

U

P

T

Análisis de imágenes de color con momentos radiales ortogonales

por

Alicia Noriega Escamilla

Tesis sometida como requisito parcial
para obtener el grado de

**MAESTRO EN
COMPUTACIÓN ÓPTICA**

en la

Universidad Politécnica de Tulancingo

Fecha Agosto 2019

Tulancingo de Bravo, Hidalgo

Supervisada por:

Dr. César Joel Camacho Bello

©UPT

El autor otorga a la UPT el permiso de reproducir y
distribuir copias en su totalidad o en partes de esta tesis.



Dedicatoria

A mi mamá por darme durante toda mi vida su cariño y apoyo incondicional.

A mis hijas Illian y Pau, por ser mi motivación para superarme cada día, y ser la alegría de mi existencia.

A mi pequeño Abel, por ser mi compañero de vida y por estar a mi lado apoyandome en cada nuevo proyecto.

Agradecimientos

Agradezco al Consejo Nacional de Ciencia y Tecnología (CONACyT) por el apoyo para realizar mis estudios de maestría a través de la beca otorgada y a la Universidad Politécnica de Tulancingo por darme la oportunidad de pertenecer al grupo de estudiantes que conforman la Maestría en Computación Óptica.

A mi asesor el Dr. César Joel Camacho Bello por su apoyo, asesoría y dirección para realizar esta tesis.

Al Dr. José Sáenz Landete, Dr. Philip Siegmann y el Dr. Hilario Gómez Moreno del Depto. de Teoría de la señal y Telecomunicaciones de la Universidad de Alcalá por permitirme realizar una estancia de investigación; brindandome su apoyo, asesoría y hospitalidad en todo momento.

Y finalmente a mis sinodales, Dr. Alfonso Padilla Vivanco, Dra. Carina Toxqui Quitl, Mtra. Lucia Gutiérrez Lazcano, por sus observaciones que contribuyeron de forma significativa a mejorar este trabajo de tesis.

Prefacio

Los momentos ortogonales de una imagen y sus invariantes han sido ampliamente utilizados en la visión por computadora y el reconocimiento de patrones. Los objetos en el mundo real muestran una complejidad natural, debido al cambio de escala, la traslación, el ruido y algunos otros factores involucrados en la naturaleza de la adquisición de imágenes. Esto nos ha llevado al desarrollo de descriptores como los momentos invariantes que pueden sobrepasar las situaciones que distorsionan la imagen, como las transformaciones geométricas y el ruido. Estos momentos invariantes son capaces de obtener suficiente información para separar objetos en clases.

Actualmente, la mayor parte de la información para el reconocimiento de objetos proviene de imágenes en color. Sin embargo, la mayor parte del estudio para el desarrollo de descriptores invariantes está en imágenes en escala de grises. Por lo general, se utilizan dos métodos en el procesamiento del color de la imagen, el primero es transformar la imagen en color en escala de grises, con la cual hay una pérdida de información de color, el segundo es descomponer la imagen en color en sus tres canales, (R, G, B) y calcular los momentos de cada canal.

Para resolver este problema, el uso del álgebra de cuaternión se ha implementado en la última década para representar imágenes en color con Momentos de Cuaternión. Por otro lado, el enfoque Multicanal surge como una propuesta alternativa a los momentos de cuaternión para la representación y el reconocimiento de imágenes en color. Las pruebas realizadas obtienen resultados superiores con los Momentos Multicanal en comparación con los Momentos de Cuaternión. Finalmente, se analiza la propuesta implementada por Sing and Sing, Clasificador de Kernel Múltiple para el reconocimiento de objetos, el cual se basa en los Momentos de Zernike, combinan como descriptores los Momentos Multicanal, los Momentos del gradiente de las imágenes en color y los histogramas de color.

Resumen

Los momentos y los invariantes de momento son funciones que describen las características de una imagen y son robustos a los cambios geométricos y al ruido. Durante los últimos años han sido ampliamente utilizados en el Reconocimiento de Patrones y el Análisis de imagen. La teoría de momentos se ha aplicado principalmente en imágenes monocromáticas y en escalas de grises. Sin embargo, el color es un descriptor que simplifica la identificación de objetos, por lo que durante la última década la teoría de momentos se aplicó a imágenes en color por medio de los momentos de cuaternión, y recientemente han surgido nuevas propuestas; los momentos multicanal y el clasificador multikernel. En este trabajo de tesis se presenta un estudio del marco teórico de estos tres métodos y se realiza también un análisis comparativo del desempeño de éstos en 4 modelos de color distintos. Se examina también su capacidad como descriptores para clasificar imágenes de melanoma. Finalmente, se presentan las conclusiones y trabajos a futuro.

Abstract

Moments and moment invariants are functions that describe the image characteristics and are robust to geometric changes and noise. In recent years they have been used in Pattern Recognition and Image Analysis. The theory of moments has been applied mainly in monochrome images and in gray scales. However, color is a descriptor that simplifies the identification of objects, so during the last decade the theory of moments was applied to color images by means of quaternion moments, and new proposals have recently emerged; the multichannel moments and the multikernel classifier. In this thesis a study of the theoretical framework of these three methods is presented and a comparative analysis of their performance in four different color models is also carried out. Their ability as descriptors to classify melanoma images is also examined. Finally, the conclusions and future works are presented.

Índice general

1. Introducción	1
1.1. Antecedentes	1
1.2. Planteamiento del problema	1
1.3. Objetivos	2
1.3.1. Objetivo general	2
1.3.2. Objetivos específicos	2
1.4. Estado del Arte	2
2. Teoría de momentos	8
2.1. Introducción	8
2.2. Momentos geométricos	8
2.3. Momentos complejos	11
2.4. Momentos ortogonales	12
2.5. Cómputo exacto de los momentos radiales	14
3. Momentos radiales ortogonales	21
3.1. Introducción	21
3.2. Invariancia a la rotación	24
3.3. Invariancia a la escala	26
3.4. Invariancia a la traslación	26
3.5. Error de reconstrucción	32
4. Momentos radiales de imágenes en color	36
4.1. Introducción	36
4.2. Momentos Cuaternión.	37
4.2.1. Invariancia a la rotación	41
4.2.2. Invariancia a la escala	43
4.2.3. Invariancia a la traslación	44
4.2.4. Error de reconstrucción de imágenes RGB	46
4.3. Momentos Multicanal	46
4.3.1. Invariancia a la rotación	47
4.3.2. Invariancia a la escala	48
4.3.3. Invariancia a la traslación	48
4.3.4. Error de reconstrucción de imágenes RGB con momentos Multicanal	49

4.4. Clasificador Multikernel	49
4.4.1. Descriptores de forma.	50
4.4.2. Descriptores de textura.	51
4.4.3. Descriptores de color	54
5. Análisis de imágenes en color con momentos radiales	61
5.1. Introducción	61
5.2. Momentos Multicanal de Zernike y Tchebichef	62
5.3. Experimentos	63
5.3.1. Reconstrucción de imágenes.	63
5.3.2. Invariancia a transformaciones geométricas	65
5.3.3. Invariancia a la rotación	65
5.3.4. Invariancia a la escala.	65
5.3.5. Invariancia a la traslación	66
5.3.6. Robustez al ruido	66
5.3.7. Reconocimiento de imágenes	70
5.4. Análisis de resultados	86
6. Análisis comparativo de los modelos de color y clasificadores	88
6.1. Introducción	88
6.2. Base de Datos y modelos de clasificación	88
6.3. Análisis modelo de color RGB	89
6.4. Análisis modelo de color CMYK	97
6.5. Análisis con momentos en el modelo de color HSI	101
6.6. Análisis con momentos en el Modelo de color YCbCr	105
6.7. Análisis de resultados	109
7. Clasificación de Melanoma	111
7.1. Introducción	111
7.2. Base de datos	113
7.3. Clasificación con el modelo de color RGB	114
7.4. Clasificación con el modelo de color HSI	117
7.5. Análisis con momentos en el modelo de color YCbCr	118
7.6. Análisis con momentos en el modelo de color CMYK	120
7.7. Análisis de resultados	123
8. Conclusiones y trabajo a futuro.	125
A. Modelos de Clasificación	127
A.1. Árboles de decisión.	127
A.2. Análisis discriminante	128
A.3. Máquinas de soporte vectorial.	128
A.4. Clasificadores de vecinos cercanos	129
A.5. Ensemble	130

B. Modelos de color	131
B.1. Modelo de color RGB	131
B.2. Modelo de color CMYK	131
B.3. Modelo de color HSI	132
B.4. Modelo de color YCbCr	134

Índice de figuras

2.1. Elipse: ϕ es el ángulo de rotación entre los ejes (x,y) y los ejes principales (x', y').	10
2.2. Aproximación de píxeles cartesianos a una región circular.	15
2.3. Esquema de píxeles polares propuestos por Xin et al.	16
2.4. Interpolación de la imagen "Baboon" de tamaño 102×102 usando el método "cubic" a través de la función "interp2" de Matlab.	18
3.1. Diagrama de bloques del método propuesto por Honarvar et al. [13] para calcular los polinomios radiales de Zernike $R_n^m(r)$	22
3.2. Gráficas de los polinomios radiales de Zernike.	23
3.3. Algunas imágenes de la base de datos Columbia Object Library (COIL 100).	24
3.4. Imagen 74-0 de la base de datos COIL-100 con diferentes ángulos de rotación.	25
3.5. Objeto 52_0 de la base de datos COIL-100 a diferentes escalas.	27
3.6. Objeto 59_19 original de la base de datos COIL-100 y trasladado.	31
3.7. Error de reconstrucción del objeto 74_0 con momentos de Zernike y distintos órdenes.	33
4.1. Objeto 74_0 en color de la base de datos COIL-100 es rotado a 30° , 60° , 90° , 120° , 150° y 180°	42
4.2. Objeto 52_0 a color de la base de datos COIL-100 a diferentes escalas.	43
4.3. Objeto 59_19 a color de la base de datos COIL-100 trasladado.	45
4.4. Error de reconstrucción del objeto 74_0 para distintos órdenes de los momentos Cuaternión de Zernike.	46
4.5. Error de reconstrucción del objeto 74_0 en color con momentos Multicanal de Zernike y distintos órdenes.	50
4.6. Gradiente del Objeto 74_0 en color de la base de datos COIL-100 rotado a 30° , 60° , 90° , 120° , 150° y 180°	52
4.7. Gradiente del Objeto 52_0 a color de la base de datos COIL-100 a diferentes escalas.	53
4.8. Gradiente del Objeto 59_19 en color de la base de datos COIL-100 con posición original y trasladado.	54
4.9. Histogramas del objeto 74__0 de la base de datos COIL-100, rotada y escalada.	55

4.10. Objeto 74_0 de la base de datos COIL-100 con diversas escalas. . . .	56
4.11. Objeto 74_0 de la base de datos Coil 100.	57
5.1. Imagen a color “Lena”.	64
5.2. Imagen de referencia y escalada del Obj_17 a color de la base de datos COIL-100.	66
5.3. Obj_23 de la base de datos COil-100 en posición original y trasladado.	67
5.4. Obj_25 de la base de datos COIL-100 libre de ruido y con ruido añadido.	68
5.5. NIRE para imagenes libre de ruido y con ruido del Obj_25 obtenidos por Hosny y Darwish [1].	68
5.6. NIRE obtenido por nosotros para los momentos Multicanal de Zernike para el Obj_25.	69
5.7. NIRE obtenido por nosotros para los momentos Multicanal de Tchebichef para el Obj_25.	69
5.8. Tasa de reconocimiento R(%) usando distintas medidas a) L_1-norm , b) L_2-norm c) χ^2 y d) <i>Canberra</i> publicadas por Hosny y Darwish [1]. .	71
5.9. Tasa de reconocimiento R(%) usando distintas medidas a) L_1-norm , b) L_2-norm c) χ^2 y d) <i>Canberra</i> obtenido por nosotros.	72
5.10. Diagrama de dispersión para las primeras 25 clases de la base de datos COIL-100 en condiciones normales con MMZ.	75
5.11. Diagrama de dispersión para las primeras 25 clases de la base de datos COIL-100 en condiciones normales con MMT.	76
5.12. Diagrama de dispersión para las primeras 25 clases de la base de datos COIL-100 en rotación aleatoria con MMZ.	78
5.13. Diagrama de dispersión para las primeras 25 clases de la base de datos COIL-100 en rotación aleatoria con MMT.	79
5.14. Diagrama de dispersión para las primeras 25 clases de la base de datos COIL-100 escala con MMZ.	81
5.15. Diagrama de dispersión para las primeras 25 clases de la base de datos COIL-100 a escala con MMT.	82
5.16. Diagrama de dispersión para las primeras 25 clases de la base de datos COIL-100 con ruido Gaussiano con MMZ.	84
5.17. Diagrama de dispersión para las primeras 25 clases de la base de datos COIL-100 con ruido Gaussiano con MMT.	85
6.1. Vocales en distintos colores y ejemplos de vocales con ruido añadido. .	89
6.2. Diagrama de dispersión para los momentos Cuaternión de Zernike en el modelo RGB.	90
6.3. Diagrama de dispersión para los momentos Multicanal de Zernike en el modelo RGB.	91
6.4. Diagrama de dispersión para el clasificador Multikernel de Zernike en el modelo RGB.	92
6.5. Matriz de confusión obtenida a partir de los momentos de Cuaternión de Zernike en el Modelo de color RGB.	94
6.6. Matriz de confusión obtenida a partir de los momentos Multicanal de Zernike del Modelo de color RGB.	95

6.7. Matriz de confusión obtenida a partir del Clasificador Multikernel de Zernike del Modelo de color RGB.	96
6.8. Diagrama de dispersión de los momentos Cuaternión de Zernike en el modelo de color CMYK.	98
6.9. Diagrama de dispersión de los Momentos Multicanal de Zernike en el modelo de color CMYK.	99
6.10. Diagrama de dispersión Clasificador Multikernel CMYK.	100
6.11. Diagrama de dispersión de los momentos Cuaternión en el Modelo de color HSI.	101
6.12. Diagrama de dispersión de los momentos Multicanal en el Modelo de color HSI.	102
6.13. Diagrama de dispersión del Clasificador Multikernel en el Modelo de color HSI.	103
6.14. Diagrama de dispersión de los momentos Cuaternión en el Modelo de color YCbCr.	105
6.15. Diagrama de dispersión de los momentos Multicanal en el Modelo de color YCbCr.	106
6.16. Modelo de dispersión del Clasificador Multikernel YCbCr.	107
7.1. Normalización de la base de datos HAM 10000.	114
A.1. Estructura básica de un árbol de decisión.	128
A.2. Hiperplano de separación entre dos clases.	129
B.1. Esquema del cubo de color RGB.	132
B.2. Modelo de color CMY de la imagen de “Lena”	133
B.3. Modelo de color HSI de la imagen de “Lena” de tamaño 220×220 . . .	135
B.4. Canales Y, Cb, Cr y el modelo de color YCbCr de la imagen de “Lena” de tamaño 220×220	136

Índice de tablas

3.1. MSE para evaluar la efectividad de la invariante de rotación.	25
3.2. momentos de Zernike para los diferentes ordenes y desviación estándar de los momentos a diferentes ángulos.	26
3.3. Error cuadrático medio MSE para evaluar la efectividad de la invariante de escala.	26
3.4. Resultado de la invariancia a la escala de los primeros diez momentos de Zernike.	27
3.5. Invariante a la traslación de acuerdo a la propuesta de Chong con momentos de Zernike para diferentes ordenes n, m	31
3.6. momentos de Zernike para diferentes ordenes n, m	32
4.1. MSE para evaluar la efectividad de la invariante de rotación de los momentos Cuaternión de Zernike.	42
4.2. momentos de Cuaternión Zernike para diferentes ordenes en diferentes rotaciones.	42
4.3. Error cuadrático medio (MSE) de la imagen de referencia y las imágenes con diferentes escalas de los momentos Cuaternión de Zernike.	43
4.4. Momentos Cuaternión de Zernike para diferentes ordenes.	44
4.5. Momentos Cuaternión de Zernike para diferentes ordenes.	45
4.6. MSE para evaluar la efectividad del invariante de rotación de los momentos Multicanal de Zernike.	48
4.7. Error cuadrático medio MSE para evaluar la efectividad de la invariante de escala de los momentos Multicanal de Zernike.	48
4.8. momentos multicanal de Zernike para diferentes ordenes de la Fig. 4.3.	49
4.9. MSE para evaluar la efectividad de la invariante de rotación de los momentos de Zernike del gradiente de la imagen en color.	52
4.10. Error cuadrático medio (MSE) para evaluar la efectividad de la invariante de escala de los momentos de Zernike del gradiente de la imagen en color.	53
4.11. momentos de cuaternión de Zernike para diferentes ordenes n, m	54
4.12. MSE para evaluar la efectividad de la invariante de rotación del histograma.	56
4.13. Error cuadrático medio MSE para evaluar la efectividad de la invariante de escala del histograma.	56
4.14. Valores normalizados del histograma del objeto 74.	57

5.1. Descriptores obtenidos de acuerdo al orden de los momentos MMZ y MMT.	62
5.2. Reconstrucción de la imagen Lena publicados por Hosny y Darwish [1].	64
5.3. Reconstrucción de la imagen Lena con momentos Multicanal de Zernike y Tchebichef.	64
5.4. MSE para evaluar la efectividad de la invariante de rotación de los momentos de Cuaternión de Zernike y los momentos Multicanal de Tchebichef.	65
5.5. Error cuadrático medio MSE para evaluar la efectividad de la invariante de escala.	65
5.6. Error cuadrático medio MSE para evaluar la efectividad de la invariante de traslación.	66
5.7. Tasa de Reconocimiento de los momentos MMZ y MMT con 60 descriptores.	71
5.8. Tasa de Reconocimiento de los momentos MMZ y MMT con 60 descriptores de la base de datos rotada aleatoriamente.	71
5.9. Promedio de la Tasa de Reconocimiento de la base de datos COIL-100 en distintas escalas.	73
5.10. Tasa de Reconocimiento de la base de datos COIL-100 con ruido Gaussiano.	73
5.11. Exactitud para los diferentes clasificadores con momentos Multicanal de Zernike, Modelo de color RGB en condiciones normales.	74
5.12. Exactitud para los diferentes clasificadores con momentos Multicanal de Zernike, Modelo de color RGB con rotación aleatoria.	77
5.13. Exactitud para los diferentes clasificadores con momentos Multicanal de Zernike, Modelo de color RGB con escala 1.5.	80
5.14. Exactitud para los diferentes clasificadores con momentos Multicanal de Zernike, Modelo de color RGB con ruido Gaussiano.	83
6.1. Exactitud para los diferentes clasificadores con momentos Cuaternión, Multicanal y Clasificador Multikernel del Modelo de color RGB.	93
6.2. Exactitud para los diferentes clasificadores con momentos Cuaternión de Zernike del Modelo de color CMYK.	97
6.3. Exactitud para los diferentes clasificadores con momentos Cuaternión, Multicanal y Clasificador Multikernel de Zernike del Modelo de color HSI.	104
6.4. Exactitud para los diferentes clasificadores con momentos Cuaternión, Multicanal y Clasificador Multikernel de Zernike del Modelo de color YCbCr.	108
6.5. Exactitud de los momentos Cuaternión, Multicanal y Clasificador Multikernel en los distintos Modelos de color.	109
7.1. Exactitud para los diferentes clasificadores con momentos Cuaternión, Multicanal y Clasificador Multikernel de Zernike del Modelo de color RGB usando 7 etiquetas.	115

7.2. Exactitud para los diferentes clasificadores con momentos Cuaternión, Multicanal y Clasificador Multikernel de Zernike de color RGB usando 2 etiquetas.	116
7.3. Exactitud para los diferentes clasificadores con momentos Cuaternión, Multicanal y Clasificador Multikernel de Zernike del Modelo de color HSI usando 7 etiquetas.	117
7.4. Exactitud para los diferentes clasificadores con momentos Cuaternión, Multicanal y Clasificador Multikernel de Zernike del Modelo de color HSI usando 2 etiquetas.	118
7.5. Exactitud para los diferentes clasificadores con momentos Cuaternión del Modelo de color YCbCr usando 7 etiquetas.	119
7.6. Exactitud para los diferentes clasificadores con momentos Cuaternión, Multicanal y Clasificador Multikernel de Zernike del Modelo de color YCbCr usando 2 clases.	120
7.7. Exactitud para los diferentes clasificadores con momentos Cuaternión, Multicanal y Clasificador Multikernel de Zernike CMYK usando 7 etiquetas.	121
7.8. Exactitud para los diferentes clasificadores con momentos Cuaternión, Multicanal y Clasificador Multikernel de Zernike CMYK usando 2 etiquetas.	122
7.9. Exactitud de clasificación con los momentos Cuaternión, Multicanal y Clasificador Multikernel de Zernike.	123
A.1. Clasificadores en la aplicación Classification Learner de Matlab.	127

Capítulo 1

Introducción

1.1. Antecedentes

Una imagen digital posee una gran cantidad de información que durante el proceso de adquisición puede sufrir degradaciones por diversos cambios geométricos (como rotación, escala, traslación, oclusión), fotométricos y ruido. Para clasificar correctamente un objeto en la clase a la cual pertenece existen tres enfoques; el primero consiste en tener un conjunto de datos con todas las deformaciones posibles que la imagen pudiera sufrir, el segundo enfoque es realizar una normalización de la imagen, y el tercero es el uso de invariantes que son insensibles a deformaciones particulares y que proporcionan suficiente información para distinguir objetos pertenecientes a diferentes clases.

Los invariantes de momento son cantidades numéricas provenientes de los momentos de una imagen, además son descriptores globales robustos a cambios geométricos y al ruido, que han sido ampliamente utilizados en el Reconocimiento de Patrones y en el Análisis de imagen. La teoría de momentos se ha aplicado principalmente en imágenes monocromáticas y en escalas de grises. Sin embargo, el color es un descriptor que simplifica la identificación de objetos, por lo que durante la última década la teoría de momentos se aplicó a imágenes en color por medio de los momentos Cuaternión, recientemente por los momentos Multicanal y sus invariantes.

1.2. Planteamiento del problema

Durante la última década los momentos Cuaternión han sido muy utilizados para el análisis de imágenes en color debido a que correlacionan los tres canales de una imagen RGB mediante la teoría de cuaterniones. Sin embargo, sólo puede analizar características de forma del objeto y no pueden discriminar el color del objeto.

1. Introducción

Recientemente han sido publicadas nuevas propuestas para el análisis de imágenes en color, los momentos Multicanal y el clasificador multikernel. En este trabajo de tesis se realiza un análisis del rendimiento de éstos tres métodos en cuatro modelos de color, RGB, CMYK, HSI y YCbCr. De igual manera se analiza su capacidad de clasificación haciendo uso de los distintos modelos de clasificación de la aplicación *ClassificationLearner* de Matlab.

1.3. Objetivos

1.3.1. Objetivo general

Realizar un análisis comparativo de la capacidad descriptiva de imágenes en color en los modelos, RGB, CMYK, HSI y YCbCr de los momentos Cuaternión, momentos Multicanal y el clasificador multikernel.

1.3.2. Objetivos específicos

- Realizar un estudio del marco teórico de los momentos ortogonales.
- Realizar un estudio del marco teórico de los momentos Cuaternión, momentos Multicanal y el clasificador multikernel.
- Análisis del código de programación en el entorno Matlab para los momentos ortogonales, momentos Cuaternión, momentos Multicanal y el clasificador multikernel usando como kernel los polinomios de Zernike.
- Análisis de los distintos modelos de color con los momentos radiales para la descripción de imágenes en color.
- Análisis de los distintos modelos de color con los momentos radiales para la descripción de melanoma.
- Clasificación de los tres métodos en los cuatro modelos de color con diversos modelos de clasificación.
- Análisis de los momentos Multicanal de Tchebichef y clasificación en el modelo de color RGB.

1.4. Estado del Arte

Los momentos han sido ampliamente utilizados con imágenes binarias o en escala de grises. No obstante, la mayoría de las imágenes hoy en día son cromáticas;

1. Introducción

generalmente el análisis de una imagen a color consiste en procesar los canales rojo, verde y azul por separado usando técnicas digitales de procesamiento de imágenes y finalmente, combinar los resultados individuales de salida. Sin embargo, los resultados del procesamiento de cada componente de color no siempre son equivalentes al procesamiento directo en el espacio de color [1]. En la última década los momentos Cuaternión han sido extensamente utilizados en el procesamiento de imágenes en color y recientemente los momentos Multicanal han demostrado tener una mejor capacidad descriptiva en imágenes RGB [2]. A continuación se presenta una breve revisión de la bibliografía existente de los momentos Cuaternión, los momentos Multicanal y la propuesta del clasificador multikernel.

- Li-Qiang Guo et al. [3] en 2011 presentan los momentos Cuaternión Fourier Mellin y sus invariantes para rotación, escala y traslación a fin de describir imágenes en color. Usan la base de datos de imágenes Coil-100 de la Universidad de Columbia y comparan sus resultados con el método basado en FFT (Transformada Rápida de Fourier).
- Chen et al.[4] en 2012, derivaron los momentos Cuaternión de Zernike y proporcionaron el marco teórico para construir un conjunto de invariantes con respecto a la rotación, escala y traslación. Para las pruebas de reconstrucción de imagen a color, invariancia de rotación, escala, traslación y reconocimiento de objetos en color utilizan imágenes de la base de datos pública Amsterdam Library of Object Images. Los resultados son comparados con momentos convencionales de Zernike obtenidos a partir de la imagen en escala de grises.
- Shao et al.[5] en 2013 presentan los momentos Cuaternión Bessel-Fourier, derivan un nuevo conjunto de descriptores invariantes basados en la magnitud y la información de fase de los momentos Cuaternión Bessel-Fourier. Para realizar las pruebas utilizan imágenes en color de la base de datos pública Amsterdam Library of Object Images y de la biblioteca de imágenes de la Universidad de Columbia. Los resultados son comparados con los momentos Cuaternión Zernike, los momentos Cuaternión pseudo-Zernike y los momentos Cuaternión ortogonal Fourier-Mellin.
- Guo et al.[6] en 2014 proponen nuevos conjuntos de descriptores de momentos de cuaternión para imagen en color. Construyen invariantes de momentos complejos para imágenes en escala de grises. Utilizan imágenes de la biblioteca de la Universidad de Columbia y hacen comparaciones con los momentos invariantes Cuaternión de Zernike.

1. Introducción

- Chen et al. [7] en 2014 utilizan el álgebra de cuaterniones, para generalizar los momentos de tipo complejo (CTM) convencionales para imágenes en escala de grises, como momentos de tipo cuaternión (QTM) a partir de la cual derivan un conjunto de invariantes cuaternión para las transformaciones de rotación, escala y traslación de imágenes. El rendimiento de los momentos propuestos se evalúa teniendo en cuenta varios marcos de aplicaciones que van desde la reconstrucción de imágenes en color y el reconocimiento facial hasta el registro de imágenes. Utilizan imágenes de la biblioteca de la Universidad de Columbia y la base de datos proporcionada por la Universidad de Essex. Además hacen comparaciones entre los resultados obtenidos y los momentos cuaternión Fourier-Mellin.
- Zhu et al. [8] en 2014 dan a conocer los momentos Cuaternión discretos de Tchebichef y construyen un conjunto de invariantes a la traslación y escala. Usaron la base de datos pública Amsterdam Library of Object Images y la base de datos de rostros en color ensamblada por el grupo de Visión por Computadora de la universidad de Essex. Los resultados son comparados con los momentos cuaternión Fourier-Mellin.
- Chen et al. [9] en 2016 presentan nuevos momentos de tipo cuaternión basado en una representación cuaternión de imágenes en RGB mejorado que combina información de color y profundidad. Utiliza la base de datos de rostros CIN 2D+3D Georgia Tech (GT) comparando los resultados obtenidos con los momentos Cuaternión Pseudo-Zernike, los momentos Cuaternión de Zernike, y los momentos Cuaternión ortogonales de Fourier-Mellin.
- Yang et al.[10] en 2016 proponen los momentos exponenciales Cuaternión para describir imágenes en color, y analizan la invariabilidad y las propiedades de robustez. Usaron 5000 imágenes en color de 150 categorías de COREL photo gallery para las pruebas realizadas y comparan los resultados con los momentos Cuaternión de Legendre y Pseudo-Zernike.
- Chen et al.[11] proponen momentos Cuaternión de Pseudo-Zernike usando una representación de cuaternión mejorada que utiliza completamente el dominio de cuaterniones de cuatro dimensiones. Realiza pruebas de detección de empalme de imágenes en color. Usa las bases de datos Columbia color DVMM, CASIA v1.0 and v2.0, and Wild Web . Usa como punto de comparación una red neuronal de quaternion back-propagation.
- Hosny et al. [2] Proponen un nuevo método para el cálculo de momentos Cuaternión Legendre-Fourier en coordenadas polares. Se derivan invariantes Cuaternión Legendre-Fourier, rotación explícita, escalado y traslación (RST)

1. Introducción

- Singh y Singh [12] proponen invariantes multicanal para imágenes en color y comparan el rendimiento de éstos con los momentos Cuaternión existentes, utilizando como kernel los momentos de Zernike. Muestra que los descriptores Cuaternión sufren una alta redundancia de información en comparación con los descriptores Multicanal, los cuales se desempeñan mejor en la tarea de reconocimiento de imágenes.
- Hosny y Darwish [13] proponen un nuevo conjunto de momentos Multicanal y sus invariantes para rotación, escala y traslación para la representación y el reconocimiento de imágenes en color. Los momentos Multicanal propuestos se basan en los polinomios radiales sustituidos de Chebyshev.
- Singh y Singh [14] presentan un marco para la fusión de los descriptores geoméricamente invariantes que representan el color, la forma y la textura para el reconocimiento de objetos de color utilizando un enfoque de aprendizaje de núcleos múltiples. Para describir la textura de las imágenes en color, utilizan un descriptor basado en los momentos de Zernike del gradiente de las imágenes en color. Para las características de la forma, utilizan momentos Multicanal y de igual manera utilizan histogramas de color como descriptores RGB.
- Camacho Bello et. al [15] presenta un análisis de los momentos cuaternión genéricos Jacobi-Fourier (QGJFM) para la descripción de la imagen en color. Se utiliza un enfoque recursivo durante el cálculo de los polinomios radiales genéricos de Jacobi y píxeles polares para aumentar la precisión numérica en el cálculo. Presentan experimentos y resultados sobre la reconstrucción de imágenes en color.

Bibliografía

- [1] Gonzalez, Rafael C. "Digital Image Processing/Richard E."Woods. Interscience, NY (2001).
- [2] Hosny, Khalid M., and Mohamed M. Darwish. "Invariant color images representation using accurate quaternion Legendre–Fourier moments."Pattern Analysis and Applications (2018): 1-18.
- [3] Guo, Li-Qiang, and Ming Zhu. "Quaternion Fourier–Mellin moments for color images."Pattern Recognition 44.2 (2011): 187-195
- [4] Chen, B. J., et al. "Quaternion Zernike moments and their invariants for color image analysis and object recognition."Signal processing 92.2 (2012): 308-318.
- [5] Shao, Zhuhong, et al. "Quaternion Bessel–Fourier moments and their invariant descriptors for object reconstruction and recognition."Pattern Recognition 47.2 (2014): 603-611.
- [6] Guo, Liqiang, Ming Dai, and Ming Zhu. "Quaternion moment and its invariants for color object classification."Information Sciences 273 (2014): 132-143.
- [7] Chen, Beijing, et al. "Color image analysis by quaternion-type moments."Journal of mathematical imaging and vision 51.1 (2015): 124-144.
- [8] Zhu, H. Q., Qian Li, and Qian Liu. "Quaternion discrete Tchebichef moments and their applications."Int. J. Signal Process. Image Process. Pattern Recogn 7.6 (2014): 149-162.
- [9] Chen, Beijing, et al. "Quaternion-type moments combining both color and depth information for RGB-D object recognition."Pattern Recognition (ICPR), 2016 23rd International Conference on. IEEE, 2016.
- [10] Yang, Hong-ying, et al. "Quaternion exponent moments and their invariants for color image."Fundamenta Informaticae 145.2 (2016): 189-205.

BIBLIOGRAFÍA

- [11] Chen, Beijing, et al. "Quaternion pseudo-Zernike moments combining both of RGB information and depth information for color image splicing detection." *Journal of Visual Communication and Image Representation* 49 (2017): 283-290.
- [12] Singh, Chandan, and Jaspreet Singh. "Multi-channel versus quaternion orthogonal rotation invariant moments for color image representation." *Digital Signal Processing* 78 (2018): 376-392.
- [13] Hosny, Khalid M. y Mohamed M. Darwish. "Nuevo conjunto de momentos ortogonales multicanal para la representación y el reconocimiento de imágenes en color". *Reconocimiento de patrones* 88 (2019): 153-173.
- [14] Singh, Chandan y Jaspreet Singh. "Características de color, forma y textura geoméricamente invariantes para el reconocimiento de objetos utilizando el enfoque de clasificación de aprendizaje de múltiples núcleos". *Ciencias de la información* 484 (2019): 135-152.
- [15] Camacho-Bello, César, et al. "Reconstruction of color biomedical images by means of quaternion generic Jacobi-Fourier moments in the framework of polar pixels." *Journal of Medical Imaging* 3.1 (2016): 014004.

Capítulo 2

Teoría de momentos

2.1. Introducción

Los momentos calculados a partir de una función imagen discreta son cantidades escalares, que generalmente representan las características globales de la forma de la imagen, y proporcionan una gran cantidad de información sobre las características geométricas de la imagen. La capacidad que poseen los momentos de representar las características de una imagen ha sido ampliamente utilizada en diversas áreas de Visión por Computadora y Robótica [1], tienen diversas aplicaciones tales como: reconocimiento de patrones [2], análisis de imagen [3], marcas de agua digitales [4], compresión de imágenes [5], reconocimiento de caracteres [6], detección de bordes [7].

Los momentos M_{nm}^f de una imagen $f(x, y)$ donde n y m son enteros no negativos y $k = n + m$ es el orden de los momentos, se definen como [8]:

$$M_{nm}^f = \iint_D P_{nm}(x, y) f(x, y) dx dy \quad (2.1)$$

donde $P_{nm}(x, y)$ es una base polinomial o kernel de los momentos y D es el dominio. Dependiendo de la base polinomial utilizada, se formaran distintas familias de momentos, tales como los momentos geométricos [9], los momentos complejos, los momentos ortogonales discretos y los momentos ortogonales continuos [10].

2.2. Momentos geométricos

Los momentos geométricos poseen una función kernel definida como un producto de las coordenadas cartesianas que representan la posición de los pixeles. La opción más común, es una base de potencia estándar, definidos de la siguiente manera [9]

2. Teoría de momentos

$$m_{pq} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x^p y^q f(x, y) dx dy, \quad (2.2)$$

Por otra parte, los momentos geométricos de una función imagen discreta $f(x, y)$ de tamaño $M \times N$ se definen de la forma [8],

$$m_{pq} = \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} x^p y^q f(x, y). \quad (2.3)$$

donde $p = 0, 1, 2, \dots$ y $q = 0, 1, 2, \dots$ son enteros. Los momentos geométricos de diferentes órdenes representan diferentes características espaciales. Por ejemplo, m_{00} es el área de una imagen binaria, los momentos de primer orden m_{10}/m_{00} y m_{01}/m_{00} definen el centroide de la imagen. Los momentos de segundo orden m_{20} y m_{02} describen la distribución de intensidad de la imagen con respecto a los ejes de coordenadas, en mecánica se les conoce como momentos de inercia. Por otra parte, si $f(x, y)$ se considera como una función de densidad probabilística, es decir, sus valores se normalizan de tal manera que $m_{00} = 1$, se pudieran considerar como m_{10} y m_{01} como los valores medios. En caso de que fueran cero, m_{20} y m_{02} serían las proyecciones horizontal y vertical, y m_{11} sería la covarianza entre ellos. De esta manera, los momentos de segundo orden caracterizan el tamaño y orientación de la imagen. Considerando sólo hasta los momentos de segundo orden, la imagen original es completamente equivalente a una elipse de intensidad constante que tiene un tamaño y una orientación centrados en el centroide de la imagen, siendo sus ejes principales, independientes de la orientación de los ejes de coordenadas originales (x, y) ; donde ϕ es el ángulo de rotación entre los ejes de coordenadas originales (x, y) y los ejes principales (x', y') , definidos como [10],

$$\phi = (1/2) \tan^{-1} \left(\frac{2m_{11}}{m_{20} - m_{02}} \right) \quad (2.4)$$

La Fig. 2.1 muestra el ángulo de rotación ϕ y los ejes principales (x', y') . Mediante los momentos geométricos se puede construir una base de momentos que sean invariantes a la escala, traslación y rotación. Los momentos calculados para describir las características de un objeto deben ser capaces de identificar el mismo objeto con un tamaño u orientación distinta [9].

La invariancia a la traslación se logra haciendo que el centroide del objeto coincida con el origen del sistema de coordenadas o viceversa. Los momentos centrales de orden $(p + q)$, pueden ser expresados en terminos de los momentos geométricos como

2. Teoría de momentos

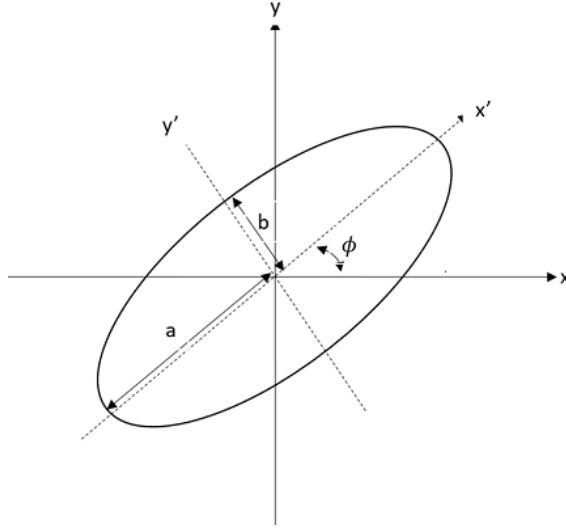


Figura 2.1: Elipse: ϕ es el ángulo de rotación entre los ejes (x, y) y los ejes principales (x', y') .

$$\mu_{pq} = \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} (x - \bar{x})^p (y - \bar{y})^q f(x, y) \quad (2.5)$$

para $p = 0, 1, 2, \dots$ y $q = 0, 1, 2, \dots$ donde las coordenadas del centroide de la imagen son

$$\bar{x} = \frac{m_{10}}{m_{00}} \text{ y } \bar{y} = \frac{m_{01}}{m_{00}}. \quad (2.6)$$

Los momentos centrales normalizados son invariantes en la traslación, se denotan por η_{pq} , y se definen como

$$\eta_{pq} = \frac{\mu_{pq}}{\mu_{pq}^{\gamma}}, \quad (2.7)$$

donde

$$\gamma = \frac{p+q}{2} + 1. \quad (2.8)$$

A partir de los momentos normalizados centrales η_{pq} , Hu en 1962 propuso invariantes derivados de cantidades algebraicas y aplicados a la función de generación de momento bajo una transformación de rotación. Consisten en grupos de expresiones de momento centralizadas no lineales. El resultado es un conjunto de momentos invariantes, que se pueden utilizar para la identificación de patrones con cambios de escala, posición y rotación [11]. Se calculan a partir de las siguientes fórmulas:

2. Teoría de momentos

$$\begin{aligned}
\phi_1 &= \eta_{20} + \eta_{02} \\
\phi_2 &= (\eta_{20} - \eta_{02})^2 + 4\eta_{11}^2 \\
\phi_3 &= (\eta_{30} - 3\eta_{12})^2 + (3\eta_{21} - \eta_{03})^2 \\
\phi_4 &= (\eta_{30} + \eta_{12})^2 + (\eta_{21} + \eta_{03})^2 \\
\phi_5 &= (\eta_{30} - 3\eta_{12})(\eta_{30} + \eta_{12})[(\eta_{30} + \eta_{12})^2 \\
&\quad - 3(\eta_{21} + \eta_{03})^2] + 3(\eta_{21} - \eta_{03})(\eta_{21} + \eta_{03}) \\
&\quad [3(\eta_{30} + \eta_{12})^2 - (\eta_{21} - \eta_{03})^2] \\
\phi_6 &= (\eta_{20} - \eta_{02})[(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2] \\
&\quad + 4\eta_{11}(\eta_{30} + \eta_{12})(\eta_{21} + \eta_{03}) \\
\phi_7 &= (3\eta_{21} - \eta_{03})(\eta_{30} + \eta_{12})[(\eta_{30} + \eta_{12})^2 \\
&\quad - 3(\eta_{21} - \eta_{03})^2] + (3\eta_{12} - \eta_{30})(\eta_{21} + \eta_{03}) \\
&\quad [3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2].
\end{aligned} \tag{2.9}$$

2.3. Momentos complejos

Los momentos geométricos y los momentos complejos llevan la misma cantidad de información. El término ‘momentos complejos’ se utiliza para abarcar todas las clases de funciones de momento que tienen núcleos complejos, se definen como [8]:

$$c_{pq} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x + jy)^p (x + jy)^q f(x, y) dx dy. \tag{2.10}$$

donde $j = \sqrt{-1}$.

Los momentos complejos también pueden expresarse en coordenadas polares (r, θ) ;

$$\begin{aligned}
x &= r \cos \theta \\
y &= r \sin \theta \\
r &= \sqrt{x^2 + y^2} \\
\theta &= \arctan(y/x)
\end{aligned}$$

$$c_{pq} = \int_0^{\infty} \int_0^{2\pi} r^{p+q+1} \exp^{j(p-q)\theta} f(r, \theta) dr d\theta. \tag{2.11}$$

La definición de los momentos complejos en coordenadas polares facilita la derivación de sus invariantes a la rotación. Sea $f'(r, \theta)$ una versión rotada de $f(r, \theta + \alpha)$ donde α es el ángulo de rotación, por lo tanto

$$c'_{pq} = \exp^{-j(p-q)\alpha} c_{pq}. \tag{2.12}$$

2. Teoría de momentos

La construcción de los invariantes de rotación se basa en la cancelación de la fase $(p-q)\alpha$, esto se logra mediante la multiplicación de potencias de momentos apropiados. La construcción de los momentos complejos invariantes a la rotación, está dada por

$$\begin{aligned}\Phi(k, 0) &= c_{k,0}c_{q_0 p_0}^k, \\ \Phi(k-1, 1) &= c_{k-1,1}c_{q_0 p_0}^{k-2},\end{aligned}\tag{2.13}$$

donde p_0 y q_0 son índices arbitrarios de tal manera que $p_0 + q_0 < k$, $p_0 - q_0 = 1$ y $c_{p_0, q_0} \neq 0$. Por otra parte, los momentos invariantes de Hu pueden ser representados a partir de los momentos complejos ,

$$\begin{aligned}\phi_1 &= c_{11} = \Phi(1, 1), \\ \phi_2 &= c_{20}c_{02} = \frac{|\Phi(2, 0)|^2}{|\Phi(2, 1)|^2}, \\ \phi_3 &= c_{30}c_{03} = \frac{|\Phi(3, 0)|^2}{|\Phi(2, 1)|^3}, \\ \phi_4 &= c_{21}c_{12} = \Phi(2, 1), \\ \phi_5 &= \text{Re } c_{30}c_{12}^3 = \text{Re}(\Phi(3, 0)), \\ \phi_6 &= \text{Re } c_{20}c_{12}^3 = \text{Re}(\Phi(2, 0)), \\ \phi_7 &= \text{Im } c_{30}c_{12}^3 = \text{Im}(\Phi(3, 0)),\end{aligned}\tag{2.14}$$

Los momentos complejos sólo son invariantes a la rotación, no son invariantes a la traslación y no están normalizados en escala.

2.4. Momentos ortogonales

Los momentos ortogonales fueron introducidos por Teague en 1980. Tienen la capacidad de capturar las características de una imagen con mínima redundancia. El kernel de los momentos tiene una base de polinomios ortogonales, los cuales pueden ser calculados mediante relaciones de recurrencia, por lo que su implementación numérica es estable y rápida. Los momentos ortogonales son frecuentemente mencionados en la literatura por su capacidad de reconstrucción de la imagen, sin embargo no son aptos para la compresión de información [8]. Los momentos ortogonales se dividen en dos grupos: los momentos ortogonales definidos en un rectángulo y los momentos ortogonales radiales definidos en una circunferencia .

Los momentos ortogonales definidos en un rectángulo de una función imagen $f(x, y)$ de tamaño $N \times M$ están dados como [8],

2. Teoría de momentos

$$q_{n,m} = \iint_D f(x,y)P_{n,m}(x,y)dx dy, \quad (2.15)$$

donde $P_{n,m}(x,y)$ es el kernel, el cual consiste en dos polinomios ortogonales $A_n(x)$ y $A_m(y)$, expresados como

$$P_{n,m}(x,y) = A_n(x)A_m(y), \quad (2.16)$$

y sus elementos cumplen la condición de ortogonalidad

$$\iint_D P_{n,m}(x,y)P_{p,q}(x,y) = \delta_{n,p}\delta_{m,q}, \quad (2.17)$$

donde $\delta_{n,p}$ es el símbolo de Kronecker,

$$\delta_{n,p} = \begin{cases} 0, & n \neq p \\ 1, & n = p \end{cases}$$

Por otra parte, los momentos ortogonales radiales usan un sistema de coordenadas polares (r, θ) , lo cual facilita la derivación de invariantes de rotación de cualquier orden [2]. La representación general para los momentos esta dada por:

$$M_{nm}^p = \int_0^{2\pi} \int_0^1 P_{nm}(r, \theta) f(r, \theta) r dr d\theta \quad (2.18)$$

donde $f(r, \theta)$ está representada en coordenadas polares y $P_{nm}(r, \theta)$ es la función kernel, la cual consiste en el producto de dos funciones: $A_n(r)$ es un polinomios ortogonal en la coordenada radial y $exp(jm\theta)$ es la exponencial compleja de Fourier en la coordenada angular. La función $P_{n,m}(r, \theta)$ es expresada como,

$$P_{n,m}(r, \theta) = A_n(r)exp(jm\theta) \quad (2.19)$$

donde n y m son enteros. La función kernel es ortogonal dentro de un círculo unitario, $0 \leq r \leq 1$, $0 \leq \theta \leq 2\pi$ y la propiedad de ortogonalidad queda expresada como,

$$\int_0^{2\pi} \int_0^1 P_{n,m}(r, \theta) P_{k,l}(r, \theta) r dr d\theta = \delta_{n,k} \delta_{m,l}. \quad (2.20)$$

Los invariantes de momento se han usado ampliamente a una variedad de aplicaciones debido a sus características invariantes a la traslación, escala y rotación de imágenes. En coordenadas polares, la rotación es una traslación del argumento angular. La principal ventaja con los momentos radiales es que la rotación de la imagen

2. Teoría de momentos

se pueden traducir directamente al momento correspondiente. La transformación, y las expresiones para invariantes de rotación se pueden derivar fácilmente. Considerando una imagen $f(r, \theta - \gamma)$ que es rotada γ grados, esto da como resultado los momentos $\phi_{n,m}^{(\theta-\gamma)}$, los cuales están relacionados con $\phi_{n,m}$ de la siguiente manera

$$\phi_{n,m}^{(\theta-\gamma)} = \phi_{n,m} \exp(-jm\gamma) \quad (2.21)$$

Por otra parte, la invariancia a la escala se obtiene al definir la imagen dentro de un círculo de radio unitario donde la imagen es remapeada independientemente de sus dimensiones. Finalmente la invariancia a la traslación de una imagen no se puede representar directamente como una transformación lineal de las coordenadas polares, ésta se logra indirectamente por medio del cálculo del punto centroide de la imagen usando momentos geométricos y luego desplazando el origen del sistema de coordenadas polares al punto centroide [8].

2.5. Cómputo exacto de los momentos radiales

Los momentos como descriptores de la imagen, deben tener la capacidad de representar y reconstruir, sin contener información redundante, lo que implica que el conjunto de funciones del núcleo sean ortogonales. En el cómputo de los momentos existe una pérdida de precisión debido al error de integración numérica que afecta la capacidad de análisis y reconstrucción de la imagen. Sin la precisión apropiada, algunas propiedades atractivas de los momentos quedan comprometidas. Se debe prestar atención a la estabilidad numérica debido a los errores de aproximación discreta de las integrales continuas y de la transformación del sistema de coordenadas de la imagen [12]. Se han desarrollado algoritmos para incrementar la exactitud de los momentos, los más representativos se mencionan a continuación:

- Zakaria et al.[13] proponen el método Delta, el cual calcula los momentos de imágenes contiguas mediante el uso de la contribución de cada línea en lugar del píxel individual, reduciendo en gran medida la cantidad de datos y procesamiento necesario para identificar un objeto.
- Yang et al.[14] plantean un algoritmo basado en el teorema de Green discreto, que evalúa la integral doble sobre el objeto mediante una integración única a lo largo del límite del objeto.
- Flusser[15] propone un método basado en la descomposición del objeto en segmentos. El método funciona para formas binarias arbitrarias.

2. Teoría de momentos

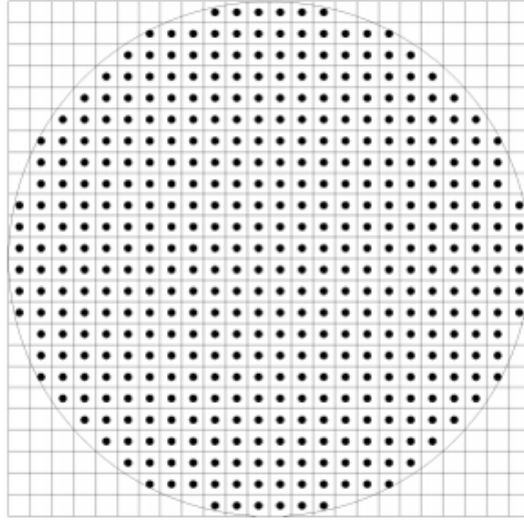


Figura 2.2: Aproximación de píxeles cartesianos a una región circular.

- Liao et al.[16] proponen nuevas técnicas para aumentar la precisión y la eficiencia de los descriptores de momento usando una regla extendida de Simpson.
- Hosny [17] propone un método para calcular los momentos utilizando la integración matemática de los términos monomiales sobre píxeles de imágenes digitales.

Por otra parte, el error geométrico se debe a la naturaleza circular del dominio de los momentos, éstos están definidos en coordenadas polares y son aplicados a una imagen representada en píxeles, es decir, en coordenadas cartesianas. No obstante, la función imagen definida sobre el disco de radio unitario, sólo toma en cuenta aquellos píxeles cuyos centros caen completamente dentro del círculo, de otra forma son descartados. En la Fig. 2.2 se muestran algunos píxeles que no están completamente dentro del círculo unitario, así como algunas partes del círculo no están cubiertas por píxeles [17]. El error geométrico se debe al hecho de que el área total cubierta por todos los píxeles cuadrados involucrados en el cálculo de los momentos no están exactamente en el disco unitario, como lo ilustra el borde irregular en la Fig. 2.2. Para solucionar este problema, Xin et al.[18] presentan un algoritmo para el cálculo numérico de alta precisión de los momentos de Zernike. El algoritmo se basa en un esquema de píxeles polares, que no manifiesta el error geométrico y el error de integración numérica que son propios de los métodos basados en coordenadas cartesianas. Esto produce un mejoramiento en la precisión de los momentos en cuanto a sus propiedades de reconstrucción e invariancia. Además, proponen utilizar un método analítico en lugar

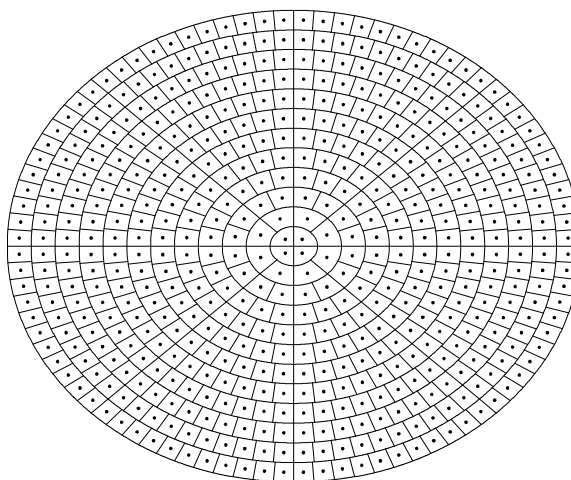


Figura 2.3: Esquema de píxeles polares propuestos por Xin et al.

de aproximaciones numéricas para la integración de funciones básicas en píxeles, y adoptar un sistema de coordenadas polares para facilitar el uso de momentos radiales. Este sistema de coordenadas polares cumple los siguientes criterios:

- Todas las áreas del sector deben ser iguales, correspondiente al hecho de que en una imagen cartesiana todos los píxeles tienen un tamaño uniforme.
- El tamaño de los píxeles polares no debe ser mayor que el de los cartesianos, de modo que se pueda mantener la resolución de imagen necesaria.
- Los píxeles polares deben ser tan “cuadrados” como sea posible, es decir, las longitudes de las fronteras de un sector debe estar lo suficientemente cerca, para que la distorsión de la imagen se pueda mantener en un nivel bajo por la conversión del sistema de coordenadas.
- Los píxeles polares deben organizarse lo más regularmente posible, para facilitar su almacenamiento y cálculo

Con las especificaciones anteriores el esquema de píxeles propuesto se ilustra en la Fig. 2.3, en donde cada píxel polar está construido bajo las siguientes restricciones:

- El disco de radio unidad se divide uniformemente a lo largo de la dirección radial en U secciones, con discos separados localizados en $\{(k/U), k = 1, \dots, U\}$

2. Teoría de momentos

- La i -ésima sección radial se divide igualmente en $(2k - 1)V$ sectores por radio comenzando desde el origen, con ángulos $\{(i - 1)(2\pi)/((2k - 1)V), i = 1, \dots, (2k - 1)V\}$, y V es el número de sectores

En la literatura se recomienda utilizar el valor de $V = 4$ en el anillo central y $(N/2) \leq U \leq N$ para una imagen de tamaño $N \times N$. Dado que las imágenes digitales se definen generalmente en coordenadas cartesianas, las ubicaciones de los píxeles entre ambas configuraciones no coinciden, por lo que se realiza una interpolación de la imagen. La interpolación es una técnica para el remuestreo de datos discretos en donde se calculan los valores de una función en aquellos puntos donde no existen relaciones exactas entre el valor de la función y el punto que lo representa. En el trabajo de Xin et al.[18] calculan el valor del pixel de la imagen en un píxel polar mediante la interpolación bicúbica propuesta por MacKay [19].

El valor del pixel de la imagen para un pixel polar Ω_{uv} puede ser calculado a través de la función “*interp2*” con el método “*cubic*” de Matlab, la cual devuelve valores interpolados de una función de dos variables en puntos específicos. En la Fig 2.4 se muestra la interpolación de la imagen "Baboon" de tamaño 102×102 usando el método, donde el valor interpolado en un punto de consulta se basa en una interpolación cúbica de los valores en puntos de cuadrícula vecinos en cada dimensión respectiva. [20].

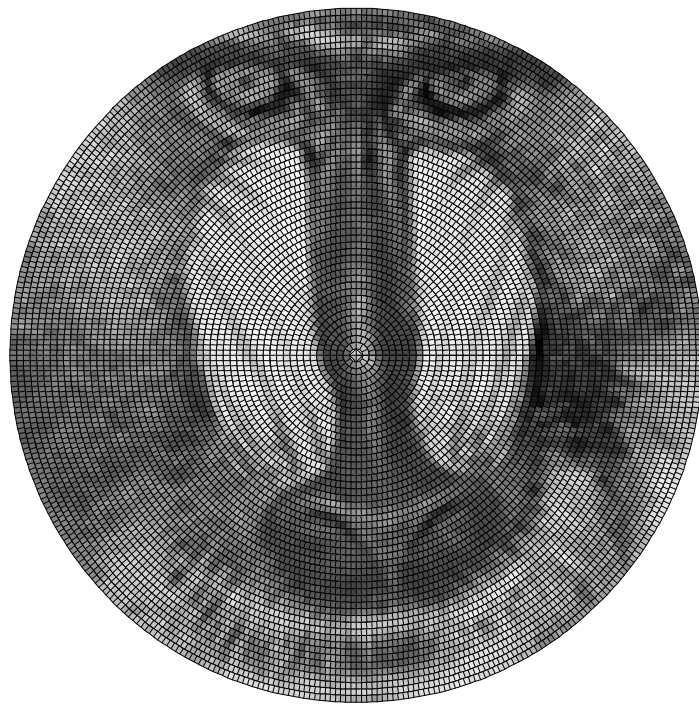


Figura 2.4: Interpolación de la imagen "Baboon" de tamaño 102×102 usando el método "cubic" a través de la función "interp2" de Matlab.

Bibliografía

- [1] Mukundan, Ramakrishnan, and K. R. Ramakrishnan. "Moment functions in image analysis-theory and applications. World Scientific", 1998.
- [2] Khotanzad, Alireza, and Yaw Hua Hong. "Invariant image recognition by Zernike moments."IEEE Transactions on pattern analysis and machine intelligence 12.5 (1990): 489-497
- [3] Teh, C-H., and Roland T. Chin. "On image analysis by the methods of moments."Computer Vision and Pattern Recognition, 1988. Proceedings CVPR'88., Computer Society Conference on. IEEE, 1988.
- [4] George A Papakostas."Over 50 years of image moments and moments invariants" In George A Papakostas, editor, Moments Moment Invariants-Theor App. Pages 3-32. Science Gate Publishing, Xanthi, Greece. (2014)
- [5] F. Ernawan N.A. Abu, N. Suryana, "An optical Tchebichef moment quantization using psychovisual threshold for image compression. Adv. Sci. Lett 20 (2014)
- [6] R. G. Keys, "Cubic convolution interpolation for digital image processing",IEEE Trans. Acoust. Speech Signal Process. 29, 1153-1160 (1981).
- [7] J. S. Rivera.C.J. Camacho-Bello "Color Image Reconstruction by Discrete Orthogonal Moment". Journal of Data Analysis and Information Processing , 5, 156-166 (2017).
- [8] Flusser, Jan, Barbara Zitova, and Tomas Suk. Moments and moment invariants in pattern recognition. John Wiley & Sons, 2009.
- [9] Hu, Ming-Kuei. "Visual pattern recognition by moment invariants."IRE transactions on information theory 8.2 (1962): 179-187
- [10] Teague, Michael Reed. "Image analysis via the general theory of moments."JOSA 70.8 (1980): 920-930

BIBLIOGRAFÍA

- [11] Hu, Ming-Kuei. "Visual pattern recognition by moment invariants." *IRE transactions on information theory* 8.2 (1962): 179-187.
- [12] Sáez-Landete, José. Comments on "fast computation of jacobi-Fourier moments for invariant image recognition". *Pattern Recognition* 67 (2017): 16-22.
- [13] Zakaria, M. F., et al. "Fast algorithm for the computation of moment invariants." *Pattern Recognition* 20.6 (1987): 639-643.
- [14] Yang, Luren, and Fritz Albreghsen. "Fast and exact computation of Cartesian geometric moments using discrete Green's theorem." *Pattern Recognition* 29.7 (1996): 1061-1073.
- [15] Flusser, Jan. "Fast calculation of geometric moments of binary images." *Pattern Recognition and Medical Computer Vision* (1998): 265-274.
- [16] Liao, Simon Xinmeng, and Mirosław Pawlak. "On image analysis by moments." *IEEE Transactions on Pattern Analysis and Machine Intelligence* 18.3 (1996): 254-266.
- [17] Hosny, Khalid M. "Exact and fast computation of geometric moments for gray level images." *Applied Mathematics and Computation* 189.2 (2007): 1214-1222.
- [18] Xin, Yongqing, Mirosław Pawlak, and Simon Liao. "Accurate computation of Zernike moments in polar coordinates." *IEEE Transactions on Image Processing* 16.2 (2007): 581-587.
- [19] D. MacKay, "Good error-correcting codes based on very sparse matrices," *IEEE Trans. Inf. Theory*, vol. 45, no. 2, pp. 399–431, Mar. 1999.
- [20] The MathWorks, Inc. "Interpolation for 2-D gridded data in meshgrid format". Recuperado de: <https://la.mathworks.com/help/matlab/ref/interp2.html>

Capítulo 3

Momentos radiales ortogonales

3.1. Introducción

El kernel de los momentos ortogonales se define a partir de los polinomios continuos ortogonales o polinomios discretos ortogonales. En los continuos tenemos: los momentos de Zernike [1], los momentos de pseudo-Zernike [2], los momentos de Fourier–Mellin [3], los momentos Chebyshev–Fourier [4], los momentos generalizados pseudo-Jacobi-Fourier [5], los momentos Radiales desplazados de Legendre [6], los momentos de Bessel-Fourier [7], los momentos Jacobi Fourier generalizados [8], los momentos de transformación exponencial del complejo polar [9], y los momentos de Chebyshev-armónicos Fourier [10]. Por otra parte, para los polinomios discretos Mukundan [11] propone los momentos radiales de Tchebichef utilizando los polinomios ortogonales discretos de Tchebichef como kernel, y con una forma radial-polar similar a la de los momentos de Zernike.

La primera definición de los momentos radiales fue dada por Tegue en 1980, él propuso los momentos de Zernike los cuales están definidos como [12],

$$A_{nl} = \frac{n+1}{\pi} \int_0^1 \int_0^{2\pi} V_{nl}^*(r, \theta) f(r, \theta) r dr d\theta, \quad (3.1)$$

donde los polinomios de enésimo grado con repetición l se definen como,

$$V_{nl}(r, \theta) = R_{nl}(r) e^{il\theta} \quad n = 0, 1, 2, \dots, l = -n, -n+2, \dots, n, \quad (3.2)$$

para $n - |l|$ siempre una cantidad par. La parte radial se define como,

$$R_{nl}(r) = \sum_{s=0}^{\frac{n-|l|}{2}} (-1)^s \frac{(n-s)!}{s! \left(\frac{n+|l|}{2} - s\right)! \left(\frac{n-|l|}{2} - s\right)!} r^{n-2s} = \sum_{k=|l|, |l|+2, \dots}^n B_{nlk} r^k. \quad (3.3)$$

3. Momentos radiales ortogonales

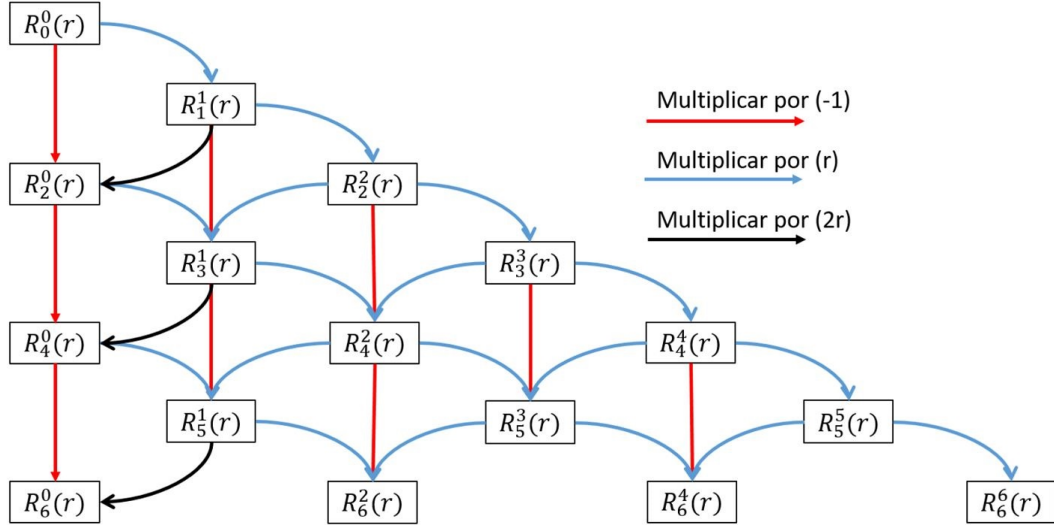


Figura 3.1: Diagrama de bloques del método propuesto por Honarvar et al. [13] para calcular los polinomios radiales de Zernike $R_n^m(r)$.

La función radial es simétrica con respecto a la repetición, es decir, $R_{n,-l}(r) = R_{n,l}(r)$. Además, los polinomios de Zernike $V_{nl}(r, \theta)$ satisfacen la condición de ortogonalidad

$$\int_0^{2\pi} \int_0^1 V_{nl}^*(r, \theta) V_{mk}(r, \theta) r dr d\theta = \frac{\pi}{n+1} \delta_{mn} \delta_{kl}. \quad (3.4)$$

Por otra parte, debido a la complejidad del cálculo de los polinomios de Zernike en la componente radial, dada por la Ec. 3.3, se han desarrollado varios métodos para agilizar su cómputo; estos métodos usan relaciones de recurrencia que dependen de n y m . Honarvar et al. [13] proponen una fórmula de recurrencia para el cálculo rápido de la parte radial de los polinomios de Zernike, utilizando una relación entre polinomios radiales y polinomios de Tchebichef de segundo clase. A diferencia de las propuestas anteriores, la relación de recurrencia derivada no depende del orden n o del orden m para reducir la complejidad de cálculo. La fórmula de recurrencia para el cálculo rápido de la parte radial propuesta por Honarvar et al, está dada por,

$$R_n^m(r) = r[R_{n-1}^{|m-1|}(r) + R_{n-1}^{m+1}(r)] - R_{n-2}^m(r), \quad (3.5)$$

con el polinomio radial de orden cero $R_0^0(r) = 1$, se puede calcular todos los polinomios radiales $R_n^m(r)$ con enteros ≥ 0 tal que $n-m$ es par y no negativo, y $R_n^m(r) \equiv 0$ cuando $n < m$. El diagrama de bloques de la relación de recurrencia de la Ec. 3.5 se muestra en la Fig. 3.1 y en la Fig. 3.2 se muestran los primeros 5 polinomios.

3. Momentos radiales ortogonales

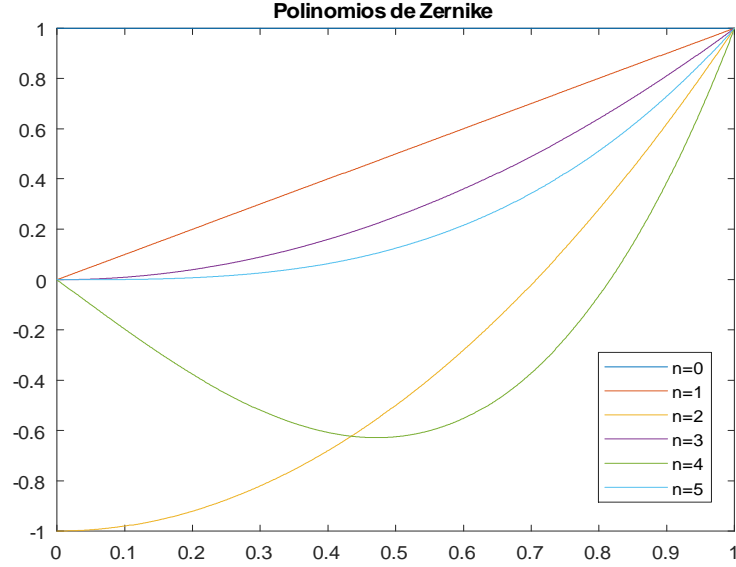


Figura 3.2: Gráficas de los polinomios radiales de Zernike.

Para los momentos Multicanal de Zernike (MMZ) el número total de momentos esta dado por,

$$\begin{aligned} & \left(\frac{2+Z_{\text{máx}}}{2} \right)^2 && Z_{\text{máx}} \text{ par,} \\ & \left(\frac{1+Z_{\text{máx}}}{2} \right) \left(\frac{3+Z_{\text{máx}}}{2} \right) && Z_{\text{máx}} \text{ impar,} \end{aligned} \quad (3.6)$$

donde $Z_{\text{máx}}$ se refiere al orden máximo. En este trabajo de tesis se examinan los momentos de Zernike considerando la aportación que estos momentos han realizado al análisis de imágenes.

La invariancia a transformaciones con respecto a la escala, traslación y rotación es muy importante en las aplicaciones de reconocimiento de patrones y es objeto de investigaciones recientes. La idea básica es describir los objetos mediante un conjunto de cantidades medibles llamadas invariantes, los cuales son insensibles a deformaciones particulares y proporcionan suficiente información para poder distinguir objetos pertenecientes a diferentes clases. Por otra parte, para evaluar la efectividad de los invariantes se utiliza el error cuadrático medio MSE, el cual es una medida cuantitativa que refleja la precisión del invariante de rotación, escala y traslación, se define como[14],

$$MSE = \frac{1}{L_{Total}} \sum_{p=0}^{\text{máx}} \sum_{q=0}^{\text{máx}} (|M_{pq}(f)| - |M_{pq}(f')|)^2, \quad (3.7)$$

3. Momentos radiales ortogonales

donde L_{Total} es el número total de momentos y $|M_{pq}(f)|$, $|M_{pq}(f')|$ son respectivamente las magnitudes de los momentos antes y después de la rotación, escala o traslación de la función imagen f .

Para implementar los invariantes, se utiliza la base de datos “Columbia Object Image Library (COIL-100)”, la cual es una base de datos pública utilizada para el propósito experimental. COIL- 100 contiene 7200 imágenes en color de 100 objetos diversos. La Fig. 3.3 se muestran algunas imágenes de la base de datos COIL-100 [15].



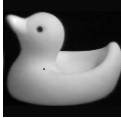

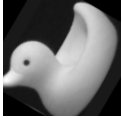

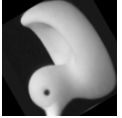
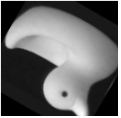
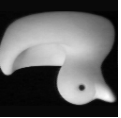
Figura 3.3: Algunas imágenes de la base de datos Columbia Object Library (COIL 100).

3.2. Invariancia a la rotación

Una de las propiedades más importantes de los momentos ortogonales es la invariancia a la traslación y escala. En el caso del invariante a la rotación, el comportamiento los momentos de Zernike es similar al de los momentos complejos. Si una imagen se rota un ángulo α , el valor de los momentos de Zernike cambia de la

3. Momentos radiales ortogonales

Tabla 3.1: MSE para evaluar la efectividad de la invariante de rotación.

Imagen original	$\alpha = 30^\circ$	$\alpha = 60^\circ$	$\alpha = 90^\circ$	$\alpha = 120^\circ$	$\alpha = 150^\circ$	$\alpha = 180^\circ$
						
MSE	1.2029e-04	1.2328e-04	1.5817e-29	1.2029e-04	1.2328e-04	2.2277e-29

siguiente forma,

$$A'_{nl} = A_{nl}e^{il\alpha}. \quad (3.8)$$

Por lo tanto, los invariantes de rotación se pueden construir tomando la magnitud $|A_{nl}|$. Para ejemplificar la invariancia a la rotación, se tomó el objeto 74_0, el cual fue rotado por 30° , 60° , 90° , 120° , 150° y 180° en el programa Matlab versión 9.1 R2016b con la función *imrotate*, usando el método *'bilinear'* y el cuadro delimitador con la opción *'crop'* para que la imagen de salida tenga el mismo tamaño que la imagen de entrada, como se muestra en la Fig. 3.4. También se calcularon los primeros diez momentos de Zernike para el canal rojo.

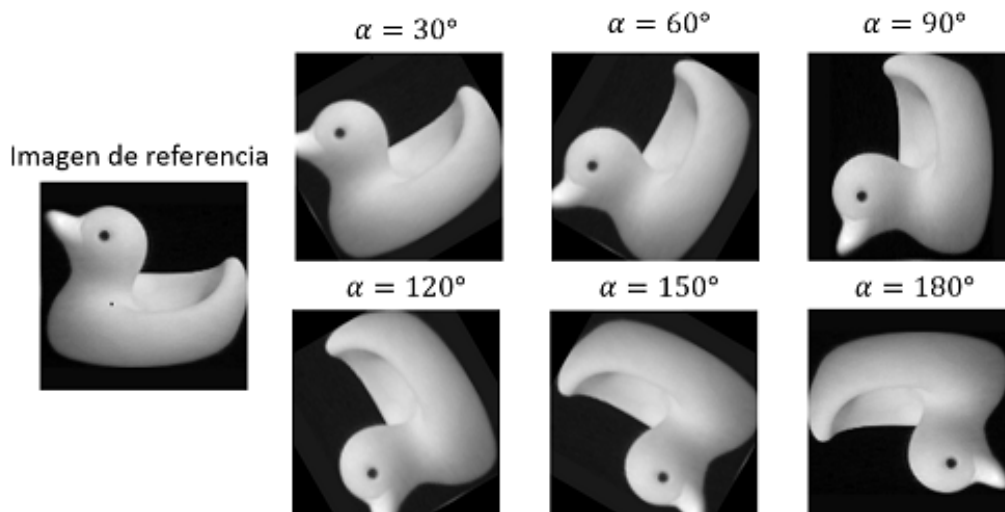


Figura 3.4: Imagen 74-0 de la base de datos COIL-100 con diferentes ángulos de rotación.

Los resultados obtenidos se muestran en las Tablas 3.1 y 3.2 donde se demuestra que los momentos de Zernike son invariantes a la rotación.

3. Momentos radiales ortogonales

Tabla 3.2: momentos de Zernike para los diferentes ordenes y desviación estándar de los momentos a diferentes ángulos.

n	m	Imagen original	$\alpha = 30^\circ$	$\alpha = 60^\circ$	$\alpha = 90^\circ$	$\alpha = 120^\circ$	$\alpha = 150^\circ$	$\alpha = 180^\circ$	σ
0	0	59.4402	59.4356	59.4361	59.4402	59.4356	59.4361	59.4402	2.3e-3
1	1	8.2196	8.2194	8.2184	8.2196	8.2194	8.2184	8.2196	5e-4
2	0	19.2657	19.2636	19.2636	19.2657	19.2636	19.2636	19.2657	1.1e-3
2	2	13.9928	13.9826	13.9816	13.9928	13.9826	13.9816	13.9928	5.7e-3
3	1	10.8174	10.8015	10.8009	10.8174	10.8015	10.8009	10.8174	8.6e-3
3	3	12.1590	12.1337	12.1340	12.1590	12.1337	12.1340	12.1590	1.134e-2
4	0	0.7490	0.7466	0.7472	0.7490	0.7466	0.7472	0.7490	1.2e-3
4	2	9.0815	9.0771	9.0786	9.0815	9.0771	9.0786	9.0815	2e-3
4	4	2.6722	2.6607	2.6600	2.6722	2.6607	2.6600	2.6722	6.3e-3
5	1	3.6798	3.6846	3.6854	3.6798	3.6846	3.6854	3.6798	2.8e-3

3.3. Invariancia a la escala

Para el caso de la invariancia a la escala, la imagen puede ser mapeada en un círculo de radio unitario de forma independiente a la escala de la imagen, este mapeo proporciona la invariancia al escalamiento. Para demostrar la invariancia a la escala, se tomó el objeto 52_0 con un tamaño de referencia de 128×128 al cual se le modificó el tamaño con la función *imresize* de MatLab R2016b, con el parámetro '*bilinear*'. En la Fig. 3.5 se muestra 7 imágenes de prueba.

Los resultados obtenidos se muestran en las Tablas 3.3 y 3.4 donde se demuestra que los momentos de Zernike son invariantes a la escala.

Tabla 3.3: Error cuadrático medio MSE para evaluar la efectividad de la invariante de escala.

Tamaño	64×64	96×96	128×128	160×160	192×192	224×224	256×256
MSE	0.2170	0.0240	0	0.0098	0.0269	0.0437	0.0589

3.4. Invariancia a la traslación

Para el invariante a la traslación Chong et al. [16] proponen un marco matemático para la derivación de invariantes a la traslación. Para derivar los momentos centrales de Zernike, primero se requiere establecer los momentos centrales radiales. En vista de la condición ($n - |l| = par$) en los momentos de Zernike, en los momentos radiales se consideran de igual manera. Los momentos centrales radiales de orden n con repetición l , \hat{D}_{nl} , se definen como

3. Momentos radiales ortogonales

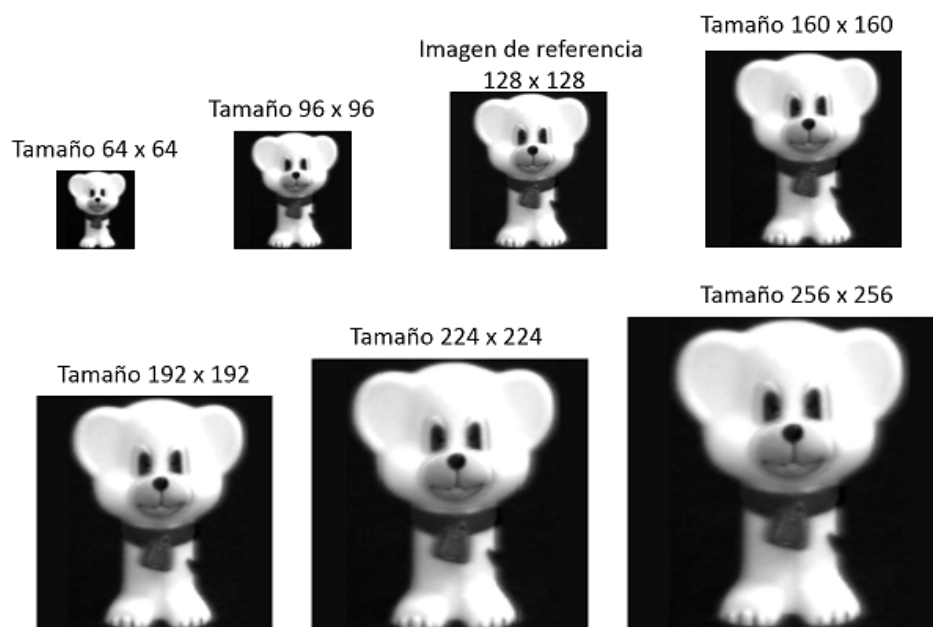


Figura 3.5: Objeto 52_0 de la base de datos COIL-100 a diferentes escalas.

Tabla 3.4: Resultado de la invariancia a la escala de los primeros diez momentos de Zernike.

n	m	64 × 64	96 × 96	128 × 128	160 × 160	192 × 192	224 × 224	256 × 256	σ
0	0	61.0536	60.7895	60.6671	60.5712	60.5050	60.4603	60.4290	0.22
1	1	18.6540	18.4887	18.4020	18.3584	18.3230	18.2986	18.2833	0.1316
2	0	3.3658	3.7664	3.9631	4.1077	4.1952	4.2561	4.3052	0.3333
2	2	30.3953	30.0780	29.9284	29.7956	29.7126	29.6567	29.6140	0.2777
3	1	7.25	7.4531	7.5666	7.6057	7.6399	7.6644	7.6799	0.1532
3	3	9.0326	9.0394	9.0476	9.0237	9.0216	9.0201	9.0131	0.0121
4	0	4.3686	4.7338	4.9190	5.0302	5.0990	5.1475	5.1864	0.2902
4	2	8.5601	9.1852	9.4964	9.6791	9.8025	9.8890	9.9530	0.4937
4	4	2.4489	2.3460	2.3119	2.2876	2.2719	2.2620	2.2558	0.680
5	1	2.1029	1.9173	1.8130	1.7936	1.7662	1.7467	1.7386	0.1306

3. Momentos radiales ortogonales

$$\hat{D}_{nl} = \int_x \int_y [(x - x_0) - \hat{j}(y - y_0)]^{(n+l)/2} [(x - x_0) + \hat{j}(y - y_0)]^{(n-l)/2} f(x, y) dx dy, \quad (3.9)$$

donde $(n - |l| = \text{par})$. Los centroides de las coordenadas x_0 y y_0 , son definidos respectivamente como

$$x_0 = \frac{\sum \sum x f(x, y)}{\sum \sum f(x, y)}, \quad y_0 = \frac{\sum \sum y f(x, y)}{\sum \sum f(x, y)}. \quad (3.10)$$

La ecuación 3.8 puede reescribirse como

$$\hat{D}_{nl} = \int_x \int_y (K - A)^u (K^* - A^*)^v f(x, y) dx dy, \quad (3.11)$$

donde $K = x - \hat{j}y$, $A = x_0 - \hat{j}y_0 = D_{11}/D_{00}$, $u = (p + q)/2$ y $v = (p - q)/2$. El kernel puede ser representado de la siguiente manera,

$$(K - A)^u = \sum_{m=0}^u \binom{u}{m} (K)^{u-m} (-A^*)^m, \quad (3.12)$$

y,

$$(K^* - A^*)^v = \sum_{n=0}^v \binom{v}{n} (K^*)^{v-n} (-A^*)^n, \quad (3.13)$$

por lo tanto, la Ec.(2.22) puede ser expresada de la siguiente manera,

$$\begin{aligned} \hat{D}_{nl} &= \sum_{m=0}^u \sum_{n=0}^v \binom{u}{m} \binom{v}{n} (-A^*)^m (-A^*)^n \\ &\int_x \int_y (K)^{u-m} (K^*)^{v-n} f(x, y) dx dy. \end{aligned} \quad (3.14)$$

Tomando la ecuación anterior, los momentos centrales radiales de orden p con repetición q se puede obtener como una serie de momentos radiales:

$$D_{pq} = \sum_{m=0}^u \sum_{n=0}^v \binom{u}{m} \binom{v}{n} (-A^*)^m (-A^*)^n D_{p'q'}, \quad (3.15)$$

donde $p' = (u - m) + (v - n)$ y $q' = (u - m) - (v - n)$. A partir de la Ec. 3.14 Chong et. al, derivan los invariantes a la traslación de los momentos radiales para $p - q = \text{par}$ como,

3. Momentos radiales ortogonales

$$\begin{aligned}
 & \text{Orden 0} \\
 \hat{D}_{00} &= D_{00} \\
 & \text{Orden 1} \\
 \hat{D}_{11} &= D_{11} - AD_{00} \\
 & \text{Orden 2} \\
 \hat{D}_{22} &= D_{22} - 2AD_{11} + A^2D_{00} \\
 \hat{D}_{20} &= D_{20} - A^*D_{11} - AD_{11}^* + AA^*D_{00} \\
 & \text{Orden 3} \\
 \hat{D}_{33} &= D_{33} - 3AD_{22} + 3A^2D_{11} - A^3D_{00} \\
 \hat{D}_{31} &= D_{31} - A^*D_{22} - 2AD_{20} + 2AA^*D_{11} \\
 & \quad + A^2D_{11}^* - A^2A^*D_{00} \\
 & \text{Orden 4} \\
 \hat{D}_{44} &= D_{44} - 4AD_{33} + 6A^2D_{22} - 4A^3D_{11} \\
 & \quad + A^4D_{00} \\
 \hat{D}_{42} &= D_{42} - A^*D_{33} - 3AD_{31} + 3AA^*D_{22} \\
 & \quad + 3A^2D_{20} - 3A^2A^*D_{11} - A^3D_{11}^* + A^3A^*D_{00} \\
 \hat{D}_{40} &= D_{40} - 2A^*D_{31} + A^{*2}D_{22} - 2AD_{31}^* + 4AA^*D_{20} \\
 & \quad - 2AA^{*2}D_{11} + A^2D_{22}^* - 2A^2A^*D_{11}^* + A^2A^{*2}D_{00} \tag{3.16}
 \end{aligned}$$

Posteriormente se derivan los invariantes a la traslación de los momentos de Zernike

3. Momentos radiales ortogonales

$$\begin{aligned}
& \text{Orden 0} \\
\hat{Z}_{00} &= Z_{00} \\
& \text{Orden 1} \\
\hat{Z}_{11} &= Z_{11} - 2AZ_{00} \\
& \text{Orden 2} \\
\hat{Z}_{22} &= Z_{22} - 3AZ_{11} + 3A^2Z_{00} \\
\hat{Z}_{20} &= Z_{20} - 3A^*Z_{11} - 3AZ_{11}^* + 6AA^*Z_{00} \\
& \text{Orden 3} \\
\hat{Z}_{33} &= Z_{33} - 4AZ_{22} + 6A^2Z_{11} - 4A^3Z_{00} \\
\hat{Z}_{31} &= Z_{31} - 4A^*Z_{22} - 4AZ_{20} + 12AA^*Z_{11} \\
& \quad - 4AZ_{00} \\
& \text{Orden 4} \\
\hat{Z}_{44} &= Z_{44} - 5AZ_{33} + 10A^2Z_{22} - 10A^3Z_{11} \\
& \quad + 5A^4Z_{00} \\
\hat{Z}_{42} &= Z_{42} - 5A^*Z_{33} - 5AZ_{31} + 20AA^*Z_{22} \\
& \quad + 10A^2Z_{20} - 5AZ_{11} - 30A^3Z_{11}^* + 15A^2Z_{00} \\
\hat{Z}_{40} &= Z_{40} - 5A^*Z_{31} - 5AZ_{31}^* + 10A^{*2}Z_{22} + 10A^2Z_{22}^* \\
& \quad + 20AA^*Z_{20} + 5AZ_{11}^* - 90A^2A^{*2}Z_{00} \tag{3.17}
\end{aligned}$$

Para verificar la Ec.3.16 se tomó el objeto 59_19 con un tamaño de 128×128 pixeles y se trasladó en varias posiciones dentro de una imagen de fondo negro con un tamaño de 300×300 pixeles como se muestra en la Fig. 3.6. Los resultados obtenidos utilizando los invariantes propuestas por Chong se muestran en la Tabla 3.5, éstos resultados no demostraron poseer invariancia a la traslación. Como puede observarse sólo en las posiciones T_0 y T_1 , los resultados son similares.

3. Momentos radiales ortogonales

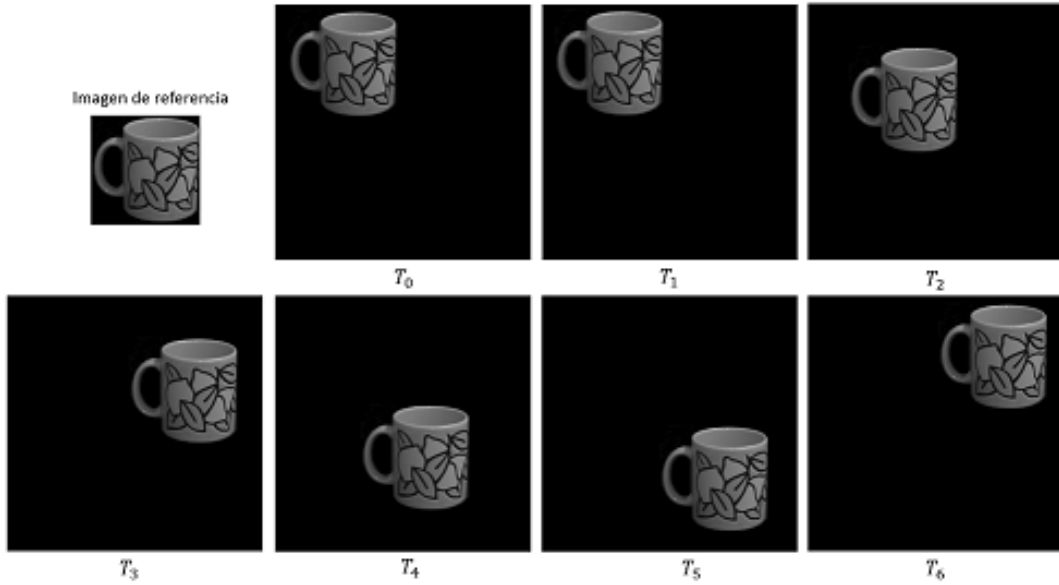


Figura 3.6: Objeto 59_19 original de la base de datos COIL-100 y trasladado.

Tabla 3.5: Invariante a la traslación de acuerdo a la propuesta de Chong con momentos de Zernike para diferentes órdenes n , m .

n	m	T_0	T_1	T_2	T_3	T_4	T_5	T_6
0	0	8.3084	9.8165	11.0657	11.0655	11.0658	10.7640	8.7765
1	1	11.7929	13.1137	7.6257	12.5478	6.1549	12.1654	12.6336
2	0	2.3031	1.0476	11.0622	3.2927	12.6507	3.4163	3.1490
2	2	10.0754	10.7817	3.5243	8.4891	2.2608	8.3224	10.5875
3	1	169.1423	168.2694	15.5258	99.3710	24.7188	93.6228	185.1495
3	3	1.8169	2.5276	1.0541	0.7907	0.1271	1.2080	0.6941
4	0	4.8524e+03	4.5506e+03	163.4650	2.3952e+03	116.4345	2.2681e+03	5.3947e+03
4	2	1.6021e+03	1.4960e+03	13.7783	705.9159	9.4470	661.6447	1.7799e+03
4	4	27.2096	25.3246	1.1025	16.8463	2.0632	15.0686	31.3558

Otra forma de obtener la invariancia a la traslación es realizando una normalización de la imagen donde el objeto se desplaza de tal manera que su centroide coincide con el origen del sistema de coordenadas. Hu [12] definió el centroide (x_c, y_c) como,

$$x_c = m_{10}/m_{00} \quad (3.18)$$

$$y_c = m_{01}/m_{00} \quad (3.19)$$

donde m_{01} , m_{10} y m_{00} son los momentos geométricos de primer orden y orden cero, respectivamente. Al ubicar el origen de coordenadas en el centroide (x_c, y_c) y mapear

3. Momentos radiales ortogonales

el radio unitario, se logra la invariancia a la traslación. Para ejemplificar lo anterior se tomó nuevamente el objeto 59_19 mostrado en la Fig. 3.6 y se calculó el centroide de todas las imágenes con la función *centroid* de Matlab R2016b. Se calcularon los momentos de Zernike con orden hasta $n = 5, m = 1$ para la imagen en escala de grises. Los resultados obtenidos se muestran en la Tabla 3.6 donde se demuestra que los momentos de Zernike son invariantes a la traslación, a diferencia del método propuesto por Chong et al. en donde los momentos no logran ser invariantes.

Tabla 3.6: momentos de Zernike para diferentes ordenes n, m .

n	m	Imagen original	T_0	T_1	T_2	T_3	T_4	T_5	T_6
0	0	0.3985	0.3985	0.3985	0.3985	0.3985	0.3985	0.3985	0.3985
1	1	0.1020	0.1020	0.1020	0.1020	0.1020	0.1020	0.1020	0.1020
2	0	0.0530	0.0530	0.0530	0.0530	0.0530	0.0530	0.0530	0.0530
2	2	0.0156	0.0156	0.0156	0.0156	0.0156	0.0156	0.0156	0.0156
3	1	0.0049	0.0049	0.0049	0.0049	0.0049	0.0049	0.0049	0.004
3	3	0.0359	0.0359	0.0359	0.0359	0.0359	0.0359	0.0359	0.0359
4	0	0.0116	0.0116	0.0116	0.0116	0.0116	0.0116	0.0116	0.0116
4	2	0.0504	0.0504	0.0504	0.0504	0.0504	0.0504	0.0504	0.0504
4	4	0.0188	0.0188	0.0188	0.0188	0.0188	0.0188	0.0188	0.0188
5	1	0.0107	0.0107	0.0107	0.0107	0.0107	0.0107	0.0107	0.0107

3.5. Error de reconstrucción

De acuerdo con la teoría ortogonal, la función imagen puede ser reconstruida por un número infinito de momentos radiales. También, podemos identificar el número de descriptores necesarios para describir la imagen. La distribución discreta de la imagen reconstruida está dada por,

$$\tilde{f}(r, \theta) = \sum_{n=0}^L \sum_{m=0}^L \phi_{n,m} A_n(r) \exp(jm\theta) \quad (3.20)$$

donde $\tilde{f}(r, \theta)$ es la versión reconstruida de $f(r, \theta)$, y L es el orden máximo de los momentos radiales utilizados en la reconstrucción. La reconstrucción de una imagen determina qué tan bien está caracterizada la imagen por un conjunto finito de momentos. Para evaluar la eficiencia de los momentos, se utiliza la métrica de reconstrucción basada en el error de reconstrucción de imagen normalizada (NIRE), que se define como el error cuadrado normalizado entre la imagen de entrada $f(x, y)$ y la imagen reconstruida $f'(x, y)$, expresado por [7],

3. Momentos radiales ortogonales

$$NIRE = \frac{\sum_{y=1}^N \sum_{x=1}^M [f(x, y) - f'(x, y)]^2}{\sum_{y=1}^N \sum_{x=1}^M f^2(x, y)}. \quad (3.21)$$

El NIRE disminuye con el aumento cantidad de momentos, para ejemplificar lo anterior se calcularon los momentos de Zernike del objeto 74_0 con distintos ordenes y posteriormente se realizó la reconstrucción de la función imagen original. El NIRE obtenido puede observarse en la Fig. 3.7.

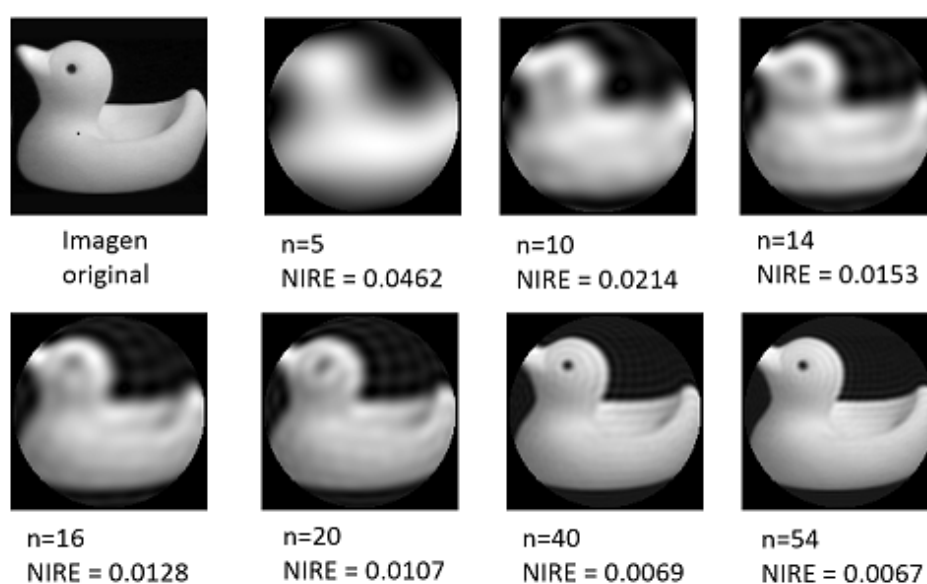


Figura 3.7: Error de reconstrucción del objeto 74_0 con momentos de Zernike y distintos órdenes.

Bibliografía

- [1] Singh, Chandan, Ekta Walia, and Rahul Upneja. "Accurate calculation of Zernike moments." *Information Sciences* 233 (2013): 255-275.
- [2] Chong, C-W., P. Raveendran, and R. Mukundan. "The scale invariants of pseudo-Zernike moments." *Pattern Analysis & Applications* 6.3 (2003): 176-184.
- [3] Sheng, Yunlong, and Lixin Shen. "Orthogonal Fourier–Mellin moments for invariant pattern recognition." *JOSA A* 11.6 (1994): 1748-1757.
- [4] Ping, ZiLiang, RiGeng Wu, and YunLong Sheng. "Image description with Chebyshev–Fourier moments." *JOSA A* 19.9 (2002): 1748-1754.
- [5] Chong, C-W., P. Raveendran, and R. Mukundan. "The scale invariants of pseudo-Zernike moments." *Pattern Analysis & Applications* 6.3 (2003): 176-184.
- [6] Xiao, Bin, Guo-yin Wang, and Wei-sheng Li. "Radial shifted Legendre moments for image analysis and invariant image recognition." *Image and Vision Computing* 32.12 (2014): 994-1006.
- [7] Xiao, Bin, Jian-Feng Ma, and Xuan Wang. "Image analysis by Bessel–Fourier moments." *Pattern Recognition* 43.8 (2010): 2620-2629.
- [8] Ping, Ziliang, et al. "Generic orthogonal moments: Jacobi–Fourier moments for invariant image description." *Pattern Recognition* 40.4 (2007): 1245-1254.
- [9] Hosny, Khalid M., and Mohamed M. Darwish. "Highly accurate and numerically stable higher order QPCET moments for color image representation." *Pattern Recognition Letters* 97 (2017): 29-36.
- [10] Upneja, Rahul, Chandan Singh, and Ajay Prashar. "Fast Computation of Chebyshev-Harmonic Fourier Moments." *Lecture Notes on Information Theory Vol 3.2* (2015).

BIBLIOGRAFÍA

- [11] Mukundan, Ramakrishnan. Radial Tchebichef invariants for pattern recognition."TENCON 2005-2005 IEEE Region 10 Conference. IEEE, 2005.
- [12] Hu, Ming-Kuei. Reconocimiento de patrones visuales por invariantes de momento". Transacciones IRE sobre la teoría de la información 8.2 (1962): 179-187.
- [13] Honarvar, B. Shakibaei, and Raveendran Paramesran. Recursive formula to compute Zernike radial polynomials..*optics letters* 38.14 (2013): 2487-2489.
- [14] Camacho-Bello, César, et al. Reconstruction of color biomedical images by means of quaternion generic Jacobi-Fourier moments in the framework of polar pixels."Journal of Medical Imaging 3.1 (2016): 014004.
- [15] Nene, Sameer A., Shree K. Nayar y Hiroshi Murase. "Biblioteca de imágenes de objetos de Columbia (bobina 20)". (1996): 7.
- [16] Chong, Chee-Way, Paramesran Raveendran, and Ramakrishnan Mukundan. "Translation invariants of Zernike moments."Pattern recognition 36.8 (2003): 1765-1773.

Capítulo 4

Momentos radiales de imágenes en color

4.1. Introducción

Los momentos y los invariantes han sido ampliamente utilizados en el reconocimiento de patrones y clasificación de objetos, debido a su capacidad para describir una imagen. Inicialmente se usaron en imágenes binarias o en escala de grises, sin embargo el uso de imágenes en color se ha generalizado gracias al desarrollo de cámaras digitales y el uso de Internet. De esta manera casi todas las imágenes adquiridas son cromáticas [1]. Usualmente para el procesamiento de imágenes en color se usan dos enfoques; el primero consiste en transformar una imagen de color en escala de grises con lo cual existe pérdida de información y la segunda es procesar los canales de forma separada. En este trabajo de tesis se analizan tres distintas propuestas para el procesamiento de imágenes en color. La primera propuesta, los momentos Cuaternión, los cuales fueron presentados por Chen et al [1], los autores hacen uso del álgebra de cuaterniones para tratar las imágenes en color de una manera holística, y a lo largo de casi una década varios autores han propuesto nuevos conjuntos de momentos Cuaternión para describir imágenes en color. La segunda propuesta analizada son los momentos Multicanal, realizada por Singh y Singh [2], los cuales describen la imagen en color por medio de los momentos de cada canal y demuestran la superioridad de los momentos Multicanal frente a los momentos Cuaternión. La tercera propuesta analizada es el clasificador multikernel, realizada por Singh y Singh [3] para el reconocimiento de objetos, que representa la imagen en color por medio de tres descriptores que aportan información sobre el color, forma, bordes y texturas. Para analizar estos tres métodos se realizaron pruebas con cuatro modelos de color para verificar su desempeño en la clasificación de imágenes.

4.2. Momentos Cuaternión.

En 1843, Hamilton [4] definió el cuaternión q , como una generalización de un número complejo. Un cuaternión tiene cuatro componentes, una parte real y tres partes imaginarias[5].

$$q = a + bi + cj + dk \quad (4.1)$$

donde a, b, c y d son números reales e i, j y k son tres unidades imaginarias. El cuaternión obedece las siguientes reglas:

$$\begin{aligned} i^2 &= j^2 = k^2 = ijk = -1 \\ ij &= -ji = k, jk = -kj = i, ki = -ik = j. \end{aligned} \quad (4.2)$$

El cuaternión q se llama cuaternión puro, cuando el componente $a = 0$. El conjugado y el módulo son definidos respectivamente como

$$q^* = a - bi - cj - dk, \quad (4.3)$$

$$|q| = \sqrt{a^2 + b^2 + c^2 + d^2}. \quad (4.4)$$

Para el análisis de las imágenes en color $f(r, \theta)$ puede ser representada usando un cuaternión puro como,

$$f(r, \theta) = f_R(r, \theta)i + f_G(r, \theta)j + f_B(r, \theta)k, \quad (4.5)$$

donde $f_R(r, \theta)$, $f_G(r, \theta)$, $f_B(r, \theta)$ representan los componentes del canal rojo, verde y azul, respectivamente. La unidad del cuaternión puro μ tiene el valor.

$$\mu = (i + j + k)/\sqrt{3} \quad (4.6)$$

debido a la propiedad no conmutativa de la multiplicación de cuaterniones, hay dos tipos de cuaternión: de lado derecho Q^R y lado izquierdo Q^L . En este trabajo de tesis trabajamos con los momentos de lado derecho. Sea una imagen RGB en coordenadas polares $(f_R(r, \theta)), (f_G(r, \theta)), (f_B(r, \theta))$ definimos los momentos Cuaternión del lado derecho de orden n con repetición m como,

$$Q_{n,m}^R(f) = \int_0^1 \int_0^{2\pi} P_n(r) f(r, \theta) e^{-\mu m \theta} r d\theta dr. \quad (4.7)$$

4. Momentos radiales de imágenes en color

Los cuaterniones de lado derecho e izquierdo para la misma imagen a color tienen la siguiente relación,

$$\begin{aligned} Q_{n,m}^R(f) &= \left(\int_0^1 \int_0^{2\pi} e^{-\mu m \theta} P_n(r) f(r, \theta) r d\theta dr \right)^* \\ &= -Q_{n,m}^L(f)^* \end{aligned}$$

Substituyendo la Ec. 4.5 en la Ec. 4.7 y teniendo en cuenta la Ec. 4.2 se obtiene [1]

$$\begin{aligned} Q_{n,m}^R(f) &= \int_0^1 \int_0^{2\pi} P_n(r) [f_R(r, \theta)i + f_G(r, \theta)j + f_B(r, \theta)k] e^{-\mu m \theta} r d\theta dr \\ &= i \int_0^1 \int_0^{2\pi} P_n(r) f_R(r, \theta) e^{-\mu m \theta} r d\theta dr \\ &\quad + j \int_0^1 \int_0^{2\pi} P_n(r) f_G(r, \theta) e^{-\mu m \theta} r d\theta dr \\ &\quad + k \int_0^1 \int_0^{2\pi} P_n(r) f_B(r, \theta) e^{-\mu m \theta} r d\theta dr \end{aligned}$$

$$\begin{aligned} Q_{n,m}^R(f) &= i \int_0^1 \int_0^{2\pi} P_n(r) f_R(r, \theta) [\cos(m\theta) - \mu \sin(m\theta)] r d\theta dr \\ &\quad + j \int_0^1 \int_0^{2\pi} P_n(r) f_G(r, \theta) [\cos(m\theta) - \mu \sin(m\theta)] r d\theta dr \\ &\quad + k \int_0^1 \int_0^{2\pi} P_n(r) f_B(r, \theta) [\cos(m\theta) - \mu \sin(m\theta)] r d\theta dr \end{aligned}$$

4. Momentos radiales de imágenes en color

$$\begin{aligned}
Q_{n,m}^R(f) &= i \left[\begin{array}{c} \int_0^1 \int_0^{2\pi} P_n(r) f_R(r, \theta) \cos(m\theta) r d\theta dr \\ -\mu \int_0^1 \int_0^{2\pi} P_n(r) f_R(r, \theta) \sin(m\theta) r d\theta dr \end{array} \right] \\
&+ j \left[\begin{array}{c} \int_0^1 \int_0^{2\pi} P_n(r) f_G(r, \theta) \cos(m\theta) r d\theta dr \\ -\mu \int_0^1 \int_0^{2\pi} P_n(r) f_G(r, \theta) \sin(m\theta) r d\theta dr \end{array} \right] \\
&+ k \left[\begin{array}{c} \int_0^1 \int_0^{2\pi} P_n(r) f_B(r, \theta) \cos(m\theta) r d\theta dr \\ -\mu \int_0^1 \int_0^{2\pi} P_n(r) f_B(r, \theta) \sin(m\theta) r d\theta dr \end{array} \right] \\
&= i \left[\operatorname{Re}(Q_{n,m}(f_R)) + \frac{i+j+k}{\sqrt{3}} \operatorname{Im}(Q_{n,m}(f_R)) \right] \\
&+ j \left[\operatorname{Re}(Q_{n,m}(f_G)) + \frac{i+j+k}{\sqrt{3}} \operatorname{Im}(Q_{n,m}(f_G)) \right] \\
&+ k \left[\operatorname{Re}(Q_{n,m}(f_B)) + \frac{i+j+k}{\sqrt{3}} \operatorname{Im}(Q_{n,m}(f_B)) \right] \\
Q_{n,m}^R(f) &= A_{n,m}^R + iB_{n,m}^R + jC_{n,m}^R + kD_{n,m}^R, \tag{4.8}
\end{aligned}$$

con

$$\begin{aligned}
A_{n,m}^R &= -\frac{1}{\sqrt{3}} [\operatorname{Im}(Q_{n,m}(f_R)) + \operatorname{Im}(Q_{n,m}(f_G)) + \operatorname{Im}(Q_{n,m}(f_B))], \\
B_{n,m}^R &= \operatorname{Re}(Q_{n,m}(f_R)) + \frac{1}{\sqrt{3}} [\operatorname{Im}(Q_{n,m}(f_G)) - \operatorname{Im}(Q_{n,m}(f_B))], \\
C_{n,m}^R &= \operatorname{Re}(Q_{n,m}(f_G)) + \frac{1}{\sqrt{3}} [\operatorname{Im}(Q_{n,m}(f_B)) - \operatorname{Im}(Q_{n,m}(f_R))], \\
D_{n,m}^R &= \operatorname{Re}(Q_{n,m}(f_B)) + \frac{1}{\sqrt{3}} [\operatorname{Im}(Q_{n,m}(f_R)) - \operatorname{Im}(Q_{n,m}(f_G))]. \tag{4.9}
\end{aligned}$$

donde $(Q_{n,m}(f_R))$, $(Q_{n,m}(f_G))$ y $(Q_{n,m}(f_B))$ son respectivamente los momentos radiales convencionales del canal rojo, verde y azul, respectivamente, $\operatorname{Re}(x)$ representa la parte real e $\operatorname{Im}(x)$, la parte imaginaria del número complejo x , es decir $\operatorname{Re}(a+bi) = a$, $\operatorname{Im}(a+bi) = b$.

4. Momentos radiales de imágenes en color

A partir de la teoría de cuaterniones y las propiedades de los polinomios ortogonales, la imagen a color $f_C(r, \theta)$ puede reconstruirse a partir de un número finito momentos Cuaternión $Q_{n,m}^R$. La imagen a color puede ser recuperada de la siguiente manera

$$\begin{aligned}
f'(r, \theta) &= \sum_{n=0}^{+\infty} \sum_{m=0}^{+\infty} Q_{n,m}^R P_n(r) e^{\mu m \theta} \\
&= \sum_{n=0}^{+\infty} \sum_{m=0}^{+\infty} (A_{n,m}^R + iB_{n,m}^R + jC_{n,m}^R + kD_{n,m}^R) P_n(r) e^{\mu m \theta} \\
&= \sum_{n=0}^{+\infty} \sum_{m=0}^{+\infty} A_{n,m}^R P_n(r) e^{-\mu m \theta} + i \sum_{n=0}^{+\infty} \sum_{m=0}^{+\infty} B_{n,m}^R P_n(r) e^{\mu m \theta} \\
&\quad + j \sum_{n=0}^{+\infty} \sum_{m=0}^{+\infty} C_{n,m}^R P_n(r) e^{-\mu m \theta} + k \sum_{n=0}^{+\infty} \sum_{m=0}^{+\infty} D_{n,m}^R P_n(r) e^{\mu m \theta} \\
&= \sum_{n=0}^{+\infty} \sum_{m=0}^{+\infty} A_{n,m}^R P_n(r) e^{\frac{i+j+k}{\sqrt{3}} m \theta} \\
&\quad + i \sum_{n=0}^{+\infty} \sum_{m=0}^{+\infty} B_{n,m}^R P_n(r) e^{\frac{i+j+k}{\sqrt{3}} m \theta} \\
&\quad + j \sum_{n=0}^{+\infty} \sum_{m=0}^{+\infty} C_{n,m}^R P_n(r) e^{\frac{i+j+k}{\sqrt{3}} m \theta} \\
&\quad + k \sum_{n=0}^{+\infty} \sum_{m=0}^{+\infty} D_{n,m}^R P_n(r) e^{\frac{i+j+k}{\sqrt{3}} m \theta} \\
f'(r, \theta) &= f'_A(r, \theta) + i f'_B(r, \theta) + j f'_C(r, \theta) + k f'_D(r, \theta) \tag{4.10}
\end{aligned}$$

donde

$$\begin{aligned}
f'_A(r, \theta) &= \operatorname{Re}(A_{n,m}^R) - \frac{1}{\sqrt{3}} [\operatorname{Im}(B_{n,m}^R) + \operatorname{Im}(C_{n,m}^R) + \operatorname{Im}(D_{n,m}^R)] \\
f'_B(r, \theta) &= \operatorname{Re}(B_{n,m}^R) + \frac{1}{\sqrt{3}} [\operatorname{Im}(A_{n,m}^R) + \operatorname{Im}(C_{n,m}^R) - \operatorname{Im}(D_{n,m}^R)] \\
f'_C(r, \theta) &= \operatorname{Re}(C_{n,m}^R) + \frac{1}{\sqrt{3}} [\operatorname{Im}(A_{n,m}^R) - \operatorname{Im}(B_{n,m}^R) + \operatorname{Im}(D_{n,m}^R)] \\
f'_D(r, \theta) &= \operatorname{Re}(D_{n,m}^R) + \frac{1}{\sqrt{3}} [\operatorname{Im}(A_{n,m}^R) + \operatorname{Im}(B_{n,m}^R) - \operatorname{Im}(C_{n,m}^R)] \tag{4.11}
\end{aligned}$$

donde $f'_A(r, \theta)$ es una matriz cercana a 0, y $f'_B(r, \theta), f'_C(r, \theta), f'_D(r, \theta)$ representan respectivamente los componentes rojo, verde y azul de la imagen a color reconstruida. También $A_{n,m}^R, B_{n,m}^R, C_{n,m}^R$ y $D_{n,m}^R$ representan respectivamente la reconstrucción de las matrices $A_{n,m}^R, B_{n,m}^R, C_{n,m}^R$ y $D_{n,m}^R$ dadas como:

4. Momentos radiales de imágenes en color

$$\begin{aligned}
 A'_{n,m} &= \sum_{n=0}^{+\infty} \sum_{m=-\infty}^{+\infty} A_{n,m}^R P_n(r) e^{-\mu m \theta} \\
 B'_{n,m} &= \sum_{n=0}^{+\infty} \sum_{m=-\infty}^{+\infty} B_{n,m}^R P_n(r) e^{-\mu m \theta} \\
 C'_{n,m} &= \sum_{n=0}^{+\infty} \sum_{m=-\infty}^{+\infty} C_{n,m}^R P_n(r) e^{-\mu m \theta} \\
 D'_{n,m} &= \sum_{n=0}^{+\infty} \sum_{m=-\infty}^{+\infty} D_{n,m}^R P_n(r) e^{-\mu m \theta}
 \end{aligned}$$

4.2.1. Invariancia a la rotación

Si denotamos una imagen en color $f(r, \theta)$, la imagen denotada como $f(r, \theta + \varphi)$ será la imagen inicial rotada por un ángulo φ . Los momentos Cuaternión de ambas imágenes tienen la siguiente relación,

$$\begin{aligned}
 Q'_{n,m}(f) &= \int_0^1 \int_0^{2\pi} P_n(r) f(r, \theta + \varphi) e^{-\mu m \theta} r d\theta dr \\
 &= \int_0^1 \int_0^{2\pi} P_n(r) f(r, \theta) e^{-\mu m (\theta + \varphi)} r d\theta dr \\
 &= e^{\mu m \varphi} \int_0^1 \int_0^{2\pi} P_n(r) f(r, \theta) e^{-\mu m \theta} r d\theta dr \\
 &= e^{\mu m \varphi} Q_{n,m}(f)
 \end{aligned} \tag{4.12}$$

donde $Q'_{n,m}(f)$ y $Q_{n,m}(f)$ son los momentos cuaternión de $f(r, \theta + \varphi)$ y $f(r, \theta)$. La rotación de una imagen a color por un ángulo φ induce un cambio de fase $e^{\mu m \varphi}$ de $Q_{n,m}$. La invariancia a la rotación se puede obtener mediante la siguiente relación,

$$|Q'_{n,m}(f)| = |e^{\mu m \varphi} Q_{n,m}(f)| = |e^{\mu m \varphi}| \cdot |Q_{n,m}(f)| = |Q_{n,m}(f)|. \tag{4.13}$$

Para ejemplificar la invariancia a la rotación, el objeto 74_0 fue rotado a 30°, 60°, 90°, 120°, 150° y 180° en el programa Matlab versión 9.1 R2016b con la función *imrotate*, con el método *'bilinear'* y el cuadro delimitador con la opción *crop*, como se muestra en la Fig. 4.1. Se calcularon los primeros 10 momentos Cuaternión de Zernike. Los datos constantes a diferentes ángulos de rotación mostrados en las Tablas 4.1 y 4.2 demuestran que los momentos de Zernike son invariantes a la rotación.

4. Momentos radiales de imágenes en color

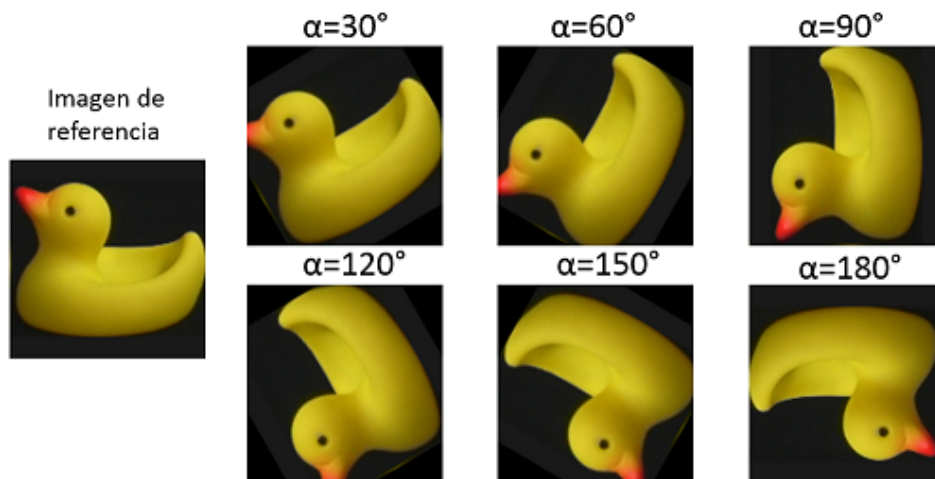


Figura 4.1: Objeto 74_0 en color de la base de datos COIL-100 es rotado a 30°, 60°, 90°, 120°, 150° y 180°.

Tabla 4.1: MSE para evaluar la efectividad de la invariante de rotación de los momentos Cuaternión de Zernike.

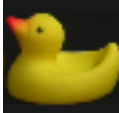


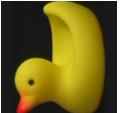
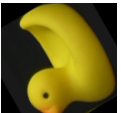

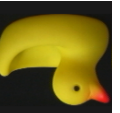
Imagen original	$\alpha = 30^\circ$	$\alpha = 60^\circ$	$\alpha = 90^\circ$	$\alpha = 120^\circ$	$\alpha = 150^\circ$	$\alpha = 180^\circ$
						
MSE	2.1898e-4	2.2319e-4	2.2266e-29	2.1898e-4	2.2319e-4	1.9958e-29

Tabla 4.2: momentos de Cuaternión Zernike para diferentes órdenes en diferentes rotaciones.

n	m	Imagen original	$\alpha = 30^\circ$	$\alpha = 60^\circ$	$\alpha = 90^\circ$	$\alpha = 120^\circ$	$\alpha = 150^\circ$	$\alpha = 180^\circ$	σ
0	0	79.7484	79.7425	79.7442	79.7484	79.7425	79.7442	79.7484	2.8.e-3
1	1	12.3023	12.2952	12.2981	12.3023	12.2952	12.2981	12.3023	3.2e-3
2	0	26.5276	26.5240	26.5238	26.5276	26.5240	26.5238	26.5276	2e-3
2	2	18.2678	18.2565	18.2527	18.2678	18.2565	18.2527	18.2678	7.2e-3
3	1	14.2171	14.1948	14.1950	14.2171	14.1948	14.1950	14.2171	1.19e-2
3	3	15.6386	15.6042	15.6030	15.6386	15.6042	15.6030	15.6386	1.87e-2
4	0	1.6971	1.6942	1.6921	1.6971	1.6942	1.6921	1.6971	2.3e-3
4	2	12.2134	12.2071	12.2088	12.2134	12.2071	12.2088	12.2134	3e-3
4	4	2.7923	2.7768	2.7800	2.7923	2.7768	2.7800	2.7923	7.5e-3
5	1	5.9090	5.9090	5.9114	5.9090	5.9090	5.9114	5.9090	1.2e-3

4. Momentos radiales de imágenes en color

4.2.2. Invariancia a la escala

De acuerdo a Hosny y Darwish [5] el módulo de los momentos son invariantes a la escala debido a que previamente la imagen queda definida dentro de un círculo de radio unitario y posteriormente se calculan los momentos Cuaternión de Zernike. Para demostrar la invariancia a la escala, se crearon imágenes sintéticas del objeto 52_0 de la base de datos COIL-100 en diferentes tamaños con la función *imresize* de MatLab *R2016b* con el parámetro *'bilinear'* como se muestra en la Fig 4.2. De la misma forma, para probar la invariancia a la escala se calcularon los primeros 10 momentos Cuaternión. Los resultados obtenidos se muestran en la Tabla 4.3 y 4.4 donde se demuestra que los momentos de Zernike son invariantes a la escala.

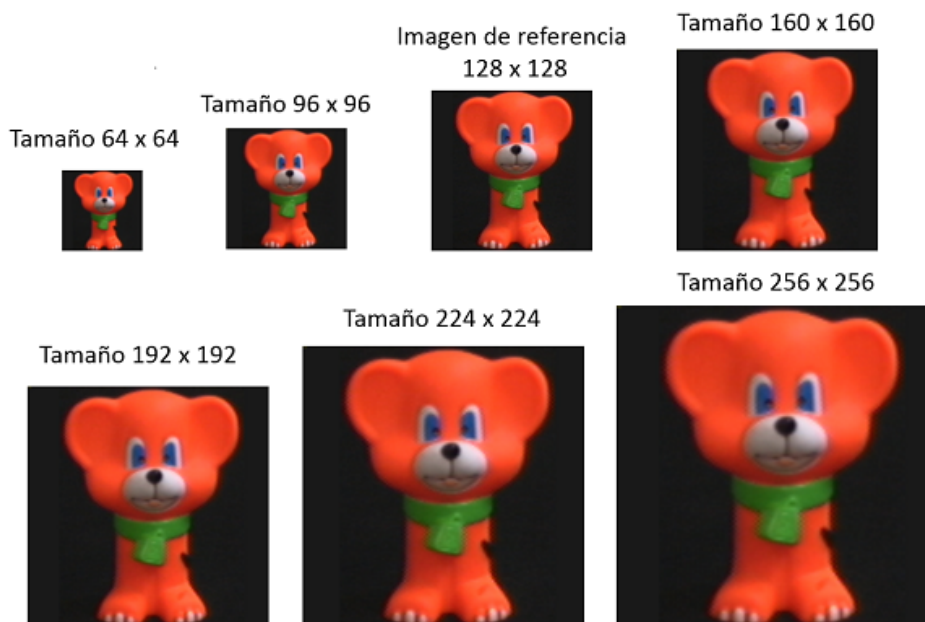


Figura 4.2: Objeto 52_0 a color de la base de datos COIL-100 a diferentes escalas.

Tabla 4.3: Error cuadrático medio (MSE) de la imagen de referencia y las imágenes con diferentes escalas de los momentos Cuaternión de Zernike.

Tamaño	64 × 64	96 × 96	128 × 128	160 × 160	192 × 192	224 × 224	256 × 256
MSE	0.1778	0.0235	0	0.0067	0.0187	0.0328	0.0414

4. Momentos radiales de imágenes en color

Tabla 4.4: Momentos Cuaternión de Zernike para diferentes órdenes.

n	m	64 × 64	96 × 96	128 × 128	160 × 160	192 × 192	224 × 224	256 × 256	σ
0	0	69.3465	69.0863	68.8590	68.8284	68.8051	68.6907	68.7592	0.2283
1	1	19.3403	19.1533	19.0582	19.0123	18.9746	18.9444	18.9238	0.1465
2	0	7.4803	7.6891	7.7797	7.8818	7.9400	7.9620	8.0184	0.1879
2	2	31.6860	31.3693	31.1991	31.0714	30.9926	30.9241	30.8965	0.2837
3	1	9.4289	9.5943	9.7016	9.7257	9.7473	9.7678	9.7740	0.1251
3	3	9.5510	9.5450	9.5518	9.5268	9.5219	9.5202	9.5163	0.0154
4	0	6.1710	6.4028	6.5212	6.5823	6.6259	6.6564	6.6762	0.1796
4	2	9.5617	10.1750	10.4981	10.6718	10.7963	10.8810	10.9392	0.4901
4	4	2.8854	2.7631	2.7176	2.6773	2.6522	2.6402	2.6208	0.0921
5	1	5.7116	5.6003	5.5317	5.5153	5.4980	5.4894	5.4773	0.0834

4.2.3. Invariancia a la traslación

La invariancia a la traslación se logra calculando el centroide de la imagen y desplazándolo de tal forma que coincida con el origen de la imagen, el centroide (x_c, y_c) de una imagen RGB queda definido como [6],

$$\begin{aligned}
 x_c &= (m_{10}(f_R) + m_{10}(f_G) + m_{10}(f_B))/m_{00}, \\
 y_c &= (m_{01}(f_R) + m_{01}(f_G) + m_{01}(f_B))/m_{00}, \\
 m_{00} &= (m_{00}(f_R) + m_{00}(f_G) + m_{00}(f_B))
 \end{aligned} \tag{4.14}$$

donde $m_{00}(f_R), m_{10}(f_R)$ y $m_{01}(f_R)$ son los momentos geométricos de orden cero y de primer orden respectivamente del canal rojo. De forma similar, $m_{00}(f_G), m_{10}(f_G)$ y $m_{01}(f_G)$ corresponde al canal verde y $m_{00}(f_B), m_{10}(f_B)$ y $m_{01}(f_B)$ son los correspondientes al canal azul. Los momentos Cuaternión centrales, invariantes a la traslación, se obtienen como

$$\bar{Q}_{n,m}(f) = \int_0^1 \int_0^{2\pi} P_n(\bar{r}) f(\bar{r}, \bar{\theta}) e^{-\mu m \bar{\theta}} \bar{r} d\bar{\theta} d\bar{r} \tag{4.15}$$

donde $(\bar{r}, \bar{\theta})$ es una imagen en coordenadas polares con origen en (x_c, y_c) . Para ejemplificar lo anterior se tomó el objeto 59_19 de la base de datos COIL-100 con un tamaño de 128×128 pixeles y se trasladó en varias posiciones dentro de una imagen de fondo con un tamaño de 300×300 pixeles como se muestra en la Fig. 4.3. Los resultados obtenidos de los primeros 10 momentos Cuaternión de Zernike de la Fig. 4.3 se muestran en la Tabla 4.5 donde los valores permanecen constantes.

4. Momentos radiales de imágenes en color

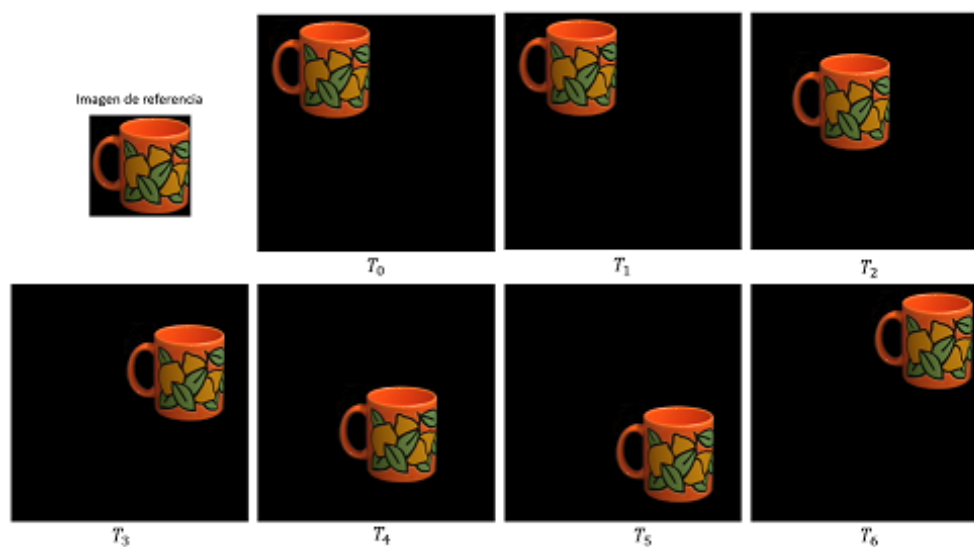


Figura 4.3: Objeto 59_19 a color de la base de datos COIL-100 trasladado.

Tabla 4.5: Momentos Cuaternión de Zernike para diferentes órdenes.

n	m	Imagen original	T_0	T_1	T_2	T_3	T_4	T_5	T_6
0	0	62.6785	62.6785	62.6785	62.6785	62.6785	62.6785	62.6785	62.6785
1	1	22.4478	22.4478	22.4478	22.4478	22.4478	22.4478	22.4478	22.4478
2	0	3.7636	3.7636	3.7636	3.7636	3.7636	3.7636	3.7636	3.7636
2	2	9.9055	9.9055	9.9055	9.9055	9.9055	9.9055	9.9055	9.9055
3	1	4.5150	4.5150	4.5150	4.5150	4.5150	4.5150	4.5150	4.5150
3	3	11.6059	11.6059	11.6059	11.6059	11.6059	11.6059	11.6059	11.6059
4	0	5.6785	5.6785	5.6785	5.6785	5.6785	5.6785	5.6785	5.6785
4	2	8.5759	8.5759	8.5759	8.5759	8.5759	8.5759	8.5759	8.5759
4	4	2.2745	2.2745	2.2745	2.2745	2.2745	2.2745	2.2745	2.2745
5	1	3.5705	3.5705	3.5705	3.5705	3.5705	3.5705	3.5705	3.5705

4. Momentos radiales de imágenes en color

4.2.4. Error de reconstrucción de imágenes RGB

El cálculo preciso de los momentos de orden superior representa un desafío para los métodos numéricos más comunes [2]. Para medir el error de reconstrucción de una imagen a color se toma en cuenta el promedio del NIRE (Ec. 3.20) de los tres canales [7],

$$MeanNIRE = \frac{NIRE_R + NIRE_G + NIRE_B}{3}. \quad (4.16)$$

La efectividad de los momentos utilizados en la reconstrucción será mayor en cuanto a los valores obtenidos del NIRE se aproxime a cero [4]. Para ejemplificar lo anterior se calcularon los momentos Zernike del objeto 74_0 con distintos órdenes en la reconstrucción de la función imagen. El NIRE obtenido puede observarse en la Fig. 4.4.

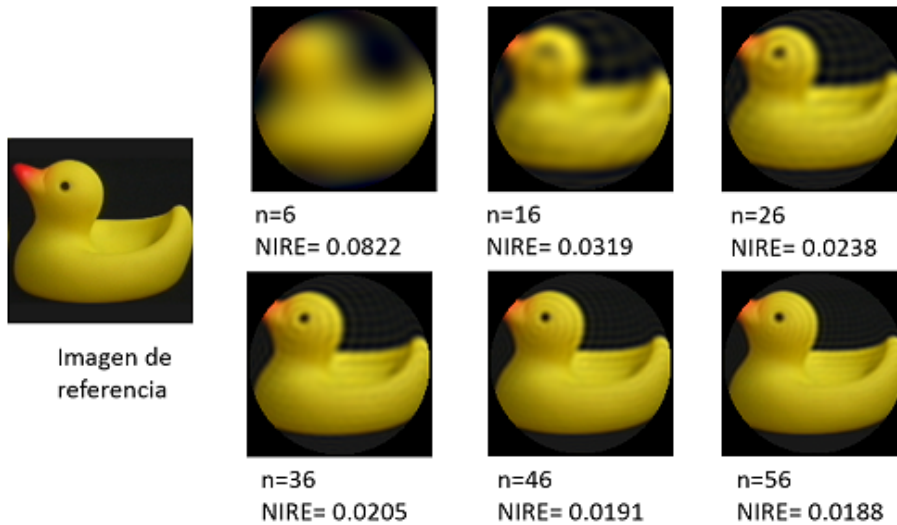


Figura 4.4: Error de reconstrucción del objeto 74_0 para distintos órdenes de los momentos Cuaternión de Zernike.

4.3. Momentos Multicanal

Los momentos ortogonales y sus invariantes han sido usados en el reconocimiento de patrones inicialmente en imágenes monocromáticas o en escalas de grises, con el avance de la tecnología hoy en día, las imágenes adquiridas son cromáticas. En los últimos años la teoría de momentos se aplicó a las imágenes de color mediante el uso de

4. Momentos radiales de imágenes en color

momentos Cuaternión. Sin embargo, los autores Singh y Singh [8] propusieron un nuevo conjunto denominado momentos Multicanal en donde concluyeron que éstos tienen un desempeño superior en tareas de reconocimiento de imágenes en comparación con los momentos Cuaternión existentes. No existen estudios comparativos entre los momentos Multicanal y los momentos Cuaternión que puedan determinar cuál conjunto de momentos representan mejor la imagen a color y si la carga computacional de los momentos Cuaternión está justificada. Los momentos Multicanal pueden representarse como,

$$M_{n,m} = \int_0^{2\pi} \int_0^1 P_{n,m}(r, \theta) f(r, \theta) r dr d\theta, \quad (4.17)$$

donde $P_{n,m}(r, \theta)$ es la función kernel y $f(r, \theta)$ representa una función imagen a color en la que es representada la información contenida en sus tres canales $fR(r, \theta), fG(r, \theta), fB(r, \theta)$, es decir $f(r, \theta) = (fR(r, \theta), fG(r, \theta), fB(r, \theta))$. Los momentos Multicanal son por lo tanto un conjunto de momentos de cada canal una imagen en color. La imagen puede ser reconstruida por el conjunto finito de momentos el cual está dado por,

$$f^{recons}(r, \theta) = \sum_{n=0}^L \sum_{m=-n_{\max}}^L M_{n,m}^p P_{n,m}(r, \theta) \quad (4.18)$$

$$= \sum_{n=0}^L \sum_{m=-n_{\max}}^L M_{n,m}^p A_n(r) \exp(jm\theta) \quad (4.19)$$

donde $f^{recons}(r, \theta)$ es la imagen reconstruida de $f(r, \theta)$ y L es el orden máximo de los momentos radiales utilizados en la reconstrucción.

4.3.1. Invariancia a la rotación

La invariancia a la rotación de los momentos Multicanal se obtiene con la Ec. 3.7 aplicada a cada canal de color. Si denotamos una imagen a color como $f_p(r, \theta)$, la imagen denotada como $f^\alpha(r, \theta)$ será la imagen rotada por un ángulo α . Los momentos Multicanal de la imagen rotada y la imagen original están dados por [3],

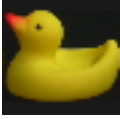
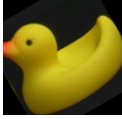
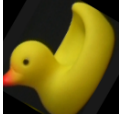
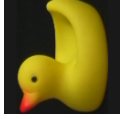
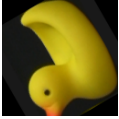

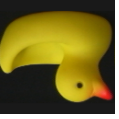
$$M_{n,l}^p(f_p) = e^{-jm\alpha} M_{n,l}^p(f_p), \quad p \in \{R, G, B\}, \quad (4.20)$$

por lo que,

$$|M_{n,l}^p(f_p)| = |M_{n,l}^p(f_p)|, \quad p \in \{R, G, B\}$$

4. Momentos radiales de imágenes en color

Tabla 4.6: MSE para evaluar la efectividad del invariante de rotación de los momentos Multicanal de Zernike.

Imagen original	$\alpha = 30^\circ$	$\alpha = 60^\circ$	$\alpha = 90^\circ$	$\alpha = 120^\circ$	$\alpha = 150^\circ$	$\alpha = 180^\circ$
						
MSE	7.9872e-05	7.7935e-05	8.3522e-30	7.9872e-05	7.7935e-05	1.0334e-29

Para ejemplificar la invariancia a la rotación de los momentos Multicanal, se utilizan los polinomios de Zernike como kernel y el objeto 74_0 de la base de datos COIL-100. La imagen de prueba es rotada a 30° , 60° , 90° , 120° , 150° y 180° en el programa Matlab versión 9.1 R2016b con la función *imrotate*, el método *'bilinear'* y el cuadro delimitador con la opción *crop* como se muestra en la Fig. 4.1. Se calculó el error cuadrático medio de los primeros 10 momentos Multicanal de Zernike. Los resultados obtenidos se muestran en la Tabla 4.6 donde se demuestra que los momentos Multicanal de Zernike son invariantes a la rotación.

4.3.2. Invariancia a la escala

Nuevamente la invariancia a la escala se obtiene mapeando previamente la imagen dentro de un círculo de radio unitario y posterior a este procedimiento se calcula el módulo de los momentos Multicanal de Zernike, para cada canal de color. Para demostrar la invariancia a la escala, tomamos el objeto 52_0 de la base de datos COIL-100 con los diferentes tamaños mostrados en la Fig. 4.2 y se calculó el error medio cuadrático para los primeros 10 momentos Cuaternión de Zernike. Los resultados obtenidos se muestran en la Tabla 4.7 donde se demuestra que los momentos de Zernike son invariantes a la escala.

Tabla 4.7: Error cuadrático medio MSE para evaluar la efectividad de la invariante de escala de los momentos Multicanal de Zernike.

Tamaño	64×64	96×96	128×128	160×160	192×192	224×224	256×256
MSE	0.0797	0.0101	0	0.0032	0.0090	0.0148	0.0196

4.3.3. Invariancia a la traslación

Como se describe en el capítulo anterior, la invariancia a la traslación se logra calculando el centroide de la imagen y desplazando el objeto de modo que coincida

4. Momentos radiales de imágenes en color

con el origen del sistema. Nuevamente mediante la Ec. 4.15 obtenemos el centroide del objeto 59_19 en color trasladado en varias posiciones como se muestra previamente en la Fig. 4.3. Con las coordenadas con origen en (x_c, y_c) se realiza cálculo de los primeros 4 momentos Multicanal de Zernike para cada canal de la imagen f_R, f_G y f_B de forma separada. Los momentos Multicanal obtenidos se muestran en la Tabla 4.8 donde se demuestra la invariancia a la traslación.

Tabla 4.8: momentos multicanal de Zernike para diferentes órdenes de la Fig. 4.3.

canal	n	m	Imagen original	T_0	T_1	T_2	T_3	T_4	T_5	T_6
R	0	0	54.2189	54.2189	54.2189	54.2189	54.2189	54.2189	54.2189	54.2189
R	1	1	21.5068	21.5068	21.5068	21.5068	21.5068	21.5068	21.5068	21.5068
R	2	0	0.2151	0.2151	0.2151	0.2151	0.2151	0.2151	0.2151	0.2151
R	2	2	8.8452	8.8452	8.8452	8.8452	8.8452	8.8452	8.8452	8.8452
G	0	0	29.5454	29.5454	29.5454	29.5454	29.5454	29.5454	29.5454	29.5454
G	1	1	12.1727	12.1727	12.1727	12.1727	12.1727	12.1727	12.1727	12.1727
G	2	0	3.7512	3.7512	3.7512	3.7512	3.7512	3.7512	3.7512	3.7512
G	2	2	2.1082	2.1082	2.1082	2.1082	2.1082	2.1082	2.1082	2.1082
B	0	0	10.7687	10.7687	10.7687	10.7687	10.7687	10.7687	10.7687	10.7687
B	1	1	3.6764	3.6764	3.6764	3.6764	3.6764	3.6764	3.6764	3.6764
B	2	0	0.2172	0.2172	0.2172	0.2172	0.2172	0.2172	0.2172	0.2172
B	2	2	0.4898	0.4898	0.4898	0.4898	0.4898	0.4898	0.4898	0.4898

4.3.4. Error de reconstrucción de imágenes RGB con momentos Multicanal

El error de reconstrucción determina cuánta información se puede describir por canal con un número determinado de momentos. Para realizar pruebas, se toma el objeto 74_0 de la base de datos COIL-100 como imagen de referencia. Se realizó la reconstrucción de la función imagen original tomando los momentos Multicanal de Zernike calculados con distintos órdenes. Los resultados obtenidos pueden observarse en la Fig. 4.5

4.4. Clasificador Multikernel

La siguiente propuesta analizada es Multiple Kernel Learning Classification, realizada por Singh y Singh [3] para el reconocimiento de objetos. Desarrollan un marco para combinar tres modalidades diferentes que representan el color, la forma y la textura del objeto. Este método se basa en los Momento de Zernike utilizado en 2 niveles; el nivel bajo proporciona información sobre la forma y en el nivel alto los

4. Momentos radiales de imágenes en color

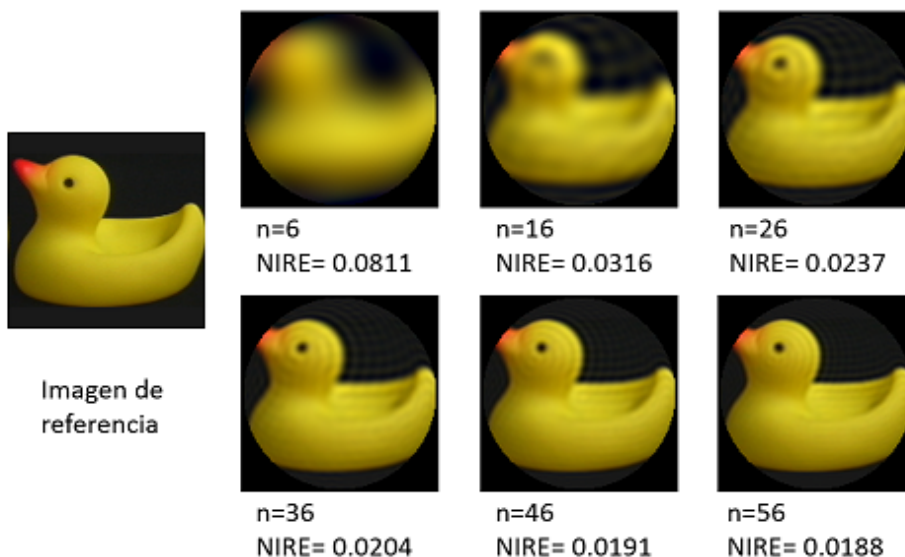


Figura 4.5: Error de reconstrucción del objeto 74_0 en color con momentos Multicanal de Zernike y distintos órdenes.

momentos de Zernike se aplican al gradiente de color de la imagen para proporcionar características de textura, como tercer descriptor se utiliza el histograma de color normalizado para describir la información de color. La técnica de Clasificador Multi Kernel permite fusionar las características de color, forma y textura proporcionando tasas de reconocimiento altas.

4.4.1. Descriptores de forma.

Utiliza los momentos Multicanal de Zernike los cuales son descriptores globales, y poseen un mejor rendimiento en el reconocimiento de la imagen en comparación con los momentos del Quaternion de Zernike. Una imagen a color $f(x, y)$ está representada por la información contenida en sus tres canales rojo, verde y azul, es decir $f_p(x, y) = (f_R(x, y), f_G(x, y), f_B(x, y))$. Los coeficientes de magnitudes de los momentos de Zernike de cada canal $|ZM_{nl}(fR)|$, $|ZM_{nl}(fG)|$ y $|ZM_{nl}(fB)|$ se concatenan para formar un conjunto de descriptores de forma de bajo nivel. Este conjunto de descriptores son invariantes a la rotación, escala y traslación como se pudo comprobar en el capítulo anterior.

4. Momentos radiales de imágenes en color

4.4.2. Descriptores de textura.

Las características de alto nivel (textura) se obtienen mediante los momentos de Zernike del gradiente de color de la imagen, que se obtiene con el método Di Zenzo. Sea $f_x(x, y)$ and $f_y(x, y)$ la representación del x-gradiente y y-gradiente, los cuales son derivados como [10],

$$\begin{aligned} f_x(x, y) &= \frac{\partial f_R(x, y)}{\partial x} i + \frac{\partial f_G(x, y)}{\partial x} j + \frac{\partial f_B(x, y)}{\partial x} k \\ f_y(x, y) &= \frac{\partial f_R(x, y)}{\partial y} i + \frac{\partial f_G(x, y)}{\partial y} j + \frac{\partial f_B(x, y)}{\partial y} k \end{aligned} \quad (4.21)$$

donde i, j, k representan los vectores unitarios a lo largo de los ejes R, G y B respectivamente de la imagen. Los términos g_{xx} , g_{yy} , y g_{xy} se definen como:

$$\begin{aligned} g_{xx} &= f_x(x, y) \cdot f_x(x, y) \\ g_{yy} &= f_y(x, y) \cdot f_y(x, y) \\ g_{xy} &= f_x(x, y) \cdot f_y(x, y) \end{aligned} \quad (4.22)$$

por lo que la magnitud del gradiente $G(x, y)$ y la dirección de la imagen $\theta(x, y)$ están dadas por,

$$\theta(x, y) = \frac{1}{2} \tan^{-1} \left(\frac{2g_{xy}}{g_{xx} - g_{yy}} \right) \quad (4.23)$$

$$G(x, y) = \left\{ \frac{1}{2} \left[(g_{xx} + g_{yy}) + \sqrt{(g_{xx} - g_{yy})^2 + 4g_{xy}^2} \right] \right\}^{1/2} \quad (4.24)$$

la magnitud de $G_{\min}(x, y)$ estará en la dirección $\theta \pm \frac{\pi}{2}$ la cual es

$$G_{\min}(x, y) = \left\{ \frac{1}{2} \left[(g_{xx} + g_{yy}) - \sqrt{(g_{xx} - g_{yy})^2 + 4g_{xy}^2} \right] \right\}^{1/2} \quad (4.25)$$

Invariancia a la rotación, escala y traslación

Para demostrar la invariancia a la escala del gradiente de la imagen, el objeto 74_0 de la base de datos COIL-100 es rotado a 30° , 60° , 90° , 120° , 150° y 180° , en el programa Matlab versión 9.1 R2016b con la función *imrotate*, el método *'bilinear'* y el cuadro delimitador con la opción *'crop'* como se muestra en la Fig. 4.6. Se calcularon los momentos de Zernike del gradiente del objeto de los primeros 10 órdenes y se calcula el MSE. Los resultados obtenidos se muestran en la Tabla 4.9 donde se demuestra que los momentos de Zernike del gradiente del objeto son invariantes a la rotación.

4. Momentos radiales de imágenes en color

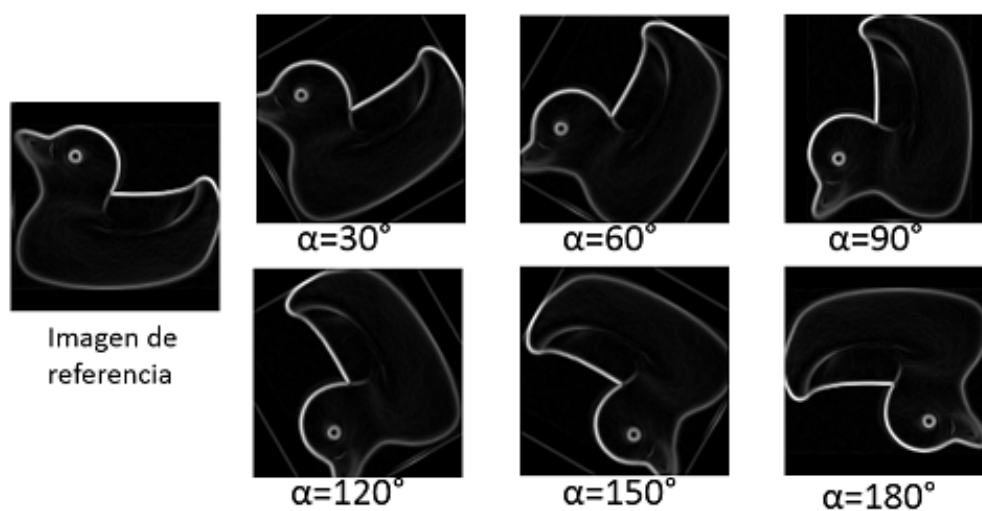
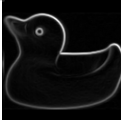

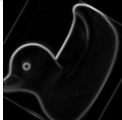
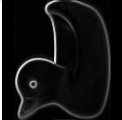
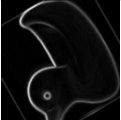
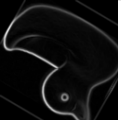
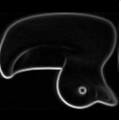


Figura 4.6: Gradiente del Objeto 74_0 en color de la base de datos COIL-100 rotado a 30°, 60°, 90°, 120°,150° y 180°.

Tabla 4.9: MSE para evaluar la efectividad de la invariante de rotación de los momentos de Zernike del gradiente de la imagen en color.

Imagen original	$\alpha = 30^\circ$	$\alpha = 60^\circ$	$\alpha = 90^\circ$	$\alpha = 120^\circ$	$\alpha = 150^\circ$	$\alpha = 180^\circ$
						
MSE	9.7639e-07	1.0451e-06	4.1328e-35	9.7639e-07	1.0451e-06	5.1327e-35

4. Momentos radiales de imágenes en color

Para demostrar la invariancia a la escala del gradiente de la imagen en color, se tomó el objeto 52_0 con un tamaño original de 128×128 y se cambió el tamaño de la imagen en el programa Matlab versión 9.1 R2016b con la función *imresize*, con el parámetro *'bilinear'*. Las imágenes escaladas se muestran en la Fig. 4.7. Se calculó los momentos Cuaternión de Zernike con orden hasta $n = 5, m = 1$. Los resultados obtenidos se muestran en la Tabla 4.10 donde se demuestra que los momentos de Zernike del gradiente de la imagen en color son invariantes a la escala.

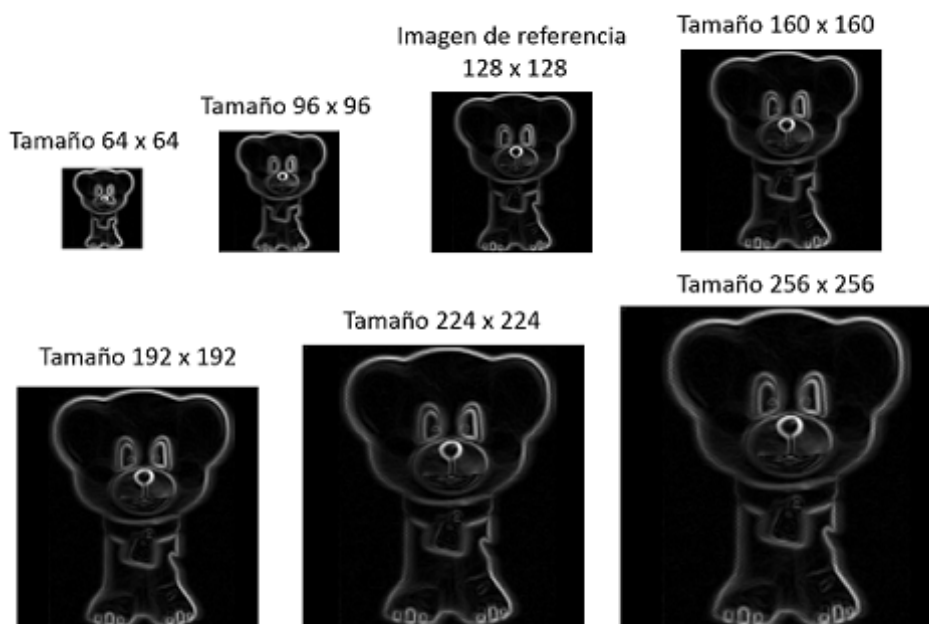


Figura 4.7: Gradiente del Objeto 52_0 a color de la base de datos COIL-100 a diferentes escalas.

Tabla 4.10: Error cuadrático medio (MSE) para evaluar la efectividad de la invariante de escala de los momentos de Zernike del gradiente de la imagen en color.

Tamaño	64×64	96×96	128×128	160×160	192×192	224×224	256×256
MSE	1.0295e-04	1.5943e-05	0	1.2508e-08	4.2379e-07	2.4194e-06	1.2215e-05

Para demostrar la invariancia a la traslación del gradiente de la imagen en color, se tomó el objeto 59_19 con un tamaño original de 128×128 y se trasladó en varias posiciones dentro de una imagen de fondo con un tamaño de 300×300 como se muestra en la Fig. 4.8 y se calculó el centroide de todas las imágenes en el programa Matlab versión 9.1 R2016b. Se calcularon los momentos Cuaternión de Zernike con orden

4. Momentos radiales de imágenes en color

hasta $n = 5, m = 1$. Los resultados obtenidos se muestran en la Tabla 4.11 donde se demuestra la invariancia a la traslación de los momentos de Zernike del gradiente de la imagen.

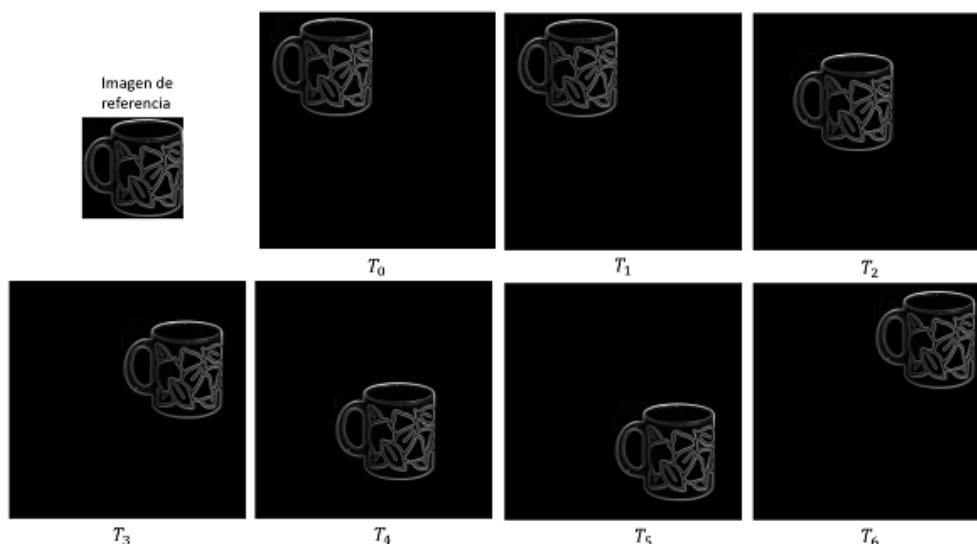


Figura 4.8: Gradiente del Objeto 59_19 en color de la base de datos COIL-100 con posición original y trasladado.

Tabla 4.11: momentos de cuaternión de Zernike para diferentes ordenes n, m .

n	m	Imagen original	T_0	T_1	T_2	T_3	T_4	T_5	T_6
0	0	0.0572	0.0572	0.0572	0.0572	0.0572	0.0572	0.0572	0.0572
1	1	0.0141	0.0141	0.0141	0.0141	0.0141	0.0141	0.0141	0.0141
2	0	0.0041	0.0041	0.0041	0.0041	0.0041	0.0041	0.0041	0.0041
2	2	0.0115	0.0115	0.0115	0.0115	0.0115	0.0115	0.0115	0.0115
3	1	0.0065	0.0065	0.0065	0.0065	0.0065	0.0065	0.0065	0.0065
3	3	0.0032	0.0032	0.0032	0.0032	0.0032	0.0032	0.0032	0.0032
4	0	0.0004	0.0004	0.0004	0.0004	0.0004	0.0004	0.0004	0.0004
4	2	0.0024	0.0024	0.0024	0.0024	0.0024	0.0024	0.0024	0.0024
4	4	0.0082	0.0082	0.0082	0.0082	0.0082	0.0082	0.0082	0.0082
5	1	0.0052	0.0052	0.0052	0.0052	0.0052	0.0052	0.0052	0.0052

4.4.3. Descriptores de color

Para imágenes en color se realiza un análisis de los tres canales R, G y B. El histograma de cada canal es normalizado en el rango de 0 a 1 y son utilizados como descriptores de color, los cuales poseen la característica de ser invariantes a la rotación

4. Momentos radiales de imágenes en color

y escala. Un histograma muestra gráficamente la distribución de una variable numérica continua midiendo la frecuencia con la que determinados valores aparecen en la imagen. En un histograma, el eje x es una línea numérica que se ha dividido en rangos de números o bins. Para cada bin, se dibuja una barra en la que el ancho de la barra representa el rango del bin y la altura de la barra representa el número de puntos de datos incluidos en ese rango [3]. Los autores dividen los bins por el tamaño de la imagen para obtener el histograma normalizado.

Invariancia a la rotación, escala y traslación

Para demostrar la invariancia a la rotación del histograma de la imagen, se usan las imágenes de

la Fig. 4.1 . Se calculó el histograma de cada objeto. La Fig. 4.9 muestra un ejemplo de los histogramas del objeto 74_0 de la base de datos COIL-100, rotada y escalada. El error medio cuadrático MSE se calculó con 10 bins para cada canal de color en el espacio de color RGB, los resultados se muestra en las Tabla 4.12 con lo que se demuestra la invariancia a la rotación del histograma. Para demostrar la invariancia a la escala el mismo objeto 74_0 de la base de datos COIL-100 con un tamaño original de 128×128 se realizó un escalado en el programa Matlab versión 9.1 R2016b con la función *imresize*, con el parámetro *'bilinear'*. La imágenes escaladas se muestran en la Fig. 4.10, y el MSE se muestra en las Tabla 4.13.

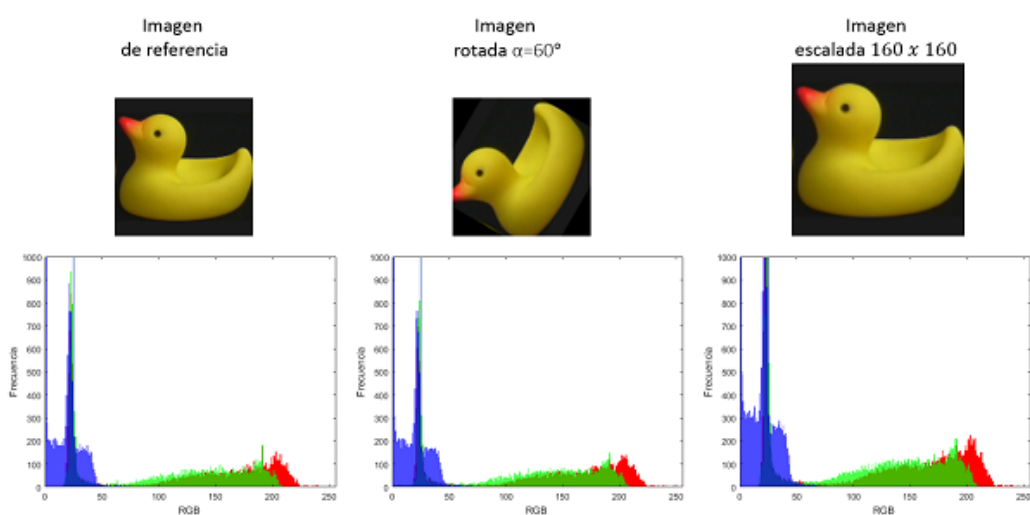
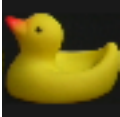


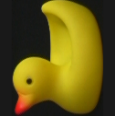
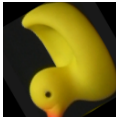




Figura 4.9: Histogramas del objeto 74__0 de la base de datos COIL-100, rotada y escalada.

4. Momentos radiales de imágenes en color

Tabla 4.12: MSE para evaluar la efectividad de la invariante de rotación del histograma.

Imagen original	$\alpha = 30^\circ$	$\alpha = 60^\circ$	$\alpha = 90^\circ$	$\alpha = 120^\circ$	$\alpha = 150^\circ$	$\alpha = 180^\circ$
						
MSE	0.0097	0.0097	0	0.0097	0.0097	0

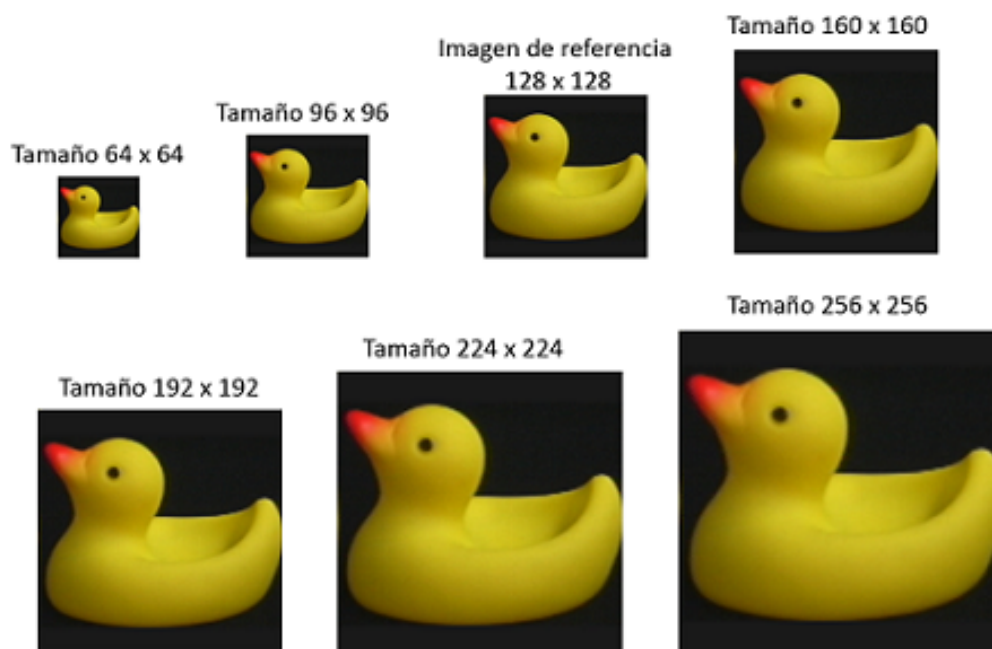


Figura 4.10: Objeto 74_0 de la base de datos COIL-100 con diversas escalas.

Tabla 4.13: Error cuadrático medio MSE para evaluar la efectividad de la invariante de escala del histograma.

Tamaño	64×64	96×96	128×128	160×160	192×192	224×224	256×256
MSE	8.5435e-05	1.8168e-05	0	6.3322e-06	7.0413e-06	5.7677e-06	8.9264e-06

4. Momentos radiales de imágenes en color

Para demostrar la invariancia a la traslación del histograma de la imagen en color, se tomó el objeto 74_0 con un tamaño original de 128×128 y se trasladó en varias posiciones dentro de una imagen de fondo con un tamaño de 300×300 como se muestra en la Fig. 4.11 posteriormente se calcula el centroide de todas las imágenes en el programa Matlab versión 9.1 R2016b y se recorta al tamaño original de 128×128 . Un ejemplo de los valores obtenidos para los histogramas normalizados se muestran en la Tabla 4.14.

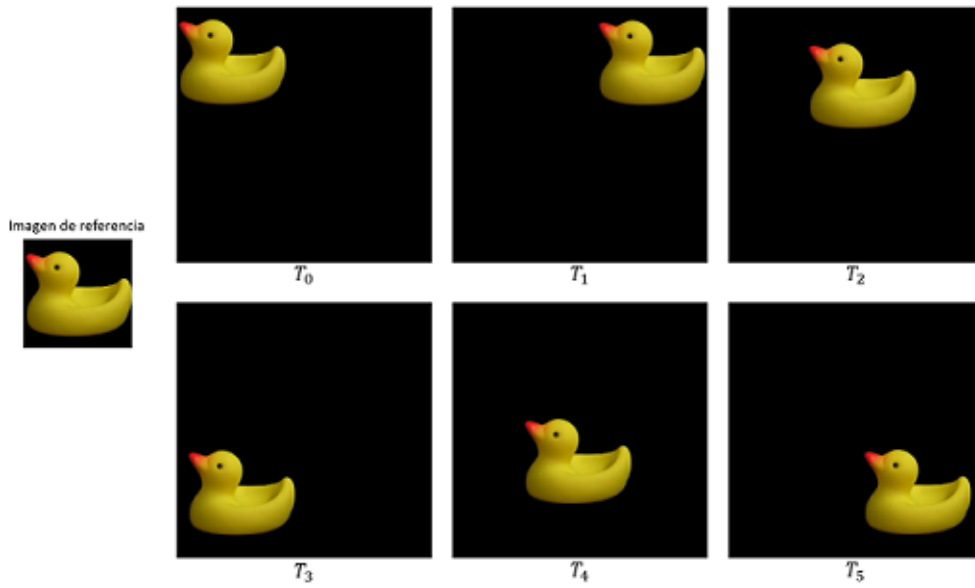


Figura 4.11: Objeto 74_0 de la base de datos Coil 100.

Tabla 4.14: Valores normalizados del histograma del objeto 74.

	Imagen original	T_0	T_1	T_2	T_3	T_4	T_5
R	0.6409	0.6409	0.6409	0.6409	0.6409	0.6409	0.6409
R	0.0146	0.0146	0.0146	0.0146	0.0146	0.0146	0.0146
R	0.0185	0.0185	0.0185	0.0185	0.0185	0.0185	0.0185
R	0.0246	0.0246	0.0246	0.0246	0.0246	0.0246	0.0246
G	0.6426	0.6426	0.6426	0.6426	0.6426	0.6426	0.6426
G	0.0312	0.0312	0.0312	0.0312	0.0312	0.0312	0.0312
G	0.0255	0.0255	0.0255	0.0255	0.0255	0.0255	0.0255
G	0.0689	0.0689	0.0689	0.0689	0.0689	0.0689	0.0689
B	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
B	0.3424	0.3424	0.3424	0.3424	0.3424	0.3424	0.3424
B	0.0312	0.0312	0.0312	0.0312	0.0312	0.0312	0.0312
B	0.0042	0.0042	0.0042	0.0042	0.0042	0.0042	0.0042

4. Momentos radiales de imágenes en color

Debido a que la concatenación de las características discriminativas que se originan de diferentes modalidades plantea una restricción en el desempeño de reconocimiento, se utilizan los métodos de kernel. Del mismo modo que las máquinas de vectores de soporte, éste método realiza un mapeo del espacio de características de entrada en un espacio de alta dimensión y construye un hiperplano para tomar una decisión binaria sobre la clase de un nuevo vector de características. Se asocia un kernel a cada modalidad de características, lo que se denomina aprendizaje de múltiple kernel.

Bibliografía

- [1] Chen, B. J., et al. "Quaternion Zernike moments and their invariants for color image analysis and object recognition." *Signal processing* 92.2 (2012): 308-318.
- [2] Singh, Chandan, and Jaspreet Singh. "Multi-channel versus quaternion orthogonal rotation invariant moments for color image representation." *Digital Signal Processing* 78 (2018): 376-392.
- [3] Singh, Chandan, and Jaspreet Singh. "Geometrically invariant color, shape and texture features for object recognition using multiple kernel learning classification approach." *Information Sciences* 484 (2019): 135-152.
- [4] Hamilton, William Rowan. "On quaternions; or on a new system of imaginaries in algebra." *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science* 25.163 (1844): 10-13.
- [5] Hosny, Khalid M., and Mohamed M. Darwish. "New set of quaternion moments for color images representation and recognition." *Journal of Mathematical Imaging and Vision* 60.5 (2018): 717-736.
- [6] Suk, Tomáš, and Jan Flusser. "Affine moment invariants of color images." *International Conference on Computer Analysis of Images and Patterns*. Springer, Berlin, Heidelberg, 2009.
- [7] Rivera-López, J. S., and C. J. Camacho-Bello. "Color Image Reconstruction by Discrete Orthogonal Moment." *Journal of Data Analysis and Information Processing* 5.04 (2017): 156.
- [8] Singh, Chandan, and Jaspreet Singh. "Multi-channel versus quaternion orthogonal rotation invariant moments for color image representation." *Digital Signal Processing* 78 (2018): 376-392.
- [9] Hosny, Khalid M., and Mohamed M. Darwish. "New set of multi-channel orthogonal moments for color image representation and recognition." *Pattern Recognition* 88 (2019): 153-173

BIBLIOGRAFÍA

- [10] Di Zenzo, Silvano. “A note on the gradient of a multi-image.” *Computer vision, graphics, and image processing* 33.1 (1986): 116-125.

Capítulo 5

Análisis de imágenes en color con momentos radiales

5.1. Introducción

En el presente capítulo se realiza una revisión del análisis de imágenes en color con momentos radiales con la finalidad de comparar los resultados publicados con una nueva implementación de los momentos Multicanal de Zernike y los momentos Multicanal de Tchebichef.

Hosny y Darwish presentan un nuevo conjunto de momentos Multicanal y sus invariantes de rotación, escala y traslación para la representación y el reconocimiento de imágenes en color [1]. Los momentos Multicanal propuestos tienen como kernel a los polinomios de Tchebichef y están definidos en coordenadas polares sobre un círculo unitario. Realizan una serie de experimentos comparando su rendimiento con los momentos Cuaternión existentes y los momentos Multicanal de Zernike propuestos por Singh y Singh [2]. Hosny y su coautor confirman nuevamente la superioridad de los momentos Multicanal frente a los momentos Cuaternión; de igual manera enfatizan que en la publicación realizada por Singh y Singh sólo se evaluaron dos métodos: los momentos Multicanal y los momentos Cuaternión de Zernike. De igual forma argumentan la inestabilidad de los momentos de Zernike a partir del orden 20 y la omisión en la reconstrucción utilizando los momentos propuestos. Los experimentos se realizan en el modelo de color RGB bajo las mismas condiciones descritas en la Ref [1] .

5.2. Momentos Multicanal de Zernike y Tchebichef

Los momentos Multicanal quedan definidos como un conjunto de momentos de cada canal de color de una imagen $f_p(r, \theta)$, donde $f_R(r, \theta), f_G(r, \theta), f_B(r, \theta)$ son las funciones de intensidad del canal rojo, verde y azul de la imagen RGB. Los momentos de cada canal se calculan como

$$M_{pq}(f_C) = \frac{1}{2\pi a_p} \int_0^{2\pi} \int_0^1 f_C(\hat{r}, \theta) P_p(\hat{r}) \bar{W}(\hat{r}) e^{-iq\theta} \hat{r} d\hat{r} d\theta, \quad (5.1)$$

donde $C \in \{R, G, B\}$ y P_p pueden ser los polinomios de Zernike R_p o los polinomios de Tchebichef \bar{R}_p . Al igual que en la publicación el número total de momentos Multicanal de Tchebichef (MMT) calculados es igual a,

$$(1 + C_{\text{máx}})^2,$$

donde $C_{\text{máx}}$ se refiere al orden máximo. Para los momentos Multicanal de Zernike (MMZ) el número total de momentos esta dado por,

$$\begin{aligned} & \left(\frac{2+Z_{\text{máx}}}{2}\right)^2 && Z_{\text{máx}} \text{ par,} \\ & \left(\frac{1+Z_{\text{máx}}}{2}\right) \left(\frac{3+Z_{\text{máx}}}{2}\right) && Z_{\text{máx}} \text{ impar,} \end{aligned} \quad (5.2)$$

donde $Z_{\text{máx}}$ se refiere al orden máximo. De acuerdo a estas fórmulas en el orden 4, por ejemplo el artículo compara 9 descriptores de cada canal de los MMZ de color contra 25 descriptores de cada canal en los MMT. En la Tabla 5.1 se muestra un ejemplo de los descriptores utilizados de acuerdo al orden.

Tabla 5.1: Descriptores obtenidos de acuerdo al orden de los momentos MMZ y MMT.

orden	Núm. total de MMZ	Núm. total de MMT
4	9R,9G,9B	25R,25G,25B
6	16R,16G,16B	49R,49G,49B
8	25R,25G,25B	81R,81G,81B

Los momentos Multicanal de Zernike han sido descritos anteriormente en el capítulo 4. Para los momentos Multicanal de Tchebichef se utilizan la propuesta por Hosny y Darwish donde las funciones radiales sustituidas $\bar{R}_p(\hat{r})$ se definen como [1],

$$\bar{R}_p(\hat{r}) = \sum_{k=0}^p \frac{(k!)^2}{(2k)!} \binom{p+k-1}{k} \binom{p}{k} (-4)^k (1-\hat{r})^k \quad (5.3)$$

para $p \geq 1$

5. Análisis de imágenes en color con momentos radiales

La función de peso sustituido esta definida como,

$$\bar{W}(\hat{r}) = \frac{(1 - \hat{r})^{3/2}}{(1 - \hat{r})^2 \sqrt{\hat{r}}} = (\hat{r} - \hat{r}^2)^{-1/2} = \frac{1}{\sqrt{(\hat{r} - \hat{r}^2)}} \quad (5.4)$$

las funciones radiales sustituidas de Tchebichef $\bar{R}_p(r)$ se definen utilizando la relación de recurrencia,

$$\bar{R}_{p+1}(\hat{r}) = 2(\hat{r} - 1)\bar{R}_p(\hat{r}) - \bar{R}_{p-1}(\hat{r}), \quad (5.5)$$

Para $p \geq 1$, donde las dos primeras funciones son:

$$\begin{aligned} \bar{R}_0(\hat{r}) &= 1, \\ \bar{R}_1(\hat{r}) &= 2\hat{r} - 1, \end{aligned}$$

siendo ortogonales sobre un disco de radio unitario y satisfacen la relacion de ortogonalidad:

$$\int_0^1 \bar{R}_p(\hat{r}) \bar{R}_q(\hat{r}) \bar{W}(\hat{r}) d\hat{r} = \frac{C_p \pi}{2} \delta_{pq}. \quad (5.6)$$

5.3. Experimentos

Los experimentos se realizan bajo las mismas condiciones detalladas en la publicación. En la primera subsección se evalúa la capacidad de reconstrucción de imágenes RGB de los momentos propuestos. En la segunda subsección se realizan los experimentos para evaluar los invariantes de rotación, escala y traslación. En la tercera subsección se evalúa la robustez frente a distintos tipos de ruido, y en la cuarta subsección se evalúa el reconocimiento de imágenes.

5.3.1. Reconstrucción de imágenes.

La reconstrucción de imágenes evalúa la precisión de los momentos para representar una imagen. Mientras un orden menor tendrá la capacidad de representar características generales de la imagen como área o forma, los órdenes superiores poseen la capacidad de representar detalles de ésta. Por lo tanto la estabilidad numérica evitará degradaciones de la imagen en órdenes superiores. La precisión en la reconstrucción se evalúa por medio del NIRE definido en la Ec. (3.20) y la Ec. (4.17). La imagen de “Lena” de la Fig 5.1 se utiliza para reconstruir la imagen utilizando los momentos Multicanal de Zernike (MMZ) y los momentos Multicanal de Tchebichef (MMT). Los resultados del artículo se muestran en la Tabla 5.2 y los obtenidos en el experimento se muestran en la Tabla 5.3.

5. Análisis de imágenes en color con momentos radiales

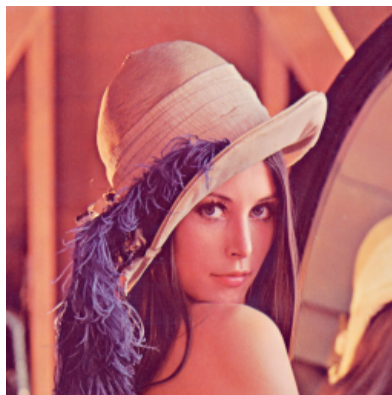


Figura 5.1: Imagen a color “Lena”.

Tabla 5.2: Reconstrucción de la imagen Lena publicados por Hosny y Darwish [1].

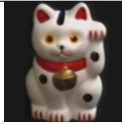


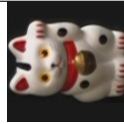

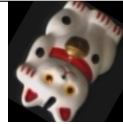
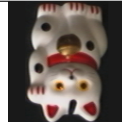
orden	10	20	30	40	50	100	140	150
MMZ								
	0.0524	0.0313	0.0228	0.0207	1.26e+05			
MMT								
	0.0356	0.0181	0.0116	0.0079	0.0056	0.0020	0.0017	0.0016

Tabla 5.3: Reconstrucción de la imagen Lena con momentos Multicanal de Zernike y Tchebichef.

orden	10	20	30	40	50	100	140	150
MMZ								
	0.0410	0.0233	0.0156	0.0109	0.0075	0.0015	7.5648e-04	7.1287e-04
MMT								
	0.0351	0.0172	0.0105	0.0068	0.0046	0.0010	8.1652e-04	8.7600e-4

5. Análisis de imágenes en color con momentos radiales

Tabla 5.4: MSE para evaluar la efectividad de la invariante de rotación de los momentos de Cuaternión de Zernike y los momentos Multicanal de Tchebichef.

Imagen original	$\alpha = 30^\circ$	$\alpha = 60^\circ$	$\alpha = 90^\circ$	$\alpha = 120^\circ$	$\alpha = 150^\circ$	$\alpha = 180^\circ$
						
MSE Hosny [1]	6.62e-05	7.01e-05	2.53e-30	7.56e-05	7.80e-05	2.13e-30
MSE MMZ	3.4638e-09	3.6029e-09	7.6117e-35	3.4638e-09	3.6029e-09	4.4749e-35
MSE MMT	4.2463e-08	4.2802e-08	1.1138e-34	4.2463e-08	4.2802e-08	1.1520e-34

5.3.2. Invariancia a transformaciones geométricas

Las transformaciones geométricas de una imagen incluyen rotación, escala y traslación. Las pruebas se realizan con la base de datos COIL-100 conformada por 7200 imágenes de tamaño 128×128 con 100 objetos distintos los cuales son rotados de 0° a 360° con un incremento de ángulo de 5° . Para evaluar la efectividad de los invariantes se utiliza el error cuadrático medio MSE, el cual es una medida cuantitativa que refleja la precisión del invariante de rotación, escala y traslación, el cual ha sido definido anteriormente en la Ec. (3.6).

5.3.3. Invariancia a la rotación

Para probar la invariancia a la rotación se utiliza el obj_14 girado en ángulos de 0° a 180° con un incremento fijo de 30° . Los resultados obtenidos y los publicados por Hosny y Darwish se muestran en la Tabla 5.4.

5.3.4. Invariancia a la escala.

El segundo experimento se realiza con el obj_17 con seis factores de escala $\lambda = 0,5, 0,75, 1,25, 1,5, 1,75$ y $2,0$ como se muestra en la Fig. 5.2. Se calculan los momentos Multicanal de Zernike y Tchebichef de orden 20. Los resultados se muestran en la Tabla 5.5

Tabla 5.5: Error cuadrático medio MSE para evaluar la efectividad de la invariante de escala.

Tamaño	64×64	96×96	128×128	160×160	192×192	224×224	256×256
MSE [1]	7.55e-05	1.17e-05	0	5.17e-06	9.03e-06	9.97e-06	1.08e-06
MSE MMZ	3.9405e-08	1.0183e-08	0	4.7337e-09	2.5015e-09	4.8182e-09	4.5425e-09
MSE MMT	1.4637e-07	3.5918e-08	0	1.5327e-08	9.8170e-09	1.5490e-08	1.5276e-08

5. Análisis de imágenes en color con momentos radiales

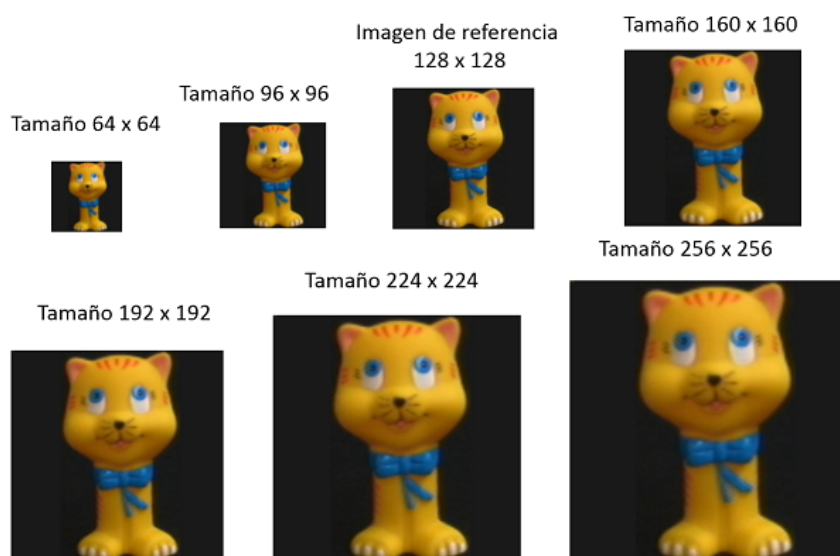


Figura 5.2: Imagen de referencia y escalada del Obj_17 a color de la base de datos COIL-100.

5.3.5. Invariancia a la traslación

En el tercer experimento se evalúa la invariancia a la traslación. El Obj_23 es trasladado en cinco posiciones distintas, la Fig. 5.3 muestra la imagen original T_0 y las trasladadas $T_1, T_2, T_3, T_4, y T_5$. Se calculan los momentos Multicanal de Zernike y Tchebichef, los resultados obtenidos y los publicados por Hosny y Darwish se muestran en la Tabla 5.6.

Tabla 5.6: Error cuadrático medio MSE para evaluar la efectividad de la invariante de traslación.

Posición	T_0	T_1	T_2	T_3	T_4	T_5
MSE Hosny [1]	0	9.93e-07	1.30e-07	1.61e-07	1.34e-07	1.93e-07
MSE MMZ	0	3.1756e-06	3.1756e-06	3.1756e-06	3.1756e-06	3.1756e-06
MSE MMT	0	3.4135e-05	3.4135e-05	3.4135e-05	3.4135e-05	3.4135e-05

5.3.6. Robustez al ruido

Se evalúa la robustez al ruido mediante el cálculo de los momentos Multicanal de Zernike y Tchebichef del Obj_25 al cual se le agrega ruido, del tipo sal y pimienta, Gaussiano, Poisson y speckle; como puede observarse en la Fig 5.4 . Mientras en el experimento realizado por Hosny y Darwish comparan imagen con ruido vs. imagen

5. Análisis de imágenes en color con momentos radiales

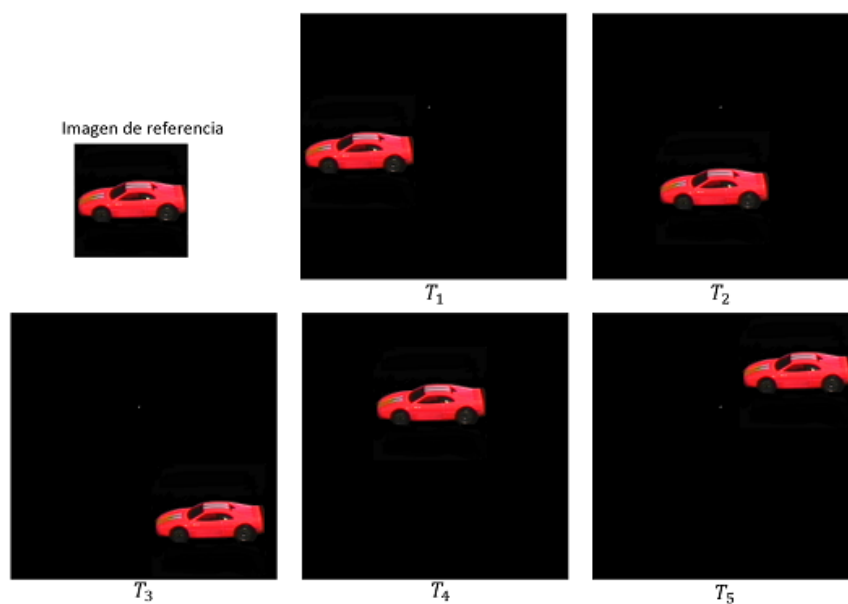


Figura 5.3: Obj_23 de la base de datos COil-100 en posición original y trasladado.

con ruido reconstruida como se muestra en la Fig. 5.5. En el experimento llevado a cabo se reconstruyen de igual forma con órdenes de 10 a 120 comparando las imágenes reconstruidas con ruido, con la imagen original del Obj_25 sin ruido. Los resultados obtenidos pueden observarse en la Fig. 5.6 para los momentos Multicanal de Zernike y en la Fig 5.7 para los momentos Multicanal de Tchebichef.

5. Análisis de imágenes en color con momentos radiales

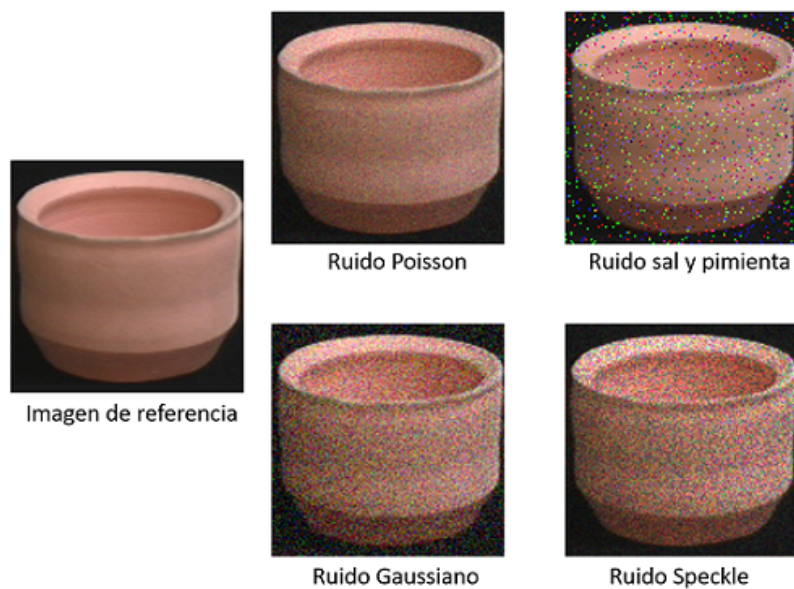


Figura 5.4: Obj_25 de la base de datos COIL-100 libre de ruido y con ruido añadido.

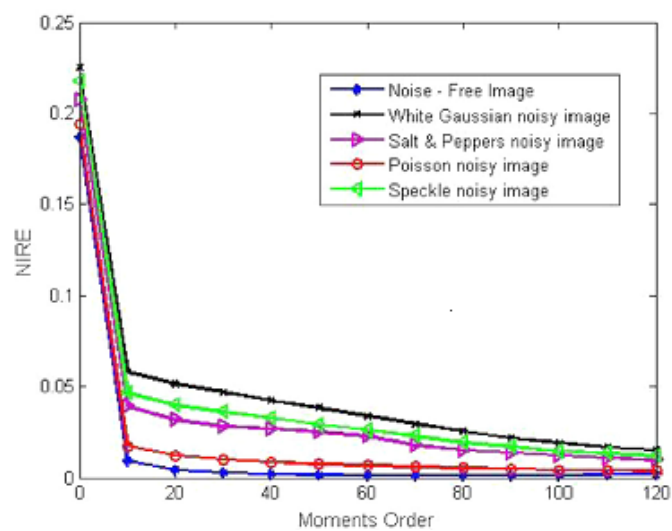


Figura 5.5: NIRE para imágenes libre de ruido y con ruido del Obj_25 obtenidos por Hosny y Darwish [1].

5. Análisis de imágenes en color con momentos radiales

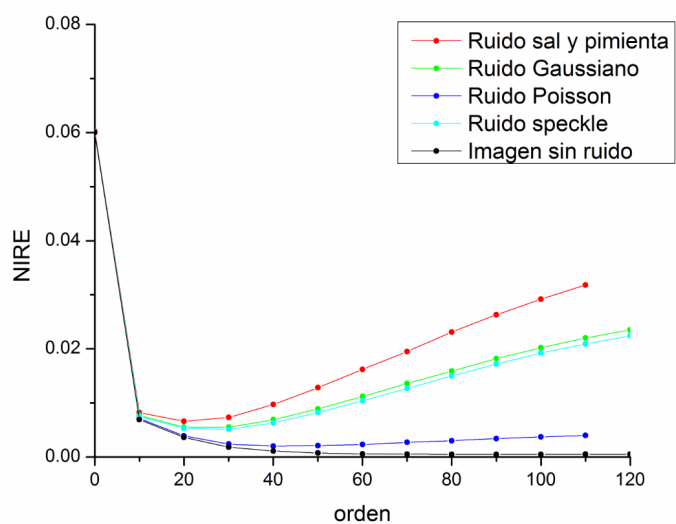


Figura 5.6: NIRE obtenido por nosotros para los momentos Multicanal de Zernike para el Obj_25.

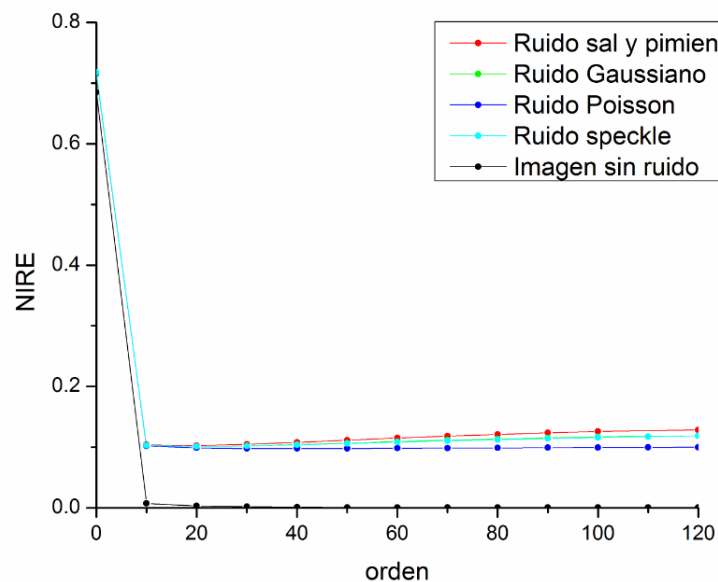


Figura 5.7: NIRE obtenido por nosotros para los momentos Multicanal de Tchebichef para el Obj_25.

5. Análisis de imágenes en color con momentos radiales

5.3.7. Reconocimiento de imágenes

Se replican los experimentos realizados con la base de datos COIL-100 separando la base de datos bajo las mismas condiciones. Se agrupan en dos conjuntos de imágenes, el primer conjunto de galería, consta de las 100 imágenes de frente de cada clase, y el segundo conjunto el cual será de prueba son los 71 objetos rotadas pertenecientes a cada clase. La tasa de reconocimiento se define como $R(\%)$ definido como $P \times 100/Q$ donde P es el número total de imágenes correctamente clasificadas y Q es el número de imágenes utilizadas en la prueba. Se usan medidas de similitud basadas en distancia entre vectores [1],

$$L_1 - norm : d_{L_1} = \sum_{i=1}^L |p_i - q_i| \quad (5.7)$$

$$L_2 - norm : d_{L_2} = \sum_{i=1}^L \sqrt{(p_i - q_i)^2} \quad (5.8)$$

$$\chi^2 : d_{\chi^2} = \sum_{i=1}^L \frac{(p_i - q_i)^2}{p_i + q_i} \quad (5.9)$$

$$Canberra : d_{Can} = \sum_{i=1}^L \frac{|p_i - q_i|^2}{|p_i| + |q_i|} \quad (5.10)$$

donde L es el tamaño de vector de características, P es la imagen del primer conjunto, Q la imagen de prueba, p_i y q_i representan los descriptores de cada vector.

El primer experimento de reconocimiento de imágenes se realiza con la base de datos COIL-100 en condiciones normales. Se calculan los MMZ y MMT para los órdenes, 4, 6, 8, 10, 12, 14 y 16, la tasa de reconocimiento se evalúa de acuerdo a las medidas de similitud descritas. Los resultados publicados se muestran en la Fig 5.8. Nuestros resultados se muestran en la Fig. 5.9 para los MMZ y MMT calculados.

El segundo experimento se realiza con la base de datos COIL-100 en condiciones normales. Para este experimento y los posteriores se igualan a 60 los descriptores tanto para los MMZ y los MMT. Los resultados del experimento realizado y los publicados por Hosny se observan en la Tabla 5.7.

El tercer experimento se realiza comparando el primer conjunto de imágenes en condiciones normales y el segundo conjunto de imágenes de prueba rotado de forma aleatoria en ángulos de 0° a 180° con un incremento fijo de 30° . Los resultados del experimento realizado y los publicados por Hosny se observan en la Tabla 5.8

En el cuarto experimento, se compara el primer conjunto en condiciones normales y el segundo conjunto con seis factores de escala $\lambda = 0,5, 0,75, 1,25, 1,5, 1,75$ y $2,0$ es decir se obtienen 6 conjuntos de prueba y se compara cada uno con el conjunto

5. Análisis de imágenes en color con momentos radiales

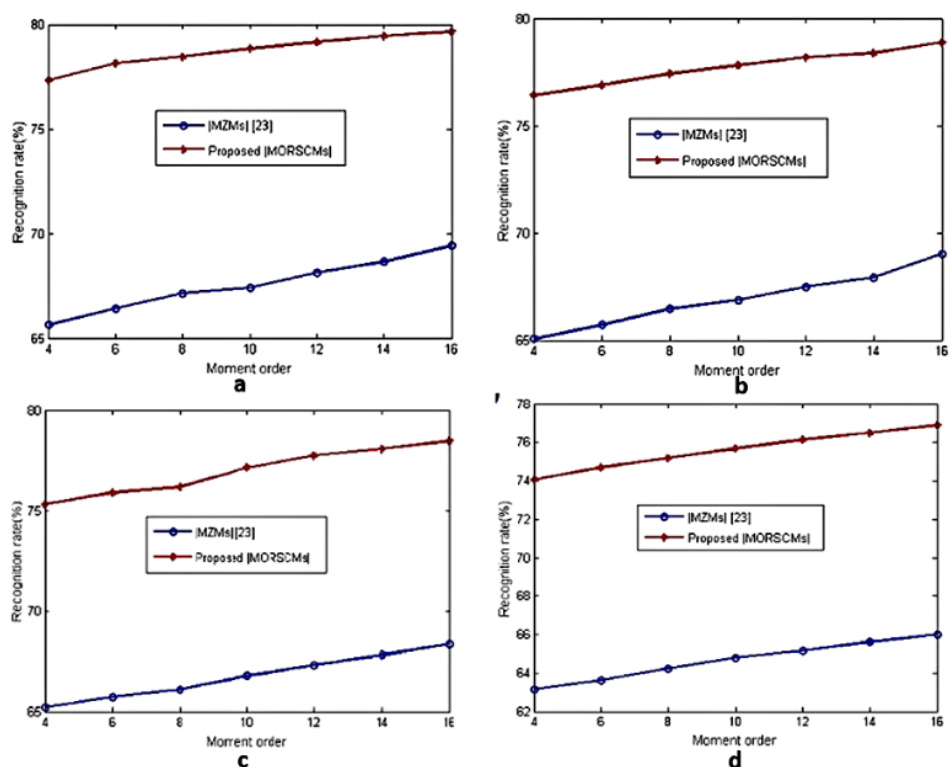


Figura 5.8: Tasa de reconocimiento $R(\%)$ usando distintas medidas a) $L_1 - norm$, b) $L_2 - norm$ c) χ^2 y d) *Canberra* publicadas por Hosny y Darwish [1].

Tabla 5.7: Tasa de Reconocimiento de los momentos MMZ y MMT con 60 descriptores.

Medida	Método			
	MMZ	MMT	MMZ Hosny y Darwish [1]	MMT Hosny y Darwish [1]
$L_1 - norm$	68.6901	59.5634	67.03	79.12
$L_2 - norm$	68.3944	54.1408	65.93	78.02
χ^2	63.1408	49.1408	66.41	79.12
<i>Canberra</i>	61.8310	49.0563	63.13	77.12

Tabla 5.8: Tasa de Reconocimiento de los momentos MMZ y MMT con 60 descriptores de la base de datos rotada aleatoriamente.

Medida	Método			
	MMZ	MMT	MMZ Hosny y Darwish [1]	MMT Hosny y Darwish [1]
$L_1 - norm$	68.7042	49.6338	66.76	78.64
$L_2 - norm$	68.3944	25.7465	65.66	77.86
χ^2	63.1690	23.2254	66.24	79.64
<i>Canberra</i>	61.8310	34.6620	62.92	76.82

5. Análisis de imágenes en color con momentos radiales

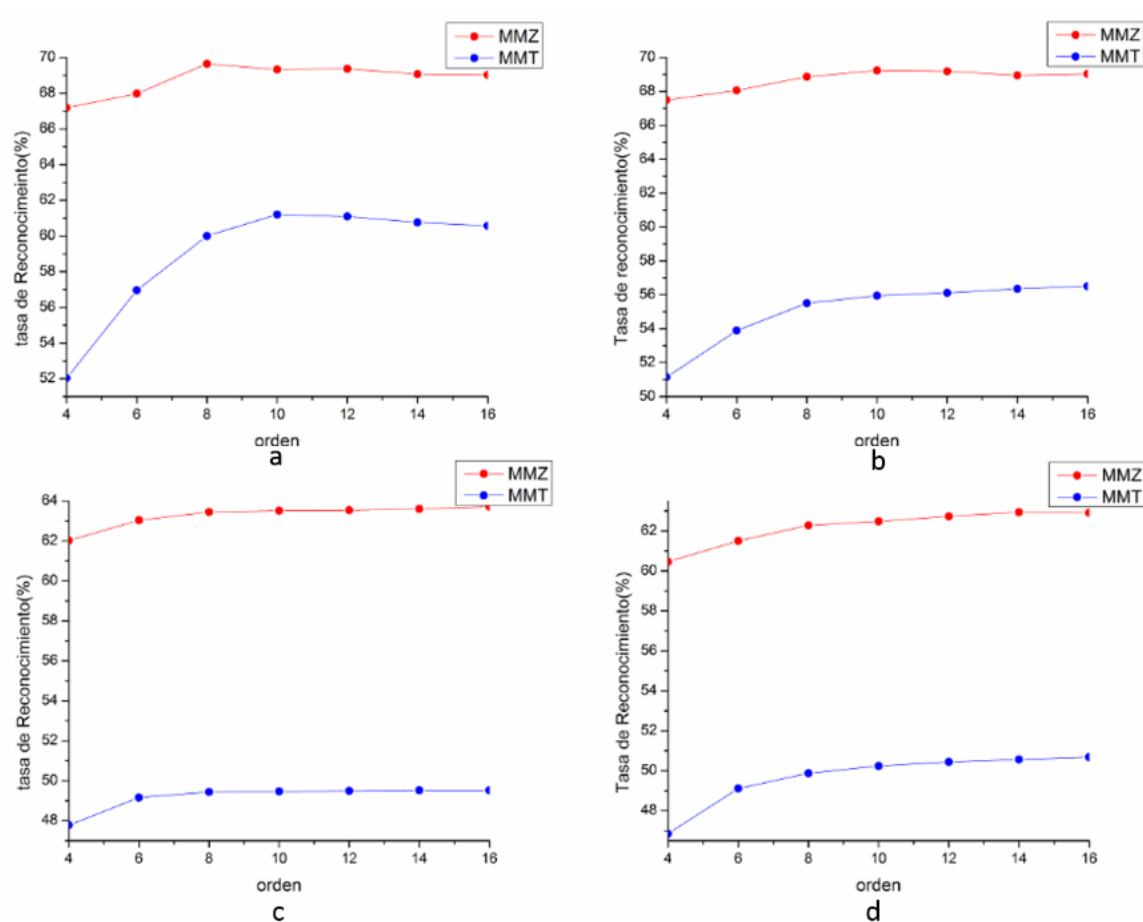


Figura 5.9: Tasa de reconocimiento R (%) usando distintas medidas a) $L_1 - norm$, b) $L_2 - norm$ c) χ^2 y d) *Canberra* obtenido por nosotros.

5. Análisis de imágenes en color con momentos radiales

galería en condiciones normales. La Tabla 5.9 muestra el promedio de las 6 mediciones realizadas y los resultados publicados.

Tabla 5.9: Promedio de la Tasa de Reconocimiento de la base de datos COIL-100 en distintas escalas.

Medida	Método			
	MMZ	MMT	MMZ Hosny y Darwish [1]	MMT Hosny y Darwish [1]
L_1 -norm	68.6408	49.6175	67.01	79.47
L_2 -norm	68.3591	25.73	65.94	78.83
χ^2	63.1714	23.2887	66.38	79.04
Canberra	61.7159	34.6713	63.14	77.52

En el quinto experimento se compara el primer conjunto en condiciones normales y el segundo conjunto con ruido Gaussiano añadido $\mu = 0$ y $\sigma^2 = 0,05, 0,10, 0,15,$ y $0,20$. La Tabla 5.10 muestra las 4 mediciones realizadas y los resultados publicados.

Tabla 5.10: Tasa de Reconocimiento de la base de datos COIL-100 con ruido Gaussiano.

Medida	Método			
	MMZ	MMT	MMZ Hosny y Darwish [1]	MMT Hosny y Darwish [1]
$\mu = 0$ y $\sigma^2 = 0,05$				
L_1 -norm	66.7746	40.1127	41.74	70.14
L_2 -norm	65.6338	21.1127	42.04	70.02
χ^2	52.0704	19.2676	39.43	69.28
Canberra	49.2535	26.1127	35.94	67.32
$\mu = 0$ y $\sigma^2 = 0,10$				
L_1 -norm	62.4225	29.7465	40.33	69.16
L_2 -norm	54.5915	16.3099	41.29	69.78
χ^2	27.8451	13.2254	39.36	68.81
Canberra	25.3099	16.5352	33.48	65.34
$\mu = 0$ y $\sigma^2 = 0,15$				
L_1 -norm	54.6620	22.4648	37.05	65.45
L_2 -norm	36.7465	12.2958	38.10	67.41
χ^2	13.1690	8.3944	36.09	66.23
Canberra	13.0845	10.7746	31.99	64.10
$\mu = 0$ y $\sigma^2 = 0,20$				
L_1 -norm	41.1127	16.9859	34.97	63.87
L_2 -norm	20.9437	9.0563	36.01	65.74
χ^2	9.0563	6.1972	33.78	64.53
Canberra	8.4789	8	31.32	63.87

Adicionalmente con los MMZ y MMT calculados se realiza la clasificación de las imágenes en la aplicación *Classification Learner* de Matlab. Los resultados obtenidos

5. Análisis de imágenes en color con momentos radiales

para la clasificación con condiciones normales con MMZ y MMT se muestra en la Tabla 5.11. El diagrama de dispersión para las primeras 25 clases puede observarse en la Fig. 5.10 para los MMZ y en la Fig. 5.11 para los MMT.

Tabla 5.11: Exactitud para los diferentes clasificadores con momentos Multicanal de Zernike, Modelo de color RGB en condiciones normales.

Clasificador	MMZ	MMT
Complex Tree	46.7 %	48.4 %
Medium Tree	15.4 %	16.3 %
Simple Tree	4.6 %	4.9 %
Linear Discriminant	94.9 %	86.8 %
Quadratic Discriminant	98.9 %	92.4 %
Linear SVM	99.0 %	98.2 %
Quadratic SVM	99.7 %	99.2 %
Cubic SVM	99.7 %	99.3 %
Fine Gaussian SVM	78.9 %	58.1 %
Medium Gaussian SVM	99.7 %	99.1 %
Coarse Gaussian SVM	96.3 %	93.9 %
Fine KNN	99.9 %	99.3 %
Medium KNN	96.2 %	94.8 %
Coarse KNN	71.8 %	72.4 %
Cosine KNN	96.2 %	94.1 %
Cubic KNN	95.5 %	93.8 %
Weighted KNN	98.9 %	97.8 %
Ensemble/Boosted Trees	39.2 %	39.2 %
Ensemble/Bagged Trees	99.2 %	98.4 %
Ensemble/Sub Discriminant	93.7 %	91.4 %
Ensemble/Subspace KNN	99.9 %	99.2 %
Ensemble/RUSBoosted Trees	16.6 %	16.4 %

Los resultados para la clasificación con rotación aleatoria se muestran en la Tabla 5.12 y en la Fig. 5.12 se observa el Diagrama de dispersión para las 25 primeras clases con MMZ y en la Fig. 5.13 con MMT.

Como ejemplo de clasificación con escala se toma el conjunto escalado con $\lambda = 1,5$ los resultados se muestran en la Tabla 5.13 y el diagrama de Dispersión correspondiente a los MMZ se muestra en la Fig. 5.14 y en la Fig. 5.15 para los MMT.

Como ejemplo de la clasificación con ruido se toma el conjunto de prueba con ruido Gaussiano $\mu = 0$ y $\sigma^2 = 0,20$ los resultados se muestran en la Tabla 5.14, el diagrama de dispersión de la Fig. 5.16 corresponde a los MMZ y la Fig. 5.17 a los MMT.

5. Análisis de imágenes en color con momentos radiales

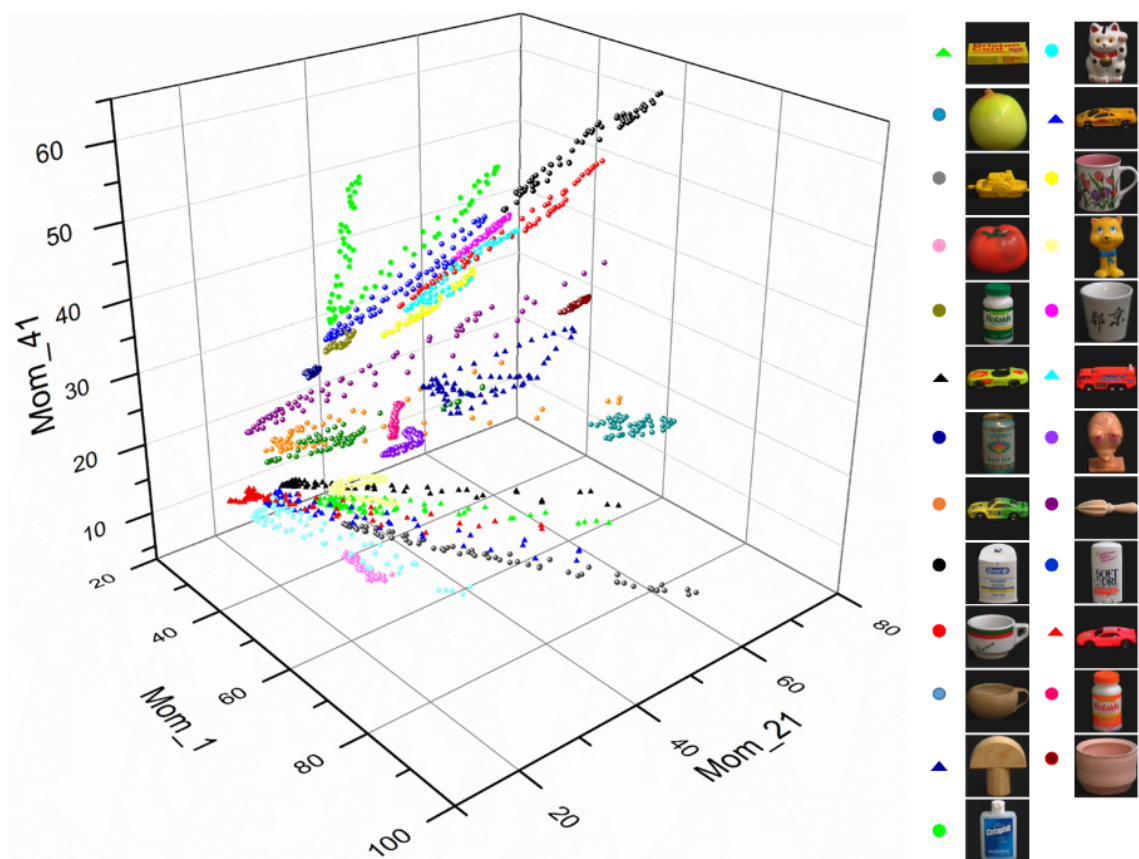


Figura 5.10: Diagrama de dispersión para las primeras 25 clases de la base de datos COIL-100 en condiciones normales con MMZ.

5. Análisis de imágenes en color con momentos radiales

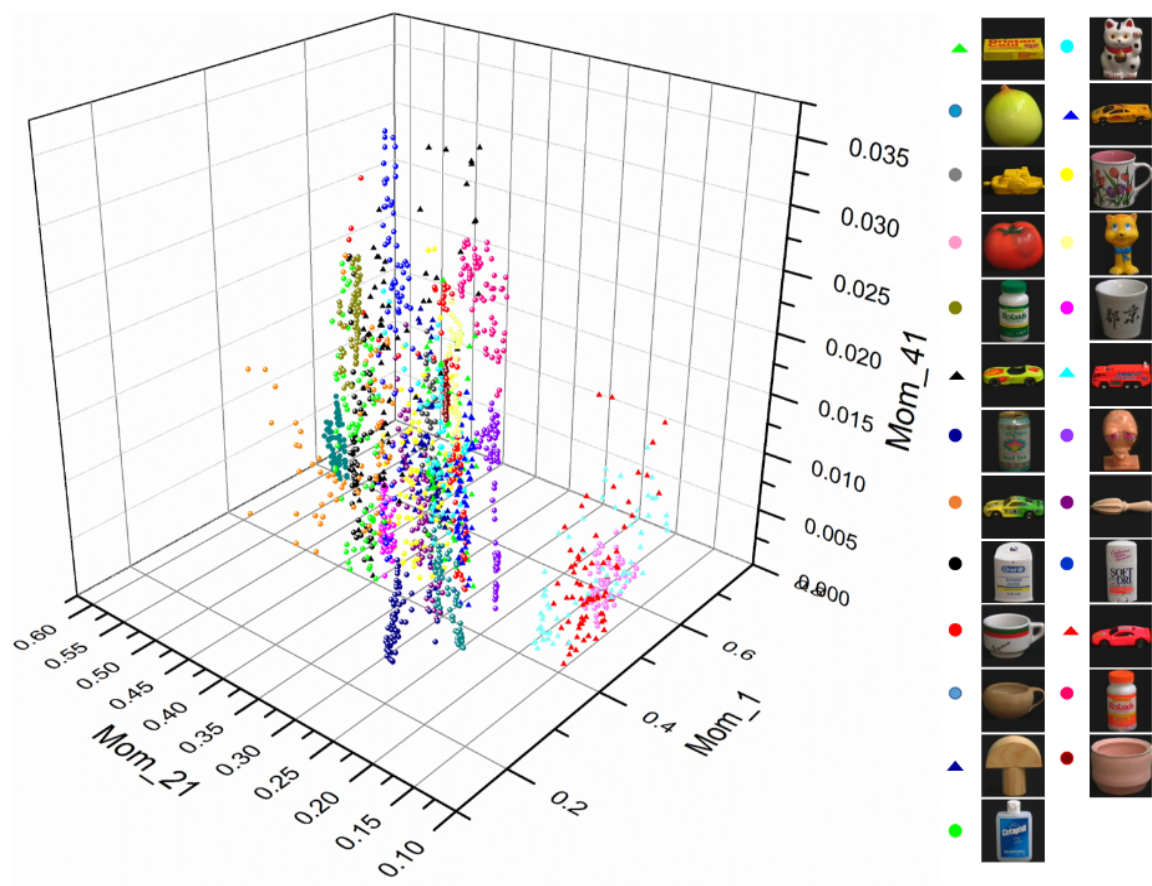


Figura 5.11: Diagrama de dispersión para las primeras 25 clases de la base de datos COIL-100 en condiciones normales con MMT.

5. Análisis de imágenes en color con momentos radiales

Tabla 5.12: Exactitud para los diferentes clasificadores con momentos Multicanal de Zernike, Modelo de color RGB con rotación aleatoria.

Clasificador	MMZ	MMT
Complex Tree	46.8 %	48.4 %
Medium Tree	15.4 %	16.3 %
Simple Tree	4.6 %	5.0 %
Linear Discriminant	89.0 %	86.5 %
Quadratic Discriminant	95.2 %	92.2 %
Linear SVM	99.0 %	98.5 %
Quadratic SVM	99.7 %	99.3 %
Cubic SVM	99.7 %	99.3 %
Fine Gaussian SVM	78.9 %	57.9 %
Medium Gaussian SVM	99.7 %	99.1 %
Coarse Gaussian SVM	96.3 %	93.9 %
Fine KNN	99.9 %	99.4 %
Medium KNN	96.2 %	94.6 %
Coarse KNN	72.0 %	72.3 %
Cosine KNN	96.2 %	94.1 %
Cubic KNN	95.5 %	93.6 %
Weighted KNN	98.9 %	97.8 %
Ensemble/Boosted Trees	39.2 %	38.0 %
Ensemble/Bagged Trees	99.2 %	98.3 %
Ensemble/Sub Discriminant	93.8 %	91.3 %
Ensemble/Subspace KNN	99.9 %	99.3 %
Ensemble/RUSBoosted Trees	16.5 %	16.8 %

5. Análisis de imágenes en color con momentos radiales

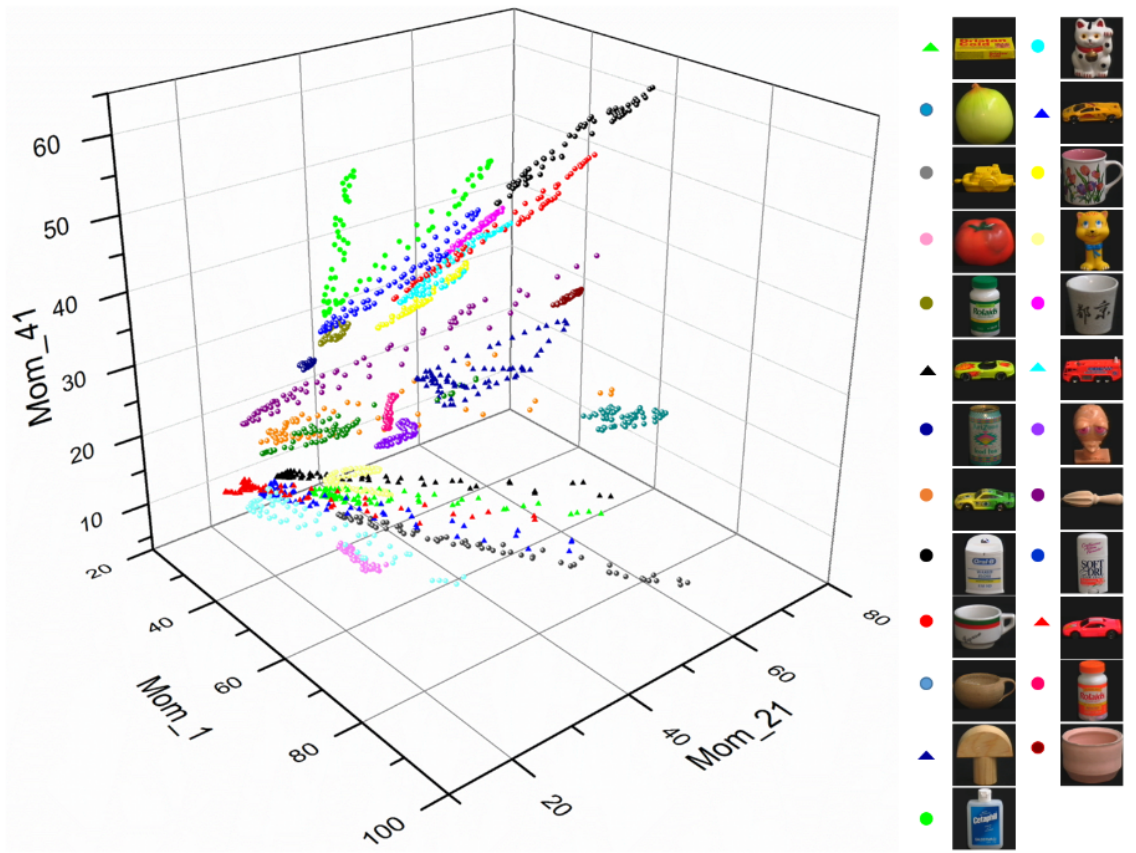


Figura 5.12: Diagrama de dispersión para las primeras 25 clases de la base de datos COIL-100 en rotación aleatoria con MMZ.

5. Análisis de imágenes en color con momentos radiales

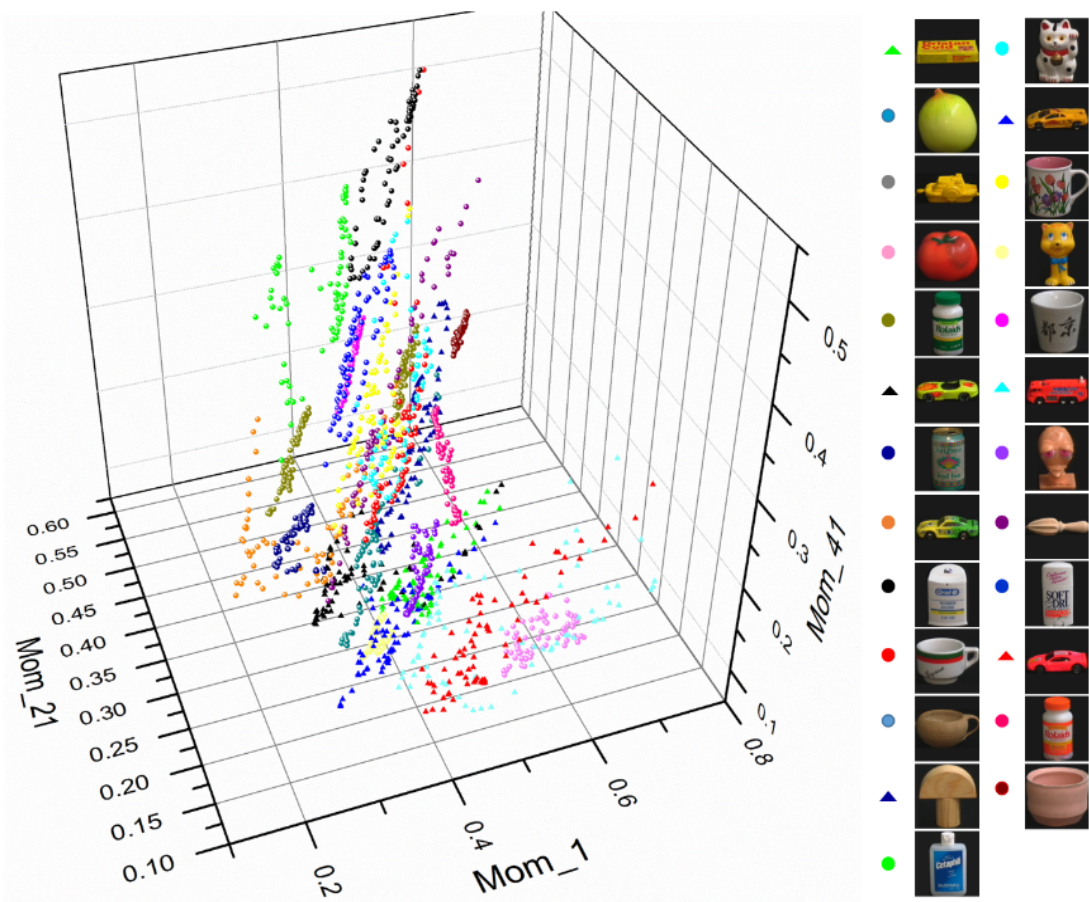


Figura 5.13: Diagrama de dispersión para las primeras 25 clases de la base de datos COIL-100 en rotación aleatoria con MMT.

5. Análisis de imágenes en color con momentos radiales

Tabla 5.13: Exactitud para los diferentes clasificadores con momentos Multicanal de Zernike, Modelo de color RGB con escala 1.5.

Clasificador	MMZ	MMT
Complex Tree	46.8 %	47.8 %
Medium Tree	16.5 %	16.4 %
Simple Tree	4.6 %	4.9 %
Linear Discriminant	88.9 %	86.6 %
Quadratic Discriminant	95.1 %	92.3 %
Linear SVM	99.1 %	98.3 %
Quadratic SVM	99.8 %	99.1 %
Cubic SVM	99.7 %	99.2 %
Fine Gaussian SVM	78.5 %	58.0 %
Medium Gaussian SVM	99.8 %	99.0 %
Coarse Gaussian SVM	96.4 %	93.9 %
Fine KNN	99.9 %	99.4 %
Medium KNN	96.0 %	94.5 %
Coarse KNN	72.2 %	72.1 %
Cosine KNN	96.2 %	94.1 %
Cubic KNN	95.7 %	93.5 %
Weighted KNN	99.0 %	97.8 %
Ensemble/Boosted Trees	38.2 %	37.9 %
Ensemble/Bagged Trees	99.1 %	98.4 %
Ensemble/Sub Discriminant	93.7 %	91.2 %
Ensemble/Subspace KNN	99.9 %	99.3 %
Ensemble/RUSBoosted Trees	16.7 %	15.9 %

5. Análisis de imágenes en color con momentos radiales

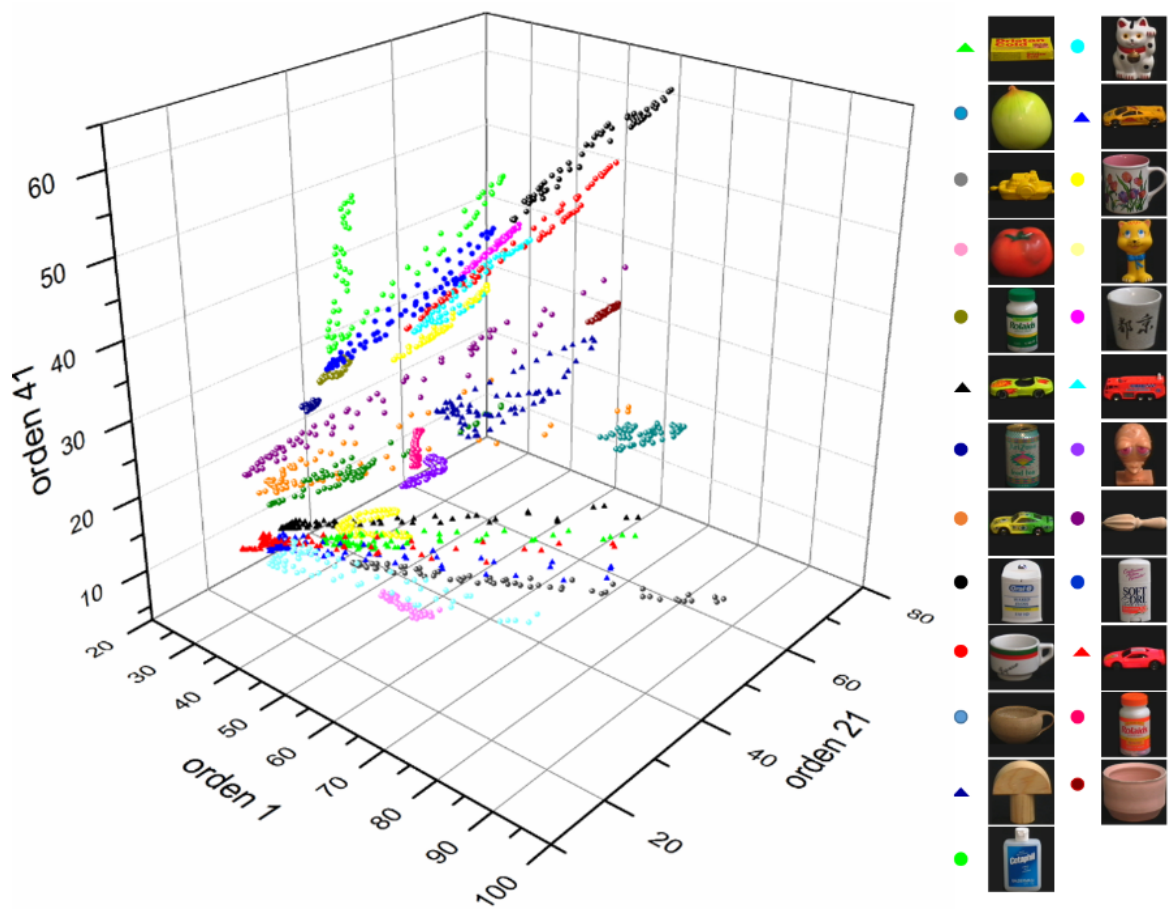


Figura 5.14: Diagrama de dispersión para las primeras 25 clases de la base de datos COIL-100 escala con MMZ.

5. Análisis de imágenes en color con momentos radiales

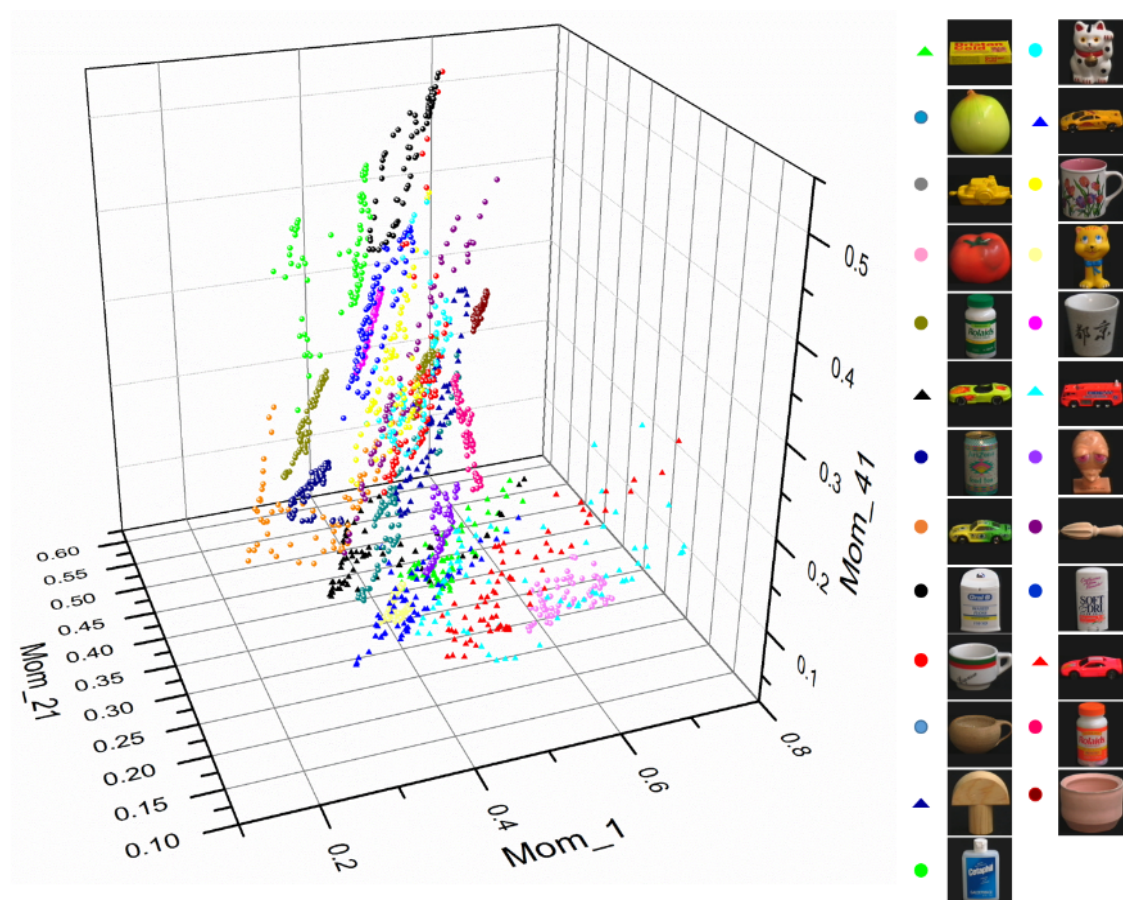


Figura 5.15: Diagrama de dispersión para las primeras 25 clases de la base de datos COIL-100 a escala con MMT.

5. Análisis de imágenes en color con momentos radiales

Tabla 5.14: Exactitud para los diferentes clasificadores con momentos Multicanal de Zernike, Modelo de color RGB con ruido Gaussiano.

Clasificador	MMZ	MMT
Complex Tree	49.0 %	34.9 %
Medium Tree	14.7 %	14.1 %
Simple Tree	4.6 %	4.2 %
Linear Discriminant	87.9 %	75.7 %
Quadratic Discriminant	94.3 %	78.2 %
Linear SVM	98.2 %	81.3 %
Quadratic SVM	99.1 %	83.2 %
Cubic SVM	99.1 %	81.9 %
Fine Gaussian SVM	5.9 %	3.5 %
Medium Gaussian SVM	98.9 %	79.0 %
Coarse Gaussian SVM	94.3 %	71.9 %
Fine KNN	99.5 %	71.3 %
Medium KNN	96.0 %	72.7 %
Coarse KNN	72.3 %	63.8 %
Cosine KNN	95.7 %	71.1 %
Cubic KNN	95.3 %	68.2 %
Weighted KNN	97.9 %	74.9 %
Ensemble/Boosted Trees	35.6 %	22.1 %
Ensemble/Bagged Trees	98.4 %	78.7 %
Ensemble/Sub Discriminant	92.6 %	80.5 %
Ensemble/Subspace KNN	99.8 %	83.8 %
Ensemble/RUSBoosted Trees	16.7 %	15.3 %

5. Análisis de imágenes en color con momentos radiales

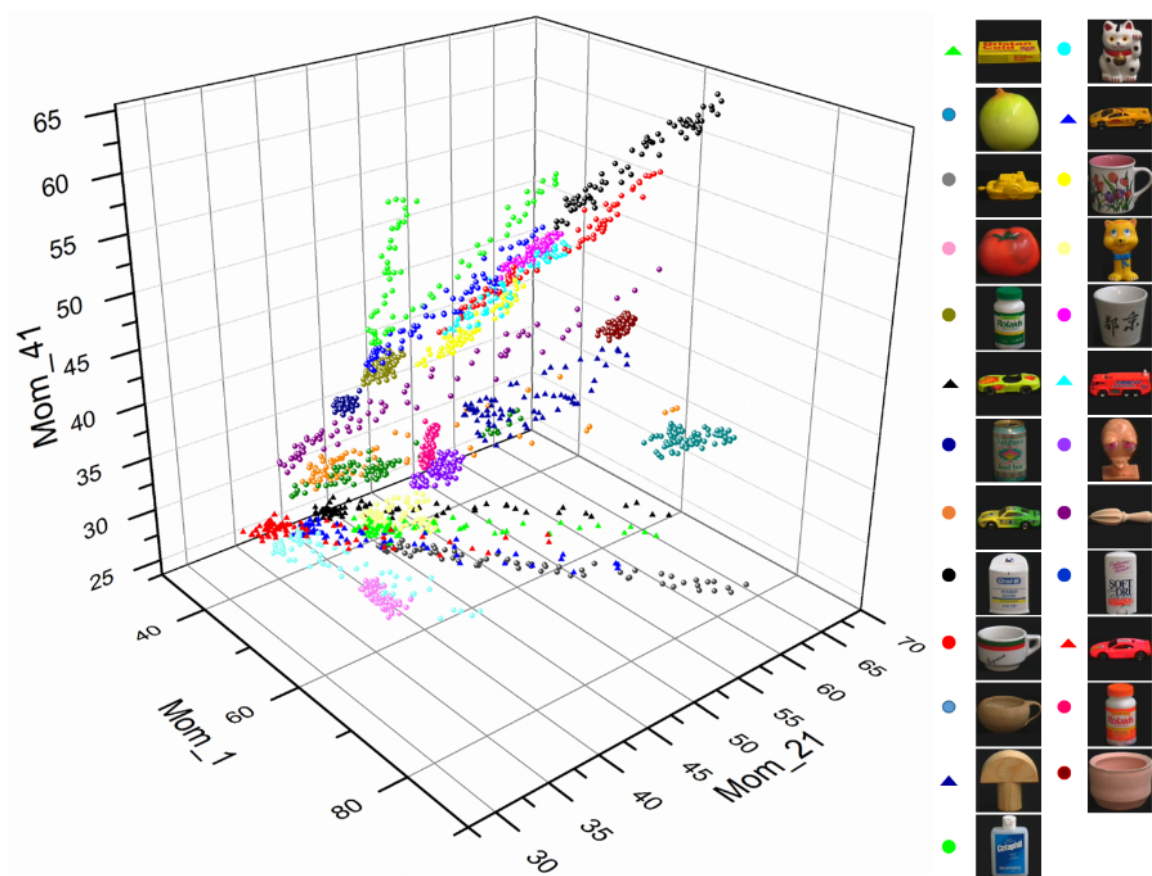


Figura 5.16: Diagrama de dispersión para las primeras 25 clases de la base de datos COIL-100 con ruido Gaussiano con MMZ.

5. Análisis de imágenes en color con momentos radiales

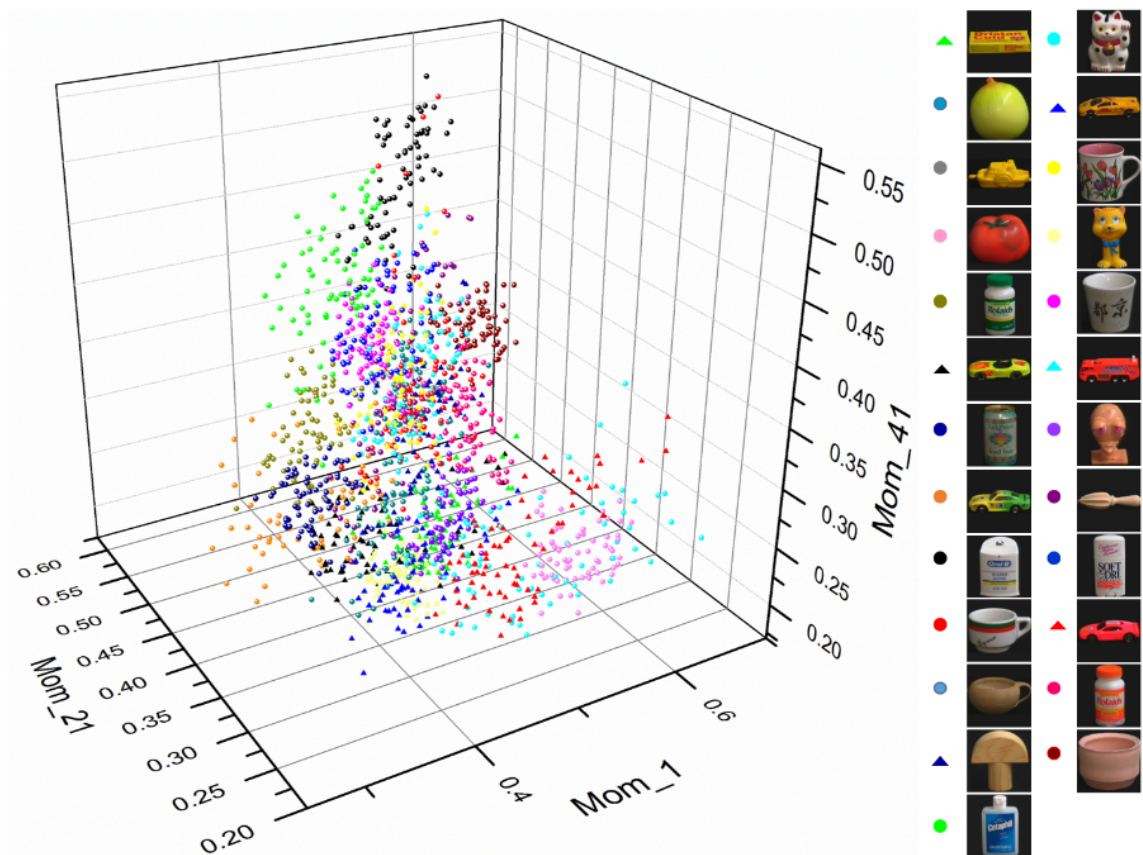


Figura 5.17: Diagrama de dispersión para las primeras 25 clases de la base de datos COIL-100 con ruido Gaussiano con MMT.

5.4. Análisis de resultados

De acuerdo a los experimentos realizados con los momentos Multicanal de Zernike y Tchebichef propuestos se obtuvo un desempeño superior en ambos casos para la reconstrucción de imágenes. Los MMZ propuestos presentan una mayor estabilidad numérica en órdenes superiores y en las pruebas de invariancia geométricas de rotación y escala obtienen un mejor desempeño en comparación con los resultados publicados. La prueba de robustez al ruido muestra el comportamiento esperado; en órdenes menores se obtiene menor detalle de la imagen y por lo tanto el error se acerca a cero, en órdenes superiores se obtiene mayor detalle de la imagen y el error se incrementa al compararse con la imagen original libre de ruido. La gráfica publicada contradice este comportamiento por lo tanto se deduce que ambas imágenes comparadas tienen ruido añadido. Los experimentos de clasificación se llevan a cabo con las medidas de similitud usadas por los autores en los MMZ, en el primer experimento con los órdenes 4-16 tienen un desempeño similar al publicado. En los siguientes experimentos el rendimiento de los MMZ es menor, sin embargo de acuerdo a la clasificación realizada por la aplicación de Matlab Classification Learner usando de igual manera el orden 7 de los MMZ y 60 descriptores se obtiene una exactitud de clasificación del 99.9% en condiciones normales con el clasificador Ensemble/Subspace KNN y este porcentaje de exactitud se mantiene con la base de datos rotada aleatoriamente, con escala $\lambda = 1,5$ e incluso con ruido Gaussiano con $\mu = 0$ y $\sigma^2 = 2,0$. Los diagramas de dispersión se realizaron con los descriptores 1, 21 y 41 lo que equivale al M_{00} de cada canal de color. En cuanto a los MMT presentan de igual manera un desempeño superior a los publicados en la pruebas de invariancia a transformaciones geométricas. En las pruebas de clasificación de acuerdo a las métricas de similitud el rendimiento es menor al publicado por Hosny y Darwish. De igual manera se realizó la prueba de clasificación con la aplicación de Matlab Classification Learner usando 60 descriptores obtenido hasta un 99.3% de exactitud en la clasificación con el modelo Ensemble/Subspace KNN en condiciones normales, rotada aleatoriamente, con escala $\lambda = 1,5$ y del 83,8% con ruido Gaussiano con $\mu = 0$ y $\sigma^2 = 2,0$ de en el mismo modelo de clasificación. En el digrama de dispersión se muestran los descriptores 1, 21 y 41.

Bibliografía

- [1] Hosny, Khalid M., and Mohamed M. Darwish. "New Set of Quaternion Moments for Color Images Representation and Recognition." *Journal of Mathematical Imaging and Vision* 60.5 (2018): 717-736.
- [2] Singh, Chandan, and Jaspreet Singh. "Multi-channel versus quaternion orthogonal rotation invariant moments for color image representation." *Digital Signal Processing* 78 (2018): 376-392
- [3] Flusser, Jan, Barbara Zitova, and Tomas Suk. *Moments and moment invariants in pattern recognition*. John Wiley & Sons, 2009.
- [4] Honarvar, B. Shakibaei, and Raveendran Paramesran. Recursive formula to compute Zernike radial polynomials. *Optics letters* 38.14 (2013): 2487-2489.
- [5] Rivera-López, J. S., and C. J. Camacho-Bello. Color Image Reconstruction by Discrete Orthogonal Moment. *Journal of Data Analysis and Information Processing* 5.04 (2017): 156.
- [6] Camacho-Bello, C., et al. "Alta precisión y cómputo rápido de los momentos Jacobi-Fourier para la descripción de la imagen". *JOSA A* 31.1 (2014): 124-134.

Capítulo 6

Análisis comparativo de los modelos de color y clasificadores

6.1. Introducción

En la actualidad no existe un estudio comparativo del rendimiento de los momentos Multicanal, los momentos Cuaternión y los momentos Multikernel que permita analizar qué opción proporciona una mejor información de la imagen con diferentes clasificadores y modelos de color. Por otro lado los momentos Cuaternión al igual que el Clasificador Multikernel utilizan los momentos Multicanal en su cálculo lo cual representa una carga computacional extra, por lo que en el presente epígrafe se realiza un análisis comparativo entre los momentos Cuaternión, los momentos Multicanal y el Clasificador Multi Kernel para comparar su rendimiento.

6.2. Base de Datos y modelos de clasificación

Para realizar las pruebas se creó una base de datos a partir de vocales en distintos colores y se realizaron 30 imágenes de cada vocal añadiendo ruido gaussiano, como se muestra en la Fig. 6.1. Se tiene por lo tanto una base de datos con un total de 1050 imágenes, con la que se calculan los momentos Cuaternión de Zernike, los momentos Multicanal, los momentos de Zernike del gradiente de la imagen en color e histogramas. La clasificación de la base de datos se hace por medio de la aplicación *Classification Learner* de Matlab versión 9.1 R2016b que permite entrenar modelos de clasificación utilizando el aprendizaje automático supervisado, al proporcionar un conjunto de datos conocidos de entrada y sus respectivas etiquetas. Un modelo de clasificación es una representación abstracta de la relación entre el conjunto de atributos y la etiqueta de clase. Existen diferentes modelos de clasificación en el *Classification Learner*, los

6. Análisis comparativo de los modelos de color y clasificadores

cuales se mencionan en el Apéndice A. Los modelos de color son un sistema para medir los colores que pueden ser percibidos por el ser humano, generalmente poseen 3 o cuatro componentes, cada espacio de color indica la manera en que un color está definido, y su objetivo es facilitar la especificación de los colores de una forma normalizada [1]. Entre los más utilizados se encuentran el RGB, HSI, YCbCR y CMYK. En el Apéndice B se detalla brevemente los Modelos de color utilizados en el análisis de rendimiento de los distintos métodos.



Figura 6.1: Vocales en distintos colores y ejemplos de vocales con ruido añadido.

6.3. Análisis modelo de color RGB

El reconocimiento se realiza con el cálculo de momentos de la base de datos de vocales. Para los momentos Cuaternión de Zernike se calcularon los primeros 90 momentos, como resultado se obtiene una matriz de descriptores de 1050×90 . En el caso de los momentos Multicanal se calculan los primeros 30 momentos Multicanal de Zernike, obteniendo una matriz de descriptores de 1050×90 . Finalmente para el clasificador multikernel, se calculan los primeros 10 momentos Multicanal de Zernike de la base de datos de vocales con ruido, de igual manera se obtienen los primeros 30 momentos de Zernike del gradiente de la imagen en color y finalmente se calcula el histograma de la imagen de la cual se toman 10 bins de cada canal en el modelo de color RGB normalizados en un rango de 0 a 1, de esta manera se obtiene una matriz

6. Análisis comparativo de los modelos de color y clasificadores

de descriptores de 1050×90 , esto es con la finalidad de obtener la misma cantidad de descriptores en cada método. Por medio de la aplicación *ClassificationLearner* de Matlab se entrenan los diferentes modelos de clasificación. La Tabla 6.1 muestra los porcentajes de clasificación y en las Figuras 6.2, 6.3 y 6.4 se observa el diagrama de dispersión de los dos mejores descriptores para los momentos Cuaternión, momentos Multicanal y clasificador Multikernel, respectivamente.

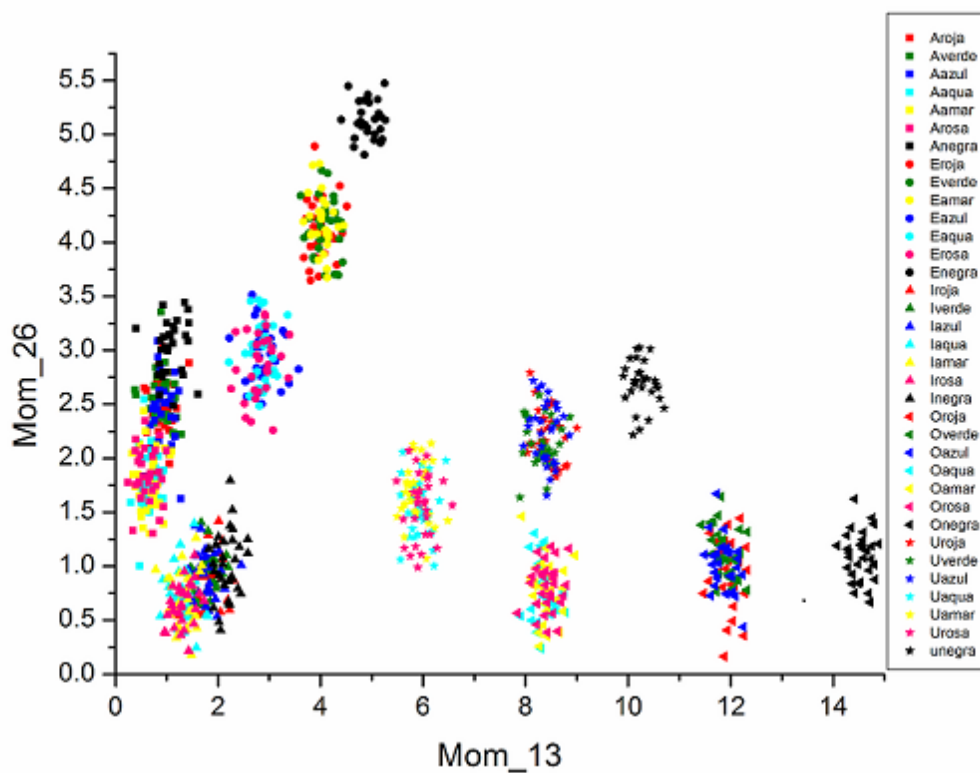


Figura 6.2: Diagrama de dispersión para los momentos Cuaternión de Zernike en el modelo RGB.

6. Análisis comparativo de los modelos de color y clasificadores

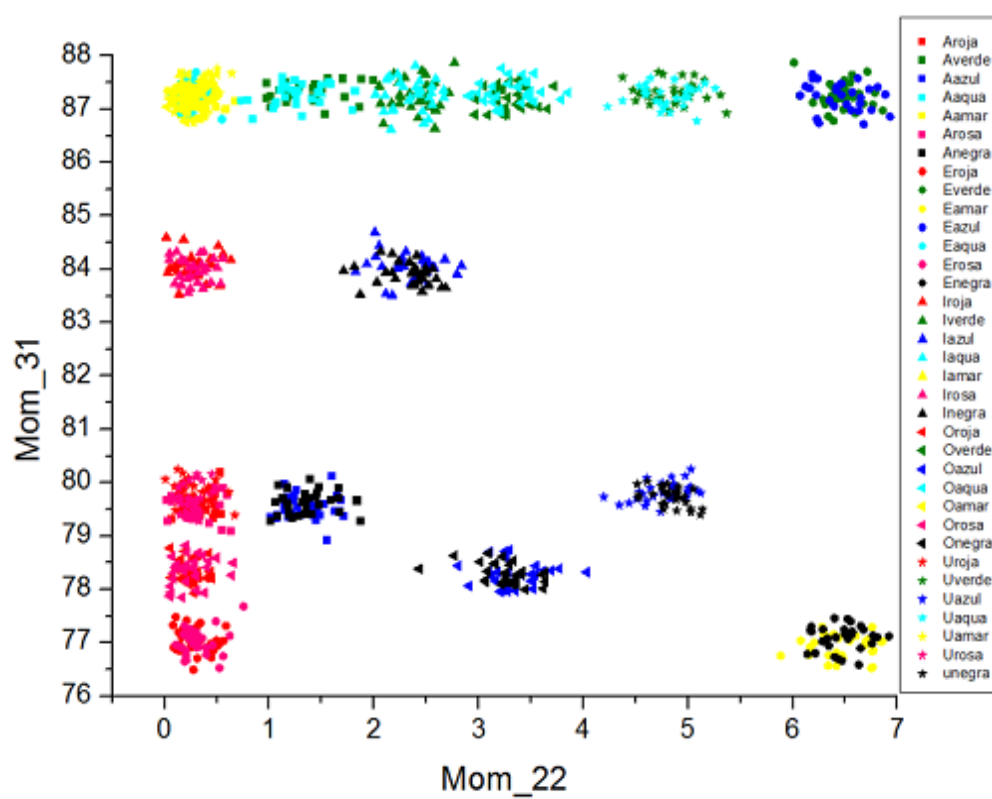


Figura 6.3: Diagrama de dispersión para los momentos Multicanal de Zernike en el modelo RGB.

6. Análisis comparativo de los modelos de color y clasificadores

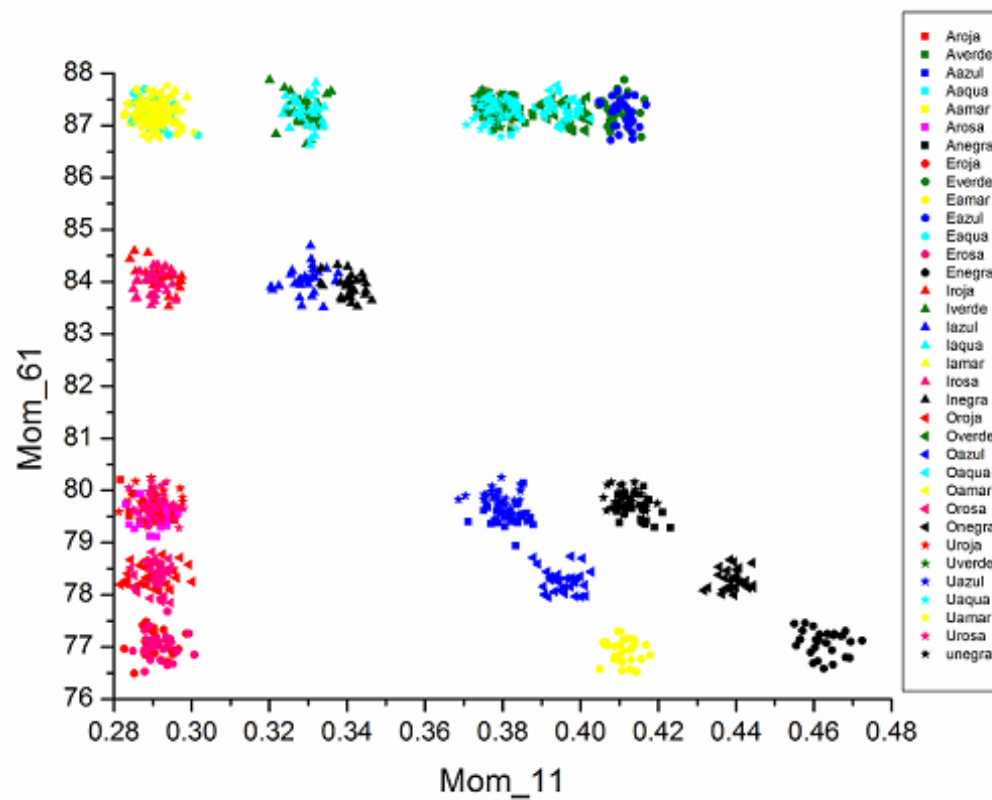


Figura 6.4: Diagrama de dispersión para el clasificador Multikernel de Zernike en el modelo RGB.

6. Análisis comparativo de los modelos de color y clasificadores

Tabla 6.1: Exactitud para los diferentes clasificadores con momentos Cuaternión, Multicanal y Clasificador Multikernel del Modelo de color RGB.

Clasificador	Cuaternión	Multicanal	Multikernel
Complex Tree	40.0 %	99.6 %	99.8 %
Medium Tree	41.6 %	60.0 %	60.0 %
Simple Tree	14.3 %	14.3 %	14.3 %
Linear Discriminant	40.7 %	100.0 %	100.0 %
Quadratic Discriminant	41.8 %	100.0 %	100.0 %
Linear SVM	40.4 %	100.0 %	99.9 %
Quadratic SVM	40.9 %	100.0 %	99.7 %
Cubic SVM	40.2 %	100.0 %	99.4 %
Fine Gaussian SVM	20.2 %	53.0 %	10.4 %
Medium Gaussian SVM	41.2 %	100.0 %	100.0 %
Coarse Gaussian SVM	42.1 %	99.7 %	99.7 %
Fine KNN	44.1 %	89.3 %	89.3 %
Medium KNN	41.8 %	95.4 %	95.4 %
Coarse KNN	40.1 %	89.5 %	92.4 %
Cosine KNN	34.6 %	100.0 %	94.0 %
Cubic KNN	40.5 %	100.0 %	85.6 %
Weighted KNN	42.7 %	100.0 %	95.3 %
Ensemble/Boosted Trees	41.0 %	99.9 %	100.0 %
Ensemble/Bagged Trees	40.8 %	100.0 %	100.0 %
Ensemble/Sub Discriminant	41.0 %	100.0 %	100.0 %
Ensemble/Subspace KNN	41.5 %	100.0 %	100.0 %
Ensemble/RUSBoosted Trees	41.6 %	60.0 %	60.0 %

6. Análisis comparativo de los modelos de color y clasificadores

La matriz de confusión es una herramienta que permite visualizar el desempeño de un algoritmo que emplea aprendizaje supervisado, en éste las filas corresponden a la clase predicha (clase de salida) y las columnas corresponden a la clase verdadera (clase de destino). Las celdas diagonales corresponden a observaciones que están correctamente clasificadas. Las celdas fuera de la diagonal corresponden a observaciones incorrectamente clasificadas. En este caso las matrices de confusión mostradas en las Figuras. 6.5, 6.6 y 6.7 demuestran el desempeño de los momentos Cuaternión, los momentos Multicanal y el Clasificador Multikernel respectivamente.

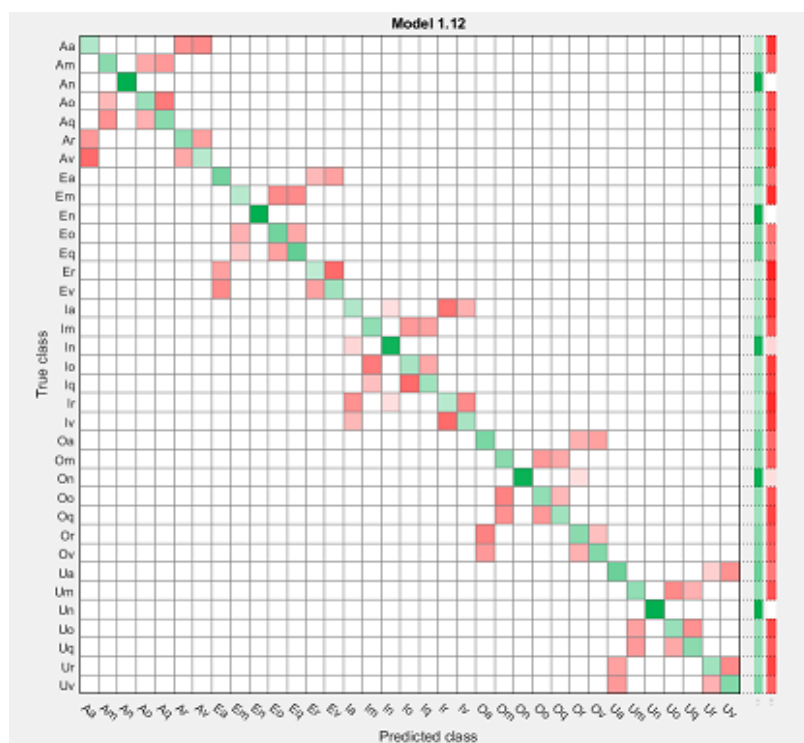


Figura 6.5: Matriz de confusión obtenida a partir de los momentos de Cuaternión de Zernike en el Modelo de color RGB.

6. Análisis comparativo de los modelos de color y clasificadores

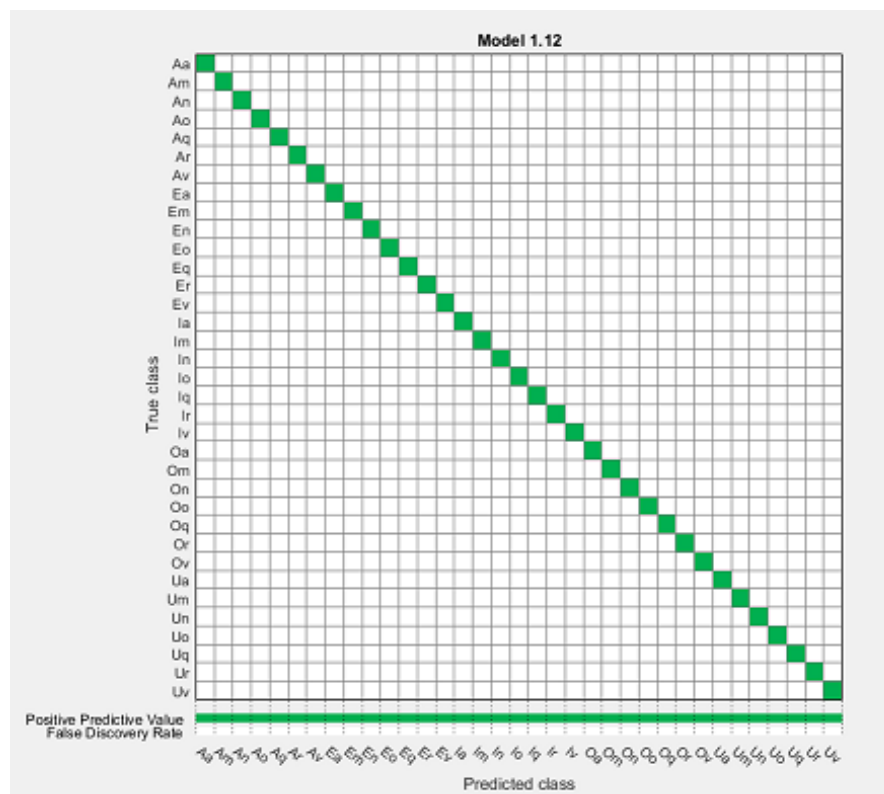


Figura 6.6: Matriz de confusión obtenida a partir de los momentos Multicanal de Zernike del Modelo de color RGB.

6. Análisis comparativo de los modelos de color y clasificadores

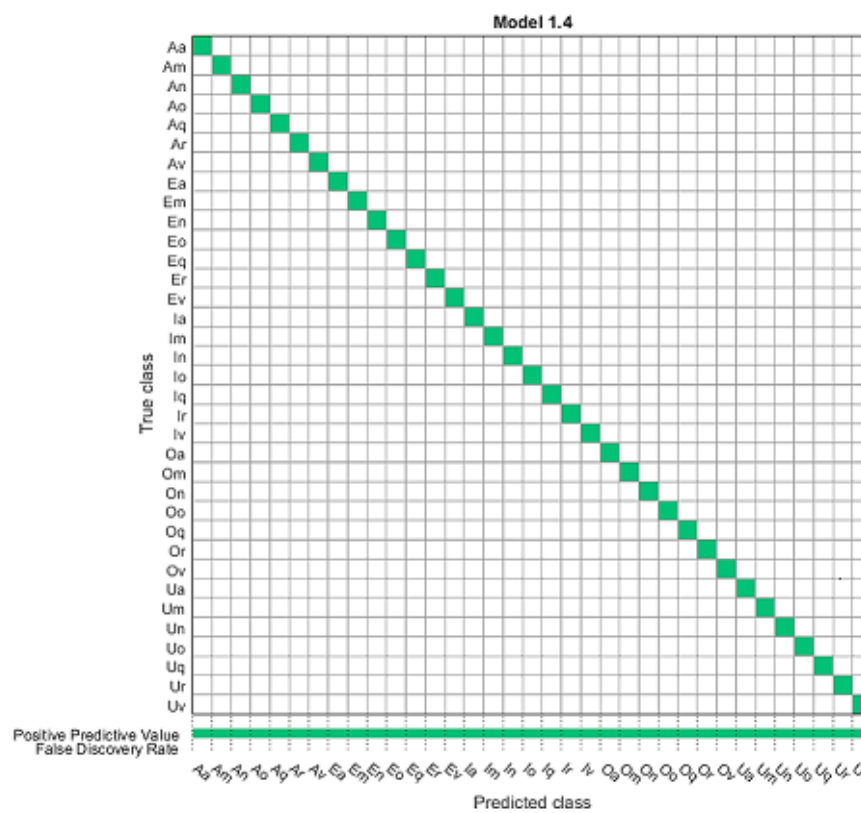


Figura 6.7: Matriz de confusión obtenida a partir del Clasificador Multikernel de Zernike del Modelo de color RGB.

6. Análisis comparativo de los modelos de color y clasificadores

6.4. Análisis modelo de color CMYK

Para analizar el desempeño de los momentos en el modelo de color CMYK, se calculan los primeros 90 momentos Cuaternión de Zernike, y se obtiene una matriz de descriptores de 1050×90 . Para los momentos Multicanal de Zernike se calculan los primeros 22 momentos Multicanal, se obtiene una matriz de descriptores de 1050×88 . Para el clasificador Multikernel se calculan los primeros 8 momentos Multicanal de Zernike para cada canal, con un total de 32 descriptores de forma, se calculan los primeros 30 momentos de Zernike del gradiente de la imagen como descriptores de textura y por último se calcula el histograma de la imagen con 10 bins para cada canal de color, siendo un total de 92 descriptores. La Tabla 6.2 muestra los porcentajes de clasificación y en las Figuras 6.8, 6.9 y 6.10 se observa el diagrama de dispersión de los dos mejores descriptores para los momentos Cuaternión, momentos Multicanal y Clasificador Multikernel respectivamente.

Tabla 6.2: Exactitud para los diferentes clasificadores con momentos Cuaternión de Zernike del Modelo de color CMYK.

Clasificador	Cuaternión	Multicanal	Multikernel
Complex Tree	42.2 %	97.7 %	97.6 %
Medium Tree	38.8 %	59.1 %	59.3 %
Simple Tree	14.3 %	14.3 %	14.3 %
Linear Discriminant	41.4 %	100.0 %	100.0 %
Quadratic Discriminant	41.2 %	100.0 %	100.0 %
Linear SVM	39.8 %	100.0 %	99.6 %
Quadratic SVM	41.0 %	100.0 %	99.3 %
Cubic SVM	40.8 %	100.0 %	98.8 %
Fine Gaussian SVM	14.6 %	14.6 %	16.6 %
Medium Gaussian SVM	40.3 %	100.0 %	100.0 %
Coarse Gaussian SVM	41.0 %	100.0 %	99.1 %
Fine KNN	34.9 %	100.0 %	92.6 %
Medium KNN	36.9 %	100.0 %	97.9 %
Coarse KNN	39.0 %	75.7 %	95.0 %
Cosine KNN	32.8 %	100.0 %	93.7 %
Cubic KNN	32.7 %	100.0 %	89.0 %
Weighted KNN	37.0 %	100.0 %	98.3 %
Ensemble/Boosted Trees	40.7 %	99.7 %	99.7 %
Ensemble/Bagged Trees	43.7 %	100.0 %	100.0 %
Ensemble/Sub Discriminant	41.5 %	100.0 %	100.0 %
Ensemble/Subspace KNN	43.0 %	100.0 %	99.2 %
Ensemble/RUSBoosted Trees	38.8 %	59.1 %	59.3 %

6. Análisis comparativo de los modelos de color y clasificadores

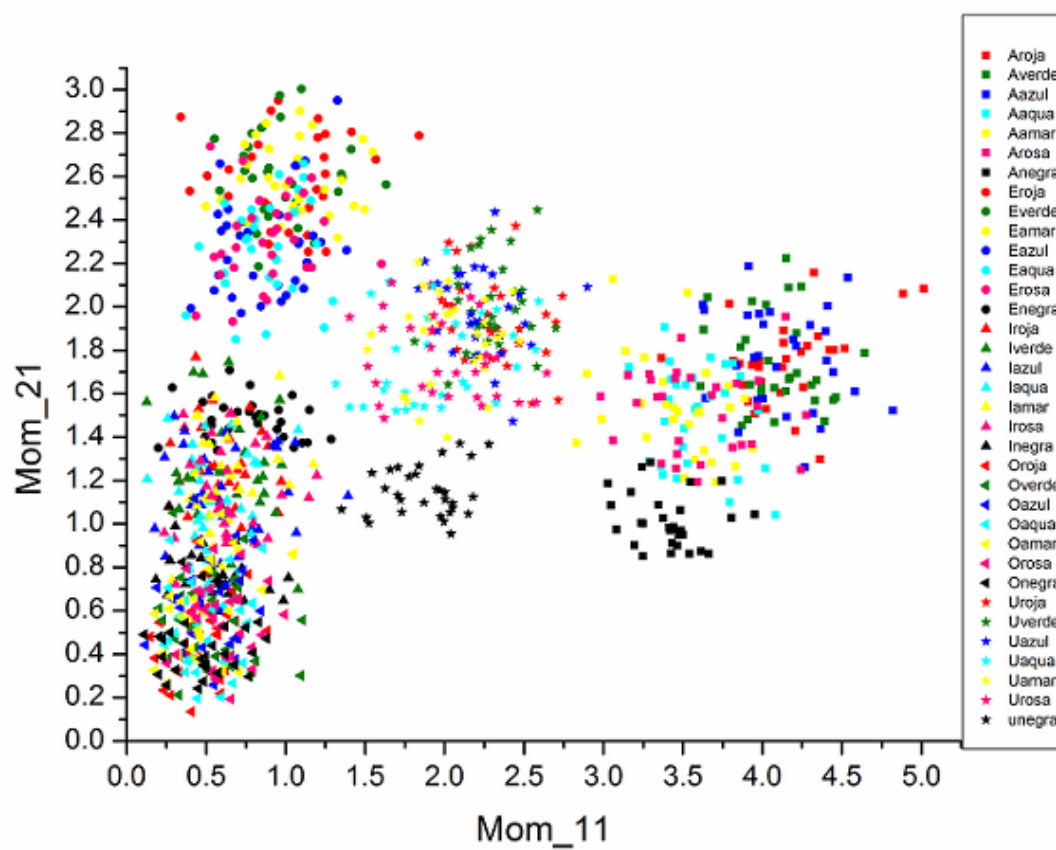


Figura 6.8: Diagrama de dispersión de los momentos Cuaternión de Zernike en el modelo de color CMYK.

6. Análisis comparativo de los modelos de color y clasificadores

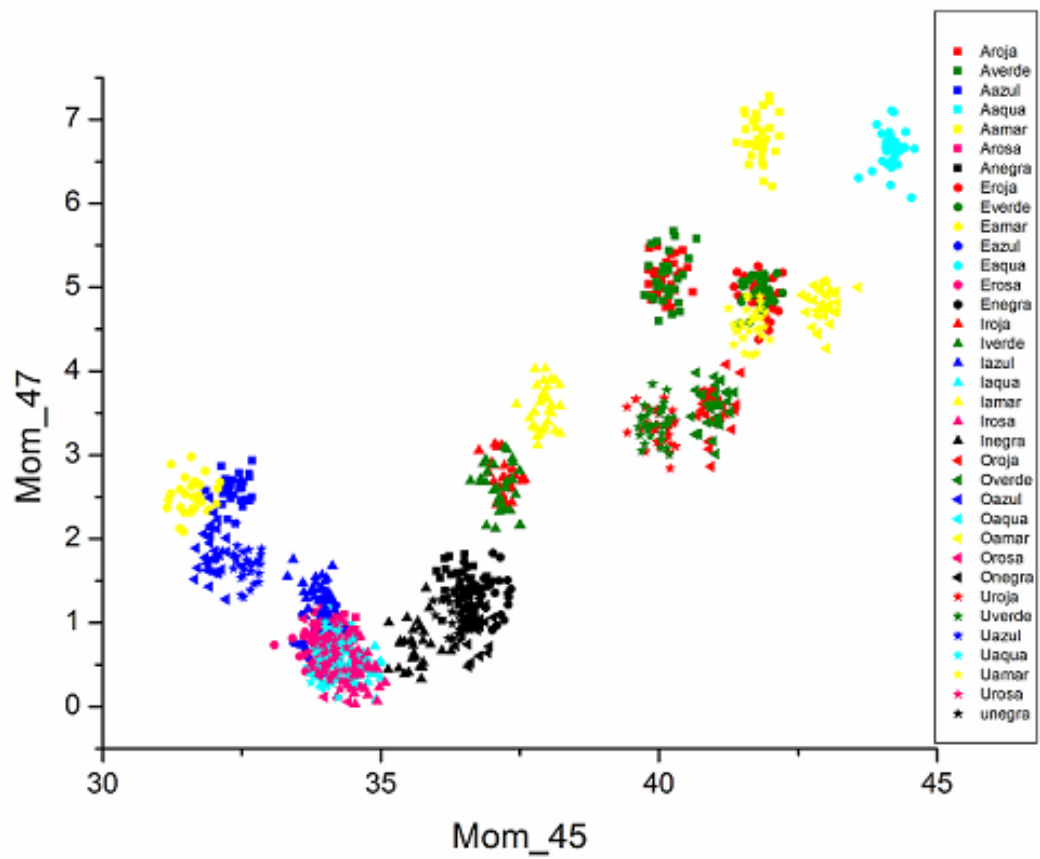


Figura 6.9: Diagrama de dispersión de los Momentos Multicanal de Zernike en el modelo de color CMYK.

6. Análisis comparativo de los modelos de color y clasificadores

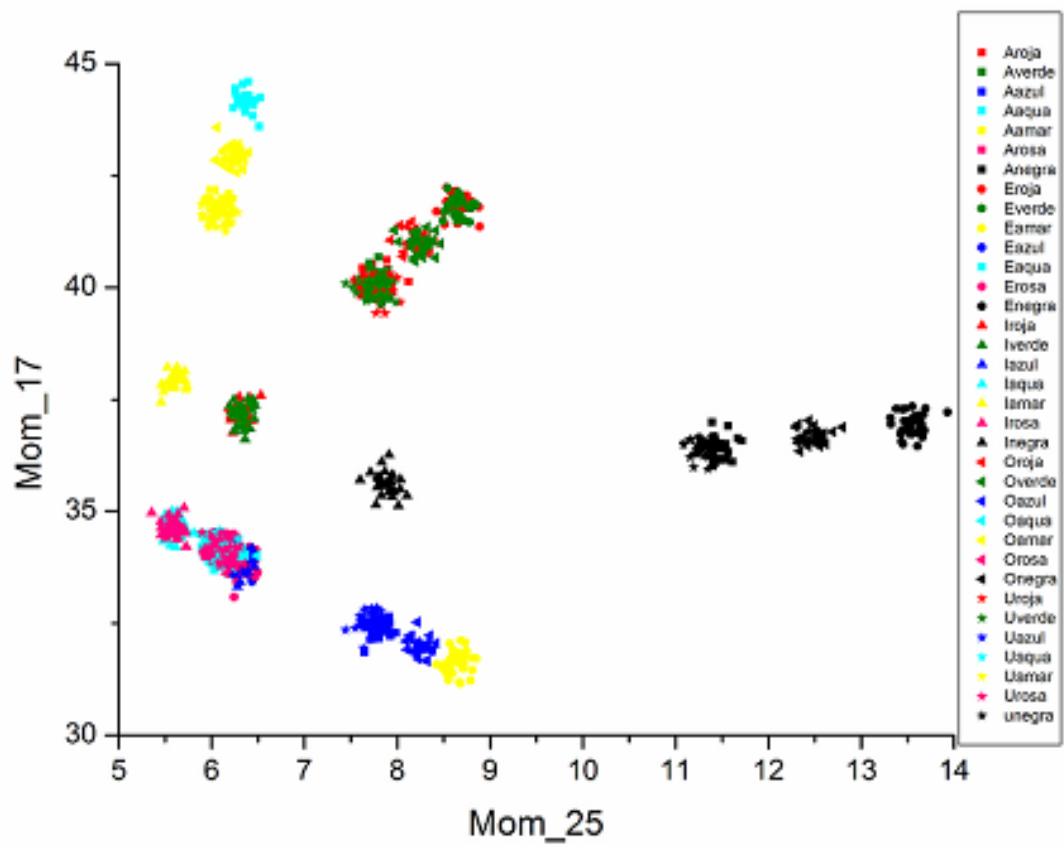


Figura 6.10: Diagrama de dispersión Clasificador Multikernel CMYK.

6.5. Análisis con momentos en el modelo de color HSI

En el Modelo HSI se calcularon nuevamente los primeros 90 momentos Cuaternión de Zernike, para los momentos Multicanal se calculan los primeros 30 momentos, y se obtiene una matriz de descriptores de 1050×90 para ambos casos. En el caso del clasificador Multikernel se calculó los primeros 10 momentos Multicanal de Zernike, los primeros 30 momentos de Zernike del gradiente de la imagen en color y finalmente se calcula el histograma de la imagen de la cual se toman 10 bins de cada canal en el modelo de color HSI normalizados en un rango de 0 a 1. La Tabla 6.3 muestra los porcentajes de clasificación y en las Figuras 6.11, 6.12 y 6.13 se observa el diagrama de dispersión de dos descriptores para los momentos Cuaternión, momentos Multicanal y Clasificador Multikernel respectivamente.

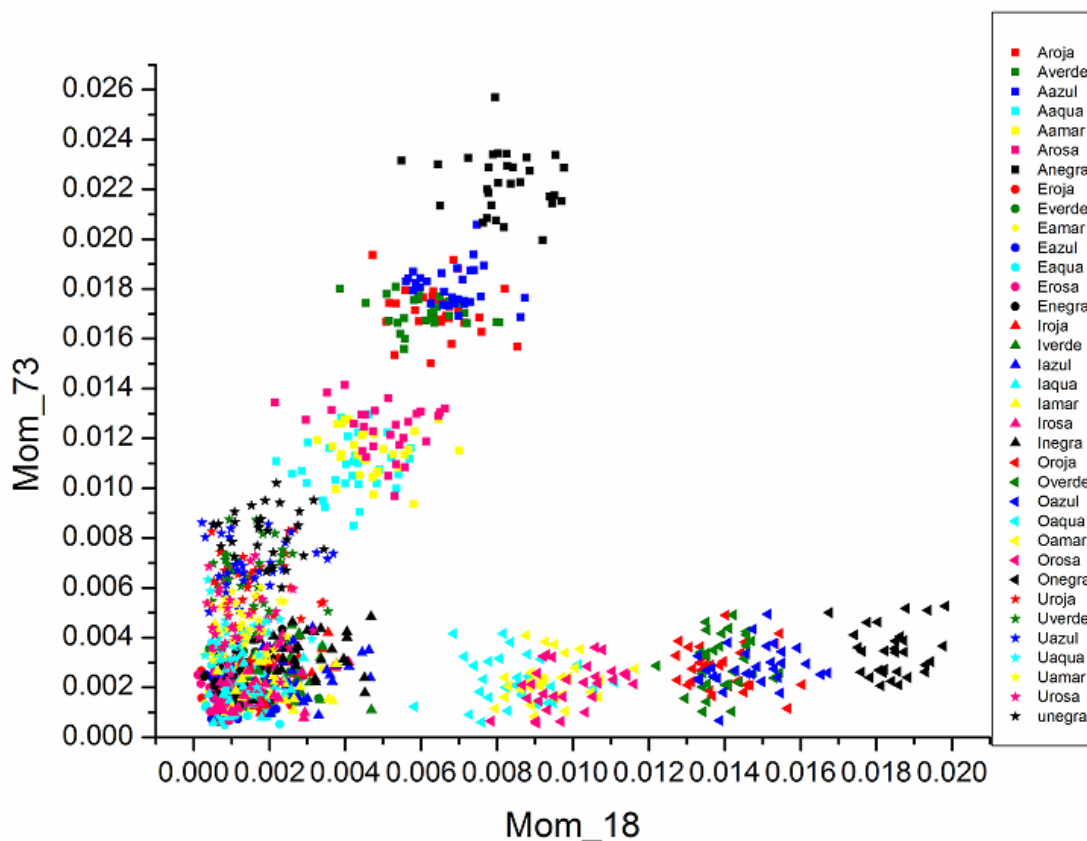


Figura 6.11: Diagrama de dispersión de los momentos Cuaternión en el Modelo de color HSI.

6. Análisis comparativo de los modelos de color y clasificadores

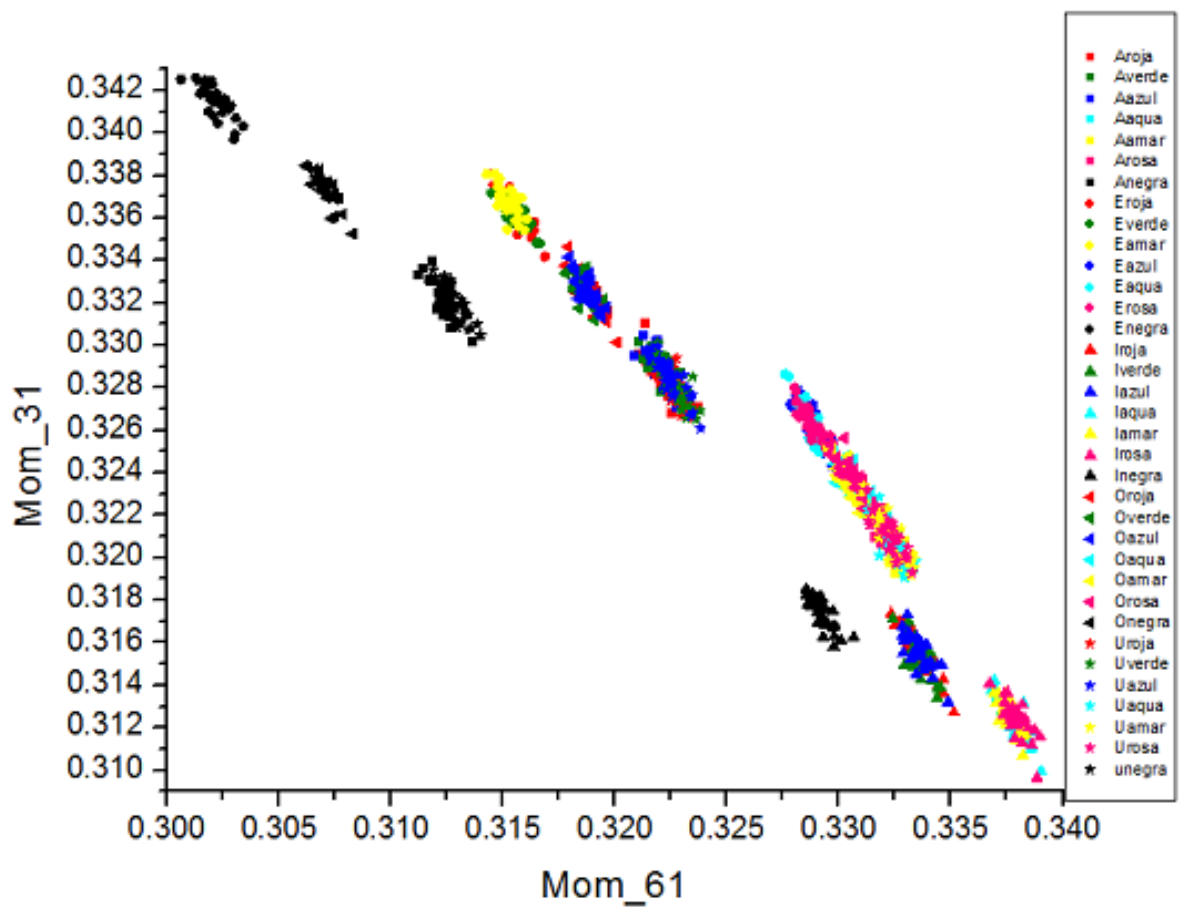


Figura 6.12: Diagrama de dispersión de los momentos Multicanal en el Modelo de color HSI.

6. Análisis comparativo de los modelos de color y clasificadores

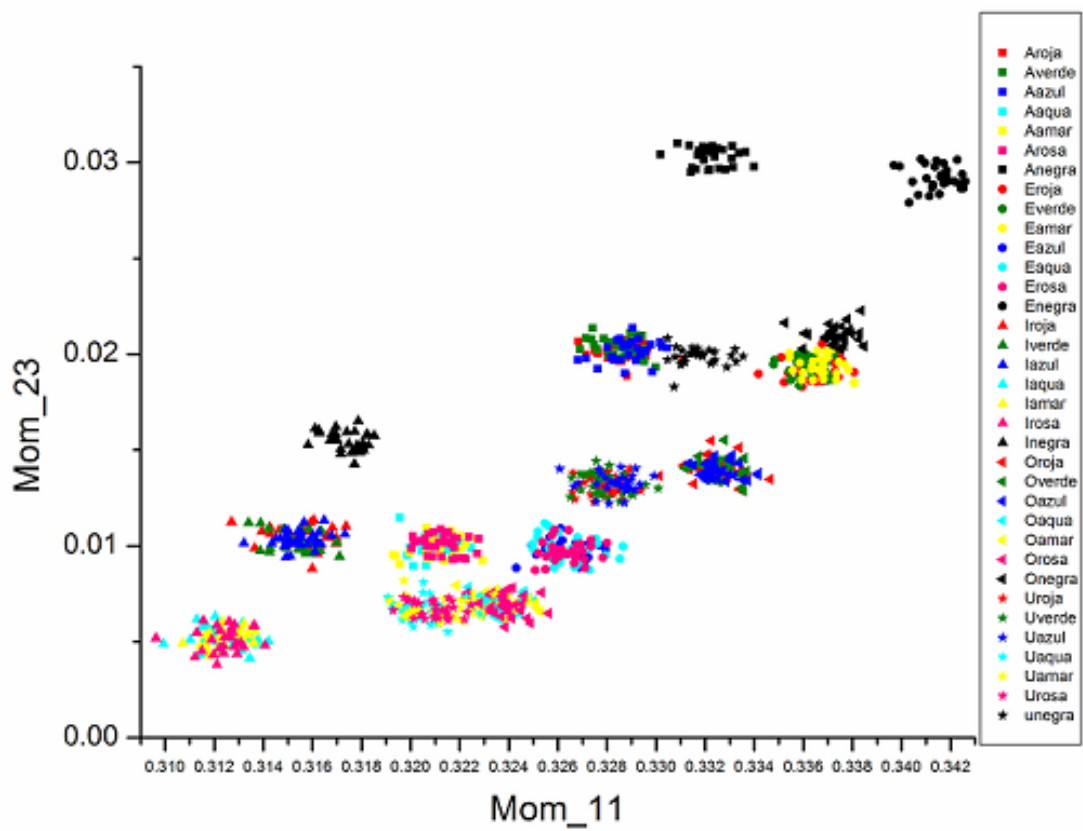


Figura 6.13: Diagrama de dispersión del Clasificador Multikernel en el Modelo de color HSI.

6. Análisis comparativo de los modelos de color y clasificadores

Tabla 6.3: Exactitud para los diferentes clasificadores con momentos Cuaternión, Multicanal y Clasificador Multikernel de Zernike del Modelo de color HSI.

Clasificador	Cuaternión	Multicanal	Multikernel
Complex Tree	86.8 %	90.5 %	98.0 %
Medium Tree	59.4 %	59.0 %	59.1 %
Simple Tree	14.3 %	14.2 %	14.3 %
Linear Discriminant	87.4 %	96.7 %	100.0 %
Quadratic Discriminant	86.5 %	95.9 %	100.0 %
Linear SVM	79.3 %	94.9 %	100.0 %
Quadratic SVM	78.8 %	93.8 %	100.0 %
Cubic SVM	77.0 %	93.3 %	100.0 %
Fine Gaussian SVM	12.7 %	10.8 %	14.5 %
Medium Gaussian SVM	84.7 %	93.7 %	100.0 %
Coarse Gaussian SVM	75.5 %	91.7 %	100.0 %
Fine KNN	58.3 %	85.8 %	99.8 %
Medium KNN	62.2 %	86.7 %	99.7 %
Coarse KNN	54.7 %	73.7 %	83.1 %
Cosine KNN	53.1 %	79.3 %	99.5 %
Cubic KNN	56.5 %	84.1 %	98.5 %
Weighted KNN	63.2 %	87.0 %	99.7 %
Ensemble/Boosted Trees	90.2 %	92.5 %	99.1 %
Ensemble/Bagged Trees	87.4 %	94.9 %	100.0 %
Ensemble/Sub Discriminant	87.4 %	97.3 %	100.0 %
Ensemble/Subspace KNN	85.5 %	94.3 %	99.6 %
Ensemble/RUSBoosted Trees	59.4 %	59.0 %	59.1 %

6.6. Análisis con momentos en el Modelo de color YCbCr

Se procede al cálculo de los primeros 90 momentos Cuaternión de Zernike de la base de datos de vocales con ruido, en el modelo de color YCbCr, para los momentos Multicanal se calculan los primeros 30 momentos de Zernike, para el Clasificador Multikernel, se obtienen los primeros 10 momentos Multicanal de Zernike, los primeros 30 momentos de Zernike del gradiente de la imagen en color y finalmente se calcula el histograma de la imagen de la cual se toman 10 bins de cada canal en el modelo de color YCbCr normalizados en un rango de 0 a 1. En los tres casos se obtiene como resultado se obtiene una matriz de descriptores de 1050×90 . Los resultados obtenidos se muestran en la Tabla 6.4, los diagramas de dispersión de dos descriptores se pueden observar en las Figuras 6.14, 6.15 y 6.16 para los momentos Cuaternión, momentos Multicanal y Clasificador Multikernel respectivamente.

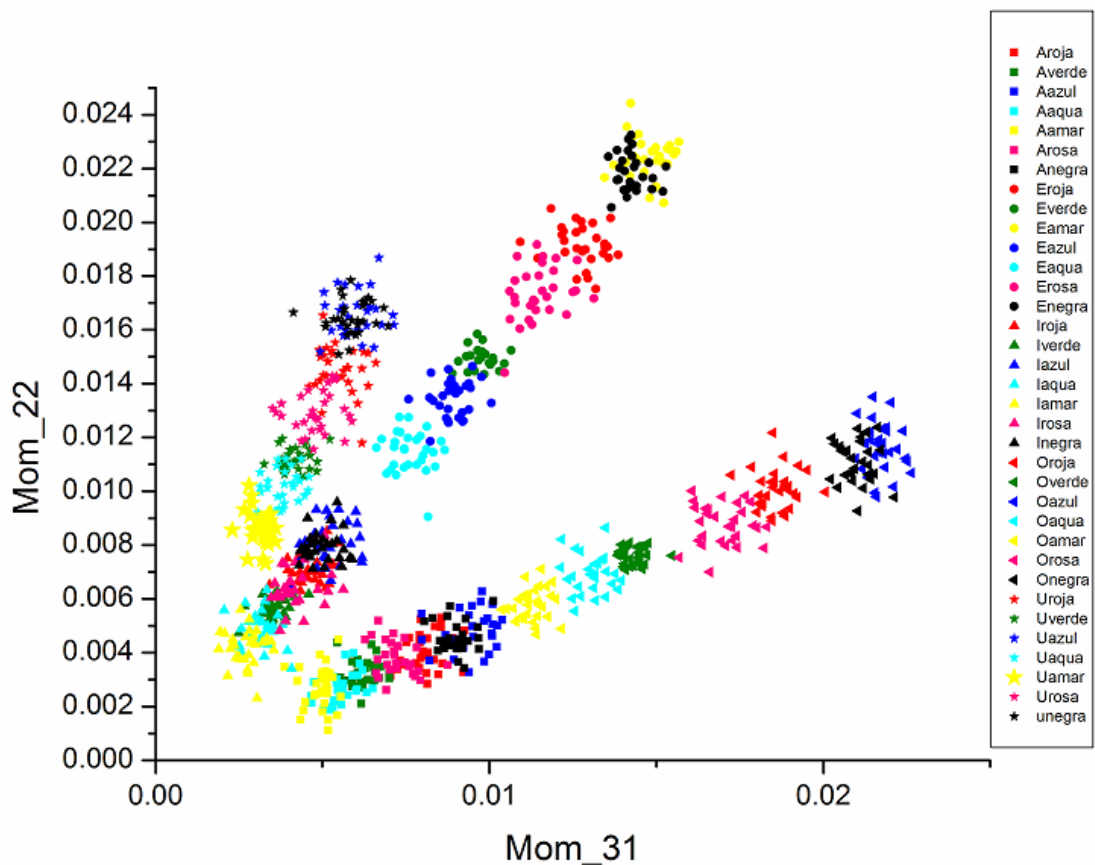


Figura 6.14: Diagrama de dispersión de los momentos Cuaternión en el Modelo de color YCbCr.

6. Análisis comparativo de los modelos de color y clasificadores

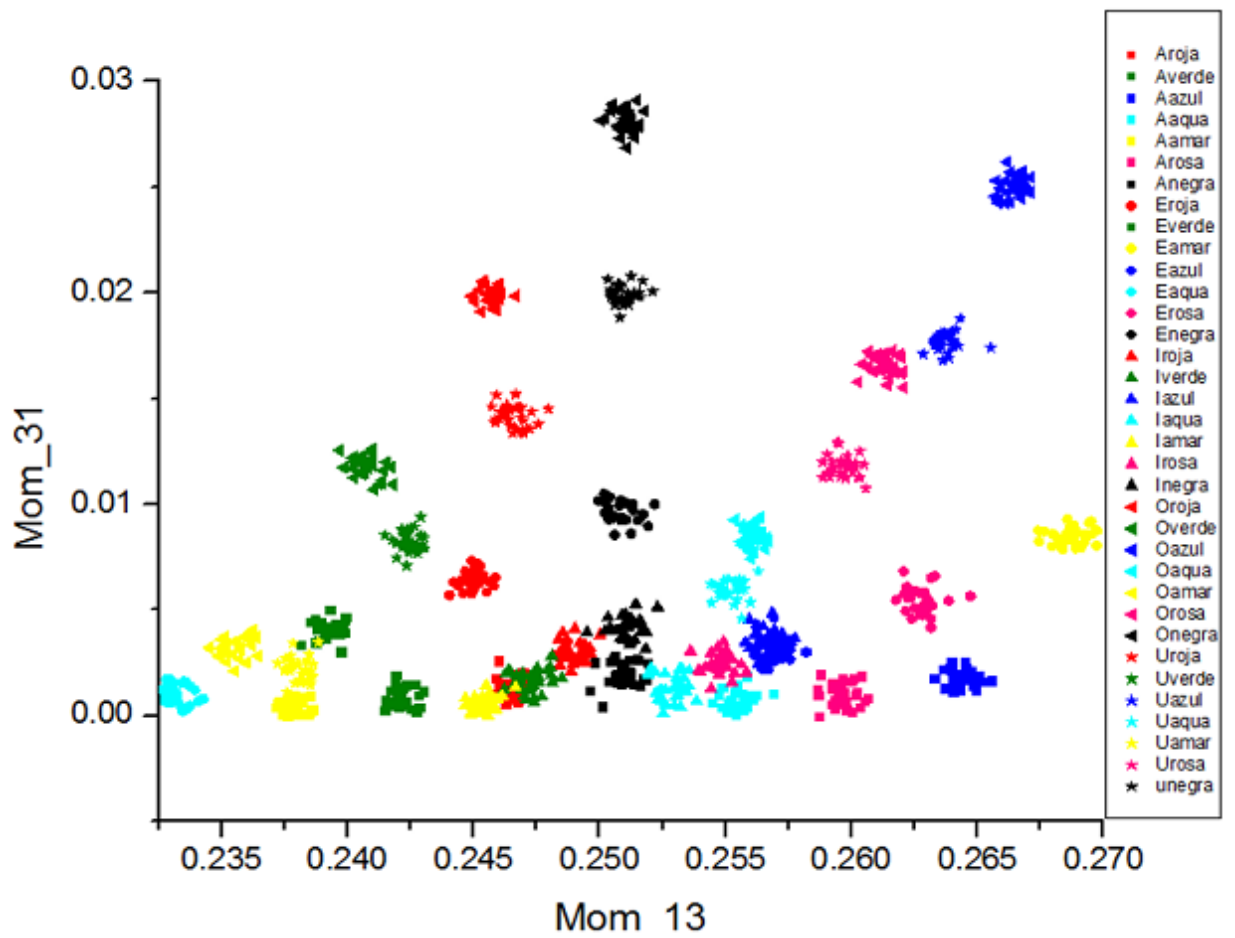
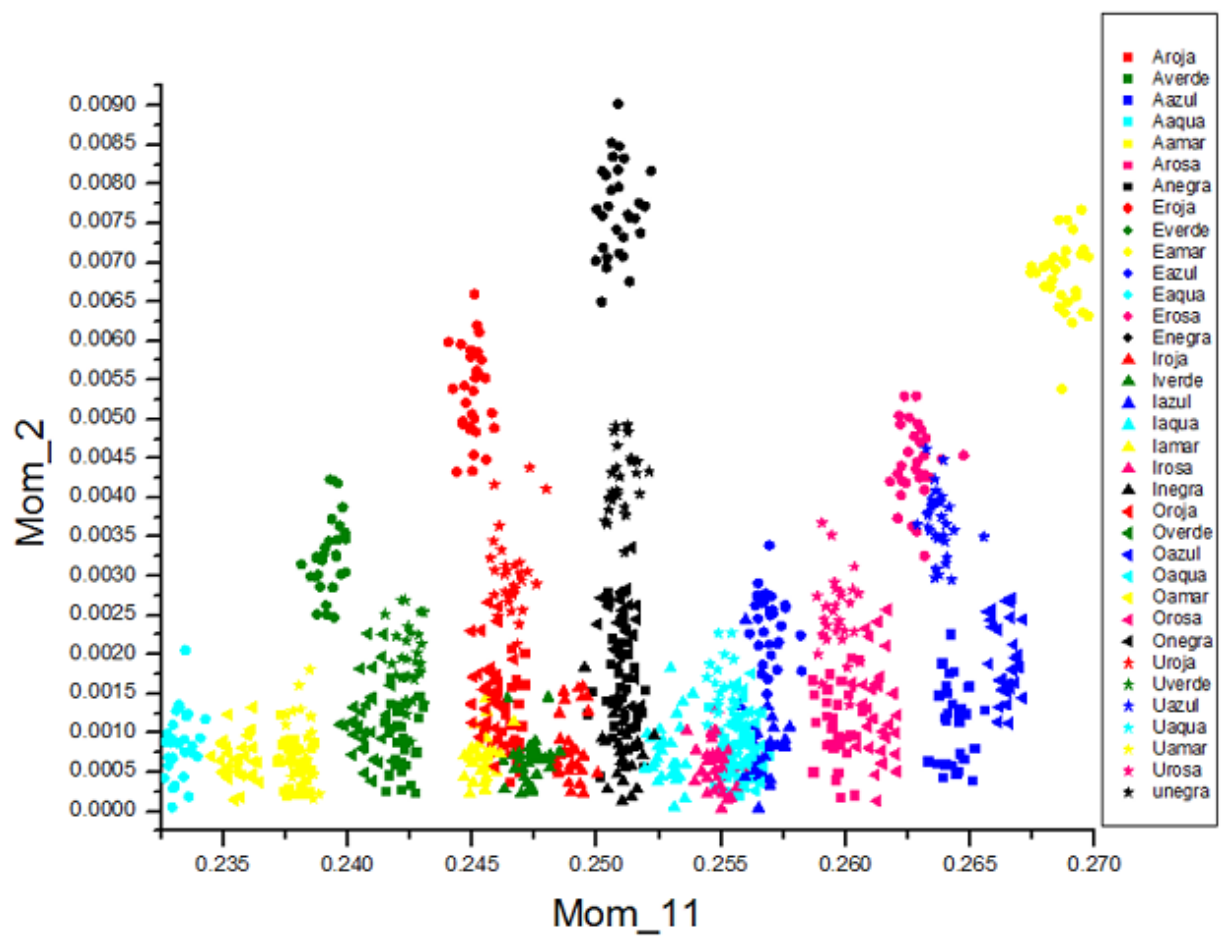


Figura 6.15: Diagrama de dispersión de los momentos Multicanal en el Modelo de color YCbCr.

6. Análisis comparativo de los modelos de color y clasificadores



6. Análisis comparativo de los modelos de color y clasificadores

Tabla 6.4: Exactitud para los diferentes clasificadores con momentos Cuaternión, Multicanal y Clasificador Multikernel de Zernike del Modelo de color YCbCr.

Clasificador	Cuaternión	Multicanal	Multikernel
Complex Tree	98.2 %	98.2 %	98.6 %
Medium Tree	59.0 %	59.6 %	59.4 %
Simple Tree	14.3 %	14.3 %	14.3 %
Linear Discriminant	99.9 %	100.0 %	100.0 %
Quadratic Discriminant	99.9 %	100.0 %	100.0 %
Linear SVM	94.2 %	100.0 %	99.8 %
Quadratic SVM	94.3 %	100.0 %	99.7 %
Cubic SVM	92.6 %	100.0 %	99.2 %
Fine Gaussian SVM	10.2 %	16.8 %	32.6 %
Medium Gaussian SVM	95.8 %	100.0 %	100.0 %
Coarse Gaussian SVM	89.0 %	100.0 %	99.5 %
Fine KNN	56.6 %	97.6 %	94.3 %
Medium KNN	63.3 %	98.9 %	98.0 %
Coarse KNN	89.0 %	75.6 %	92.1 %
Cosine KNN	51.0 %	97.9 %	95.4 %
Cubic KNN	53.0 %	95.9 %	88.3 %
Weighted KNN	63.4 %	99.0 %	98.3 %
Ensemble/Boosted Trees	98.1 %	99.7 %	99.4 %
Ensemble/Bagged Trees	99.0 %	100.0 %	100.0 %
Ensemble/Sub Discriminant	100.0 %	100.0 %	100.0 %
Ensemble/Subspace KNN	99.2 %	100.0 %	95.8 %
Ensemble/RUSBoosted Trees	59.0 %	59.6 %	59.4 %

6. Análisis comparativo de los modelos de color y clasificadores

6.7. Análisis de resultados

La Tabla 6.5 muestra el porcentaje de clasificación más alta obtenidos mediante la aplicación *Classification Learner* con cada uno de los tres métodos utilizados en los distintos modelos de color.

Los momentos Cuaternión se desarrollaron con la visión de que proporcionarían dependencia intercanal para obtener la fusión de forma y color para su mejor desempeño [1]. A pesar de que han demostrado poseer invariancia a la traslación, rotación y escala, tienen una mayor carga computacional porque en su cálculo se utilizan los momentos Multicanal. Su rendimiento es inferior en comparación con los momentos Multicanal y el clasificador multikernel en los modelos de color RGB y CMYK. El clasificador multikernel tiene un buen desempeño, sin embargo los resultados son muy semejantes a los obtenidos con los momentos Multicanal.

Tabla 6.5: Exactitud de los momentos Cuaternión, Multicanal y Clasificador Multikernel en los distintos Modelos de color.

	momentos Cuaternión	momentos Multicanal	Clasificador Multikernel
RGB	44.1 %	100.0 %	100.0 %
CMYK	43.7 %	100.0 %	100.0 %
HSI	90.2 %	97.3 %	100.0 %
YCbCr	100 %	100.0 %	100.0 %

Bibliografía

- [1] Pérez, Marco Antonio Alonso. Espacios de color RGB, HSI y sus Generalizaciones a n-Dimensiones . Diss. Instituto Nacional de Astrofísica, Óptica y Electrónica, 2009.

Capítulo 7

Clasificación de Melanoma

7.1. Introducción

El análisis del capítulo anterior demuestra la capacidad descriptiva de los momentos con diferentes modelos de color. Por tal motivo, nos dimos a la tarea de buscar una aplicación que requiera un análisis descriptivo tanto de forma como de color. Entre las aplicaciones que requiere este tipo de estudio se encuentra la detección temprana del melanoma cutáneo.

El cáncer de piel es la neoplasia maligna más común entre los distintos tipos de cáncer. La piel es el órgano más grande del cuerpo, que representa el 16% de la masa corporal, cumple con distintas funciones como: recubrir los órganos internos y protegerlos de lesiones, como barrera contra patógenos, evita la pérdida excesiva de agua, controla la temperatura corporal y protege contra los rayos ultravioleta. La piel consta de tres capas, la epidermis, la dermis y la hipodermis. La epidermis es la capa superior de la piel con un espesor promedio de 1/100 pulgadas, los principales tipos de células presentes en esta capa son las células escamosas, las células basales y los melanocitos. Los melanocitos son las células que pueden convertirse en melanoma. El cáncer de piel es nombrado de acuerdo con la célula de la que surgen y a su comportamiento clínico. Los tres tipos más comunes son [1]:

- Carcinomas de células basales
- Carcinomas de células escamosas
- Melanomas malignos cutáneos

Los carcinomas de células basales y escamosas superan en incidencia al melanoma, no obstante son más fáciles de tratar y tienen un mejor pronóstico a largo plazo.

7. Clasificación de Melanoma

Por otro lado, el melanoma es mucho menos común pero es la forma más mortal de cáncer de piel por su gran capacidad de producir metástasis y su alta quimiorresistencia. Puede presentarse en cualquier parte de la piel, usualmente posee un color café o negro por la producción de melanina y según estimaciones presenta una incidencia creciente, por lo que es objeto de numerosas investigaciones. Debido a que el melanoma es curable en sus etapas iniciales se ha hecho hincapié en la prevención y diagnóstico oportuno, ya que de no detectarse oportunamente este tipo de cáncer tiene una mayor probabilidad de propagarse a otras partes del cuerpo.

Los factores de riesgo para el melanoma incluyen la exposición a los rayos ultravioleta, nevos múltiples, nevos clínicamente atípicos o displásicos, pecas y antecedentes familiares. Para evaluar una lesión pigmentada se consideran cinco características: asimetría, borde irregular, coloración, diámetro (mayor de 6 mm) y evolución, es decir si el tamaño o la forma del lunar está cambiando. Por lo general, la ulceración, el sangrado y la supuración son signos de melanoma tardío [2]. La recomendación o el estándar de oro para el tratamiento es la biopsia por escisión de cualquier lesión pigmentada que se sospeche sea un melanoma.

Debido a que la detección temprana de lesiones aumenta la tasa de supervivencia en un 90 % se han propuesto distintos sistemas de clasificación para lesiones de la piel.

Codella et al presentan un enfoque que combina el aprendizaje profundo, codificación dispersa y algoritmos de aprendizaje de máquina de vectores de soporte SVM, evalúan su método con la base de datos obtenida de International Skin Imaging Collaboration obteniendo una precisión del 93.1 % al discriminar entre melanoma y lesiones sin melanoma y 73.9 % discriminando entre melanoma y lesiones atípicas [3].

Hosny et. al proponen un método de clasificación utilizando una red neuronal convolucional profunda para clasificar las imágenes en color del cáncer de piel en 3 tipos: melanoma, nevus atípico y nevo común. Usan transferencia de aprendizaje y aumento de imagen aplicada a una AlexNet pre-entrenada. Obteniendo hasta un 98.93 % de precisión en la clasificación [4].

Barros y Correia presentan un modelo capaz de clasificar 12 lesiones cutáneas con una arquitectura Res Net-152, utilizan bases de datos proporcionadas por el Departamento de Dermatología del Centro Médico Universitario MED-NODE y Edinburgh Dermofit Image Library logrando una precisión hasta del 80 % [5].

En este trabajo de tesis los momentos Multicanal de Zernike han demostrado en experimentos anteriores una excelente capacidad de describir tanto la forma como el color del objeto, razón por lo cual en el presente epigrafe se lleva a cabo la clasificación de melanoma mediante momentos Multicanal de Zernike.

7.2. Base de datos

Se utiliza la base de datos HAM10000 (“Human Against Machine”) [6] con 10015 imágenes dermatoscópicas de diferentes poblaciones que está disponible públicamente a partir de febrero del 2018 e incluye metadatos. Las imágenes provienen del Departamento de Dermatología de la Universidad Médica de Viena, Austria, y la práctica de Cliff Rosendahl en Queensland, Australia. En esta base de datos las imágenes han sido validadas mediante diagnósticos histopatológicos realizados por dermatólogos especializados, excluyendo de la base los casos ambiguos. Algunas queratosis benignas faciales se corroboraron mediante un estudio con microscopía confocal realizado en el Departamento de Dermatología de la Universidad Médica de Viena. También se hizo seguimiento de nevos monitoreados evaluados también por Harald Kittler quien cuenta con más de 20 años de experiencia en el seguimiento de dermatoscopia digital y por último por consenso de los autores para casos típicos benignos sin histopatología o seguimiento. El conjunto de entrenamiento HAM10000 incluye lesiones pigmentadas de diferentes poblaciones, presenta categorías diagnósticas importantes en el ámbito de lesiones de piel.

- **nv** Nevos melanocíticos que son neoplasias benignas de melanocitos y aparecen en una gran variedad de variantes.
- **mel** Neoplasia maligna derivada de melanocitos que puede aparecer en diferentes variantes. Se incluyeron todas las variantes de melanoma, incluido el melanoma in situ, pero se excluyeron el melanoma no pigmentado, subungueal, ocular o mucoso.
- **bkl** Queratosis benigna.
- **df** El dermatofibroma es una lesión benigna de la piel considerada como una proliferación benigna o una inflamación.
- **akiec** Son queratosis actínicas (queratosis solares) y el carcinoma intraepitelial (enfermedad de Bowen) son comunes no invasivos, variantes del carcinoma de células escamosas que pueden tratarse localmente sin cirugía.
- **bcc** Carcinoma de células basales variante común del cáncer de piel epitelial que rara vez hace metástasis.
- **vasc** Son lesiones vasculares de la piel, las hemorragias y angiomas están incluidas en esta categoría.

7. Clasificación de Melanoma

La base de datos consta de un total de 10015 imágenes de las cuales 6705 pertenecen a la clase de nevos, 1113 a melanoma, 1099 a queratosis benigna, 115 a dermatofibroma, 327 a queratosis actínica, 514 a carcinoma de células basales y por último 142 pertenecen a la clase de lesiones vasculares. Las imágenes originalmente poseen un tamaño de 600×450 píxeles, dado que el cálculo de los momentos requiere una imagen de tamaño $N \times N$ el primer paso para el análisis de la base de datos fue realizar una normalización de la misma. Mediante la Ec. 4.15 se obtuvo el centroide de cada imagen y se procedió a recortar las imágenes en el programa MatLab versión 9.1 R2016b mediante la función *imcrop* a un tamaño de 350×350 píxeles, obteniendo de esta manera la mayor información sin alterar la forma de las imágenes dermatoscópicas como puede observarse en la Fig 7.1. Debido a que en la base de datos la clase Nevos representa más del 50% de los datos se optó por elegir aleatoriamente 1113 imágenes para evitar el sobreajuste en los algoritmos de clasificación. La clasificación de la base de datos se hizo nuevamente por medio de la aplicación del programa MatLab *ClassificationLearner* en los 4 diferentes Modelos de color.

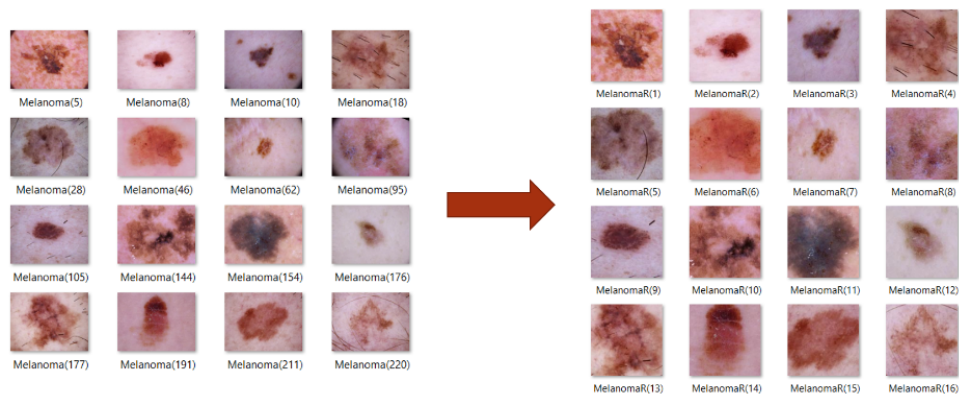


Figura 7.1: Normalización de la base de datos HAM 10000.

7.3. Clasificación con el modelo de color RGB

La clasificación se realizó de dos diferentes formas, en la primera se tomó en cuenta las 7 distintas clases presentes en la base de datos. Se procedió al cálculo de los primeros 90 momentos Cuaternión, en el caso de los momentos Multicanal se calculó los primeros 30 momentos Multicanal de Zernike, para los Momento Multikernel se calcula los primeros 10 momentos Multicanal de Zernike, así como los primeros 30 momentos de Zernike del gradiente de la imagen en color y finalmente se calcula el histograma de la imagen de la cual se toman 10 bins de cada canal normalizados en un rango de 0

7. Clasificación de Melanoma

a 1. Se obtiene una matriz de descriptores de 4423×90 para cada uno de los tres métodos. Los resultados obtenidos se muestran en la Tablas 7.1, para los momentos Cuaternión, momentos Multicanal y Clasificador Multikernel respectivamente.

Tabla 7.1: Exactitud para los diferentes clasificadores con momentos Cuaternión, Multicanal y Clasificador Multikernel de Zernike del Modelo de color RGB usando 7 etiquetas.

Clasificador	Cuaternión	Multicanal	Multikernel
Complex Tree	41.8 %	48.3 %	51.6 %
Medium Tree	40.2 %	47.6 %	50.2 %
Simple Tree	37.1 %	42.1 %	45.4 %
Linear Discriminant	36.8 %	35.4 %	43.3 %
Quadratic Discriminant	30.9 %	29.7 %	30.3 %
Linear SVM	48.6 %	57.1 %	58.5 %
Quadratic SVM	49.0 %	59.4 %	61.2 %
Cubic SVM	48.3 %	57.3 %	60.1 %
Fine Gaussian SVM	38.3 %	40.3 %	37.9 %
Medium Gaussian SVM	50.0 %	57.8 %	59.8 %
Coarse Gaussian SVM	46.1 %	50.6 %	55.1 %
Fine KNN	38.1 %	43.9 %	50.1 %
Medium KNN	44.4 %	49.7 %	53.9 %
Coarse KNN	44.7 %	49.1 %	54.3 %
Cosine KNN	44.7 %	49.9 %	53.9 %
Cubic KNN	43.7 %	50.0 %	53.6 %
Weighted KNN	45.0 %	50.7 %	54.5 %
Ensemble/Boosted Trees	44.3 %	50.2 %	53.2 %
Ensemble/Bagged Trees	46.2 %	53.4 %	58.7 %
Ensemble/Sub Discriminant	46.9 %	56.3 %	57.3 %
Ensemble/Subspace KNN	45.4 %	53.8 %	52.7 %
Ensemble/RUSBoosted Trees	37.3 %	45.2 %	50.6 %

Para todas las pruebas en los distintos Modelos de color se realizó una segunda clasificación tomando en cuenta los datos pertenecientes sólo a 2 clases, Nevo con 1113 elementos y Melanoma con el mismo número de elementos. Los resultados obtenidos en el Modelo de color RGB se pueden observar en las Tablas 7.2, para los momentos Cuaternión, momentos Multicanal y Clasificador Multikernel respectivamente.

7. Clasificación de Melanoma

Tabla 7.2: Exactitud para los diferentes clasificadores con momentos Cuaternión, Multicanal y Clasificador Multikernel de Zernike de color RGB usando 2 etiquetas.

Clasificador	Cuaternión	Multicanal	Multikernel
Complex Tree	69.0 %	75.4 %	78.8 %
Medium Tree	70.1 %	76.9 %	78.8 %
Simple Tree	67.7 %	73.5 %	75.2 %
Linear Discriminant	76.2 %	72.2 %	72.2 %
Quadratic Discriminant	74.9 %	71.8 %	64.2 %
Linear SVM	77.5 %	83.8 %	81.9 %
Quadratic SVM	77.5 %	82.5 %	82.0 %
Cubic SVM	75.9 %	80.5 %	82.2 %
Fine Gaussian SVM	69.7 %	69.6 %	67.9 %
Medium Gaussian SVM	79.1 %	83.9 %	83.5 %
Coarse Gaussian SVM	78.5 %	81.7 %	81.3 %
Fine KNN	72.3 %	73.8 %	77.6 %
Medium KNN	75.2 %	78.1 %	82.3 %
Coarse KNN	73.9 %	74.8 %	80.1 %
Cosine KNN	77.3 %	78.2 %	79.6 %
Cubic KNN	76.1 %	78.6 %	81.3 %
Weighted KNN	74.0 %	76.6 %	81.9 %
Ensemble/Boosted Trees	76.9 %	82.9 %	84.2 %
Ensemble/Bagged Trees	77.4 %	80.3 %	84.3 %
Ensemble/Sub Discriminant	77.5 %	82.5 %	81.1 %
Ensemble/Subspace KNN	74.5 %	80.0 %	78.9 %
Ensemble/RUSBoosted Trees	70.0 %	77.0 %	79.2 %

7. Clasificación de Melanoma

7.4. Clasificación con el modelo de color HSI

En el modelo de color HSI, se calculan los primeros 90 momentos Cuaternión. De igual manera se calculó los primeros 30 momentos Multicanal de Zernike y en el caso del Clasificador Multikernel se cálculo los primeros 10 momentos Multicanal de Zernike de la base de datos HAM10000 en esta ocasión el modelo de color HSI, así como los primeros 30 momentos de Zernike del gradiente de la imagen en color y finalmente se calcula el histograma de la imagen de la cual se toman 10 bins de cada canal normalizados en un rango de 0 a 1. Para todos los métodos se obtiene una matriz de descriptores de 4423×90 . Los resultados obtenidos en la primera clasificación realizada con las 7 etiquetas se muestran en las Tablas 7.3, para los momentos Cuaternión, los momentos Multicanal y el Clasificador Multikernel respectivamente.

Tabla 7.3: Exactitud para los diferentes clasificadores con momentos Cuaternión, Multicanal y Clasificador Multikernel de Zernike del Modelo de color HSI usando 7 etiquetas.

Clasificador	Cuaternión	Multicanal	Multikernel
Complex Tree	34.5 %	67.7 %	52.7 %
Medium Tree	34.8 %	68.5 %	52.2 %
Simple Tree	33.7 %	66.7 %	50.5 %
Linear Discriminant	24.2 %	41.2 %	45.4 %
Quadratic Discriminant	23.5 %	39.8 %	25.5 %
Linear SVM	36.4 %	68.6 %	55.9 %
Quadratic SVM	36.3 %	72.1 %	59.4 %
Cubic SVM	36.4 %	71.3 %	58.7 %
Fine Gaussian SVM	31.0 %	67.3 %	32.1 %
Medium Gaussian SVM	37.6 %	72.0 %	59.5 %
Coarse Gaussian SVM	37.1 %	67.0 %	54.0 %
Fine KNN	29.4 %	65.1 %	48.3 %
Medium KNN	35.4 %	69.6 %	53.9 %
Coarse KNN	37.5 %	68.6 %	52.6 %
Cosine KNN	34.7 %	69.1 %	54.9 %
Cubic KNN	35.0 %	69.2 %	51.6 %
Weighted KNN	36.6 %	69.9 %	54.8 %
Ensemble/Boosted Trees	36.6 %	69.7 %	53.3 %
Ensemble/Bagged Trees	36.5 %	70.6 %	59.2 %
Ensemble/Sub Discriminant	37.3 %	69.3 %	54.4 %
Ensemble/Subspace KNN	38.5 %	70.5 %	53.1 %
Ensemble/RUSBoosted Trees	30.4 %	46.4 %	49.8 %

Los resultados obtenidos en la segunda clasificación con 2 etiquetas Nevo y Melanoma en el Modelo de color HSI se pueden observar en las Tablas 7.4, para los momentos Cuaternión, momentos Multicanal y Clasificador Multikernel

7. Clasificación de Melanoma

respectivamente.

Tabla 7.4: Exactitud para los diferentes clasificadores con momentos Cuaternión, Multicanal y Clasificador Multikernel de Zernike del Modelo de color HSI usando 2 etiquetas.

Clasificador	Cuaternión	Multicanal	Multikernel
Complex Tree	62.0 %	74.2 %	78.2 %
Medium Tree	61.8 %	76.0 %	81.5 %
Simple Tree	60.8 %	71.7 %	80.5 %
Linear Discriminant	61.3 %	71.2 %	75.1 %
Quadratic Discriminant	62.9 %	71.1 %	71.4 %
Linear SVM	64.2 %	81.6 %	81.0 %
Quadratic SVM	64.4 %	81.9 %	84.5 %
Cubic SVM	64.9 %	80.8 %	82.7 %
Fine Gaussian SVM	57.7 %	61.2 %	56.4 %
Medium Gaussian SVM	66.9 %	82.6 %	84.7 %
Coarse Gaussian SVM	64.6 %	79.4 %	81.4 %
Fine KNN	59.6 %	72.8 %	76.1 %
Medium KNN	64.3 %	76.8 %	81.2 %
Coarse KNN	67.0 %	73.0 %	79.6 %
Cosine KNN	64.1 %	77.0 %	80.7 %
Cubic KNN	64.8 %	76.9 %	79.8 %
Weighted KNN	64.2 %	75.6 %	81.4 %
Ensemble/Boosted Trees	67.7 %	80.7 %	84.3 %
Ensemble/Bagged Trees	65.7 %	80.4 %	84.5 %
Ensemble/Sub Discriminant	65.0 %	80.4 %	81.1 %
Ensemble/Subspace KNN	67.6 %	76.5 %	82.3 %
Ensemble/RUSBoosted Trees	61.6 %	75.8 %	81.5 %

7.5. Análisis con momentos en el modelo de color YCbCr

En el Modelo de color YCbCr se calcula los primeros 90 momentos Cuaternión de Zernike. De igual manera se obtienen los primeros 30 momentos Multicanal de Zernike y en el caso el clasificador Multikernel en el Modelo de color YCbCr se calculó los primeros 10 momentos Multicanal de la base de datos HAM10000, así como los primeros 30 momentos de Zernike del gradiente de la imagen en color y por último el histograma de la imagen de la cual se toman 10 bins de cada canal normalizados en un rango de 0 a 1. En todos los casos se obtuvo una matriz de descriptores de 4423×90 . Los resultados de la primera clasificación con 7 etiquetas se muestran en la Tabla 7.5, para los momentos Cuaternión, momentos Multicanal y Clasificador Multikernel respectivamente.

7. Clasificación de Melanoma

Tabla 7.5: Exactitud para los diferentes clasificadores con momentos Cuaternión del Modelo de color YCbCr usando 7 etiquetas.

Clasificador	Cuaternión	Multicanal	Multikernel
Complex Tree	40.2 %	48.0 %	53.0 %
Medium Tree	40.5 %	48.3 %	52.9 %
Simple Tree	37.1 %	44.1 %	49.6 %
Linear Discriminant	35.5 %	44.4 %	47.6 %
Quadratic Discriminant	29.6 %	39.2 %	27.0 %
Linear SVM	48.5 %	57.2 %	57.5 %
Quadratic SVM	48.6 %	58.6 %	60.2 %
Cubic SVM	49.0 %	58.3 %	59.8 %
Fine Gaussian SVM	38.2 %	38.6 %	37.4 %
Medium Gaussian SVM	49.8 %	58.7 %	59.9 %
Coarse Gaussian SVM	45.4 %	52.4 %	55.5 %
Fine KNN	37.7 %	48.1 %	50.2 %
Medium KNN	43.0 %	51.5 %	53.8 %
Coarse KNN	44.9 %	50.7 %	54.0 %
Cosine KNN	44.1 %	51.7 %	53.7 %
Cubic KNN	44.0 %	51.1 %	53.1 %
Weighted KNN	43.4 %	52.3 %	55.0 %
Ensemble/Boosted Trees	43.2 %	51.8 %	54.7 %
Ensemble/Bagged Trees	46.5 %	54.6 %	58.4 %
Ensemble/Sub Discriminant	47.2 %	54.9 %	56.1 %
Ensemble/Subspace KNN	46.5 %	55.8 %	50.4 %
Ensemble/RUSBoosted Trees	35.4 %	48.0 %	49.7 %

7. Clasificación de Melanoma

Los resultados para la clasificación con las clases, Nevo y Melanoma se muestran en las Tablas 7.6, para los momentos Cuaternión, momentos Multicanal y Clasificador Multikernel respectivamente.

Tabla 7.6: Exactitud para los diferentes clasificadores con momentos Cuaternión, Multicanal y Clasificador Multikernel de Zernike del Modelo de color YCbCr usando 2 clases.

Clasificador	Cuaternión	Multicanal	Multikernel
Complex Tree	68.6 %	77.0 %	78.8 %
Medium Tree	68.6 %	78.9 %	81.1 %
Simple Tree	67.8 %	79.7 %	81.8 %
Linear Discriminant	75.0 %	75.8 %	78.1 %
Quadratic Discriminant	74.6 %	74.3 %	52.6 %
Linear SVM	78.3 %	82.7 %	82.9 %
Quadratic SVM	77.2 %	83.0 %	83.2 %
Cubic SVM	76.3 %	81.9 %	81.9 %
Fine Gaussian SVM	68.2 %	68.1 %	67.0 %
Medium Gaussian SVM	79.3 %	83.8 %	84.1 %
Coarse Gaussian SVM	78.0 %	82.6 %	81.4 %
Fine KNN	70.8 %	77.5 %	77.9 %
Medium KNN	74.3 %	80.5 %	81.5 %
Coarse KNN	72.9 %	76.4 %	79.1 %
Cosine KNN	76.5 %	80.0 %	80.4 %
Cubic KNN	74.8 %	80.5 %	80.7 %
Weighted KNN	73.1 %	80.1 %	81.6 %
Ensemble/Boosted Trees	77.6 %	82.4 %	83.3 %
Ensemble/Bagged Trees	77.3 %	82.2 %	83.6 %
Ensemble/Sub Discriminant	77.8 %	82.1 %	81.7 %
Ensemble/Subspace KNN	74.6 %	82.1 %	81.9 %
Ensemble/RUSBoosted Trees	68.9 %	78.9 %	81.0 %

7.6. Análisis con momentos en el modelo de color CMYK

En el Modelo de color CMYK, se calcularon los primeros 90 momentos Cuaternión de Zernike obteniendo una matriz de descriptores de 4423×90 . Para el caso de los momentos Multicanal se calcularon los primeros 23 momentos Multicanal de Zernike de cada canal de color, y se obtuvo una una matriz de descriptores de 4423×92 . El último caso analizado, Clasificador Multikernel de Zernike se calculó los primeros 8 momentos Multicanal de cada canal de color, de igual manera se calcularon los primeros 30 momentos de Zernike de la imagen y el histograma de la imagen de la cual se toman 10 bins de cada canal normalizado en un rango de 0 a 1 ambos en el modelo de color CMY. Los resultados de la primera clasificación con las 7 etiquetas

7. Clasificación de Melanoma

se muestran en la Tablas 7.7 para los momentos Cuaternión, momentos Multicanal y Clasificador Multikernel respectivamente.

Tabla 7.7: Exactitud para los diferentes clasificadores con momentos Cuaternión, Multicanal y Clasificador Multikernel de Zernike CMYK usando 7 etiquetas.

Clasificador	Cuaternion	Multicanal	Multikernel
Complex Tree	45.7 %	49.0 %	51.6 %
Medium Tree	47.9 %	49.0 %	52.7 %
Simple Tree	44.9 %	44.7 %	49.1 %
Linear Discriminant	37.7 %	45.5 %	46.6 %
Quadratic Discriminant	30.3 %	24.4 %	44.1 %
Linear SVM	49.7 %	58.3 %	55.8 %
Quadratic SVM	49.9 %	59.5 %	59.2 %
Cubic SVM	48.0 %	58.7 %	57.9 %
Fine Gaussian SVM	38.7 %	40.8 %	39.2 %
Medium Gaussian SVM	50.8 %	59.9 %	58.2 %
Coarse Gaussian SVM	49.2 %	53.4 %	54.9 %
Fine KNN	38.1 %	47.9 %	47.3 %
Medium KNN	44.9 %	52.9 %	53.1 %
Coarse KNN	48.2 %	51.8 %	53.7 %
Cosine KNN	46.7 %	53.0 %	53.0 %
Cubic KNN	44.9 %	53.4 %	52.2 %
Weighted KNN	45.5 %	54.1 %	53.9 %
Ensemble/Boosted Trees	49.5 %	51.3 %	54.5 %
Ensemble/Bagged Trees	48.7 %	55.5 %	56.3 %
Ensemble/Sub Discriminant	48.6 %	55.8 %	55.0 %
Ensemble/Subspace KNN	49.2 %	57.0 %	48.7 %
Ensemble/RUSBoosted Trees	35.4 %	47.4 %	46.6 %

Los resultados obtenidos en la segunda clasificación con las clases Nevo y Melanoma se muestran en la Tabla 7.8 para los momentos Cuaternión, momentos Multicanal y Clasificador Multikernel respectivamente.

7. Clasificación de Melanoma

Tabla 7.8: Exactitud para los diferentes clasificadores con momentos Cuaternión, Multicanal y Clasificador Multikernel de Zernike CMYK usando 2 etiquetas.

Clasificador	Cuaternión	Multicanal	Multikernel
Complex Tree	75.1 %	77.8 %	77.9 %
Medium Tree	78.4 %	80.1 %	80.4 %
Simple Tree	77.4 %	80.8 %	80.3 %
Linear Discriminant	75.0 %	73.4 %	76.6 %
Quadratic Discriminant	74.1 %	66.5 %	76.8 %
Linear SVM	80.6 %	82.3 %	76.8 %
Quadratic SVM	78.2 %	81.9 %	84.8 %
Cubic SVM	77.7 %	80.2 %	82.4 %
Fine Gaussian SVM	68.5 %	71.5 %	68.5 %
Medium Gaussian SVM	81.1 %	83.9 %	83.1 %
Coarse Gaussian SVM	79.7 %	82.7 %	80.2 %
Fine KNN	71.3 %	76.8 %	76.8 %
Medium KNN	76.8 %	81.0 %	80.7 %
Coarse KNN	74.8 %	79.4 %	79.8 %
Cosine KNN	77.1 %	80.5 %	79.6 %
Cubic KNN	76.1 %	80.8 %	80.0 %
Weighted KNN	75.1 %	80.7 %	80.6 %
Ensemble/Boosted Trees	81.4 %	82.6 %	82.3 %
Ensemble/Bagged Trees	80.1 %	81.9 %	83.6 %
Ensemble/Sub Discriminant	79.5 %	82.3 %	80.5 %
Ensemble/Subspace KNN	80.5 %	82.9 %	79.6 %
Ensemble/RUSBoosted Trees	78.3 %	80.2 %	79.7 %

7. Clasificación de Melanoma

7.7. Análisis de resultados

Los resultados obtenidos con los momentos Cuaternión, los momentos Multicanal y el Clasificador Multikernel en los distintos Modelos de color se resumen en la Tabla 7.9. Como puede observarse, la mayor exactitud de clasificación para 7 etiquetas se obtuvo con los momentos Multicanal en el Modelo de color HSI siendo el modelo Quadratic SVM el que presentó la exactitud más alta, clasificando el 72,1%. Para la clasificación con las dos etiqueta Nevo y Melanoma la mayor exactitud de clasificación se obtuvo con el Clasificador Multikernel en el Modelo de color HSI y CMYK con una exactitud del 84.8% en los modelos Medium Gaussian SVM y Quadratic SVM. Cabe mencionar que tanto los momentos Multicanal como el Clasificador Multikernel tienen un rendimiento similar en la clasificación de Melanoma. Los resultados obtenidos muestran el desempeño de los momentos ante un problema real donde las formas que presentan las imágenes son completamente distintas entre sí, y de igual manera es importante considerar que se usó un orden bajo que caracteriza forma sin llegar a detalles que podrían obtenerse con órdenes superiores.

Tabla 7.9: Exactitud de clasificación con los momentos Cuaternión, Multicanal y Clasificador Multikernel de Zernike.

momentos	Modelo	Clasificador	7 etiquetas	2 etiquetas
Cuaternión	RGB	Medium Gaussian SVM / Coarse Gaussian SVM	50.0 %	78.5 %
Multicanal	RGB	Quadratic SVM / Medium Gaussian SVM	59.4 %	83.9 %
Multikernel	RGB	Quadratic SVM / Ensemble/Bagged Trees	61.2 %	84.3 %
Cuaternión	HSI	Ensemble/Subspace KNN / Ensemble/Boosted Trees	38.5 %	67.7 %
Multicanal	HSI	Quadratic SVM / Medium Gaussian SVM	72.1 %	82.6 %
Multikernel	HSI	Medium Gaussian SVM / Medium Gaussian SVM	59.5 %	84.8 %
Cuaternión	YCbCr	Medium Gaussian SVM / Medium Gaussian SVM	49.8 %	79.3 %
Multicanal	YCbCr	Medium Gaussian SVM / Medium Gaussian SVM	58.7 %	83.8 %
Multikernel	YCbCr	Quadratic SVM / Medium Gaussian SVM	60.2 %	84.1 %
Cuaternión	CMYK	Medium Gaussian SVM / Ensemble/Boosted Trees	50.8 %	81.4 %
Multicanal	CMYK	Medium Gaussian SVM / Medium Gaussian SVM	59.9 %	83.9 %
Multikernel	CMYK	Quadratic SVM / Quadratic SVM	59.2 %	84.8 %

Bibliografía

- [1] Simões, M. C. F., J. J. S. Sousa, and A. A. C. C. Pais. "Skin cancer and new treatment perspectives: A review." *Cancer letters* 357.1 (2015): 8-42.
- [2] Herrera González, Norma Estela, and Aramara Yasmín Aco Flores. "El melanoma en México." *Revista de Especialidades Médico-Quirúrgicas* 15.3 (2010): 161-164.
- [3] Codella, Noel, et al. "Deep learning, sparse coding, and SVM for melanoma recognition in dermoscopy images." *International workshop on machine learning in medical imaging*. Springer, Cham, 2015
- [4] Hosny, Khalid M., Mohamed A. Kassem, and Mohamed M. Foad. "Skin Cancer Classification using Deep Learning and Transfer Learning." *2018 9th Cairo International Biomedical Engineering Conference (CIBEC)*. IEEE, 2018.
- [5] Mendes, Danilo Barros, and Nilton Correia da Silva. "Skin Lesions Classification Using Convolutional Neural Networks in Clinical Images." *arXiv preprint arXiv:1812.02316* (2018).
- [6] Tschandl, Philipp, Cliff Rosendahl, and Harald Kittler. "The HAM10000 dataset, a large collection of multi-source dermatoscopic images of common pigmented skin lesions." *Scientific data* 5 (2018): 180161.

Capítulo 8

Conclusiones y trabajo a futuro.

En este trabajo de tesis se presenta un panorama general de los momentos ortogonales radiales y sus invariantes geométricos en los principales métodos para el análisis de imágenes en color; momentos Cuaternión, momentos Multicanal y clasificador multikernel de Zernike. También se analiza su desempeño en cuatro distintos modelos de color. Las aportaciones más relevantes de este trabajo de tesis se presentan a continuación:

- Los experimentos realizados con los momentos Cuaternión, los momentos Multicanal y el clasificador multikernel de Zernike demostraron su robustez en las pruebas de invariancias geométricas de rotación escala y traslación en los cuatro modelos de color. Siendo muy similar el desempeño del clasificador multikernel y los momentos Multicanal, los cuales son capaces de reconocer tanto la forma del objeto como el color, por lo que alcanzan porcentajes de clasificación altos. Los momentos Cuaternión en el modelo de color RGB presentan una limitada capacidad de clasificación. Sin embargo, en otros modelos de color, por ejemplo en YCbCr, los resultados obtenidos son comparables con los momentos Multicanal. Además, en este caso también se debe considerar la carga computacional adicional que presentan los momentos Cuaternión.
- El cómputo de los momentos Multicanal de Zernike en este trabajo de tesis presentan una mayor estabilidad numérica en órdenes superiores que otros métodos. También, obtienen un mejor desempeño en las pruebas realizadas de invariancia geométrica, robustez al ruido y reconstrucción de imagen en comparación con resultados publicados en la literatura usando la base de datos COIL-100.
- La clasificación lograda en el experimento realizado con la base de datos de imágenes de melanoma presenta una clasificación menor a la reportada en

8. Conclusiones y trabajo a futuro.

publicaciones donde se usa redes neuronales. Sin embargo, en este trabajo de tesis aporta información sobre el comportamiento de los descriptores obtenidos por los tres métodos usados en los diferentes modelos de color. El análisis presentado es susceptible de mejorar considerando que el orden de los momentos usado fue $Z_{\text{máx}} = 7$, lo cual es un orden bajo que representa características generales de la imagen. Cabe mencionar que este tipo de aplicaciones de los momentos a bases de datos de problemas reales no han sido reportados en la literatura.

- El análisis realizado mediante la aplicación de *Classification Learner* de Matlab provee modelos de clasificación para problemas multiclase y permite decidir qué algoritmo puede generar una mejor clasificación. De acuerdo a los experimentos realizados los modelos basados en algoritmos de máquinas de soporte vectorial y vecinos cercanos funcionan mejor clasificando los descriptores.
- Como trabajo a futuro y en conjunto con el equipo de la Universidad de Alcalá, se está trabajando en la elaboración de una publicación con los resultados obtenidos de los momentos Multicanal de Zernike y Tchebichef. Actualmente se está mejorando los métodos de clasificación, en especial la máquina de soporte vectorial. También, se trabaja en el contenido una publicación posterior en la cual se desarrolla en una nueva implementación de momentos Multicanal de Tchebichef con transformada discreta del coseno y con el análisis de los resultados obtenidos en los diferentes modelos de color.

Apéndice A

Modelos de Clasificación

La aplicación de Matlab Classification Learner entrena modelos para clasificar datos por medio de aprendizaje automático supervisado. El aprendizaje supervisado, se basa en el conocimiento previo de las clases o etiquetas de un conjunto de datos que son utilizados para entrenar un modelo que genera predicciones sobre nuevos datos de entrada. Classification Learner utiliza una variedad de algoritmos de aprendizaje para realizar la clasificación, los cuales se mencionan en la Tabla 1

Tabla A.1: Clasificadores en la aplicación Classification Learner de Matlab.

Clasificador	
Árboles de decisión	Complex Tree, Medium Tree, Simple Tree
Análisis discriminante	Linear Discriminant, Quadratic Discriminant
Máquinas de soporte vectorial	Linear SVM, Quadratic SVM, Cubic SVM, Fine Gaussian SVM, Medium Gaussian SVM, Coarse Gaussian SVM
Vecinos cercanos	Fine KNN, Medium KNN, Coarse KNN, Cosine KNN, Cubic KNN, Weighted KNN
Ensemble	Boosted Trees, Bagged Trees, Sub Discriminant Subspace KNN, RUSBoosted Trees

A.1. Árboles de decisión.

El árbol de decisión es un algoritmo de agrupamiento jerárquico, tiene tres tipos de nodos [2]:

- Un nodo raíz, sin enlaces de entrada y cero o más enlaces de salida.
- Nodos internos, los cuales tienen un enlace entrante y dos o más de salida.
- Nodos hojas o terminales con un enlace de entrada y dos o más de salida.

A. Modelos de Clasificación

Cada nodo hoja está asociado a una etiqueta de clase. Los nodos no terminales, es decir la raíz y los nodos internos contienen atributos. La clasificación comienza desde el nodo raíz aplicando una condición de prueba de atributo, siguiendo por el nodo interno apropiado en función del resultado de la prueba hasta otro nodo interno donde nuevamente una condición de prueba de atributos es aplicada este proceso continúa hasta llegar a un nodo hoja, un ejemplo básico de su estructura se observa en la Fig. 1. Una vez que se alcanza el nodo hoja se le asigna la etiqueta de clase asociada.

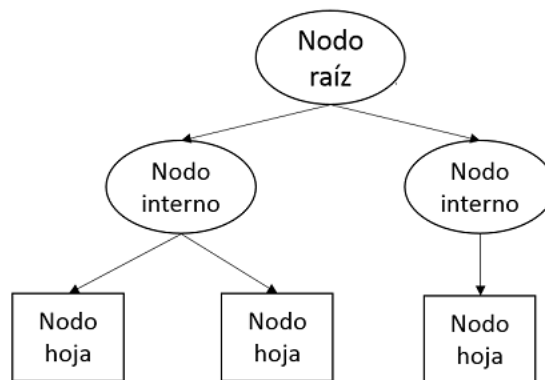


Figura A.1: Estructura básica de un árbol de decisión.

A.2. Análisis discriminante

El análisis discriminante es un método de clasificación supervisada que se utiliza para la clasificación de distintos datos en grupos o clases. Este método equivale a un análisis de regresión donde la variable categórica es la clase y las variables independientes son continuas y determinan a qué clase pertenece un objeto [3]. El análisis discriminante lineal también se conoce como el discriminante Fisher, llamado así por su inventor, y es ampliamente utilizado en aplicaciones que involucran datos de alta dimensión como reconocimiento facial y recuperación de imágenes [4].

A.3. Máquinas de soporte vectorial.

Una máquina de soporte vectorial es un método de aprendizaje supervisado para la clasificación, que genera funciones de mapeo de entrada-salida a partir de un conjunto de datos de entrenamiento etiquetados. Se tiene una función de mapeo $y = f(x)$ entre un vector de entrada x y un escalar y de salida donde la única información disponible es un conjunto de entrenamiento $D = \{(x_i, y_i) \in X \times Y\}$, $i = 1, 2, 3, \dots, l$,

A. Modelos de Clasificación

donde l representa el número de datos de entrenamiento. La máquina de vectores de soporte utilizan un hiperplano de separación lineal para crear un clasificador con un margen máximo [5]. Si los datos no son linealmente separables encuentra el hiperplano que maximiza el margen de separación entre dos clases como se puede observar en la Fig 2. Primero, la máquina de vectores transforma los predictores o datos de entrada en un espacio de características dimensionales superior; esta transformación se logra mediante el uso de kernels como: polinomial, lineal, sigmoideal como en los perceptrones multicapa y las asignaciones que tienen como base las funciones de carácter radial o gaussiana. De esta manera los datos de entrada se vuelven más separables en comparación con el espacio de entrada original. Después de este paso se crean los hiperplanos de margen máximo. La máquina de soporte vectorial ha demostrado un rendimiento altamente competitivo en numerosos aplicaciones del mundo real, como la bioinformática, la minería de textos, el reconocimiento facial, y el procesamiento de imágenes [5].

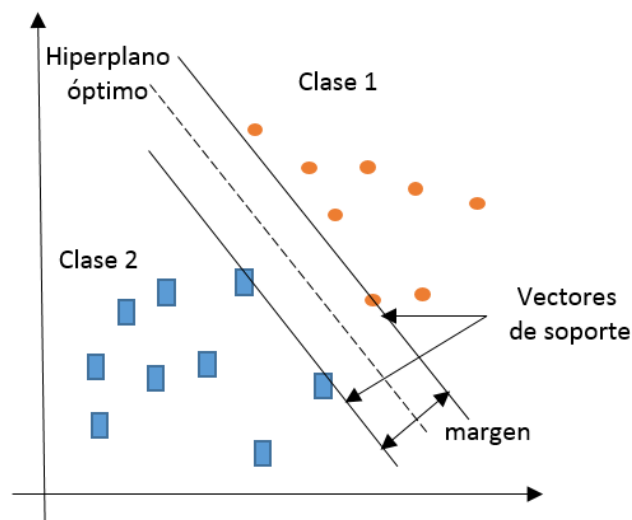


Figura A.2: Hiperplano de separación entre dos clases.

A.4. Clasificadores de vecinos cercanos

El método de vecinos más cercanos es un método de clasificación supervisada, el cual almacena los datos disponibles, y clasifica un objeto según una medida de similitud dada por una función de distancia. El objeto es asignado a la clase más común entre sus K vecinos cercanos, donde k es un entero positivo, si $k=1$ el objeto se asigna a la clase

A. Modelos de Clasificación

de su vecino más cercano. La clasificación KNN se realiza midiendo la distancia entre el objeto con los objetos conocidos. Para variables continuas las funciones distancia a utilizar son,

$$\begin{aligned} \textit{Euclideana} & \sqrt{\sum_{i=1}^k (x_i - y_i)^2} \\ \textit{Manhattan} & \sum_{i=1}^k |x_i - y_i| \\ \textit{Minkowski} & \left(\sum_{i=1}^k (|x_i - y_i|^q) \right)^{1/q} \end{aligned} \quad (\text{A.1})$$

en el caso de variables categóricas se debe usar la distancia de Hamming,

$$\begin{aligned} D_H &= \sum_{i=1}^k |x_i - y_i| \\ x &= y \Rightarrow D = 0 \\ x &\neq y \Rightarrow D = 1 \end{aligned} \quad (\text{A.2})$$

El método de clasificación de K vecinos cercanos es simple y es uno de los más utilizados en diversas áreas como por ejemplo: bioinformática, procesamiento de imágenes, recuperación de documentos y visión por computadora [6].

A.5. Ensemble

El método ensemble o de conjunto son algoritmos de aprendizaje que construyen un conjunto de clasificadores y luego clasifican nuevos datos de entrada tomando un voto ponderado de sus predicciones. Se han desarrollado muchos métodos para construir conjuntos, originalmente utiliza el promedio bayesiano sin embargo los algoritmos recientes realizan por ejemplo una manipulación del conjunto de entrenamiento llamado *bagging*. En cada ejecución el *bagging* presenta el algoritmo de aprendizaje con un set de entrenamiento de m elementos, este set es una replica del conjunto original de entrenamiento y es llamado *bootstrap*; cada réplica *bootstrap* contiene en promedio 63,2% del set de entrenamiento original [7].

Apéndice B

Modelos de color

El color es la sensación resultante de la estimulación visual por parte de determinadas longitudes de onda. Un modelo de color o sistema de color es una especificación de un sistema de coordenadas y un subespacio, donde cada color está representado por un solo punto. En este anexo se realiza una breve descripción de los cuatro modelos más utilizados.

B.1. Modelo de color RGB

El sistema, de color RGB es el más común en sistemas digitales en donde cada color aparece en sus componentes espectrales primarios rojo, verde y azul, los cuales se suman para reproducir una amplia gama de colores. Este modelo se basa en un sistema de coordenadas cartesiano mostrado en la Fig 1B [8], los valores de R, G, y B están en el rango [0-255] esto debido a que los colores se representan en bytes y estos son colecciones de 8 bits donde 2^8 es 256, cada pixel de color es un triplete de valores (R,G,B) es decir que posee una profundidad de 24 bits. El número total de colores en una imagen RGB de 24 bits es $(2^8)^3 = 16,777,216$ [8].

B.2. Modelo de color CMYK

El modelo de color CMY es un modelo de color sustractivo; sus componentes son el cian, magenta y amarillo, los cuales tienen un valor en el rango [0-255] . La conversión de RGB a CMY se realiza utilizando la operación

$$\begin{bmatrix} C \\ M \\ Y \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} - \begin{bmatrix} R \\ G \\ B \end{bmatrix} \quad (\text{B.1})$$

B. Modelos de color

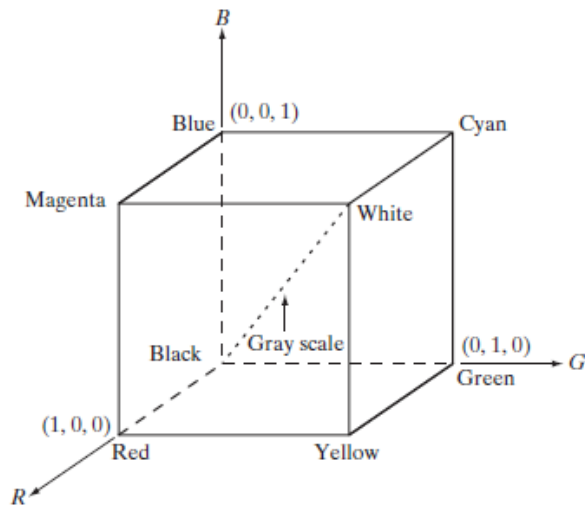


Figura B.1: Esquema del cubo de color RGB.

en el modelo CMY el color negro se obtiene mezclando cian, magenta y amarillo pero no genera un negro puro, por lo que CMY se modifica a CMYK. CMYK es un acrónimo de cian, magenta, amarillo donde k es el color negro que varía en un rango $[0-1]$. para realizar la conversión de CMY a CMYK se realiza la operación [9],

$$\begin{aligned} K &= \min\{C, M, Y\} \\ C &= (C - K)/(1 - K) \\ M &= (M - K)/(1 - K) \\ Y &= (Y - K)/(1 - K) \end{aligned} \tag{B.2}$$

el modelo CMYK es usado comúnmente para la impresión con tintas.

En la Fig 2B podemos observar los canales C, M, Y y el modelo de color CMY de la imagen de “Lena”.

B.3. Modelo de color HSI

El modelo HSI se asemeja a las propiedades de detección de color de la visión humana, tiene tres componentes principales: matiz, saturación y brillo. El componente Hue (H) define el color, la saturación (S) representa la cantidad por la cual el color se diluye con luz blanca, la intensidad o brillo (I) es un descriptor subjetivo que es muy difícil de medir y éste es desacoplado a partir de información de color en una imagen

B. Modelos de color

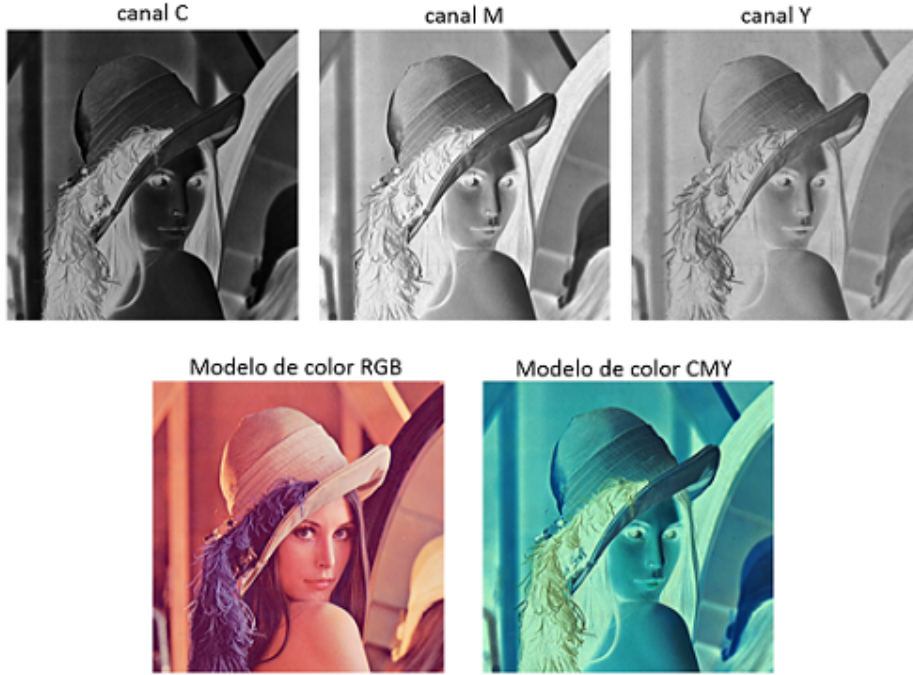


Figura B.2: Modelo de color CMY de la imagen de “Lena” .

[8]. Para transformar del espacio de color RGB a HSI, donde se asume que los valores RGB son normalizados en un rango [0-1], el valor de la intensidad I queda definido como [9],

$$I = \frac{1}{3}(R + G + B) \quad (\text{B.3})$$

El componente de matiz H y la saturación S se describen como

$$H = \begin{cases} \alpha, & B \leq G; \\ 2\pi - \alpha & B > G, \end{cases} \quad (\text{B.4})$$

donde

$$\alpha = \cos^{-1} \frac{(R - G) + (R - B)}{2\sqrt{(R - G)^2 + (R - B)^2 + (G - B)^2}} \quad (\text{B.5})$$

$$S = 1 - \frac{3 \min(R, G, B)}{R + G + B} \quad (\text{B.6})$$

del espacio de color HSI al RGB asumiendo el rango de saturación [0-1], se tienen tres sectores de interés correspondiente a los intervalos de 120° , el primer paso es multiplicar H por 360° para tener el rango original de $[0^\circ - 360^\circ]$. Cuando el valor de H está en el sector $(0^\circ \leq H < 120^\circ)$ se tiene [10]

B. Modelos de color

$$B = I(1 - S), \quad (\text{B.7})$$

$$R = I \left[1 + \frac{S \cos H}{\cos(60^\circ - H)} \right], \quad (\text{B.8})$$

$$G = 3I - (R + B), \quad (\text{B.9})$$

cuando el valor de H está en el sector ($120^\circ \leq H < 240^\circ$), se sustrae 120° de H ,

$$H = H - 120^\circ \quad (\text{B.10})$$

$$R = I(1 - S), \quad (\text{B.11})$$

$$G = I \left[1 + \frac{S \cos H}{\cos(60^\circ - H)} \right], \quad (\text{B.12})$$

$$B = 3I - (R + G), \quad (\text{B.13})$$

finalmente si el valor de H está en el sector ($240^\circ \leq H \leq 360^\circ$), se sustrae 240° de H ,

$$H = H - 240^\circ \quad (\text{B.14})$$

$$G = I(1 - S), \quad (\text{B.15})$$

$$B = I \left[1 + \frac{S \cos H}{\cos(60^\circ - H)} \right], \quad (\text{B.16})$$

$$R = 3I - (G + B), \quad (\text{B.17})$$

En la Fig 3B podemos observar los canales H , S , I y el modelo de color HSI de la imagen de "Lena".

B.4. Modelo de color YCbCr

El espacio de color YCbCr es un sistema de color digital, ampliamente utilizado en video digital que separa los componentes RGB en luminancia o brillo (Y) en un rango [16-235] y cromaticidad (Cb) como concentración de azul y (Cr) en un rango [16-240] como concentración de rojo de la imagen [11]. La conversión del modelo de color RGB al modelo de color YCbCr es

$$\begin{pmatrix} R \\ G \\ B \end{pmatrix} = \begin{pmatrix} +1,000 & -0,000 & +1,403 \\ +1,000 & -0,344 & -0,714 \\ +1,000 & +1,773 & +0,000 \end{pmatrix} \begin{pmatrix} Y \\ Cb - 128 \\ Cr - 128 \end{pmatrix} \quad (\text{B.18})$$

B. Modelos de color



Figura B.3: Modelo de color HSI de la imagen de “Lena” de tamaño 220×220 .

La conversión del modelo de color YCbCr al modelo de color RGB es

$$\begin{pmatrix} Y \\ Cb \\ Cr \end{pmatrix} = \begin{pmatrix} +0,299 & +0,587 & +0,114 \\ -0,169 & -0,331 & +0,500 \\ +0,500 & -0,419 & -0,081 \end{pmatrix} \begin{pmatrix} R \\ G \\ B \end{pmatrix} + \begin{pmatrix} 0 \\ 128 \\ 128 \end{pmatrix} \quad (\text{B.19})$$

En la Fig 4B podemos observar los canales Y, Cb, Cr y el modelo de color YCbCr de la imagen de “Lena”.

B. Modelos de color

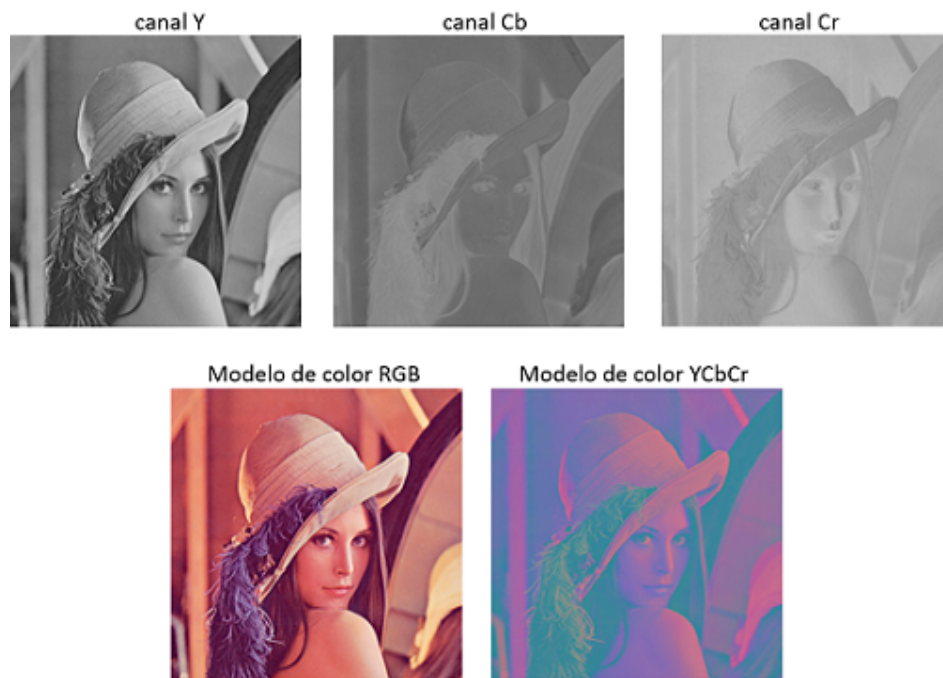


Figura B.4: Canales Y, Cb, Cr y el modelo de color YCbCr de la imagen de “Lena” de tamaño 220×220 .

Bibliografía

- [1] Schrider, Daniel R., y Andrew D. Kern. "Supervised Machine Learning for Population Genetics: A New Paradigm". Trends in Genetics 34.4 (2018): 301-312.
- [2] Tan, Pang-Ning. Introduction to data mining. Pearson Education India, 2018.
- [3] Ye, Jieping, Ravi Janardan, and Qi Li. "Two-dimensional linear discriminant analysis..Advances in neural information processing systems. 2005.
- [4] The MathWorks, Inc.(2019)recuperado de <https://www.mathworks.com/help/stats/feature-selection-and-feature-transformation.html>
- [5] Wang, Lipo, ed. Support vector machines: theory and applications. Vol. 177. Springer Science & Business Media, 2005.
- [6] The MathWorks, Inc.(2019)recuperado de <https://www.mathworks.com/help/stats/classification-using-nearest-neighbors.html>
- [7] Dietterich, Thomas G. "Métodos de conjunto en el aprendizaje automático". Taller internacional sobre sistemas clasificadores múltiples . Springer, Berlín, Heidelberg, 2000.
- [8] Gonzalez, Rafael C. "Digital Image Processing/Richard E."Woods. Interscience, NY (2001).
- [9] Ibraheem, Noor A., et al. Understanding color models: a review..RPN Journal of science and technology 2.3 (2012): 265-275.
- [10] Gong, Li-Hua, et al. "Single channel quantum color image encryption algorithm based on HSI model and quantum Fourier transform."International Journal of Theoretical Physics 57.1 (2018): 59-73.
- [11] Patvardhan, C., Pragyesh Kumar, and C. Vasantha Lakshmi. Effective Color image watermarking scheme using YCbCr color space and QR code."Multimedia Tools and Applications (2018): 1-23.