



UNIVERSIDAD NACIONAL
AUTÓNOMA DE
MÉXICO

UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

**PROGRAMA DE MAESTRIA Y DOCTORADO EN
INGENIERIA**

FACULTAD DE INGENIERÍA ELÉCTRICA

*Descriptores de imágenes digitales
con momentos de Zernike*

T E S I S

QUE PARA OPTAR POR EL GRADO DE:

MAESTRO EN INGENIERIA

PROCESAMIENTO DIGITAL DE SEÑALES

P R E S E N T A :

Ing. Karina Viridiana Sánchez López

TUTOR:

Dr. Francisco García Ugalde



2011



Universidad Nacional
Autónoma de México

Dirección General de Bibliotecas de la UNAM

Biblioteca Central



UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

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

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

Resumen

En este trabajo se estudian los momentos de Zernike como extractores de características como una herramienta del proceso de reconocimiento de patrones. Los momentos de Zernike han demostrado un buen desempeño en imágenes reales (satelitales, de ultrasonido, microscópicas) y digitales. Entre otras de sus cualidades, teóricamente los momentos de Zernike son invariantes a la traslación, escala y rotación. Sin embargo, debido a errores de precisión y a la diferencia de métodos de invariancia, en la elección del clasificador no se puede suponer una invariancia incondicional, existiendo un porcentaje de variación en los cálculos.

Con el fin de medir este porcentaje, en este trabajo se realizan pruebas de invariancia con el algoritmo rápido propuesto en [16] en imágenes sometidas a transformaciones de traslación, escala y rotación. También se efectúan pruebas con el algoritmo desarrollado en [3] que, aprovechándose de la naturaleza de los momentos de Zernike y con la ayuda de la distancia euclidiana obtiene el ángulo de rotación de una imagen con respecto a otra. En este caso las pruebas se enfocan a comprobar la precisión del algoritmo con imágenes con invariancia en traslación, escala y rotación; mismas que se usan en la prueba anterior.

Agradecimientos

A CONACyT, por brindarme el apoyo económico para realizar mis estudios de posgrado.

A UNAM, por dejarme ser parte de su comunidad y mejor mi estilo de vida en el aspecto intelectual, físico y cultural.

A mis profesores, por su paciencia y sabiduría.

A mi familia y amigos, por su apoyo incondicional y su fe en mí.

Índice general

1. Introducción	7
1.1. Antecedentes	9
1.2. Objetivo	11
1.3. Justificación	11
1.4. Contenido	12
2. Generalidades sobre el reconocimiento de patrones	13
2.1. Reconocimiento supervisado	15
2.1.1. Clasificadores Bayesianos	15
2.1.2. Clasificadores lineales	16
2.1.3. Clasificadores no lineales	17
2.2. Reconocimiento no supervisado	17
2.2.1. Agrupamiento jerárquico	19
2.2.2. Agrupamiento secuencial	19
2.2.3. Agrupamiento basado en optimización de la función de costo	19
2.3. Problemas de clasificación	20
2.4. Descriptores	20
2.4.1. Descripción por medio de contornos	20
2.4.2. Descripción por medio de regiones	22
3. Teoría de los Momentos Geométricos	25
3.1. Momentos no ortogonales	27
3.2. Momentos ortogonales	29
3.3. Momentos de Zernike	30
3.3.1. Uso de un núcleo geométrico	32
3.3.2. Comparación de momentos usando un ángulo de rotación	33
4. Evaluación de algoritmos	37
4.1. Validación de la invariancia en los momentos de Zernike (MZ)	37
4.1.1. Invariancia respecto a la traslación	39
4.1.2. Invariancia respecto a la escala	43
4.1.3. Invariancia respecto a la rotación	46
4.1.4. Invariancia respecto a la traslación y a la escala	47
4.1.5. Invariancia respecto a la traslación y a la rotación	49

Índice general

4.1.6.	Invariancia respecto a la escala y a la rotación	51
4.1.7.	Invariancia respecto a la traslación, la escala y la rotación	53
4.2.	Invariancia respecto al ángulo de rotación entre dos imágenes	55
4.2.1.	Prueba con diferentes rotaciones	56
4.2.2.	Prueba con diferentes rotaciones y escalas	59
4.2.3.	Prueba con diferentes rotaciones y traslaciones	60
4.2.4.	Prueba con diferentes rotaciones, escalas y traslaciones	61
5.	Conclusión	65
A.		69
A.1.	Momentos de Hu invariantes.	69
A.2.	Momentos invariantes de Zernike	69
B.		73
B.1.	Algoritmo de cálculo de matriz de MZ	73
B.2.	Algoritmo de cálculo de matriz de H's	74
B.3.	Algoritmo de cálculo de MZ	76
C.		79
C.1.	Algoritmo de búsqueda de distancia mínima entre dos imágenes	79
C.2.	Algoritmo de cálculo de función F y Fd	81

1. Introducción

Uno de los objetivos de la ingeniería, consiste en buscar la automatización de tareas repetitivas, como puede ser la separación de frutas de acuerdo a su tamaño, o la identificación y separación de billetes falsos.

La tarea de distinguir frutas de acuerdo a su tamaño puede resolverse con el uso de sensores que operen en función de umbrales de tamaño, previamente establecidos, y de acuerdo a esos umbrales se podría clasificar a las frutas en diferentes tamaños. En el caso de los billetes sin embargo, es evidente que un solo tipo de sensor no es suficiente. Para esta categoría de problemas resulta muy eficaz el uso de la visión artificial, un conjunto de herramientas diseñadas para automatizar la identificación de objetos en una imagen. Inclusive en la aplicación de la clasificación de las frutas por tamaño, donde el problema de clasificarlas podría parecer sencillo, la visión artificial es capaz de alcanzar varios objetivos: tamaños, colores, formas, calidad de la fruta, etc.

En un proceso de automatización se utilizan herramientas y procesos que en su conjunto forman un *sistema artificial inteligente*. Los sistemas artificiales inteligentes necesitan interactuar con el medio ambiente en el que se encuentran a través de sensores, que ayudan al sistema a entender como está conformado el medio. En otras palabras, un sensor es un elemento a través del cuál el sistema puede reconocer el medio ambiente con el que interactúa.

Frecuentemente, para identificar al medio el sistema considera ciertos parámetros clave de referencia, tales como la temperatura, el olor, la luminosidad, algunos objetos u otras características del medio. Cada sensor está diseñado para “capturar” el valor de uno o varios de estos parámetros, para que el sistema interprete estos valores y pueda identificar el medio.

A grandes rasgos, los sistemas artificiales inteligentes realizan el proceso de reconocimiento de la siguiente manera. El sistema recibe información a través de los sensores y/o transductores con los que se encuentre dotado. Esta información es analizada y comparada con una base de datos, si la información adquirida tiene congruencia con algún elemento de la base de datos, se dice que el sistema ha reconocido un elemento.

Los sensores son dispositivos capaces de transformar magnitudes físicas o químicas (llamadas variables de instrumentación) en magnitudes eléctricas. La temperatura, intensidad lumínica, distancia, aceleración, inclinación, desplazamiento, presión, fuerza, torsión, humedad, pH, etc. [18], son algunos ejemplos.

En particular existe una gran variedad de sensores construidos para analizar diferentes tipos y rangos de radiación electromagnética, generando gran cantidad de información del

1. Introducción

ambiente que en muchos casos el ojo humano es incapaz de ver. Toda la información obtenida por estos sensores es dispuesta en matrices de valores conocidas comúnmente como imágenes, las cuales pueden provenir de señales acústicas, magnéticas, gravimétricas, de radar, neurográficas y micrográficas, entre otras.

Por lo tanto cuando se desea analizar un medio, basta con analizar las imágenes que proyecta, Figura 1.0.1. El nivel de confianza que se tiene en las señales emitidas por el ambiente y los sensores electromagnéticos que las capturan es muy alto. Sólo basta con mencionar que en la medicina, infinidad de diagnósticos se basan en radiografías, ultrasonidos y resonancias magnéticas. También la industria pone en manos de la visión artificial procesos importantes como son el control de calidad. Finalmente, un ejemplo que vemos comúnmente en los televisores es el seguimiento con un puntero sobre un jugador en tiempo real en un partido de fútbol.

La adquisición de imágenes es sólo el primer paso para el reconocimiento de objetos. En líneas generales, el proceso de interpretación de imágenes involucra las etapas de segmentación, extracción de características y clasificación. El reconocimiento de objetos se hace más difícil cuando los objetos de interés en una imagen vienen acompañadas de información extra del medio o del instrumento de captura que no sabemos interpretar, y que es conocida comúnmente como ruido.

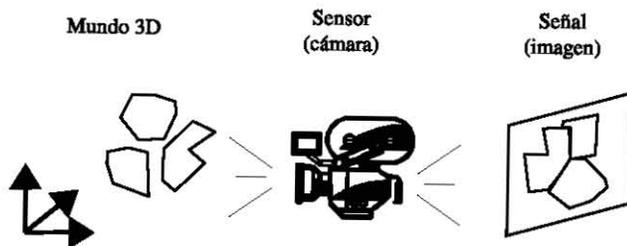


Figura 1.0.1.: Representación de una escena 3D a una de 2D.

Por consiguiente, el reconocimiento es una correspondencia de la representación propia del medio con la información sensorial, obtenida en este caso por medio de un sistema de visión. De manera general, un descriptor es un parámetro representativo de una imagen que se utiliza para realizar su correspondencia con uno o varios objetos. Existen varios métodos de hacer corresponder los descriptores con los objetos, dependiendo de la información que se precise de cada objeto, o grupo de objetos.

Resumiendo, tenemos que por medio de sensores se capturan imágenes del entorno. Estas imágenes nos interesan porque son una representación de nuestro ambiente, y poseen descriptores con los cuales podemos clasificar la información que contienen y/o inclusive las propias imágenes, Figura 1.0.2.

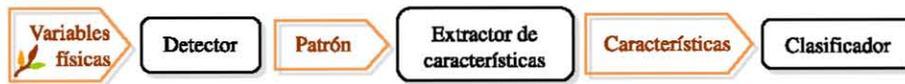


Figura 1.0.2.: Proceso de clasificación

La elección de los descriptores depende del objetivo, respecto a lo que buscamos clasificar. Por ejemplo, el área del objeto es un buen descriptor si se busca clasificar objetos por su tamaño. Cuando se busca describir la forma de los objetos se cuenta con descriptores dispuestos en dos grandes tipos: los descriptores de contornos y los descriptores de regiones.

Algunos de los problemas a los que se enfrentan este tipo de descriptores se presentan cuando las imágenes son de diferente escala, cuando tienen diferente orientación, o cuando están recubiertas (oclusión).

Con la finalidad de seleccionar este tipo de problemas, en esta tesis se trabaja con un tipo de descriptores de regiones conocidos como *momentos geométricos*. Estos descriptores tienen la ventaja de ser invariantes ante la escala, traslación y rotación, mediante operaciones simples.

De acuerdo a la teoría de señales, los momentos geométricos son el producto punto entre dos funciones: la imagen y una función de base. Los momentos pueden usarse con diferentes tipos de bases, según la naturaleza de las base, ortogonal o no ortogonal, se define el tipo de momento. Como es de esperarse, las bases no ortogonales originan momentos con información redundante. En esta tesis no se desea redundancia en información, por lo cual se ha decidido trabajar con una base ortogonal que utiliza los polinomios de Zernike. Por ende, los descriptores aquí empleados son los Momentos Geométricos de Zernike [21].

1.1. Antecedentes

Históricamente fue en los años veinte la primera experiencia de la transmisión de una imagen por medio de un cable submarino entre Londres y Nueva York. Sin embargo fue hasta los años cincuenta y sesenta en que aparecen las primeras computadoras digitales y con ellas la necesidad de procesar y transmitir imágenes satélites, dando pie al desarrollo de técnicas para visión, tratamiento y análisis de imágenes digitales. Por mencionar algunos proyectos, se tiene el caso del National Institute of Health donde se ha trabajado con imágenes de rayos X, microscopía óptica y electrónica. Por su parte en el MIT y en la universidad de Stanford se ha trabajado en visión aplicada en robótica, generando el Pattern Information Processing System (PIPS) y el Understanding Image System (UIS), proyectos que desarrollaron importantes técnicas de visión. En l'Ecole Nationale Supérieure des Mines de Francia, G. Matheron y J. Serra, trabajaron en la caracterización y medición de formas en base al método denominado "morfología matemática" [6].

La madurez en el diseño de arquitecturas de computadoras en la década de los ochenta impulsa el desarrollo de las técnicas de análisis de imágenes digitales, dando origen a nu-

1. Introducción

merosas aplicaciones. Uno de los investigadores que definieron el rumbo del procesamiento de imágenes fue David Courtnay Marr [19] quién desarrolló algoritmos para la detección de características (bordes, líneas, texturas), así como técnicas de segmentación. En el área industrial se han desarrollado técnicas para reconocimiento, haciendo uso del área de inteligencia artificial.

A mediados de los ochentas la atención se centró en nuevas teorías y algoritmos para la interpretación de imágenes en 2D, resultado de la proyección en un plano de escenas en 3D. Así como el uso de flujo óptico para la reconstrucción de escenas con base en sus diferentes proyecciones. La caracterización de superficies a partir del estudio del movimiento, el estudio de las formas a partir de las sombras, el estudio de la orientación a partir de texturas y técnicas de representación y búsqueda de objetos.

Paralelamente al desarrollo de las técnicas y métodos matemáticos, se desarrollan diferentes arquitecturas computacionales específicas para el procesamiento de imágenes digitales. Lo que posibilita la experimentación con imágenes más complejas de interpretar.

Una herramienta utilizada a principios del siglo XX es la transformada de ondula, la cual trabaja con una amplia variedad de ondulas madre, siendo la ondeleta Haar la más sencilla. Para el reconocimiento de patrones, en imágenes de un solo objeto. Se hace una extracción de bordes, ya que este tipo de imágenes se caracterizan por ser imágenes de componente localizable o componente transitorio, es decir, los bordes del objeto de la imagen ante una transformada de ondula son localizables. Pero la textura del objeto es desechada. Si se requiere hacer un análisis de textura el procesamiento de la imagen es diferente. De hecho la transformada de ondula también se usa para clasificación y eliminación de ruido. En otros casos trabaja en conjunto con otras herramientas de procesamiento de señales para expandir su aplicación. Por ejemplo para invariancia en rotación combina los momentos geométricos con la ondeleta madre, conocida como transformada de ondeleta de momentos invariantes. Aprovechándose de la multiresolución que brinda la transformada de ondeleta es posible hacer reconocimiento de siluetas de objetos con invariancia en escala [15].

En 1999 David Lowe publica en [14] un algoritmo llamado SIFT (Scale-Invariant Feature Transform) que es una transformada que proporciona características invariantes en escala. Adicionalmente es invariante a traslación, rotación y oclusiones parciales. Esto es posible gracias a una red de puntos clave que toma como referencia en cada imagen que son transformados en un conjunto de vectores invariantes. Sin embargo esto sólo funciona en imágenes texturizadas, quedando fuera de su alcance imágenes de texturas muy suaves, como dibujos hechos a mano.

W. G. Lin y Sh. Wang en 1994 publican [20], donde presentan una forma discreta de calcular momentos, intentando corregir algunos errores, que no presenta la forma continua. Posteriormente Yang y Albregtsen en [1] y [2], encuentran una técnica para el cálculo de momentos en forma discreta. Utilizan el teorema de Green, obteniendo una versión eficaz, rápida y precisa.

En el 2002 Tomas Suk y Jan Flusser publican [8], donde declaran que: encontrar un sistema de descriptores invariantes es un paso dominante a reconocer objetos, sin importar

las deformaciones particulares. En 2004 François Chaumette presenta [5] donde propone utilizar los momentos de imágenes en control Visual Servo (sistema de visión para controlar el movimiento de un robot), ya que los momentos son una representación genérica de cualquier objeto.

En particular los momentos de Zernike ofrecen invariancia en traslación, escala y rotación [21], ofreciendo una alternativa diferente de SIFT que también ofrece invariancia ante las mismas transformaciones. El estudio comparativo realizado en [13] entre descriptores globales como: SIFT, momentos de Hu, momentos de Zernike y momentos de Fourier-Mellin. Concluye que SIFT es mejor herramienta para reconocer objetos solo si la base de datos no contiene pocas imágenes. Sin embargo cuando las imágenes tienen ruido los momentos de Zernike son mejores.

Los trabajos posteriores al cálculo de momentos se han centrado en mejorar la eficiencia de los momentos, ya sea que su cálculo sea más rápido y/o que se obtenga mayor cantidad de información con menor número de coeficientes.

En esta tesis se utiliza los resultados obtenidos por Mohammed Al-Rawi en [16], publicado en 2008, para el cálculo de momentos. Y para comparar la similitud entre los momentos de dos imágenes se utiliza [3], publicado en el 2009 por Jerome Revaud, Guillaume Lavoue, y Atila Baskurt.

1.2. Objetivo

Verificar la robustez de los momentos de Zernike y su invariancia ante las transformaciones de traslación, escala y rotación.

1.3. Justificación

Debido el avance tecnológico tanto en el *hardware* como en el *software*, en la era actual de las tecnologías de la información y las telecomunicaciones, es posible llevar a cabo procedimientos con la ayuda de las computadoras personales que en décadas anteriores no era posible, o que requerían de diversos y numerosos recursos. Debido a esto, gran parte de la teoría desarrollada años atrás en diferentes áreas del conocimiento hoy puede ponerse en la práctica. El área de procesamiento de señales no es la excepción, históricamente los momentos geométricos se ubican dentro de las primeras propuestas como descriptores para el reconocimiento de patrones. Con el paso del tiempo se han desarrollado diversos algoritmos con el fin de hacerlos más eficientes, haciendo de estos descriptores una herramienta muy eficiente y moderna, capaz de emplearse en tiempo real.

Dado lo anterior se decide utilizar a los momentos de Zernike como una herramienta prometedora para la extracción de descriptores de una imagen. Una vez que se eligieron a los momentos de Zernike como descriptores se observó que no son invariantes al cien por ciento, lo que ocasionó dificultades al momento de hacer una clasificación. Debido

1. Introducción

a esto el objetivo del presente trabajo fue modificado, enfocándose finalmente en hacer pruebas de porcentaje de invariancia de los momentos de Zernike en imágenes expuestas a transformaciones de traslación, escala y rotación.

1.4. Contenido

En el capítulo 2 se presentan generalidades sobre el reconocimiento de patrones haciendo mención de varios procedimientos para reconocer un objeto dentro de una imagen, tomando en cuenta los procedimientos más conocidos y eficientes que existen hasta el momento en la literatura científica. El capítulo 3 continúa con la descripción matemática de los momentos geométricos y de los momentos geométricos de Zernike. También se mencionan algunas de las aplicaciones de los momentos geométricos, de los momentos geométricos de Zernike, así como de la descripción de los principales artículos y los algoritmos que se proponen para el cálculo de los momentos en tiempo real (en el primer artículo), y otro artículo que describe un algoritmo encargado de hacer la comparación entre los momentos. En el capítulo 4 se encuentran las pruebas que se realizaron para la validación de los algoritmos del capítulo 3. Finalmente se tienen las conclusiones que se hacen en referencia a los resultados plasmados en el capítulo 4, y los resultados de los artículos del capítulo 3.

2. Generalidades sobre el reconocimiento de patrones

El objetivo de este capítulo es el de situar a los momentos de Zernike (MZ) dentro del reconocimiento de patrones. Se inicia con la descripción del concepto de patrón y su relación con el área de reconocimiento de patrones en general. Posteriormente se describe la naturaleza de las características de los objetos que se van a manejar y su influencia en el tipo de enfoque con el que se utiliza el reconocimiento de patrones. Se prosigue con una división en la clasificación entre supervisada y no supervisada para exponer algunos clasificadores y tipos de agrupación. En este punto se hace también énfasis en el problema de selección de variables con el fin de subrayar la importancia del papel que juegan los descriptores dentro del reconocimiento de patrones. Finalmente se hace un sumario de los descriptores más populares dentro de los cuáles se encuentran los momentos MZ motivo de esta tesis.

Considerando las imágenes digitales como nuestro soporte de análisis, un patrón es un conjunto de características de una imagen, y un conjunto de patrones similares forman una clase, Figura 2.0.1. En el procesamiento de imágenes el objetivo del reconocimiento de patrones es clasificar patrones, es decir, asignar cada patrón a la clase a la que pertenece. Por su parte las características pueden ser de carácter cualitativo o cuantitativo. Las primeras son cualidades que comúnmente no se miden numéricamente, más bien se categorizan. Pueden ser dicotómicas, que toman uno de dos valores posibles, o politómicas, cuando los posibles valores son tres o más. Algunas características de este tipo pueden ser la forma de los objetos y la regularidad de una imagen. Las características cuantitativas son las que podemos medir numéricamente como el área, la intensidad lumínica, etc. En consecuencia la naturaleza de las características aunada al tipo de aplicación, van a definir el enfoque con el que se manejen los patrones, Figura 2.0.2.

2. Generalidades sobre el reconocimiento de patrones

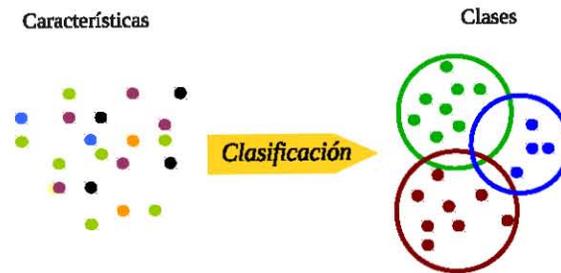


Figura 2.0.1.: A partir de las características de las imágenes se inicia la clasificación.

Debido a que las primeras investigaciones realizadas en reconocimiento de patrones fueron derivadas de resultados obtenidos en estadística [7], uno de los primeros enfoques del reconocimiento de patrones es a su vez estadístico, conocido como teoría de la decisión, basado en las características estadísticas de los patrones: media, varianza, etc.

El enfoque sintáctico o estructural, busca relaciones entre las características de los objetos. El caso sintáctico se apoya en la estructura matemática de los lenguajes formales, donde las características de los objetos son elementos de un alfabeto. Las combinaciones resultantes de dichos elementos forman palabras que pueden pertenecer o no a un lenguaje dado. El caso estructural ve a las características de los objetos como puntos que forman parte de estructuras geométricas. Teniendo nuevamente elementos que son considerados como palabras que pueden pertenecer o no a un lenguaje.

El enfoque lógico combinatorio [17] se basa en el cometido de modelar los problemas de forma real, tomando en consideración todas las variables (cuantitativas y cualitativas) que interfieren en el problema. Para posteriormente modelarlas matemáticamente y clasificarlas con una metodología sistemática.

Finalmente las redes neuronales [4] también se utilizan para clasificación explotando su similitud con el funcionamiento del cerebro humano, “aprendiendo” a colocar patrones de prueba en su correspondiente grupo o clase. Durante el proceso de aprendizaje las regiones de decisión van tomando formas no lineales y complejas, que se ajustan lo más posible al problema.

Amén de los diferentes enfoques, la teoría del reconocimiento de patrones se puede dividir en dos grandes rubros, el supervisado y el no supervisado. Aunque sus nombres nos permitan inferir correctamente como trabajan, es necesario subrayar algunos detalles de cada uno.

El reconocimiento supervisado se utiliza cuando se tiene información a priori al proceso de reconocimiento. También es conocido como clasificación con aprendizaje, debido a que ciertas clases se encuentran ya definidas. Inclusive se pueden encontrar algunos objetos previamente clasificados.

En el caso del reconocimiento no supervisado, no se tiene información a priori sobre las clases, los patrones no se encuentran etiquetados. Sin embargo, en algunos casos es necesario definir el número de clases.

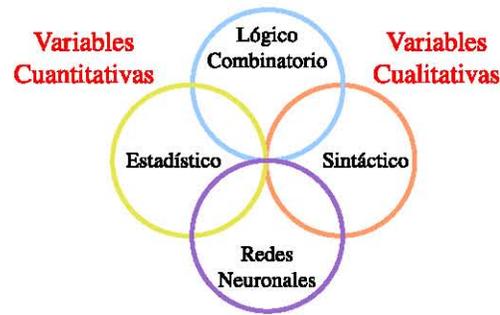


Figura 2.0.2.: La naturaleza de las características redundante en el tipo de clasificación.

2.1. Reconocimiento supervisado

Los clasificadores supervisados pueden representarse a través de un modelo genérico, en donde los clasificadores son una función discriminante, la encargada de dividir el espacio de representación en regiones de decisión, Figura 2.1.1. El tipo de función discriminante que emplea cada clasificador, define al clasificador en cuestión.

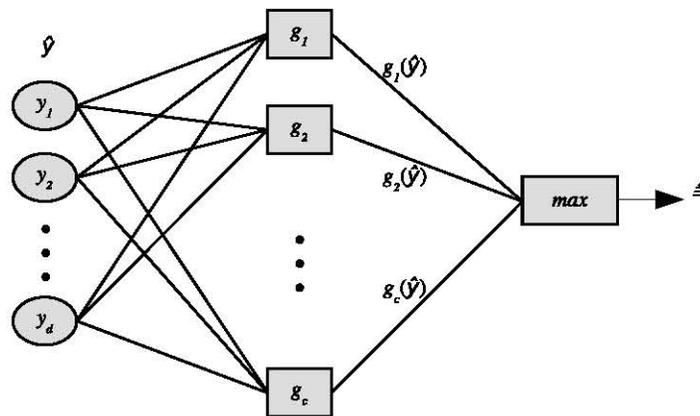


Figura 2.1.1.: Discriminantes

2.1.1. Clasificadores Bayesianos

Estos clasificadores se basan en la teoría de la probabilidad y la estadística, utilizan análisis de varianzas, covarianzas, dispersión y distribución, principalmente. Resuelven la pertenencia de un elemento a cierta clase, calculando la mayor probabilidad de que un elemento pertenezca a una clase.

Se representan por medio de una red bayesiana, que es un grafo acíclico dirigido en el que cada nodo representa una variable y cada conexión una dependencia probabilística,

2. Generalidades sobre el reconocimiento de patrones

Figura 2.1.2.

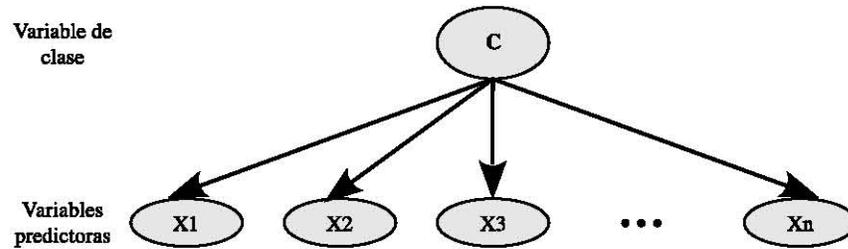


Figura 2.1.2.: Estructura gráfica de un modelo Bayesiano

2.1.2. Clasificadores lineales

Las funciones discriminantes que rigen este tipo de clasificadores son combinaciones lineales de las componentes del vector de características.

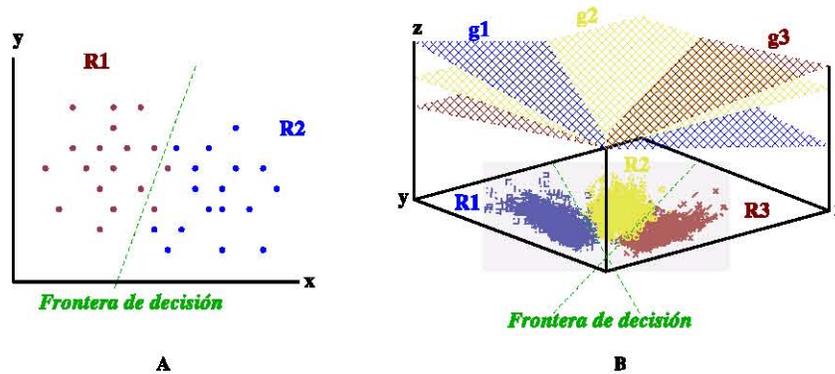


Figura 2.1.3.: División lineal en un espacio de características de dos y tres dimensiones.

Normalmente el espacio de características tiene un número elevado de dimensiones (del orden de decenas), se considera que es N-dimensional. Las regiones divididas por las fronteras de decisión se conocen como hiperplanos, por lo tanto existen tantos hiperplanos como clases de objetos. En el caso de $N=1$, el hiperplano es una recta, para $N=2$ es un plano, etc., Figura 2.1.3.

Matemáticamente la función de discriminantes lineal tiene la siguiente forma:

$$g(x) = \sum_{i=1}^d w_i x_i \quad (2.1.1)$$

donde:

x : vector de características
 x_i : componentes del vector de características
 w : vector de peso
 d : número de parámetros (proporcional con la dimensión)

2.1.3. Clasificadores no lineales

En este caso las fronteras de decisión están expresadas con funciones cuadráticas, las cuales pueden ser cónicas: parábolas, círculos, elipses, hipérbolas, Figura 2.1.4.

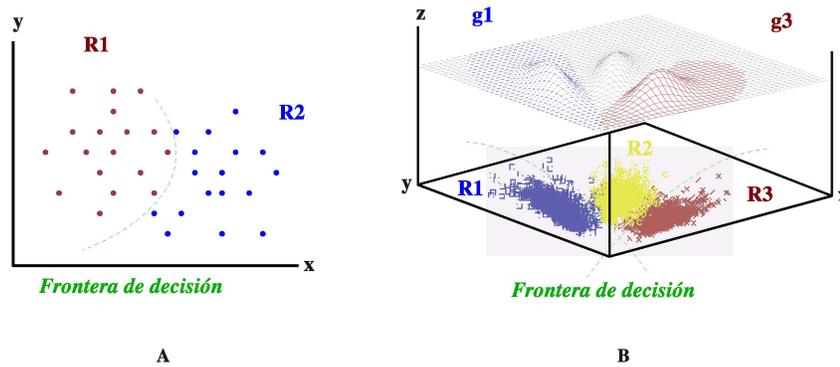


Figura 2.1.4.: División lineal en un espacio de características de dos y tres dimensiones.

En este caso la función discriminante se muestra en forma matricial:

$$g(x) = x^T A x + x^T B + C, C = w_{n+1} \quad (2.1.2)$$

donde:

A : matriz de pesos de $d \times d$

B : vector de pesos de $1 \times d$

2.2. Reconocimiento no supervisado

El reconocimiento no supervisado es también conocido como clasificación sin aprendizaje. En donde el objetivo es identificar las formas que describen los grupos de patrones, de los cuales no se tiene información a priori. Usualmente esto se logra construyendo diferentes clases en base a características afines, es decir agrupar patrones similares entre sí, y a su vez diferentes a los patrones de otros agrupamientos.

Los patrones procedentes de imágenes están caracterizados por vectores de características, es decir cada uno de estos vectores es equivalente a un patrón. Estos vectores se encuentran en espacios n-dimensionales, donde el número de dimensiones depende del número de

2. Generalidades sobre el reconocimiento de patrones

características. La población de estos patrones se distribuye en diferentes formas, a las que se les puede clasificar por sus propiedades cualitativas como: formas alargadas, circulares huecas, circulares densas, etc. La forma de cada nube (grupo de patrones) depende no solo de los patrones, sino también del algoritmo de agrupamiento seleccionado, y de la medida de similitud utilizada para compararlos.

Un algoritmo de agrupamiento es un procedimiento lógico de agrupación de vectores conforme a un criterio de cercanía. Esta cercanía se define en términos de una determinada función de distancia, Figura 2.2.1. Los vectores que no satisfagan con esta función de distancia, serán parte de otro grupo, o excepcionalmente no pertenecen a grupo alguno. En la teoría de reconocimiento de patrones el algoritmo de agrupamiento tiene otros nombres, como función de decisión, o clasificador.

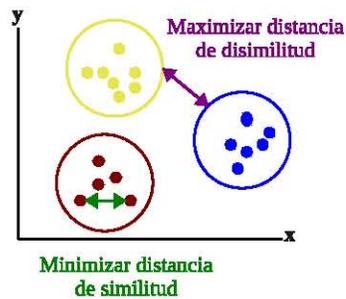


Figura 2.2.1.: Distancias de agrupamiento.

El criterio de cercanía se obtiene con una medida de similitud, sin embargo también es necesario tener un criterio de lejanía, el cual se calcula con una medida de disimilitud. Para cuantificar estas distancias de disimilitud se pueden usar varias métricas como las distancias de Minkowski, Ecuación 2.2.1, también conocida como métrica L_p .

$$d_q(x, y) = \left(\sum_{i=1}^l w_i |x_i - y_i|^q \right)^{\frac{1}{q}} \quad (2.2.1)$$

donde

w_i : coeficiente de peso

q : tipo de norma

De esta distancia se derivan la distancia euclidiana ($q=2$), la de Manhattan ($q=1$) y la Chebychev ($q=\infty$).

En el caso de valores reales las medidas de similitud que se utilizan comúnmente son el producto punto, los coeficientes de correlación de Pearson, y la distancia de Tanimoto. Sin embargo la métrica se elige de acuerdo con el tipo de datos que se estén manejando y de

la semántica que se les asocie. Esto sugiere que en la similitud entre objetos subyace una noción subjetiva.

A continuación se describen brevemente algunos de los algoritmos de agrupamiento.

2.2.1. Agrupamiento jerárquico

Este agrupamiento no supervisado tiene la estructura progresiva de un árbol, llamado dendrograma. Admite tantas variantes como funciones de similitud puedan definirse. La agrupación se realiza mediante un proceso de fases de agrupación, o desagrupación sucesivas. En cada fase de agrupación, o desagrupación, usa la distancia mínima, la distancia máxima y la distancia media entre todos los pares posibles.

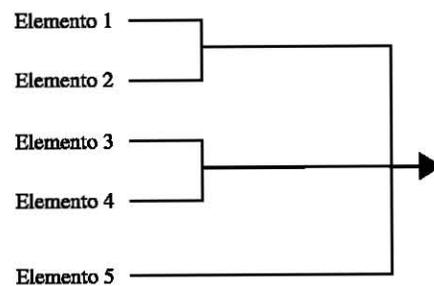


Figura 2.2.2.: Método jerárquico aglomerativo

En la Figura 2.2.2 se muestra el método jerárquico aglomerativo, el inverso del método aglomerativo es el método jerárquico disociativo, su representación esquemática es el espejo del esquema jerárquico aglomerativo.

2.2.2. Agrupamiento secuencial

Este método comienza con un vector el cual representa un grupo, cuando un nuevo vector es evaluado y no se encuentra lo suficientemente alejado de ese vector inicial, en términos de una medida preestablecida, se integra a ese grupo. Entonces este grupo cambia sus características debido a este nuevo vector asociado. Cuando un vector nuevo está lo suficientemente alejado de los grupos existentes se crea un nuevo grupo. La lejanía también depende de una distancia previamente definida. En algunos algoritmos se requiere como criterio de paro de una cota superior del número de grupos a clasificar.

2.2.3. Agrupamiento basado en optimización de la función de costo

La frontera de los grupos de este tipo de agrupamiento se definen con una función de costo. Esta función es paramétrica, ya sea lineal, cuadrática, etc., y está regida por una

2. Generalidades sobre el reconocimiento de patrones

variable desconocida. En estas condiciones, el objetivo de estos algoritmos es encontrar dicha variable, haciendo uso del cálculo diferencial, generando grupos sucesivos y tomando en cuenta el número máximo de grupos, en el caso de que éstos estén preestablecidos.

2.3. Problemas de clasificación

La primera interrogante para comenzar el proceso de clasificación consiste en saber cuáles son los patrones más adecuados para representar a determinados objetos, este problema se conoce como de selección de variables (descriptores).

Una vez que se ha superado el problema anterior, inmediatamente surge una nueva interrogante, ¿cuál es el número de patrones a utilizar? De antemano se sabe que comúnmente entre más patrones (variables) se tengan es mejor, pero el número de ellos depende de la aplicación y de la complejidad que se decida manejar. Ya que entre más variables se tengan el tiempo de procesado es mayor, así que para aplicaciones de tiempo real por lo general se tiene que encontrar un compromiso sobre el número de variables.

Ya que han sido seleccionados el tipo y el número de variables, se necesita diseñar un clasificador, buscando que la clasificación de cada elemento sea lo más clara posible, para evitar que las variables pertenezcan a más de un grupo, y con esto minimizar errores. Es decir, se debe de encontrar un algoritmo que genere grupos definidos, de tal forma que sea difícil que se traslapen entre ellos.

Otros problemas secundarios (porque dependen del proceso de captura) pero importantes, son la oclusión de los objetos a reconocer. Ocasionada por las variadas y diferentes proyecciones en dos dimensiones (2D), captadas por las cámaras del mundo tridimensional (3D).

2.4. Descriptores

Para poder clasificar una imagen es necesario tener un grupo de características que las defina, estas características también se conocen como descriptores, en el proceso de clasificación son vectores de características. Los descriptores son el primer paso para encontrar el vínculo entre los píxeles contenidos en una imagen digital y lo que representan (sentido semántico). Entre los más comúnmente utilizados en análisis y procesamiento de imágenes para reconocimiento, están los descriptores de contorno y de regiones, por su significado respecto al sistema de visión.

2.4.1. Descripción por medio de contornos

Este tipo de descriptores se emplean cuando basta con conocer información a cerca de la forma de los objetos. El proceso general de extracción de características consiste primero en reducir el efecto del ruido, el siguiente paso es la detección de regiones. Posteriormente se determinan los que se consideran contornos de interés, y el último paso es refinar el resultado,

eliminando puntos aislados y conectando los pertenecientes a un contorno deseado.

Una vez extraídos los contornos de los objetos hay que analizar la forma geométrica de los mismos utilizando para su representación una estructura de datos compacta, es decir, un esquema de representación. Dentro de los esquemas de representación externa, que son los que usan el contorno de los objetos y sus rasgos característicos, se encuentran los códigos de cadena, los descriptores de Fourier, las firmas, las aproximaciones poligonales, entre otros, Figura 2.4.1.

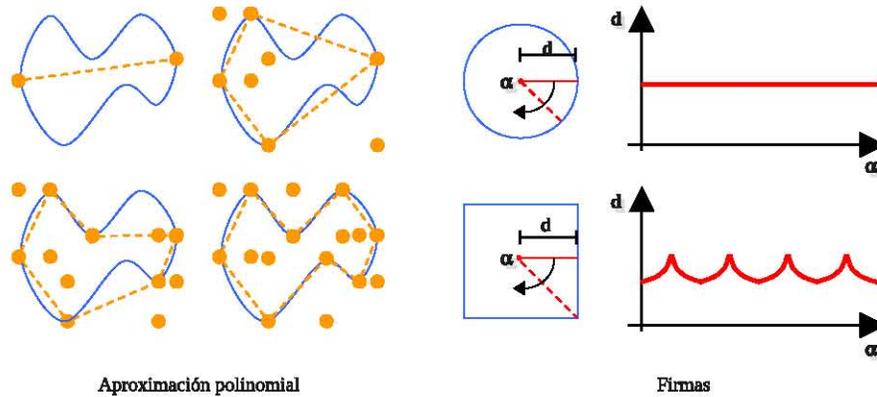


Figura 2.4.1.: Ejemplificación de algunos descriptores de contorno

Algunos tipos de estos descriptores se describen brevemente a continuación:

Los *códigos de cadena* describen un objeto mediante una sucesión de números que representan segmentos de líneas de tamaño unitario (con conectividad 4 u 8), con una determinada orientación. Este descriptor es comúnmente utilizado para contornos.

Los *descriptores de Fourier* se usan cuando se considera un contorno como una función periódica, con el fin de aplicar a este contorno (función) la transformada de Fourier. Los coeficientes de la transformada se usan como descriptores de la forma del contorno. Las bajas frecuencias corresponden a detalles burdos y las altas frecuencias a los detalles finos. Este otro descriptor es comúnmente utilizado para contornos y texturas.

La *firma* es una representación de un contorno mediante una función real unidimensional, más sencilla que la función bidimensional que define el contorno. Una de las maneras más simples de definirlas es a través de la distancia como una función del ángulo, partiendo desde un punto interior (por ejemplo el centroide del contorno) hasta alcanzar cada uno de los puntos del contorno.

La *representación de polilíneas* o *aproximaciones poligonales* consiste en una descripción de contornos en base a segmentos de líneas, donde cada segmento se especifica mediante el punto inicial y final (coordenadas). La concatenación de estos puntos, con el mismo punto inicial y final, describe un contorno. Estos dos últimos descriptores son también utilizados

2. Generalidades sobre el reconocimiento de patrones

para la descripción de contornos.

El *área* se calcula como el número de píxeles que pertenecen a una región. Para calcular el área real de objeto se debe tener en cuenta la resolución del píxel en el mundo real. Este descriptor es comúnmente usado para el método de regiones.

2.4.2. Descripción por medio de regiones

Se utilizan cuando el interés principal es describir las propiedades de la región de los objetos detectados, como la intensidad, el color, la textura, etc.

Existen varios descriptores que pueden ser calculados de forma muy sencilla sobre las regiones, conocidos como *descriptores simples*. Algunos de ellos son el área, el número de Euler y la densidad.

También están los *descriptores geométricos*, estudian la forma geométrica de los contornos de las regiones (objetos). En esta categoría se encuentra el alargamiento, la espesura, la dirección y la compacidad, Figura 2.4.2.

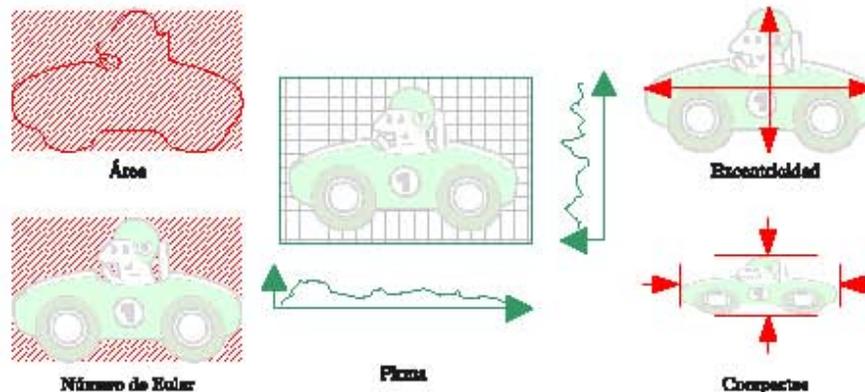


Figura 2.4.2.: Ejemplificación de algunos descriptores de región

Es importante mencionar que algunos de estos descriptores presentan invariancia a las transformaciones lineales (alargamiento y rectangularidad), e invariancia a la rotación (dirección).

Finalmente están los *descriptores con base matemática*, en donde la interpretación de las funciones imagen se suponen como densidades de probabilidad de variables aleatorias bidimensionales. Las propiedades de estas variables aleatorias se pueden describir usando características estadísticas como son los momentos de diferente orden.

Un momento de orden $(p+q)$ es dependiente de la escala, de las translaciones, de las rotaciones y de cualquier transformación realizada sobre los niveles de gris. Sin embargo la invariancia a la escala y a las rotaciones se puede alcanzar con ayuda de las expresiones de los momentos y los momentos centrales. De forma un poco más elaborada se puede

conseguir la invariancia a la rotación y a la traslación.

Del conjunto de rasgos descriptivos, existen algunos que representan las características esenciales del objeto, mientras que otros representan características marginales. Los rasgos esenciales no pueden ser eliminados de la descripción de los objetos sin confundirlos. Mientras que los rasgos marginales, pueden ser ignorados en una descripción, sin que ello implique la confusión de los objetos.

Son precisamente los momentos (aunque no estadísticos, sino geométricos), el tipo de descriptores que se utilizan en este trabajo de tesis. Esta elección se debe a que ofrecen cierta invariancia y por lo tanto mayor estabilidad al trabajar con ellos. Además de que permiten obtener mucha información útil del objeto en cuestión.

3. Teoría de los Momentos Geométricos

Este capítulo está dedicado a exponer los principales conceptos de la teoría detrás de los momentos geométricos. Para desarrollar el tema de manera concisa y clara, se inicia utilizando los momentos estadísticos como analogías de los momentos geométricos, dada la popularidad de los primeros. Matemáticamente los momentos se desarrollan bajo la operación del producto punto entre una función y una base, esta última es la que da la posibilidad de clasificar los momentos en: momentos de base ortogonal y momentos de base no ortogonal. Dentro de los momentos geométricos con bases ortogonales se encuentran los momentos de Zernike (MZ), que constituyen la esencia de esta tesis. El capítulo concluye con una breve descripción de los artículos utilizados para implementar un algoritmo eficaz de cálculo de momentos.

Los momentos geométricos de una imagen digital son descriptores de regiones con base matemática que proporcionan información sobre diferentes estructuras geométricas. Gracias a la información que proporciona el reconocimiento de patrones es posible tener aplicaciones como: reconocimiento de formas, clasificación de objetos, reconstrucción, acoplamiento de escenas, reconocimiento de expresiones faciales, marcas de agua, análisis de texturas, determinación de orientación, recuperación de imágenes basada en contenido, registro de imágenes, estimación de posición, etc.

Una manera de estudiar los momentos geométricos es a través de sus similitudes con los momentos estadísticos, fundamentados en la noción de *momento* que se define en mecánica [12].

En estadística los momentos se utilizan como descriptores de propiedades de una variable aleatoria, y en reconocimiento de patrones se usan para describir una imagen.

En estadística, los datos recolectados son resumidos para facilitar su interpretación de manera gráfica o numérica. Las medidas estadísticas que sirven para caracterizar estos datos se conocen con el nombre de *momentos* de diferente orden. El momento de primer orden es la media, el de segundo orden es la varianza. Y los momentos subsecuentes, conocidos como momentos de orden superior son menos utilizados, debido a la dificultad de su interpretación y a la carga computacional que representa su cálculo. Estos momentos estadísticos pueden servir para describir un buen número de características de un conjunto de datos.

Ahora bien lo que en estadística se considera un conjunto de datos, en el área de reconocimiento de patrones corresponde a una imagen, o a una región de una imagen. Así como los momentos estadísticos caracterizan conjuntos, los momentos geométricos se utilizan para caracterizar estructuras en imágenes. Sin embargo, como es de esperarse y se verá más adelante, la interpretación de los momentos en cada caso es diferente.

3. Teoría de los Momentos Geométricos

Desde un punto de vista matemático los momentos geométricos son un producto punto (Ecuación 3.0.1), que representa una medida de similitud entre vectores, puesto que son la proyección de un vector sobre otro. Permitiendo medir cuánto se parece una señal (modelada como un vector), a una función dada (considerada como una base de un espacio vectorial). Es decir, se considera el producto punto:

$$\langle r, s \rangle = \int_{-\infty}^{\infty} w^2(x)r(x)s^*(x)dx \quad (3.0.1)$$

donde:

s^* : complejo

$r(x), s(x)$: vectores complejos

$w(x)$: función de peso, número real¹

Considerando que las imágenes se representan comúnmente en un plano es conveniente generalizar el producto punto en dos dimensiones (Ecuación 3.0.1).

$$\langle f, V \rangle = w^2 \int_0^{\infty} \int_0^{\infty} f(x, y)V^*(x, y)dxdy \quad (3.0.2)$$

donde:

$V^*(x, y)$: polinomio complejo de 2 dimensiones (imagen)

$f(x, y)$: función de densidad de la imagen

w : función de peso

Prácticamente la Ecuación 3.0.2 es la ecuación general con la que se obtienen los momentos geométricos, sin embargo se reescribe de la siguiente forma para lograr una mejor comprensión de los elementos que la componen.

$$M_{pq} = \lambda_{pq} \int_0^{\infty} \int_0^{\infty} f(x, y)V_{pq}^*(x, y)dxdy \quad (3.0.3)$$

donde

λ_{pq} : constante que depende de la función de momentos utilizada

$V_{pq}^*(x, y)$: polinomio complejo de momentos en 2 dimensiones

$f(x, y)$: función de densidad de la imagen

Se observa que los momentos son una sumatoria infinitesimal de la proyección de una imagen sobre una base polinomial. La elección de esta base es muy importante debido a que las cualidades que la definen formarán parte de las características de los momentos, y es de acuerdo a la naturaleza de su base como se pueden clasificar. En este trabajo los dividiremos en dos categorías: momentos con base no ortogonal y momentos con base ortogonal.

¹La función de peso asegura la convergencia del producto interno.

3.1. Momentos no ortogonales

Cronológicamente, los primeros momentos usados para el tratamiento de imágenes fueron no ortogonales y fueron propuestos por Hu en 1962 [10]. En el plano cartesiano se encuentran definidos de la siguiente forma.

$$M_{pq} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x^p y^q f(x, y) dx dy \quad (3.1.1)$$

Y en su forma discreta se expresan mediante las siguientes sumatorias.

$$m_{pq} = \sum_{x=0}^{\infty} \sum_{y=0}^{\infty} x^p y^q f(x, y) \quad (3.1.2)$$

En el reconocimiento de patrones se buscan básicamente tres propiedades importantes en los descriptores, las cuales son: la invariancia a la traslación, la invariancia a la escala y la invariancia a la rotación. La invariancia a la traslación se logra con la ayuda de los momentos centrales. Los cuales se extraen de una imagen centrada con respecto a su centroide (centro de masa de la imagen).

$$MC_{pq} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x - \bar{x})^p (y - \bar{y})^q f(x, y) dx dy \quad (3.1.3)$$

$$\mu_{pq} = \sum_{x=0}^{\infty} \sum_{y=0}^{\infty} (x - \bar{x})^p (y - \bar{y})^q f(x, y) \quad (3.1.4)$$

El momento simple de orden cero, m_{00} , en imágenes binarias, representa el área de la figura, y en imágenes en escala de grises, la superficie. Es decir, es la suma de los valores de todos los píxeles.

Los momentos simples de orden uno, m_{10}, m_{01} , se emplean para ubicar el centro de masa de una imagen.

$$\bar{x} = \frac{m_{10}}{m_{00}}, \bar{y} = \frac{m_{01}}{m_{00}} \quad (3.1.5)$$

La invariancia en escala se logra a través de los momentos centrales, y consecutivamente a partir de ellos se obtienen los momentos centrales normalizados, mejor conocidos como momentos normalizados. Esta normalización se hace en función del área, empleando para ello el momento central de orden cero. Es importante señalar que la normalización puede realizarse con cualquier momento, sin embargo comúnmente se utilizan los momentos de bajo orden porque son más estables respecto al ruido y son más fáciles de calcular.

Entonces si un momento invariante en traslación es dividido entre uno de sus momentos elevado a un factor de escala, se obtiene un momento invariante tanto a la traslación como

3. Teoría de los Momentos Geométricos

a la escala, siempre y cuando se respete la siguiente desigualdad $(p + q) > 1$ [10]. Sin embargo por definición para una imagen centrada, tanto los momentos centrales de orden uno, μ_{10}, μ_{01} , como los momentos normalizados de orden uno, η_{10}, η_{01} , tienen un valor nulo, o cero. Pero si la imagen no está centrada con respecto al su centroide este valor es diferente de cero. En consecuencia, para escalar una imagen y evitar una indeterminación, no se utilizan los momentos centrales de orden uno, al igual que otros momentos que tengan valor cero. En la Ecuación 3.1.6 se emplea el momento de orden cero para obtener un factor de escala, el cual también se utilizará a lo largo de este trabajo.

$$f'(x, y) = f\left(\frac{x}{\lambda}, \frac{y}{\lambda}\right), \lambda = \frac{m_{00}}{m'_{00}} \quad (3.1.6)$$

donde

λ : factor de escalamiento

m_{00} : momento de orden cero

el apóstrofe, en el caso de f' denota la señal escalada y en el caso de m'_{00} respectivamente, el momento cero de f' .

De esta manera, recalculando la Ecuación 3.1.3 se pueden obtener los momentos centrales normalizados, dados en la Ecuación 3.1.7.

$$MCN_{pq} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x^p y^q f\left(\frac{x}{\lambda}, \frac{y}{\lambda}\right) dx dy = \frac{MC_{pq}}{\lambda^{2+p+q}} \quad (3.1.7)$$

Correspondiendo en el dominio discreto en la Ecuación 3.1.8.

$$\eta_{pq} = \frac{\mu_{pq}}{\lambda^{\frac{p+q}{2}} + 1}, p + q = 2, 3, \dots \quad (3.1.8)$$

En mecánica los tres momentos centrales de orden dos, $\mu_{20}, \mu_{02}, \mu_{11}$, son los momentos de inercia que forman las componentes del tensor de inercia, o matriz de rotación. Un momento de inercia es una magnitud escalar que refleja la distribución de masas de un cuerpo, o un sistema de partículas en rotación, respecto al eje de giro. Y depende de la geometría del cuerpo y de la posición del eje de giro; pero no de las fuerzas que intervienen en el movimiento.

$$R = \begin{bmatrix} \mu_{02} & -\mu_{11} \\ \mu_{11} & \mu_{20} \end{bmatrix} \quad (3.1.9)$$

A partir de la excentricidad de la figura y de estas componentes se podrá obtener el ángulo de rotación de la figura alrededor de su centro de masa. La orientación, o ángulo de rotación de la figura se define como el ángulo entre el eje de las abscisas, y el eje alrededor del cual la figura puede rotar con mínima inercia.

En [10] Ming Kuei Hu, en la sección intitulada “el método de los ejes principales” determina que la orientación de una imagen se define con la Ecuación 3.1.10, sin embargo esto es

válido únicamente para imágenes que cumplen ciertas restricciones. Para obtener el ángulo de rotación de una imagen específica a partir de la Ecuación 3.1.9 es necesario analizar detenidamente la relación que hay entre su geometría, específicamente su excentricidad, y los momentos que la describen.

$$\theta = \frac{1}{2} \arctan \frac{2\mu_{11}}{\mu_{20} - \mu_{02}} \quad (3.1.10)$$

Finalmente, para obtener los momentos invariantes a la rotación, conocidos simplemente como momentos invariantes, se trabaja con los momentos normalizados y haciendo uso de la teoría de invariantes algebraicas, Hu [10] obtiene un conjunto de siete momentos invariantes a la traslación, a la escala y a la rotación, en el plano. Ver apéndice A.

3.2. Momentos ortogonales

Los momentos obtenidos a través de bases ortogonales ofrecen la ventaja de no presentar redundancia. Esto implica que la información que se extrae de una imagen es mayor, utilizando un número de momentos igual o menor, que la que se pudiese extraer utilizando momentos con base no ortogonal.

La ausencia de redundancia se debe a que los polinomios ortogonales son linealmente independientes. En consecuencia al realizar un producto punto entre una imagen (función) y un polinomio se obtiene cierta información de la imagen, cuando se realiza un nuevo producto punto con la misma imagen pero con un polinomio ortogonal al primero, se obtiene información adicional de la imagen. Si sucesivamente se hace un producto punto entre la misma imagen y otro polinomio ahora ortogonal a los dos anteriores, una vez más se consigue información adicional de la imagen en cuestión. En un espacio n dimensional, donde n tiende a infinito, este proceso puede continuar infinitamente, o hasta que tengamos toda la información necesaria del imagen. Las bajas frecuencias se encuentran representadas en su mayoría en los momentos de orden bajo. Lo anterior puede corroborarse con ayuda de la Figura 3.2.1 donde se proyectan sobre diferentes planos ortogonales las diferentes caras de un poliedro. En la realidad los objetos existentes son muy complejos y se utilizan diferentes tipos de bases para proyectar características específicas de estos objetos.

La desventaja que presenta este tipo de bases es que son polinomios más complejos que los no ortogonales y por ende más complejos de implementar. Normalmente, por simplicidad están expresados en coordenadas polares. Algunos ejemplos de polinomios con características importantes son los de Legendre, Tchebyshev y Zernike. En este trabajo de tesis y considerando su aportación en el análisis de imágenes utilizaremos los polinomios de Zernike.

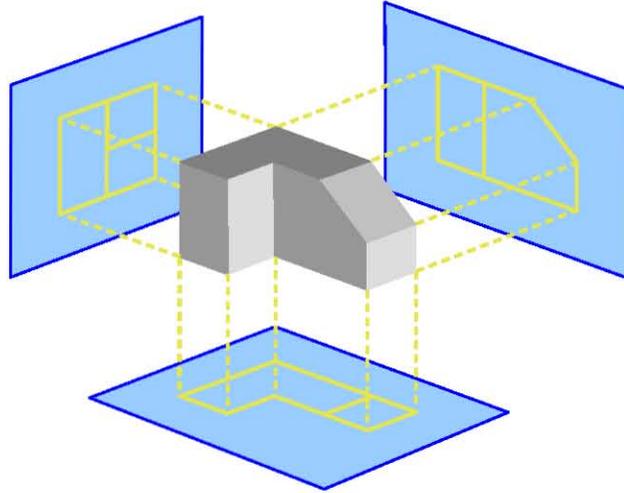


Figura 3.2.1.: Proyección de un objeto, sobre una base canónica de tres dimensiones.

3.3. Momentos de Zernike

Estos momentos utilizan como base los polinomios del mismo nombre, que fueron desarrollados en 1934 por el holandés Frits Zernike a través del proceso de ortogonalización de Gram Schmidt y se definen dentro del círculo unitario del plano complejo [22]. En 1979 Michael Reed Teague decide usarlos como base para los momentos, es decir reemplaza el núcleo no ortogonal de Hu por otro ortogonal (polinomios de Zernike) y los nombra momentos de Zernike [21].

Las principales características de los momentos de Zernike son: robustez al ruido, invariancia a la rotación y fácil reconstrucción. En los años recientes la reconstrucción de imágenes digitales es uno de los usos más populares. El procedimiento de reconstrucción se basa en la hipótesis de que una función definida en intervalos puede descomponerse en un conjunto de funciones ortogonales. Esta descomposición trae como resultado una serie infinita de coeficientes que representan fielmente a la función original. Adicionalmente una función discreta, como una imagen, puede reconstruirse con un número finito de coeficientes. Usando los polinomios de Zernike para hacer este tipo de descomposición se tiene una mejor representación que con el mismo número de coeficientes de una descomposición con polinomios no ortogonales.

Considerando los antecedentes presentados, la ecuación para el cálculo de los momentos de Zernike puede resultar familiar, dado que es bastante similar a las ya utilizadas. Con la variante de que el núcleo y la constante de peso, son diferentes.

$$A_{nl} = \frac{n+1}{\pi} \int_0^{2\pi} \int_0^1 f(r, \theta) Z_{nl}^*(r, \theta) r dr d\theta \quad (3.3.1)$$

donde

$n + l$: grado del momento

El polinomio de Zernike está dado por:

$$Z_{nl}(r, \theta) = R_{nl}(r)e^{il\theta} \quad (3.3.2)$$

donde

n : entero positivo o cero

l : entero positivo o negativo

r : distancia del vector que va desde el origen hasta el punto (r, θ)

θ : ángulo entre r y el eje real, en sentido anti-horario

i : $\sqrt{-1}$

$$R_{nl}(r) = \sum_{s=0}^{\frac{n-|l|}{2}} (-1)^s \frac{(n-s)!r^{n-2s}}{s! \left(\frac{n+|l|}{2} - s\right)! \left(\frac{n-|l|}{2} - s\right)!} \quad (3.3.3)$$

Debido a la forma de la Ecuación 3.3.3 se tienen las siguientes restricciones:

$$|l| \leq n$$

$n - |l|$ debe ser par.

Cuando se aplican operaciones simples a una imagen, como por ejemplo cambio de tamaño, traslación, rotación, o reflexión, la función de densidad de la imagen cambia, así como sus momentos correspondientes. Para evitar que los momentos cambien cuando la imagen sufre cambios es necesario normalizarlos.

Para mantener la invariancia a la traslación y a la escala se retoman los conceptos propuestos por Hu, es decir, se centra la imagen con respecto a su centroide y se escala.

Al expresar los momentos de Zernike como en la Ecuación 3.3.4 se observa que $|Z_{pq}e^{iq\alpha}| = |Z_{pq}|$ lo que implica que la magnitud de los momentos de Zernike es invariante a la rotación.

$$Z_{pq}^\alpha = Z_{pq}e^{iq\alpha} \quad (3.3.4)$$

donde

α : ángulo de rotación al rededor del centro de masa

En el Apéndice A y para hacer una relación, se encuentran los momentos invariantes de Zernike expresados en función de los momentos de Hu.

Para hacer eficiente la implementación computacional de los momentos de Zernike, se han desarrollado varios métodos. En este trabajo se emplean dos de ellos. Con el propósito de contar con los antecedentes, a continuación se hace una breve descripción de ambos métodos.

3. Teoría de los Momentos Geométricos

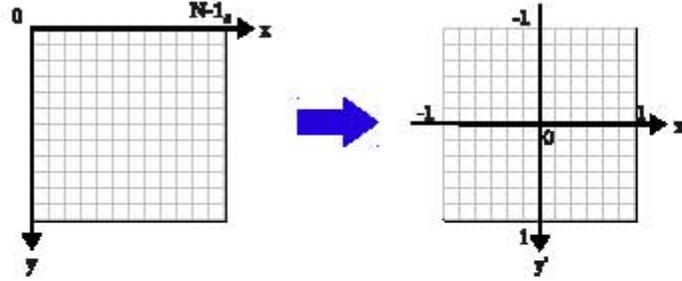


Figura 3.3.1.: Centrado de imagen.

3.3.1. Uso de un núcleo geométrico

En este método se hace uso de las propiedades de los polinomios de Zernike para hacer eficiente su implementación. La primera propiedad es que los polinomios de Zernike son polinomios denominados de núcleo simétrico, esto significa que se encuentran definidos en un intervalo de $[-1, 1]$ en vez de $[0, \infty]$. Como consecuencia los productos que resulten de este tipo de núcleo serán valores dentro del mismo intervalo y tendrán valores pequeños entre $[-1, 1]$, permitiendo hacer más rápidas las operaciones y ocupando menos memoria que cuando se manejan valores entre $[-\infty, \infty]$. Por esta razón es necesario normalizar una imagen con dimensiones de $M \times N$, reubicando el origen de la misma y limitando a uno el valor máximo de sus coordenadas, Figura 3.3.1.

El cambio de los índices de los ejes se hace utilizando la Ecuación 3.3.5 y la Ecuación 3.3.6.

$$x' = -1 + \frac{2(x + 0,5)}{N} \quad (3.3.5)$$

$$y' = -1 + \frac{2(y + 0,5)}{M} \quad (3.3.6)$$

donde

x' y y' : coordenadas transformadas

x y y : coordenadas originales

N y M : dimensiones de la imagen

$$g(x') = \left[\frac{2(x + 0,5)}{N} - 1 \right]^p \quad (3.3.7)$$

$$g(y') = \left[\frac{2(y + 0,5)}{M} - 1 \right]^q \quad (3.3.8)$$

Los momentos de Zernike se pueden calcular a través de la Ecuación 3.0.3, sin embargo si se utilizan tal como están definidos se pierde la propiedad de simetría. Buscando no

perder esta propiedad los monomios del núcleo de Hu son redefinidos en el intervalo $[-1, 1]$ expresándose ahora como en la Ecuación 3.3.7 y la Ecuación 3.3.8. Luego, en la Ecuación 3.3.1 se puede expresar de la siguiente forma:

$$Z_{pq} = \frac{p+1}{\pi} \sum_{k=q}^p B_{pqk} \sum_{m=0}^s \sum_{n=0}^q w^2 \binom{s}{m} \binom{q}{n} \quad (3.3.9)$$

$$\times \sum_{a=0}^{N-1} \sum_{b=0}^{M-1} f(x'_a, y'_b) H_{k-2m-n}(x'_a) H_{2m+n}(y'_b) \quad (3.3.10)$$

donde

$$B_{pqk} = \frac{(-1)^{\frac{p-k}{2}} (\frac{p+k}{2})!}{(\frac{p-k}{2})! (\frac{k+q}{2})! (\frac{k-q}{2})!}$$

$$w = \begin{cases} -i, & q > 0 \\ +i, & q \leq 0 \end{cases}$$

$$i = \sqrt{-1}$$

$$s = \frac{k-q}{2}$$

$$H_p(x'_a) = \int_{x'_a - \frac{\Delta x'_a}{2}}^{x'_a + \frac{\Delta x'_a}{2}} x'^p dx' = \left[\frac{x'^{p+1}}{p+1} \right]_{x'_a - \frac{\Delta x'_a}{2}}^{x'_a + \frac{\Delta x'_a}{2}}$$

$$H_q(y'_b) = \int_{y'_b - \frac{\Delta y'_b}{2}}^{y'_b + \frac{\Delta y'_b}{2}} y'^q dy' = \left[\frac{y'^{q+1}}{q+1} \right]_{y'_b - \frac{\Delta y'_b}{2}}^{y'_b + \frac{\Delta y'_b}{2}}$$

Esta expresión permite que $H_p(x'_a)$ y $H_q(y'_b)$ puedan ser precalculados, de esta manera se calculan una sola vez, y posteriormente se reutilizan sus valores almacenados tantas veces como sea necesario.

3.3.2. Comparación de momentos usando un ángulo de rotación

En este otro método, una vez que se obtienen los momentos de Zernike de una determinada imagen, se comparan con los respectivos momentos de Zernike de una imagen prototipo. Esta comparación y con el propósito de mantener la invariancia a la rotación, normalmente se enfoca a la magnitud de los momentos, es decir se desecha la información de fase. En nuestro trabajo como se verá en el capítulo siguiente, y respecto a este segundo método, se propone hacer uso tanto de la magnitud, como de la fase.

El objetivo que se desarrolla en la referencia propia a este método consiste en encontrar una forma de comparar los momentos por medio de una función en la que se relacionan dos imágenes a través del ángulo de desfase entre ellas. Sabiendo a priori que los momentos de una imagen cambian si esta se encuentra rotada, y utilizando la distancia euclidiana entre la imagen de referencia y la rotada, para obtener el ángulo que se busca.

3. Teoría de los Momentos Geométricos

$$d_{I,J}^2(\phi) = \sum_{x^2+y^2 \leq 1} \sum |I(x,y) - (J * \mathfrak{R})(x,y)|^2 \quad (3.3.11)$$

donde

$I(x,y)$: imagen de referencia

$(J * \mathfrak{R})(x,y)$: imagen rotada

ϕ : ángulo de desfase entre el par de imágenes

Después de desarrollar y simplificar la Ecuación 3.3.11, se llega a la Ecuación 3.3.12. Al ser evaluada, el valor más cercano a cero que se obtenga de esta expresión, indica que las dos imágenes son muy parecidas o iguales. Y el valor de ϕ con el que se ubicó el valor cercano a cero, señala el ángulo de rotación de una imagen con respecto a la otra.

$$d_{I,J}^2(\phi) = \sum_{(p,q) \in D} \sum \frac{\pi}{p+1} \left[|Z_{pq}^I|^2 + |Z_{pq}^J|^2 - 2 |Z_{pq}^I Z_{pq}^J| \cos(q\phi + [Z_{pq}^J] - [Z_{pq}^I]) \right] \quad (3.3.12)$$

donde

$|Z_{pq}^I|, |Z_{pq}^J|$: magnitud del momento pq de las imágenes I y J , respectivamente

$[Z_{pq}^J], [Z_{pq}^I]$: fase del momento pq de las imágenes I y J , respectivamente.

La ecuación en función de ϕ que relaciona las dos imágenes, (Ecuación 3.3.12), se reduce eliminando los términos que no dependen de ϕ , ya que son términos constantes. Con la ayuda de la Ecuación 3.3.13 se llega a la Ecuación 3.3.14, función que relaciona a un par de imágenes en función del desfase que existe entre ellas.

$$|C| \cos(q\phi + [C]) = |A_1| \cos(q\phi + [B_1]) + |A_2| \cos(q\phi + [B_2]) + \dots + |A_n| \cos(q\phi + [B_n]) \quad (3.3.13)$$

donde

$$C = A_1 e^{iB_1} + A_2 e^{iB_2} + \dots + A_n e^{iB_n}$$

$$f(\phi) = \sum_{q=1}^n |Z_{pq}^I Z_{pq}^J| \cos(q\phi + ([Z_{pq}^J] - [Z_{pq}^I])) \quad (3.3.14)$$

donde

$$A_q \in \mathbb{R}^+$$

$$B_q \in [-\pi, \pi]$$

Si en la Ecuación 3.3.14 se hace variar a ϕ de 0 a 2π , se traza una gráfica donde el punto mínimo corresponde al ángulo con el que la imagen J esta rotada con respecto a la imagen I.

Una forma eficiente de encontrar este mínimo es conociendo los mínimos de la Ecuación 3.3.14 a través de su derivada. Se puede demostrar que no es necesario evaluar la derivada con

infinidad de valores ya que la Ecuación 3.3.14 tiene un espectro de Fourier discreto y está limitado por una frecuencia máxima $f^{MAX} = \frac{N}{2\pi}$, es decir, que la función tiene como máximo N mínimos locales en el intervalo de 0 a 2π . De acuerdo al teorema de Nyquist-Shannon dos muestras consecutivas no cambian abruptamente, si la señal fue muestreada a la frecuencia de Nyquist, por lo que la distancia mínima entre dos mínimos consecutivos es $\frac{2\pi}{2N} = \frac{\pi}{N}$. Por lo tanto los valores que debe tomar ϕ deben ser múltiplos de $\frac{\pi}{N}$, pero para asegurarse que solo un mínimo (o un máximo) está presente en un intervalo en vez de dividir la Ecuación 3.3.14 en $2N$ intervalos, se divide en $4N$. Por lo que la función va a ser evaluada cada $\frac{2\pi}{4N} = \frac{\pi}{2N}$.

Por último la ecuación que se propone en [3] para conocer el ángulo con el que una imagen se encuentra rotada con respecto a otra es la Ecuación 3.3.15.

$$x_{min} = x_n + \frac{\pi f'(x_n)}{2N (f'(x_n) - x_n f'(x_{n+1}))} \quad (3.3.15)$$

donde

$$\{x_n = \frac{n\pi}{2N} \mid 0 \leq n \leq 4N\}$$

4. Evaluación de algoritmos

En este capítulo se muestran resultados de pruebas realizadas a los algoritmos propuestos en [3, 16], para validar su desempeño.

Primero se muestran las pruebas de invariancia del algoritmo rápido de [16]. Para confirmar que la implementación de este algoritmo es correcta se compararon sus resultados con los resultados de una lista de equivalencias, Apéndice Apéndice A, entre los momentos de Hu, y los momentos de Zernike (MZ). Tales equivalencias fueron propuestas por Teague en [21]. Se prosigue con las pruebas de invariancia de los momentos MZ, observando el comportamiento de los MZ tanto sin invariancia, como con invariancia. De esta manera las primeras tres pruebas están diseñadas para analizar de forma independiente la invariancia en traslación, escala y rotación. Las siguientes tres pruebas son combinaciones de dos de las tres invariancias, es decir: a) invariancia en traslación y escala, b) invariancia en traslación y rotación e c) invariancia en escala y rotación. La última prueba combina las tres invariancias simultáneamente: traslación, escala y rotación. Con el fin de observar claramente el comportamiento de los momentos MZ ante estas transformaciones, en cada una de las pruebas se utilizan imágenes de Lenna, que se obtienen a partir de la misma imagen original, pero se encuentran modificadas en traslación, rotación y escala.

En [3] con propósitos de encontrar similitudes entre dos imágenes, así como el ángulo de rotación que tiene una imagen con respecto a la otra se plantea la distancia euclidiana entre sus momentos MZ. Esto obtenido a partir de la ejecución de un solo algoritmo, para que haya consistencia en los resultados. En esta etapa para el estudio de la rotación se utilizaron imágenes de Lenna con las mismas dimensiones, sin trasladado alguno, pero rotadas con ciertos ángulos pre-establecidos. Por último también para efectos de comparación, con el algoritmo propuesto en [3], tomando un par de imágenes se calculó la distancia euclidiana y su desfase.

4.1. Validación de la invariancia en los momentos de Zernike (MZ)

La invariancia es una característica que se indica que una propiedad no cambia cuando la función, estructura o sistema a la que pertenece es sometida a una transformación. Cuando esto sucede se dice que tal propiedad es invariante respecto a la transformación.

Haciendo referencia a los momentos, en este trabajo se estableció que bajo ciertas circunstancias los momentos geométricos son invariantes a cambios de escala, traslación y

4. Evaluación de algoritmos

rotación. Sin embargo al calcular los momentos por medio de una computadora los valores numéricos de los momentos que se obtienen, no son del todo invariantes debido al problema de precisión numérica. A saber, los momentos cambian dentro de un cierto rango dinámico, lógicamente entre menor sea el valor de este rango dinámico la invariancia de los momentos es mayor (considerando el valor numérico). Dicho de otra manera, entre menor sea la diferencia entre los momentos de referencia y los momentos de las imágenes utilizadas para evaluar la invariancia en cuestión, la precisión numérica es mayor.

De acuerdo con [3] los dieciséis primeros ordenes son representativos de los momentos MZ pero a partir del doceavo orden la información subsecuente no es tan significativa. Tomando esto en cuenta en las pruebas realizadas en este trabajo se utilizan momentos MZ del orden cero al quince, haciendo un total de 72 momentos MZ.

$q \backslash p$	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
0	0	2	4	6	8	10	12	14	16	18	20	22	24	26	28	30
1		2	4	6	8	10	12	14	16	18	20	22	24	26	28	30
2			4	6	8	10	12	14	16	18	20	22	24	26	28	30
3				6	8	10	12	14	16	18	20	22	24	26	28	30
4					8	10	12	14	16	18	20	22	24	26	28	30
5						10	12	14	16	18	20	22	24	26	28	30
6							12	14	16	18	20	22	24	26	28	30
7								14	16	18	20	22	24	26	28	30
8									16	18	20	22	24	26	28	30
9										18	20	22	24	26	28	30
10											20	22	24	26	28	30
11												22	24	26	28	30
12													24	26	28	30
13														26	28	30
14															28	30
15																30

Figura 4.1.1.: Momentos MZ calculados en cada prueba.

En cada subsección se presentan gráficas de las diferencias euclidianas entre los 72 momentos MZ. Con el objetivo de comparar en la misma gráfica las diferencias de los 72 momentos MZ de todas las imágenes involucradas en la prueba, se emplea una gráfica de barras acumulativas. También se presenta un cuadro en donde se sintetizan los resultados obtenidos utilizando valores promedio. El cálculo del promedio suma las 72 diferencias de

4.1. Validación de la invariancia en los momentos de Zernike (MZ)

los momentos MZ calculados de cada una de las imágenes a prueba y el total se divide en el número total de sumas. El eje de las ordenadas de estas gráficas representa el orden de los momentos MZ, la graduación de este eje es no lineal debido a que el orden de los momentos MZ definido como $p + q$ tiene las siguientes restricciones:

1. $|q| \leq p$
2. $p - |q|$ debe ser par.

Arbitrariamente se ordenan los ordenes de los momentos MZ calculados de menor a mayor, iniciando con los ordenes de la diagonal. Es decir inicia con $p = q$ y los valores consecutivos de p van en aumento mientras que los de q en decremento. Ver la Figura 4.1.1.

4.1.1. Invariancia respecto a la traslación

La propiedad de invariancia a la traslación, indica que el valor de los momentos se mantiene casi constante, independientemente de la ubicación de un objeto bidimensional de un área determinada, en un plano.

De acuerdo a sus propiedades, se sabe de antemano que por sí solos los MZ sólo ofrecen invariancia a la rotación, no obstante para lidiar con la invariancia en traslación se tienen dos opciones bastante simples. La primera consiste en centrar la imagen, y la segunda en hacer sus momentos invariantes a la traslación. A continuación se analizan ambos casos haciendo uso de cuatro imágenes diferentes, ver Figura 4.1.2¹, ubicadas en las esquinas de un rectángulo con el fin de apreciar más claramente los casos extremos.



Figura 4.1.2.: Imágenes varias de Lenna de 1024×1024 píxeles.

Invariancia haciendo un centrado de la imagen

El algoritmo de esta prueba consiste en modificar la imagen de tal forma que el centroide y el centro de la imagen coincidan. Lo cual por lo general no es posible puesto que las coordenadas del centroide generalmente son números fraccionarios, y las posiciones de los píxeles se manejan en coordenadas enteras.

¹Las imágenes mostradas en este trabajo no tienen la escala real, están ajustadas de acuerdo a los márgenes del documento.

4. Evaluación de algoritmos

Invariancia con centroide

La idea de este método de invariancia es similar a la anterior, pero en vez de centrar la imagen físicamente, el centrado se realiza simultánea e intrínsecamente al cálculo de los momentos MZ. En este caso el centrado puede resultar más preciso, puesto que depende directamente de la exactitud numérica que trabaje la computadora en la que se realizan los cálculos.

Comparación de los resultados de cada método

Para comparar gráficamente los resultados de cada método se ilustran en la Figura 4.1.3 tres gráficas, en forma de barras acumulativas, las diferencias que guardan cada uno de los ordenes de los momentos MZ de las imágenes 4.1.2b, 4.1.2c, 4.1.2d, 4.1.2e con respecto a una imagen de referencia. La Imagen 4.1.2a es la que funge como referencia, donde sus MZ no son sometidos a ningún tipo de invariancia. Las imágenes de estas gráficas tienen dimensiones de 1024×1024 píxeles, donde las dimensiones que ocupa la imagen de Lenna es de un cuarto del área, es decir 512×512 .

En la primera Gráfica 4.1.3a únicamente se utiliza la (Ecuación 3.3.1) sin utilizar ningún método de invariancia. Para la siguiente Gráfica 4.1.3b, se ilustran los momentos MZ invariantes a la traslación utilizando el centroide. En la última Gráfica 4.1.3c, los momentos obtenidos no son invariantes a la traslación, pero cada imagen se centra previamente al calculo de sus momentos MZ.

Esta prueba se repitió para tres escalas más, 100×100 , 192×192 y 512×512 píxeles. Es decir, a partir de cinco imágenes de 512×512 píxeles se obtuvieron cuatro grupo de diferencias de momentos MZ, con gráficas muy similares a las mostradas en la Figura 4.1.3.

	100×100	192×192	512×512	1024×1024
sin invariancia	$1,3750 \times 10^4$	$1,3538 \times 10^4$	$8,6045 \times 10^3$	$9,0499 \times 10^3$
inv. c/centroide	0.2872	0.2897	0.2556	0.3166
centrado manual	0.3513	0.2962	0.2685	0.3350

Tabla 4.1.: Diferencias de invariancia a traslación con grupos de imágenes de diferente tamaño, la imagen de referencia es la de la posición 0.

Con el fin de comparar numéricamente estos resultados se promedian todas las diferencias en cada uno de los métodos. En Tabla 4.1 se concentran los valores promedio de las pruebas realizada en esta subsección.

Observando los resultados en la Tabla 4.1, no obstante que las dos técnicas en esencia realizan el mismo proceso: centrar la imagen haciendo coincidir el centro geométrico con el centro de masa, los resultados son diferentes.

Esta diferencia se debe en parte a un error de precisión. En ocasiones el valor que ocupa el centro geométrico es fraccionario, por ejemplo en una imagen de 50×79 píxeles, el centro

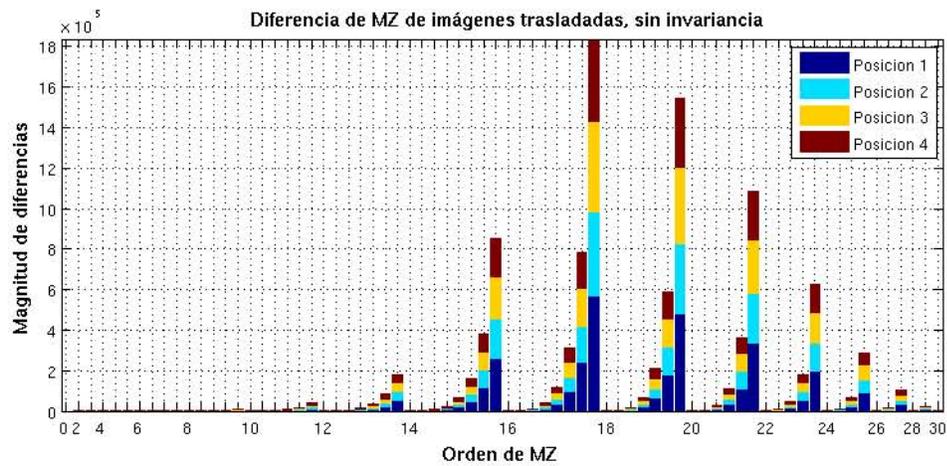
4.1. Validación de la invariancia en los momentos de Zernike (MZ)

geométrico se localiza en el píxel 25 en el eje de las abscisas, sin embargo el valor de la ordenada esta en 38.5 píxeles. Al calcular el centroide, éste casi siempre se encuentra en valores fraccionarios, que de hecho son valores irracionales en muchos casos. En ambos casos estos valores tienen que modificarse con un redondeo, o truncamiento, para poder mover los valores de los píxeles al centro de la imagen. Como consecuencia las coordenadas resultantes de la diferencia entre el centro geométrico y el centroide, son diferentes en cada caso, haciendo que los MZ sean ligeramente diferentes entre sí.

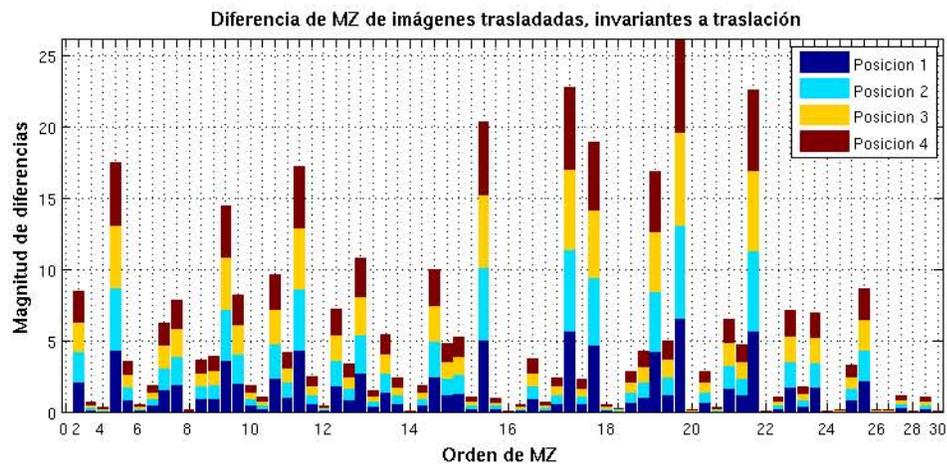
En cuanto a la invariancia respecto a la traslación haciendo uso del centroide de manera simultánea, e intrínseca, al cálculo de los momentos, los resultados también difieren. En este caso la diferencia se justifica con un error de exactitud debido a la limitante del *hardware* en el que se hacen los cálculos. Lo que conlleva a resultados diferentes entre ambos métodos.

Conforme a la definición de invariancia, se busca que las diferencias entre momentos MZ sean cero. Entonces cuanto más se acerquen los valores de las diferencias al cero, indica mayor invariancia y mejores resultados. En Tabla 4.1 el valor más pequeño se encuentra en la columna de las imágenes de 192×192 y en la fila del *centra manual*. Pese a este resultado se considera que el mejor método es el que utiliza el centroide, debido a que a excepción de este resultado el resto son menores que los resultados del centrado manual.

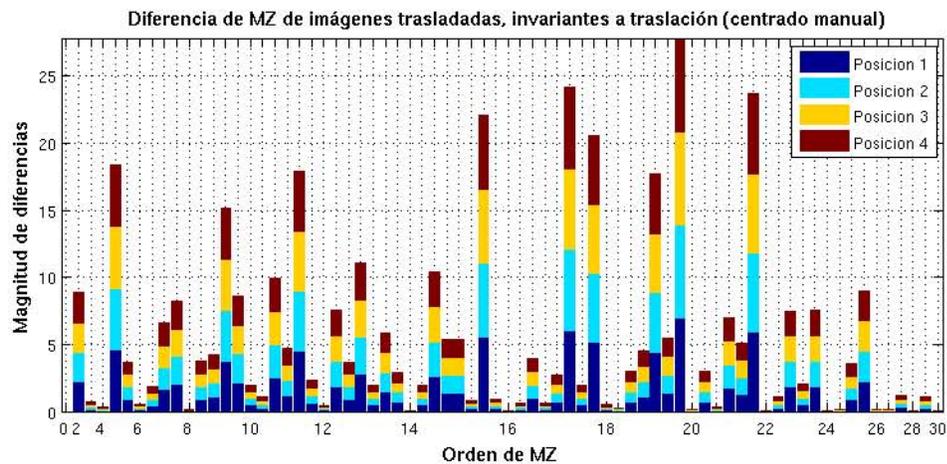
4. Evaluación de algoritmos



(a) Diferencia de MZ sin invariancia



(b) Diferencia de MZ con invariancia en traslación usando el centroide



(c) Diferencia de MZ con invariancia en traslación, centrando manual

Figura 4.1.3.: Gráficas comparativas de MZ invariantes a la traslación, la imagen de referencia es la de la posición 0.

4.1.2. Invariancia respecto a la escala

La invariancia respecto a la escala se refiere a la propiedad de que permanecen sin cambio cuando las dimensiones de un objeto son multiplicadas por un determinado factor diferente de cero. No obstante en el contexto de una representación digital se tienen algunas restricciones cuando los números que se manejan no son enteros.

Específicamente en el caso de imágenes con texturas complejas, como son las imágenes reales, el producto por un factor no existe como tal. Por lo general se utiliza un submuestreo cuando el factor de escala es menor que la unidad, o un sobremuestreo cuando el factor de escala es mayor a la unidad. En el caso del submuestreo tiene un límite inferior el cual depende de las dimensiones de cada imagen, de tal forma que existe un valor donde la imagen submuestreada ya no es capaz de representar la información contenida en la imagen original.

Para las pruebas de invariancia a la escala en este trabajo se utilizan ocho imágenes de Lenna en diferentes escalas dando como resultado las dimensiones en píxeles siguientes: 50×50 , 96×96 , 256×256 , 360×360 , 480×480 , 500×500 , 512×512 y 1024×1024 . A las que se les aplican dos métodos de invariancia en escala, el primero utiliza el momento de orden cero elevado a una potencia variable λ y el segundo utiliza únicamente el momento de orden cero.

Factor de escalamiento λ

La ecuación de momentos invariantes a la traslación y a la escala propuesta en [10] es $\frac{\mu_{pq}}{\mu_{00}^{(p+q)/2+1}}$, $p + q = 2, 3, \dots$, donde μ representa los momentos invariantes a la traslación. Hu llega a este valor basándose en el método de invariantes algebraicas.

Posteriormente en [9] los momentos invariantes a la traslación y a la escala tienen la siguiente forma $\frac{\mu_{pq}}{\mu_{00}^{(p+q+2)/2}}$. Este resultado es producto de normalizar la imagen, primero con respecto a la traslación y posteriormente con respecto a la escala.

En el presente trabajo se hacen pruebas con un factor de escala $\lambda = (p + q + 2)/2$, que es el mismo que Teague utiliza en [9]. En esta parte del estudio se excluye la invariancia a la traslación dado que el objetivo particular es analizar el comportamiento de la invariancia respecto a la escala. Siendo entonces el propósito analizar el rango de invariancia de dos de los métodos en estudio, para efectos de precisión en los cálculos y dado que los momentos utilizados no son invariantes a la traslación $\frac{m_{pq}}{m_{00}^{(p+q+2)/2}}$.

Si se observa la ecuación $\frac{m_{pq}}{m_{00}^{(p+q+2)/2}}$, misma con la que se analiza la invariancia a la traslación, se aprecia que existe la posibilidad de poder desarrollar el denominador y manipularlo independiente a los valores de p y q , lo que puede aprovecharse para un cálculo más eficiente.

4. Evaluación de algoritmos

En cuanto al rango de invariancia en el que fluctúan los resultados obtenidos con las versiones de escalado propuestas tanto por Hu, como por Teague la discrepancia es mínima, alcanzando una diferencia cuadrática del orden de centésimas.

Momento de orden cero

En [11] se plantea que para hacer los momentos invariantes a la escala basta con dividirlos entre su área. Sin embargo el momento de orden cero para imágenes en escala de grises representa la superficie de la imagen y no el área, como ocurre con las imágenes binarias. Se propone entonces el momento de orden cero como único factor de escalamiento.

Comparación de los métodos

A continuación se hace un análisis comparativo para cuantificar las diferencias entre los métodos. Para visualizar estas diferencias se emplean las gráficas de la Figura 4.1.4 donde en la Gráfica 4.1.4a se ilustran los momentos MZ sin ningún método de invariancia, en la Gráfica 4.1.4b se muestran los momentos escalados usando $\lambda = (p + q + 2)/2$, y en la Gráfica 4.1.4c el escalamiento se hace dividiendo los MZ entre el momento de orden cero.

La imagen más grande (1024×1024 píxeles) es la que tiene mayor resolución y por lo tanto sus MZ son los que sirven como referencia.

En la Gráfica 4.1.4b no aparecen la mayoría de las diferencias debido a que los valores son menores en escala logarítmica. También se aprecia que las mayores diferencias se obtienen en la imagen más pequeña 50×50 , seguidas de las correspondientes a la imagen de 96×96 píxeles y así consecutivamente. Este comportamiento se justifica porque entre más pequeña es la imagen menor información contiene lo que implica una mayor diferencia.

También se observa que conforme se incrementa el orden de los momentos MZ el valor de su magnitud es menor. Esto puede resultar conveniente al momento de clasificar, puesto que al comparar los momentos MZ de una imagen con respecto a los momentos MZ de otra, los momentos de orden alto van a influir en menor proporción que los de orden bajo, mismos que contienen mayor información.

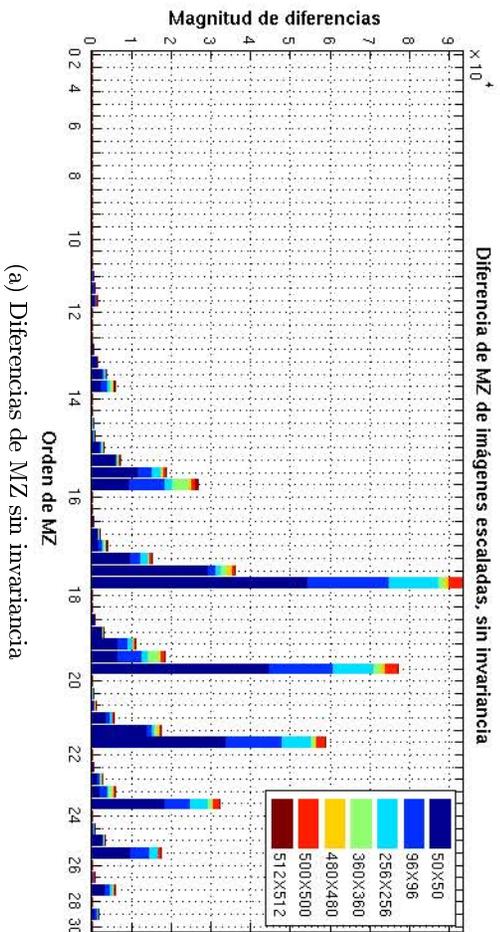
En Tabla 4.2 se proporciona el promedio de las diferencias de los momentos MZ de las tres imágenes de la Figura 4.1.4:

sin invariancia	425.5319
inv. en traslación, usando mto. 0	0.0117
inv. en traslación, usando λ	$5,9347 \times 10^{-7}$

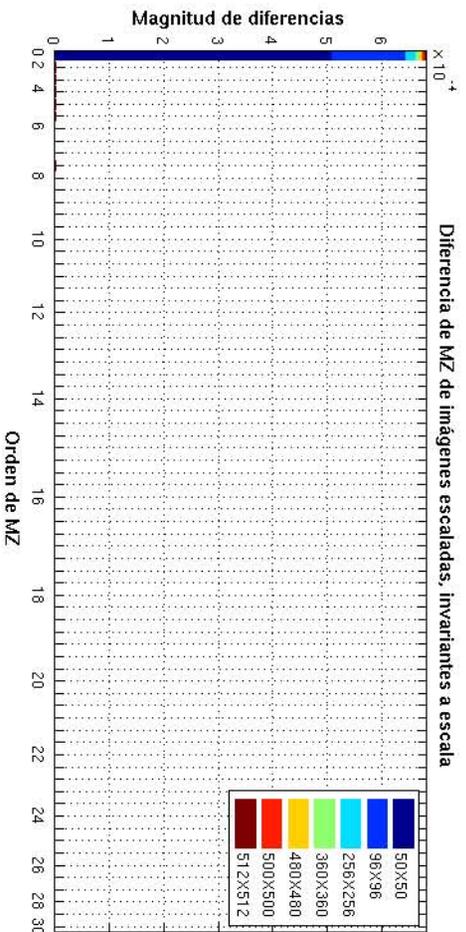
Tabla 4.2.: Diferencias de invariancia a escala, la imagen de referencia es la de mayor escala.

En esta ocasión indudablemente el menor y mejor resultado es el método que utiliza λ como factor de escalamiento. Cabe mencionar que cuando los momentos MZ están normalizados el uso del momento cero como factor de escalamiento no es muy útil.

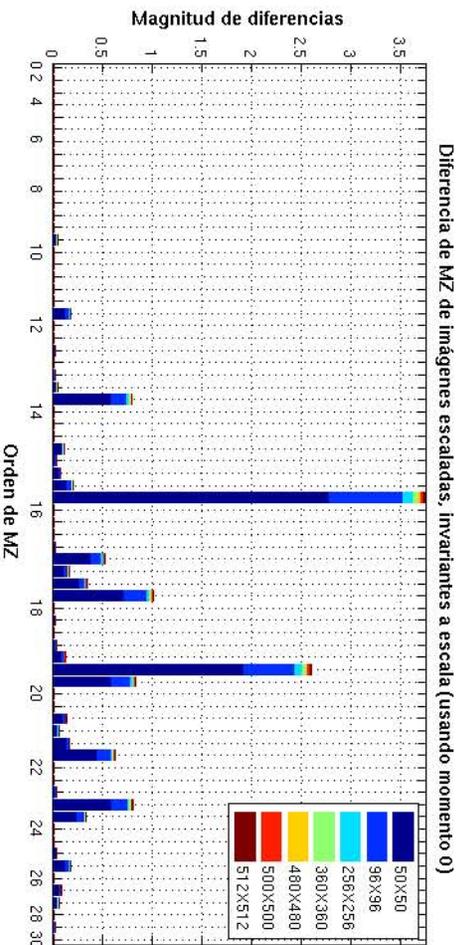
4.1. Validación de la invariancia en los momentos de Zernike (MZ)



(a) Diferencias de MZ sin invariancia



(b) Diferencias de MZ con invariancia a escala, usando λ



(c) Diferencias de MZ con invariancia a escala, usando el momento de orden cero

Figura 4.1.4.: Gráficas comparativas de los momentos MZ invariantes a la escala, la imagen de referencia es la de mayor escala.

4.1.3. Invariancia respecto a la rotación

Cuando se gira una imagen respecto a un eje y una de sus propiedades se conserva, se dice que esa propiedad es invariante a la rotación. La magnitud de los momentos MZ se caracterizan por tener esta propiedad.

Con la finalidad de observar la invariancia que hay entre los momentos MZ de dos imágenes rotadas, se hacen pruebas utilizando cuatro imágenes rotadas sucesivamente, cada $\frac{\pi}{2}$ radianes respecto a una de ellas que se considera como referencia. Ver Figura 4.1.5.

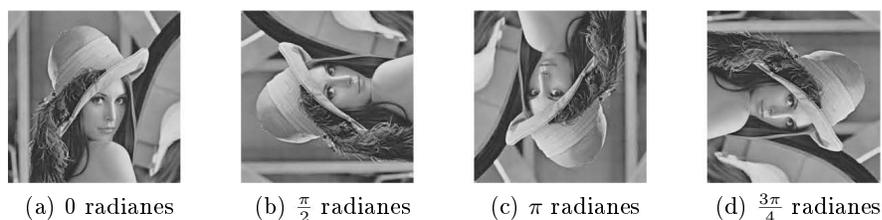


Figura 4.1.5.: Imágenes de 256X256 rotadas cada $\frac{\pi}{2}$ radianes, utilizadas para probar la invariancia a la traslación en los MZ.

En la Gráfica 4.1.6a se muestran las diferencias de los momentos MZ de las imágenes de la Figura 4.1.5 cuando no se tiene invariancia, es decir, la magnitud de las diferencias entre números complejos. En la misma figura se muestran los momentos MZ con invariancia respecto a la rotación, en este caso primero se obtienen las magnitudes de los momentos MZ y posteriormente se efectúa la diferencia.

Arbitrariamente se elige como referencia la Imagen 4.1.5a, entonces se tienen tres grupos de diferencias de momentos MZ que corresponden a las otras Imágenes 4.1.5b, 4.1.5c y 4.1.5d. En la Gráfica 4.1.6a se ilustran las diferencias entre los momentos MZ de la imagen de referencia y las otras imágenes giradas.

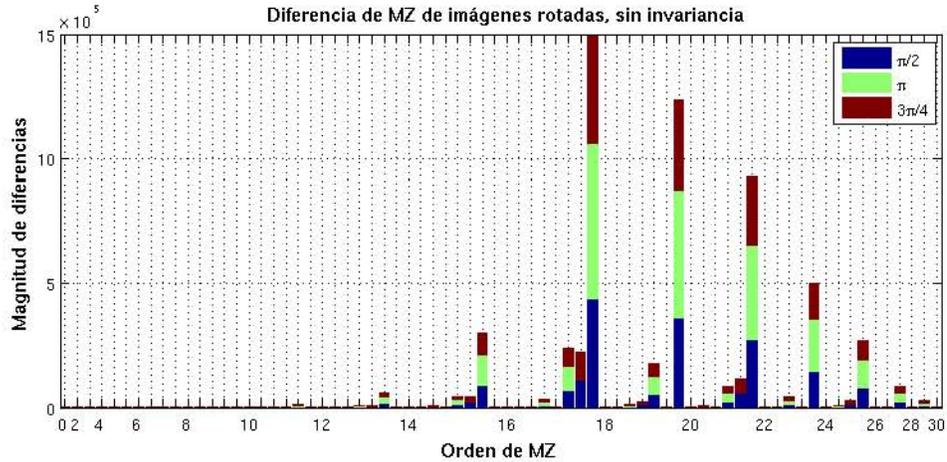
Se puede notar que en este caso la invariancia respecto a la rotación cambió en forma lineal en la medida de las magnitudes de las diferencias, comparadas con la invariancia en escala que altera las magnitudes en escala logarítmica, o bien respecto a la invariancia en traslación que altera las magnitudes de las diferencias de manera un tanto aleatoria.

sin invariancia	$5,2651 \times 10^3$
invariancia en rotación	0.4729

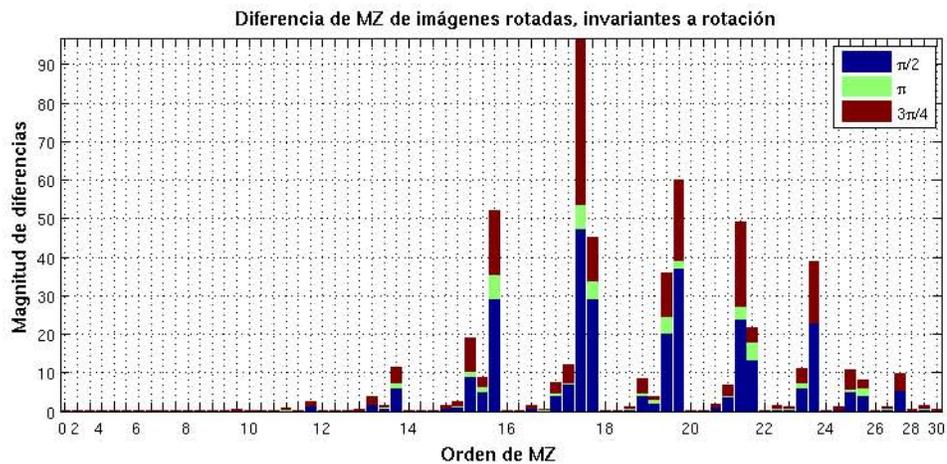
Tabla 4.3.: Diferencias de invariancia a rotación, la referencia es la Imagen 4.1.5a.

4.1. Validación de la invariancia en los momentos de Zernike (MZ)

En este caso no hay otro método contra el cual comparar, sin embargo se aprecia una diferencia notable en cuanto a los promedios, ver Tabla 4.3.



(a) Diferencias de MZ sin invariancia



(b) Diferencias de MZ con invariancia a rotación

Figura 4.1.6.: Gráficas comparativas de momentos MZ invariantes a rotación, la referencia es la Imagen 4.1.5a.

4.1.4. Invariancia respecto a la traslación y a la escala

A partir de estas pruebas se utilizan los métodos de invariancia que presentaron mayor invariancia con respecto a la referencia. En el caso de la invariancia a la traslación se utilizan

4. Evaluación de algoritmos

los momentos centrales basados en el centroide, y en el caso de la invariancia a la escala se utiliza el parámetro λ como factor de escalamiento propuesto por Teague.

Para esta prueba se requiere de imágenes no centradas y de diferente tamaño. En total se utilizan nueve imágenes diferentes, ocho de ellas trasladadas con respecto al centro. Cuatro de 256×256 píxeles y cuatro de 512×512 píxeles; en lo sucesivo se les denominará “imágenes pequeñas” e “imágenes grandes”, respectivamente. La novena es una imagen grande centrada originalmente que se utiliza como referencia.

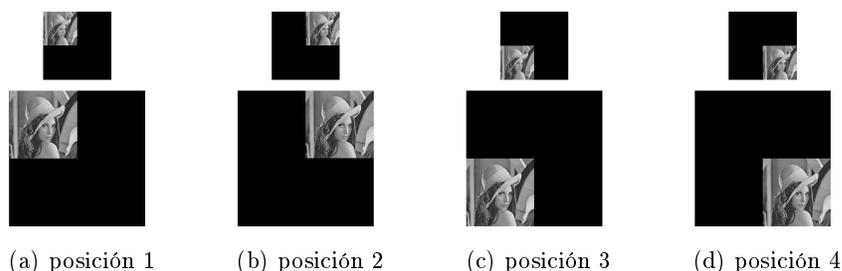


Figura 4.1.7.: Imágenes utilizadas para prueba de la invariancia a la traslación y a la escala.

Para estas pruebas se calculan nuevamente momentos MZ sin ninguna invariancia, y posteriormente con invariancia en rotación y traslación. En la Gráfica 4.1.8a se muestran las diferencias de los momentos sin utilizar invariancia en escala, ni traslación. Y en la Gráfica 4.1.8b se muestran las diferencias de los momentos escalados y centrados.

En la Gráfica 4.1.8b se obtiene igualmente que las magnitudes de las diferencias se decrecientan en escala logarítmica. Esto se debe al hecho de imponer a los momentos la invariancia a escala. Por otro lado se puede observar una diferencia entre las magnitudes de los MZ de las imágenes grandes y las pequeñas, siendo la diferencia con respecto a las imágenes pequeñas mayor debido a que la referencia es una imagen grande, y por lo tanto no corresponden en tamaño.

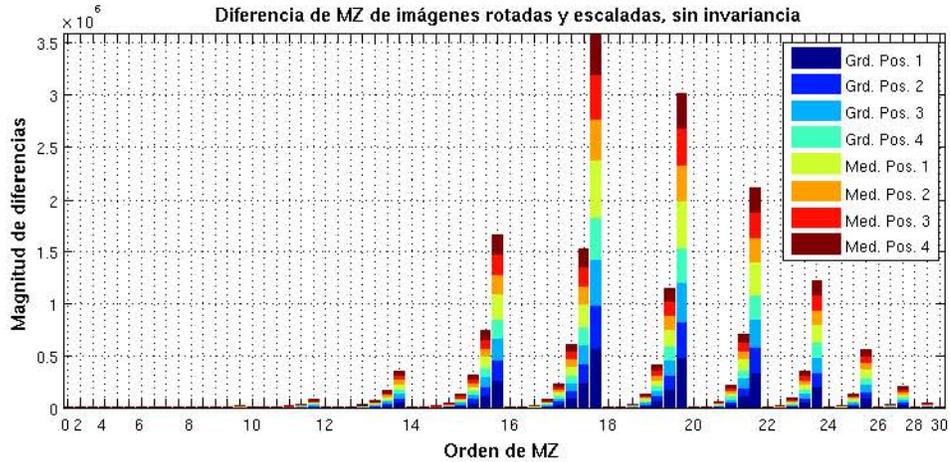
sin invariancia	$1,7654 \times 10^4$
invariancia en traslación y escala	$1,2649 \times 10^{-8}$

Tabla 4.4.: Diferencias de invariancia a traslación y escala, la referencia es una imagen grande (512×512) centrada originalmente.

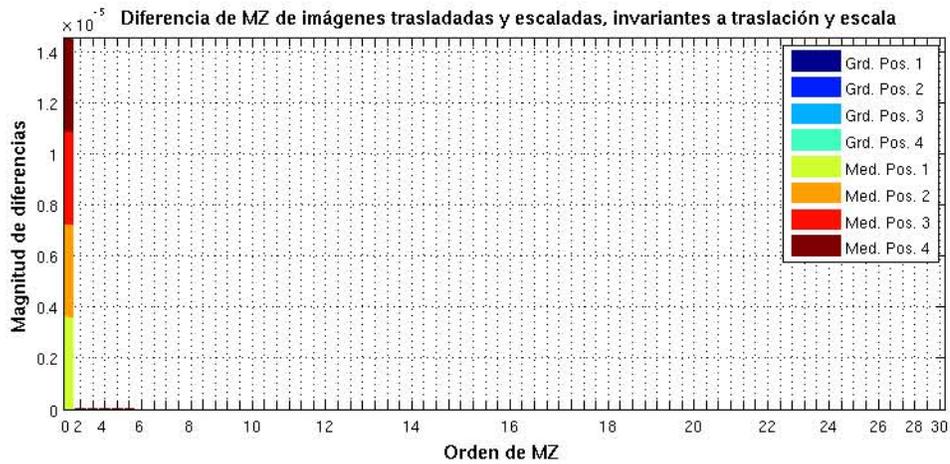
Los valores promedio en Tabla 4.4 muestran que la invariancia es mayor que en los casos que solo se aplica un método de invariancia. Al aplicar invariancia en traslación el valor promedio tiene una reducción de cinco ordenes, en la invariancia a la escala la reducción de

4.1. Validación de la invariancia en los momentos de Zernike (MZ)

orden es de nueve. En este caso donde se aplican ambas invariancias (traslación y escala) la reducción es de doce ordenes, mayor a los anteriores.



(a) Diferencias de MZ sin invariancia



(b) Diferencias de MZ con invariancia a traslación y a escala

Figura 4.1.8.: Gráficas comparativas de momentos MZ invariantes a la traslación y a la escala, la referencia es una imagen grande (512×512) centrada originalmente.

4.1.5. Invariancia respecto a la traslación y a la rotación

En esta prueba se analiza el comportamiento de los momentos MZ con las Figura 4.1.9 que se encuentran en diferentes ubicaciones respecto al área global que las contiene y si-

4. Evaluación de algoritmos

multáneamente están giradas con diferentes ángulos de rotación. Las dimensiones globales de las imágenes son de 512×512 píxeles, mientras que las dimensiones de la sección de imagen de Lenna son de 256×256 píxeles.

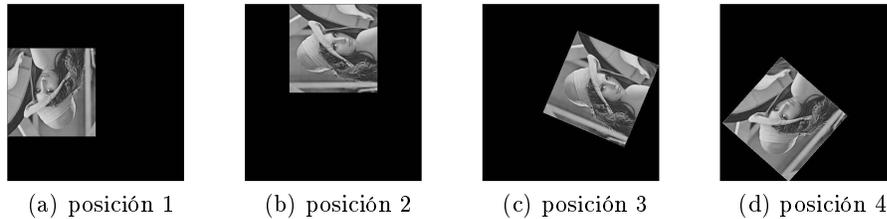


Figura 4.1.9.: Imágenes utilizada para probar la invariancia a la traslación y a la rotación.

En las gráficas de la Figura 4.1.10 se tienen las diferencias de los momentos MZ de las imágenes de la Figura 4.1.9 menos los MZ de la imagen de referencia, imagen centrada originalmente y sin rotación.

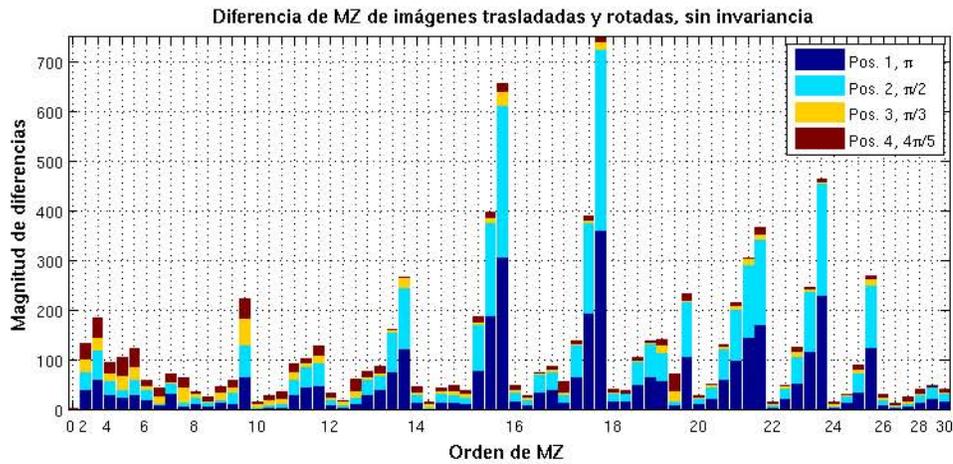
En la Gráfica 4.1.10a se observa que las diferencias presentadas por los MZ de las imágenes que se encuentran rotadas en un ángulo que no es múltiplo de $\frac{\pi}{2}$ presentan mayor diferencia en comparación con aquellas en donde el ángulo de giro si es múltiplo de $\frac{\pi}{2}$. Sin embargo en la Gráfica 4.1.10b esta mayor diferencia no se tiene para ángulos de giro que no son múltiplos de $\frac{\pi}{2}$ evidenciando que la invariancia en rotación implementada es general, no hace distinción entre ángulos de giro.

sin invariancia	7.5544
invariancia a traslación y rotación	0.1679

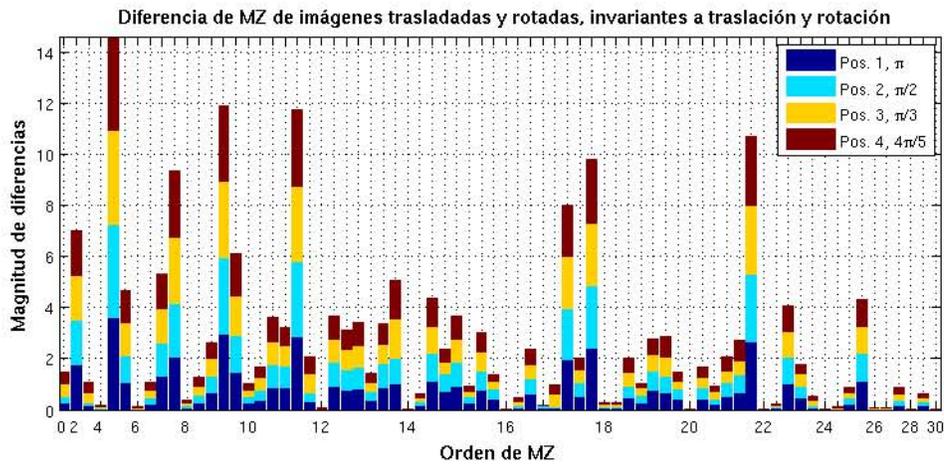
Tabla 4.5.: Diferencias de invariancia a la traslación y a la rotación, referencia imagen centrada originalmente y sin rotación.

La Tabla 4.5 muestra un promedio de las diferencias entre ambas pruebas. En este caso la reducción es de solo un orden.

4.1. Validación de la invariancia en los momentos de Zernike (MZ)



(a) Diferencias de MZ sin invariancia



(b) Diferencias de MZ con invariancia a traslación y a rotación

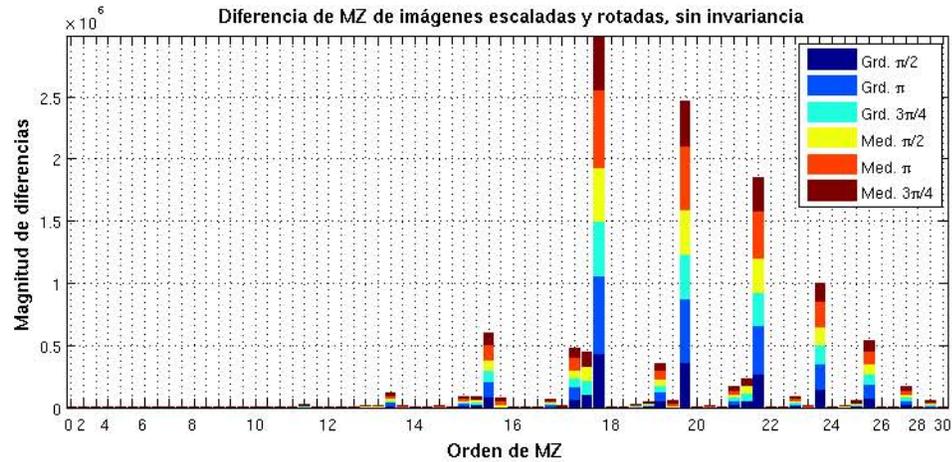
Figura 4.1.10.: Gráficas comparativas de momentos MZ invariantes a la traslación y a la rotación, referencia imagen centrada originalmente y sin rotación.

4.1.6. Invariancia respecto a la escala y a la rotación

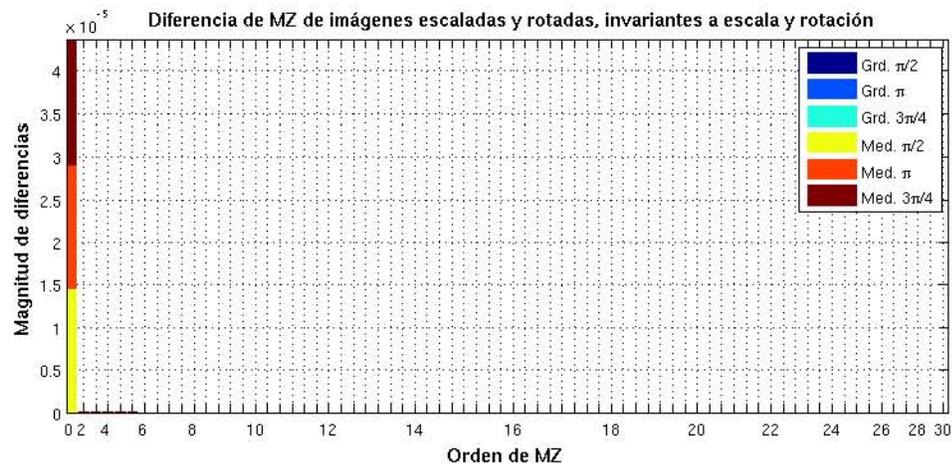
Con el propósito de llevar a cabo esta prueba se requirieron imágenes en diferentes escalas y con diferentes ángulos de rotación, por lo que se utilizan tres de las imágenes rotadas, que tienen ángulo de rotación diferente de 0. Son imágenes de Lenna con dimensiones 256×256 píxeles, ver Figura 4.1.5, y de 512×512 píxeles, estas últimas imágenes con los mismos ángulos de rotación. Generando de esta manera un total de seis imágenes de prueba, de las

4. Evaluación de algoritmos

cuales por conveniencia para facilitar la comparación, se utiliza como referencia una imagen extra de Lenna con 0 radianes de rotación y de dimensiones 512×512 píxeles.



(a) Diferencias de MZ sin invariancia



(b) Diferencias de MZ invariancia a rotación y a escala

Figura 4.1.11.: Gráficas comparativas de momentos MZ invariantes a la escala y a la rotación, referencia imagen grande (512×512) y sin rotación.

De acuerdo a la Gráfica 4.1.11a se observa que no hay distinción entre las diferencias de los MZ de las imágenes pequeñas y las diferencias de los MZ de las imágenes grandes. En cambio en la Gráfica 4.1.11b se tienen cambios en las magnitudes de las diferencias en escala logarítmica, esto generado por la invariancia a la escala. Para imágenes medianas

4.1. Validación de la invariancia en los momentos de Zernike (MZ)

las diferencias de los MZ son mayores comparadas con las diferencias de las magnitudes de los MZ de las imágenes grandes que en la Gráfica 4.1.11b no son visibles (por ser tan pequeñas). Los valores promedios de las diferencias se resumen en la Tabla 4.6, nuevamente los promedios decrecientan significativamente en once ordenes.

sin invariancia	$0,654 \times 10^4$
invariancia a escala y rotación	$3,7946 \times 10^{-8}$

Tabla 4.6.: Diferencias de invariancia a la escala y a la rotación, referencia imagen grande (512×512) y sin rotación.

4.1.7. Invariancia respecto a la traslación, la escala y la rotación

En esta subsección se hacen las pruebas necesarias para estudiar el comportamiento de las tres invariancias simultáneamente: escala, traslación y rotación. Para esta prueba, en total se utilizan diez imágenes, de las cuales cuatro son imágenes pequeñas 512×512 píxeles, Figura 4.1.9, otras cuatro corresponden a imágenes grandes 1024×1024 píxeles, Figura 4.1.12 y dos más están centradas originalmente y sin rotar, similares a la Imagen 4.1.2a, una es pequeña y otra grande. Para facilitar la comparación, la referencia es una imagen grande, centrada originalmente y sin rotar.

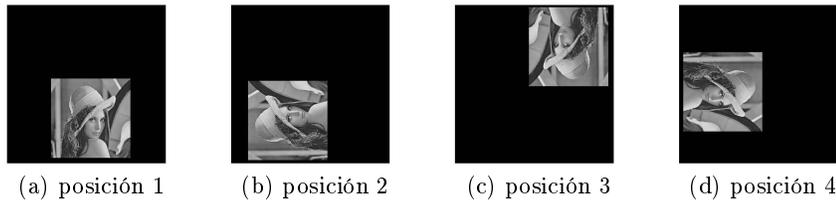
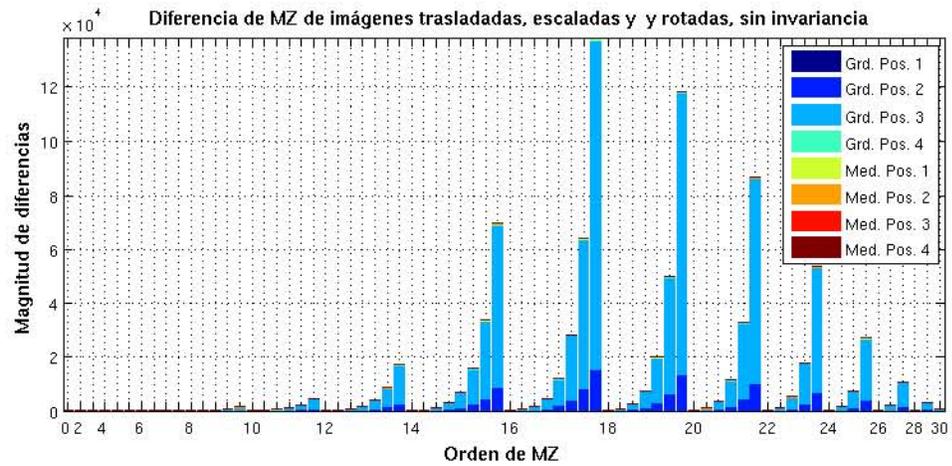


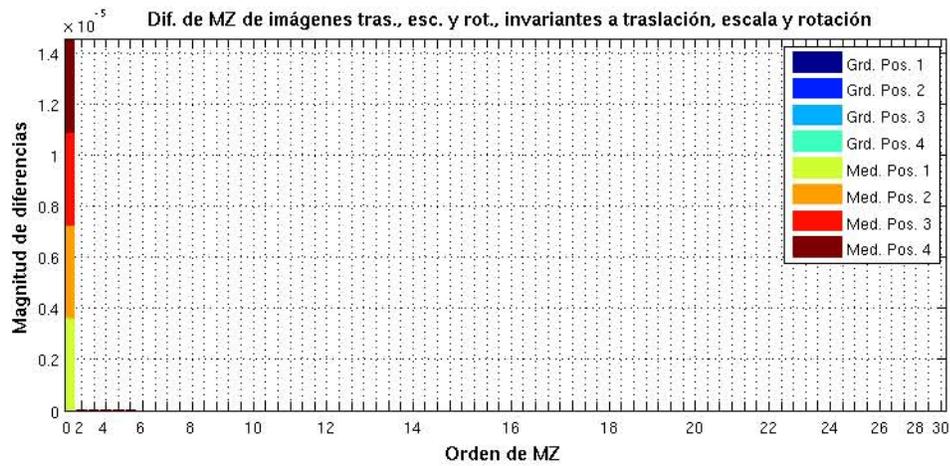
Figura 4.1.12.: Imágenes utilizadas para la prueba de invariancia respecto a la traslación, a la escala y rotación.

En la Gráfica 4.1.13a se ilustra la evolución de los momentos MZ cuando no hay invariancia a la traslación, a la escala y a la rotación. Resaltando el predominio con la mayor diferencia en los MZ la que corresponde a la imagen grande de Lenna en la posición 3, que presenta un desfase de π radianes respecto a la imagen de referencia. La Gráfica 4.1.13b presenta las diferencias de los MZ con invariancia en traslación, escala y rotación.

4. Evaluación de algoritmos



(a) Diferencias de MZ sin invariancia



(b) Diferencias de MZ invariancia a traslación, a escala y a rotación

Figura 4.1.13.: Gráficas comparativas de momentos MZ invariantes respecto a la traslación, a la escala y a la rotación, referencia imagen grande (512×512) centrada originalmente y sin rotación.

Observando nuevamente la influencia en los valores de las diferencias por el hecho de utilizar la invariancia en escala, la cual modifica la magnitud de las diferencias en escala logarítmica, además de una marcada diferencia entre las escalas. En la Tabla 4.7 se resumen los valores promedio de cada prueba.

4.2. Invariancia respecto al ángulo de rotación entre dos imágenes

sin invariancia	777.9836
invariancia a escala, traslación y rotación	$1,2649 \times 10^{-8}$

Tabla 4.7.: Diferencias de invariancia a escala, traslación y rotación, referencia imagen grande (512×512) centrada originalmente y sin rotación.

4.2. Invariancia respecto al ángulo de rotación entre dos imágenes

En esta sección se estudia el comportamiento del algoritmo que calcula el ángulo de rotación, y distancia euclidiana entre dos imágenes. El objetivo principal es verificar que el ángulo en el que se encuentra una imagen rotada respecto a otra sea el correcto. Así que todas las imágenes utilizadas serán las de Lenna (las mismas que se emplearon en la sección anterior), y siempre estarán sometidas a diferentes ángulos de rotación.

En esta prueba se consideran los MZ de dos imágenes, los correspondientes a la imagen de referencia, y aquellos obtenidos a partir de otra imagen que se encuentra rotada, considerando diversos valores de ángulos de rotación en radianes. Los MZ de la imagen de referencia y los MZ de las imágenes rotadas son comparados respecto a la distancia euclidiana, y los valores de los ángulos en radianes (de 0 a 2π) relacionan los dos grupos de MZ.

El algoritmo que calcula la distancia euclidiana está diseñado de tal forma que el resultado no sólo entrega la distancia euclidiana entre los MZ de las imágenes que recibe como entrada, sino que utiliza el valor de los ángulos en radianes para simular una rotación de una de las imágenes. Teniendo como resultado la distancia euclidiana entre los MZ de las dos imágenes, donde una de ellas esta rotada por un ángulo dado en radianes.

Para calcular el ángulo en el que difieren dos imágenes se utiliza el algoritmo de distancia euclidiana, en conjunto con un algoritmo de búsqueda con el método de gradiente descendiente, esto con la finalidad de hacer una búsqueda del mínimo de manera eficiente, utilizando solamente un grupo de MZ donde el valor del orden p es fijo y el valor del orden q varia desde 0 hasta su valor máximo. Finalmente como resultado del cálculo se tiene un par de valores, el primero es el valor más pequeño del conjunto de diferencias de distancias euclidianas calculadas, y el segundo es el ángulo en radianes en el que se encuentra rotada numéricamente la imagen de referencia. Es decir, se obtiene el valor del ángulo en radianes que hay que rotar a la imagen de referencia para obtener la distancia euclidiana mínima entre las dos imágenes, la distancia euclidiana será mínima cuando las dos imágenes son más parecidas.

En una primera prueba se utilizan imágenes en las cuales la única diferencia entre ellas es el ángulo de rotación. En otra prueba las imágenes son sometidas a dos tipos de transformaciones: a) la de rotación y escala, y b) la de rotación y traslación. En una última prueba las imágenes son sometidas a tres diferentes transformaciones simultáneamente: rotación, escala y traslación.

4. Evaluación de algoritmos

4.2.1. Prueba con diferentes rotaciones

En esta sección se usan cuatro imágenes semejantes, diferenciándose entre sí únicamente por el ángulo de rotación, 0 , $\frac{\pi}{2}$, π y $\frac{3\pi}{2}$ radianes. Las cuatro imágenes son sometidas al algoritmo de búsqueda del ángulo de desfase, se efectuaron en total cuatro pruebas usando como referencia la imagen de ángulo de rotación de 0 radianes, y comparándola con las otras, incluida la misma de referencia (0 , $\frac{\pi}{2}$, π y $\frac{3\pi}{2}$ radianes).

Para graficar el comportamiento de la distancia euclidiana entre dos imágenes se utiliza un algoritmo diferente al que se utiliza para calcular la distancia mínima y el ángulo de rotación. En este caso en lugar de llevar a cabo la búsqueda del mínimo, se hace variar el ángulo de rotación de 0 a 2π , en pasos de $\frac{\pi}{180}$ radianes.

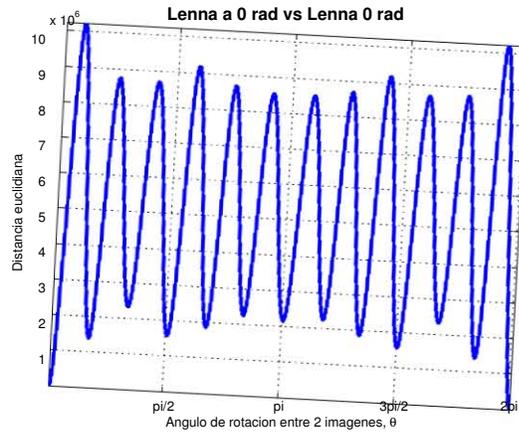


Figura 4.2.1.: Variación de la distancia euclidiana de lo MZ con respecto a ϕ , en base a Ecuación 3.3.12, entre una imagen rotada 0 radianes y otra rotada 0 radianes.

De la gráfica de la Figura 4.2.1 se observan dos mínimos globales en 0 y 2π , lo cual nos indica que cuando más se parecen las imágenes entre sí es cuando la imagen a comparar no tiene ninguna rotación. Esta gráfica es el resultado de la suma de varios cosenos, los cuales dependen del valor de q (variable que define junto con p el grado de los MZ). En la gráfica de la Figura 4.2.2 se ilustran estos cosenos donde se aprecia que el número de ciclos depende del valor de q .

Del conjunto de gráficas de la Figura 4.2.2 la de mayor magnitud (ver escala) es la correspondiente al coseno de $q = 12$, por lo tanto esa es la componente predominante en la figura Figura 4.2.1. En las gráficas de la Figura 4.2.1 y Figura 4.2.2 el valor de la ordenada está definido por la distancia euclidiana, y en la abscisa por los ángulos de rotación en radianes.

4.2. Invariancia respecto al ángulo de rotación entre dos imágenes

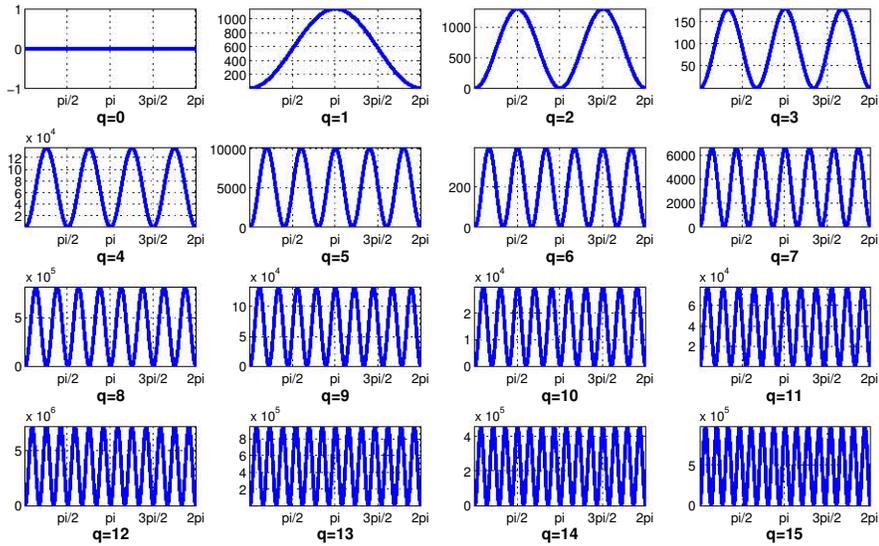


Figura 4.2.2.: Cosenos formados al variar el valor de q , imagen 0 radianes e imagen 0 radianes.

A continuación en la Figura 4.2.3 y en la Figura 4.2.4 donde ahora las dos imágenes a comparar corresponden a la de referencia (0 radianes), y otra con un ángulo de giro de $\frac{\pi}{2}$. Nuevamente la componente del coseno para $q = 12$ es la que tiene mayor magnitud.

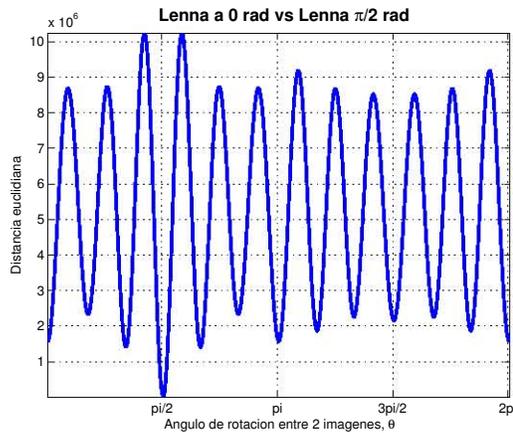


Figura 4.2.3.: Coseno generado por la suma de los cosenos de la Gráfica Figura 4.2.4

4. Evaluación de algoritmos

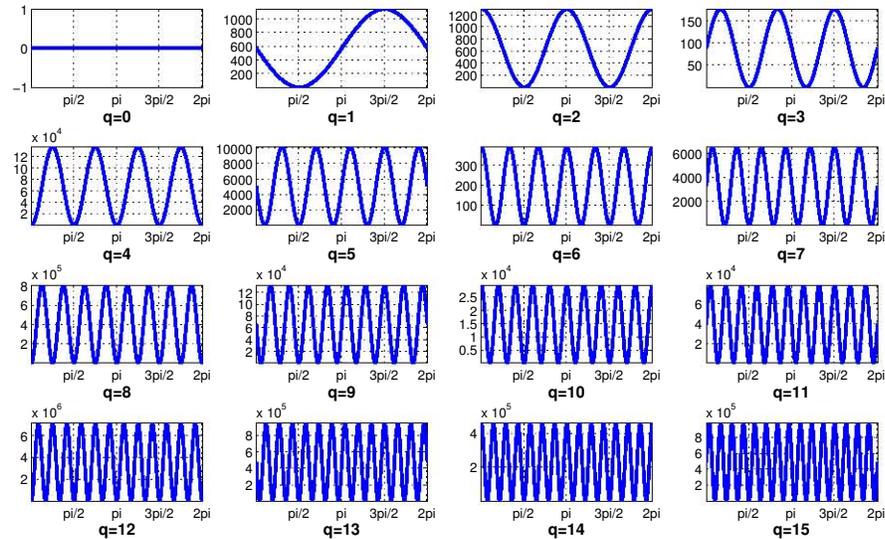


Figura 4.2.4.: Cosenos formados al variar el valor de q , imagen 0 radianes e imagen $\frac{\pi}{2}$ radianes.

La gráfica de la Figura 4.2.5 muestra las pruebas realizadas entre una imagen y su versión desfasada de π radianes y $\frac{3\pi}{2}$ radianes.

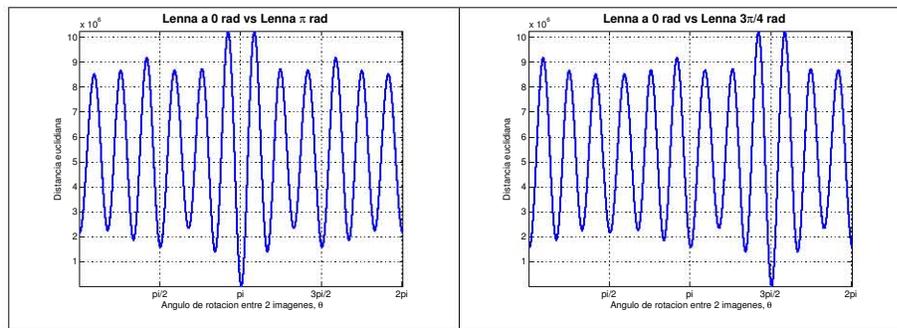


Figura 4.2.5.: Cambio de distancia euclidiana con respecto a ϕ , en base a Ecuación 3.3.14. Izquierda imagen rotada 0 radianes contra imagen rotada a π . Derecha imagen rotada 0 radianes contra imagen rotada a $\frac{3\pi}{4}$.

4.2. Invariancia respecto al ángulo de rotación entre dos imágenes

En la Tabla 4.9 se resumen los valores obtenidos al realizar esta prueba.

ϕ	dis. euc. mín.	grd. calculados
0	$-8,3494 \times 10^{-6}$	0
$\frac{\pi}{2}$	$9,6798 \times 10^3$	89.9998
π	293,1766	180.0001
$\frac{3\pi}{2}$	$9,5808 \times 10^3$	269.9998

Tabla 4.9.: Diferentes rotaciones.

4.2.2. Prueba con diferentes rotaciones y escalas

Para las pruebas de esta sección se utilizan dos grupos de imágenes rotadas de diferentes dimensiones, las mismas imágenes utilizadas en la subsección 4.1.6. Primero se realiza la prueba sin invariancia en escala, las gráficas resultantes son similares a las de la subsección 4.2.1 y los resultados son satisfactorios. Sin embargo al repetir la prueba con invariancia en escala los resultados son modificados ligeramente, teniendo un error promedio de dos grados. En la Figura 4.2.6 se tienen gráficas cuya forma es diferente puesto que la curva con mayor magnitud es la correspondiente al parámetro $q = 2$.

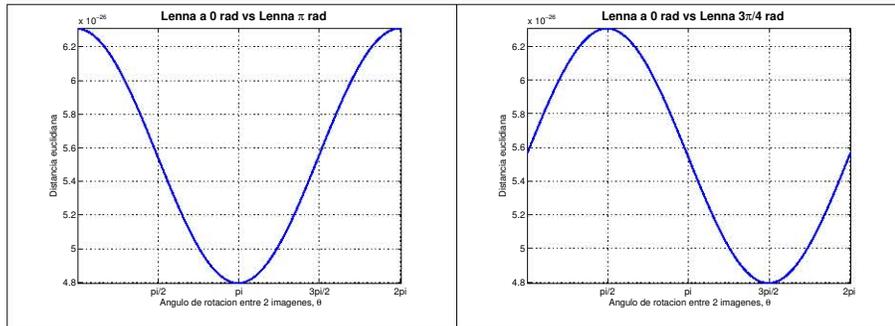


Figura 4.2.6.: Cambio de distancia euclidiana con respecto a ϕ , en base a (Ecuación 3.3.14).

Izquierda imagen rotada $\phi = 0$ radianes contra imagen rotada $\phi = \pi$ radianes.

Derecha imagen rotada $\phi = 0$ radianes contra imagen rotada a $\phi = \frac{3\pi}{4}$ radianes.

En la Tabla 4.10 se resumen los valores recabados en esta prueba.

Cabe señalar que en la Tabla 4.11 se observa que los valores mínimos son muy similares entre sí y muy pequeños en comparación con los registrados en la Tabla 4.10.

Es importante resaltar que los valores de la distancia euclidiana obtenidos usando una imagen pequeña como referencia son del mismo orden de magnitud que los valores de distancia euclidiana que resultan de utilizar una imagen de referencia grande. Y esto sucede

4. Evaluación de algoritmos

Ref. imagen grande vs pequeñas			Ref. imagen pequeña vs grandes		
ϕ	dis. euc. mín.	grd. calculados	ϕ	dis. euc. mín.	grd. calculados
0	$2,4618 \times 10^7$	0.0103	0	$2,4618 \times 10^7$	359.9897
$\frac{\pi}{2}$	$2,3877 \times 10^7$	90.0101	$\frac{\pi}{2}$	$2,4629 \times 10^7$	89.9893
π	$2,4733 \times 10^7$	180.0104	π	$2,4531 \times 10^7$	179.9897
$\frac{3\pi}{2}$	$2,3881 \times 10^7$	270.0101	$\frac{3\pi}{2}$	$2,4587 \times 10^7$	269.9891

Tabla 4.10.: Diferentes rotaciones, sin invariancia a escala.

Ref. imagen grande vs pequeñas			Ref. imagen pequeña vs grandes		
ϕ	dis. euc. mín.	grd. calculados	ϕ	dis. euc. mín.	grd. calculados
0	$7,1982 \times 10^{-39}$	358.3193	0	$7,1982 \times 10^{-39}$	1.6807
$\frac{\pi}{2}$	$7,1970 \times 10^{-39}$	88.3185	$\frac{\pi}{2}$	$7,1981 \times 10^{-39}$	91.6768
π	$7,1992 \times 10^{-39}$	178.3190	π	$7,1981 \times 10^{-39}$	181.6807
$\frac{3\pi}{2}$	$7,1971 \times 10^{-39}$	268.3200	$\frac{3\pi}{2}$	$7,1981 \times 10^{-39}$	271.6761

Tabla 4.11.: Diferentes rotaciones, con invariancia a escala.

en ambos casos, sin y con invariancia a la escala. Indicando que en este caso la diferencia en escala no produce cambios en los resultados.

4.2.3. Prueba con diferentes rotaciones y traslaciones

Se hicieron pruebas donde las imágenes se encuentran rotadas y trasladadas, como las imágenes de la subsección 4.1.5. En este caso la referencia es la Imagen 4.1.2a. Se obtienen las gráficas de la Figura 4.2.7 donde no hay invariancia en traslación, en la que los resultados aunque son cercanos a los valores esperados, no son satisfactorios. Al repetir la prueba pero ahora con invariancia en traslación se tienen las gráficas de la Figura 4.2.8 donde a pesar de que se observa gráficas más homogéneas los resultados tampoco son satisfactorios, de hecho se alejan de los valores reales esperados. Esto se puede atribuir al algoritmo de invariancia a traslación que son sometidos los momentos.

4.2. Invariancia respecto al ángulo de rotación entre dos imágenes

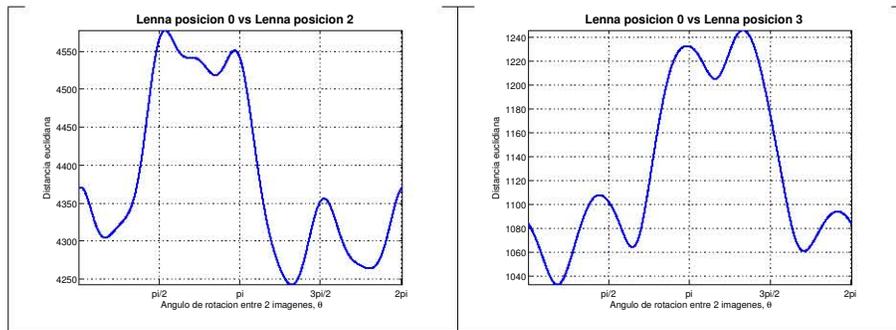


Figura 4.2.7.: Cambio de distancia euclidiana con respecto a ϕ , en base a (Ecuación 3.3.14). Izquierda imagen centrada contra imagen en posición 2. Derecha imagen centrada contra imagen en posición 3.

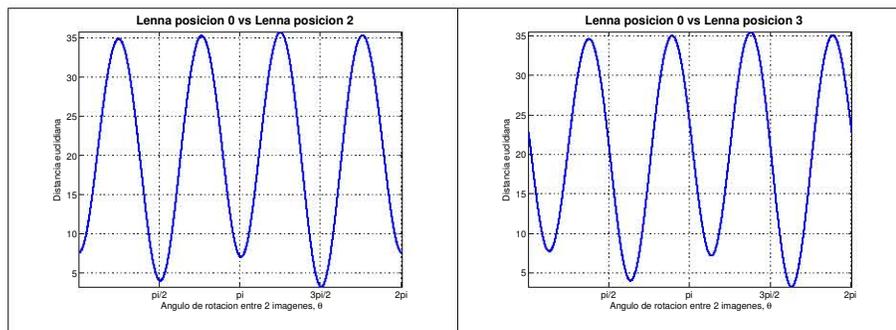


Figura 4.2.8.: Cambio de distancia euclidiana con respecto a ϕ , en base a Ecuación 3.3.14. Izquierda imagen centrada contra imagen en posición 2. Derecha imagen centrada contra imagen en posición 3.

En seguida se tabulan en la Tabla 4.12 y la Tabla 4.13 los resultados obtenidos con el algoritmo de búsqueda del mínimo. Al comparar los valores asentados en la Tabla 4.12 con los correspondientes en las gráficas de la Figura 4.2.7 se constata que estos valores son muy diferentes, esto se debe a que el algoritmo de búsqueda del mínimo supone curvas cosenos, y este no es el caso en la gráfica de la Figura 4.2.7.

En la Tabla 4.13 los valores de los ángulos obtenidos son más próximos a lo esperado, aunque no exactos.

4.2.4. Prueba con diferentes rotaciones, escalas y traslaciones

La última prueba se lleva a cabo con las imágenes que sufren las tres transformaciones simultáneamente, corresponden a las imágenes de la subsección 4.1.7, con las cuales se obtienen las gráficas de la Figura 4.2.9. Se observa que la gráfica sin invariancia en traslación

4. Evaluación de algoritmos

Ref. imagen centrada vs trasladadas		
ϕ	dis. euc. mín.	grd. calculados
π	$4,7764 \times 10^4$	319.7437
$\frac{3\pi}{2}$	$4,7811 \times 10^4$	49.7417
$\frac{5\pi}{3}$	58.0982	140.8749
$\frac{2\pi}{9}$	79.2225	267.1628

Tabla 4.12.: Diferentes rotaciones y traslaciones , sin invariancia a traslación.

Ref. imagen centrada vs trasladadas		
ϕ	dis. euc. mín.	grd. calculados
π	13.6249	182.0020
$\frac{3\pi}{2}$	13.6385	272.0051
$\frac{5\pi}{3}$	13.6708	294.9996
$\frac{2\pi}{9}$	13.7561	223.6055

Tabla 4.13.: Diferentes rotaciones y traslaciones , con invariancia a traslación.

y escala sufre deformaciones ocasionadas por las traslaciones de las imágenes, y el valor mínimo no se acerca al valor esperado. Mientras que en la gráfica en la que se utiliza los MZ invariantes en traslación y escala se consigue suavizar la curva, pero sin lograr una mejora en cuanto a encontrar el valor del ángulo de desfaseamiento entre las dos imágenes en cuestión.

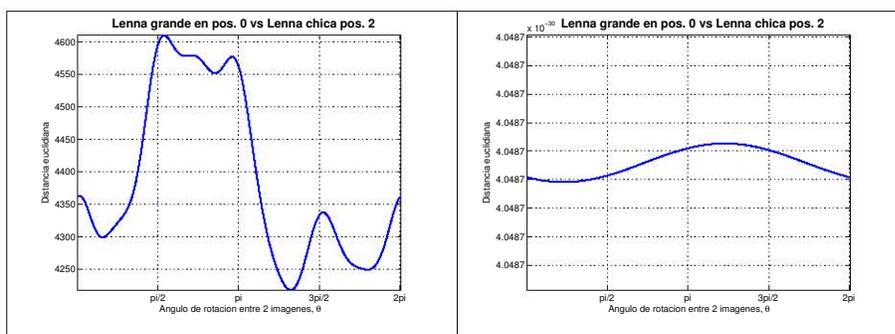


Figura 4.2.9.: Cambio de distancia euclidiana con respecto a ϕ , en base a (Ecuación 3.3.14). Izquierda imagen grande en posición 0 contra imagen chica en posición 2, sin invariancia en traslación, ni escala. Derecha imagen grande en posición 0 contra imagen chica en posición 2, con invariancia en traslación y escala.

Para apreciar la diferencia numérica entre el valor real del ángulo y el que se calcula con este algoritmo, se concentran los valores calculados en la Tabla 4.14 y la Tabla 4.15.

4.2. Invariancia respecto al ángulo de rotación entre dos imágenes

Ref. imagen grande vs pequeñas			Ref. imagen pequeña vs grandes		
ϕ	dis. euc. mín.	grd. calculados	ϕ	dis. euc. mín.	grd. calculados
π	$4,7452 \times 10^4$	319.8330	0	$1,1343 \times 10^3$	221.8236
$\frac{3\pi}{2}$	$4,7499 \times 10^4$	49.8311	$\frac{3\pi}{2}$	$1,1836 \times 10^8$	290.8050
$\frac{5\pi}{3}$	61.7772	140.9222	π	$6,8043 \times 10^9$	116.0525
$\frac{2\pi}{9}$	81.4122	267.3181	$\frac{\pi}{2}$	$2,9356 \times 10^4$	311.2743

Tabla 4.14.: Diferentes rotaciones, traslaciones y escala, sin invariancia a escala.

Ref. imagen grande vs pequeñas			Ref. imagen pequeña vs grandes		
ϕ	dis. euc. mín.	grd. calculados	ϕ	dis. euc. mín.	grd. calculados
π	$6,0792 \times 10^{-43}$	40.0251	0	$1,3300 \times 10^{-40}$	223.9193
$\frac{3\pi}{2}$	$6,0792 \times 10^{-43}$	34.0389	$\frac{3\pi}{2}$	$1,3300 \times 10^{-40}$	217.6649
$\frac{5\pi}{3}$	$6,0792 \times 10^{-43}$	230.7972	π	$1,3300 \times 10^{-40}$	42.6651
$\frac{2\pi}{9}$	$6,0792 \times 10^{-43}$	33.1793	$\frac{\pi}{2}$	$1,3300 \times 10^{-40}$	202.4414

Tabla 4.15.: Diferentes rotaciones, traslaciones y escala, con invariancia a escala.

5. Conclusión

Respecto a los resultados obtenidos en la sección 4.1, se concluye en primer lugar que la invariancia que ofrecen los momentos geométricos no es numéricamente exacta, teniendo un rango de variación que depende del método de invariancia utilizado e implícitamente de la transformación respecto a la cual se vuelve invariante. Comparando los casos “sin invariancia” con respecto a aquellos “con invariancia” referidos en las Gráficas 4.1.4a y 4.1.4b, en el caso de la invariancia en escala obtenida con el parámetro λ los valores de las diferencias promedio se reducen en un orden de magnitud de 9, mientras que las diferencias promedio de invariancia en traslación referidas en las Gráficas 4.1.3a y 4.1.3b tienen una reducción de 5 ordenes de magnitud, y la invariancia a la rotación genera cambios en un orden decreciente de magnitud de 4, con referencia a las Gráficas 4.1.6a y 4.1.6b.

En las pruebas en donde la imagen sufre mas de una transformación, se observa que cuando los MZ se modifican para obtener la invariancia en traslación y escala, Gráficas 4.1.8a y 4.1.8b, se tiene una reducción en un orden de magnitud de 12. Cuando se combina la invariancia en traslación y rotación, Gráficas 4.1.10a y 4.1.10b, el cambio obtenido es de una reducción de 1 ordenes de magnitud. Cuando se combinó la invariancia en escala y rotación, Gráficas 4.1.11a y 4.1.11b, se logró una reducción de 12 ordenes de magnitud. En el caso donde intervienen las tres transformaciones: traslación, escala y rotación, Gráficas 4.1.13a y 4.1.13b, se obtiene una reducción de 10 ordenes. De aquí se concluye que las invariancias son acumulativas en la mayoría de los casos, es decir que el decrecimiento en ordenes de magnitud aumenta cuando se consideran simultáneamente más tipos de invariancias en el cálculo de los momentos.

En la sección 4.2 se realizaron pruebas para estudiar el desempeño del algoritmo que se utiliza para encontrar la distancia mínima entre los MZ de dos imágenes, además dicho algoritmo también proporciona el ángulo de rotación respecto al cual ambas imágenes son lo más parecidas en términos de la distancia euclidiana. Los resultados de las pruebas muestran que este algoritmo aporta resultados precisos, siempre y cuando las imágenes en cuestión no hayan sufrido transformaciones de traslación y/o escala.

En la Subsección 4.2.2 las transformaciones en escala no afectan significativamente la precisión de los resultados del ángulo de giro medido. Pero, cuando las imágenes se someten al proceso de invariancia en escala, el ángulo de giro que se obtiene sufre cambios en la precisión de aproximadamente ± 2 radianes. Sin embargo, los valores mínimos de la distancia euclidiana entre la imagen de referencia y las imágenes rotadas se reducen en un orden de magnitud de 46 respecto a los valores obtenidos sin utilizar la invariancia en escala.

5. Conclusión

De la Subsección 4.2.3 se observa que las transformaciones en traslación afectan el cálculo de los ángulos de giro, obteniendo resultados erróneos. Afortunadamente al utilizar los MZ invariantes en traslación se consiguen resultados relativamente precisos con un error de ± 2 grados.

Finalmente cuando una imagen es afectada por transformaciones en escala, traslación y rotación los valores de los ángulos de giro no son correctos. Esto ocurre tanto empleando MZ invariantes como no invariantes, en escala y traslación.

Los procesos de invariancia que modifican las curvas que generan los momentos, ocasionan que el algoritmo diseñado para la búsqueda de la distancia euclidiana mínima dada en [3] no opere correctamente. Concluyendo que los algoritmos de invariancia propuestos en este trabajo afectan la forma de las curvas, por lo que no se recomienda su uso para el cálculo de la distancia euclidiana mínima requerida para obtener el ángulo de giro.

Por lo que de acuerdo con estos resultados (Subsección 4.2.2) es factible trabajar con imágenes de diferentes escalas (sin invariancia a la escala), y aún así obtener valores de ángulo de giro aceptables. Otra posibilidad de lidiar con problemas de escala es utilizar el momento de orden cero propuesto en [11] para proporcionar invariancia en escala, debido a que este algoritmo es una división entre un escalar y no afecta la forma de las curvas de los momentos.

Por último para evitar utilizar la invariancia en traslación es aconsejable centrar la imagen previamente al cálculo de sus momentos.

Trabajo futuro

La elección de trabajar con los MZ se justifica argumentando su extenso y satisfactorio uso en otros trabajos del área de análisis de imágenes. Sin embargo, existen otros grupos de polinomios ortogonales que prometen ser mejores como por ejemplo los de Fourier-Mellin. Dado que el algoritmo rápido que se describe en el Capítulo 4 se puede utilizar para cualquier conjunto de momentos que empleen un núcleo ortogonal, se propone modificar e implementar el algoritmo rápido dado en [16] con los momentos de Fourier-Mellin. Así como realizar un análisis similar al desarrollado en [3] para encontrar el ángulo de giro entre dos imágenes.

Se propone también caracterizar exhaustivamente los momentos seleccionados, empleando diferentes tipos de imágenes, imágenes sintéticas y reales, con problemas de oclusión, con ruido, etc.

Por último diseñar un clasificador que explote la información extraída por los momentos, tal como: escala, traslación y rotación. Tomando en cuenta en esta clasificación que los primeros momentos contienen la mayor parte de información de la imagen, para evitar hacer cálculos innecesarios de todos los momentos de una imagen. Es decir, que con solo unos cuantos momentos poder decidir si la imagen a identificar pertenece o no a un cierto grupo. Y utilizar los momentos de orden alto para ubicar a la imagen dentro de un subgrupo.

Por ejemplo, si con los momentos de orden bajo se decidió que la imagen pertenece a un cierto objeto, es tarea de los momentos de orden alto decidir si ese objeto tiene ciertas propiedades geométricas particulares.

A.

A.1. Momentos de Hu invariantes.

Conjunto de siete momentos invariantes a la traslación, rotación y escala. Derivados de los momentos centrales normalizados de orden dos y tres [10].

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

A.2. Momentos invariantes de Zernike

Estos momentos de Zernike invariantes, son invariantes a traslación, escala y rotación. Los primeros cinco están expresados en términos de los momentos de Hu [21]. Donde

cc: complejo conjugado del término anterior

S_n : momentos invariantes de Zernike

A_{nl} : momentos de Zernike

μ_{mn} : momentos centrales normalizados de Hu.

Segundo orden

$$\begin{aligned}S_1 &= A_{20} = \frac{3}{\pi} [2(\mu_{20} - \mu_{02}) - 1] \\ S_2 &= |A_{22}|^2 = \frac{9}{\pi^2} [(\mu_{20} - \mu_{02})^2 + 4\mu_{11}^2]\end{aligned}$$

Tercer orden

A.

$$\begin{aligned}
S_3 &= |A_{33}|^2 = \frac{16}{\pi^2} \left[(\mu_{03} - 3\mu_{21})^2 + (\mu_{30} - 3\mu_{12})^2 \right] \\
S_4 &= |A_{31}|^2 = \frac{144}{\pi^2} \left[(\mu_{03} + \mu_{21})^2 + (\mu_{30} + \mu_{12})^2 \right] \\
S_5 &= A_{33}^* A_{31}^2 + cc = \frac{13824}{\pi^4} \left\{ (\mu_{03} - 3\mu_{21})(\mu_{03} + \mu_{21}) \right. \\
&\quad \times \left[(\mu_{03} + \mu_{21})^2 - 3(\mu_{30} + \mu_{12})^2 \right] \\
&\quad \left. - (\mu_{03} - 3\mu_{21})(\mu_{03} + \mu_{21}) \right. \\
&\quad \left. \times \left[(\mu_{30} + \mu_{12})^2 - 3(\mu_{03} + \mu_{21})^2 \right] \right\} \\
S_6 &= A_{31}^2 A_{22}^* + cc = \frac{864}{\pi^3} \left\{ (\mu_{02} - \mu_{20}) \left[(\mu_{03} + \mu_{21})^2 - (\mu_{30} + \mu_{12})^2 \right] \right. \\
&\quad \left. + 4\mu_{11}(\mu_{03} + \mu_{21})(\mu_{30} + \mu_{12}) \right\}
\end{aligned}$$

Cuarto orden

$$\begin{aligned}
S_7 &= |A_{44}|^2 = \frac{25}{\pi^2} \left[(\mu_{40} - 6\mu_{22} + \mu_{04})^2 - 16(\mu_{31} - \mu_{13})^2 \right] \\
S_8 &= |A_{42}|^2 = \frac{25}{\pi^2} \left\{ (\mu_{04} - \mu_{40}) + 3(\mu_{20} - \mu_{02})^2 \right. \\
&\quad \left. + 4[4(\mu_{31} + \mu_{13}) - 3\mu_{11}]^2 \right\} \\
S_9 &= A_{40} = \frac{5}{\pi} [6(\mu_{40} + 2\mu_{22} + \mu_{04}) - 6(\mu_{20} + \mu_{02}) + 1] \\
S_{10} &= A_{44} A_{42}^2 + cc = \frac{250}{\pi^3} \left\{ [(\mu_{40} - 6\mu_{22} + \mu_{04})] \{ [4(\mu_{04} - \mu_{40}) \right. \right. \\
&\quad \left. \left. + 3(\mu_{20} - \mu_{02})^2 - 4[4(\mu_{31} + \mu_{13}) - 3\mu_{11}]^2 \right\} \right. \\
&\quad \left. - 16[4(\mu_{04} - \mu_{40}) + 3(\mu_{20} - \mu_{02})] \right. \\
&\quad \left. [4(\mu_{31} + \mu_{13}) - 3\mu_{11}](\mu_{31} + \mu_{13}) \right\} \\
S_{11} &= A_{42} A_{22}^* + cc = \frac{30}{\pi^2} \left\{ [4(\mu_{40} - \mu_{04}) + 3(\mu_{20} - \mu_{02})] (\mu_{02} - \mu_{20}) \right. \\
&\quad \left. 4\mu_{11} [4(\mu_{31} + \mu_{13}) - 3\mu_{11}] \right\}
\end{aligned}$$

$$\begin{aligned}
S_{12} &= |A_{55}|^2 \\
S_{13} &= |A_{53}|^2 \\
S_{14} &= |A_{51}|^2 \\
S_{15} &= A_{51}^* A_{31}^2 + cc \\
S_{16} &= A_{53}^* A_{33}^2 + cc \\
S_{17} &= A_{55}^* A_{31} + cc
\end{aligned}$$

A partir del sexto orden, se consideran momentos de Zernike de alto orden.

Sexto orden

$$\begin{aligned}
S_{18} &= |A_{66}|^2 \\
S_{19} &= |A_{64}|^2 \\
S_{20} &= |A_{62}|^2 \\
S_{21} &= A_{60} \\
S_{22} &= A_{66}^* A_{33}^2 + cc \\
S_{23} &= A_{64}^* A_{44} + cc \\
S_{24} &= A_{62}^* A_{22} + cc
\end{aligned}$$

Séptimo orden

A.2. Momentos invariantes de Zernike

$$\begin{aligned} S_{25} &= |A_{77}|^2 \\ S_{26} &= |A_{75}|^2 \\ S_{27} &= |A_{73}|^2 \\ S_{28} &= |A_{71}|^2 \\ S_{29} &= A_{77}^* A_{31}^7 + cc \\ S_{30} &= A_{75}^* A_{55} + cc \\ S_{31} &= A_{73}^* A_{33} + cc \\ S_{32} &= A_{71}^* A_{31} + cc \\ \text{Octavo orden} \\ S_{33} &= |A_{88}|^2 \\ S_{34} &= |A_{86}|^2 \\ S_{35} &= |A_{84}|^2 \\ S_{36} &= |A_{82}|^2 \\ S_{37} &= A_{80} \\ S_{38} &= A_{88}^* A_{44}^2 + cc \\ S_{39} &= A_{86}^* A_{66} + cc \\ S_{40} &= A_{84}^* A_{44} + cc \\ S_{41} &= A_{82}^* A_{22} + cc \end{aligned}$$

B.

El algoritmo de calculo de MZ que a continuación se describe esta basado en el propuesto en [16].

B.1. Algoritmo de cálculo de matriz de MZ

El conjunto de MZ que caracteriza una determinada imagen se distribuyen dentro de una matriz. El algoritmo que construye dicha matriz se describe a continuación con la ayuda del diagrama de flujo de la Figura B.1.1.

Como entrada recibe la imagen denotada con I , y los valores de P y Q que determinan el grado de los MZ. Con estos datos y la ayuda de otro programa se obtiene H_p y H_q , que son el esqueleto de la matriz de los MZ. La matriz que se va a calcular es de $P + 1 \times Q + 1$ dimensiones, el elemento extra corresponde al momento cero.

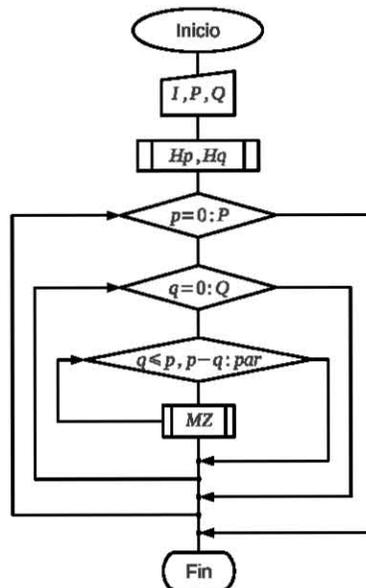


Figura B.1.1.: Diagrama de flujo de matriz de momentos

B.

Dentro de dos ciclos anidados se calcula la matriz de MZ, un condicional previa que restringe el acceso a algunas celdas de la matriz. Esta restricción es parte del algoritmo de cálculo de MZ que a su vez hace que se efectúen solo una cuarta parte de la matriz, es decir $\frac{(P+1)(Q+1)}{4}$ cálculos. El programa que hace el cálculo de cada uno de los MZ se explica más adelante.

B.2. Algoritmo de cálculo de matriz de H's

Los valores de H_p y H_q pueden ser utilizados para varias imágenes siempre y cuando compartan las mismas dimensiones y el orden requerido sea igual o menor. Esto se debe gracias a que este algoritmo trabaja sólo con las dimensiones de la imagen I y el orden de los MZ, P y Q .

En el diagrama de flujo de la Figura B.2.1 se ilustra el cálculo de las variables x y y , las cuales a partir de ese punto representan las nuevas coordenadas contenidas en dos vectores con valores que van de -1 a 1 , tanto en la ordenada como en la abscisa.

El siguiente bloque permite la elección de centrar o no la imagen, cabe aclarar que esta opción y la de escalado no son parte del algoritmo propuesto en [16], son mejoras realizadas a partir de este trabajo. La opción de centrado proporciona el valor del centroide en las variables Y_h y X_h , en caso que se requieran momentos invariantes en traslación. De lo contrario se les asigna el valor de cero a ambas variables.

B.2. Algoritmo de cálculo de matriz de H's

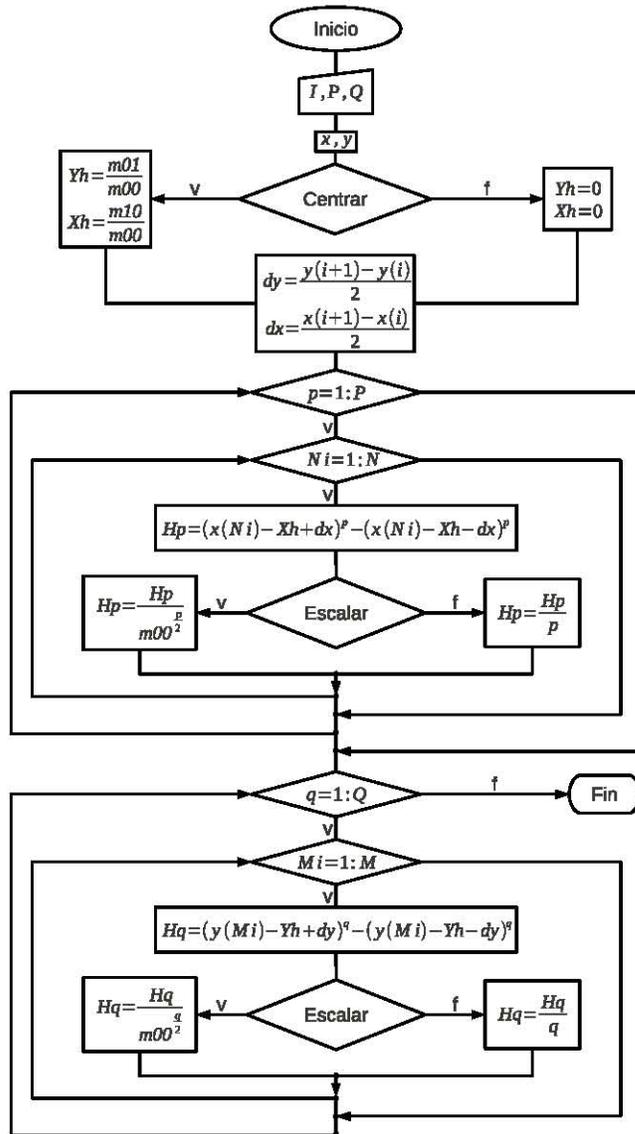


Figura B.2.1.: Diagrama de flujo de matriz de H

Las siguientes variables que se obtienen son dy y dx , que son los límites de integración utilizados para una exacta estimación de los valores de H_p y H_q . Son la mitad de la diferencia de dos valores de coordenadas consecutivos.

A continuación se tienen dos ciclos anidados, el primero corre de la unidad al valor de P , el segundo también inicia con valor uno hasta N que indica la longitud del vector x . En

B.

el núcleo de estos ciclos se ejecuta el cálculo de H_p y posteriormente se elige o no hacer invariantes a escala los MZ por medio del uso de λ .

Lo siguiente es determinar H_q , su algoritmo es semejante al que se utilizó con H_p , solo que ahora se utiliza M para indicar la longitud del vector que indica una de las dimensiones de la imagen, en este caso es y . Este algoritmo concluye cuando $q = Q$, valor que complementa el grado de los MZ a calcular.

B.3. Algoritmo de cálculo de MZ

La siguiente descripción corresponde al algoritmo de evaluación de un solo momento MZ, representado con el diagrama de flujo de la Figura B.3.1. Este es utilizado sucesivamente dentro del diagrama de flujo de la Figura B.1.1. La información de entrada que alimenta este algoritmo es la imagen I , P y Q que determinan el grado de los MZ y las matrices H_p y H_q .

La primera condicional determina el valor de w , quien a su vez define si el valor complejo del MZ es positivo o negativo. El siguiente es el ciclo principal dentro del cual se ejecuta el cálculo del MZ, comienza con el cálculo de $B_p q k$ seguido por cuatro ciclos anidados que como variable medular tienen a $m p q$, valor que involucra directamente a la imagen y las matrices H_p y H_q . Los índices de las ordenadas de las matrices H_p y H_q se calculan dos ciclos antes con el fin de no calcularlos más de las veces necesarias, mientras que los índices j e i recorren la ubicación de todos los píxeles de la imagen.

Al terminar de recorrer toda la imagen se hace una sumatoria z del producto entre $m p q$, w y los valores de dos combinaciones. Finalmente fuera de los cuatro ciclos anidados se realiza otra sumatoria del producto entre $B_p q k$ y z .

B.3. Algoritmo de cálculo de MZ

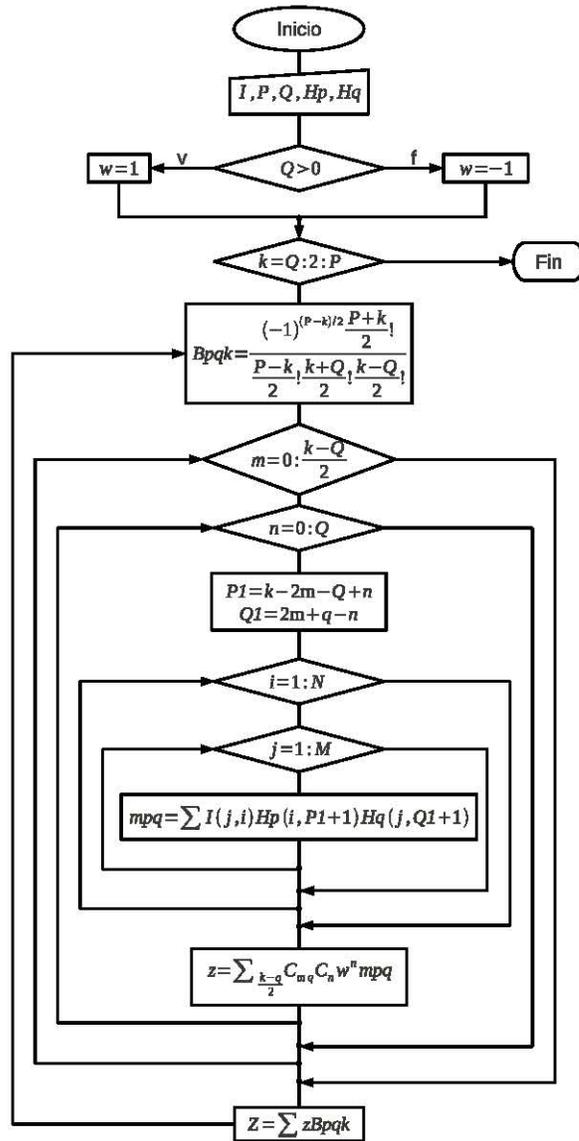


Figura B.3.1.: Diagrama de flujo de MZ

C.

En este apéndice se describe el algoritmo basado en [3], el cual a partir de la distancia euclidiana mínima puede indicar cuál es el par de imágenes más parecidas dentro de un conjunto de varias imágenes, además de encontrar el ángulo de desfase que hay entre ellas.

C.1. Algoritmo de búsqueda de distancia mínima entre dos imágenes

La idea sobre la que se basa este algoritmo es obtener la distancia euclidiana entre los momentos MZ de dos imágenes, si esta distancia es lo suficientemente pequeña se dice que ambas imágenes son similares. La novedad en este algoritmo es que obtiene diferentes distancias euclidianas debido a que es posible cambiar el ángulo de rotación que tiene una imagen con respecto a la otra.

Como argumento de entrada este algoritmo recibe dos imágenes I y J . El valor de entrada Q es utilizado para obtener el grado máximo de los momentos MZ, en este caso es $2Q$ porque P tiene el mismo valor que Q . El cuarto argumento ARG se utiliza para indicar, ya sea si se desea centrar y/o escalar los momentos.

En el primer bloque de cálculo del diagrama de flujo de la Figura C.1.1 se indica que el inicio del algoritmo consiste en la extracción de los momentos MZ de las imágenes I y J , sean ZI y ZJ , así como de las variables rad y aux . La variable rad se emplea para calcular la amplitud con la que cambia el ángulo que relaciona los momentos MZ de ambas imágenes, esta operación puede verse como si a una de las imágenes se girara, permaneciendo la otra fija, y se calcula la distancia euclidiana entre sus momentos MZ. La otra variable aux se utiliza más adelante como índice para el vector B .

Enseguida se tiene un ciclo que recorre los 360° a partir de cero. Dejando como límite $4Q$ al sustituir el valor mínimo y máximo en xi , se pone en evidencia que los límites van de cero a 2π . El proceso que envuelve este ciclo es el cálculo de los vectores xi y Fd . El primero es simplemente un vector que contiene los valores de los ángulos con los que se calcularon las diferentes distancias euclidianas de los momentos. Mientras que Fd contiene los valores de la derivada de la función que describe el comportamiento de la distancia euclidiana de los MZ entre ambas imágenes.

La razón por la que se calcula un índice por adelantado es porque el siguiente bloque de proceso así lo requiere; de aquí que a la condicional se le reste una unidad.

C.

El siguiente bloque es una condicional, su objetivo es buscar los puntos de inflexión, específicamente los puntos mínimos de la función Fd . Los valores que cruzan esta condicional son utilizados para ejecutar el algoritmo de gradiente descendiente y construir un nuevo vector B .

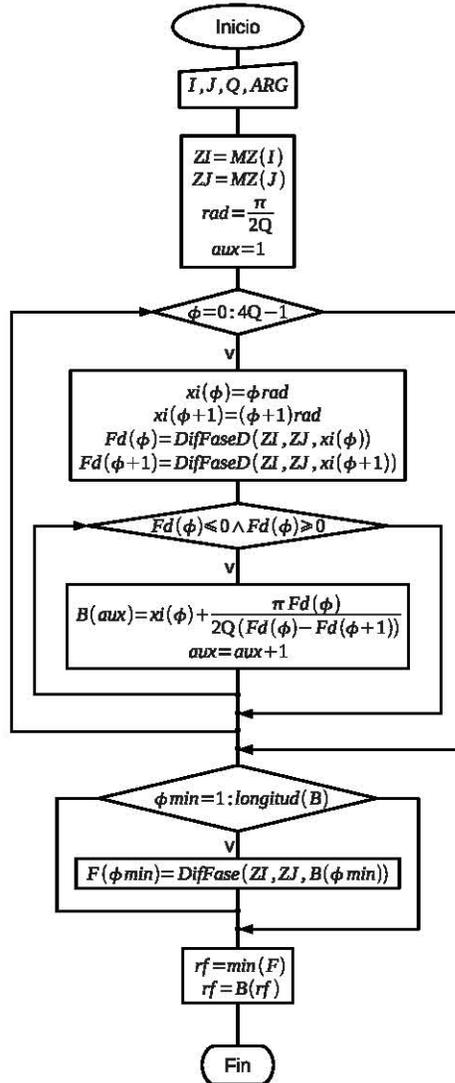


Figura C.1.1.: Diagrama de flujo de búsqueda de distancia mínima

Una vez finalizada la recopilación de los valores de B , son retomados para evaluarse en

la función F que describe el comportamiento de la distancia euclidiana de los MZ entre ambas imágenes. Y finalmente el valor mínimo de F indica el valor en el que las dos imágenes presentaron la distancia euclidiana más mínima. Para averiguar el ángulo en el que se presenta esta distancia mínima basta con utilizar la posición del valor mínimo dentro del vector F en B , que es un vector con los valores de los ángulos de desfase correspondientes a los valores mínimos de F .

C.2. Algoritmo de cálculo de función F y Fd

En el diagrama de flujo de la Figura C.1.1 se hace uso de una función F , así como de su derivada, que describe el comportamiento de la distancia euclidiana de los MZ entre dos imágenes. Con la ayuda del diagrama de flujo de la Figura C.2.1 se describe el comportamiento de esta función.

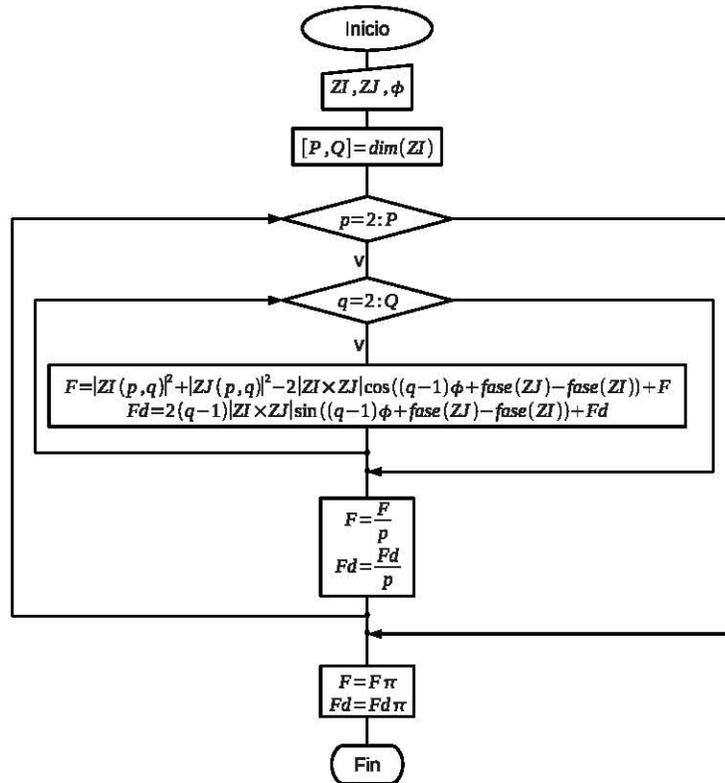


Figura C.2.1.: Diagrama de flujo de matriz de H

C.

Para iniciar el proceso se necesitan tres elementos de entrada para este algoritmo, esto son las matrices de momentos ZI y ZJ de las dos imágenes y un valor de ángulo ϕ . Esta información es procesada dentro de dos ciclos anidados que dependen de los valores de P y Q . Puesto que las matrices ZI y ZJ tienen como dimensiones los valores de P y Q , es factible extraer estos datos, adicional al esto $P = Q$, entonces con sólo saber una de las dimensiones de una de las matrices de MZ de las imágenes se conocen los valores de P y Q .

La ecuación que se efectúa en el núcleo de los dos ciclos anidados es una sumatoria desarrollada en [3] construida en base a la distancia euclidiana existente entre ZI y ZJ . Gracias a la estructura de los polinomios de Zernike es posible manipular el ángulo que relaciona los dos conjuntos de momentos y obtener una serie de distancias euclidianas correspondientes a ángulos conocidos.

Debido a que las matrices ZI y ZJ incluyen el momento cero, los ciclos inician desde el valor dos, que es la ubicación del momento uno dentro de las matrices. Los procesos que se ejecutan fuera de el núcleo del los ciclos anidados ayudan a hacer más eficiente el algoritmo, ya que se ejecutan menos número de operaciones.

La salida de este algoritmo es el valor de F o Fd , dependiendo del caso. El valor de Fd se necesita para ejecutar el algoritmo de gradiente descendiente y el valor de F se usa cuando se tiene el conjunto de valores mínimos de distancias euclidianas y es necesario evaluarlos directamente en la función F para definir el valor final de la distancia euclidiana mínima entre los dos conjuntos de MZ .

Bibliografía

- [1] Luren Yang/Fritz Albrechtsen. Fast computation of invariant geometric moments: A new method giving correct results. *Proceedings of the 12th IAPR International Conference on Pattern Recognition, Conference A: Computer Vision & Image Processing*, 1:201–204, 1994. 10
- [2] Luren Yang/Fritz Albrechtsen. Fast and exact computation of Cartesian geometric moments using discrete Green’s theorem. *Pattern Recognition*, 29(7):1061–1073, 1996. 10
- [3] Revaud Jérôme/Lavoué Guillaume/Baskurt Atilla. Improving zernike moments comparison for optimal similarity and rotation angle retrieval. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 31(4):627–636, 2009. 1, 11, 35, 37, 38, 66, 79, 82
- [4] Sing-Tze Bow. *Pattern recognition and image preprocessing*. CRC, 2002. 14
- [5] François Chaumette. Image moments: a general and useful set of features for visual servoing. *Robotics, IEEE Transactions on*, 20(4):713–723, 2004. 11
- [6] Gerhard X. Ritter/Joseph N. Wilson/J. L. Davidson. Image algebra: An overview. *Computer Vision, Graphics, and Image Processing*, 49(3):297–331, 1990. 9
- [7] Rutovitz Denis. Pattern Recognition. *Journal of the Royal Statistical Society. Series A (General)*, 129:504–530, 1966. 14
- [8] Tomas Suk/Jan Flusser. Blur and affine moment invariants. *Pattern Recognition*, 4:339–342, 2002. 10
- [9] Alireza Khotanzad/Yaw Hua Hong. Invariant image recognition by Zernike moments. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 12(5):489–497, 1990. 43
- [10] Ming-Kuei Hu. Visual pattern recognition by moment invariants. *IRE Transactions on Information Theory*, 8(2):179–187, 1962. 27, 28, 29, 43, 69
- [11] Ye Bin/Peng Jia-Xiong. Invariance analysis of improved Zernike moments. *Journal of Optics A: Pure and Applied Optics*, 4:606–614, 2002. 44, 66

Bibliografía

- [12] Obagi Araujo Juan José. *Elementos de teoría de probabilidad para ingenieros*. Pontificia Universidad Javeriana, 2003. 25
- [13] Anant Choksuriwong/Bruno Emile/Christophe Rosenberger/Helene Laurent. Comparative study of global invariant descriptors for object recognition. *Journal of Electronic Imaging*, 17, 2008. 11
- [14] David G. Lowe. Object recognition from local scale-invariant features. 2:1150–1157, 1999. 10
- [15] Tang Yuan Yan/L. H. Yang/Jiming Liu/H. Ma. *Wavelet theory and its application to pattern recognition*, volume 36. World Scientific Pub Co Inc, 2000. 10
- [16] Chong-Yaw Wee/Raveendran Paramesran/Ramakrishnan Mukundan. Fast computation of geometric moments using a symmetric kernel. *Pattern Recognition*, 41(7):2369–2380, 2008. 1, 11, 37, 66, 73, 74
- [17] Jesús Ariel Carrasco Ochoa. Enfoque Lógico Combinatorio al Reconocimiento de Patronos, I. Selección de Variables y Clasificación Supervisada. *Computación y Sistemas*, 4(1):68–69. 14
- [18] Pallas Areny Ramón. *Sensores y acondicionadores de señal*. Marcombo, 1994. 7
- [19] Edmund T. Rolls. David Marr’s Vision: floreat computational neuroscience. *Brain*, 134(3):913–916, 2011. 10
- [20] Lin Wen-Gou/Wang Shuenn-shyang. A note on the calculation of moments. *Pattern recognition letters*, 15(11):1065–1070, 1994. 10
- [21] Michael Reed Teague. Image analysis via the general theory of moments*. *Journal of the optical society of America*, Vol.70(8):920–930, 1980. 9, 11, 30, 37, 69
- [22] Frits Zernike. Phase contrast, a new method for the microscopic observation of transparent objects part II. *Physica*, 9(10):974–986, 1942. 30