

**U**

**P**

**T**

# Reconocimiento de la marcha humana basado en la correlación de historias de momentos de Jacobi-Fourier

por

**César Joel Camacho Bello**

Tesis sometida como requisito parcial para obtener el grado de

**MAESTRO EN COMPUTACIÓN ÓPTICA**

en la

**Universidad Politécnica de Tulancingo**

Marzo 2011

Tulancingo, Hidalgo

Supervisada por:

**Dra. Carina Toxqui Quitl  
Dr. Alfonso Padilla Vivanco**

©UPT 2011

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



**UNIVERSIDAD  
POLITÉCNICA  
DE  
TULANCINGO**



# Dedicatoria

A mi madre

A mi hermano



# Agradecimientos



Al Proyecto FOMIX CONACyT - ESTADO DE HIDALGO 96792 bajo el título *“Clasificación mediante la técnica de momentos circulares para el control de calidad de piezas usando sistemas óptico digitales para la adquisición de imágenes multidistorsionadas”* por los recursos otorgados para la realización de esta Tesis.



**UNIVERSIDAD  
POLITÉCNICA  
DE  
TULANCINGO**

Al Laboratorio de Óptica y Visión por Computadora de la Universidad Politécnica de Tulancingo por las facilidades otorgadas.

A mis Asesores los Drs. Carina Toxqui Quitl y Alfonso Padilla Vivanco por haber dirigido este proyecto de Tesis y por permitirme colaborar en su grupo de trabajo en la UPT.

A mis sinodales los Drs. José Javier Báez Rojas, César Santiago Tepantlán y el M. C. José Francisco Solis Villareal por sus valiosas sugerencias que ayudaron a mejorar este trabajo.

Al Rector Dr. Luis Téllez Reyes y al Secretario Académico Dr. César Santiago Tepantlán por el apoyo brindado para el desarrollo de este trabajo de Tesis.

A la Dra. Rocío Ruiz de la Barrera, Secretaria de Educación Pública en el Estado de Hidalgo, por el apoyo brindado durante todo este tiempo a la Universidad Politécnica de Tulancingo

# Prefacio

Con la aparición de las primeras computadoras en la década de los cincuenta, inicia el análisis y tratamiento digital de imágenes. Como consecuencia de ello, surgen áreas como la Visión por Computadora, que tiene por objetivo modelar matemáticamente los procesos de percepción visual en los seres vivos y generar programas que permitan simular estas capacidades visuales por computadora.

Lo anterior, nos muestra un desarrollo y evolución de técnicas o algoritmos para el procesado de información. Además, avances en el análisis de métodos matemáticos para la representación de señales o funciones digitales. En la actualidad podemos encontrar sistemas de visión en múltiples entornos, como en la industria, donde se requiere de la inspección automática de productos o bien en el control de calidad de alimentos. Esto en función de su color, tamaño o textura. Algunas otras áreas que se benefician de estos sistemas de visión son la Astronomía, la Medicina, la Seguridad, las Ciencias Foreneces y la Biometría. Ésta última con el propósito de identificar personas a partir de imágenes del rostro, retina, huella dactilar, iris, rubrica o forma de caminar, por mencionar algunas.

En el laboratorio de Óptica y Visión por Computadora de la UPT, se implementó un sistema de visión para la adquisición de imágenes de 15 personas caminando bajo condiciones variables de iluminación y amplificación. Lo anterior, con el propósito de reconocerlas por su marcha o forma de caminar.

Este proyecto de Tesis, propone un método para la descripción y clasificación de la marcha humana basado en el análisis y procesamiento de una secuencia de  $N$  imágenes digitales. Una técnica de extracción de características usada es la de momentos circulares de una imagen digital. El núcleo

de ésta transformación está basada en los polinomios complejos de Jacobi-Fourier.

A partir de una serie de imágenes, se generan historias de momentos de Jacobi-Fourier, cuyo perfil depende de la forma de caminar de cada persona. Una forma de evaluar el algoritmo propuesto es aplicándolo sobre bases de datos estándar de personas que están caminando. Los datos obtenidos son comparados con otros métodos publicados en la bibliografía.

En el capítulo uno, se plantean los objetivos y aportaciones del trabajo realizado. En el capítulo dos, se hace una revisión de las funciones momento base del algoritmo propuesto; que va desde los momentos geométricos y centrales, hasta los Jacobi-Fourier. Asimismo se revisan algunos invariantes a la escala. En el capítulo tres, se describe el método propuesto para la clasificación de la marcha mediante la correlación de las historias de momentos, para así obtener un conjunto de descriptores de la marcha humana. También, se propone un mapa para localizar los mejores descriptores y un método para clasificarlos. En el capítulo cuatro se compara y prueba la capacidad de clasificación con las tres bases de datos obtenidas. En el capítulo cinco se presentan los resultados de la clasificación usando las distintas bases y se llevo acabo una comparación del funcionamiento del algoritmo sobre ellas. En el capítulo seis y último se presenta las conclusiones del trabajo. Finalmente, se describe el trabajo a futuro.

# Resumen

Se presenta un método para la descripción y clasificación de la marcha humana basado en la correlación temporal de funciones o historias momento de Jacobi Fourier. Cada historia de momentos describe la forma de un individuo en una secuencia de  $N$  imágenes. El núcleo de los momentos está compuesto por el factor de Fourier y la función Generadora de Jacobi. Variaciones en los parámetros de la función de Jacobi genera infinito número de familias de polinomios ortogonales. Con esto, un análisis del funcionamiento de diferentes momentos circulares es hecho. El enfasamiento automático de las historias momento es usada para la calibración temporal del sistema de adquisición de imágenes. Mapas para la selección de momentos son generados. Usando la distancia mínima como criterio de clasificación de  $M$  clases, el método ajusta sólo los descriptores seleccionados para obtener el máximo de clasificación. El método propuesto logra una clasificación completa con el menor número de descriptores, comparado con otros conjuntos como Zernike. Bases de datos estándar fueron usadas para comparar la efectividad del Algoritmo.



# Abstract

A new method to calculate the temporal correlation in a moment function representation of lower body images for gait recognition is shown. A Jacobi- Fourier moment history is obtained from  $N$ -frames to describe the history shape of individuals. The kernel of the moments is composed by the Fourier factor  $e^{il\theta}$  of order  $l$  and the Generic Jacobi function  $J_n(\alpha, \beta, r)$  of order  $n$ . Value variations for  $\alpha$  and  $\beta$  generate infinity number of orthogonal polynomial families. With this in mind a performance analysis of different circular moments is done. The automatic phasing of several moment histories is used for video temporal calibration. A discriminative method of moments by means some selective  $n - l$  maps is proposed. By using a minimum distance classifier  $M$ -classes, our method fitting only the selected features in order to obtain the highest correct classification rate for all sets tested. The Jacobi polynomials  $J_n(10, 10, r)$  achieve a full recognition percentage with the smallest number of descriptors compared with others as Zernike. The databases used here are MoBo, CASIA A, and our home database.



# Índice general

<b>1. Introducción</b>	<b>19</b>
1.1. Antecedentes . . . . .	22
1.2. Objetivos . . . . .	24
1.3. Estado del Arte del Reconocimiento de la Marcha . . . . .	25
1.4. Aportaciones . . . . .	26
<b>2. Revisión Matemática</b>	<b>27</b>
2.1. Momentos Geométricos de una Imagen . . . . .	28
2.1.1. Descriptores Invariantes . . . . .	31
2.2. Momentos de Jacobi-Fourier de una Imagen . . . . .	35
2.2.1. Descriptores Invariantes . . . . .	41
<b>3. Bases de datos y Sistemas de Adquisición de Imágenes Digitales para la marcha</b>	<b>43</b>
3.1. Adquisición de Imágenes (Base UPT) . . . . .	46
3.1.1. Preprocesamiento . . . . .	48
3.2. Base MoBo . . . . .	51
3.2.1. Preprocesamiento . . . . .	57
3.3. Base CASIA . . . . .	59
3.3.1. Preprocesamiento . . . . .	62
<b>4. Clasificación de la marcha a partir de la extracción de descriptores invariantes</b>	<b>65</b>
4.1. Historia de momentos . . . . .	66
4.2. Problema de poner en fase . . . . .	67
4.3. Descriptores de la marcha . . . . .	72
4.4. Mapa de la medida discriminatoria $Q$ . . . . .	74

4.5. Método de clasificación . . . . .	79
<b>5. Discusión de Resultados</b>	<b>85</b>
5.1. Clasificación de la marcha de la base UPT . . . . .	87
5.2. Clasificación de la marcha usando la base MoBo . . . . .	94
5.3. Clasificación de la marcha usando la base CASIA-A . . . . .	98
<b>6. Conclusiones Generales</b>	<b>103</b>
<b>A. Algoritmos Genéticos</b>	<b>107</b>
<b>B. Marcha humana</b>	<b>109</b>
<b>C. Software de adquisición de imágenes</b>	<b>111</b>
<b>D. Momentos de Zernike</b>	<b>113</b>
<b>E. Trabajos de investigación</b>	<b>115</b>

# Índice de figuras

1.1.	El hombre de la imagen tiene en las pierna un arco distintivo.	20
1.2.	Esquema general de un sistema de reconocimiento de la marcha.	21
2.1.	(a) Mapa de intensidades, (b) Función imagen, (c) Representación del plano matricial de una imagen digital.	29
2.2.	a) Imágenes binarias de dos objetos bajo rotación y cambio de escala. b) Invariante de Hu $\phi_1$ de las imágenes.	34
2.3.	Transformación de coordenadas de una imagen en el círculo de radio unidad: (a) Imagen fuera, $a = -1, b = 1$ ; (b) Imagen dentro, $a = -\sqrt{\frac{1}{2}}, b = \sqrt{\frac{1}{2}}$ .	36
2.4.	Polinomios radiales de Jacobi $J_n(\alpha, \beta, r)$ para ordenes $n = 0, 1, 2, 9, 10$ . a) $\alpha = \beta = 4$ , b) $\alpha = \beta = 6$ , c) $\alpha = \beta = 8$ , d) $\alpha = \beta = 10$ .	39
2.5.	Polinomios radiales de Jacobi $J_n(\alpha, \beta, r)$ para: a) $\alpha = \beta = 4$ , b) $\alpha = \beta = 6$ , c) $\alpha = \beta = 8$ y d) $\alpha = \beta = 10$ .	40
2.6.	a) Imágenes binarias de dos personas con cambio de escala aleatorio. b) Invariantes de Jacobi-Fourier usando $\alpha = \beta = 4$ .	42
3.1.	Tres tipos de imágenes de entrada para el algoritmo propuesto.	44
3.2.	Esquema de la adquisición de imágenes.	46
3.3.	Marcos de tres secuencias a tres diferentes distancias. El número de marcos capturados en cada secuencia es 50.	47
3.4.	Proceso del preprocesamiento.	49
3.5.	49 Imágenes de una secuencia procesada.	50
3.6.	Resolución de imágenes procesadas.	51
3.7.	Posición de las cámaras de la captura de datos.	52
3.8.	Vistas de la base MoBo con sus respectivas siluetas.	52
3.9.	Vistas de una secuencia caminando lento.	53

3.10. Vistas de una secuencia caminando rápido. . . . .	54
3.11. Vistas de una secuencia caminando con una inclinación. . . . .	55
3.12. Vistas de una secuencia caminando con un balón. . . . .	56
3.13. Preprocesamiento de la base MoBo. . . . .	57
3.14. Imágenes de una secuencia de la base MoBo procesada. . . . .	58
3.15. Base CASIA-A. a) Vista frontal. b) Vista a $45^\circ$ . c) Vista a $90^\circ$ . . . . .	59
3.16. Base CASIA-B. Vistas de $0^\circ, 18^\circ, 36^\circ, \dots, 180^\circ$ . . . . .	60
3.17. Base CASIA-B, con distintos tipos de ropa: a) mochila, b) abrigo y c) normal. . . . .	61
3.18. Base CASIA-C, a) caminando lento y b) caminando con mochila. . . . .	61
3.19. Preproceso para obtener el cuadro de la parte inferior. . . . .	62
3.20. 49 Imágenes de una secuencia de la base CASIA-A procesada. . . . .	63
4.1. Historia de momentos de una secuencia de 16 imágenes. . . . .	66
4.2. Historia de momentos e intervalo seleccionado $[w_s, w_s + H]$ . En donde $H$ es un valor arbitrario. . . . .	68
4.3. a) Historia de momentos $s = 1$ . b) Historia de momentos $s = 2$ . c) Sección de referencia para la secuencias $s = 1$ y $s = 2$ . . . . .	69
4.4. Historia de momentos. a) En forma inicial. b) Puestas en fase. . . . .	71
4.5. Historia de momentos de dos personas. . . . .	72
4.6. Ajuste de la distribución de probabilidad Normal de la simulación de 100 secuencias para $D_{2,1}^{s,1}$ . . . . .	74
4.7. Ajuste de la distribución de probabilidad Normal de la simulación de 100 secuencias para $D_{30,1}^{s,1}$ . . . . .	75
4.8. Mapa $Q$ , para clasificar dos clases. . . . .	76
4.9. a) Gráfica del mejor descriptor $n = 1, m = 2$ para $\alpha = \beta = 2$ . b) Gráfica del peor descriptor $n = 10, m = 9$ para $\alpha = \beta = 2$ . . . . .	77
4.10. a) Gráfica de los dos mejores descriptores. b) Gráfica de los tres mejores descriptores. Las clases siempre se separan y están agrupadas. . . . .	78
4.11. Métodos para la clasificación . . . . .	81
4.12. Gráfica de porcentajes de clasificación con diferente posición de los mejores descriptores. . . . .	83
5.1. Proceso de clasificación del método propuesto. . . . .	86
5.2. Mapa $Q$ para $\alpha = \beta = 4$ usando la base UPT . . . . .	87
5.3. Mapas $Q$ para $\alpha = \beta = 6, 8, 10$ . . . . .	88

5.4. Mejor ajuste de la simulación del descriptor con parámetros $n = 1$ y $m = 0$ para $\alpha = \beta = 4$ .El descriptor se ajusta a la distribución normal. . . . .	89
5.5. Mejor ajuste de la simulación del descriptor con parámetros $n = 10$ y $m = 10$ para $\alpha = \beta = 4$ .El descriptor no se ajusta a una distribución normal. . . . .	90
5.6. Gráfica de dispersión de los tres mejores descriptores para $\alpha = \beta = 4$ , de cada una de las 4 secuencias de imágenes para cada uno de los 14 individuos. . . . .	91
5.7. Porcentajes de clasificación usando la base UPT . . . . .	92
5.8. Mapa de la posición de los descriptores para clasificar para $\alpha = \beta = 4$ usando la base UPT . . . . .	93
5.9. Mapas $Q$ para $\alpha = \beta = 4, 6, 8, 10$ , usando la base MoBo. . . . .	95
5.10. Diagramas de dispersión para $\alpha = \beta = 4, 6, 8, 10$ .usando la base MoBo. . . . .	96
5.11. Porcentajes de clasificación de la combinación de los mejores descriptores usando la base MoBo. . . . .	97
5.12. Mapas $Q$ para $\alpha = \beta = 4, 6, 8, 10$ , usando la base CASIA-A. . . . .	99
5.13. Diagramas de dispersión para $\alpha = \beta = 4, 6, 8, 10$ .usando la base CASIA-A. . . . .	100
5.14. Porcentajes de clasificación usando la base CASIA-A . . . . .	101
B.1. Ciclo de la marcha normal. El ciclo de la marcha se divide en dos fases: de apoyo y de oscilación. Además, el ciclo de la marcha puede dividirse en las fases de soporte simple y soporte doble. . . . .	109
C.1. Pantalla principal del programa para capturar una secuencia de imágenes. . . . .	111
D.1. Polinomios Radiales de Zernike $R_n(r)$ . . . . .	114



# Índice de cuadros

3.1. Tiempos cómputo de los datos de estrada usados en la Fig. 3.1	44
3.2. Bases de datos usadas en el contexto internacional para el reconocimiento de la marcha. . . . .	45
6.1. Cuadro comparativo de los resultados de las tres bases: UPT, MoBo y CASIA-A . . . . .	104
D.1. Polinomios radiales de Zernike . . . . .	114



# Capítulo 1

## Introducción

Existen ciertas características biológicas o conductuales singulares e inalterables, que pueden ser analizadas y medidas para crear una huella irrepetible. Estas características son difíciles de perder, transferir u olvidar y son perdurables en el tiempo.

La biometría informática es la aplicación de técnicas matemáticas y estadísticas sobre los rasgos físicos o de conducta de un individuo, para verificar identidades o para identificar individuos.

Aunque las técnicas biométricas usan una combinación de factores corporales y de comportamiento, se clasifican de manera separada. La medición de las características corporales de la persona es conocida como biometría estática. Los principales estudios y aplicaciones de la biometría estática están basados en la medición de huellas digitales, geometría de la mano, iris, forma de la cara, retina y venas del dorso de la mano. Existen también, pero menos usadas, técnicas biométricas basadas en la forma de la oreja, temperatura corporal y la forma del cuerpo.

Las mediciones de las características de comportamiento de las personas es conocida como biometría dinámica. Los principales estudios y aplicaciones de la biometría dinámica están basados en el patrón de voz, firma manuscrita, dinámica del teclado, la marcha y el análisis de gestos.

En los últimos años existe un gran interés sobre la biometría de la marcha. El avance de la tecnología brinda la oportunidad de procesar imágenes de videos en cuestión de segundos, creando nuevos enfoques en el análisis de la marcha.

La marcha definida como el paso bípedo que utiliza la raza humana para desplazarse de un lugar a otro, con bajo esfuerzo y mínimo consumo energéti-

co, constituye un elemento característico que identifica a todo sujeto, pues permite relevar aspectos individuales de la personalidad, la autoestima, la condición de salud y las situaciones emocionales de cada persona [1].

El reconocimiento y análisis de la marcha humana es una tecnología prometedora, con posibles aplicaciones en numerosos sectores de nuestra sociedad, incluyendo las aplicaciones de vigilancia, tales como la caracterización de un sistema para identificación y autenticación de las personas, la detección de comportamiento sospechoso a partir de un vídeo de vigilancia. El reconocimiento de la marcha ha contribuido a las pruebas para condenados, como el caso del ladrón de Lancashire del Reino Unido [2]. Un podólogo, que se especializa en el análisis de la marcha, adaptó la marcha de un individuo y junto con pruebas de ADN pudo ser acusado. En la Fig. 1.1 se muestra una imagen en infrarrojo del circuito cerrado donde el individuo es visto. El análisis de la marcha puede ser una herramienta de apoyo para un especialista.



Figura 1.1: El hombre de la imagen tiene en las pierna un arco distintivo.

Sin embargo, existen importantes aplicaciones relacionadas con el análisis de la marcha, por ejemplo, la detección de alteraciones posturales, debido a los trastornos de la movilidad o el envejecimiento, también, en el desarrollo de estrategias óptimas en la técnica de algunos deportes.

Por otra parte, el análisis del patrón de la marcha también se utiliza en dispositivos para la rehabilitación automática robotizada de los pacientes.

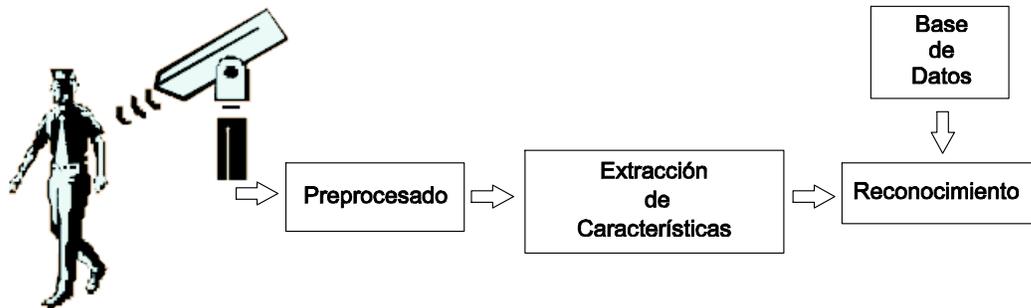


Figura 1.2: Esquema general de un sistema de reconocimiento de la marcha.

En este caso, el dispositivo reproduce el papel de un fisioterapeuta. Si la persona lesionada es capaz de contribuir activamente al patrón de marcha que determina el movimiento del dispositivo. El cual, se puede adaptar a la postura de siempre del paciente. Otra aplicación de análisis de la marcha es en el diseño de robots bípedos. Se busca lograr la estabilidad en superficies diferentes, mientras suben y descienden escaleras. Por último, el análisis de la marcha con aplicaciones en la industria de la animación y video juegos, en los que la presentación realista de las personas al caminar es muy importante.

El reconocimiento de patrones en la marcha, busca extraer características que las diferencian del resto. El sistema de reconocimiento debe asignar a cada individuo en su categoría o clase. Por lo cual es necesario generar características únicas que puedan ser usadas en la clasificación. La extracción de las características en ocasiones viene precedido, por un preprocesado de la señal, necesario para corregir posibles deficiencias en los datos debido a errores del sensor, o bien para preparar los datos de cara a posteriores procesos de las etapas de extracción de características o clasificación. En la Fig. 1.2 se muestran las etapas del reconocimiento de la marcha.

El reconocimiento de la marcha depende de una secuencia de video tomadas en ambientes controlados o incontrolados. Existen limitaciones evidentes en la captura de la marcha que la hace extremadamente difícil de identificar y de registrar todos los parámetros que afectan a la marcha. Incluso si mejora la precisión con la que es posible medir algunos parámetros de la marcha, todavía no se sabe si el conocimiento actual de estos parámetros permita una clasificación adecuada para lograr un despliegue a gran escala de tecnologías

de la marcha [3]. Por otra parte, estudios informan que los cambios en el tiempo de la marcha se ve afectada por la ropa, el calzado, la superficie, la velocidad al caminar, y el estado anímico [4]. Los hechos imponen limitaciones en la precisión inherente a la marcha y la cuestión de su despliegue como discriminativo biométricos.

Las investigaciones realizadas hasta ahora en el ámbito del reconocimiento de la marcha han demostrado que esta puede ser confiable en combinación con otros datos biométricos. Si suponemos que la palma de la mano, huellas digitales, iris y otros métodos pertenecen a una clase de biométricos invasivos. La biometría adicional que podría ser utilizada en conjunción con la marcha de un sistema multibiométrico sería la cara y la presión del pie [5] (este último requiere alguna equipos especializados para medir la fuerza de reacción del suelo). Se esperaba que se utilizara como único medio de identificación de los individuos en grandes bases de datos, sin embargo se ha visto como un componente de utilidad en un sistema biométrico multimodal.

## 1.1. Antecedentes

El estudio de la marcha como una forma de identificación se intentó por primera vez hace algunas décadas desde un punto de vista médico-conductual [6], [7]. Más tarde, se hicieron varios intentos para investigar el problema de reconocimiento de la marcha desde la perspectiva de la captura y análisis de señales [8] - [12].

El despliegue de los campos de movimiento en el reconocimiento de la marcha fue investigada en las referencias [11] y [13]. Aunque ambos métodos reportaron buenos resultados en sus propias bases de datos, suponen la disponibilidad de información de la textura, que debe ser utilizada para el cálculo del movimiento. En las referencias [14] - [17], se proporcionan diversas técnicas para la extracción de características, los métodos están basados en el cálculo de las proyecciones, los contornos, o de otras características como la marcha de la silueta. En la referencia [18], se hace la comparación de diferentes técnicas, para mejorar la calidad de las siluetas extraídas a partir de secuencias de vídeo.

En el año 2000, la Universidad de Southampton publicó la base de datos SOTON, con diez sujetos caminando. Un número relativamente grande de documentos han utilizado la base SOTON, y alcanzan el 100 % de clasificación. La mayoría de los métodos son estadísticos, como el método de los

momentos de velocidad de Shutler et al. [19], el operador de simetría de Hayfron-Acquah et al. [20], la máscara de la marcha de Foster et al. [21], el análisis de la forma de la silueta de Wang et al. [22], o la descripción en planos polares de la silueta de Chen y Gao [23].

En 2001, el Centro de Biometría y de investigación sobre seguridad (CB-SR) de la Academia China de Ciencias (CASIA) presentó un conjunto de datos con diferentes ángulos de visión, llamada CASIA-A compuesta por 20 sujetos y tres vistas diferentes. Un gran número de trabajos han sido probados en esta base de datos, obteniendo un 92,5 % de nivel de reconocimiento en la vista lateral. La mayoría de ellos están basados en estadísticas, incluido el enfoque de Yang et al. [25], el método de vectores de masa de Hong et al. [26], la metodología de la fusión de Nizami et al. [27], la descripción en planos polares de la silueta de Chen y Gao [23], y la silueta del espectro espacio-temporal de Lam et al. [28].

Recientemente, se publicó una base de datos CASIA multivista marcha grande llamado CASIA-B [24], en 124 sujetos que han sido capturados en 11 cámaras sincronizadas desde diferentes puntos de vista.

La representación polar de un sujeto caminando propuesto por Chen y Gao [23] se ha probado sólo en la vista lateral, mientras que Bashir et al. [29] y Yu et al. [30] probaron su enfoque vista independiente mediante la comparación de diferentes ángulos de visión.

## 1.2. Objetivos

Este trabajo de Tesis está enfocado al análisis de los descriptores de la marcha y su correcta clasificación. De manera particular, las tareas descritas son:

- A.** Análisis de una secuencia de imágenes para crear descriptores de la marcha a través de los momentos de Jacobi-Fourier.
- B.** Implementación de un método para poner en fase las secuencias de momentos de una misma persona.
- C.** Normalización de una secuencia de momentos para que sean invariantes a la escala.
- D.** Implementación de una medida discriminatoria de momentos para seleccionar los  $k$  momentos que mejor clasifiquen.
- E.** Implementación de un método para la clasificación con la mínima cantidad de descriptores.
- F.** Programación de algoritmos que calculan los momentos de una secuencia de imágenes digitales, sus descriptores de la marcha y el método de clasificación.
- G.** Análisis comparativo del número de descriptores que son necesarios para una correcta clasificación de los momentos de Jacobi-Fourier y los momentos de Zernike.

### 1.3. Estado del Arte del Reconocimiento de la Marcha

Un paso crítico en el reconocimiento de la marcha es la extracción de características, es decir, la extracción a partir de secuencias de vídeo que representan la marcha de una persona. Este paso es muy importante ya que hay numerosas maneras concebibles para extraer la señal de la marcha a partir de una secuencia de video, por ejemplo, espacial o temporal, su combinación espacial y temporal, o la extracción de características en el dominio de frecuencias de Fourier.

Por lo tanto, la extracción de las características de la marcha debe asegurar la mayor cantidad posible de información. El reconocimiento tiene como objetivo comparar una señal desconocida de la marcha con una señal almacenada en una base de datos. Las técnicas utilizadas para el reconocimiento de la marcha se pueden dividir en dos categorías: holístico y basado en modelos.

El método holístico opera directamente en las secuencias de la marcha sin asumir ningún modelo en específico para la marcha humana. Una clase muy interesante de técnica holística, es emplear mapas binarios (siluetas) de la marcha humana. Estas técnicas son especialmente adecuadas para la mayoría de aplicaciones prácticas, ya que el color o la textura de la información puede no estar disponible o ser extraíbles.

El enfoque basado en modelos se basa en la identificación de los parámetros de la marcha que se emplean mediante el procesamiento de secuencias, estos métodos por lo general requieren secuencias de imágenes de alta calidad para ser útiles. A diferencia de los enfoques holísticos, los enfoques basados en modelos son invariantes a la vista y a la escala. Esto es una ventaja considerable sobre el enfoque holístico. Sin embargo tienen la desventaja de que la oclusión de los sujetos caminando con objetos, puede hacer que el cálculo de los parámetros del modelo sean imposibles. Por esta razón un sistema de adquisición multicámara de la marcha sería más apropiado para esta técnica.

## 1.4. Aportaciones

Con esta investigación se contribuye al desarrollo de los sistemas de visión por computadora basados en los momentos de Jacobi-Fourier y la correlación de la historia de momentos para el análisis y clasificación de la marcha, con aplicaciones en seguridad. El método puede ser extendido en otros campos en el contexto de evaluar el movimiento corporal humano para determinar la condición del funcionamiento y discapacidad de un individuo. También, en el análisis de objetos o individuos en movimiento, que tengan cierto patrón repetitivo. Por ejemplo, la corrección de la técnica en un atleta.

La aportación más importante esta en lograr que todas las historias de momentos del movimiento de todas las secuencias de imágenes queden en fase, la cual permitirá extraer un patrón de referencia para poderlo comparar con otras secuencias

Otra aportación importante esta en el diseño de una medida discriminatoria para  $N$  número de clases distintas, que permita encontrar los parámetros de los momentos que mejor clasifican.

El método de clasificación de la marcha propuesto en este trabajo de Tesis, es la combinación de dos métodos de clasificación, lo cual permite aumentar el rango de clasificación.

Se propone un método que minimice el número de descriptores que al ser combinados clasifiquen al 100%, mediante técnicas heurísticas modernas.

Se comprueba que las familias de momentos de Jacobi-Fourier clasifican con un número menor de descriptores que los momentos de Zernike.

## Capítulo 2

# Revisión Matemática

El problema general a ser considerado, es la caracterización, evaluación y manipulación de información visual que está presente en el plano imagen, de un sistema óptico. Para el método de momentos ofrece una solución al problema. Esto se debe a que los momentos de una función imagen, son valores numéricos que describen la distribución de intensidad en un arreglo de píxeles. También son descriptores globales y locales de la forma. Es por ello que, las funciones momentos han sido usadas en áreas como reconocimiento de patrones, clasificación de objetos, estimación de la posición, codificación de imágenes y reconstrucción.

Además, este método es potencialmente útil en problemas como el sensado en tiempo real de aberraciones en sistemas ópticos, a causa de la turbulencia atmosférica o cambios de la temperatura. Los momentos de una función imagen han sido calculados usando procesadores ópticos coherentes o bien procesadores digitales.

Históricamente, Hu publicó el primer artículo significativo en el uso de momentos geométricos de la imagen como, el centroide, el tamaño y la orientación. Sin embargo, las funciones base de los momentos geométricos no son ortogonales, lo cual indica una dependencia entre descriptores y una redundancia en información.

En general para el análisis de imágenes, el conjunto base en el cual la función imagen es proyectada, debe ser ortogonal. La ventaja de usar bases ortogonales está en la independencia de los descriptores y la consecuente reducción del cálculo numérico.

En este capítulo se describen los dos tipos de funciones momento: ortogonales y no ortogonales. Los primeros para la extracción de características en una imagen digital y los segundos para el cálculo de parámetros tales el centroide y la escala.

## 2.1. Momentos Geométricos de una Imagen

Una imagen digital es una función  $f(x, y)$  que se ha discretizado tanto en coordenadas espaciales como en el brillo, donde el valor de la función denota la intensidad en el pixel  $(x, y)$ . La Fig. 2.1 muestra un ejemplo de esta distribución.

Los momentos geométricos son los más simples entre los momentos de una imagen, con el núcleo definido como un producto de las coordenadas del píxel. La ventaja más importante de los momentos geométricos es que las transformaciones de las coordenadas de la imagen son fácilmente expresadas y analizadas en términos de las transformaciones correspondientes en el espacio de los momentos. Los momentos geométricos están definidos con el conjunto base de transformación  $\{x^p y^q\}$ . Los momentos geométricos de orden  $(p + q)$  en dos dimensiones, están denotados por  $m_{pq}$ , y se definen como [31],

$$m_{pq} = \int \int_{\zeta} x^p y^q f(x, y) dx dy, \quad (2.1)$$

en donde  $p, q = 0, 1, 2, 3, \dots$  y  $\zeta$  es la región del espacio de pixeles en la cual la función de intensidad de la imagen  $f(x, y)$  está definida. La Ec. (2.1) tiene la forma de la proyección de la función  $f(x, y)$  sobre el monomio  $x^p y^q$ . El conjunto base  $\{x^p y^q\}$ , es completo y no ortogonal. A continuación se establecen los teoremas de existencia y unicidad [31], del conjunto de momentos geométricos de una imagen.

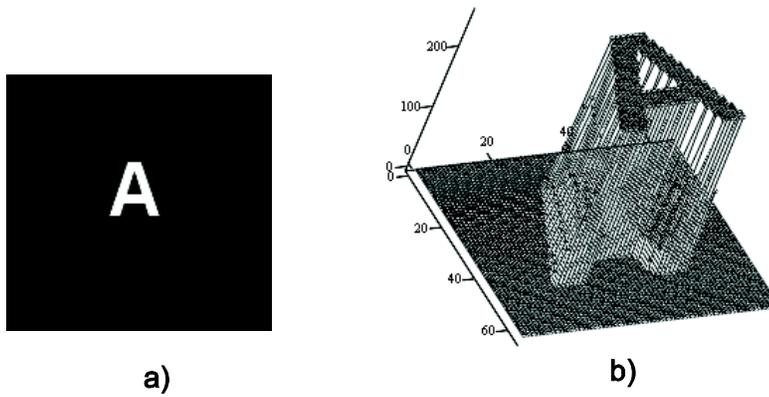


Figura 2.1: (a) Mapa de intensidades, (b) Función imagen, (c) Representación del plano matricial de una imagen digital.

**Teorema 1** (*Existencia*): Asumiendo que la función de intensidad  $f(x, y)$  es continua a trozos y acotada en la región, entonces los momentos de todos los órdenes existen y son finitos.

**Teorema 2** (*Unicidad*): Asumiendo que la función de intensidad  $f(x, y)$  es continua a trozos y acotada en la región, la sucesión de momentos esta únicamente determinada por la función  $f(x, y)$  y viceversa.

Los momentos geométricos de diferentes órdenes representan diferentes características espaciales de la distribución de intensidades en la imagen. Un conjunto de momentos puede formar un descriptor de la forma global de la imagen.

El momento de orden cero  $m_{00}$ , representa la intensidad total de una función imagen y para imágenes binarias representa el área,

$$m_{00} = \int \int_{\zeta} f(x, y) dx dy. \quad (2.2)$$

Los momentos de primer orden,  $m_{01}$ ,  $m_{10}$ , dan respectivamente, el momento de intensidad alrededor del eje  $x$  y el eje  $y$  en la imagen. El centroide de intensidad  $(x_0, y_0)$  se define como,

$$x_0 = \frac{m_{10}}{m_{00}}, \quad y_0 = \frac{m_{01}}{m_{00}}. \quad (2.3)$$

Si  $f(x, y)$  representa una función imagen discreta o la distribución de intensidades de una imagen digital, la Ec.(2.1) se representa mediante una suma discreta como se muestra a continuación,

$$m_{pq} = \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} x^p y^q f(x, y), \quad (2.4)$$

en donde  $M \times N$  es el tamaño de la función imagen discreta.

Es conveniente evaluar los momentos en el origen del sistema de referencia desplazado al centroide de intensidad de la imagen. Esta traslación hace el cálculo de momentos independiente de la posición del sistema de referencia de la imagen. Los momentos calculados con respecto a la intensidad del centroide son llamados momentos centrales y están definidos como,

$$\mu_{pq} = \int \int_{\zeta} (x - x_0)^p (y - y_0)^q f(x, y) dx dy, \quad (2.5)$$

en donde  $p, q = 0, 1, 2, 3 \dots \infty$ . Para los momentos centrales de orden cero y uno; resulta que,

$$\begin{aligned}\mu_{00} &= m_{00}, \\ \mu_{10} &= \mu_{01} = 0.\end{aligned}\tag{2.6}$$

En el caso discreto, la Ec. (2.5) se convierte en,

$$\mu_{pq} = \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} (x - x_0)^p (y - y_0)^q f(x, y).\tag{2.7}$$

### 2.1.1. Descriptores Invariantes

Los momentos invariantes constituyen un conjunto de descriptores que se pueden usar bajo las siguientes transformaciones geométricas y en intensidad,

$$f_1(x, y) = u f_2(x, y)\tag{2.8}$$

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = g \begin{bmatrix} \cos \theta & \text{sen } \theta \\ -\text{sen } \theta & \cos \theta \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} a \\ b \end{bmatrix}\tag{2.9}$$

donde  $\theta$  es una rotación arbitraria,  $(a, b)$  el desplazamiento respectivamente en  $x$  e  $y$ ,  $u$  el cambio en el contraste y  $g$  un cambio de escala arbitrario. Siendo  $f_1$  y  $f_2$  funciones imagen discreta, llamadas imagen de referencia y distorsionada respectivamente.

Por definición los momentos centrales dados en la Ec. (2.5) son invariantes a la traslación. Esto se debe a que el centroide de la imagen  $(x_0, y_0)$ , se mueve cuando ésta se traslada, y porque los momentos centrales están definidos con respecto a este punto como origen.

Una transformación de las coordenadas del píxel por un factor de escala  $g$ , está dada por,

$$\begin{aligned}x' &= gx \\ y' &= gy.\end{aligned}\tag{2.10}$$

La transformación anterior nos conduce a la siguiente expresión para el área del elemento escalado,

$$dx' dy' = g^2 dx dy.\tag{2.11}$$

Los momentos de la imagen escalada pueden ser expresados en términos de los momentos de la imagen original como [31],

$$m'_{pq} = g^{p+q+2} m_{pq}. \quad (2.12)$$

De lo anterior se tiene que si  $p = q = 0$ , entonces,

$$m'_{00} = g^2 m_{00}. \quad (2.13)$$

De esta manera, la escala  $g$  se obtiene usando la relación,

$$g = \sqrt{\frac{m_{00}^{\text{Referencia}}}{m_{00}^{\text{Objeto}}}}. \quad (2.14)$$

Elevando a la potencia  $\frac{p+q+2}{2}$ , a ambos miembros de la Ec. (2.13) resulta,

$$(m'_{00})^{\frac{p+q+2}{2}} = g^{p+q+2} (m_{00})^{\frac{p+q+2}{2}}. \quad (2.15)$$

Dividiendo la Ec. (2.12) entre la Ec. (2.15) y eliminando el factor de escala  $g$  desconocido, de las dos ecuaciones se tiene que,

$$\frac{m'_{pq}}{(m'_{00})^{\frac{p+q+2}{2}}} = \frac{m_{pq}}{(m_{00})^{\frac{p+q+2}{2}}}. \quad (2.16)$$

Podemos intercambiar a  $m_{pq}$  por los momentos centrales  $\mu_{pq}$  en la Ec. (2.16) que son además invariantes al desplazamiento. De esta manera se define el término  $\eta_{pq}$  como,

$$\eta_{pq} = \frac{\mu_{pq}}{(\mu_{00})^\gamma}, \quad (2.17)$$

donde

$$\gamma = \frac{p+q+2}{2}. \quad (2.18)$$

El término es invariante a la traslación y variación en la escala de una imagen. Un método para derivar invariantes a la rotación usando la teoría de invariantes algebraicos fue propuesto por Hu [32]. Tales invariantes son llamados *momentos invariantes*. Un conjunto de siete momentos invariantes puede ser derivado del segundo y tercer momento, mediante las siguientes expresiones,

$$\phi_1 = \eta_{20} + \eta_{02}, \quad (2.19)$$

$$\phi_2 = (\eta_{20} + \eta_{02})^2 + 4(\eta_{11})^2, \quad (2.20)$$

$$\phi_3 = (\eta_{30} - 3\eta_{12})^2 + (3\eta_{21} - \eta_{03})^2, \quad (2.21)$$

$$\phi_4 = (\eta_{30} + \eta_{12})^2 + (\eta_{21} + \eta_{03})^2, \quad (2.22)$$

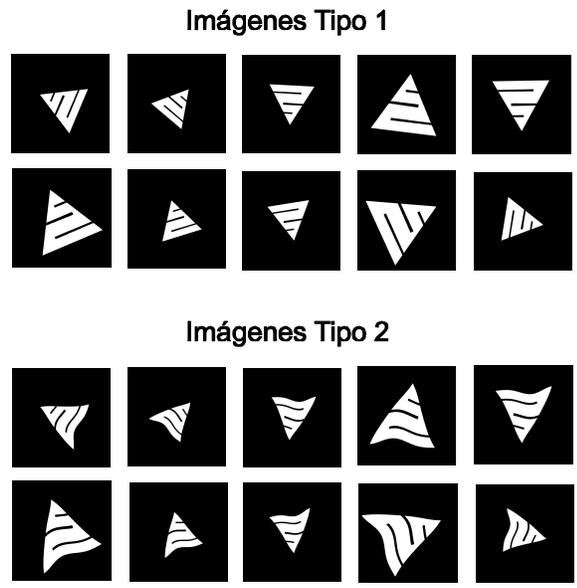
$$\phi_5 = (\eta_{30} - 3\eta_{12})(\eta_{30} + \eta_{12}) [(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2] \\ + (3\eta_{21} + \eta_{03})(\eta_{21} + \eta_{03}) [(3\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2], \quad (2.23)$$

$$\phi_6 = (\eta_{20} + \eta_{02}) [(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2] \\ + 4\eta_{11}(\eta_{30} + \eta_{12})(\eta_{21} + \eta_{03}), \quad (2.24)$$

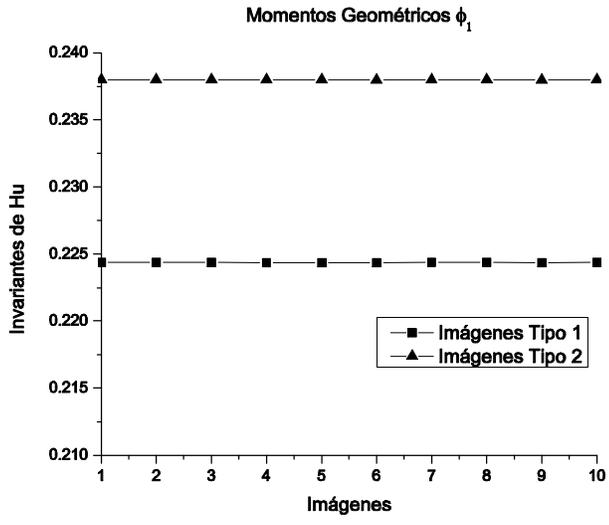
$$\phi_7 = (3\eta_{21} - \eta_{03})(\eta_{30} + \eta_{12}) [(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2] \\ + (3\eta_{12} + \eta_{30})(\eta_{21} + \eta_{03}) [(3\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2]. \quad (2.25)$$

Este conjunto de momentos  $\{\phi_1, \phi_2, \phi_3, \dots, \phi_7\}$  son invariantes a la traslación, rotación y cambio de escala.

Para probar la invariancia a rotación y a escala de los momentos Invariantes de Hu, se toman dos imágenes y se generan imágenes rotadas y escaladas de forma aleatoria, como se muestra en la Fig.2.2. Se calcula el primer Invariante de Hu ( $\phi_1$ ) de la Ec.(2.19) para cada una de las imágenes generadas y se grafican en la Fig.2.2 . Como se puede ver, las imágenes del mismo tipo tienen el mismo valor y su valor es distinto entre tipos de imágenes.



a)



b)

Figura 2.2: a) Imágenes binarias de dos objetos bajo rotación y cambio de escala. b) Invariante de Hu  $\phi_1$  de las imágenes.

## 2.2. Momentos de Jacobi-Fourier de una Imagen

Los momentos ortogonales de una imagen digital son adecuados para aplicaciones en el campo de reconocimiento de patrones, clasificación de objetos, codificación y reconstrucción de una imagen. El conjunto de momentos ofrece generalmente la representación global de las características de una imagen [34], los cuales se usan como técnicas de identificación en áreas de visión por computadora.

Los momentos están construidos sobre una base que consiste de un polinomio ortogonal radial multiplicado por una exponencial compleja angular de Fourier. Este producto de funciones se denota por  $P_{nm}(r, \theta)$ . Los momentos circulares  $\phi_{nm}$  están dados por,

$$\phi_{nm} = \int \int_{\zeta} f(r, \theta) P_{nm}(r, \theta) r dr d\theta \quad (2.26)$$

en donde  $f(r, \theta)$  es la función imagen discreta en coordenadas polares. Las funciones base ortogonales válidas dentro del círculo de radio unidad, requieren hacer una transformación del espacio de coordenadas  $\zeta$  a la región  $[-1, 1] \times [1, -1]$  o a  $[-\sqrt{\frac{1}{2}}, \sqrt{\frac{1}{2}}] \times [\sqrt{\frac{1}{2}}, -\sqrt{\frac{1}{2}}]$ , como se muestra en la Fig. 2.3.

Teague [35] utilizó por primera vez los momentos de Zernike [36] para la descripción de la forma. El kernel de los Momentos de Zernike está compuesto de los polinomios de Zernike y un factor complejo angular de Fourier en coordenadas polares.

Sheng et al [37], propuso los Momentos ortogonales de Fourier-Mellin, que son construidos por el proceso de ortogonalización de Gram-Schmidt, de un conjunto de monomios de bajas potencias. Otro tipo de momentos ortogonales, se debe a los Momentos de Chebyshev-Fourier [38].

Cada uno de los momentos anteriores es independiente y no asociados entre sí. Se forman varios tipos de Momentos ortogonales de Jacobi-Fourier al variar los parámetros  $\alpha$  y  $\beta$  en la definición de los polinomios de Jacobi. Ejemplo de estos, son los Momentos de Legendre-Fourier ( $\alpha=1, \beta=1$ ); Momentos de Chebyshev-Fourier ( $\alpha=2, \beta=3/2$ ); Momentos de Fourier-Mellin ( $\alpha=2, \beta=2$ ); Momentos de Zernike; Momentos de Pseudo-Zernike, y así sucesivamente.

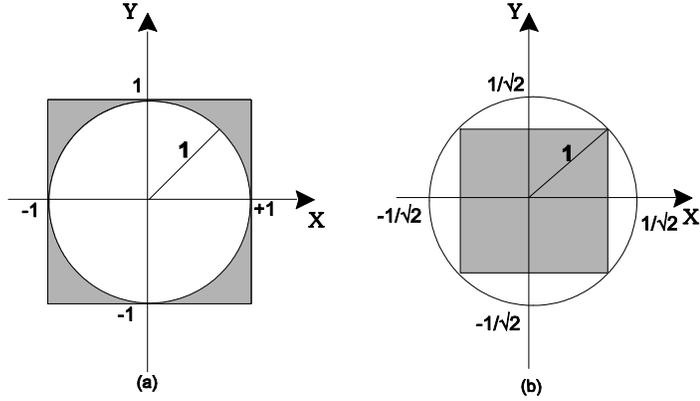


Figura 2.3: Transformación de coordenadas de una imagen en el círculo de radio unidad: (a) Imagen fuera,  $a = -1, b = 1$  ; (b) Imagen dentro,  $a = -\sqrt{\frac{1}{2}}, b = \sqrt{\frac{1}{2}}$ .

El kernel de los Momentos de Jacobi-Fourier  $P_{nm}(r, \theta)$  consiste en dos familias de funciones separables: el polinomio deformable de Jacobi  $J_n(\alpha, \beta, r)$  que es la función radial y la exponencial compleja de Fourier  $\exp(jm\theta)$  que es la función angular,

$$P_{nm}(r, \theta) = J_n(\alpha, \beta, r) \exp(jm\theta), \quad (2.27)$$

donde  $n$  y  $m$  son enteros. El kernel de la función de Jacobi-Fourier es un conjunto ortogonal en el interior del círculo unitario,  $0 \leq r \leq 1, 0 \leq \vartheta \leq 2\pi$ .

$$\int_0^{2\pi} \int_0^1 P_{nm}(r, \theta) P_{kl}(r, \theta) r dr d\theta = \delta_{nk} \delta_{ml}, \quad (2.28)$$

donde  $\delta_{nk} \delta_{ml}$  son los símbolos de Kronecker y  $r = 1$  es la máxima escala del objeto. Las funciones  $\exp(jm\theta)$  son ortogonales y las funciones radiales  $J_n(\alpha, \beta, r)$  también son ortogonales en el intervalo,

$$\int_0^1 J_n(\alpha, \beta, r) J_k(\alpha, \beta, r) r dr = \delta_{nk}(\alpha, \beta). \quad (2.29)$$

En la teoría de polinomios ortogonales el polinomio Genérico de Jacobi  $G_n(\alpha, \beta, r)$  está definido como,

$$G_n(\alpha, \beta, r) = \frac{n!(\beta-1)!}{(\alpha+n-1)!} \sum_{s=0}^n (-1)^s \times \frac{(\alpha+n+s-1)!}{(n-s)!s!(\beta+s-1)!} r^s. \quad (2.30)$$

Los polinomios de Jacobi  $G_n(\alpha, \beta, r)$  cumplen con la propiedad de ortogonalidad en el intervalo  $0 \leq r \leq 1$ ,

$$\int_0^1 G_n(\alpha, \beta, r) G_m(\alpha, \beta, r) w(\alpha, \beta, r) r dr = b_n(\alpha, \beta) \delta_{nm}, \quad (2.31)$$

donde  $w(\alpha, \beta, r)$  es una función de peso y el  $b_n$  es llamada constante de normalización, dada por

$$b_n = \frac{n! [(\beta-1)!]^2 (\alpha-\beta+n)!}{(\beta+n-1)! (\alpha+n-1)! (\alpha+2n)}, \quad (2.32)$$

qué es una función de los parámetros  $\alpha$  y  $\beta$ . La función de peso general es,

$$w(\alpha, \beta, r) = (1-r)^{\alpha-\beta} r^{\beta-1} \quad \alpha - \beta > -1, \quad \beta > 0. \quad (2.33)$$

En las fórmulas anteriores,  $\alpha$  y  $\beta$  son parámetros reales, y su variación permiten formar familias de polinomios de Jacobi diferentes. Comparando las Ecs. (2.29) y (2.31), se puede obtener el conjunto de funciones radiales mediante,

$$J_n(\alpha, \beta, r) = \sqrt{\frac{w(\alpha, \beta, r)}{b_n(\alpha, \beta) r}} G_n(\alpha, \beta, r). \quad (2.34)$$

En el sistema de coordenadas polares, una función imagen  $f(r, \theta)$  se puede sintetizar en la superposición de componentes ortogonales tal como,

$$f(r, \theta) = \sum_{n=0}^{\infty} \sum_{m=-\infty}^{\infty} \phi_{nm} J_n(r) \exp(jm\theta), \quad (2.35)$$

donde  $\phi_{nm}$  son los coeficientes de la descomposición y llamados Momentos de Jacobi-Fourier, estos se pueden calcular mediante,

$$\phi_{nm} = \int_0^{2\pi} \int_0^1 f(r, \theta) J_n(\alpha, \beta, r) \exp(jm\theta) r dr d\theta. \quad (2.36)$$

Las formas explícitas de las funciones radiales de los polinomios de Jacobi, dependen de los valores de los parámetros,  $\alpha$  y  $\beta$ . A continuación se listan algunas expresiones donde  $\alpha = \beta$ .

$$J_n(4, 4, r) = (-1)^n r^2 \sqrt{4 + 2n} \sum_{s=0}^n (-1)^s \frac{(3 + n + s)!}{(n - s)! s! (3 + s)!} r^s. \quad (2.37)$$

$$J_n(6, 6, r) = (-1)^n r^4 \sqrt{6 + 2n} \sum_{s=0}^n (-1)^s \frac{(5 + n + s)!}{(n - s)! s! (5 + s)!} r^s. \quad (2.38)$$

$$J_n(8, 8, r) = (-1)^n r^6 \sqrt{8 + 2n} \sum_{s=0}^n (-1)^s \frac{(7 + n + s)!}{(n - s)! s! (7 + s)!} r^s. \quad (2.39)$$

$$J_n(10, 10, r) = (-1)^n r^8 \sqrt{10 + 2n} \sum_{s=0}^n (-1)^s \frac{(9 + n + s)!}{(n - s)! s! (9 + s)!} r^s. \quad (2.40)$$

En la Fig. 2.4 se muestran algunas gráficas de los polinomios radiales de Jacobi. La distribución de ceros es casi uniforme en el intervalo  $0 \leq r \leq 1$ . Las curvas oscilan y normalmente decrecen conforme  $r$  tiende a 1. Aunque para altos ordenes sucede que; pueden presentar inestabilidad numérica y dispararse a infinito [39]. Esto último puede ser observado en la grafica de la Fig. 2.5.

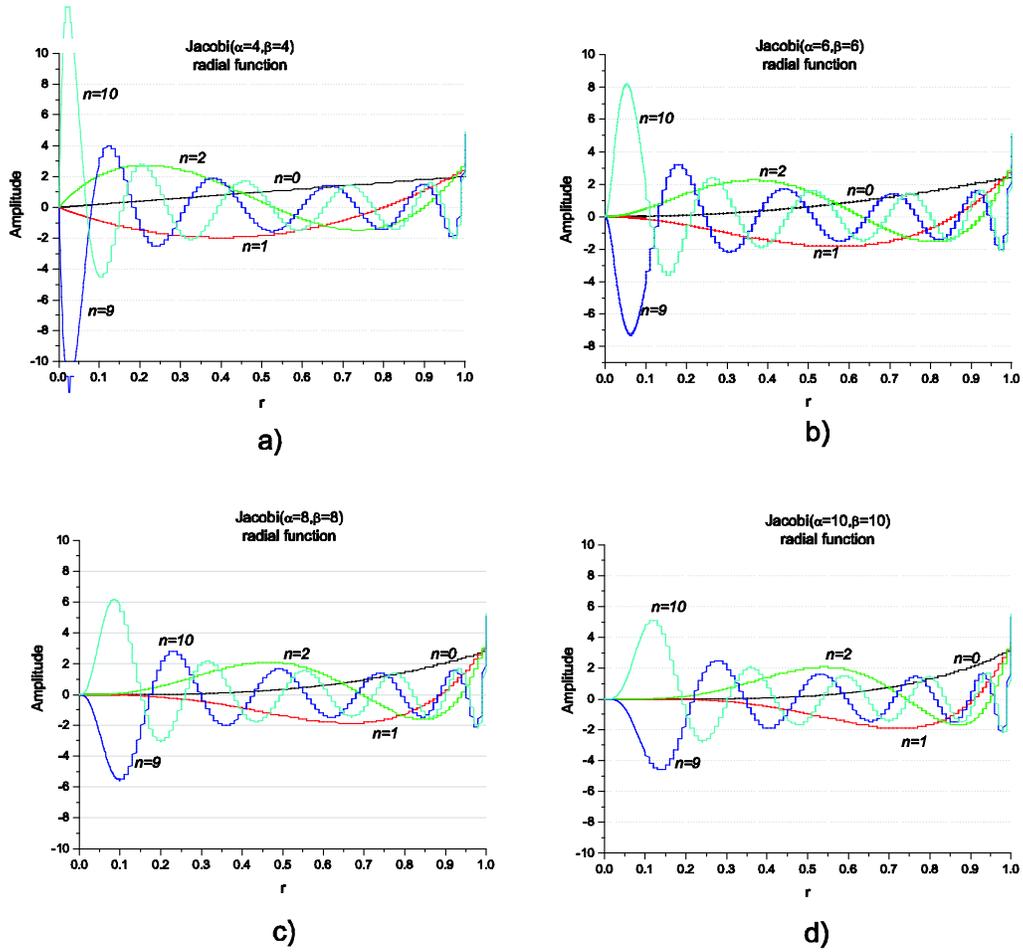


Figura 2.4: Polinomios radiales de Jacobi  $J_n(\alpha, \beta, r)$  para ordenes  $n = 0, 1, 2, 9, 10$ . a)  $\alpha = \beta = 4$ , b)  $\alpha = \beta = 6$ , c)  $\alpha = \beta = 8$ , d)  $\alpha = \beta = 10$ .

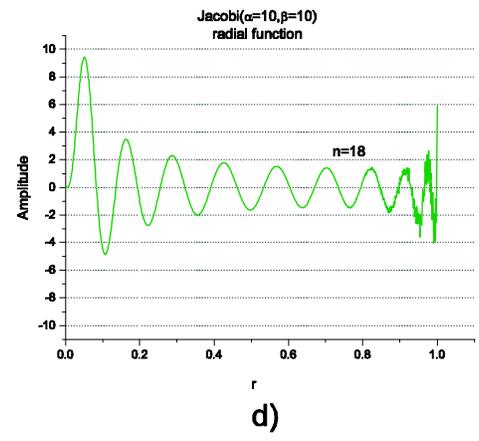
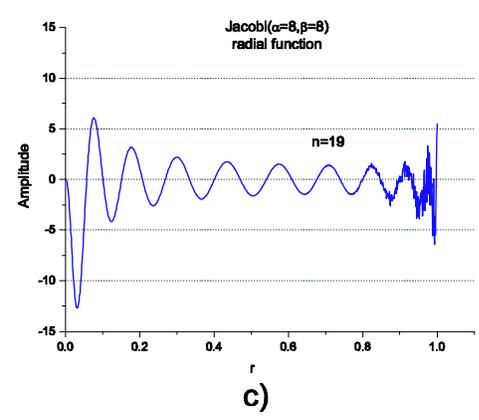
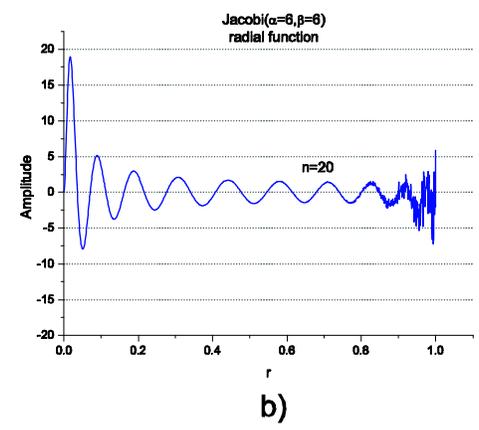
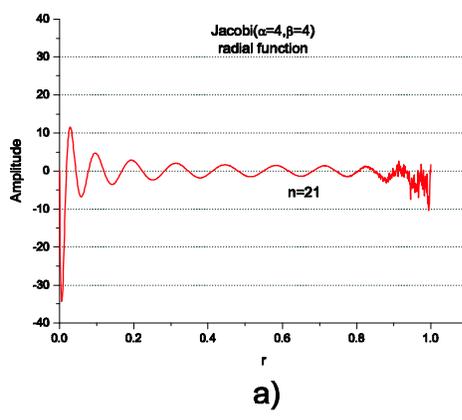


Figura 2.5: Polinomios radiales de Jacobi  $J_n(\alpha, \beta, r)$  para: a)  $\alpha = \beta = 4$ , b)  $\alpha = \beta = 6$ , c)  $\alpha = \beta = 8$  y d)  $\alpha = \beta = 10$

### 2.2.1. Descriptores Invariantes

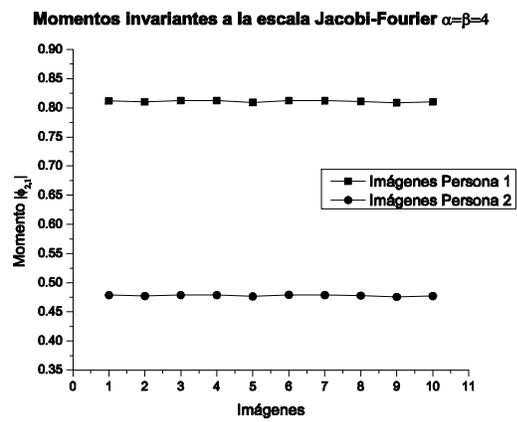
Los momentos de Jacobi-Fourier  $\phi_{nm}$  pueden ser normalizados para ser invariantes a transformaciones geométricas como desplazamiento o cambio de escala. Para lograr la invariancia al desplazamiento, usamos los momentos geométricos, primero se determina el centroide de la imagen con la Ec. (2.3). Se define con esto un nuevo origen de coordenadas. Todos los momentos calculados son ahora invariantes al desplazamiento.

Dada la función imagen discreta en coordenadas polares  $f(r, \theta)$ , está puede ser escalada mediante el factor  $g$ . Con estas consideraciones, la nueva definición de los momentos  $A_{nm}$  resulta ser:

$$A_{nm} = g^2 \int_0^{2\pi} \int_0^g f\left(\frac{r}{g}, \theta\right) P_{nm}\left(\frac{r}{g}, \theta\right) r dr d\theta = g^2 A'_{nm} \quad (2.41)$$

en donde  $A'_{nm}$  son los momentos de la imagen de referencia y el factor de escala  $g$  se obtienen mediante la Ec. (2.14).

Para probar la invariancia a la escala de los momentos Invariantes de Jacobi-Fourier, se toman dos imágenes y se generan imágenes escaladas de forma aleatoria, como se muestra en la Fig. 2.6a. Se calculan momentos Invariante Jacobi-Fourier con los parámetros  $\alpha = \beta = 4$  y así como su factor de escala  $g$  para cada una de las imágenes generadas, y se grafican en la Fig.2.6b. Como se puede ver, las imágenes de la misma clase tienen el mismo valor del momento invariante. Existe una diferencia en el valor del invariante entre clases.



**b)**

Figura 2.6: a) Imágenes binarias de dos personas con cambio de escala aleatorio. b) Invariantes de Jacobi-Fourier usando  $\alpha = \beta = 4$ .

## Capítulo 3

# Bases de datos y Sistemas de Adquisición de Imágenes Digitales para la marcha

La visión por computadora y el procesamiento de imágenes generalmente inician con la adquisición de la imagen. Su misión es resaltar los aspectos del objeto que interesen en una aplicación concreta, ocultar aquellas otras que impidan verlos y garantizar que la cámara esté registrando los objetos siempre con intensidad iguales o similares. La calidad de la imagen digitalizada depende de su resolución y rangos dinámicos del valor.

En este caso, la información que nos interesa registrar son secuencias de imágenes de individuos, para que a partir de estas se obtengan patrones de la forma de caminar.

La marcha como actividad rítmica y cíclica que compromete todo el cuerpo en especial las dos extremidades inferiores [1], ha sido sujeta de estudios sobre la contribución de las diferentes componentes del cuerpo para encontrar patrones que sirvan como métodos de reconocimiento de la marcha [40]. Para nuestro estudio, se ha hecho una simplificación de la información que fue adquirida inicialmente en niveles de gris hacia siluetas binarias. Tomando además tres posibles caminos para el procesamiento de los datos. Esto se observa en la Fig. 3.1.

En este trabajo de tesis se consideró tomar la parte inferior del cuerpo,

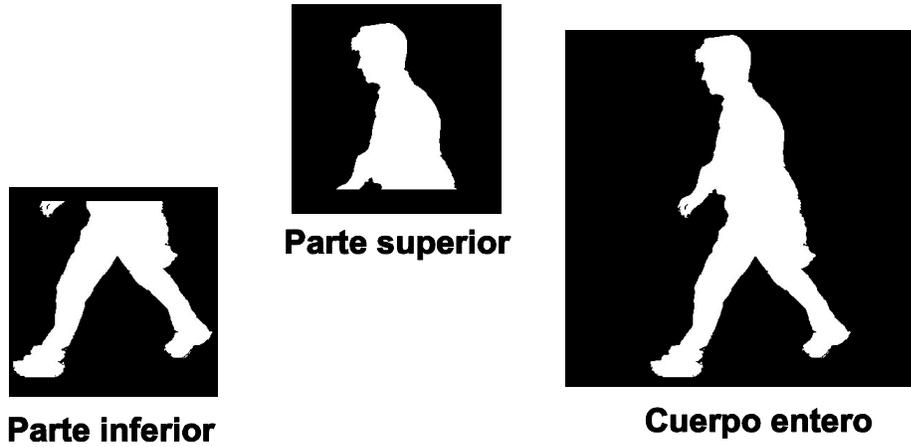


Figura 3.1: Tres tipos de imágenes de entrada para el algoritmo propuesto.

la cual contiene gran cantidad de información sobre el ciclo de la marcha. También, se reducen los tiempos de cómputo en el cálculo de momentos, como se muestra en el siguiente Cuadro 3.1.

<b>Orden <math>n</math></b>	<b>Parte inferior</b>	<b>Parte superior</b>	<b>Cuerpo entero</b>
2	0.47 s.	0.47 s.	1.34 s.
4	0.55 s.	0.55 s.	1.64 s.
6	0.66 s.	0.66 s.	1.95 s.
8	0.95 s.	0.95 s.	2.79 s.
10	1.42 s.	1.42 s.	4.28 s.
12	1.90 s.	1.90 s.	5.92 s.
14	2.57 s.	2.57 s.	7.36 s.
16	2.86 s.	2.86 s.	8.76 s.

Cuadro 3.1: Tiempos cómputo de los datos de estrada usados en la Fig. 3.1

La primera columna señala el orden de los momentos empleados en el cálculo. Las resoluciones de las imágenes de las siluetas de cuerpo entero y medio cuerpo para el cálculo de los  $A_{n,l}$  momentos es de  $512 \times 512$ px y  $300 \times$

300px, respectivamente. Los tiempos fueron tomados con una computadora con procesador de 2.8 Ghz y 1.00GB de RAM.

En la referencia [33] el cálculo de su descriptor de una imagen de  $250 \times 245$ px es de 0.97s, es el trabajo más reciente donde indican tiempos de cómputo. El método propuesto busca reducir o mantener este tiempo de procesamiento.

La mayoría de trabajos recientes que investigan la marcha como una medida biométrica para identificación humana han tenido lugar en el contexto del proyecto *HumanID* patrocinado por la *Defense Advanced Research Project Agency (DARPA)*. Cada una de las instituciones participantes ha establecido su propia base de datos de secuencias de la marcha. Una lista de las bases más conocidas de secuencias de la marcha se presentan en la Cuadro 3.2.

País	Base	URL	Sujetos	Parámetros
USA	USF/NIST	<a href="http://www.gaitchallenge.org">www.gaitchallenge.org</a>	71	Varias vistas, superficies, zapatos
USA	CMU	<a href="http://www.hid.ri.cmu.edu">www.hid.ri.cmu.edu</a>	25	Varias vistas, velocidad al caminar
U.K.	SOTON	<a href="http://www.gait.ecs.soton.ac.uk">www.gait.ecs.soton.ac.uk</a>	118	Varias vistas, interior / exterior
USA	SM	<a href="http://www.cpsc.ucalgary.ca/gait.html">www.cpsc.ucalgary.ca/gait.html</a>	6	—
USA	GATECH	<a href="http://www.cc.gatech.edu/cpl/projects/">www.cc.gatech.edu/cpl/projects/</a>	20	Varias vistas, interior / exterior
CHINA	CASIA	<a href="http://www.sinobiometrics.com">www.sinobiometrics.com</a>	20	Varias vistas
USA	MIT	<a href="http://www.ai.mit.edu/projects/gait/">www.ai.mit.edu/projects/gait/</a>	25	Tiempo

Cuadro 3.2: Bases de datos usadas en el contexto internacional para el reconocimiento de la marcha.

En este trabajo se tuvo acceso a las bases de MoBo y CASIA-A, mediante una solicitud formal con los responsables de las bases. Las bases mencionadas contienen al menos una base de vista frontal a una misma escala de la persona. En el Laboratorio de Visión por Computadora de la Universidad Politécnica de Tulancingo (UPT) se generó una base para la evaluación de la biometría de la marcha a tres diferentes distancias. Para el desarrollo de nuestro método fue necesario que las secuencias de imágenes pasaran por un pre procesamiento y así obtener un cuadro que contenga medio cuerpo de la cintura a los pies. La descripción de las bases se trata en las siguientes secciones así como su preprocesamiento.

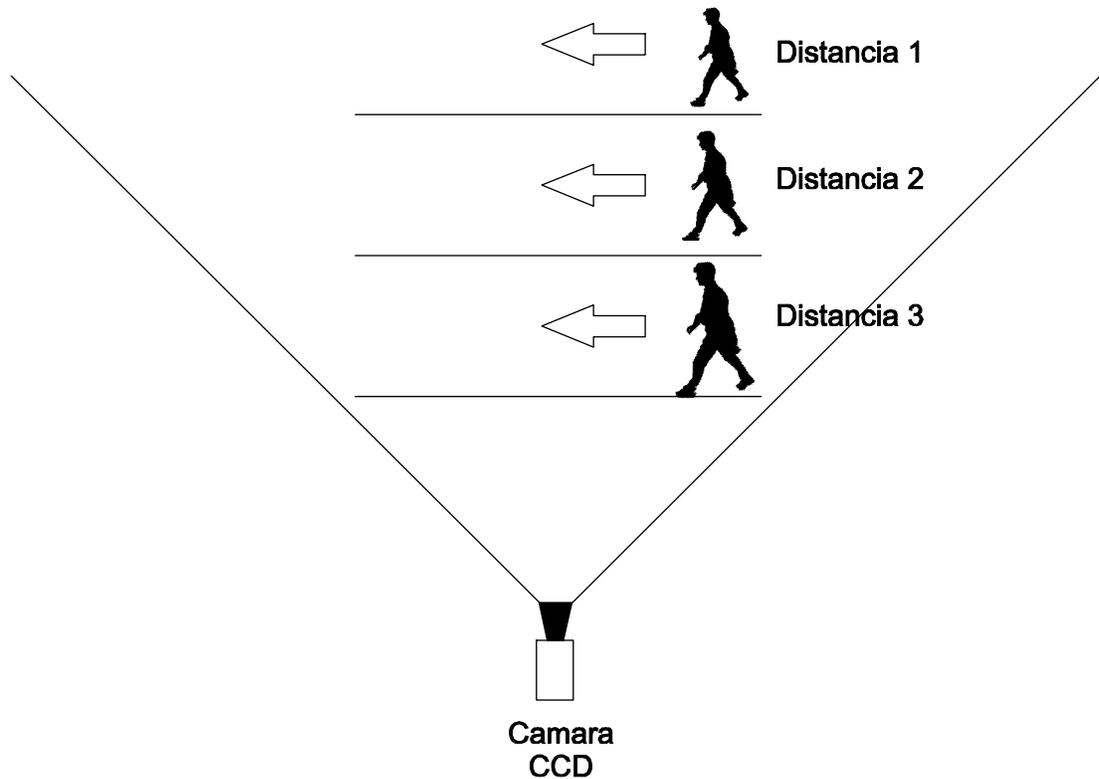
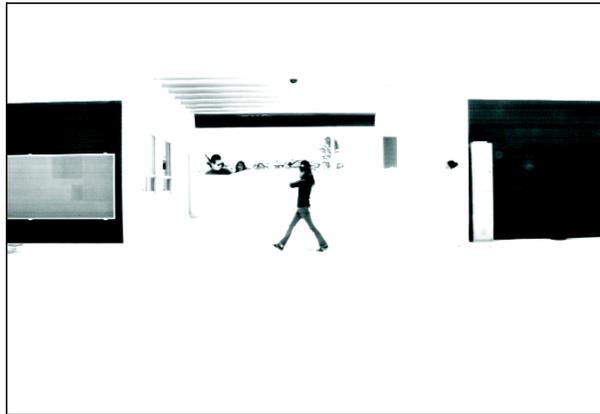


Figura 3.2: Esquema de la adquisición de imágenes.

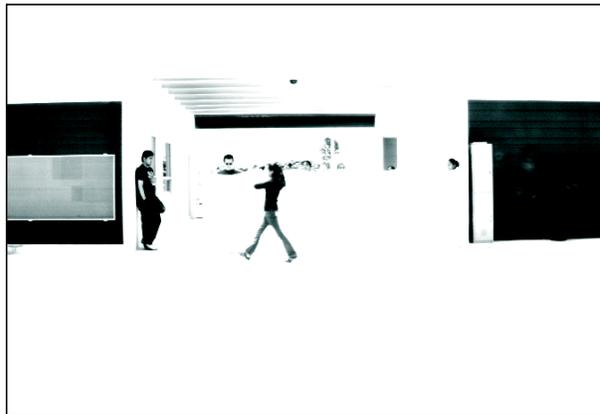
### 3.1. Adquisición de Imágenes (Base UPT)

Se capturó una base de datos para el análisis de la marcha y además como datos de entrada para probar el algoritmo propuesto en esta tesis. Las condiciones de captura de las secuencias de imágenes fueron en exteriores con iluminación natural y a tres diferentes distancias. Esto con el fin de probar el método de reconocimiento invariante a la escala. La base está compuesta por 15 personas, cada persona tiene seis secuencias, dos por cada distancia. Cada secuencia está compuesta por 50 marcos. En la Fig. 3.2 se muestra el esquema para generar la base y en la Fig. 3.3 se muestran algunos marcos de una secuencia a tres diferentes distancias.

**Distancia 1**



**Distancia 2**



**Distancia 3**

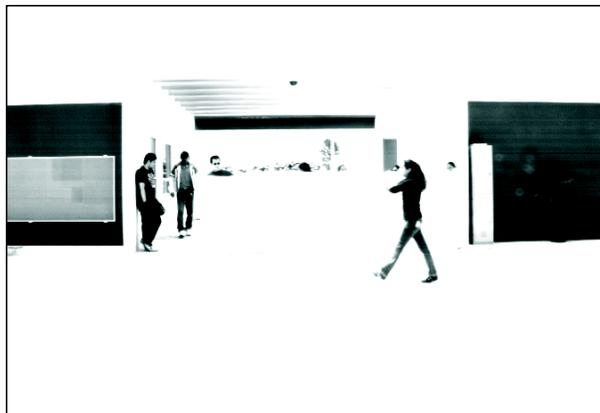


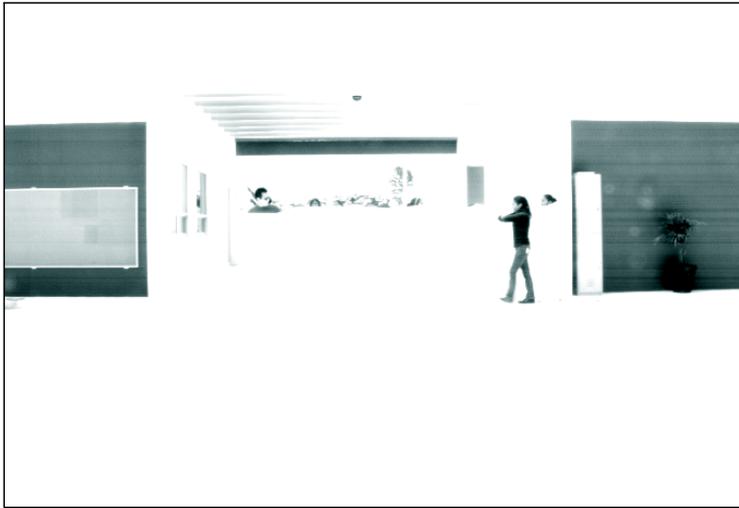
Figura 3.3: Marcos de tres secuencias a tres diferentes distancias. El número de marcos capturados en cada secuencia es 50.

El sistema óptico digital de captura consta de una cámara CCD monocromática con una resolución de 1392x1040px, la frecuencia de captura es de 16 marcos/segundo. Se empleó una tarjeta de adquisición de imágenes de la empresa Foresight modelo I-50. La captura es controlada a través de una interface grafica (Apéndice C) programada en el lenguaje de programación de Visual Basic 6. Las ampliaciones fueron realizadas con un sistema de lentes zoom, con rango focal 35-75mm. Las personas que colaboraron en la toma de imágenes portaban ropa oscura, en fondo blanco. Las imágenes fueron preprocesadas, pasando de 256 a 2 niveles de gris. El pre procesamiento se describe en la siguiente sección.

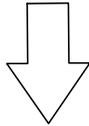
### **3.1.1. Preprocesamiento**

Al calcular los momentos de Jacobi-Fourier para el reconocimiento de la marcha, la imagen de medio cuerpo del individuo marchando tiene que estar dentro del círculo de radio unidad. Se usan los momentos geométricos para determinar el centroide de la imagen con la Ec.(2.14) y ubicar un nuevo origen de coordenadas.

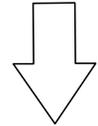
Para realizar el preprocesamiento se utilizo el lenguaje de programación MATLAB. En la Fig. 3.4 se muestran el proceso para el preprocesamiento. En la Fig.3.5 se muestran algunas imágenes procesadas de la secuencia de una persona.



**Imagen en niveles de gris.**



**Segmentación del medio cuerpo**



**Binarización de la imagen.**

Figura 3.4: Proceso del preprocesamiento.

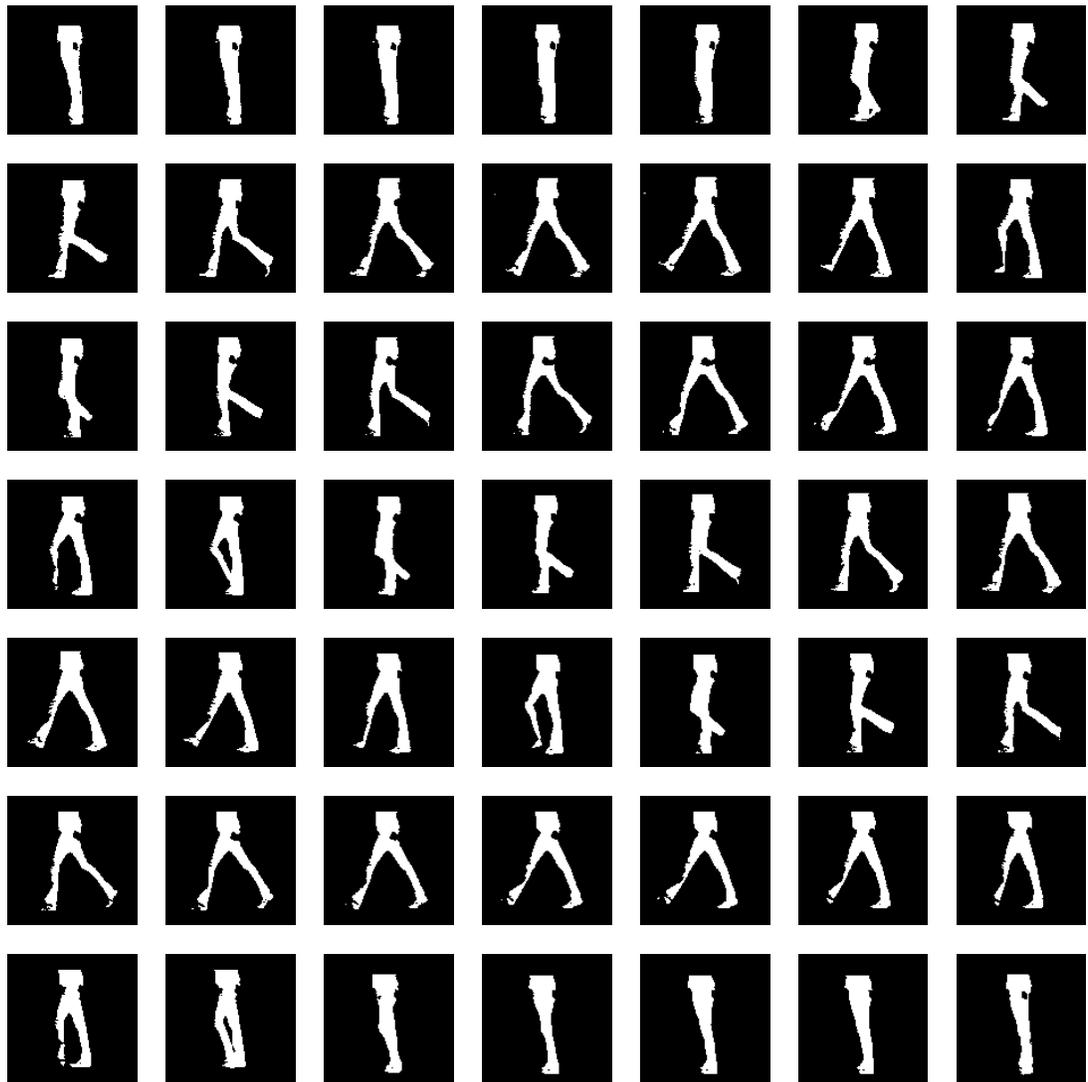


Figura 3.5: 49 Imágenes de una secuencia procesada.

La resolución de la imagen del individuo marchando depende de la escala de las imágenes capturadas. En la Fig. 3.6 se muestra las imágenes preprocesadas de tres escalas y su resolución.

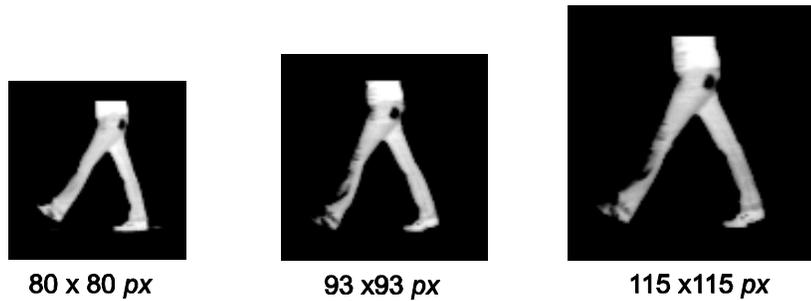


Figura 3.6: Resolución de imágenes procesadas.

## 3.2. Base MoBo

En 2001, la Universidad Carnegie Mellon (CMU) para avanzar en la investigación sobre la marcha humana como una biométrica, presentó una base de datos del movimiento del cuerpo (MoBo). Contiene 25 individuos marchando en una caminadora. Los sujetos realizan cuatro rutinas diferentes: caminar lento, rápido, inclinado y caminar con una pelota. Los sujetos fueron capturados con seis cámaras de alta resolución en color distribuidas uniformemente alrededor de la cinta caminadora, como se muestra en la Fig.3.7. Se capturaron más de 8000 imágenes por rutina. Las secuencias tienen una duración de 11 segundos, grabado en una frecuencia de imagen máxima de 30 marcos / segundo, cada secuencia grabada tiene 340 marcos. En la Fig.3.8 se muestra las seis vistas de la base de datos con las siluetas extraídas.

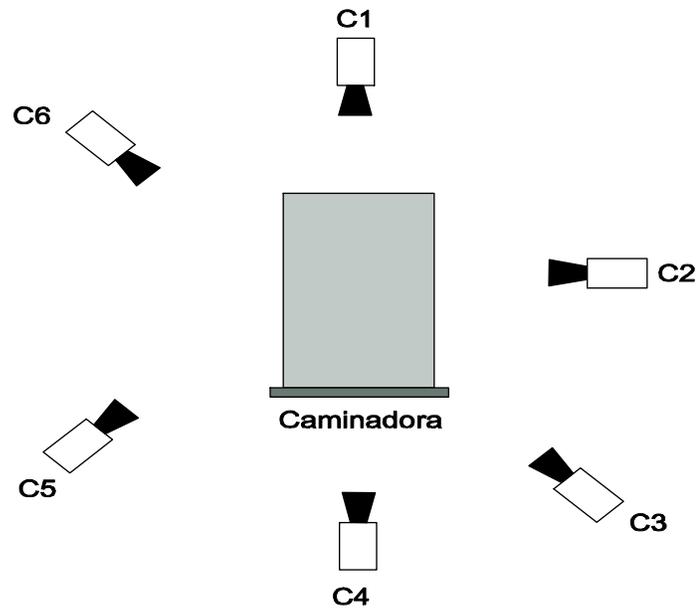


Figura 3.7: Posición de las cámaras de la captura de datos.

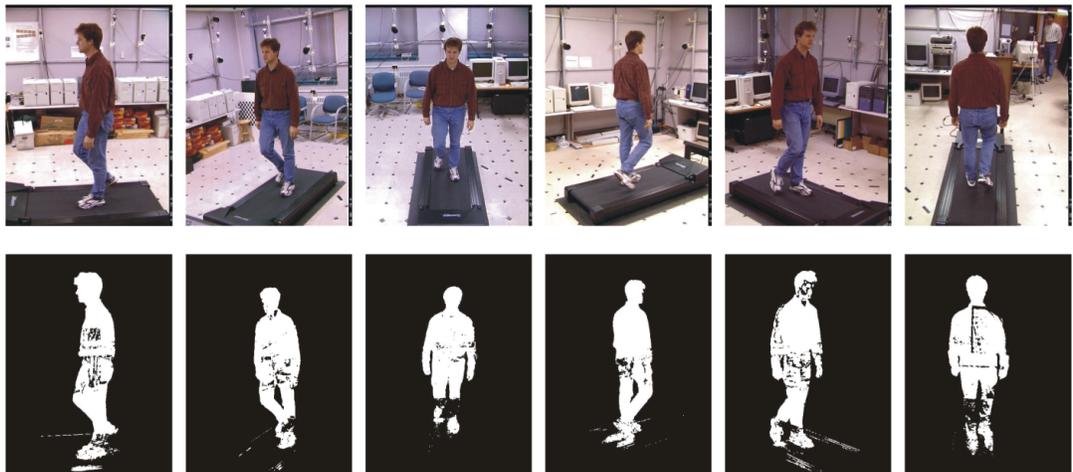


Figura 3.8: Vistas de la base MoBo con sus respectivas siluetas.

Antes de la grabación de los sujetos se les dio tiempo para acostumbrarse a caminar en la cinta. La mayoría de los sujetos, se sienten cómodos para caminar en el dispositivo después de practicar durante unos minutos. Las rutinas que fueron grabadas se describen a continuación:

### Caminar lento

La velocidad de la caminadora se ajustó a la velocidad de los sujetos. La Fig. 3.9 muestra ejemplos de las seis cámaras. Velocidad promedio de la marcha: 3.3 km/h.



Figura 3.9: Vistas de una secuencia caminando lento.

### **Caminar rápido**

Se les pidió que fueran más rápido, pero a un paso cómodo. La Fig. 3.10 muestra ejemplos de las seis cámaras. Velocidad promedio de la marcha: 4.5 km/h.



Figura 3.10: Vistas de una secuencia caminando rápido.

### **Caminar inclinado**

La caminadora se ajusto a una pendiente máxima de 15 grados. La velocidad se ajusto para que fuera cómoda para el sujeto. La Fig. 3.11 muestra ejemplos de las seis cámaras. Velocidad promedio de la marcha: 3.15 km/h.

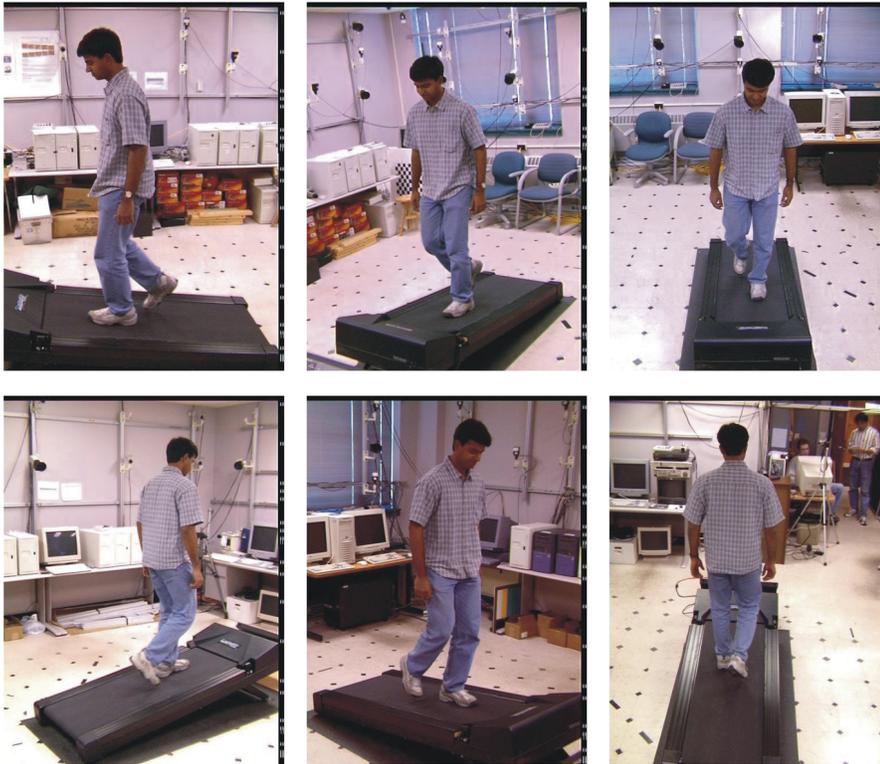


Figura 3.11: Vistas de una secuencia caminando con una inclinación.

### Caminar con una pelota

Finalmente, a cada persona se le pidió llevar una pelota frente a su cuerpo mientras camina a una velocidad cómoda. El objetivo de esta secuencia fue para inmovilizar los brazos y analizar cómo afecta a su forma de andar. La Fig. 3.12 muestra ejemplos de las seis cámaras. Velocidad promedio de la marcha: 3.28 km/h.

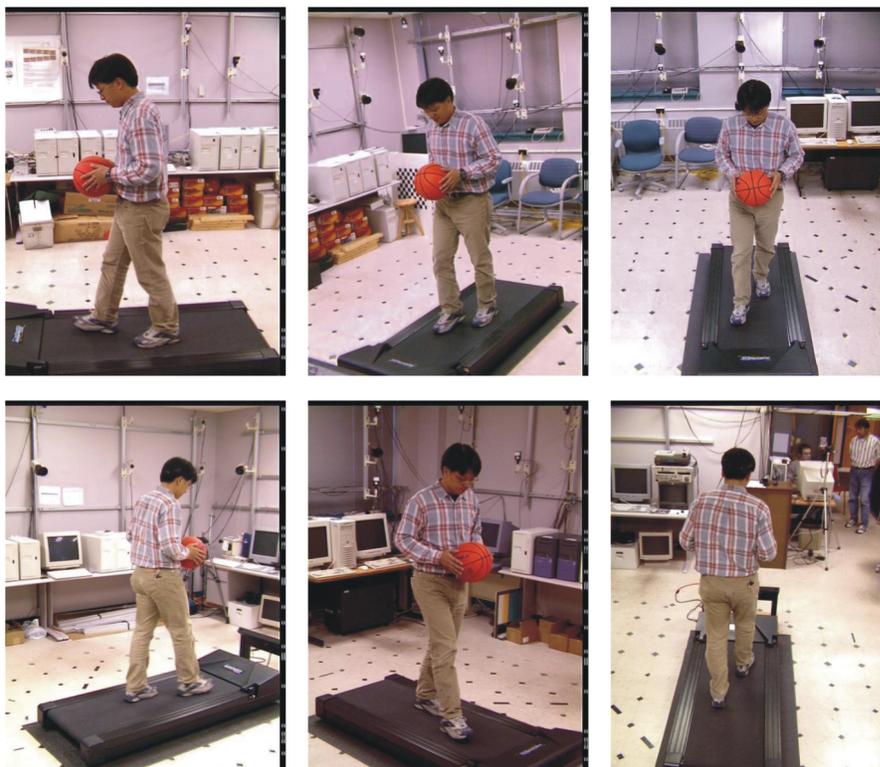


Figura 3.12: Vistas de una secuencia caminando con un balón.

### 3.2.1. Preprocesamiento

Como en la base de la UPT, es necesario sólo obtener un cuadro que contenga la parte inferior del cuerpo. La base MoBo proporciona imágenes binarias que contienen la silueta de la persona. Esto facilita el preprocesamiento. Se utilizó las secuencias de la cámara C2 con caminado lento. En la Fig. 3.13 se muestra el proceso de pre procesamiento de la base MoBo. La resolución de la imagen procesada es de  $120 \times 120$  px.

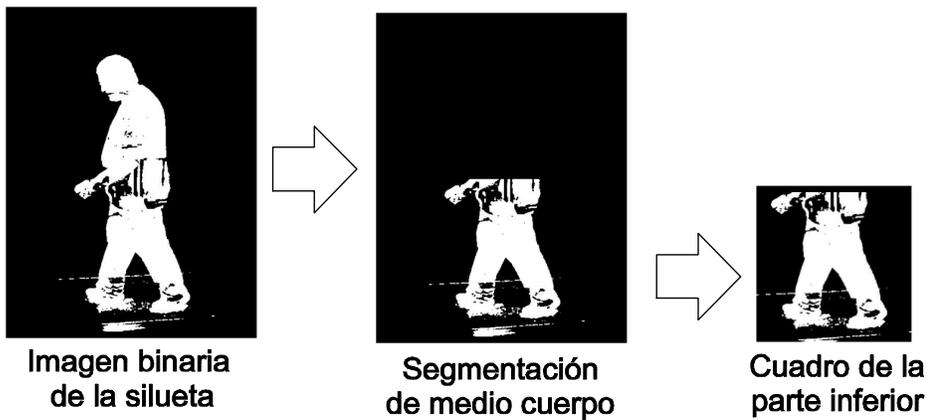


Figura 3.13: Preprocesamiento de la base MoBo.

La secuencia de cada persona contiene 340 marcos, para el algoritmo propuesto en este trabajo son necesarias por lo menos 4 secuencias. Tomamos 50 marcos como una secuencia, de esta forma tenemos 6 secuencias de cada persona. En la Fig. 3.14 se muestra una secuencia de imágenes procesadas de una persona.



Figura 3.14: Imágenes de una secuencia de la base MoBo procesada.

### 3.3. Base CASIA

Debido a que, el reconocimiento de la marcha ha sido un tema de investigación activo en los últimos años, el Centro de Biometría y de investigación sobre seguridad (CBSR) del Instituto de Automatización de la Academia China (CASIA) proporciona la base de datos CASIA, con el fin de promover la investigación relacionada con el reconocimiento y análisis de la marcha. En la base de datos CASIA hay tres conjuntos de datos: base A, base B (multivistas) y base C (por infrarrojos).

Base de datos A. Fue creada el 10 de diciembre del 2001, incluye 20 personas. Cada persona tiene 12 secuencias de imágenes, cuatro secuencia en tres direcciones, es decir, a 45 y 90 grados con el plano de la imagen perpendicular a la cámara. La longitud de la secuencia no es idéntica depende de la velocidad del caminante, pero hay un rango de 37 a 127 marcos. El tamaño de la base de datos es de aproximadamente 2.2 GB e incluye 19,139 imágenes. En la Fig. 3.15 se muestran algunas imágenes de la base.

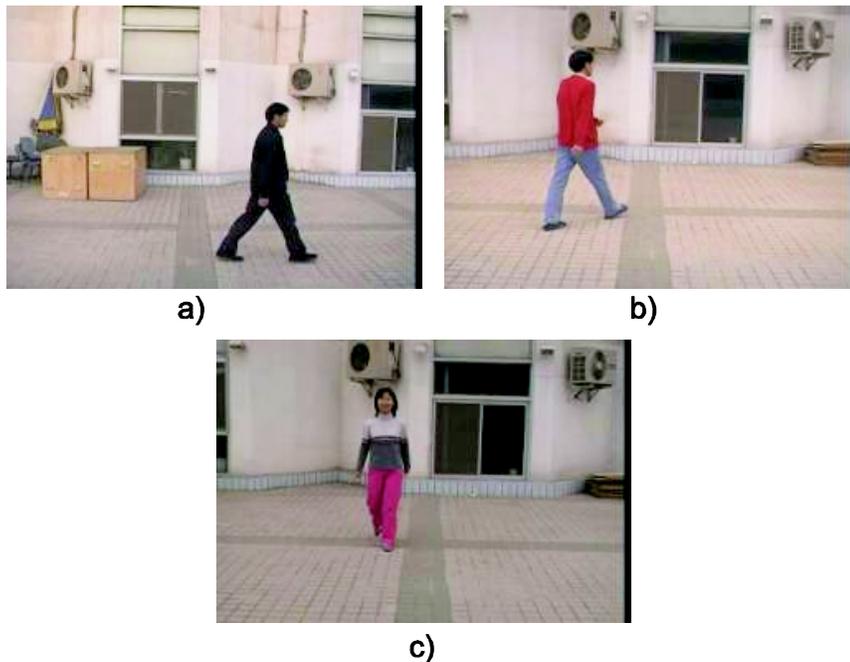


Figura 3.15: Base CASIA-A. a) Vista frontal. b) Vista a 45°. c) Vista a 90°.

Base de datos B. Es una base de datos grande de secuencias multivista, se creó en Enero del 2005, incluye 142 persona. Los datos de la marcha fueron tomados desde 11 puntos de vista, como se muestra en la Fig. 3.16. También se incluyo en la base, tomas con distintos cambios de ropa, esto se puede ver en la Fig. 3.17.

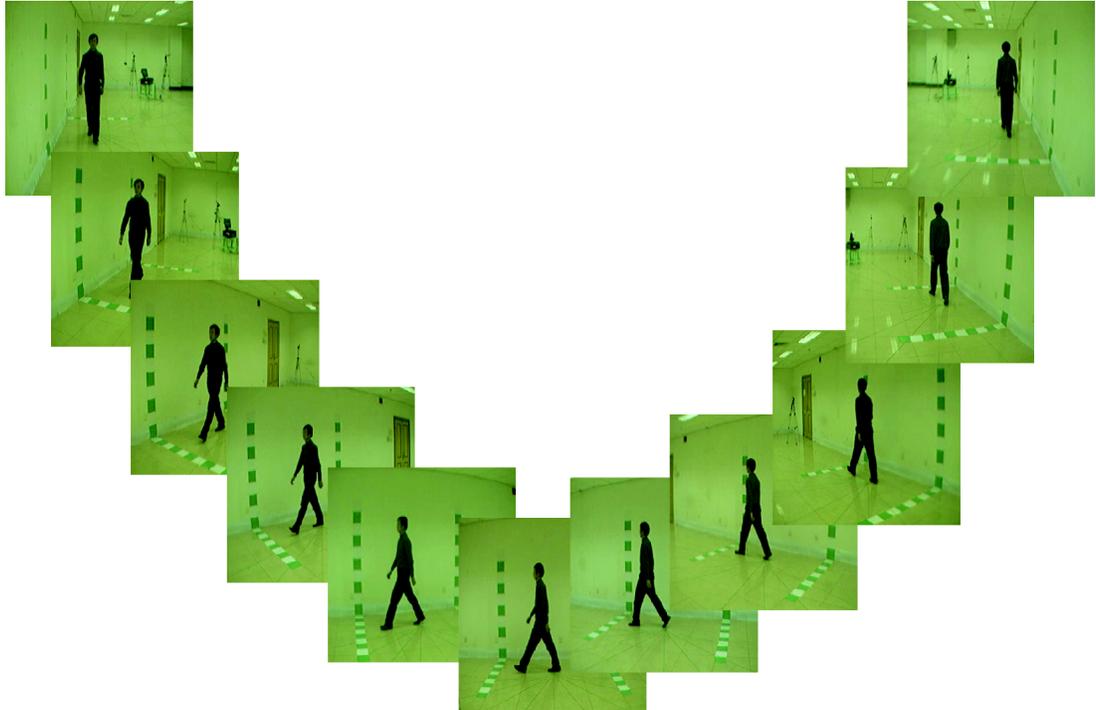


Figura 3.16: Base CASIA-B. Vistas de  $0^\circ$ ,  $18^\circ$ ,  $36^\circ$ , ...,  $180^\circ$ .

Base de datos C. El video fue capturado con una cámara infrarrojo en Junio del 2005. Contiene 153 sujetos y cuenta con 4 distintas formas de caminar; normal, lento, rápido y con una mochila. Las tomas fueron capturadas de noche. En la Fig. 3.18 se muestra algunas imágenes de la base.

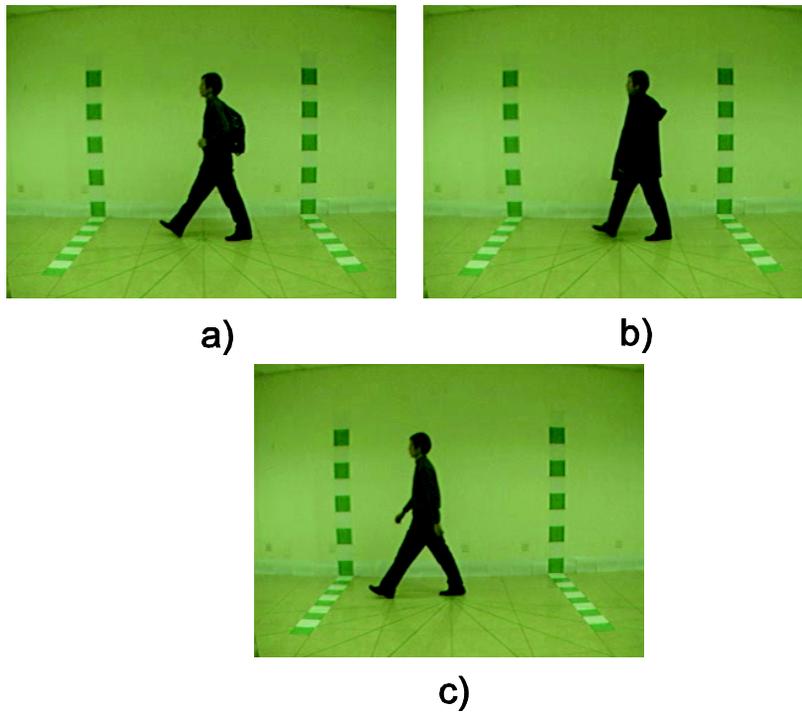


Figura 3.17: Base CASIA-B, con distintos tipos de ropa: a) mochila, b) abrigo y c) normal.

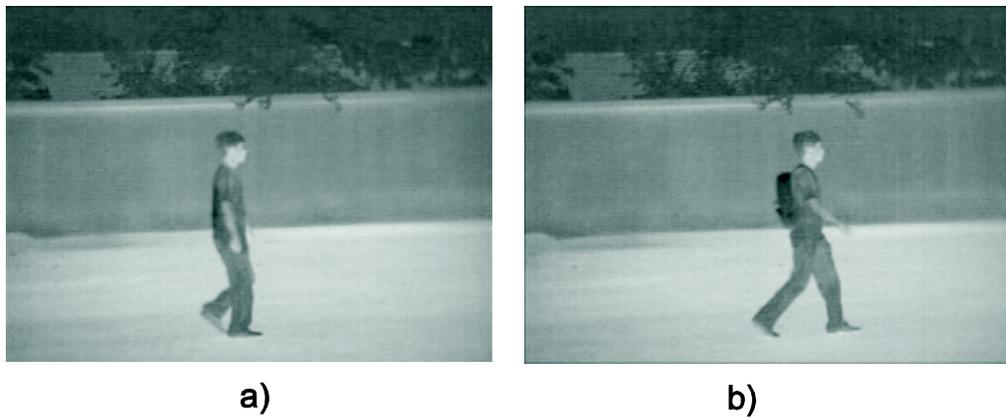


Figura 3.18: Base CASIA-C, a) caminando lento y b) caminando con mochila.

### 3.3.1. Preprocesamiento

Para este trabajo de tesis se utilizó las siluetas y la vista frontal de la base CASIA-A. El método propuesto necesita por lo menos 50 imágenes por secuencia. En la base CASIA-A hay un rango de 37 a 127 imágenes por secuencia. Solamente 12 sujetos cumplieron con la condición de secuencias mayores a 50 imágenes. Es necesario obtener un cuadro que contenga la parte inferior del sujeto. Se usan los momentos geométricos para determinar el centroide de la imagen con la Ec.(2.14) y definir un nuevo origen de coordenadas para obtener el cuadro centrado en el medio cuerpo. En la Fig.3.19 se muestra el proceso para el pre procesamiento.

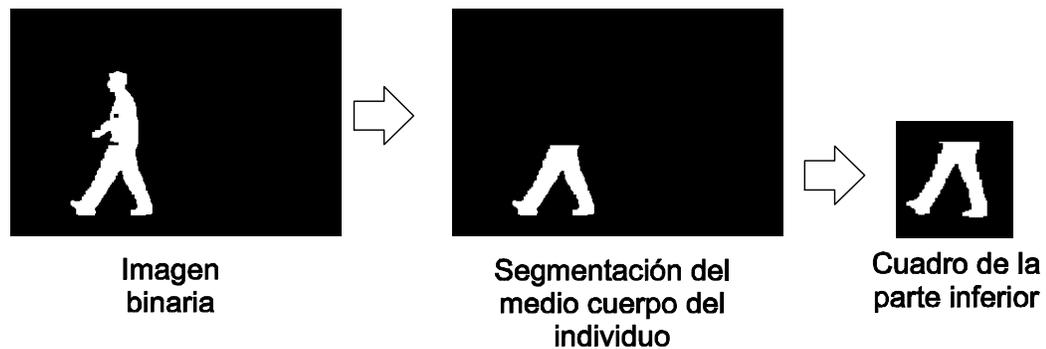


Figura 3.19: Preproceso para obtener el cuadro de la parte inferior.

En la Fig. 3.20 se muestra una secuencia de imágenes preprocesadas de cada persona.

Las bases de datos para las Figs. 3.5, 3.14 y 3.20 son una versión preprocesada de las bases originales. En ellas se tiene sólo una región de la imagen que es útil en la caracterización de la marcha.

En el siguiente capítulo se analizan cada una de estas secuencias usando el método de momentos de Jacobi-Fourier de una imagen descrito en el capítulo anterior.

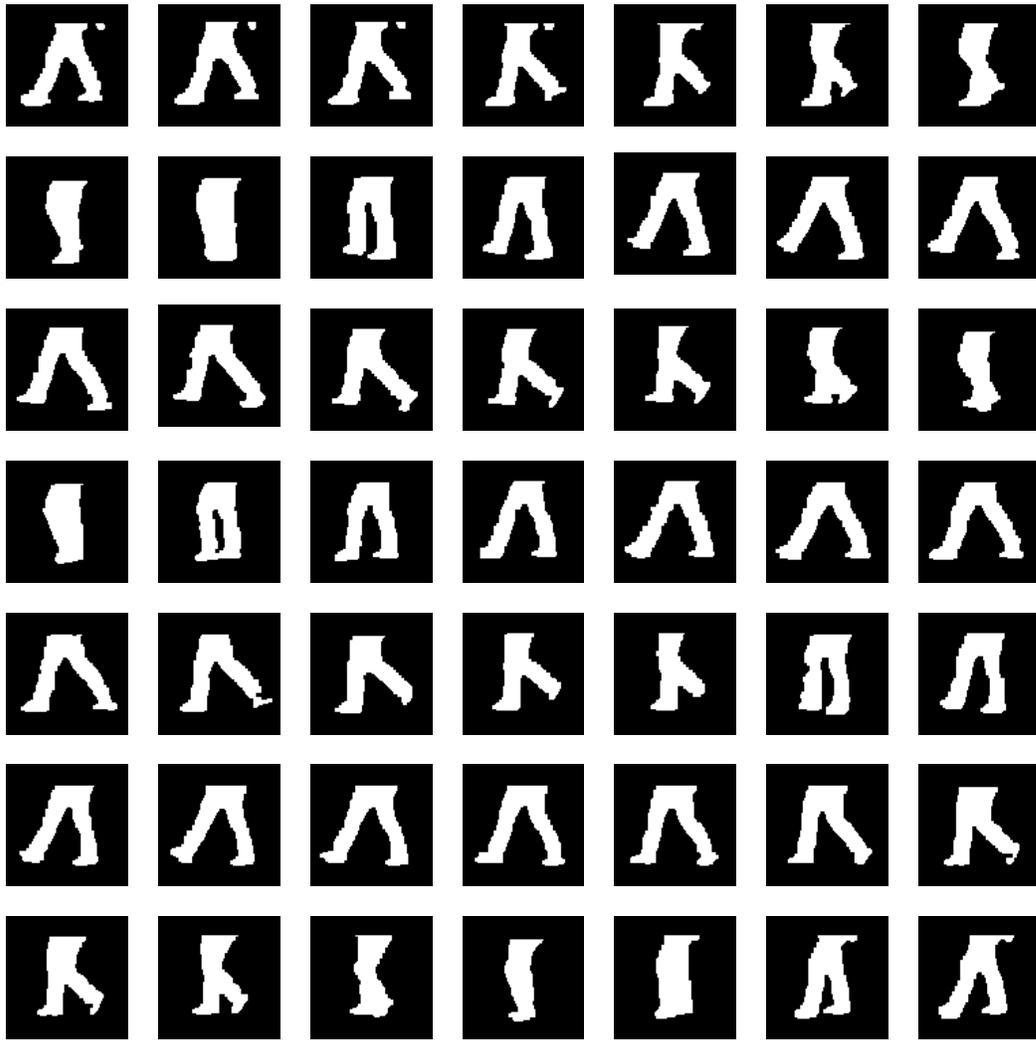


Figura 3.20: 49 Imágenes de una secuencia de la base CASIA-A procesada.



## Capítulo 4

# Clasificación de la marcha a partir de la extracción de descriptores invariantes

Existe un creciente interés en el procesamiento de secuencias de imágenes que motiva el desarrollo de técnicas para el análisis de la marcha basadas en la secuencia. El análisis de la marcha basado en modelos por lo general implica ajustar a un modelo que representa los diversos aspectos de la anatomía humana a los datos de vídeo para después extraer y analizar sus parámetros. Por el contrario, existe el enfoque estadístico el cual está basado en la forma de una silueta o la distribución estadística de los píxeles de la silueta y la forma en que ésta varía con el tiempo. En consecuencia, se desarrolló un método basado en la correlación de las historias de momentos de Jacobi-Fourier que permiten una descripción estadística, de la forma de la marcha de una persona. Para poder clasificar a cada individuo se propone un procedimiento matemático, de tal forma que este utiliza un número mínimo de descriptores para una correcta clasificación.

## 4.1. Historia de momentos

La *historia de momentos* es una sucesión cronológica de observaciones de momentos de imágenes. En la Fig. 4.1 se muestra el cálculo de los momentos  $|A_{n,m}^{f,s,p}|$  de  $f$  imágenes, de una secuencia  $s$ , de una persona  $p$ . A menudo los datos de las historias de momentos se examinan con la idea de encontrar un patrón permanente que se pueda aprovechar para la descripción de las imágenes que se desee analizar.

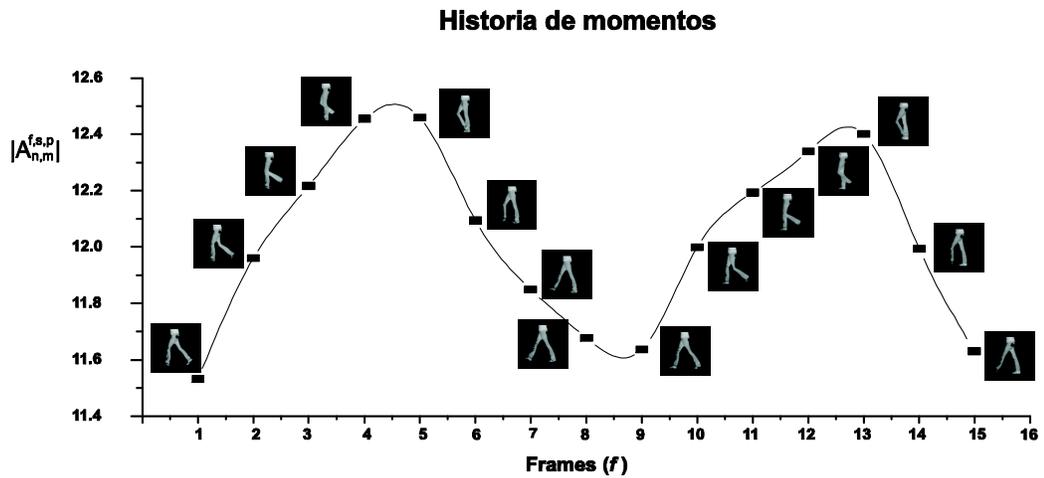


Figura 4.1: Historia de momentos de una secuencia de 16 imágenes.

También, facilita el estudio relacionado con el análisis del movimiento del cuerpo humano, al obtener parámetros temporales de la marcha.

En el apéndice B, se revisan brevemente las componentes de la marcha humana desde un punto de vista biomecánico.

La *historia de momentos* puede ser normalizada para ser invariante a la escala mediante el factor de escala  $g$  definida en la Ec.(2.14) y puede ser generalizado para la *historia de momentos*, buscando el área de referencia,

$$m_{00}^{\text{Referencia}} = \underset{f}{\text{máx}} \left( m_{00}^f \right), \quad (4.1)$$

donde  $m_{00}^f$  es el momento geométrico de orden cero de los marcos  $f$  de la *historia de momentos*. La Ec.(2.41) se puede redefinir para los momentos normalizados de Jacobi Fourier como,

$$A_{n,m}^f = \frac{m_{00}^{\text{Referencia}}}{m_{00}^f} \int_0^{2\pi} \int_0^g I \left( \frac{r}{g}, \theta \right) P_{n,m} \left( \frac{r}{g}, \theta \right) r dr d\theta = \frac{m_{00}^{\text{Referencia}}}{m_{00}^f} A'_{n,m}{}^f \quad (4.2)$$

en donde  $A'_{n,m}{}^f$  son los momentos de Jacobi-Fourier para cada marco  $f$ .

## 4.2. Problema de poner en fase

No es desconocido que la extracción de la marcha característica de una persona es un problema difícil. El punto de partida de cada muestra de las *historias de momentos* es diferente para cada persona. Este tipo de problema, de poner en fase la *historia de momentos* se puede resolver utilizando técnicas heurísticas como son los *Algoritmos Genéticos*, una breve revisión de esta técnica se lleva a cabo en el Apéndice A. El objetivo es buscar maximizar el coeficiente de correlación de una *sección de referencia* con varias *historias de momentos* de la misma clase. Una clase incluye las secuencias de imágenes de una misma persona. El procedimiento consiste en seleccionar un intervalo  $H$  de la *historia de momentos*  $|A_{n,m}^{f,s,p}|$  para cada una de las secuencias  $s$  y persona  $p$ , el inicio de la sección esta dado por  $w_s$  donde  $s = 1, 2, \dots, S$ . En la Figura 4.2 se muestra el intervalo seleccionado de una secuencia.

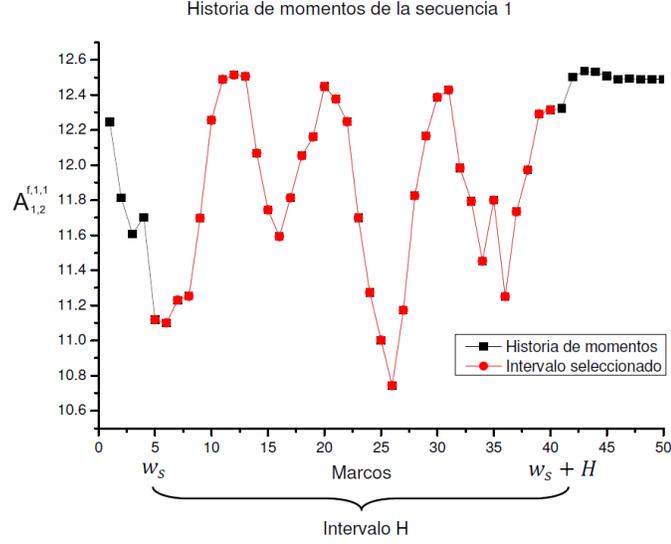


Figura 4.2: Historia de momentos e intervalo seleccionado  $[w_s, w_s + H]$ . En donde  $H$  es un valor arbitrario.

La *sección de referencia* se calcula con el promedio de secuencias con valor inicial  $w_s$ , esto es;

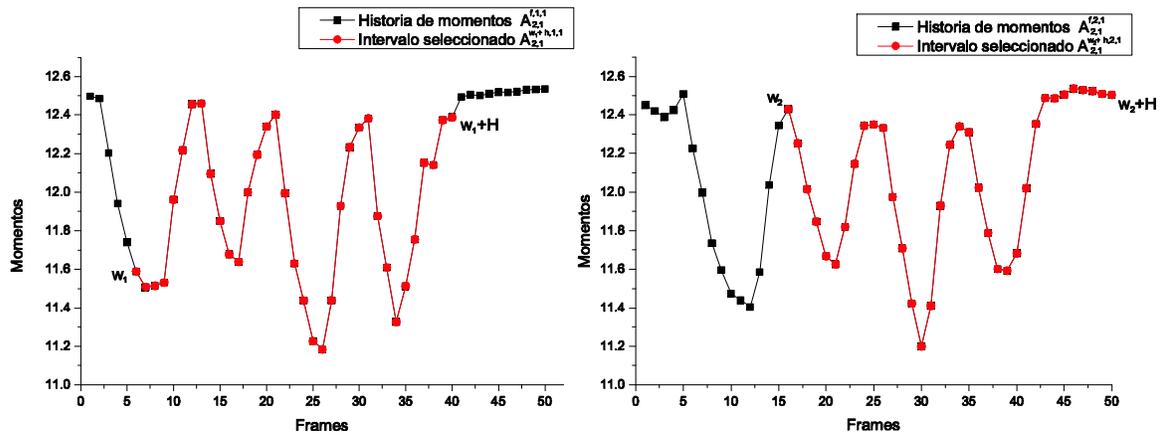
$$P_h = \frac{1}{S} \sum_{s=1}^S A_{n,m}^{w_s+h,s,p}, \quad (4.3)$$

donde  $h = 0, 1, , 2 \dots H$ . En la Fig.4.3 se muestran dos historias de momentos para  $s = 1$  y  $s = 2$ .

En la Fig.4.3 se muestra como obtener la *sección de referencia*, a partir de *historia de momentos* de las secuencias  $s = 1$  y  $s = 2$ .

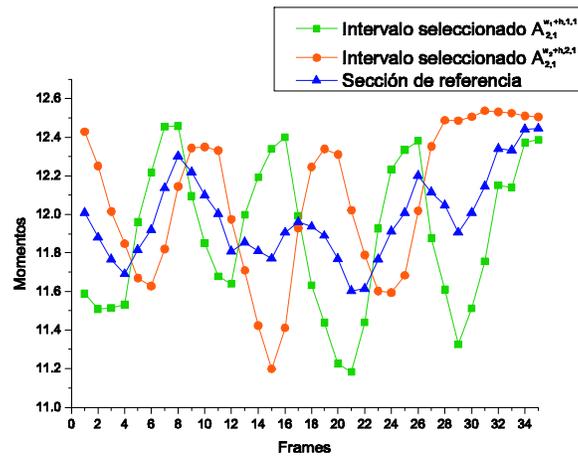
Se busca secciones de las secuencias que estén correlacionadas con la *sección de referencia* (o promedio). La medida de asociación lineal entre  $P_h$  y  $A_{n,m}^{w_s+h,s,p}$  se estima con el coeficiente de correlación  $c_s^p$ , definido como;

$$c_s^p = \frac{n \sum P_h \cdot A_{n,m}^{w_s+h,s,p} - \sum P_h \sum A_{n,m}^{w_s+h,s,p}}{\sqrt{[n \sum P_h^2 - (\sum P_h)^2] \left[ n \sum \left( A_{n,m}^{w_s+h,s,p} \right)^2 - \left( \sum A_{n,m}^{w_s+h,s,p} \right)^2 \right]}} \quad (4.4)$$



a)

b)



c)

Figura 4.3: a) Historia de momentos  $s = 1$ . b) Historia de momentos  $s = 2$ .  
c) Sección de referencia para la secuencias  $s = 1$  y  $s = 2$ .

La función objetivo está dada por:

$$\text{máx } Z^p = \sum_{s=1}^S c_s^p(w_s) \quad (4.5)$$

Al solucionar el problema de la Ec. (4.5) se obtienen los valores de las posiciones  $w_s$  iniciales de las historias de momentos. Esto mediante un *Algoritmo Genético*, cuyo pseudocódigo se muestra en el Algoritmo 1.

---

**Algoritmo 1** Algoritmo Genético para poner en fase historias de momentos

---

**Inicio:** se genera una población inicial aleatoria.

**repite**

**Aptitud:** se evalúa la función objetivo  $Z^p$  de cada individuo.

**Selección:** se hace al azar dependiendo la aptitud.

**if** Probabilidad de cruza  $\geq$  Aleatorio **then**

        Se realiza la cruza.

**else**

        Se deja como esta.

**end if**

**if** Probabilidad de mutación  $\geq$  Aleatorio **then**

        Se deforma la nueva descendencia.

**else**

        Se deja como esta.

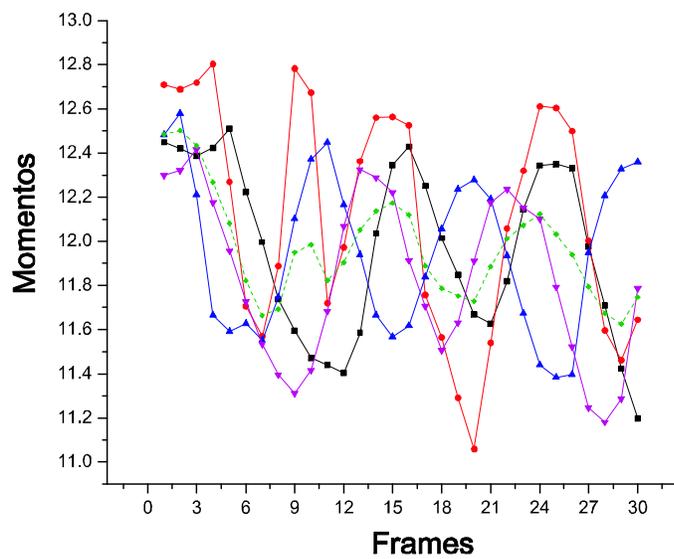
**end if**

**Reemplazo:** la nueva población generada reemplaza a la vieja población.

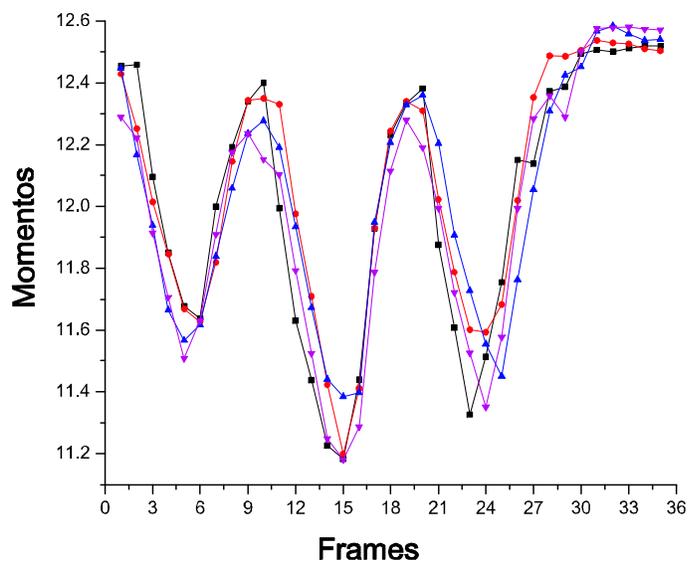
**hasta** contador  $\geq$  número de iteraciones

---

Al obtener la solución con *Algoritmos Genéticos* y graficar el inicio de los intervalos  $w_s$  de las diferentes secuencias, se puede observar que el intervalo de las *historias de momentos* queda en fase. En la Fig.4.4a se muestra las historias de momentos como se encuentran inicialmente y en la Fig.4.4b se muestra las historias de momentos en fase.



a)



b)

Figura 4.4: Historia de momentos. a) En forma inicial. b) Puestas en fase.

### 4.3. Descriptores de la marcha

El problema de poner en fase las *historias de momentos* de las diferentes secuencias de un mismo individuo ayuda a que la estadística de las secuencias de una misma clase sea similar, permitiendo diferenciarse entre otras clases y así poder identificarse con un margen de error menor. En la Fig. 4.5 se muestra la historia de momentos de dos individuos en tres secuencias distintas. Se puede observar que existe una similitud entre los perfiles de las secuencias de un mismo individuo, pero son diferentes entre individuos.

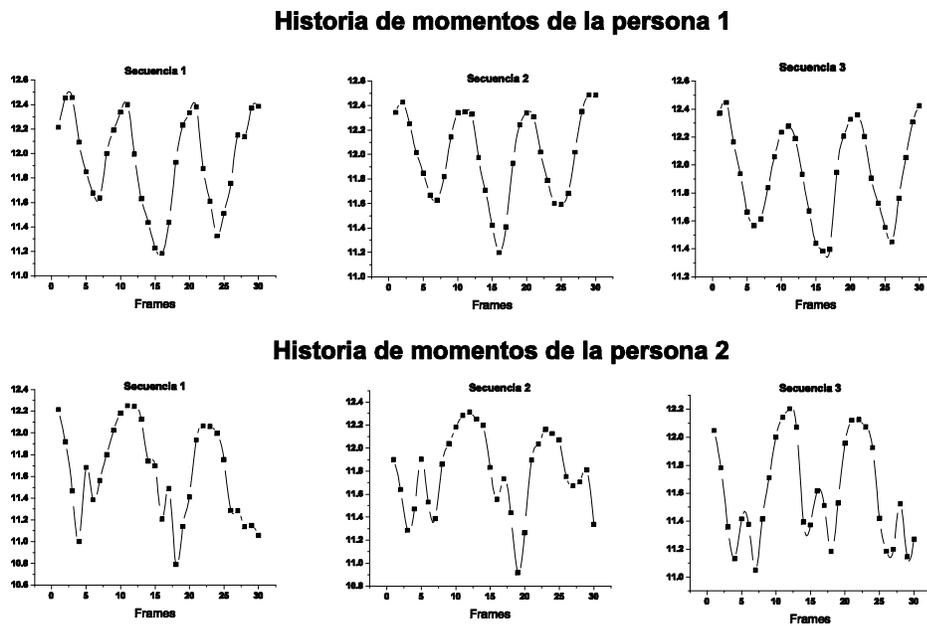


Figura 4.5: Historia de momentos de dos personas.

Se define la media de la *historia de momentos* como una medida descriptiva de cada historia, mediante la expresión,

$$D_{n,m}^{s,p} = \frac{1}{H} \sum_{h=0}^H |A_{n,m}^{w_s+h,s,p}|, \quad (4.6)$$

donde  $A_{n,m}^{f,s,p}$  son los valores de los momentos de la *historia de momentos* y  $w_s$  es el valor inicial óptimo de cada secuencia, que se encuentra en fase con el resto de las secuencias.

En la Fig.4.3a se mostró la historia de momentos  $A_{n,m}^{f,s,p}$  de la secuencia  $s = 1$ , la persona  $p = 1$ , del orden radial  $n = 2$ , orden armónico  $m = 1$ , tomando en cuenta que  $w_1 = 5$  y  $H = 35$ . Se puede calcular el descriptor  $D_{2,1}^{1,1}$ , mediante la expresión,

$$D_{2,1}^{1,1} = \frac{1}{35} (|A_{1,2}^{5+0,1,1}| + |A_{1,2}^{5+1,1,1}| + |A_{1,2}^{5+2,1,1}| + \dots + |A_{1,2}^{5+35,1,1}|).$$

El descriptor  $D_{n,m}^{s,p}$  se utiliza como una estimación puntual de la marcha de cada individuo.

Existe una gran variedad de descriptores dependiendo del orden  $n$  y la repetición  $m$  en el cálculo de los momentos de Jacobi-Fourier. Asimismo como cuando se combinan los valores de los parámetros  $\alpha$  y  $\beta$ . Este espacio de búsqueda de los mejores descriptores  $D_{n,m}^{s,p}$  puede quedar acotado por dos vertientes diferentes. La primera, los valores asignados a los parámetros  $\alpha$  y  $\beta$ . Y la segunda por el valor del orden radial  $n$ , de los polinomios elegidos. En la Fig. 2.5 del capítulo dos, se muestra como, dependiendo de algunos parámetros  $\alpha, \beta$  y  $n$ . Las oscilaciones de los polinomios de Jacobi podrían ser muy severas y volverse inestables, ocasionando un descriptor  $D_{n,m}^{s,p}$  que no pueda ser adecuado para clasificar. Por ejemplo si  $\alpha = \beta = 4$  y  $n \geq 20$ , implica que  $D_{20,20}^{s,p}$  no es adecuada para clasificar.

## 4.4. Mapa de la medida discriminatoria $Q$

La selección de descriptores es un paso crucial en cualquier sistema de reconocimiento de formas. Los descriptores que recuperan aspectos globales de un objeto son a menudo sensibles al ruido [41].

Para saber el comportamiento del descriptor  $D_{2,1}^{s,1}$ , se realizó una simulación de 100 secuencias, cambiando aspectos de forma aleatoria como: ruido Gaussiano [42] y cambio de escala [42]. En la Fig. 4.6 se muestra que algunos descriptores tiene una distribución de probabilidad normal. Es decir, el comportamiento de estos descriptores esta normalmente distribuido.

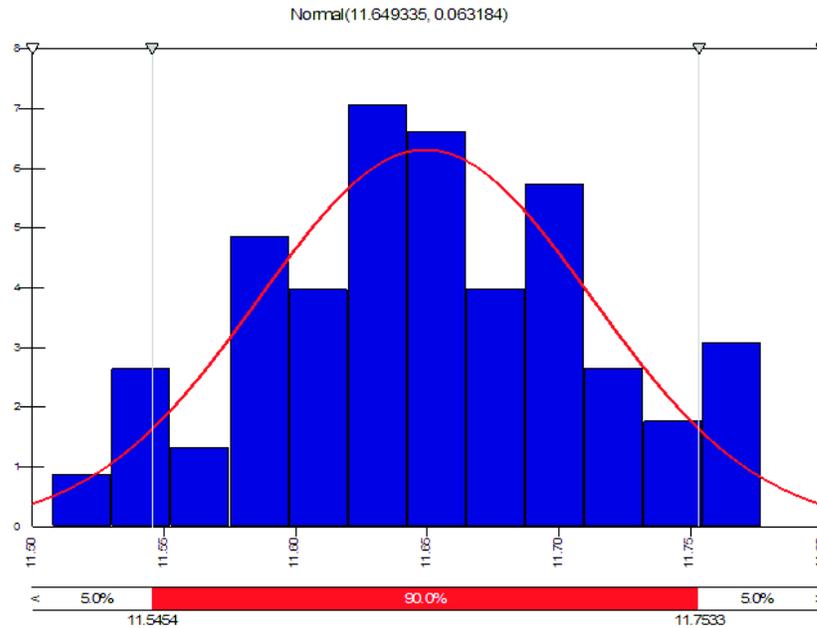


Figura 4.6: Ajuste de la distribución de probabilidad Normal de la simulación de 100 secuencias para  $D_{2,1}^{s,1}$ .

Un parámetro que puede cambiar la distribución de los descriptores, es el orden radial  $n$  de los polinomios de Jacobi. En la Fig. 4.7 se muestra la distribución de probabilidad que mejor se ajusta al descriptor de orden  $n = 30, m = 1$ . Una distribución de probabilidad que no se ajusta a una distribución Normal, por ejemplo el polinomio de Jacobi de orden  $n = 30$  el

cual presenta oscilaciones abruptas en algunas regiones en el intervalo radial  $[0,1]$ . Esto se observa en la Fig. 2.5.

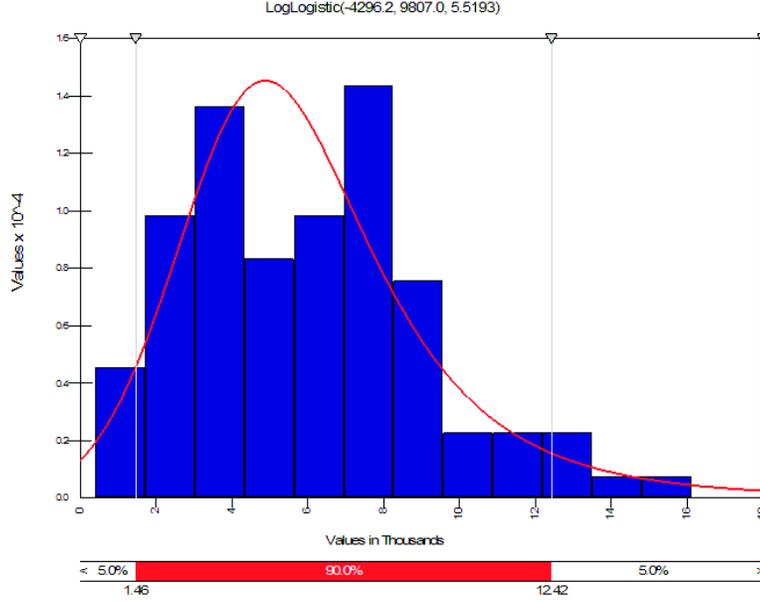


Figura 4.7: Ajuste de la distribución de probabilidad Normal de la simulación de 100 secuencias para  $D_{30,1}^{s,1}$ .

Es importante diseñar procedimientos de alto rango de discriminación de tal forma que la varianza intraclases sea pequeña y la separación entre clase sea grande. Shen [43] desarrolló una medida discriminatoria para dos clases, para encontrar la máxima separación entreclase y la mínima varianza intraclase. El presente trabajo de tesis, propone una medida heurística que generaliza el método de Shen para múltiples clases. Busca la máxima separación entreclases de todas las posibles combinaciones y la mínima varianza intreclase. La medida discriminatoria para múltiples clases que se propone en este trabajo (considerando que cada clase corresponde a una persona  $p$ ) está dada por,

$$Q_{n,m} = \sum_{p=1}^{P-1} \sum_{p'=p-1}^P \left| \mu_{n,m}^p - \mu_{n,m}^{p'} \right| - \sum_{p=1}^P \sigma_{n,m}^p \quad (4.7)$$

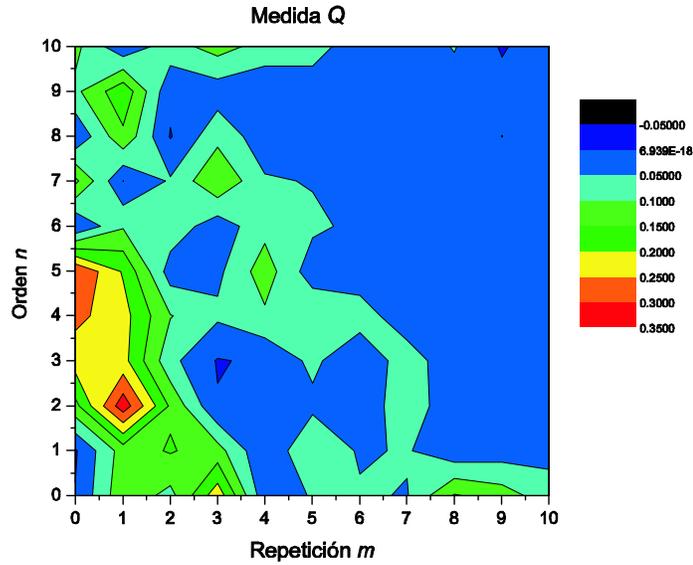


Figura 4.8: Mapa  $Q$ , para clasificar dos clases.

donde la  $\mu_{n,m}^p$  y  $\sigma_{n,m}^p$  son respectivamente la media y la desviación estándar para cada descriptor  $D_{n,m}^{s,p}$ . La media y la desviación estándar pueden ser estimadas por una cantidad suficiente de secuencias  $s$  para las  $p$  personas. La medida discriminativa  $Q$  proporciona una selección heurística de los mejores descriptores, suponiendo a priori que tienen una distribución normal como la indicada en la Fig. 4.6. En la Fig. 2.6a se muestran 10 imágenes de dos personas a diferentes escalas, con estas imágenes de entrada se puede buscar los mejores descriptores. En la Fig. 4.8 se muestra un mapa  $Q$  con  $n = m = 1, 2, 3, \dots, 10$ , se puede observar que el valor más alto se encuentra en la combinación  $n = 2$  y  $m = 1$ . Los valores *más altos* del mapa  $Q$  indican que estos, están más separados entreclases y agrupados intraclase, como se observa en la Fig. 4.9a. Contrariamente, si graficamos la combinación de los valores de  $n$  y  $m$  de los valores más bajos del mapa  $Q$ , los valores de los momentos de Jacobi-Fourier de las clases se juntarían y sería imposible clasificar, esto se puede observar en la Fig. 4.9b. En la Fig. 4.10 se muestra cómo al graficar la combinación de los valores más altos, siguen agrupados y a su vez están separados entre clases.

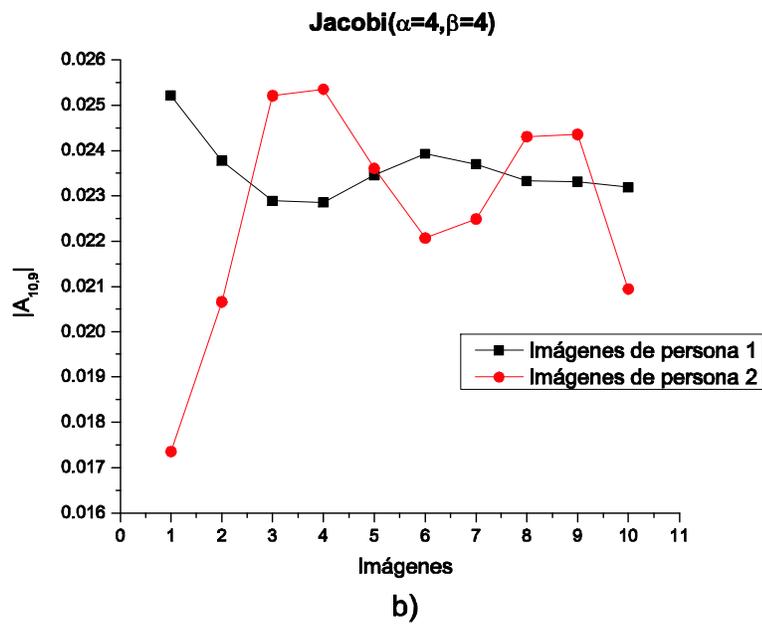
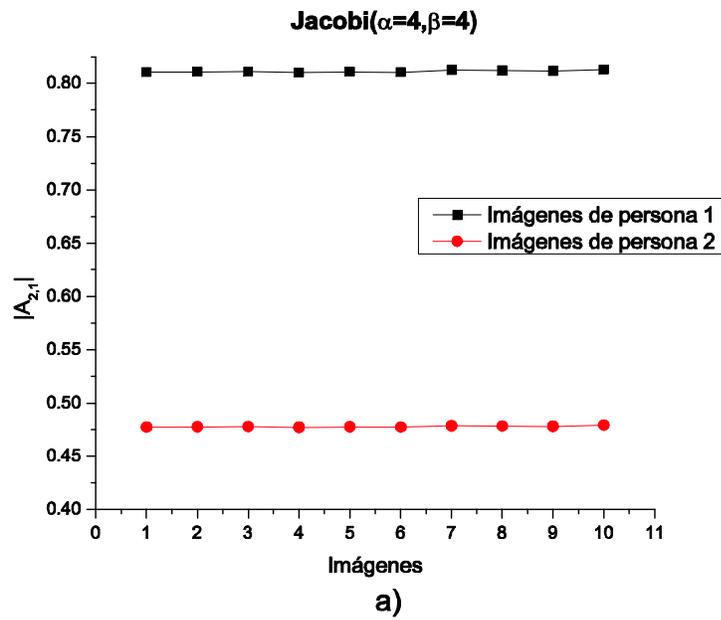


Figura 4.9: a) Gráfica del mejor descriptor  $n = 1, m = 2$  para  $\alpha = \beta = 2$ . b) Gráfica del peor descriptor  $n = 10, m = 9$  para  $\alpha = \beta = 2$ .

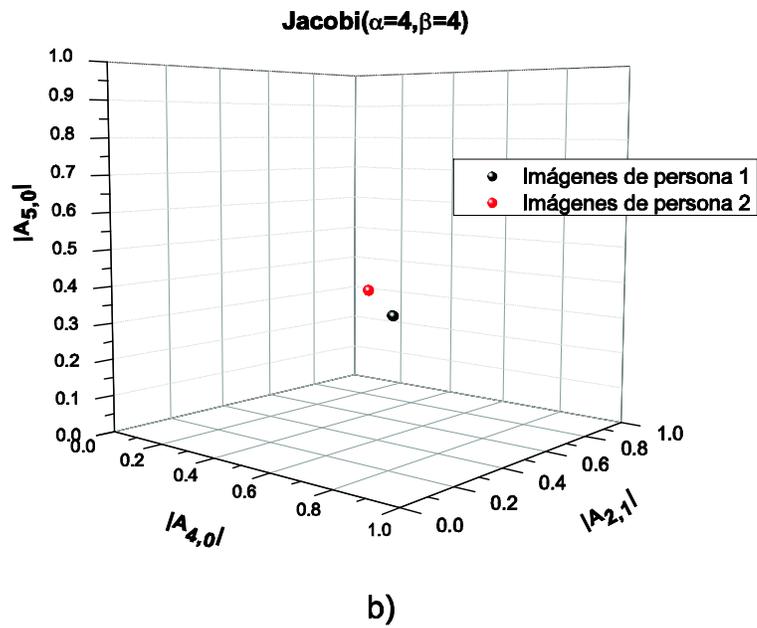
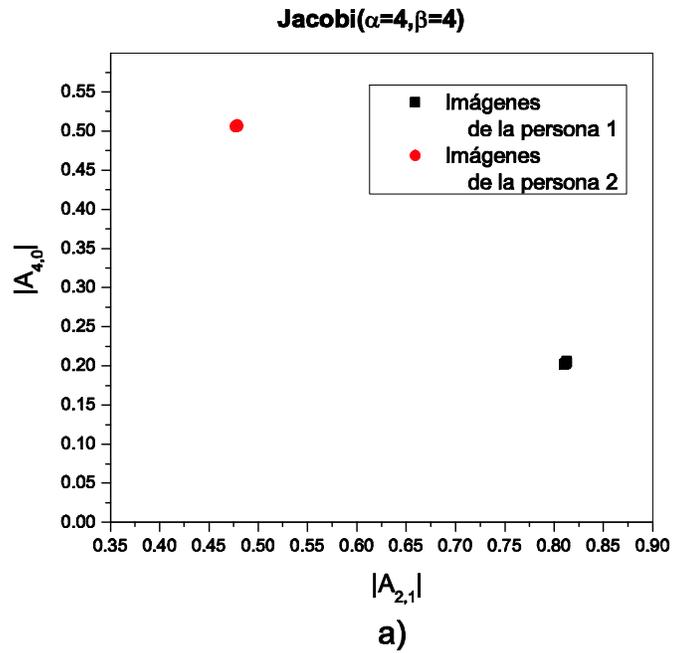


Figura 4.10: a) Gráfica de los dos mejores descriptores. b) Gráfica de los tres mejores descriptores. Las clases siempre se separan y están agrupadas.

## 4.5. Método de clasificación

La información contenida en los datos de la *historia de momentos* con orden  $n$  y repetición  $m$ , pueden reflejar atributos únicos de la forma de la secuencia y se diferencian entre sí de cierta manera. Sin embargo la cantidad de descriptores encontrados puede ser muy grande. Esto en parte debido a la falta de conocimientos sobre qué elementos de la marcha son más adecuados para distinguir entre personas.

El análisis que se presenta en este trabajo consiste en reducir considerablemente la cantidad de descriptores calculados de la marcha, para obtener un mayor porcentaje de clasificación con un número menor de descriptores.

Se combinan dos métodos de clasificación: la distancia mínima y correlación. Estos métodos tienen condiciones similares en los datos de entrada y en la forma de clasificar aunque tengan enfoques diferentes. La distancia mínima busca la mínima distancia entre la referencia y la media de las clases y la correlación indica la relación lineal entre dos variables y su dependencia.

A continuación se definen estos métodos para la clasificación de los descriptores de las historias de momentos de la marcha. Sea  $\{D_k | k = 1, 2, 3, \dots, K\}$  el descriptor seleccionado, donde  $K$  es el número de descriptores que se consideran para clasificar. Cuando un descriptor objeto con vector  $X = [X_1, X_2, \dots, X_K]$  ha sido clasificado, la distancia mínima o la correlación de  $X$  se lleva a cabo todas las muestras de las personas  $p$ . La distancia entre  $X$  y la  $p$ -ésima muestra, se calcula como:

$$d(X, Personap, K) = \sum_{k=1}^K \left[ \frac{(X_k - \mu_k^p)^2}{\sigma_k^p} \right] \quad (4.8)$$

La correlación entre  $X$  y la  $p$  - ésima muestra, está dada por

$$\text{Corr}(X, \text{Persona}^p, K) = \sum_{k=1}^K \left[ \frac{(X_k - \mu_K^X) ( \text{Persona}_k^p - \mu_K^p )}{\sigma_K^X \sigma_K^p} \right], \quad (4.9)$$

donde:

$$\begin{aligned} \mu_N^p &= \frac{1}{K} \sum_{k=1}^K \mu_k^p, \quad \sigma_N^p = \frac{1}{K} \sum_{k=1}^K \sigma_k^p, \\ \mu_N^X &= \frac{1}{K} \sum_{k=1}^K \mu_k^X, \quad \sigma_N^X = \frac{1}{K} \sum_{k=1}^K \sigma_k^X. \end{aligned}$$

Al combinar los dos métodos se pretende buscar la mínima distancia y la máxima correlación, se calcula como,

$$d(X, \text{Persona}^p, K) = \sum_{k=1}^K \left[ \frac{(X_k - \mu_k^p)^2}{\sigma_k^p} - \frac{(X_k - \mu_K^X) ( \text{Persona}_k^p - \mu_K^p )}{\sigma_K^X \sigma_K^p} \right]. \quad (4.10)$$

El objetivo de  $X$  es clasificar dentro de la persona  $p^*$  la cual satisface la condición de los  $K$  descriptores:

$$d(X, \text{Persona}^{p^*}, K) = \min_p d(X, \text{Persona}^p, K). \quad (4.11)$$

Se puede obtener el vector de descriptores  $X$  ordenando los valores de mayor a menor del mapa  $Q$  para obtener la posición de la combinaciones de  $n$  y  $m$  que mejor clasifiquen. En la Fig.4.11 se muestra el porcentaje de clasificación de la distancia mínima, la correlación y el método combinado.

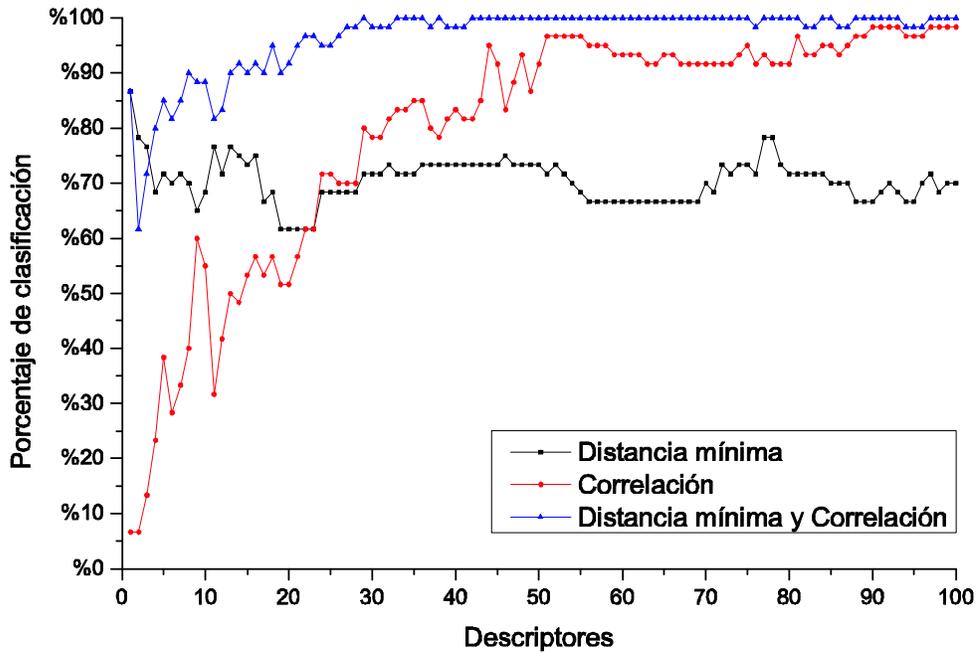


Figura 4.11: Métodos para la clasificación

El porcentaje de la correcta clasificación depende del número de descriptores y la posición que son considerados en el Ec. (4.10). Se define un vector permutable  $\pi_k = [1, 2, \dots, K]$  el cual puede modificar la posición de los descriptores seleccionados  $D_k$ , y se incluye en una función  $\eta(D_{\pi_k}, k)$  la cual indica el porcentaje de correcta clasificación dependiendo el número de descriptores, como se muestra en la Fig.4.11 . El problema radica en encontrar la posición de los descriptores de tal manera que clasifiquen correctamente, se plantea el problema como,

$$\text{máx } Z(\pi_k) = \frac{1}{K} \sum_{k=1}^K \eta(D_{\pi_k}, k). \quad (4.12)$$

El problema de la Ec.(4.12) utiliza una representación de un vector permutable  $\pi_k$ , el cual es utilizado frecuentemente en problemas de Optimización Combinatoria. El mapa  $Q$  indica donde están ubicados los mejores descriptores que clasifican de forma individual. Para los *Algoritmos Genéticos* (Apéndice A) puede facilitar la vecindad de búsqueda indicando el orden de los mejores descriptores en la solución inicial. Se muestra en el Algoritmo 2.

---

**Algoritmo 2** Algoritmo Genético para buscar el mínimo número de descriptores que clasifiquen correctamente

---

**Inicio:** se genera una población y se incluye la posición de los mejores descriptores del mapa  $Q$

**repite**

**Aptitud:** se evalúa la función objetivo  $Z(\pi_k)$  de cada individuo.

**Selección:** se hace al azar dependiendo la aptitud.

**if** Probabilidad de cruza  $\geq$  Aleatorio **then**

        Se realiza la cruza.

**else**

        Se deja como esta.

**end if**

**if** Probabilidad de mutación  $\geq$  Aleatorio **then**

        Se deforma la nueva descendencia.

**else**

        Se deja como esta.

**end if**

**Reemplazo:** la nueva población generada reemplaza a la vieja población.

**hasta** contador  $\geq$  número de iteraciones

---

En la Fig.4.12 se muestra como al combinar la búsqueda de los *Algoritmos Genéticos* con el mapa  $Q$  clasifica correctamente con menos descriptores y de forma individual clasifican con un número mayor de descriptores.

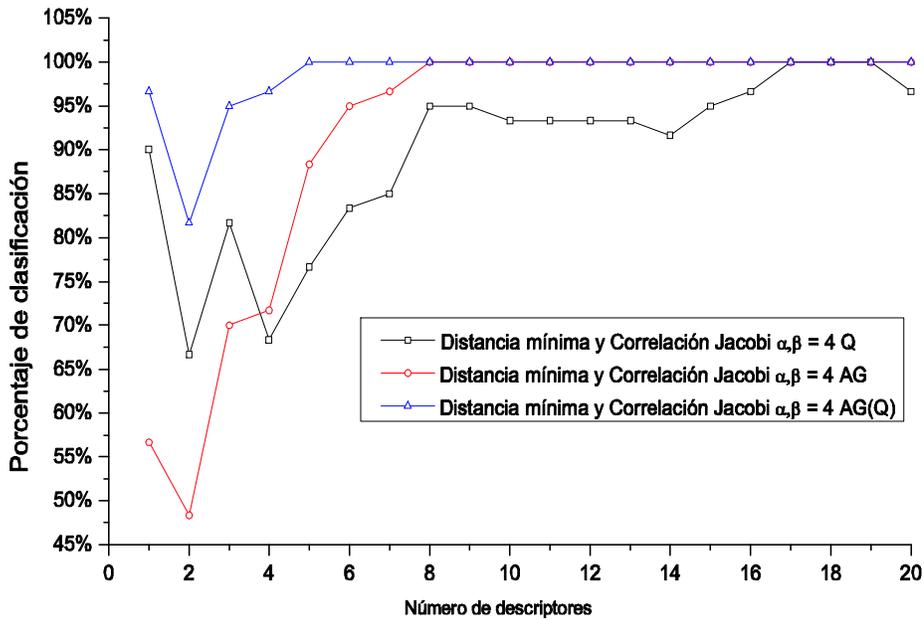


Figura 4.12: Gráfica de porcentajes de clasificación con diferente posición de los mejores descriptores.

La gráfica anterior también muestra que con tan sólo 4 momentos es posible obtener un porcentaje de clasificación de 100%. Un análisis comparativo y evaluativo del algoritmo propuesto en éste capítulo se realiza en el capítulo 5 con el propósito de identificar y clasificar un individuo a través de la marcha con el mínimo número de descriptores.



# Capítulo 5

## Discusión de Resultados

En este capítulo se aplica el método de correlación de historia de momentos de Jacobi-Fourier para clasificación de las tres bases preprocesadas: UPT, MoBo y CASIA-A. La descripción de las bases y su preprocesamiento fue revisado en el Capítulo 3. Los parámetros de los polinomios de Jacobi utilizados son:  $\alpha = \beta = 4, 6, 8, 10$ . El orden  $n$  y la repetición  $m$  para las familias de los momentos de Jacobi-Fourier utilizadas para la clasificación de este capítulo, está dado por  $n = m = 0, 1, 2, \dots, 10$ . También, el método es puesto a prueba con los momentos de Zernike. Una breve explicación de estos momentos se revisan en el Apéndice D.

En la Fig. 5.1 se muestra un esquema para determinar los porcentajes de clasificación de las personas usando el método propuesto, previamente descrito en el Capítulo 4.

La efectividad del algoritmo es determinada por el máximo porcentaje de correcta clasificación y el mínimo número de descriptores de la marcha.

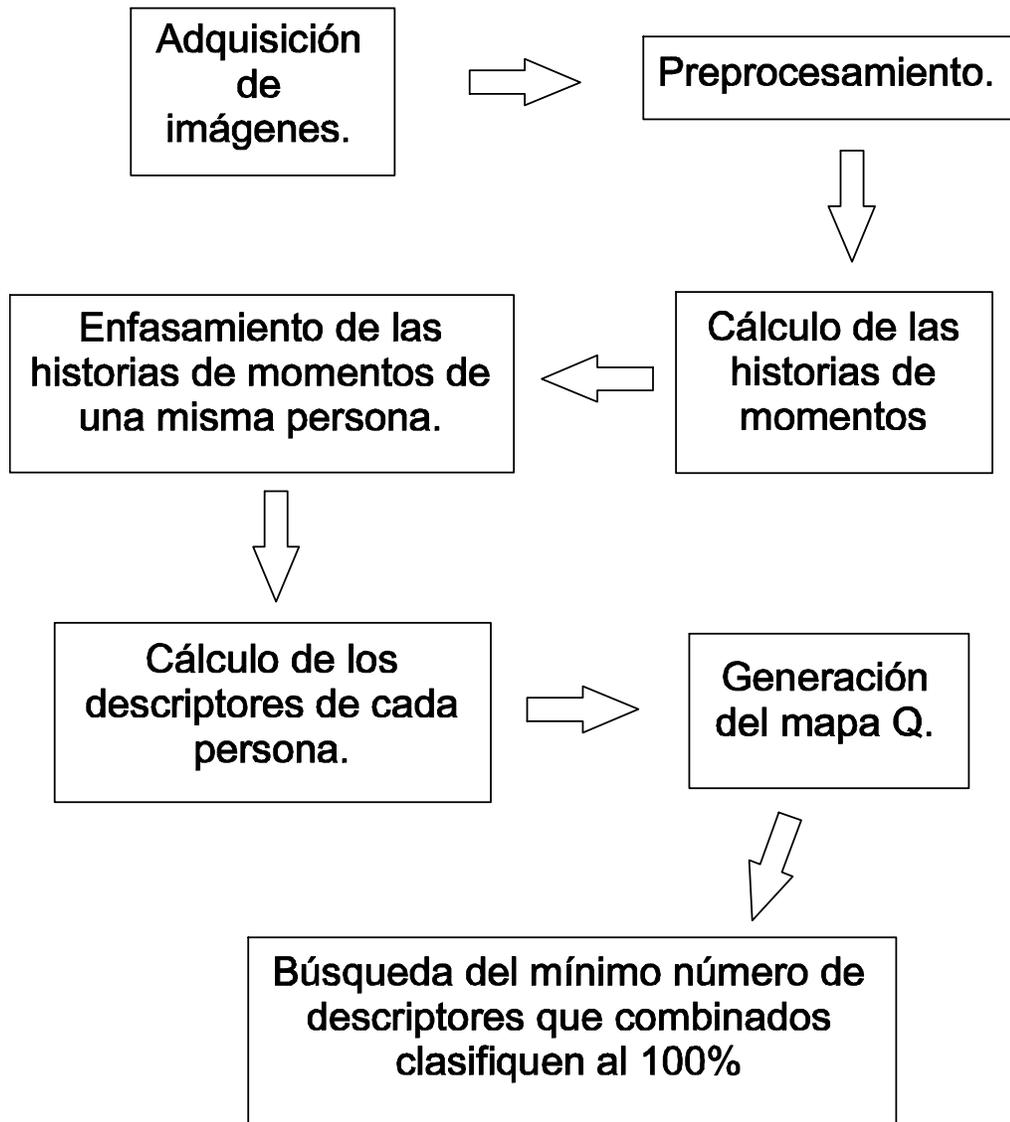


Figura 5.1: Proceso de clasificación del método propuesto.

## 5.1. Clasificación de la marcha de la base UPT

Una vez calculados los descriptores de las historias de momentos de Jacobi-Fourier, se procede a medir la capacidad de discriminación a través del uso de una medida propuesta por la Ec. (4.7). Esta medida genera un conjunto de valores que se registran de acuerdo con el mapa que se observa en la Fig. 5.2.

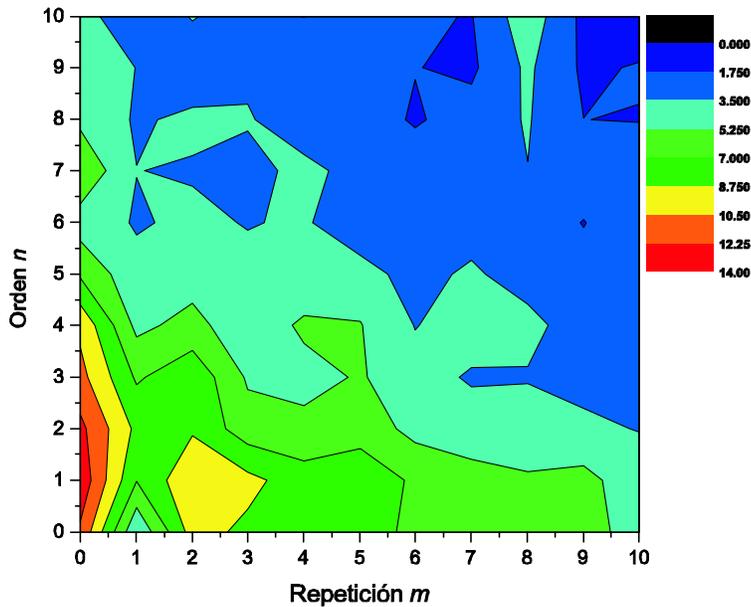


Figura 5.2: Mapa  $Q$  para  $\alpha = \beta = 4$  usando la base UPT

En este mapa, se han agrupado los valores de la medida  $Q$ , conforme al valor absoluto más alto. Esto es, las zonas en rojo son aquellas en donde los valores para  $Q$ , suelen ser mayores. Conforme se tienen valores más bajos, estos van decreciendo de izquierda a derecha. De color rojo a azul intenso. A manera de comparación con otras familias en la Fig. 5.3, se muestran los mapas  $Q$  para  $\alpha = \beta = 6, 8, 10$ .

Dentro de las regiones de color rojo, anaranjado y amarillo, se encuentran los descriptores que mejor se adaptan a los perfiles de las historias de

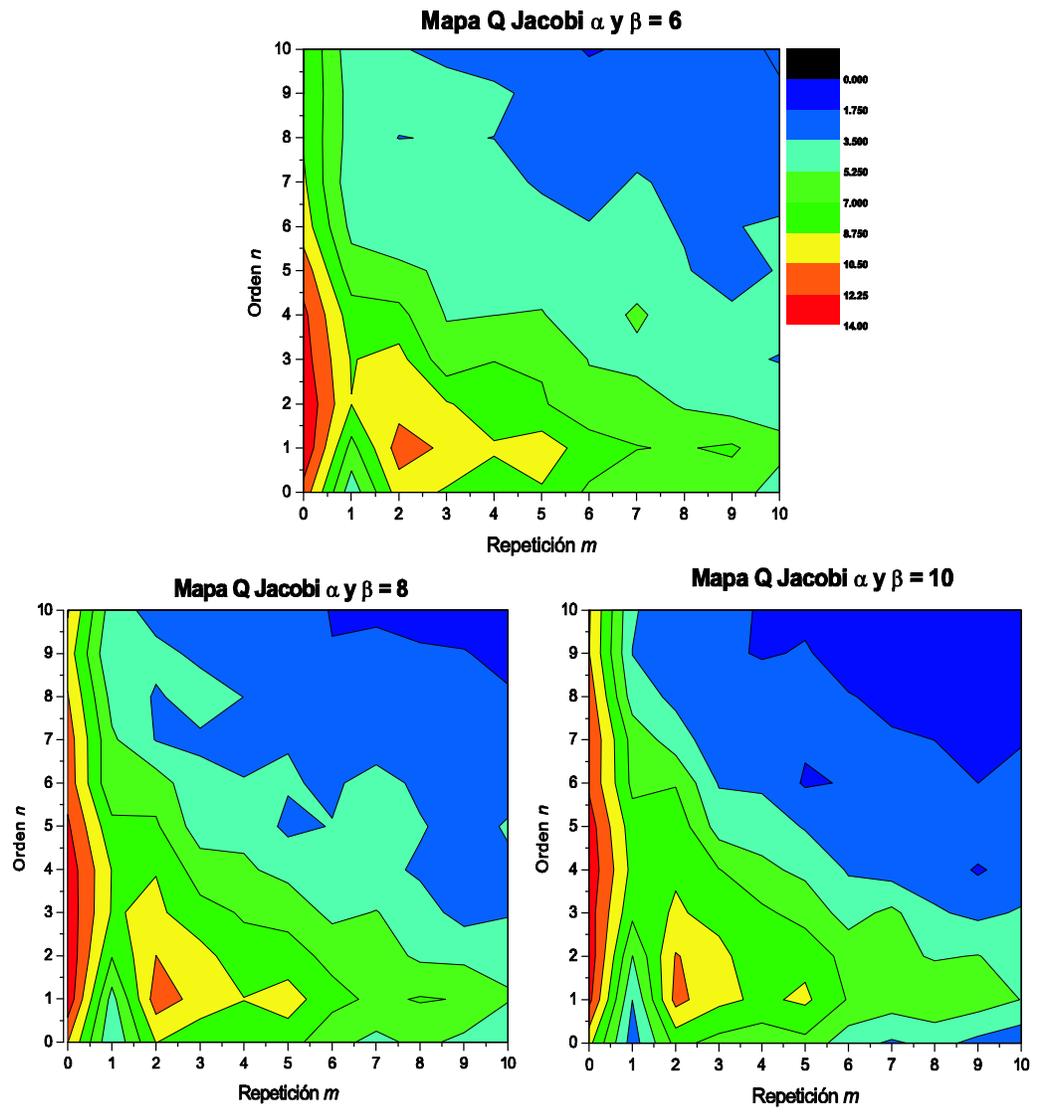


Figura 5.3: Mapas  $Q$  para  $\alpha = \beta = 6, 8, 10$ .

momentos. Una forma de verificar el comportamiento de algunos de los momentos de estas regiones, es calculando la distribución de probabilidad bajo una simulación con condiciones de ruido aleatorio y cambios de escala. Un ejemplo del ajuste de una distribución para el mejor descriptor se observa en la grafica de la Fig. 5.4.

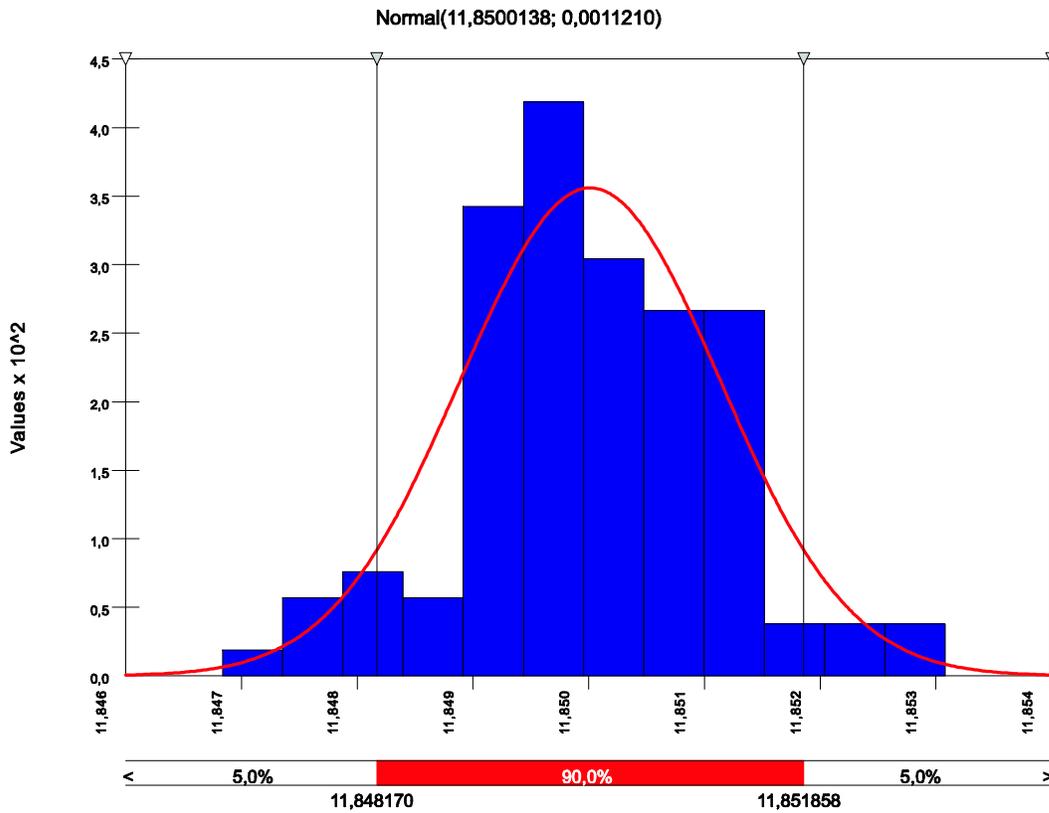


Figura 5.4: Mejor ajuste de la simulación del descriptor con parámetros  $n = 1$  y  $m = 0$  para  $\alpha = \beta = 4$ . El descriptor se ajusta a la distribución normal.

Asimismo como contra ejemplo, se verifica el comportamiento de uno de los momentos de las regiones con los valores más bajos (en azul) en la Fig. 5.2. Generando este último una distribución como la observada en la Fig. 5.5. Comparando las Fig. 5.4 y 5.5; se puede observar que ambas distribuciones se comportan de manera diferente. La primera se ajusta a una distribución nor-

mal. Esto en concordancia con la definición de la medida  $Q$ . Contrariamente, la distribución de la Fig. 5.5 no está normalmente distribuida (se ajusta a una distribución LogLogística). Esto refuerza el criterio de que los mejores descriptores de la historia de momentos, deben ser buscados en las regiones de valores más altos (en rojo, anaranjado, amarillo o en combinaciones).

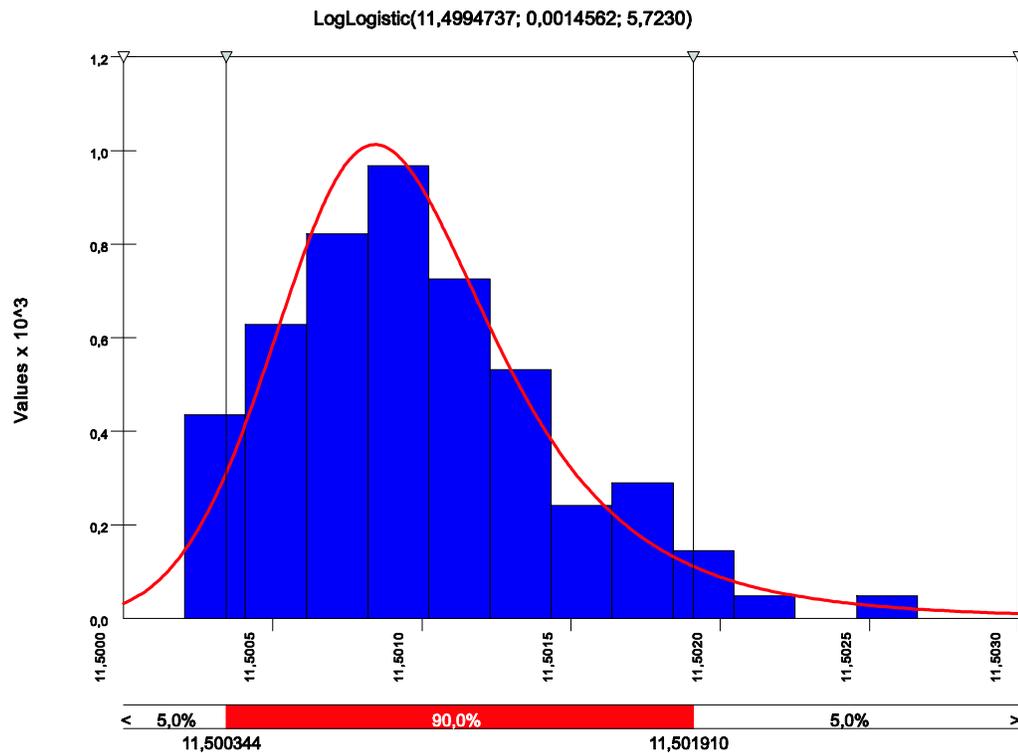


Figura 5.5: Mejor ajuste de la simulación del descriptor con parámetros  $n = 10$  y  $m = 10$  para  $\alpha = \beta = 4$ . El descriptor no se ajusta a una distribución normal.

De acuerdo con el mapa de la Fig. 5.2. Se consideran los tres mejores descriptores  $D_{1,0}$ ,  $D_{2,0}$  y  $D_{0,0}$  para ser graficados en un diagrama de dispersión como se muestra en la Fig. 5.6. Cada punto representa los descriptores de una persona en el espacio de los tres mejores.

Como se puede observar en la Fig. 5.6 los puntos de los descriptores de

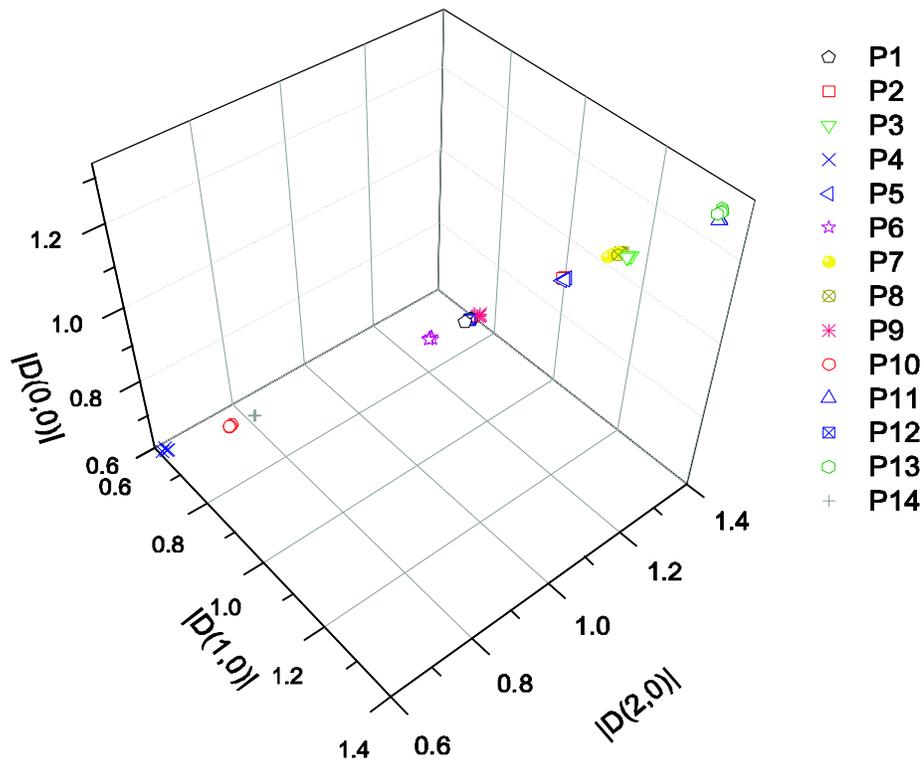


Figura 5.6: Gráfica de dispersión de los tres mejores descriptores para  $\alpha = \beta = 4$ , de cada una de las 4 secuencias de imágenes para cada uno de los 14 individuos.

las personas están agrupados intraclase, pero no en todos los casos están separados entreclases. Esto se debe a que el mapa  $Q$  ayuda a identificar a los descriptores que mejor clasifican de manera individual. Es necesario buscar los mejores descriptores que combinados clasifique correctamente con el mínimo número de descriptores, esto se logra resolviendo el problema de la Ec. (4.12) mediante *Algoritmos Genéticos*. La solución del problema anterior se mejora incluyendo como solución inicial el mapa  $Q$ . Facilita la búsqueda para encontrar los descriptores que combinados clasifique al 100%. En la Fig.5.7 se muestra la solución del problema de la Ec. (4.12) para  $\alpha = \beta = 4, 6, 8, 10$  y los momentos de Zernike. Se puede observar que con tan sólo 4 momentos de Jacobi  $J_n(10, 10, r)$  se puede clasificar al 100% de los individuos.

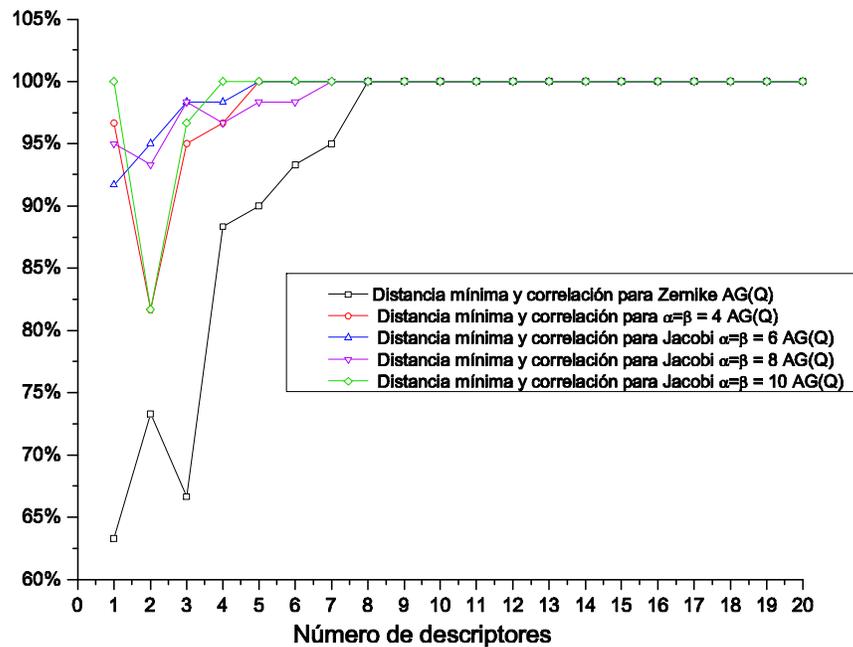


Figura 5.7: Porcentajes de clasificación usando la base UPT

En la Fig. 5.8 se muestra el mapa de la posición del resultado de la clasificación con los *Algoritmos Genéticos*. Las tonalidades de la barra de colores mostrada a la derecha del mapa, de rojo oscuro (valor de 120) a azul oscuro (valor de 1) indica que el rojo intenso del mapa es la combinación del orden  $n$

y repetición  $m$  del primer descriptor que se toma para clasificar, las posiciones siguen sucesivamente hasta llegar al azul intenso. Como se puede observar los primeros descriptores que son tomados para clasificar corresponden a los mejores descriptores del mapa  $Q$  mostrado en la Fig.5.2. Esto demuestra que al tomar como solución inicial el mapa  $Q$  en los *Algoritmos Genéticos* facilita la búsqueda de la combinación de los mejores descriptores para una correcta clasificación.

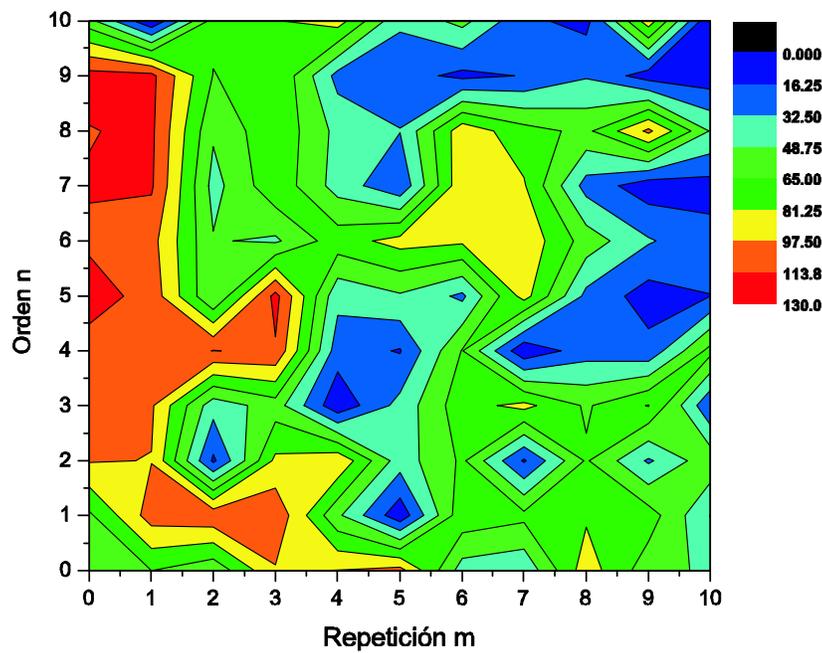


Figura 5.8: Mapa de la posición de los descriptores para clasificar para  $\alpha = \beta = 4$  usando la base UPT

## 5.2. Clasificación de la marcha usando la base MoBo

La base MoBo contiene 25 sujetos en 6 ángulos de visión distintos con 4 rutinas diferentes marchando en una caminadora. Para la clasificación usando el método propuesto se utilizaron 15 sujetos con el ángulo de visión lateral (cámara C2 de la Fig.3.7) y caminando lento, esto con el fin de tener condiciones similares entre bases.

Los mapas  $Q$  de la base MoBo para los parámetros  $\alpha = \beta = 4, 6, 8, 10$ , se muestran en la Fig.5.9. Tienen cierta similitud con los mapas de la base UPT. Los valores más altos van decreciendo de rojo a azul intenso. Algunos de los mejores descriptores de la base MoBo están ubicados para el orden angular  $m$  cuando  $n = 0$ . Esto quiere decir que para estos casos en particular los polinomios de Jacobi no se pudieron ajustar a la forma de la imagen.

Se grafican los tres mejores descriptores de cada mapa  $Q$ . En algunos casos las clases están agrupadas intraclase pero ligeramente separadas entraclase ó ligeramente agrupadas intraclase pero separadas entreclases. Esto se puede observar en la Fig.5.10.

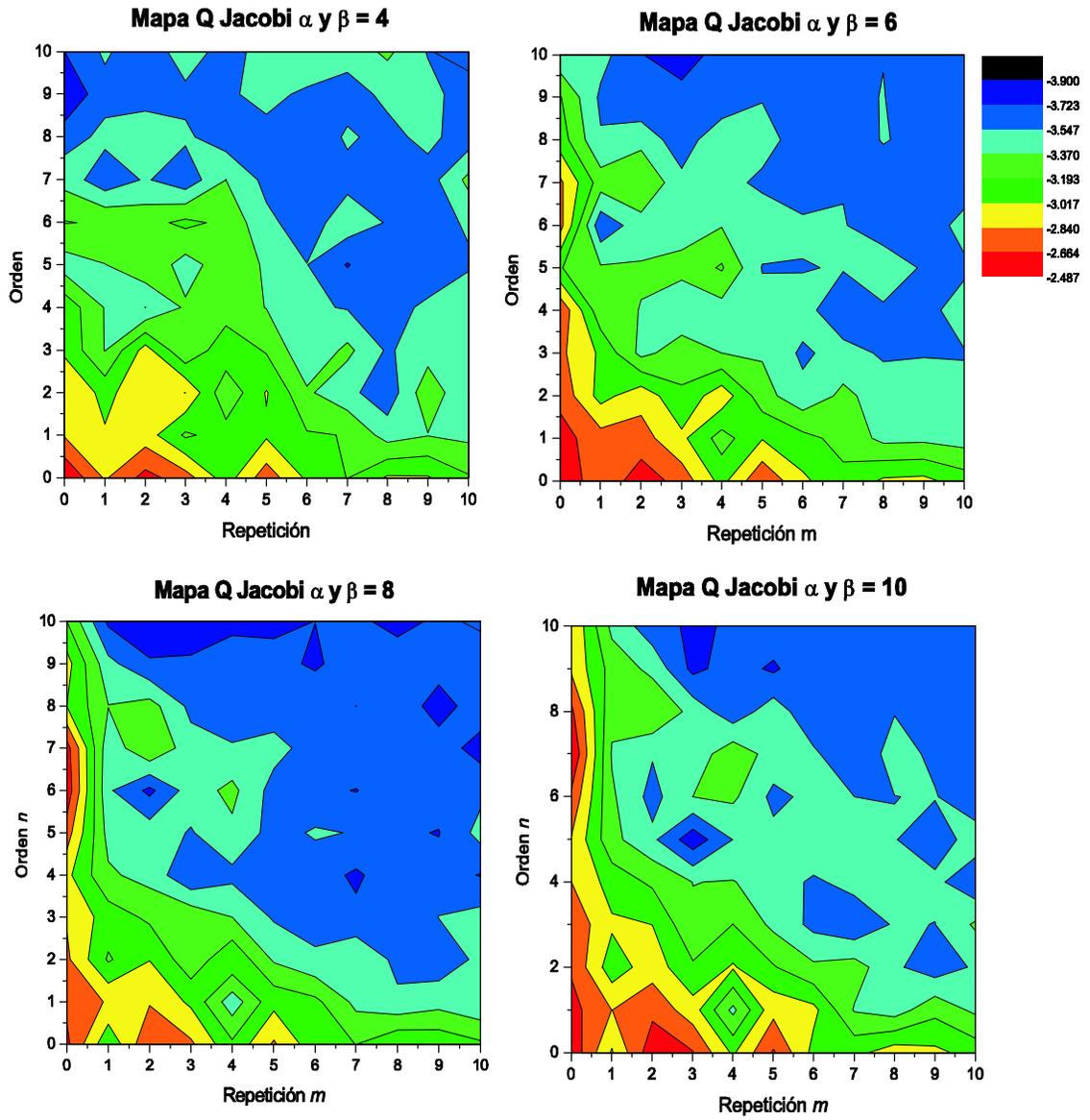


Figura 5.9: Mapas  $Q$  para  $\alpha = \beta = 4, 6, 8, 10$ , usando la base MoBo.

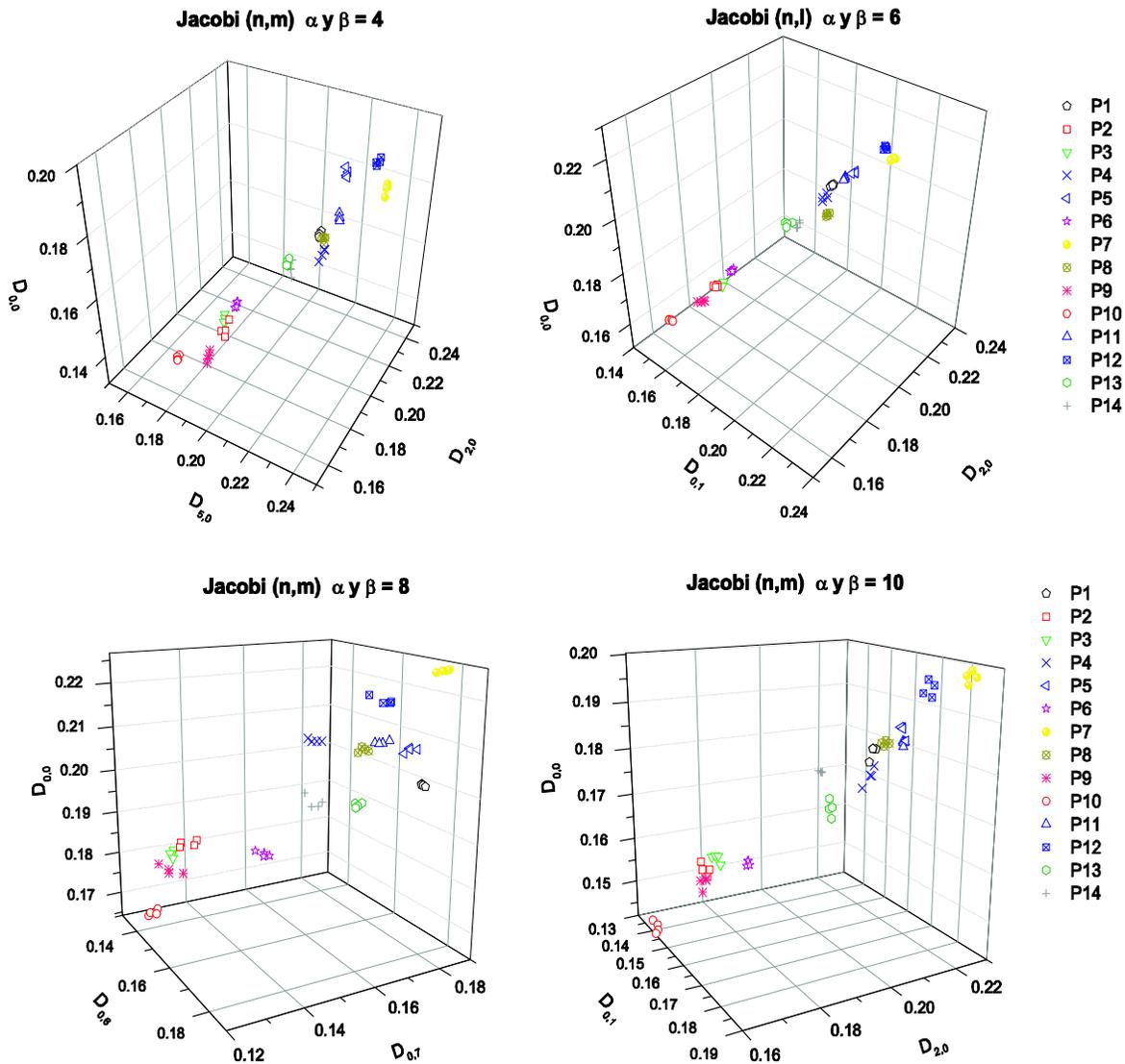


Figura 5.10: Diagramas de dispersión para  $\alpha = \beta = 4, 6, 8, 10$ . usando la base MoBo.

Una vez calculados los mapas  $Q$ , se procede a buscar los descriptores que combinados clasifiquen al 100%. El resultado de la búsqueda de los mejores descriptores se muestra en la Fig. 5.11.

A diferencia de los resultados de la base UPT, la base MoBo necesita un mayor número de descriptores para una correcta clasificación. Algunas condiciones que pudieron haber afectado los resultados fueron el ruido de las siluetas, el hecho de marchar en una caminadora puede modular los tiempos de la marcha.

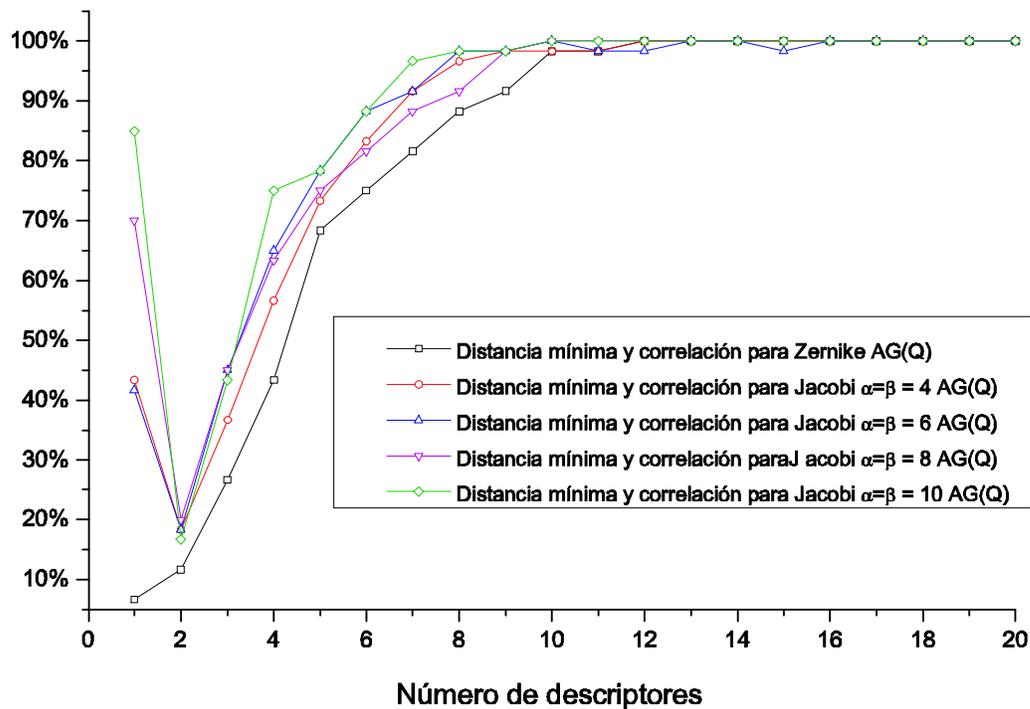


Figura 5.11: Porcentajes de clasificación de la combinación de los mejores descriptores usando la base MoBo.

### 5.3. Clasificación de la marcha usando la base CASIA-A

Finalmente para la base CASIA-A fue necesario reducir el número de sujetos de 20 a 12 debido a que no contienen el número mínimo de 50 marcos para implementar el método propuesto. Como en el caso de las bases anteriores se utiliza la vista frontal.

El hecho de contar un número menor de los 15 sujetos de las pruebas anteriores y que las secuencias no estén afectadas por un cambio de escala, puede ocasionar que facilite la clasificación.

Los mapas  $Q$  de la base CASIA-A para los parámetros  $\alpha = \beta = 4, 6, 8, 10$ , se muestran en la Fig.5.12. Los mapas tienen cierta similitud con los mapas de la base anteriores, los valores más altos van decreciendo de rojo a azul intenso.

Se gráfica los tres descriptores de cada mapa  $Q$  para tener una idea del comportamiento de los mejores descriptores observando si están agrupados intracalse y separados entreclase. Esto se puede ver en la Fig.5.13.

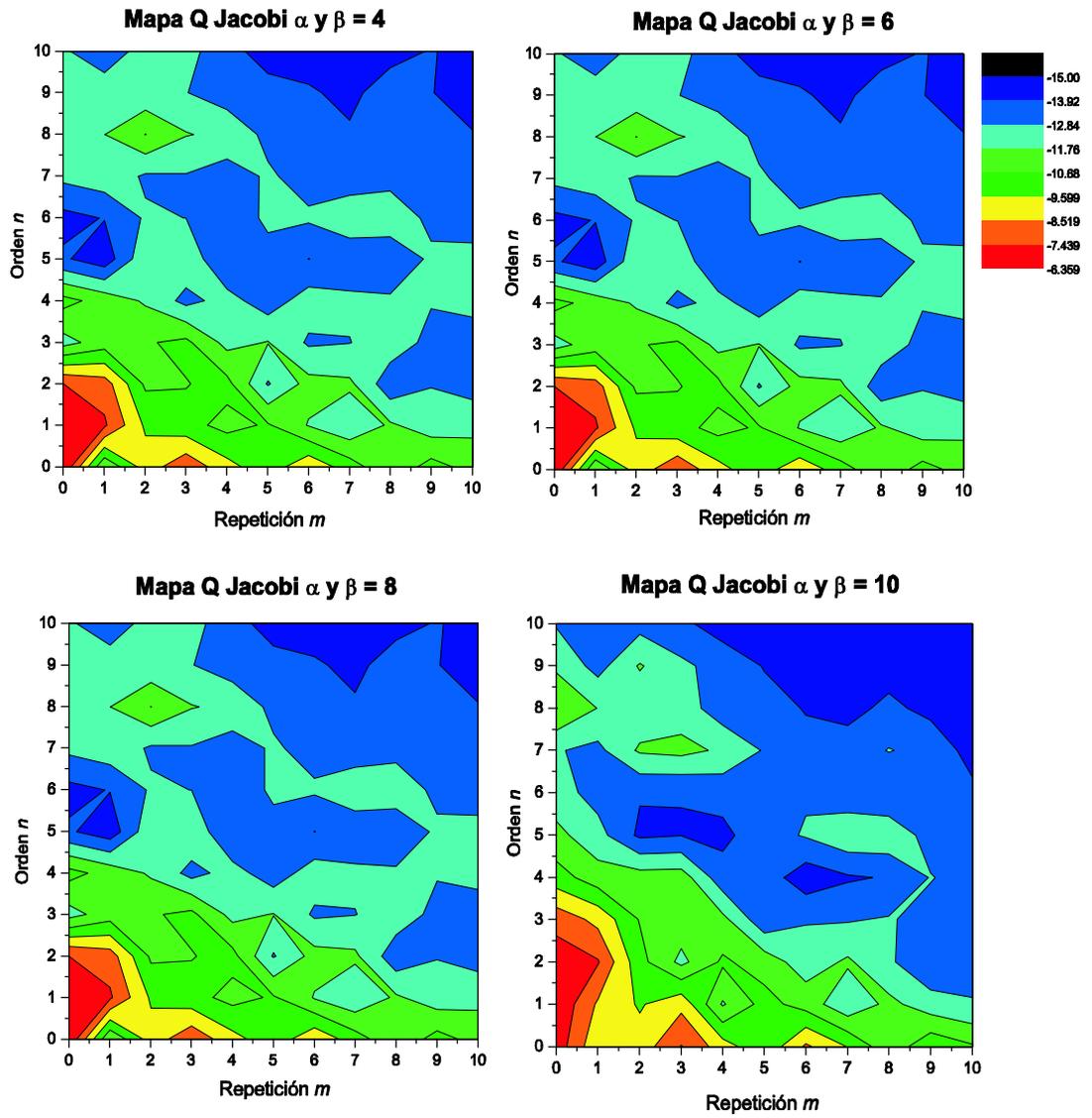


Figura 5.12: Mapas  $Q$  para  $\alpha = \beta = 4, 6, 8, 10$ , usando la base CASIA-A.

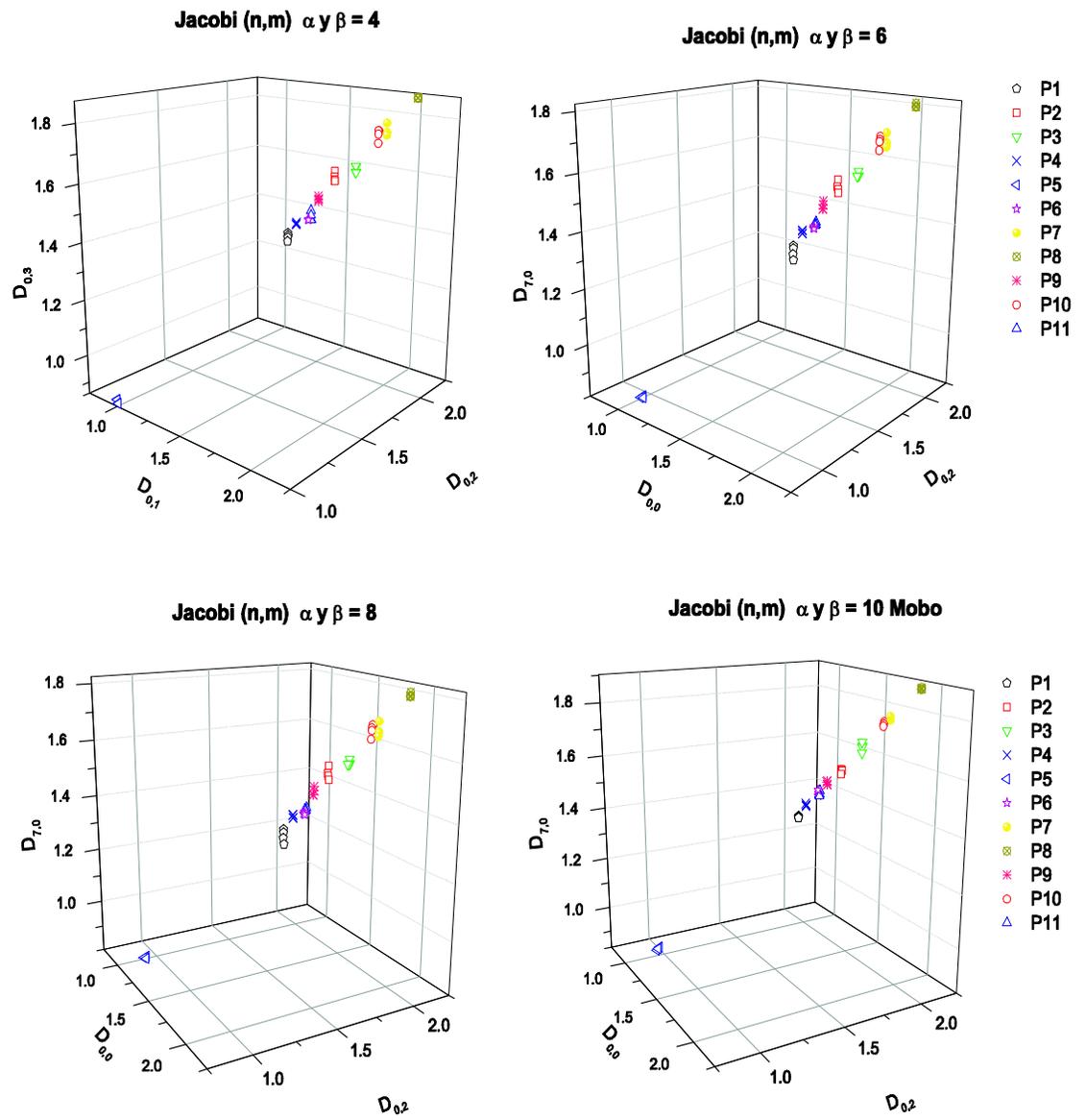


Figura 5.13: Diagramas de dispersión para  $\alpha = \beta = 4, 6, 8, 10$ . usando la base CASIA-A.

Una vez calculados los mapas  $Q$ , se buscan los mejores descriptores que combinados clasifiquen correctamente. La grafica de los descriptores seleccionados y su clasificación se muestra en la Fig. 5.14.

Como se puede observar en la gráfica de la Fig.5.14 son necesario una mínima cantidad entre 3 o 4 descriptores para la familia de los Jacobi salvo el caso contrario de los momentos de Zernike que necesita 12 descriptores para clasificar al 100 %. También se puede observar que para el caso de  $\alpha = \beta = 8$  un solo descriptor clasifica al 100 %.

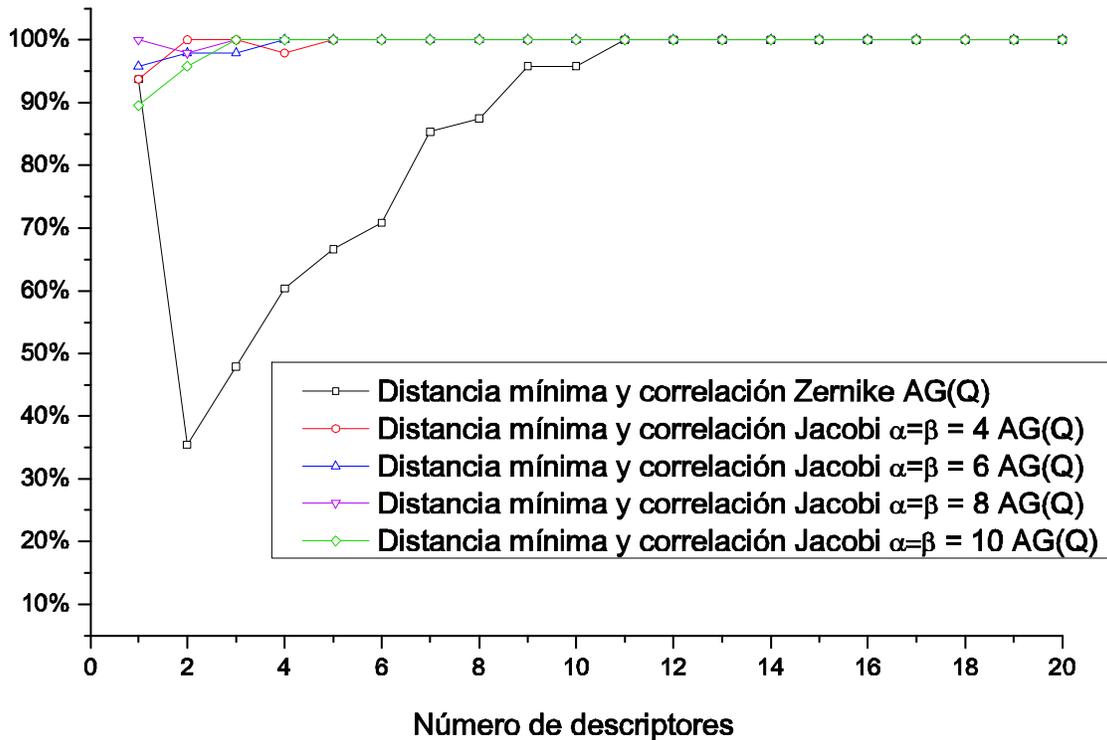


Figura 5.14: Porcentajes de clasificación usando la base CASIA-A



# Capítulo 6

## Conclusiones Generales

El método implementado en esta tesis tiene la ventaja de analizar la marcha codificada como una historia de momentos y su relación con otras historias de momentos de la misma clase. De esta forma se obtienen descriptores de la marcha con la sección de referencia característica de una misma persona. A diferencia de otros métodos que analizan todos los marcos de una secuencia como un todo [10],[19],[26],[28],[45]. Además, en este tipo de representaciones se relacionan con parámetros temporales de la marcha.

En general en éste trabajo se implementó un método para la clasificación de la marcha de personas a partir de secuencias de imágenes usando la correlación de la historia de momentos de Jacobi-Fourier. Se calibran las tomas iniciales de secuencias de imágenes poniendo en fase las historias de momentos de una misma persona. Ésto se logró maximizando la correlación de una sección de referencia y los intervalos seleccionados de las historias de momentos. El promedio de la sección de referencia es considerado como un descriptor que engloba la información de la historia de momentos.

El problema de poner en fase las historias de momentos ayuda a que el descriptor de una misma clase sea similar, permitiendo diferenciarse entre otras clases y así poder identificar con un margen de error menor.

El método de clasificación es la combinación de dos métodos: la distancia mínima y la correlación. La cual busca la mínima distancia y la relación lineal entre los descriptores de entrada con la referencia de forma simultánea; mejorando así el porcentaje de clasificación.

Se logró reducir el número de descriptores que combinados clasifican correctamente, mediante una búsqueda heurística de las posibles combinaciones de los descriptores. La vecindad de búsqueda de los descriptores, se realiza a través de una medida discriminatoria de momentos  $Q$ . La cual, a través de mapas de calibración entre bases nos indica el orden  $n$  y la repetición  $m$  de los momentos que mejor clasifican. Figs. 5.3, 5.9 y 5.12.

Se comprobó la efectividad de la medida  $Q$  mediante la simulación de los descriptores de la marcha con cambios de escala y ruido aleatorio. Los datos arrojados de la simulación de los mejores descriptores tiene un comportamiento de tipo Gaussiano, esto quiere decir que el error del descriptor es siempre pequeño. Si se realiza un número grande de observaciones en la simulación, puede esperarse que los errores positivos se presenten con la misma frecuencia que los errores negativos, de modo que al tomar la media, existe una compensación del error.

Por lo anterior se propuso probar el método con tres bases de datos: UPT, MoBo y CASIA-A. El resumen de los resultados obtenidos se presenta en el Cuadro 6.1.

Base	Condiciones	Momentos	Descriptores	Clasificación
UPT 15 sujetos 120 × 120px	Aire libre Fondo blanco Medio cuerpo	Zernike	8	100 %
		J(4, 4, r)	5	100 %
		J(6, 6, r)	5	100 %
		J(8, 8, r)	7	100 %
		J(10, 10, r)	4	100 %
MoBo 15 sujetos 120 × 120px	Caminadora En habitación Medio cuerpo	Zernike	11	100 %
		J(4, 4, r)	12	100 %
		J(6, 6, r)	13	100 %
		J(8, 8, r)	10	100 %
		J(10, 10, r)	10	100 %
CASIA-A 12 sujetos 120 × 120px	Aire libre Medio cuerpo	Zernike	11	100 %
		J(4, 4, r)	5	100 %
		J(6, 6, r)	4	100 %
		J(8, 8, r)	3	100 %
		J(10, 10, r)	3	100 %

Cuadro 6.1: Cuadro comparativo de los resultados de las tres bases: UPT, MoBo y CASIA-A

Como se muestra en el Cuadro 6.1 todos los momentos clasifican al 100 %. A manera de comparación con el método de Shulter [45] para la base MoBo que requiere 90 descriptores de imágenes de  $486 \times 640$ px para obtener una correcta clasificación, para el caso de los polinomios  $J_n(10, 10, r)$  sólo necesita 10 descriptores con imágenes de  $120 \times 120$ px. Lo cual significa que, con menor información se incrementó la velocidad de respuesta y la eficacia del método. En general los polinomios de Jacobi  $J_n(10, 10, r)$  logran una correcta clasificación con un número menor de descriptores en comparación con otros como Zernike.

El propósito de normalizar la historia de momentos para que sea invariante a la escala, de poner en fase la historia de momentos, de proponer métodos de clasificación y de buscar los descriptores que combinados clasifiquen correctamente, es para tener un método robusto que pueda reconocer un número grande de personas. Las limitaciones de qué tan robusto es el método son aún desconocidas.

Como trabajo a futuro se plantea, extender el número de individuos a clasificar, estudiar la problemática de los cambios de ropa para cambiar el contraste en las imágenes y estudiar los diferentes ángulos de visión.

Un punto importante es el estudio de los polinomios de tipo circular ortogonal y cartesiano, que plantean nuevos conjuntos como los Bessel y Krawchouk. Los cuales recientemente han demostrado ser adecuados en la descripción de imágenes digitales y sus aplicaciones en los diferentes contextos del tratamiento digital de imágenes y la visión por computadora.

El estudio realizado refleja la necesidad de llevar a cabo investigación en el área de la biometría de la marcha, su estudio es relativamente nuevo. No existe un enfoque único por el cual se pueda resolver el problema del reconocimiento y análisis de la marcha, brindando la oportunidad de proponer nuevos métodos y ser pioneros en el estudio de la marcha a través de medios no invasivos. Esto último es muy útil en cuestiones de seguridad, al reconocer a un individuo a través de su marcha, sin importar que se cubra el rostro y tenga guantes para no dejar huellas dactilares.

Este trabajo de tesis es el puntero para realizar otras investigaciones como: el obtener parámetros temporales de la marcha como apoyo en un análisis clínico, al resolver el problema de poner en fase la historia de momentos mediante la Ec.(4.5) se puede identificar la historia de momentos característica de un atleta de alto rendimiento, la cual sirve como una herramienta de apoyo para mejorar su técnica.



# Apéndice A

## Algoritmos Genéticos

Un problema de optimización puede parecer muy simple. Uno puede formular todas las posibles soluciones que corresponden al problema, el juego de estas posibles soluciones corresponde a un espacio de búsqueda. El problema consiste en averiguar la solución que encaja mejor, es decir la que nos arroje un mejor beneficio de todas las posibles soluciones. Si fuera posible enumerar todas las soluciones rápidamente, el problema no provoca mucha dificultad. Pero, cuando el espacio de búsqueda es bastante grande, la enumeración no parece ser muy factible, tomaría demasiado tiempo. Para este tipo de problemas, es necesario utilizar una técnica en específico para encontrar la solución óptima. Los *Algoritmos Genéticos* proporcionan uno de estos métodos.

Los *Algoritmos Genéticos* se ocupan de una población de posibles soluciones. Cada solución se representa a través de un cromosoma que simplemente es una representación abstracta, donde codifica todas las posibles soluciones. Se tiene que determinar un juego de operadores para la reproducción y la mutación de los cromosomas. La representación apropiada, y los operadores; son un paso determinante en la conducta de los *Algoritmos Genéticos*. Con frecuencia, puede ser sumamente difícil de encontrar una representación que respeta la estructura del espacio de búsqueda, también es difícil encontrar operadores que sean coherentes y pertinentes, según las propiedades del problema.

La selección se realiza comparando cada individuo de la población a través de una función de aptitud. Cada cromosoma tiene un valor asociado que corresponde a una función de aptitud. La aptitud debe de corresponder a la evaluación de un candidato, para ver si es una buena solución. La selección puede adaptarse para buscar los individuos más aptos, ya sea buscando los

valores máximos o mínimos de un problema de optimización. Una vez que se ha definido apropiadamente la reproducción y la función de aptitud, el algoritmo genético evoluciona según su estructura básica. Empieza generando una población inicial de cromosomas. La población debe de ofrecer una gran diversidad de material genético. Generalmente, la población inicial se genera al azar.

Un *Algoritmo Genético* es un proceso de iterativo para que una población de individuos evolucione. El Algoritmo 3 es un esquema de un algoritmo genético simple

---

**Algoritmo 3** Algoritmo Genético Simple

---

**Inicio:** se genera una población aleatoria

**repite**

**Aptitud:** se evalúa la función objetivo  $Z$  de cada individuo.

**Selección:** se hace al azar dependiendo la aptitud.

**if** Probabilidad de cruza  $\geq$  Aleatorio **then**

        Se realiza la cruza.

**else**

        Se deja como esta.

**end if**

**if** Probabilidad de mutación  $\geq$  Aleatorio **then**

        Se deforma la nueva descendencia.

**else**

        Se deja como esta.

**end if**

**Reemplazo:** la nueva población generada reemplaza a la vieja población.

**hasta** contador  $\geq$  número de iteraciones

---

Para implementar de forma adecuada un algoritmo genético, es necesario el cumplimiento de una serie de condiciones, que puedan resumirse en la correcta definición de los siguientes aspectos: la codificación, función de adaptación, criterios de selección, operadores genéticos, criterio de parada. Los detalles teóricos y prácticos pueden encontrarse en [46], [47].

# Apéndice B

## Marcha humana

La marcha humana, como actividad rítmica y cíclica que compromete todo el cuerpo y, en especial las dos extremidades inferiores, comprende un inicio y un final para cada ciclo. El ciclo de la marcha normal consta de dos fases: de soporte y de balanceo. La *fase de soporte* o estático constituye el 60 % de cada ciclo de la marcha, ocurre cuando una pierna sufre carga y está en contacto con la superficie de apoyo, y la *fase de balanceo* o dinámica que constituye el 40 % del ciclo y que se produce cuando se avanza una pierna para dar el paso siguiente. El ciclo de la marcha normal se ilustra en la Fig. B.1 .

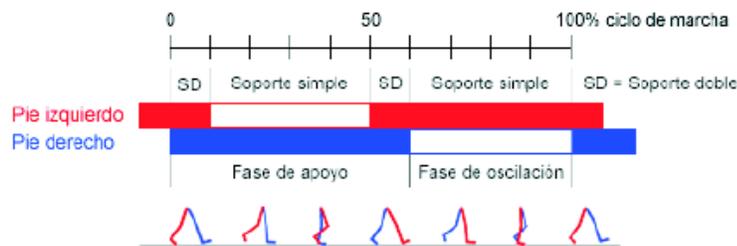


Figura B.1: Ciclo de la marcha normal. El ciclo de la marcha se divide en dos fases: de apoyo y de oscilación. Además, el ciclo de la marcha puede dividirse en las fases de soporte simple y soporte doble.

Los componentes de la marcha humana se dividen en dos grupos, los *componentes temporales* y los *componentes espaciales*. Entre los parámetros

relacionados con el tiempo, o temporales que pueden valorarse con relativa facilidad, encontramos:

- **Periodo de Paso:** referido al intervalo de tiempo entre el contacto inicial de un pie y el mismo evento del pie contralateral.

- **Periodo de soporte:** es el tiempo que pasa entre el instante en el que el pie hace contacto con el piso (talón), hasta el momento de despegue de los dedos del mismo.

- **Periodo de balanceo:** es el tiempo transcurrido entre el instante del despegue de los dedos del pie que pierden por completo la relación con el piso, hasta el punto del contacto del talón con el suelo.

- **Frecuencia o cadencia:** este parámetro hace referencia al número de pasos por unidad de tiempo, que generalmente se fija en un minuto.

- **Velocidad:** es la relación de la distancia recorrida en la dirección de la marcha por unidad de tiempo, se expresa en m/s.

- **Aceleración:** es el cambio en la velocidad, se expresa en  $m/s^2$ .

Entre los parámetros espaciales que se deben explorar bilateralmente, encontramos:

- **Longitud de paso largo o zancada:** es la distancia lineal en metros entre dos eventos iguales y sucesivos de la misma extremidad.

- **Longitud de paso o paso corto:** es la distancia lineal en metros desde un evento específico, que, generalmente es el contacto inicial de una extremidad hasta el mismo evento de la extremidad opuesta.

- **Ancho de paso:** esta dado por la distancia lineal en centímetros entre dos puntos iguales de los pies.

- **Angulo de paso:** se refiere a la orientación del pie durante el apoyo.

Los componentes temporales y espaciales son conocido como parámetros biomédicos de la marcha humana. El cálculo de estos parámetros biomédicos es necesario para el análisis clínico de la marcha.

La *Historia de Momentos de Jacobi-Fourier* muestran patrones particulares para cada fase del ciclo de la marcha. Los máximos y mínimos de la *Historia de Momentos Jacobi-Fourier*, pueden estimar algunos parámetros biomédicos de interés.

# Apéndice C

## Software de adquisición de imágenes

La adquisición de imágenes fue controlada mediante una interfaz programada en Visual Basic 6 la cual captura en tiempo real un número de  $N$  imágenes a razón de 16 cuadros por segundo. La Fig.C.1 muestra la interfaz que se programo.



Figura C.1: Pantalla principal del programa para capturar una secuencia de imágenes.



# Apéndice D

## Momentos de Zernike

Los momentos de Zernike han sido exitosamente usados en el reconocimiento de patrones y en el análisis de imágenes debido a sus propiedades de ortogonalidad e invariancia a la rotación.

El kernel de los Momentos de Zernike  $P_{nl}(r, \theta)$  consiste en dos familias de funciones separables: el polinomio de Zernike  $R_{n,l}(r)$  que es la función radial y la exponencial compleja de Fourier  $\exp(jl\theta)$  que es la función angular,

$$P_{nl}(r, \theta) = R_{n,l}(r) \exp(jl\theta), \quad (\text{D.1})$$

donde  $n \geq 0$ ,  $l = 0, \pm 1, \pm 2, \dots$ . Sujetas a las condiciones de que  $n - |l|$  debe ser un número par y  $|l| \leq n$ . El polinomio  $R_{nl}(r)$  está definido como,

$$R_{n,l}(r) = \sum_{s=0}^{\lfloor \frac{n-|l|}{2} \rfloor} (-1)^s \frac{(n-s)!}{s! \binom{n+|l|}{2-s}! \binom{n-|l|}{2-s}!} r^{n-2s}, \quad (\text{D.2})$$

La Fig.D.1 muestra las gráficas para los primeros 8 polinomios radiales y en el Cuadro D.1 las expresiones de estos mismos polinomios,  $R_{nl}(r)$ , hasta orden 6.

La fórmula para calcular el número de polinomios es;  $\sum_{n=0}^{n_{max}} \lfloor \frac{n}{2} \rfloor + 1$ . Para el caso particular  $n_{max} = 4$ , la suma da 9. Los Momentos de Zernike, se calculan mediante,

$$A_{nl} = \int_0^{2\pi} \int_0^1 f(r, \theta) R_{n,l}(r) \exp(jl\theta) r dr d\theta. \quad (\text{D.3})$$

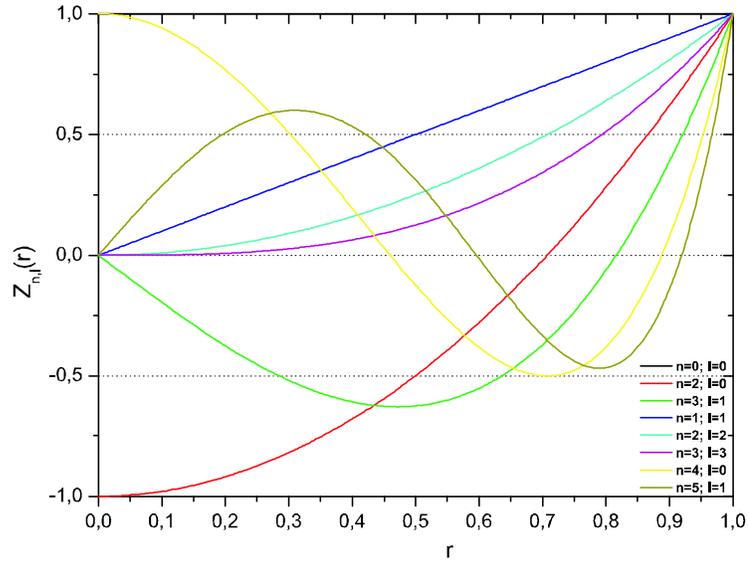


Figura D.1: Polinomios Radiales de Zernike  $R_n(r)$

$n/l$	0	1	2	3	4	5	6
0	1		$2r^2-1$		$6r^4-6r^2+1$		$20r^6-30r^4+12r^2-1$
1		$r$		$3r^3-2r$		$10r^5-12r^3+3r$	
2			$r^2$		$4r^4-3r^2$		$15r^6-20r^4+6r^2$
3				$r^3$		$5r^5-4r^3$	
4					$r^4$		$6r^6-5r^4$
5						$r^5$	
6							$r^6$

Cuadro D.1: Polinomios radiales de Zernike

# Apéndice E

## Trabajos de investigación

- [1] C. Camacho-Bello, C. Toxqui-Quitl, A. Padilla-Vivanco and C. Santiago-Tepantlán, "*Gait recognition analysis based on the correlated history of Jacobi-Fourier moments*". En proceso de envío.



# Bibliografía

- [1] Daza Lesmes, J. (2007). *Evaluación clínico-funcional del movimiento corporal humano*. Bogota: Ed. Panamericana.
- [2] <http://news.bbc.co.uk/1/hi/england/lancashire/7343702.stm>
- [3] N. V. Boulgouris, D. Hatzinakos, and K. N. Plataniotis, “Gait recognition: a challenging signal processing technology for biometric identification,” *IEEE Signal Processing Magazine*, vol. 22, no. 6, 78 pages, 2005.
- [4] L. Sloman, M. Berridge, S. Homatidis, D. Hunter, and T. Duck, “Gait patterns of depressed patients and normal subjects,” *Am. J. Psychiatry*, vol. 139, no. 1, pp. 94–97, 1982
- [5] M.D. Addlesee, A. Jones, F. Livesey, and F. Samaria, “ORL active floor,” *IEEE Pers. Commun.*, vol. 4, no. 5, pp. 35–41, Oct. 1997.
- [6] G. Johansson, “Visual perception of biological motion and a model for its analysis,” *Percept. Psychophysics*, vol. 14, no. 2, pp. 201–211, 1973.
- [7] J.E. Cutting and L.T. Kozlowski, “Recognizing friends by their walk: Gait perception without familiarity cues,” *Bulletin Psychonomic Soc.*, vol. 9, no. 5, pp. 353–356, 1977.
- [8] S.A. Niyogi and E.H. Adelson, “Analyzing and recognizing walking figures in XYT,” in *Proc. Computer Vision and Pattern Recognition*, Seattle, WA, June 1994, pp. 469–474.
- [9] H. Murase and R. Sakai, “Moving object recognition in eigenspace representation: Gait analysis and lip reading,” *Pattern Recognit. Lett.*, vol. 17, no. 2, pp. 155–162, 1996.

- [10] D. Cunado, M.S. Nixon, and J.N. Carter, "Using gait as a biometric, via phaseweighted magnitude spectra," in Proc. Int. Conf. Audio- and Video-Based Biometric Person Authentication, Crans-Montana, Switzerland, Mar. 1997, vol. LNCS 1206, pp. 95–102.
- [11] J. Little and J. Boyd, "Recognizing people by their gait: The shape of motion," *Videre, Int. J. Computer Vision*, vol. 14, no. 6, pp. 83–105, 1998.
- [12] P.S. Huang, C.J. Harris, and M.S. Nixon, "Visual surveillance and tracking of humans by face and gait recognition," in Proc. 7th IFAC Symp. Artificial Intelligence in Real-Time Control, Grand Ganyon National Park, AZ, Oct. 1998, pp. 43–44.
- [13] C. BenAbdelkader, R. Cutler, and L. Davis, "Motion-based recognition of people in eigengait space," in Proc. IEEE Int. Conf. Automatic Face and Gesture Recognition, Washington, DC, May 2002, pp. 254–259.
- [14] Y. Liu, R. Collins, and Y. Tsin, "Gait sequence analysis using Frieze patterns," in Proc. Eur. Conf. Computer Vision, Copenhagen, May 2002, vol. LNCS 2351, pp. 657–671.
- [15] J.P. Foster, M.S. Nixon, and A. Prugel-Bennett, "Automatic gait recognition using area-based metrics," *Pattern Recognit. Lett.*, vol. 24, no. 14, pp. 2489–2497, 2003.
- [16] L. Wang, T. Tan, H. Ning, and W. Hu, "Silhouette analysis-based gait recognition for human identification," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 25, no. 12, pp. 1505–1518, Dec. 2003.
- [17] N.V. Boulgouris, K.N. Plataniotis, and D. Hatzinakos, "An angular transform of gait sequences for gait assisted recognition," in Proc. IEEE Int. Conf. Image Processing, Singapore, Oct. 2004, pp. 857–860.
- [18] L. Lee, G. Dalley, and K. Tieu, "Learning pedestrian models for silhouette refinement," in Proc. Int. Conf. Computer Vision, Nice, France, Oct. 2003, pp. 663–670.
- [19] J. D. Shutler, M. S. Nixon, and C. J. Harris, "Statistical gait description via temporal moments," in Proc. 4th IEEE Southwest Symp. Image Anal. Interpretation, Mar. 2000, pp. 291–295.

- [20] J. B. Hayfron-Acquah, M. S. Nixon, and J. N. Carter, "Automatic gait recognition by symmetry analysis," in *Audio-and-Video-Based Biometric Person Authentication*, vol. 2091, LNCS, J. Bigun and F. Smeraldi, Eds. New York: Springer-Verlag, Jun. 2001, pp. 272–277.
- [21] J. P. Foster, M. S. Nixon, and A. Prudel-Bennett, "New area based metrics for automatic gait recognition," in *Proc. Brit. Mach. Vis. Conf.*, 2001, pp. 233–242.
- [22] L. Wang, T. Tan, W. Hu, and H. Ning, "Automatic gait recognition based on statistical shape analysis," *IEEE Trans. Image Process.*, vol. 12, no. 9, pp. 1120–1131, Sep. 2003.
- [23] S. Chen and Y. Gao, "An invariant appearance model for gait recognition," in *Proc. IEEE Conf. Multimedia Expo*, 2007, pp. 1375–1378.
- [24] <http://www.cbsr.ia.ac.cn/english/Gait%20Databases.asp>
- [25] J. Yang, X. Wu, and Z. Peng, "Gait recognition based on difference motion slice," in *Proc. 8th Int. Conf. Signal Process.*, Nov. 16–20, 2006, vol. 4, pp. 2428–2431.
- [26] S. Hong, H. Lee, I. Nizami, and E. Kim, "A new gait representation for human identification: Mass vector," in *Proc. 2nd IEEE ICIEA*, May 2007, pp. 669–673.
- [27] I. Nizami, S. Hong, H. Lee, S. Ahn, K.-A. Toh, and E. Kim, "Multiview gait recognition fusion methodology," in *Proc. 3rd IEEE ICIEA*, Jun. 2008, pp. 2101–2105.
- [28] T. H. W. Lam, T. W. H. A. Jeong, and R. S. T. Lee, "Silhouette spatiotemporal spectrum (SStS) for gait-based human recognition," in *Proc. Int. Conf. Adv. Pattern Recog.*, 2005, pp. 309–315.
- [29] K. Bashir, T. Xiang, and S. Gong, "Feature selection on gait energy image for human identification," in *Proc. IEEE ICASSP*, Mar. 31–Apr. 4, 2008, pp. 985–988.
- [30] S. Yu, D. Tan, and T. Tan, "A framework for evaluating the effect of view angle, clothing and carrying condition on gait recognition," in *Proc. 18th ICPR*, 2006, vol. 4, pp. 441–444.

- [31] R. Mukundan and K.R. Ramakrishnan. “Moment functions in image analysis: Theory and Applications”.World Scientific.(1998).
- [32] M.K. Hu, “ Visual pattern recognition by moment invariants,” IRE. Trans. Inform.Theory 8, 179-187(1962).
- [33] Michela Goffredo, Imed Bouchrika, John N. Carter, “Self-Calibrating View-Invariant Gait Biometrics”, IEEE Transactions on Systems, Vol. 40, No. 4, pp. 997-1008, 2010
- [34] R. Mukundan, S.H. Ong, P.A.Lee, “Discrete vs. Continuous orthogonal moments for image analysis”, International Conference, (2001)
- [35] M.R. Teague, Image analysis via the general theory of moments, J. Opt. Soc. Am. 70 (1980) 920–930.
- [36] M. Born, E. Wolf, Principle of Optics, sixth ed., Pergamon Press, Oxford, 1980, pp. 767–772.
- [37] Y. Sheng, L. Shen, Orthogonal Fourier–Mellin moments for invariant pattern recognition, J. Opt. Soc. Am. A 11 (6) (1994) 1748–1757.
- [38] Z.L. Ping, R. Wu, Y.L. Sheng, Image description with Chebyshev–Fourier moments, J. Opt. Soc. Am. A 19 (9) (2002)1748–1754.
- [39] C. Toxqui Quitl. Clasificación de objetos en movimiento usando momentos de Jacobi-Fourier y la MTF del sistema óptico digital". Tesis doctoral. Instituto Nacional de Astrofísica, Óptica y Electrónica, 2010.
- [40] Shiqi Yu, Tieniu Tan, Kaiqi Huang, A Study on Gait-Based Gender Classification. IEEE Trans. on Image Processing. pp:1905-1910,V18(8),2009.
- [41] C. Toxqui-Quitl, A. Padilla-Vivanco, and J. Báez-Rojas “Classification of motion-blurred images by means circular moments and a controlled vibration system.” Applications of Digital Image Processing XXXII Proc. of SPIE, Vol. 7389, (2009).
- [42] R. C. Gonzalez and R. E. Woods, Digital Image Processing, 3rd ed., Prentice Hall, Upper Saddle River, NJ, 2008

- [43] Shen, D. and Horace, H.S. “Discriminative wavelet shape descriptor for recognition of 2-D patterns”, *Pattern. Recognition* 32, pp 151 – 165, 1999.
- [44] R. Gross and J. Shi, “The cmu motion of body (MoBo) database,” Robotics Institute, Carnegie Mellon University, Pittsburgh, Pa, USA, 2001.
- [45] J. Shutler and M. S. Nixon, “Zernike velocity moments for sequencebased description of moving features,” *Image Vis. Comput.*, vol. 24, no. 4, pp. 343–356, Apr. 2006.
- [46] Goldberg D. E. (1989). *Genetic Algorithms in Search, Optimisation and Machine Learning*. Addison-Wesley Publishing Company, Reading.
- [47] Sivanandam S.N., Deepa S.N. (2008) *Introduction to Genetic Algorithms*, Springer