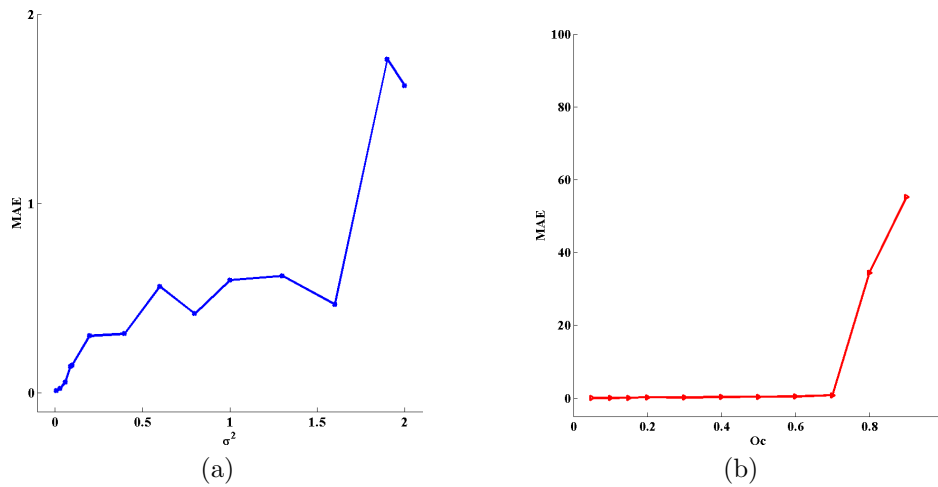


Figura 4.8: Error MAE en la detección de la orientación. Con ruido Gaussiano en (a) y con ruido impulsivo en (b).

4.2. Estimación del Texel

El proceso de estimación del texel se basa en encontrar la ventana ideal que capture la cantidad suficiente de información considerando que existen texturas cuyos patrones se ubican repetidamente sobre la imagen a cierta distancia y obedecen ciertas reglas de correspondencia que los vuelve estadísticamente similares de forma local. Existen diversas maneras en que esta ventana es calculada, en la sección de *Antecedentes* se explica de manera general algunos de los trabajos de mayor impacto, sin embargo el problema de rotación es común para la mayoría de ellos. Dicho problema consiste en hallar el texel a cualquier orientación que pudiera tener la imagen de textura, así que la estimación del tamaño de ventana varía cuando se desconoce si los patrones poseen una tendencia o si simplemente la imagen ha sido rotada.

4.2.1. Valor de contraste

El SVH percibe la textura de acuerdo con seis características importantes según [Tamura et al., 1978], estos son: regularidad, contraste, direccionalidad, línea de semejanza, aspereza y rugosidad. A través de experimentos que relacionan mediciones psicológicas y computacionales tres características sobresalen del resto por la manera

de describir en forma amplia las texturas, las cuales son **aspereza, contraste y direccionalidad**, evaluadas en [Rao and Lohse, 1993] y utilizadas por [Abbadeni et al., 2000] con ayuda de la función de autocovarianza para mostrar la efectividad en experimentos psicológicos y tareas de adquisición.

- Aspereza, *-Coarseness-*. (Áspero - fino): Es una de las características más importantes de la textura. Es comúnmente asociada a la presencia de patrones a cierto nivel de escala. En este sentido una textura fina corresponde a pequeñas primitivas mientras que una textura áspera corresponde a grandes y pronunciadas primitivas.
- Contraste, *-Contrast-*. (Alto - bajo): Es una característica dependiente de los niveles de intensidad. Se presenta cuando dos o más regiones en la imagen difieren en su distribución de intensidad entonces dicha diferencia se conoce como contraste. Sin embargo, más factores influyen en la percepción de esta medida entre patrones de diferente estructura.
 - Rango dinámico de niveles de intensidad.
 - Polarización en la distribución para negros y blancos sobre el histograma o la relación entre áreas negras y blancas.
 - Forma de los contornos.
 - Periodo fundamental de patrones repetidos.
- Direccionalidad, *-Directionality-*. (Dirección - Sin dirección): Incluye elementos de cierta forma y reglas de posicionamiento en dos grupos, unidireccional y bidireccional, o simplemente puede ser visto como el grado total de direccionalidad.

Como se mencionó en la sección *Percepción visual*, los estadísticos de segundo orden operan de manera eficiente para la discriminación entre texturas de la misma manera que el SVH lo lleva a cabo. Obtener el modelo adecuado supone el uso de métodos estadísticos como GLCM o las métricas de SDH descritas en la sección *Características de segundo orden*.

Para considerar un nivel de semejanza entre pares de píxeles y con esto, hallar un valor de similitud entre estructuras sobre la imagen, las matrices de coocurrencia

son capaces de capturar en posiciones específicas sobre uno o más arreglos de GLCM, cantidades acumulativas que simulan probabilidades de ocurrencia entre elementos de la imagen.

Sea un pixel sobre la imagen de valor p_i en la posición (x_i, y_i) y otro de valor p_j sobre (x_j, y_j) son iguales si $p_i = p_j$ separados una distancia $d = (\Delta_x, \Delta_y)$ dentro de una vecindad, posiblemente del tamaño de la imagen $N \times M$, entonces el elemento de la matriz de coocurrencia $C(i, j|d, \theta)$ está descrita por los valores $C(p_i, p_j)$ en posición similar tal que $i = j$. De manera generalizada, una textura f es periódica si y sólo si $f(x, y) = f(x + v_x, y + v_y)$ dónde el vector de desplazamiento $v = (v_x, v_y)$ tiende al periodo del patrón elemental; la matriz de coocurrencia concentra los valores $C(i, j)$ en posiciones similares en el arreglo bidimensional, lo que dibuja una matriz diagonal.

La concentración de valores alrededor de los elementos de la diagonal principal en una matriz GLCM, indica el nivel de similitud entre dos vecindades separadas a cierta distancia, entonces es posible detectar si existe algún tipo de periodicidad cuando en un intervalo de distancias $[d_a, d_b]$ las matrices $C(i, j|\{d_a, \dots, d_b\})$ tiendan a un máximo global de concentración diagonal, dicha concentración también puede verse de forma recíproca a un parámetro de dispersión, llámese dispersión diagonal.

El valor de contraste es el resultado de aplicar la métrica de contraste en GLCM, Eq.(2.15) a un arreglo o conjunto de arreglos de co-ocurrencia y determinar la tendencia de concentrar ciertos elementos en una región de acuerdo a los niveles de intensidad G . El parámetro $(i - j)^2$ para $i, j \in [0, G - 1]$ de la métrica tiene una relación directa con la forma de una matriz diagonal de Toeplitz $W_n = w_{i-j}$, cuyos elementos obedecen la regla $w_{ij} = w_{|i-j|}$, Eq.(4.6).

$$W_n = \begin{pmatrix} w_0 & w_1 & w_2 & \dots & w_{G-1} \\ w_1 & w_0 & w_1 & \dots & w_{G-2} \\ w_2 & w_1 & w_0 & \dots & w_{G-3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ w_{G-1} & w_{G-2} & w_{G-3} & \dots & w_0 \end{pmatrix}, \quad (4.6)$$

la matriz W_n es una función de G parámetros $W_n = W_n(w_0, w_1, w_2, \dots, w_{G-1})$. La matriz de Toeplitz puede utilizarse como modelo que genere el valor de contraste y como

medida de dispersión diagonal para cada arreglo de GLCM. La Figura 4.9 muestra los pesos de contraste cuando se ocupa una etapa de cuantización a 8 niveles de gris.

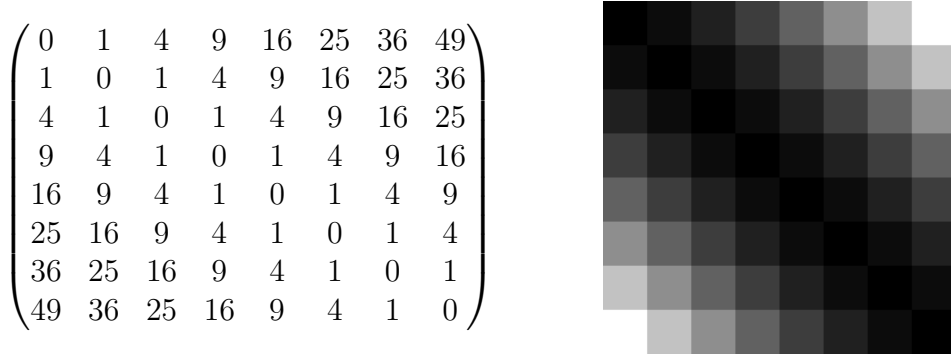


Figura 4.9: Toeplitz de 8 niveles de cuantización para la métrica de contraste.

Si una región Ω sobre la imagen es completamente uniforme, realizar el cálculo de $C(i, j|d, \theta)$ a cualquier distancia y orientación dibujaría una matriz cuyos elementos de mayor concentración se ubicarían a lo largo de la diagonal principal. Analizar sobre $\text{diag}(C)$ equivale a hallar patrones repetitivos cuyo comportamiento es completamente homogéneo. En tales casos, la Eq.(4.7) tiende a cero, $\text{diag}(W_n) \cdot \text{diag}(C) = 0$.

$$\text{Contraste} = \sum_i \sum_j W_n \cdot C(i, j) \quad (4.7)$$

Dado lo anterior, la estimación del texel involucra el valor de contraste para cada arreglo acumulativo de GLCM desde la distancia d_a hasta d_b con el fin de detectar el intervalo adecuado donde los elementos tienden a presentar el menor grado de dispersión con respecto a sus elementos centrales. El espacio de búsqueda donde es posible detectar una periodicidad sobre la imagen será a distancias $d = \{2, 3, 4, \dots, N/2\}$ dónde N es el número total de pixeles a analizar sobre una orientación dominante.

Es común que los pixeles pertenecientes a una sección cercana de la imagen presenten alta correlación, por lo que el nivel de intensidad en una vecindad pequeña puede ser similar, lo que implica un menor nivel de contraste, sin embargo mientras la distancia de análisis crece, la curva de contraste toma valores que dependen de la regularidad de las imágenes, dicho de otra forma, mientras la regularidad de una textura decrece, la curva de contraste perderá claridad, (ver Figura (4.10)). Por esta razón se requiere

un algoritmo capaz de detectar hasta el mínimo cambio de contraste como solución a un problema de optimización.

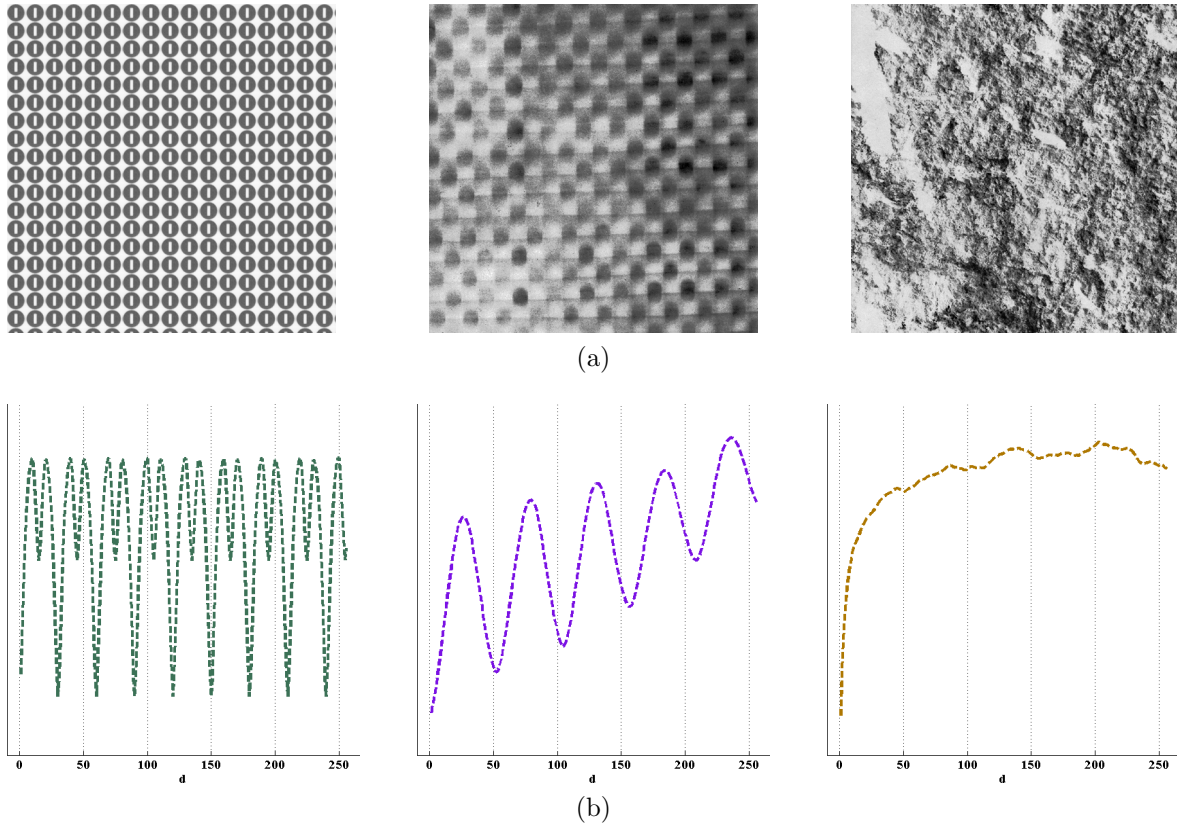


Figura 4.10: Curvas de contraste. Se muestran tres texturas en (a), periódica sintética, regular e irregular de dimensiones 512×512 píxeles. En (b) sus respectivas curvas de contraste.

4.2.2. Picos y picos prominentes

Algoritmos para la identificación de picos han sido desarrollados en diferentes áreas del procesamiento. Recientemente [Palshikar, 2009] brinda una descripción matemática de los picos y valles en las series de tiempo y propone métodos computacionalmente simples para la optimización de aplicaciones que requieran la detección de los mismos en tiempo real. Aplicaciones como en detección sobre ruido [Scholkmann et al., 2012], detección de dispositivos de RF [Rayanchu et al., 2011], estimación de la presión arterial [Chandrasekaran et al., 2013] o bien el monitoreo de signos vitales con ayuda de

téfonos inteligentes [Mitra and Poellabauer, 2012]; se basan en éste y otros criterios propuestos en la literatura.

Los picos denotan variaciones significativas de un evento representado por una serie de datos. Estos eventos pueden describirse por un incremento o decremento abrupto del comportamiento de una serie cuya forma se expresa como $\Gamma = \{(t_1, x_1), \dots, (t_N, x_N)\}$. Por lo que un pico P se describe de la siguiente manera.

$$P = \{(t_i, x_i) | (x_{i-1} < x_i > x_{i+1}) \vee (x_1 > x_2) \vee (x_n > x_{N-1})\}, \forall i = 2, \dots, N - 1, \quad (4.8)$$

donde N es el total de observaciones. Un pico se vuelve prominente P' cuando sobresale con respecto a los valores vecinos y posee una característica estadística particular. Una manera de establecer esta relación de pertenencia es mediante un criterio de decisión ζ , que puede depender de un conjunto local o global de los datos. Por lo que aquellos valores mayores o menores que ζ se convierten en buenos candidatos.

$$P' = \{P' \subset P | P_i(\zeta)\}, \forall i = 1, \dots, m, \quad (4.9)$$

donde m es la longitud de P . Establecer en ζ un valor de umbral debe considerar el comportamiento de la curva de datos, así como determinar si el pico pertenece a un máximo o mínimo local.

4.2.3. Identificación de picos basado en múltiples escalas

La metodología utilizada en este trabajo propone una modificación al algoritmo Automatic Multiscale-based Peak Detection (AMPD) de [Scholkmann et al., 2012], con el fin de obedecer un cierto criterio de decisión para la identificación de picos prominentes y brindar una solución adecuada a un modelo visto como un problema de optimización.

Dada una serie Γ discreta, los máximos y mínimos locales pueden determinarse utilizando una ventana deslizante β_k con un comportamiento de longitud variable simulando cambios de escala de tal forma que $\{\beta_k = 2k | k = 1, 2, \dots, L\}$ donde k representa

la escala de la señal y $L = (N/2) - 1$ la mayor escala posible. Para cada escala un valor $m_{k,i}$ es calculado en el cual se intenta resaltar exclusivamente los cambios locales mínimos cuando $i = k + 2, \dots, N - k + 1$:

$$m_{k,i} = \begin{cases} 0 & x_{i-1} < x_{i-k-1} \wedge x_{i-1} < x_{i+k-1} \\ r + \alpha & e.o.c \end{cases} \quad (4.10)$$

donde r es un número aleatorio con distribución uniforme de rango $[0, 1]$ y α en este caso un factor constante.

Estas operaciones dan como resultado una matriz \tilde{M} de elementos $m_{k,i}$ de dimensión $L \times N$, la cual es llamada escalograma. Cada fila de la matriz \tilde{M} indica el contenido de elementos asociados a sus mínimos ubicados en su respectiva escala. De tal manera que un vector $\gamma = \{\gamma_1, \gamma_2, \dots, \gamma_L\}$ representa el promedio de los elementos de escala similar, $\gamma_k = \sum_k m_{k,i}$. Finalmente, esto genera un valor que resalta la escala con contenido de mayor concentración de mínimos locales $\lambda = \arg \min(\gamma_k)$, donde la señal es analizada en el espacio adecuado del escalograma para la identificación de picos.

El parámetro λ reescala la matriz \tilde{M} removiendo aquellos valores de escala cuyo criterio es $k > \lambda$ produciendo una nuevo arreglo \tilde{M}_r de dimensión $\lambda \times N$. Cada columna i de la matriz \tilde{M}_r expresa una cantidad en la ubicación t_i de un valor x_i sobre la señal, de tal manera que la mínima variación para la i -ésima columna tiende a contener los mínimos locales.

$$\sigma_i = \sqrt{\frac{1}{\lambda} \sum_k \left(m_{k,i} - \frac{1}{\lambda} \sum_k m_{k,i} \right)^2}. \quad (4.11)$$

Aquellos índices i para el cual $\sigma_i \rightarrow 0$ son considerados candidatos como solución global al problema de optimización. Tales puntos están ordenados en el vector de picos P_{min} .

Los picos mínimos prominentes P'_{min} son el resultado de establecer un criterio propuesto como lo indica la Eq.(4.9) para el conjunto de valores obtenidos, así pues $P'_{min} = \{P'_{min} \subset P_{min} | P_{min}^{(j)}(\zeta)\}, \forall j = 1, \dots, m$, cuando j es el total de picos mínimos detectados. En este estudio se requiere proponer una única solución de las posibles. La

más adecuada se genera midiendo el valor de los picos en la señal original con respecto a un nivel de escala. Ya que la metodología provee de un análisis resultante de diversas escalas, se consideran directamente los valores de P_{min} para emitir una óptima solución, llámese T .

$$T = \arg \underset{\Gamma}{\min} (P_{min}) . \quad (4.12)$$

Finalmente, la curva de contraste, llámese Γ , posee una tendencia creciente que depende del nivel de regularidad de la textura, como en la Figura 4.10, este algoritmo es capaz de identificar aquellas variaciones notables de contraste y con ello estimar el mejor valor que esté directamente relacionado con el patrón repetitivo de textura, proponiendo así un tamaño de ventana texel cuya forma es rectangular de dimensiones $T_x \times T_y$, puesto que se analiza sobre dos ejes mutuamente ortogonales. Se detalla el procedimiento en el Algoritmo 1.

Algorithm 1 Estimación del texel mediante identificación de picos prominentes

```

1: Input: Textura orientada
2: Output: Texel
3: procedure PICOS
4:    $\Gamma \leftarrow$  Curva de contraste;
5:    $M \leftarrow$  Escalograma;
6:    $\lambda \leftarrow$  Escala apropiada de análisis;
7:    $M_r \leftarrow$  Escalograma reescalado;
8:    $P_{min} \leftarrow$  Identificación de picos mínimos;
9:    $P'_{min} \leftarrow$  Picos mínimos prominentes mediante criterio;
10:   $T \leftarrow$  Solución optima;
11: end procedure

```

4.2.4. Base experimental 2. Ventanas óptimas

Experimento 3. Consideremos una textura sintética cuyos patrones se repitan exactamente a determinada distancia sobre dos direcciones en la imagen. La forma más simple de representar tal evento es simulando un tablero de ajedrez, ver Figura 4.11a, con bloques blancos y negros intercalados, cada uno de dimensiones $k \times k$ pixeles. El análisis se realiza sobre dos ejes y de dos formas, cuando se omite (1) y cuando se

incluye la etapa de detección de orientación (DO) (2).

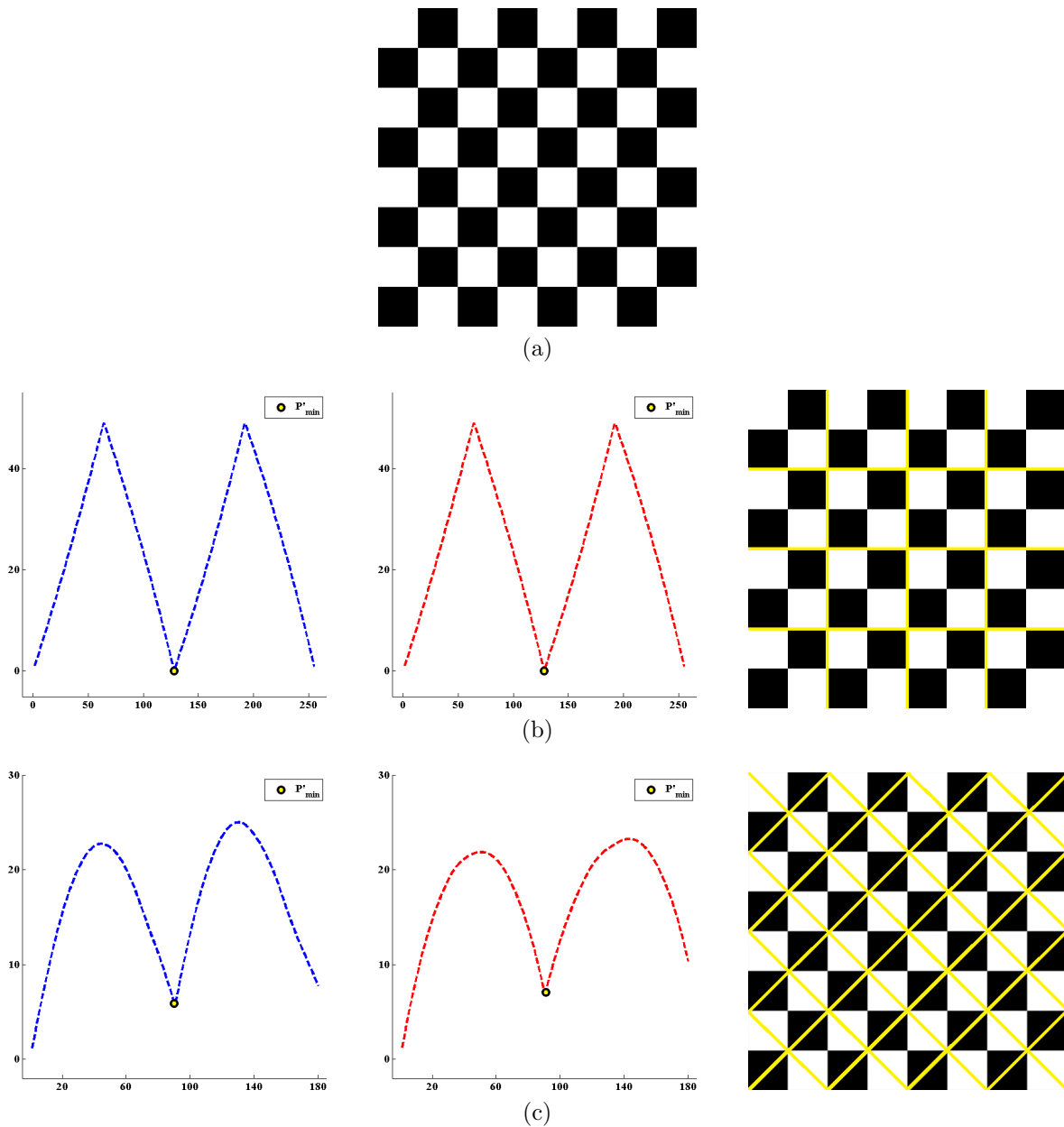


Figura 4.11: Estimación de texel en tablero. Tablero de ajedrez de 512×512 píxeles en (a). (b) Curvas de contraste para la estimación de texel sin etapa previa de DO, arroja un texel de 128×128 píxeles. (c) Estimación de texel considerando etapa previa de DO con un texel resultante de 90×90 píxeles.

- (1) De manera intuitiva se puede notar que el patrón repetitivo es una vecindad que envuelve dos cuadros blancos y dos cuadros negros, repetidos a distancia $2k, 4k$

y $6k$ sobre los ejes x e y . Esta suposición es demostrada por el algoritmo de estimación de texel de la sección anterior, Figura 4.11b. El texel estimado es de tamaño 128×128 pixeles, como se percibe visualmente.

- (2) Por otro lado, se muestra el resultado de estimar el texel en conjunto con la detección de orientación, Figura 4.11c. El algoritmo en este caso arroja un resultado de orientación de 45° , lo que significa que se detecta una alineación aún más elemental de los cuadros blancos y negros. En consecuencia se pueden hallar patrones más básicos sobre esa dirección. Finalmente se debe considerar la posibilidad de que exista un texel de menor tamaño que cuando se omite una etapa previa para detectar la orientación, en tal caso el texel es de dimensiones 90×90 pixeles.

Experimento 4. Esta prueba se basa en encontrar el texel invariante a rotación para imágenes envueltas en varios niveles de degradación. Esto con el fin de medir que tan robusta es la estimación de la ventana óptima en presencia de alteraciones comunes. El tablero de ajedrez, Figura 4.11a, se hace rotar en $\theta^r = 0, 10, 20, \dots, 170$, similar a lo presentado en el *Experimento 2*. La estimación del texel se lleva a cabo considerando una etapa previa de DO presentando dos casos de alteración.

- Para cada imagen rotada, se simula el ruido Gaussiano de parámetros $\mu = 0$ y $\sigma^2 \in [0, 2]$, Figura 4.12a.
- Para cada imagen rotada, ruido impulsivo de parámetro $Oc \in [0, 1]$, dónde $Oc = 1$ equivale a un nivel de ruido del 100 % de ocupación sobre la imagen. Figura 4.12b

Las gráficas del error MAE se muestran en la Figura 4.12c para ruido Gaussiano e Impulsivo respectivamente. Dado que el tamaño del texel involucra dos valores que dan forma a una ventana rectangular, la estimación considera dos ejes, llámense primario y secundario los cuales son ortogonales. El error se muestra para cada eje.

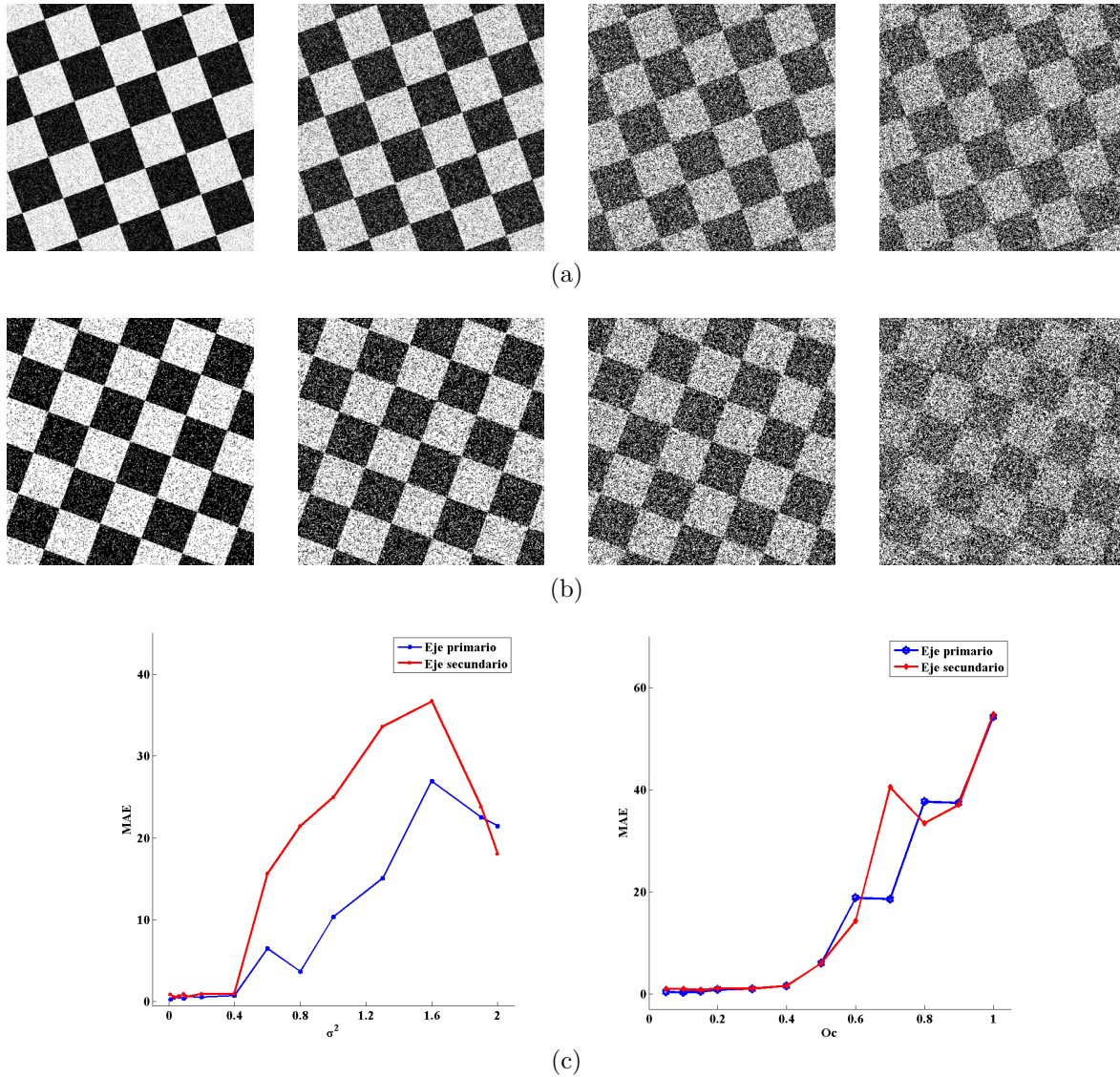


Figura 4.12: Estimación de Texel para el tablero con ruido. En (a) se presenta el tablero rotado a $\theta^r = 20$ en sentido opuesto al avance de las manecillas del reloj, con ruido Gaussiano $\sigma^2 = \{0.06, 0.2, 1, 1.9\}$ respectivamente. Sobre (b) existe ruido impulsivo de niveles $O_c = \{20\%, 40\%, 60\%, 80\%\}$ para la imagen de rotación a $\theta^r = 70$. Gráficas del error MAE para la estimación sobre ruido Gaussiano e Impulsivo (c).

Capítulo 5

Clasificación de estructuras

En este capítulo se presentan los resultados de clasificación de texturas. Dichas pruebas se realizaron con un conjunto seleccionado de texturas de la base de datos [Brodatz, 1966] y [Klette, 2010], ambas colecciones de imágenes representan escenas naturales. La metodología de extracción y clasificación presentada en este trabajo se compara con resultados de metodologías habituales de clasificación las cuales emplean técnicas de ventaneo de tamaño fijo. Ésto con el fin de demostrar el resultado de analizar texturas con ventanas de tamaño texel cuya información espacial es suficiente en comparación a ventanas amplias de abundante información que en ocasiones se vuelven innecesarias.

5.1. Extracción de características

La metodología presentada en el Capítulo 4, brinda la descripción para extraer el texel invariante a rotación de determinada imagen de textura. El estudio presentado hasta ahora es utilizado como etapa de pre procesamiento. Consecuentemente se efectúa la clasificación de las estructuras obtenidas mediante técnicas de ventaneo, véase Figura 2.2. Durante la extracción de características, es útil encontrar aquellas medidas que ayuden a definir una representación particular de las imágenes, esto es, colocar los valores del vector de características lo más separado posibles en el espacio de variables con el fin de identificar fácilmente cada clase. Este trabajo realiza un análisis en mul-

tiescala utilizando los coeficientes de la transformada Hermite. Dos metodologías de extracción de características son evaluadas a continuación.

5.1.1. Vector de características con la Transformada Hermite cartesiana en multiescala

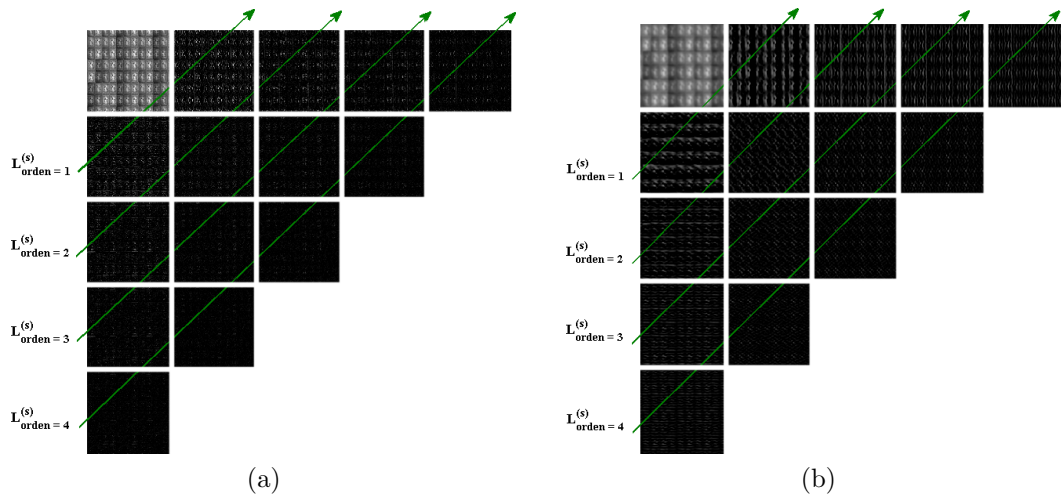


Figura 5.1: HT-multiescala aplicada a la imagen Kufu09e del álbum [Klette, 2010]. Coeficientes en (a) para $n = 4, s = 1$, los coeficientes en (b) son resultado de aplicar $n = 4, s = 2$. Todos los coeficientes de las escalas son procesados para generar V_{HT} .

Los coeficientes de la Transformada Hermite cartesiana (HT) representan una serie de filtros dirigidos hacia los dos ejes principales, vertical y horizontal. Dicha propiedad es útil para resaltar elementos en tales direcciones.

1. Se define un orden de expansión $n = \{0, 1, \dots, N\}$ y un determinado número de escalas $s = \{s_1, s_2, \dots, S\}$. Se sigue el esquema piramidal de descomposición mostrado en la Figura 3.4. El máximo orden de la derivada N estará limitado por el soporte del filtro F de tal manera que se establece una relación de $N \leq F$, siendo la igualdad el primer valor de escala s_1 .
2. Se aplica HT-multiescala para cada ventana, con lo que se obtienen versiones filtradas para cada una de las escalas. La Figura 5.1 muestra una imagen procesada a dos escalas distintas.

3. La extracción de estadísticos media μ y desviación estándar σ para los coeficientes de HT-multiescala obtenidos en $L_{n-m,m}^{(s)}$.
4. Construcción del vector de características mediante la concatenación de los valores estadísticos de $L_{orden}^{(s)}$

$$V_{HT} = \{\mu_{1,0}^{s_1}, \sigma_{1,0}^{s_1}, \mu_{0,1}^{s_1}, \sigma_{0,1}^{s_1}, \dots, \mu_{p,t}^{s_q}, \sigma_{p,t}^{s_q}, \dots, \mu_{0,N}^S, \sigma_{0,N}^S\}, \quad (5.1)$$

donde s_q representa un valor en s y p, t valores en n . La longitud del vector V_{HT} es igual a: $length(V_{HT}) = N(N + 3)S$.

5.1.2. Vector de características con la Transformada Hermite rotada en multiescala

Para el uso de SHT-multiescala, el conjunto de funciones de Hermite es proyectado sobre las ventanas de análisis formando coeficientes con respecto al orden de expansión y una escala establecida. La concentración máxima de energía se localiza en los primeros ordenes de análisis $L_{n,0}^\theta$. La metodología de construcción del vector de características es la siguiente:

1. Se define un orden de expansión $n = \{0, 1, \dots, N\}$ y un determinado número de escalas $s = \{s_1, s_2, \dots, S\}$. Se sigue el esquema piramidal de descomposición mostrado en la Figura 3.4. De la misma forma que en la extracción de características con HT-multiescala, el máximo orden de la derivada N estará limitado por el soporte del filtro F de tal manera que se mantiene la misma relación $N \leq F$, siendo la igualdad el primer valor de escala s_1 .
2. Se aplica la SHT-multiescala para cada ventana o muestra de textura, con lo que se obtienen versiones filtradas y orientadas para cada escala. La Figura 5.2 muestra una imagen procesada a dos escalas distintas y la manera en que se consideran los coeficientes de primeros ordenes.
3. La extracción de estadísticos media μ y desviación estándar σ para cada coeficiente de SHT-multiescala en la representación $L_{n,0}^{(\theta,s)}$.

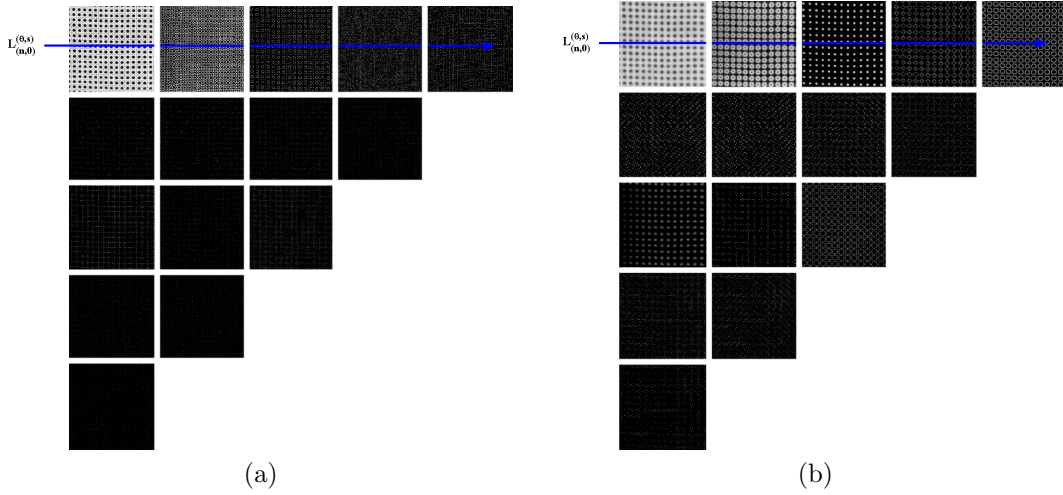


Figura 5.2: SHT-multiescala aplicada a la imagen D101 del álbum [Brodatz, 1966]. Coeficientes en (a) para $n = 4, s = 1$, los coeficientes en (b) son resultado de aplicar $n = 4, s = 2$. La concentración de energía sobre los primeros ordenes de análisis disminuye la longitud del vector V_{SHT} sustancialmente.

4. Se construye del vector de características mediante la concatenación de los valores

$$V_{SHT} = \{\mu_1^{s_1}, \sigma_1^{s_1}, \mu_2^{s_1}, \sigma_2^{s_1}, \dots, \mu_p^{s_q}, \sigma_p^{s_q}, \dots, \mu_N^S, \sigma_N^S\}, \quad (5.2)$$

donde s_q representa un valor en s y p un valor en n . La longitud del vector V_{SHT} es igual a: $length(V_{SHT}) = 2NS$.

5.2. Proceso de clasificación

Se generó un conjunto de 23 imágenes para el proceso de clasificación, cada imagen de tamaño 640×640 representa una clase distinta. Para la base de datos de [Brodatz, 1966] se extrajeron algunas texturas que visualmente presentaban cierto grado de regularidad: *D1, D101, D102, D18, D20, D34, D46, D47, D52, D56, D64, D65, D66, D8*. A su vez, la colección de [Klette, 2010] proporcionó las siguientes texturas naturales: *101Taipei05, BurgHohenzollern06, EierDeko06, KentingNightMarkt05b, KentingNightMarkt05c, Knoxville06, YellowCraneWuhan06, Kufu09e, Minneapolis06a*. Cada textura se hizo rotar desde 0° a 170° formando imágenes de tamaño 453×453 que evitan

el borde oscuro al rotarlas y obtener un conjunto de tamaño consistente. Los vectores generados se someten a una clasificación del vecino más cercano (NN) utilizando en los experimentos, a las distancias *Euclidean* y *CityBlock* como parámetros. Se siguieron dos procesos diferentes de validación.

5.2.1. Holdout validation

Un proceso de rotación cada 10° genera 18 imágenes rotadas por clase. Se seleccionaron intercaladamente 9 imágenes para tareas de entrenamiento y 9 imágenes para prueba. Se realizó posteriormente un análisis de la textura a través de ventanas que genera parches de textura. El utilizar ventanas de tamaño fijo, por ejemplo de 128×128 con 0% de traslape produce un número total de parches de $9 \times 9 \times 23 = 1863$ para entrenamiento y otros 1863 para clasificar. Si el tamaño de ventana es de 64×64 sin considerar alguna cantidad de traslape, esto genera $49 \times 9 \times 23 = 10143$ parches para entrenamiento y 10143 para clasificación. El número de parches cuando se analiza bajo la metodología de texel varia en proporción al nivel de regularidad y periodicidad que la textura posee. Cada texel tiende a capturar la estructura fundamental de la imagen.

El objetivo de esta validación es generar un vector de características mediante el análisis de Hermite considerando cuatro ordenes de expansión en 4 escalas. Por cada parche de textura se decide su clase de procedencia. La clase se vincula manualmente para aquellos vectores que sean identificados como entrenamiento. Con el fin de evitar posible sesgo entre parches de distintas versiones rotadas, se eligen aleatoriamente el 50% del total de parches para entrenar y clasificar, véase Figura 5.3a. Se repite el experimento de clasificación 10 veces obteniendo como resultado un promedio de clasificación.

5.2.2. Random-fold cross-validation (RFCV)

Considerar una rotación de textura cada 30° produce 6 imágenes por cada clase. Se sabe que el número de parches a clasificar depende del tamaño de ventana con el que se extrae cierto contenido de textura y el nivel de traslape entre ventanas. Para ventanas fijas sin traslape de 128×128 el número total de parches posibles a generar

de la base de datos es de $9 \times 6 \times 23 = 1242$, mientras que utilizando ventanas de 64×64 forma un conjunto de $49 \times 6 \times 23 = 6762$. Finalmente, la metodología texel construye un conjunto variable dependiente de las dimensiones de la ventana.

Esta forma de validación divide el total de elementos en K subconjuntos cuyos elementos son elegidos de forma aleatoria, con el objetivo de evaluar aleatoriamente cada subconjunto con tareas de entrenamiento y prueba, véase Figura 5.3b. Así, una validación aleatoria cruzada de 10 iteraciones separa el total de datos en 10 grupos por clase donde uno de ellos hace la función de prueba. Con el fin de eliminar posible sesgo entre los parches extraídos de las versiones rotadas, se eligió aleatoriamente el 50% del total de parches para ser clasificados bajo esta metodología. Los vectores de características son construidos siguiendo la descripción de *Extracción de características* con parámetros fijos de cuarto orden de expansión en 4 escalas.

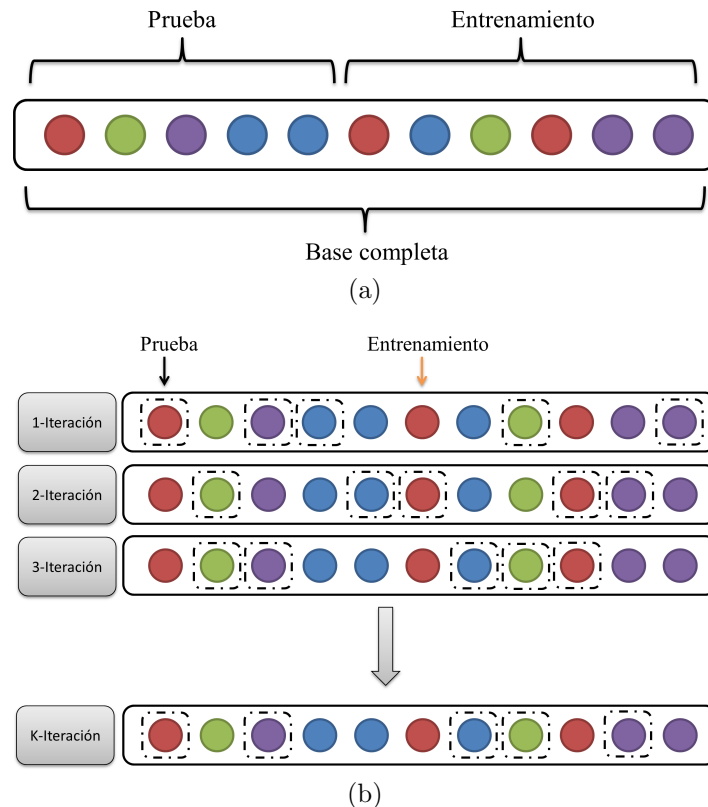


Figura 5.3: Validación Holdout en (a), los elementos de entrenamiento y prueba son separados en grupos determinados. La validación RFCV (b) elige aleatoriamente el conjunto de datos para entrenamiento y prueba, se efectúan K -iteraciones.

5.3. Evaluación de la predicción en clasificación

Una matriz de confusión tiene por objetivo evaluar el desempeño del sistema de clasificación. Esto significa medir el nivel de exactitud o imprecisión del sistema para encontrar una concordancia entre las clases procesadas y las clases estimadas según los datos recolectados por las tareas de entrenamiento. Los métodos supervisados de aprendizaje nos permiten acceder a las etiquetas de los datos durante las etapas de entrenamiento y prueba. Cuando las entradas de los datos x_1, \dots, x_k deben ser asignados a las clases correspondientes C_1, \dots, C_k , entonces la clasificación se vuelve multiclase.

Las matrices de confusión en la Figura 5.4 y Figura 5.5 muestran los valores como un promedio de predicción (AP) por clase para los parches clasificados, dada la relación que existe entre el número de parches clasificados $N_c^{(k)}$ entre el número total de parches por clase $N_T^{(k)}$. (Para mayor detalle de los valores véase el Apéndice A)

$$AP = \frac{N_c^{(k)}}{N_T^{(k)}}. \quad (5.3)$$

En el primer caso, las matrices de la Figura 5.4 indican el porcentaje del nivel de predicción cuando la metodología de extracción HT-multiescala es aplicado. Cada predicción está representada por un porcentaje de AP en escala de 0 a 100. Un cuadro completamente negro es proporcional a un valor de 100 %, lo cual indica que todos los parches de dicha clase han sido clasificados correctamente. En el caso ideal, la matriz de confusión mostraría solamente una diagonal de cuadros negros. Se observa sobre la primera fila de matrices (Figuras 5.4a, (b) y (c)) el rendimiento de clasificación con validación *Holdout*, mientras que en la segunda fila (Figuras 5.4d, (e), y (f)) expone el resultado de la validación RFCV.

Por otro lado, los valores de AP considerando la extracción de características mediante SHT-multiescala se muestra en la Figura 5.5. La primera fila de matrices (Figuras 5.5a, (b) y (c)) se obtuvieron al aplicar la validación *Holdout*. La segunda fila de matrices (Figuras 5.5d, (e) y (f)) son resultado de validar a través de la metodología RFCV. De la misma forma, un tono de intensidad es asociado al nivel AP siendo el color negro un 100 % y blanco un valor de 0. Una manera subjetiva de evaluar el promedio de predicción correcto en forma global es percibir el nivel de claridad de la matriz de

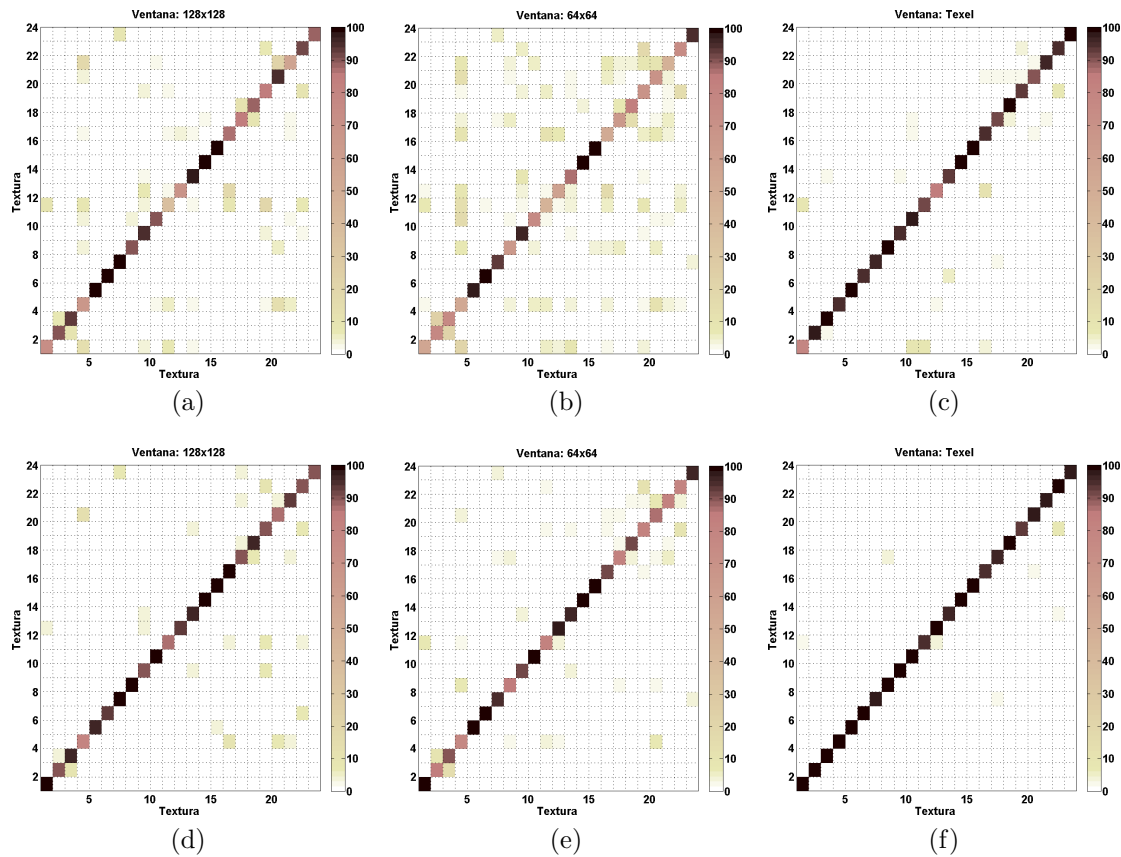


Figura 5.4: Matrices de Confusión de AP para HT-multiescala. En (a), (b) y (c) se muestra el valor AP por clase de ventanas fijas de 128×128 , fijas de 64×64 y ventanas texel para validación *Holdout*. En (d), (e) y (f) se muestra las matrices del proceso *RFCV* para ventanas fijas de 128×128 , ventanas fijas de 64×64 y ventanas texel, respectivamente.

confusión en los valores adyacentes a su diagonal. Mientras más alta sea la saturación de intensidad de los bloques diagonales, mayor es el nivel de predicción correcta del sistema.

5.4. Medidas de rendimiento

Es necesario representar de manera numérica los niveles de clasificación correcta y precisión para brindar un estimado de efectividad y realizar una comparación directa entre las técnicas. Los niveles de clasificación correcta son medibles a través del

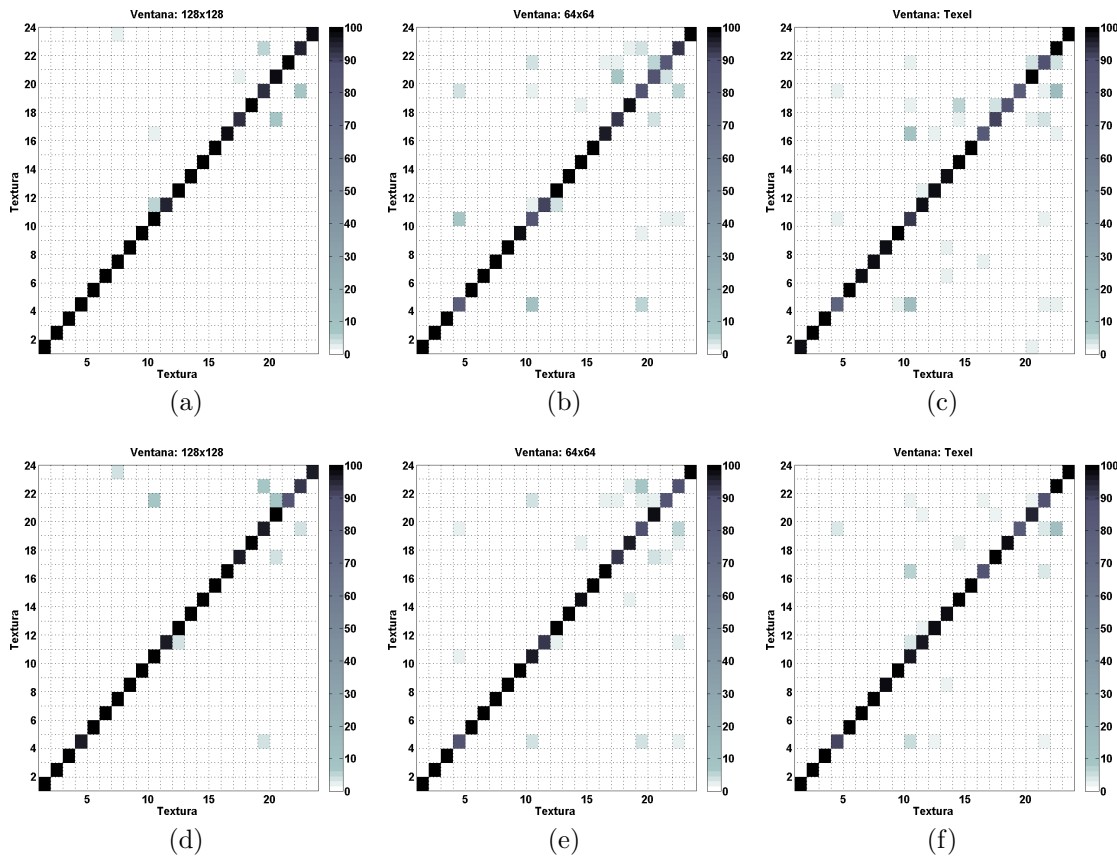


Figura 5.5: Matrices de Confusión de AP para SHT-multiescala. En (a), (b) y (c) se muestra el valor AP por clase de ventanas fijas de 128×128 , fijas de 64×64 y ventanas texel para validación *Holdout*. En (d), (e) y (f) se muestra las matrices del proceso *RFCV* para ventanas fijas de 128×128 , fijas de 64×64 y ventanas texel, respectivamente.

cálculo de la tasa de clasificación correcta (CCR), como una relación entre el número de texturas totales que han sido correctamente clasificadas y el número total de texturas evaluadas. La tasa CCR es útil cuando varias clases tiene un número distinto de elementos a clasificar, por lo que es un valor que considera la totalidad de los parches evaluados. Si el número de elementos a clasificar es el mismo para todas las clases, el valor de CCR tiende a un valor promedio de sensibilidad (*Sensibility*), esto es $CCR \approx Sensibility$.

La sensibilidad es un parámetro que expresa la cantidad de elementos que han sido clasificados de manera correcta con respecto a su clase. El valor de especificidad (*Specificity*) es una medida de efectividad para la identificación de clases negativas. La

Cuadro 5.1: Medidas HT-multiescala

	Ventana	CCR	Sensitivity	Specificity	Precision
Holdout	128x128	85.80	85.81	99.36	86.26
	64x64	75.48	75.26	98.88	75.91
	Texel	95.71	94.58	99.75	94.74
RFCV	128x128	93.91	93.91	99.72	94.14
	64x64	90.87	90.87	99.58	90.84
	Texel	98.79	98.64	99.92	98.40

precisión (*Precision*) indica el número de elementos que han sido clasificados correctamente con respecto a los valores de predicción de una clase. Las medidas de rendimiento obtenidas de las matrices de confusión usualmente expresan valores de clasificación en forma binaria para una clase en particular. Los Cuadros 5.1 y 5.2 expresan el valor promedio de las medidas individuales de las clases.

En el Cuadro 5.1 se muestran las medidas de rendimiento para las validaciones hechas por la herramienta HT-multiescala. La metodología que utiliza ventanas texel exhibe un mejor desempeño a comparación de ventanas de longitud fija, superando la ventana más grande de 128×128 píxeles hasta en un 9% de CCR, precisión y sensibilidad. Esto se debe a que la etapa de estimación de texel invariante a rotación logra ubicar el contenido de textura sobre una dirección significativa. Por consiguiente, la extracción de un vector de características también es invariante al considerar que las ventanas están alineadas. Esto es útil cuando se utiliza una herramienta anisotrópica para la construcción de una firma de textura que depende de la orientación.

El Cuadro 5.2 señala los porcentajes de CCR por metodología y ventana de análisis aplicando SHT-multiescala en la construcción de características. La clasificación *NN* validada por *Holdout* muestra gran porcentaje de clasificación cuando ventanas relativamente grandes de 128×128 píxeles son utilizadas, seguido de ventanas texel sólo 1% por debajo en escala CCR. Por otro lado, la validación RFCV indica que se obtiene una tasa mayor de clasificación correcta utilizando ventanas texel que ventanas de otros tamaños, aunque las otras medidas de desempeño son menores. Estas medidas pueden ser variables debido al distinto número de elementos a clasificar por la metodología presentada.

Se sabe que un valor de CCR cambia de acuerdo al número de elementos utilizados

Cuadro 5.2: Medidas SHT-multiescala

	Ventana	CCR	Sensitivity	Specificity	Precision
Holdout	128x128	98.62	98.62	99.94	98.66
	64x64	95.43	95.44	99.79	95.48
	Texel	97.43	96.24	99.77	95.24
RFCV	128x128	98.41	98.41	99.93	98.50
	64x64	95.55	95.54	99.88	96.58
	Texel	98.63	96.64	99.85	96.81

en el entrenamiento y prueba. Entre menor sea el número de vectores asignados para entrenamiento el coeficiente de clasificación correcta decrece al aumentar el nivel de incertidumbre debido al establecer pocos parches o pocas imágenes que definan clases completas.

Las gráficas de la Figura 5.6 muestran ambas herramientas de análisis para los dos métodos de validación cuando se varía el porcentaje de datos de entrenamiento. En la validación Holdout, el número de vectores de prueba se mantiene constante, sin embargo los datos de entrenamiento se hacen variar desde 10% hasta 70% del total disponible. Por otra parte, en la validación RFCV, al disminuir el número de datos de entrenamiento, aumentan los datos de prueba.

Las Figuras 5.6a y 5.6b utilizan la versión HT-multiescala para la validación *Holdout* (a) y RFCV (b). En este caso, la ventana texel presenta el mayor nivel de CCR aún cuando el porcentaje de entrenamiento disminuye. Este resultado es comparable con las otras ventanas de análisis de tamaño heurístico que no son dependientes de la orientación por lo que el nivel de CCR es menor hasta en un 20%.

La versión rotada SHT-multiescala es útil para describir texturas invariante a rotación por la propiedad de direccionalidad de sus filtros. Esto se puede percibir en las gráficas para validación *Holdout* en la Figura 5.6c y RFCV en la Figura 5.6d. La cercanía de las curvas de CCR indica que la generación de un descriptor es eficiente y no es completamente dependiente de la ventana de análisis. Aunque con ventanas texel los resultados tienden a ser mejores y consistentes cuando no se dispone del total de elementos que entrenen al clasificador. Una comparación directa entre las ventanas texel y las ventanas de 128×128 señalan una ligera variación entre 3% de CCR, considerando que todas las ventanas texel son de menor tamaño (Véase Figura 5.7).

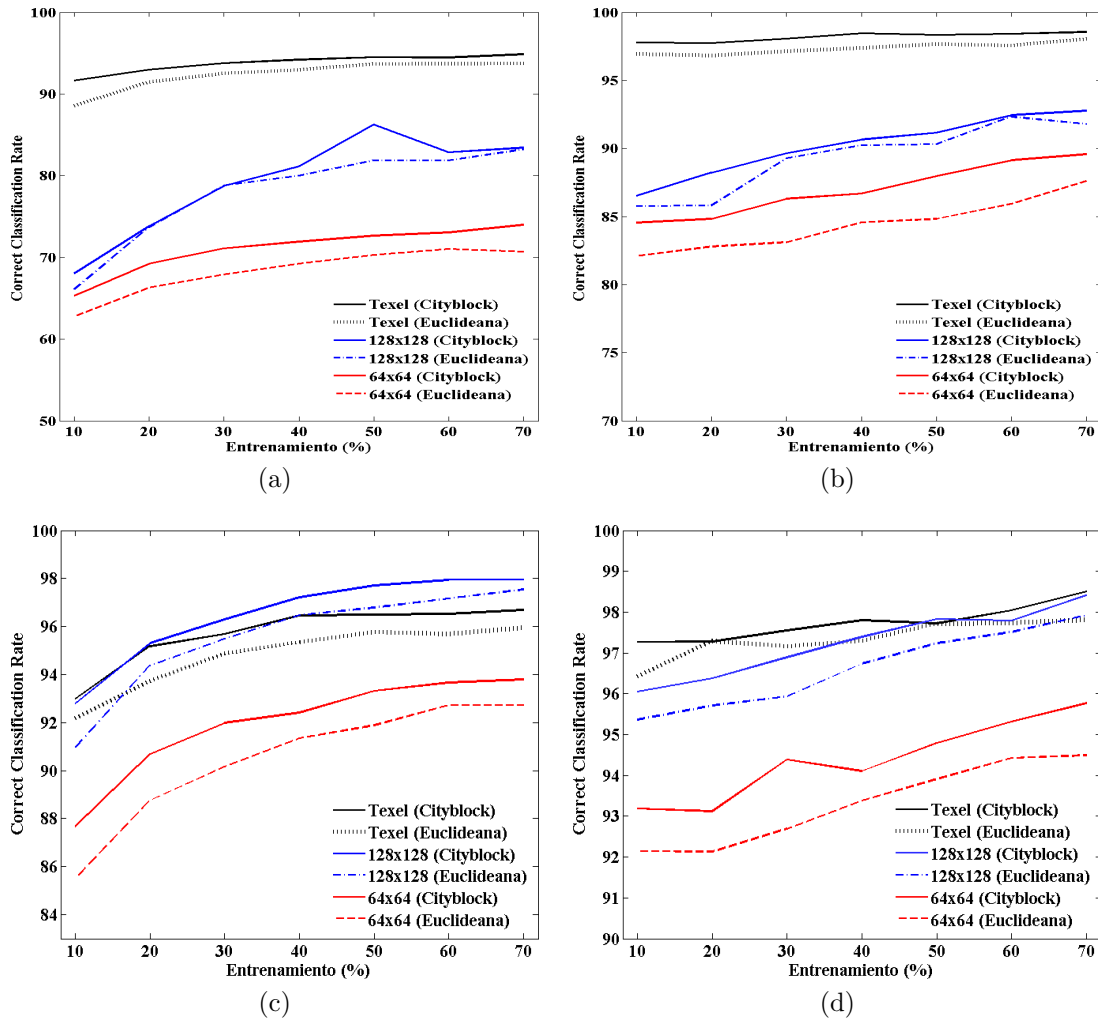


Figura 5.6: Niveles de clasificación CCR. Se presentan los resultados de la herramienta HT-multiescala para la validación Holdout en (a) y RFCV en (b). La herramienta SHT-multiescala arroja resultados de Holdout en (c) y RFCV en (d). Las gráficas muestran los niveles CCR por ventana de análisis y parámetro de distancia para cierto porcentaje de vectores de entrenamiento.

Considerar que el tamaño promedio del texel por imagen en estos experimentos contiene un número menor de pixeles comparado con ventanas fijas de 128×128 las cuales analizan vecindades de 16384 pixeles. Por otro lado, la mayoría de los texel se encuentran por debajo del umbral que definen las ventanas de 64×64 las cuales analizan vecindades en cantidades de 4096 pixeles véase Figura 5.7.

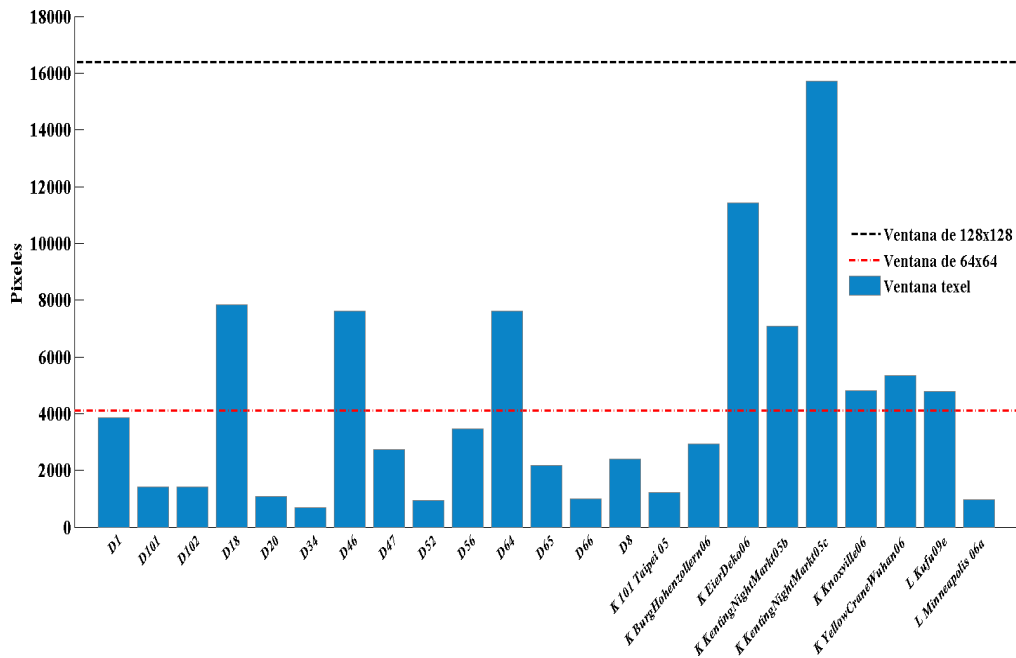


Figura 5.7: Promedio de pixeles para cada clase ocupando ventanas texel y ventanas de tamaño fijo.

En las imágenes de textura de la Figura 5.8 se muestra una parrilla de ventanas de tamaño fijo y ventanas texel. Las ventanas de tamaño fijo actúan de manera independiente a alguna propiedad o parámetro de textura, por lo que el número de patrones capturados puede variar y sus valores estadísticos tienden a ser diferentes. Las ventanas texel ayudan a identificar los elementos repetitivos y con ello producir un análisis generalizado que represente clases completas con la menor cantidad de información.

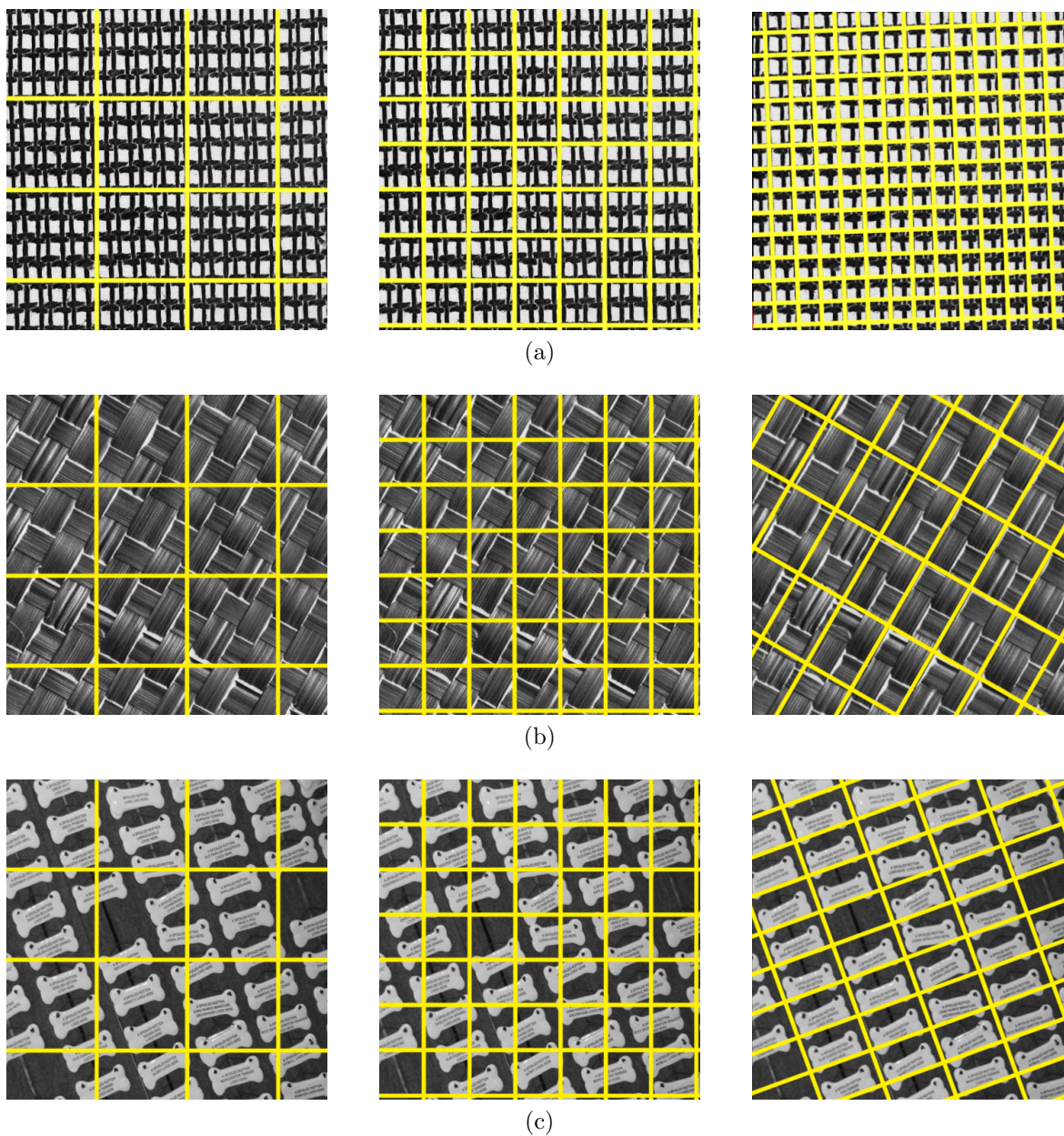


Figura 5.8: Ventanas de análisis para las imágenes D20 y D64 en (a) y (b) del álbum [Brodatz, 1966] y Knoxville del álbum [Klette, 2010] en (c). Los tamaños de ventana por columna de imagen son de 128×128 , 64×64 y texel respectivamente.

Capítulo 6

Conclusiones

La metodología presentada en este trabajo parte considerando que existe un gran conjunto de texturas regulares que pueden ser representadas de manera óptima en el sentido de reducir el número de cálculos, evitar inestabilidad por el uso de polinomios de orden superior así como evitar analizar las imágenes completas para generar una firma que identifique sus propiedades, esto con el fin de distinguir clases completas con gran exactitud y con la menor cantidad de información requerida.

Tres de las propiedades de textura han mostrado ser de mayor utilidad para la percepción y comparación entre imágenes. Aspereza, contraste y direccionalidad, las cuales fueron analizadas con herramientas distintas para generar una base consistente que tienda a una clasificación mejorada. Los niveles de contraste calculados como estadísticos de segundo orden permitieron estimar el nivel de similitud entre vecindades a cierta distancia y a su vez plantear el análisis como solución a un problema de optimización. Considerar la detección de una dirección dominante mediante la transformada de Radón hace a la herramienta invariante a rotación. Los experimentos llevados a cabo para la estimación de orientación y tamaño de texel mostraron buena precisión incluso si las imágenes presentan un nivel medio de ruido tanto impulsivo como Gaussiano.

El uso del análisis de Hermite en multiescala asemeja el funcionamiento del sistema de visión humano, cambios de escala simulan cambios de perspectivas para ubicar cierto contenido de aspereza. La versión cartesiana de la transformada Hermite provee una firma de textura que no considera un nivel de direccionalidad para producir un

descriptor, por lo tanto la etapa de preprocesamiento propuesta para estimar la orientación es de gran ayuda al demostrar resultados más consistentes con respecto a la tasa clasificación correcta. Por otro lado, la versión rotada de Hermite permite ubicar la mayor concentración de energía hacia una dirección dominante. Esto produce que tal herramienta se considere invariante al generar un descriptor de texturas isotrópicas o anisotrópicas. Lo anterior se percibe en los niveles altos de clasificación correcta.

Es común encontrar en la literatura de procesamiento de imágenes métodos heurísticos con ventanas de análisis que capturen la mayor cantidad de elementos representativos para una posterior clasificación. Si se considera una textura cuyo comportamiento obedece reglas oscilatorias, algunas vecindades de pixeles obtendrán un alto nivel de correlación ya que se muestra un patrón repetitivo a lo largo del espacio de la imagen. Capturar dicho patrón a través de una ventana texel de análisis evita redundancia de información comparado con ventanas de tamaño fijo de dimensión arbitraria, en tales casos se tornan innecesarias. Lo anterior se demostró mediante tareas de clasificación, donde el objetivo fue obtener la mayor cantidad de imágenes correctamente clasificadas. El índice de clasificación correcta para la metodología presentada es consistente aún cuando disminuye el número de valores de entrenamiento. El tamaño de los texeles tiende a ser ideal y por lo tanto un número reducido de pixeles es analizado.

Finalmente, esta metodología sienta las bases para un estudio posterior sobre el uso de texeles, el cual que puede ser aplicado como un modelo estadístico para solucionar problemas comunes de clasificación o segmentación de imágenes. Se podrían mencionar algunas áreas como Imagenología médica, Percepción remota o Inspección industrial donde las imágenes con cierto contenido de regularidad pueden ser sometidas a este proceso con el fin de evaluar cada una de sus estructuras y definir aquellas mayor de relevancia para tareas posteriores de caracterización.

Apéndice A

Matrices de confusión de valores AP

Una matriz de confusión es una herramienta que permite visualizar el desempeño de los sistemas de extracción, descripción y clasificación de procesos supervisados [Kohavi and Provost, 1998]. Los valores en la matriz refieren a las clases evaluadas y pronosticadas con el fin de identificar el nivel de acierto del clasificador. La exactitud en la clasificación se evalúa mediante el cálculo de ciertos parámetros que expresan el nivel de acierto para reconocer los positivos (*True Positives*) y negativos (*True Negatives*), así como el fallo de reconocer positivos como negativos (*False Negatives*) y negativos que eran positivos (*False Positives*). Estos cuatro valores constituyen una matriz de confusión como se muestra a continuación para clases binarias, Cuadro A.1.

Cuadro A.1: Tabla de confusión binaria

Clase actual	Clasificado como positivo	Clasificado como negativo
Positivo	Verdadero Positivo (TP)	Falso Negativo (FN)
Negativo	Falso Positivo (FP)	Verdadero Negativo (TN)

Cuando se estudia el caso de multiclase, cada valor de una clase actual x_1, \dots, x_k debe asignarse a una clase de predicción y_1, \dots, y_k donde k representa el total de clases. Es posible generar una tabla de confusión binaria para cada clase definida por sus

medidas individuales TP_i, TN_i, FP_i, FN_i . Estos valores individuales sirven de apoyo para generar las llamadas medidas de rendimiento, las cuales expresan niveles de acierto y fallo del clasificador. Dado un desarrollo multiclase, las medidas de rendimiento del sistema completo se representan mediante un promedio estadístico de las medidas de rendimiento individuales, [Sokolova and Lapalme, 2009]. Algunas de las más importantes se presentan a continuación:

- Sensitivity (Recall): Mide la efectividad del clasificador para identificar etiquetas positivas.

$$Sensitivity = \frac{1}{k} \sum_i^k \frac{TP_i}{TP_i + FN_i}.$$

- Specificity: Que tan efectivo el clasificador identifica las etiquetas negativas.

$$Specificity = \frac{1}{k} \sum_i^k \frac{TN_i}{FP_i + TN_i}.$$

- Precision: Indica la proporción de los elementos positivos relevantes en las etiquetas clasificadas de positivo.

$$Precision = \frac{1}{k} \sum_i^k \frac{TP_i}{TP_i + FP_i}.$$

Ventana: 128x128

D1	72.43	0.00	0.00	13.17	0.00	0.00	0.00	4.63	0.00	6.59	0.49	1.95	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.24	0.00	0.00	0.00	0.00	0.00	
D101	0.00	89.76	10.24	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	
D102	0.00	6.59	93.41	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	
D18	0.73	0.00	0.00	67.07	0.00	0.00	0.00	0.24	0.00	2.44	5.12	0.73	1.95	0.00	0.00	1.46	0.00	0.00	0.00	2.20	13.17	4.88	0.00	0.00	0.00	0.00	0.00	0.00	0.00	
D20	0.00	0.24	0.00	0.00	99.76	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	
D34	0.00	0.00	0.00	0.00	0.00	99.51	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.49	0.00	
D46	0.00	0.00	0.00	0.00	0.00	0.00	100.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	
D47	0.24	0.00	0.00	3.41	0.00	0.00	0.00	0.00	0.24	0.00	0.00	0.00	0.00	0.00	0.00	0.98	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	5.61	0.00	0.00	
D62	0.00	0.00	0.00	0.00	0.00	0.00	0.00	93.90	0.00	0.73	0.00	0.24	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	3.41	0.00	0.00	1.71	0.00	
D66	0.00	0.00	0.00	3.17	0.00	0.00	0.00	3.41	0.00	99.51	0.00	0.00	0.00	0.00	0.00	0.24	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.73	2.93	0.00	0.00	
D64	10.73	0.00	0.00	8.54	0.00	0.00	0.00	3.90	0.00	36.59	0.00	2.44	0.00	0.00	0.00	9.27	0.00	0.00	0.00	20.00	0.00	0.73	7.80	0.00	0.00	0.00	0.00	0.00	0.00	0.00
D65	0.24	0.00	0.00	0.98	0.00	0.00	0.00	7.07	0.00	2.20	69.76	0.49	0.00	0.00	0.00	18.29	0.00	0.00	0.00	0.00	0.00	0.98	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
D66	0.00	0.00	0.00	0.00	0.00	0.00	0.00	2.44	0.00	0.00	0.00	0.00	0.00	0.00	0.00	97.32	0.00	0.00	0.00	0.24	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
D8	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	100.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
K 101Taipei05	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
K BurgHohenzollern06	0.00	0.00	0.00	2.20	0.00	0.00	0.00	0.00	0.00	2.20	4.15	2.20	0.00	0.00	0.00	67.32	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.95	0.00	0.00	0.00	0.00	0.00
K EierDeko06	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.95	0.00	0.00	84.63	10.73	0.00	0.73	1.95	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
K KentingNightMarkt05b	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	12.44	87.66	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
K KentingNightMarkt05c	0.00	0.00	0.00	0.00	0.00	0.00	0.00	3.17	0.00	2.44	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	84.39	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	10.00	0.00
K Knoxville06	0.00	0.00	0.00	3.66	0.00	0.00	0.00	1.22	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.49	0.49	0.00	94.15	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
K YellowCraneWuhan06	0.00	0.00	0.00	17.80	0.00	0.00	0.00	0.98	0.00	1.95	0.00	0.00	0.00	0.00	0.00	0.49	0.00	0.24	0.00	22.20	66.34	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
L Kufu09e	0.00	0.00	0.00	0.24	0.00	0.00	0.00	0.24	0.00	1.22	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	7.07	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	91.23	0.00	0.00
L Minneapolis06a	0.00	0.00	0.00	0.00	0.00	0.00	9.27	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.98	0.73	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	89.62	0.00	0.00
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23							

(a)

Ventana: 64x64

D1	65.43	0.00	0.00	20.36	0.00	0.00	0.00	0.45	0.00	4.98	3.62	7.92	0.00	0.00	1.58	0.00	0.00	0.00	4.30	0.00	0.00	1.36	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	
D101	0.00	75.05	21.95	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	
D102	0.00	23.53	76.47	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	
D18	2.94	0.00	0.00	53.85	0.00	0.00	0.00	4.07	0.45	4.75	4.75	1.36	4.07	0.00	0.00	3.39	0.23	0.68	1.81	11.31	4.07	2.26	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	
D20	0.00	1.36	0.45	0.00	98.19	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00		
D34	0.00	0.00	0.00	0.00	0.00	99.32	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.45	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00		
D46	0.00	0.00	0.00	0.45	0.00	0.00	92.53	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.68	0.90	0.00	1.13	0.00	0.00	4.30	0.00	0.00	0.00	0.00	0.00	4.30	0.00	0.00		
D47	0.68	0.00	0.00	7.47	0.00	0.00	65.36	0.00	3.85	0.45	0.00	0.90	0.00	4.30	3.39	4.98	1.13	0.45	6.11	0.90	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	
D62	0.23	0.00	0.00	0.00	0.00	0.00	0.00	95.48	0.00	0.00	0.00	2.71	0.00	0.00	0.00	0.00	0.00	1.58	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	
D66	0.00	0.00	0.00	16.06	0.00	0.00	1.81	0.00	75.57	1.58	0.00	0.00	0.00	0.00	0.90	0.23	0.00	0.00	1.58	1.81	0.45	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	
D64	7.01	0.00	0.00	13.12	0.00	2.26	0.00	0.45	2.71	0.90	45.93	2.49	3.85	0.00	0.00	6.11	0.00	0.00	7.92	0.90	0.90	5.43	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	
D65	2.24	0.00	0.00	9.05	0.00	0.00	0.00	3.85	0.00	2.94	58.60	3.17	0.00	0.00	0.00	16.06	0.00	0.00	0.68	0.00	0.00	3.62	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	
D66	0.03	0.00	0.68	1.36	0.00	0.00	0.00	8.60	0.00	0.23	86.43	0.00	0.00	0.00	0.00	0.00	0.00	0.00	2.49	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	
D8	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	100.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	
K 101Taipei05	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.68	0.00	0.00	0.00	0.00	0.00	0.00	0.00	99.10	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	
K BurgHohenzollern06	0.45	0.00	0.00	10.86	0.00	0.00	0.00	0.90	0.00	0.23	11.54	6.56	1.13	0.00	0.00	51.81	0.00	0.00	4.98	7.24	3.62	0.68	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
K EierDeko06	0.00	0.00	0.00	3.39	0.00	0.00	1.58	5.88	0.00	1.13	0.00	0.00	0.00	0.00	0.00	65.38	14.93	0.68	1.58	5.20	0.00	0.23	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
K KentingNightMarkt05b	0.00	0.00	0.00	0.23	0.00	0.00	0.68	0.90	0.00	0.00	0.00	0.00	0.00	0.00	0.00	7.24	83.71	0.45	0.90	2.04	0.23	0.23	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
K KentingNightMarkt05c	0.45	0.00	0.00	2.26	0.00	0.00	0.00	3.17	0.00	4.98	0.00	1.81	0.00	0.00	0.00	1.36	0.00	0.00	69.76	0.00	1.58	15.61	0.00									

Bibliografía

- [Abbadeni et al., 2000] Abbadeni, N., Ziou, D., and Wang, S. (2000). Autocovariance-based perceptual textural features corresponding to human visual perception. In *Pattern Recognition, 2000. Proceedings. 15th International Conference on*, volume 3, pages 901–904 vol.3.
- [Arivazhagan et al., 2006] Arivazhagan, S., Ganesan, L., and Bama, S. (2006). Fault segmentation in fabric images using gabor wavelet transform. *Machine Vision and Applications*, 16(6):356–363.
- [Asha et al., 2013] Asha, V., Bhajantri, N. U., and Nagabhushan, P. (2013). Periodicity extraction using superposition of distance matching function and one-dimensional haar wavelet transform. *CoRR*, abs/1311.3808.
- [Belkasim et al., 1991] Belkasim, S. O., Shridhar, M., and Ahmadi, M. (1991). Pattern recognition with moment invariants: A comparative study and new results. *Pattern Recogn.*, 24(12):1117–1138.
- [Bishop, 1996] Bishop, C. M. (1996). *Neural Networks for Pattern Recognition*. Oxford University Press, UK, UK.
- [Bovik et al., 1990] Bovik, A. C., Clark, M., and Geisler, W. S. (1990). Multichannel texture analysis using localized spatial filters. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 12(1):55–73.
- [Brodatz, 1966] Brodatz, P. (1966). *Texture: a photographic album for artist and designers*.

- [Carbajal-Degante et al., 2016] Carbajal-Degante, E., Nava, R., Olveres, J., Escalante-Ramírez, B., and Kybic, J. (2016). Texel-based image classification with orthogonal bases. *Proc. SPIE*, 9896:98961K–98961K–11.
- [Chandrasekaran et al., 2013] Chandrasekaran, V., Dantu, R., Jonnada, S., Thiyagaraja, S., and Subbu, K. P. (2013). Cuffless differential blood pressure estimation using smart phones. *IEEE Transactions on Biomedical Engineering*, 60(4):1080–1089.
- [Chang and Kuo, 1993] Chang, T. and Kuo, C. C. J. (1993). Texture analysis and classification with tree-structured wavelet transform. *IEEE Transactions on Image Processing*, 2(4):429–441.
- [Charalampidis and Kasparis, 2002] Charalampidis, D. and Kasparis, T. (2002). Wavelet-based rotational invariant roughness features for texture classification and segmentation. *IEEE Transactions on Image Processing*, 11(8):825–837.
- [Connors et al., 1984] Connors, R. W., Trivedi, M. M., and Harlow, C. A. (1984). Segmentation of a high-resolution urban scene using texture operators. *Computer Vision, Graphics, and Image Processing*, 25(3):273 – 310.
- [Dunn et al., 1994] Dunn, D., Higgins, W. E., and Wakeley, J. (1994). Texture segmentation using 2-d gabor elementary functions. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 16(2):130–149.
- [Estudillo-Romero and Escalante-Ramirez, 2011] Estudillo-Romero, A. and Escalante-Ramirez, B. (2011). Rotation-invariant texture features from the steered hermite transform. *Pattern Recognition Letters*, 32(16):2150 – 2162. Advances in Theory and Applications of Pattern Recognition, Image Processing and Computer Vision.
- [Freeman and Adelson, 1991] Freeman, W. T. and Adelson, E. H. (1991). The design and use of steerable filters. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 13(9):891–906.
- [Gibson, 1950] Gibson, J. (1950). *The perception of the visual world*. The riverside Press, Boston, USA.
- [Grigorescu and Petkov, 2003] Grigorescu, S. E. and Petkov, N. (2003). A dynamical system approach to texel identification in regular textures. In *Image and Signal*

- Processing and Analysis, 2003. ISPA 2003. Proceedings of the 3rd International Symposium on*, volume 1, pages 66–71 Vol.1.
- [Gui et al., 2011] Gui, Y., Chen, M., Ma, L., and Chen, Z. (2011). Texel based regular and near-regular texture characterization. In *Multimedia and Signal Processing (CMSP), 2011 International Conference on*, volume 1, pages 266–270.
- [Han et al., 2008] Han, J., McKenna, S. J., and Wang, R. (2008). *Regular Texture Analysis as Statistical Model Selection*, pages 242–255. Springer Berlin Heidelberg, Berlin, Heidelberg.
- [Haralick, 1979] Haralick, R. M. (1979). Statistical and structural approaches to texture. *Proceedings of the IEEE*, 67(5):786–804.
- [Haralick et al., 1973] Haralick, R. M., Shanmugam, K., and Dinstein, I. (1973). Textural features for image classification. *IEEE Transactions on Systems, Man, and Cybernetics*, SMC-3(6):610–621.
- [Hays et al., 2006] Hays, J. H., Leordeanu, M., Efros, A. A., and Liu, Y. (2006). Discovering texture regularity as a higher-order correspondence problem. In *9th European Conference on Computer Vision*.
- [Hu, 1962] Hu, M.-K. (1962). Visual pattern recognition by moment invariants, computer methods in image analysis. *IRE Transactions on Information Theory*, 8.
- [Jafari-Khouzani and Soltanian-Zadeh, 2005] Jafari-Khouzani, K. and Soltanian-Zadeh, H. (2005). Radon transform orientation estimation for rotation invariant texture analysis. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 27(6):1004–1008.
- [Jain and Farrokhnia, 1991] Jain, A. K. and Farrokhnia, F. (1991). Unsupervised texture segmentation using gabor filters. *Pattern recognition*, 24(12):1167–1186.
- [Jain et al., 1995] Jain, R., Kasturi, R., and Schunck, B. (1995). *Machine vision*. McGraw-Hill, first edition.
- [Jan and Hsueh, 1998] Jan, S.-R. and Hsueh, Y.-C. (1998). Window-size determination for granulometrical structural texture classification. *Pattern Recognition Letters*, 19(5–6):439 – 446.

- [Julesz, 1962] Julesz, B. (1962). Visual pattern discrimination. *IRE Transactions on Information Theory*, 8(2):84–92.
- [Julesz et al., 1978] Julesz, B., Gilbert, E., and Victor, J. (1978). Visual discrimination of textures with identical third-order statistics. *Biomedical Cybernetics*, 1(31):137–140.
- [Klette, 2010] Klette (2010). Basic multimedia imaging. ONLINE available, <http://www.mi.auckland.ac.nz/DATA/CCV/>.
- [Kohavi and Provost, 1998] Kohavi, R. and Provost, F. (1998). Glossary of terms. *Machine Learning*, 30(2-3):271–274.
- [Lee and Chen, 2002] Lee, K.-L. and Chen, L.-H. (2002). A new method for extracting primitives of regular textures based on wavelet transform. *International Journal of Pattern Recognition and Artificial Intelligence*, 16(01):1–25.
- [Leu, 2001] Leu, J.-G. (2001). On indexing the periodicity of image textures. *Image and Vision Computing*, 19(13):987 – 1000.
- [Liao and Pawlak, 1996] Liao, S. X. and Pawlak, M. (1996). On image analysis by moments. *IEEE Trans. Pattern Anal. Mach. Intell.*, 18(3):254–266.
- [Lin et al., 1997] Lin, H.-C., Wang, L.-L., and Yang, S.-N. (1997). Extracting periodicity of a regular texture based on autocorrelation functions. *Pattern Recogn. Lett.*, 18(5):433–443.
- [Lizarraga-Morales et al., 2013] Lizarraga-Morales, R. A., Sanchez-Yanez, R. E., and Ayala-Ramirez, V. (2013). Fast texel size estimation in visual texture using homogeneity cues. *Pattern Recognition Letters*, 34(4):414 – 422. *Advances in Pattern Recognition Methodology and Applications*.
- [Malik and Perona, 1990] Malik, J. and Perona, P. (1990). Preattentive texture discrimination with early vision mechanisms. *J. Opt. Soc. Am. A*, 7(5):923–932.
- [Mallat, 1989] Mallat, S. G. (1989). A theory for multiresolution signal decomposition: the wavelet representation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 11(7):674–693.

- [Manjunath and Ma, 1996] Manjunath, B. S. and Ma, W. Y. (1996). Texture features for browsing and retrieval of image data. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 18(8):837–842.
- [Manthalkar et al., 2003] Manthalkar, R., Biswas, P., and Chatterji, B. (2003). Rotation and scale invariant texture features using discrete wavelet packet transform. *Pattern Recognition Letters*, 24(14):2455 – 2462.
- [Marcos and Cristóbal, 2013] Marcos, J. V. and Cristóbal, G. (2013). Texture classification using discrete Tchebichef moments. *J. Opt. Soc. Am. A*, 30(8):1580–1591.
- [Marr, 1982] Marr, D. (1982). *Vision: A computational investigation into the human representation and processing of visual information*. MIT Press, San Francisco, CA, USA.
- [Martens, 1990] Martens, J. B. (1990). The Hermite transform-theory. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 38(9):1595–1606.
- [Matsuyama et al., 1983] Matsuyama, T., Miura, S.-I., and Nagao, M. (1983). Structural analysis of natural textures by fourier transformation. *Journal CVGIP*, 24(3):347–362.
- [Mirmehdi et al., 2009] Mirmehdi, M., Xie, X., and Suri, J. (2009). *Handbook of Texture Analysis*. Imperial College Press, London, UK, UK.
- [Mitra and Poellabauer, 2012] Mitra, P. and Poellabauer, C. (2012). Emergency response in smartphone-based mobile ad-hoc networks. In *2012 IEEE International Conference on Communications (ICC)*, pages 6091–6095.
- [Mukundan et al., 2001] Mukundan, R., Ong, S. H., and Lee, P. A. (2001). Image analysis by Tchebichef moments. *IEEE Transactions on Image Processing*, 10(9):1357–1364.
- [Nava et al., 2012] Nava, R., Escalante-Ramírez, B., and Cristóbal, G. (2012). *Texture Image Retrieval Based on Log-Gabor Features*, pages 414–421. Springer Berlin Heidelberg, Berlin, Heidelberg.

- [Oh et al., 1999] Oh, G., Lee, S., and Shin, S. Y. (1999). Fast determination of textural periodicity using distance matching function. *Pattern Recognition Letters*, 20(2):191–197.
- [Olshausen and Field, 1997] Olshausen, B. A. and Field, D. J. (1997). Sparse coding with an overcomplete basis set: A strategy employed by v1? *Vision Research*, 37(23):3311–3325.
- [Palshikar, 2009] Palshikar, G. K. (2009). Simple algorithms for peak detection in time-series. In *First Int. Conf. Advanced Data Analysis, Business Analytics and Intelligence (ICADABAI2009), Ahmedabad, India, 6-7 June 2009*.
- [Parkkinen et al., 1990] Parkkinen, J., Selkäinaho, K., and Oja, E. (1990). Detecting texture periodicity from the co-occurrence matrix. *Pattern Recogn. Lett.*, 11(1):43–50.
- [Petrou and Garcia-Sevilla, 2006] Petrou, M. and Garcia-Sevilla, P. (2006). *Image processing: dealing with texture*. Wiley, first edition.
- [Rao and Jain, 1992] Rao, A. R. and Jain, R. C. (1992). Computerized flow field analysis: oriented texture fields. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 14(7):693–709.
- [Rao and Lohse, 1993] Rao, A. R. and Lohse, G. L. (1993). Towards a texture naming system: Identifying relevant dimensions of texture. In *Visualization, 1993. Visualization '93, Proceedings., IEEE Conference on*, pages 220–227.
- [Rayanchu et al., 2011] Rayanchu, S., Patro, A., and Banerjee, S. (2011). Airshark: Detecting non-wifi rf devices using commodity wifi hardware. In *Proceedings of the 2011 ACM SIGCOMM Conference on Internet Measurement Conference, IMC '11*, pages 137–154, New York, NY, USA. ACM.
- [Scholkmann et al., 2012] Scholkmann, F., Boss, J., and Wolf, M. (2012). An efficient algorithm for automatic peak detection in noisy periodic and quasi-periodic signals. *Algorithms*, 5(4):588–603.

- [Silvan-Cardenas and Escalante-Ramirez, 2006] Silvan-Cardenas, J. L. and Escalante-Ramirez, B. (2006). The multiscale Hermite transform for local orientation analysis. *IEEE Transactions on Image Processing*, 15(5):1236–1253.
- [Sokolova and Lapalme, 2009] Sokolova, M. and Lapalme, G. (2009). A systematic analysis of performance measures for classification tasks. *Information Processing and Management*, 45(4):427 – 437.
- [Starovoitov et al., 1998] Starovoitov, V. V., Jeong, S.-Y., and Park, R.-H. (1998). Texture periodicity detection: features, properties, and comparisons. *IEEE Transactions on Systems, Man, and Cybernetics - Part A: Systems and Humans*, 28(6):839–849.
- [Stevens, 1979] Stevens, K. A. (1979). Surface perception from local analysis of texture and contour. Technical report.
- [Tamura et al., 1978] Tamura, H., Mori, S., and Yamawaki, T. (1978). Textural features corresponding to visual perception. *IEEE Transactions on Systems, Man, and Cybernetics*, 8(6):460–473.
- [Teh and Chin, 1988] Teh, C. H. and Chin, R. T. (1988). On image analysis by the methods of moments. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 10(4):496–513.
- [Terzopoulos, 1985] Terzopoulos, D. (1985). Co-occurrence analysis of speech waveforms. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, ASSP-33:52–96.
- [Unser, 1986] Unser, M. (1986). Sum and difference histograms for texture classification. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, PAMI-8(1):118–125.
- [Unser, 1995] Unser, M. (1995). Texture classification and segmentation using wavelet frames. *IEEE Transactions on Image Processing*, 4(11):1549–1560.
- [Valois et al., 1982] Valois, R. L. D., Albrecht, D. G., and Thorell, L. G. (1982). Spatial frequency selectivity of cells in macaque visual cortex. *Vision Research*, 22(5):545 – 559.

- [van Dijk and Martens, 1997] van Dijk, A. M. and Martens, J.-B. (1997). Image representation and compression with steered hermite transforms. *Signal Processing*, 56(1):1 – 16.
- [Van Rullen and Thorpe, 2001] Van Rullen, R. and Thorpe, S. J. (2001). Rate coding versus temporal order coding: What the retinal ganglion cells tell the visual cortex. *Neural Comput.*, 13(6):1255–1283.
- [Witkin, 1983] Witkin, A. P. (1983). Scale-space filtering. In *Proceedings of the Eighth International Joint Conference on Artificial Intelligence - Volume 2, IJCAI'83*, pages 1019–1022, San Francisco, CA, USA. Morgan Kaufmann Publishers Inc.
- [Wu et al., 2000] Wu, P., Manjunath, B., Newsam, S., and Shin, H. (2000). A texture descriptor for browsing and similarity retrieval. *Journal of Signal Processing: Image Communication*, 16(1-2):33–43.
- [Yap et al., 2003] Yap, P. T., Paramesran, R., and Ong, S.-H. (2003). Image analysis by krawtchouk moments. *IEEE Transactions on Image Processing*, 12(11):1367–1377.
- [Young, 1991] Young, R. A. (1991). Oh say, can you see? the physiology of vision.
- [Young and Lesperance, 2001] Young, R. A. and Lesperance, R. M. (2001). The gaussian derivative model for spatial-temporal vision. *I. Cortical Model. Spatial Vision*, 2001:3–4.