Universidad Nacional de Colombia
Facultad de Medicina

Extracción y Seguimiento de los Miembros Inferiores Durante la Marcha

Lorenza Henao Murillo

Director:
Eduardo Romero Castro, Ph.D.
Profesor Asociado, Depto. de Imágenes Diagnósticas
Facultad de Medicina

Bogotá D.C., 2 de febrero de 2011

Nota de aceptación

_____

_____

_____

_____

_____

_____

_____

Firma del jurado

_____

Firma del jurado

Bogotá D.C., 2 de febrero de 2011

A los amores de mi vida:
Mis papás, Melquisedec y Emma.
Mis hermanos, Marcelo y Andrea.
Mi esposo, Andrés Julián.

Agradecimientos

Expreso mi profunda gratitud a mi director de tesis, Eduardo Romero, por incentivar en mí la búsqueda del conocimiento.
A Antoine Manzanera, por sus valiosos aportes en el desarrollo de mi investigación.
A mis jurados, Carlos Julio Cortés y Fabio Martínez por su apreciación y evaluación en este trabajo.
A mis compañeros de la maestría y del grupo de investigación *Bioingenium*, por su apoyo y colaboración.
A Iván Jaramillo y Gracielita del Depto. de Ingeniería Electrónica, por estimularme a seguir adelante.
A Alba Patricia Arias, Luis Alberto Toro, Alejandro Gómez, Carlos Barco Gómez y la Universidad Autónoma de Manizales, por el respaldo brindado.
A mis grandes amigos y colaboradores, Jeimy, Luigi, Nancy, Boni, Vicentico y Aleja, por su constante acompañamiento.
A quienes con su buena voluntad participaron en todo el proceso de capturas de video.
A Edgar Carreño, por impulsarme a hacer esta maestría.
A mi familia, por su amor, paciencia y apoyo.
A todos los demás que estuvieron conmigo durante este proceso.

Contenido

1. Introducción **1**

2. Captura de Movimiento sin Marcadores **4**
2.1. Modelos 2D 6
2.1.1. Segmentos 6
2.2. Modelos 3D 7
2.2.1. *Superquadrics* 7
2.2.2. Cilindros y poliedros 8
2.2.3. Modelos deformables 8
2.2.4. Estrategia de identificación de las partes del cuerpo humano (HBPIS) 9
2.2.5. Algoritmo iterativo del punto más cercano (ICP) . . . 9
2.2.6. Cadenas cinemáticas 9
2.2.7. Superelipsoides 11
2.2.8. Voxeles 11
2.3. Métodos sin modelo 11

3. Conceptos Preliminares **13**
3.1. Análisis de marcha 13
3.1.1. Análisis dinámico del movimiento 15
3.1.2. Captura de movimiento 16
3.2. Estimación de movimiento 21
3.2.1. Detección del objeto 21
3.2.2. Seguimiento de puntos clave en un objeto 29

4. Extracción y Seguimiento de los Miembros Inferiores **34**
4.1. Captura de video 36
4.2. Segmentación 36
4.3. Transformación geodésica y esqueletización 37

4.4. Detección y seguimiento de centros articulares 37
4.5. Resultados generales y análisis 39
4.6. Validación de resultados 46
4.6.1. Discusión 56

5. Productos de la Tesis **60**

6. Conclusiones y Perspectivas Futuras **62**

A. Algoritmos Implementados **64**
A.1. Estimación de Fondo $\Sigma - \Delta$ 64
A.2. Filtro *Kalman* 65
A.3. Flujograma Método Completo 66

B. Transformación de Distancia **67**
B.0.1. Distancia Geodésica 69

C. Representación del Cuerpo Humano **70**

Bibliografía **73**

Lista de Figuras

2.1. Diagrama estado del arte 5
2.2. Representaciones del cuerpo humano 6

3.1. Ciclo de marcha 16
3.2. Estimación de movimiento 22
3.3. Sustracción de cuadros consecutivos 23
3.4. Máscara de detección de cambio 24
3.5. Sustracción de fondo utilizando filtro exponencial recursivo clásico 25
3.6. Diagrama sistema de modulación $\Sigma - \Delta$ 27
3.7. Diagrama cíclico del filtro *Kalman* 32

4.1. Diagrama de la metodología utilizada 35
4.2. Resultado estimación Σ-Δ, secuencia 1 en laboratorio de marcha 39
4.3. Diferentes secuencias de marcha y la estimación Σ-Δ correspondiente 40
4.4. Proceso y resultados de la aplicación con iluminación semi-controlada en un niño. 42
4.5. Proceso y resultados de la aplicación con iluminación semi-controlada en hombre adulto 43
4.6. Proceso y resultados de la aplicación con iluminación natural en mujer adulta 1 44
4.7. Proceso y resultados de la aplicación con iluminación natural en mujer adulta 2 45
4.8. Comparación entre algunos métodos de detección de objeto . 47
4.9. Representación ángulos articulares 48
4.10. Desplazamientos angulares teóricos de cadera y rodilla 49
4.11. Desplazamientos angulares de cadera y rodilla en un niño . . 52
4.12. Desplazamientos angulares de cadera y rodilla en hombre adulto 5 53

4.13. Desplazamientos angulares de cadera y rodilla en mujer adulta 1 54
4.14. Desplazamientos angulares de cadera y rodilla en mujer adulta 2 55

A.1. Diagrama filtro *Kalman* 65
A.2. Flujograma método completo 66

B.1. Transformada de distancia 67
B.2. Carta de distancias 68
B.3. Métrica euclidiana 69

C.1. Modelo articulado 71

Lista de Tablas

4.1. Coeficientes de similitud de *Jaccard*. 46
4.2. Distancia de *Fréchet. Iluminación semicontrolada.* 56
4.3. Distancia de *Fréchet. Iluminación natural* 57

Resumen

El propósito principal de esta investigación ha sido el desarrollo de métodos que permitan el análisis de la marcha sin uso de marcadores. Los marcadores alteran el gesto natural del movimiento, son incómodos y su ubicación varía entre experimentos sucesivos. El método desarrollado captura el movimiento en el plano sagital mediante una cámara convencional, estima el fondo y segmenta la silueta de los miembros inferiores, y mediante una transformación de distancia que esqueletoniza la silueta, detecta y sigue los puntos articulares que corresponden a la cadera, rodillas y tobillos. El método se evaluó en 22 diferentes videos, 12 capturados en condiciones semicontroladas y 10 videos en condiciones ambientales. Los resultados obtenidos se compararon con aquellos obtenidos por el método convencional (con marcadores), mediante un estimador llamado distancia de *Fréchet* encontrándose una similitud del 65 % entre los desplazamientos angulares de cadera y rodilla. Este método podría ser utilizado en aplicaciones clínicas de análisis de marcha.

Abstract

The main goal of this investigation has been the development of methods which allow markerless gait analysis. Markers alter the natural movement gesture, they are uncomfortable and their location vary between consecutive experiments. The developed method captures the movement in sagittal plane using a conventional camera, estimates the background and segmentates the lower limb silhouette. By means of a distance transformation it skeletonizes the silhouette, detects and tracks the joint points corresponding to the hip, knees and ankles. The method was evaluated in 22 different videos, 10 captured in semi-controlled conditions and 12 in environmental conditions. The obtained results were compared with those obtained by the conventional method (with markers), by means of *Fréchet* distance estimation, finding a 65 % similarity between angular displacements of hip and knee. This method could be easily used in clinical applications.

Capítulo 1

Introducción

El análisis de marcha es la medición, descripción y evaluación sistemática de los parámetros cinemáticos y cinéticos que caracterizan el movimiento humano [17]. Este proceso consta de tres pasos:

- Medición y adquisición de los datos cinemáticos y cinéticos.
- Análisis de los datos que proporcionan información importante de la marcha.
- Interpretación de los diferentes tipos de información de la marcha como una forma de evaluación.

El análisis clínico de marcha es una herramienta de evaluación y no de diagnóstico, la cual le permite al especialista obtener los parámetros necesarios para realizar comparaciones y distinguir una marcha patológica de una normal [17]. El conjunto de datos cuantitativos utilizados para la descripción de la marcha del individuo incluyen entre otros, orientación de los segmentos de los miembros a través de un ciclo de marcha (cinemática), fuerzas de reacción generadas durante la fase de apoyo (cinética), actividad bioeléctrica de los músculos involucrados en la marcha (electromiografía) y consumo energético. Estas evaluaciones son muy importantes debido a que proporcionan descripciones objetivas de la marcha, como es el caso del análisis cinemático, que permite obtener datos generales como velocidad, cadencia y longitud del paso. Por ejemplo en la patología de hemiplejía espástica cualquier reducción que se presente en la fase de apoyo monopodal, refleja inestabilidad, aspecto que conduce a una reducción de la longitud del paso y de la duración del apoyo simple [60].

El seguimiento del movimiento para el análisis de marcha se agrupa en dos grandes categorías [50]: sistemas que utilizan marcadores y sistemas que no lo hacen.

Los sistemas que utilizan marcadores, como VICON [58] tienen la particularidad de que la persona debe usar múltiples marcadores, situados en puntos claves de los miembros inferiores del paciente, con lo cual se obtiene información acerca del centro de masa de cada segmento articular [14]. Estas soluciones son normalmente más robustas que aquellas que no requieren de marcadores porque soportan la variación rápida y compleja de posturas. Sin embargo, las soluciones que utilizan marcadores tienen diversos inconvenientes [9, 14, 20, 21, 24, 28, 29, 31, 53, 55]:

- Ubicar marcadores en el cuerpo de la persona puede ser difícil de manejar, molesto e incómodo.
- Introducen artefactos experimentales.
- Alteran el gesto natural.
- La información del movimiento se limita a ciertos puntos característicos.
- La ubicación de los marcadores varía entre experimentos sucesivos.
- Los marcadores pueden ubicarse en tejido blando, con lo cual no se obtiene un dato preciso de la localización de las articulaciones.
- La superficie de la piel cambia cuando la persona está en movimiento, con lo cual se modifica la posición de los marcadores.
- Puede presentarse oclusión.
- Implica el uso de equipo especializado y costoso.
- Requieren que las medidas sean tomadas en un ambiente especialmente adecuado para las evaluaciones.

Las soluciones sin marcadores intentan eliminar los inconvenientes generados por los métodos que requieren de estos dispositivos, utilizando solamente visión por computador para observar el gesto. La tendencia actual en captura de movimiento sin marcadores es unir el contenido de una o múltiples vistas de una persona en modelos 2D ó 3D, los cuales se usan para estimar y analizar la postura humana en cada instante de tiempo.

El análisis de movimiento sin marcadores requiere primero la extracción del objeto de interés en una secuencia de video [26]. Segundo, el seguimiento que consiste en detectar y encontrar la correspondencia de los segmentos y articulaciones del cuerpo humano en la secuencia de video, conociendo su posición en un cuadro previo.

Después de haber estudiado los distintos inconvenientes que tiene el análisis de marcha con marcadores, el objetivo en esta tesis fue desarrollar e implementar un método que permite extraer la silueta de los miembros inferiores, detectar los centros articulares y seguirlos a lo largo de una secuencia de video sin utilizar marcadores, con el fin de evitar las incomodidades que implica el uso de estos dispositivos.

Este documento está organizado como sigue: en el Capítulo 1 se hace un preámbulo sobre el presente documento, en el Capítulo 2 se presenta el estado del arte sobre la captura de movimiento sin marcadores; algunos conceptos preliminares sobre el análisis de marcha y las técnicas involucradas en el desarrollo de este trabajo se muestran en el Capítulo 3.

En el Capítulo 4 se explica la metodología desarrollada. En la sección 4.1 se describen las características de la captura de video, la sección 4.2 define la extracción del objeto binario mediante un método de estimación de fondo estático, en la sección 4.3 se detalla el cálculo de una transformación geodésica sobre la máscara binaria y un proceso de adelgazamiento mediante operadores morfológicos para obtener el esqueleto de la silueta. La detección y seguimiento de los puntos articulares sobre toda la secuencia de video se describen en la sección 4.4. Los resultados del proceso y la validación del método se muestran en las secciones 4.5 y 4.6 respectivamente.

En el Capítulo 5 se exponen los productos derivados de esta tesis y finalmente en el Capítulo 6 se presentan las conclusiones y perspectivas futuras.

Capítulo 2

Captura de Movimiento sin Marcadores

La captura de movimiento sin marcadores es un área de interés para muchos investigadores con aplicaciones en vigilancia, biometría, interacción hombre-máquina, animación, desempeño de deportistas, clínicas y biomecánicas. En esta última se han desarrollado sistemas sin marcadores para estimación de centros articulares, seguimiento de segmentos de los miembros inferiores, análisis de discapacidades del movimiento y estimación de posturas en el trabajo. Estos estudios demuestran la aplicabilidad de técnicas en visión por computador para análisis automático de movimiento humano, sin embargo carecen de validación contra datos basados en marcadores [31].

Los sistemas basados en captura de movimiento sin marcadores varían de acuerdo al número de cámaras utilizadas, la representación de los datos capturados, tipos de algoritmos y la aplicación a regiones específicas o a todo el cuerpo humano. La captura de movimiento humano se divide normalmente en [55]:

- Inicialización, el sistema comienza su operación con una correcta interpretación de la escena actual.
- Seguimiento, obtiene la segmentación y una posterior correspondencia en uno o más cuadros.
- Estimación de postura en uno o más cuadros.
- Reconocimiento del gesto para identificación de los individuos, así como las acciones, actividades y comportamientos desempeñados por una

o varias personas en uno o más cuadros.

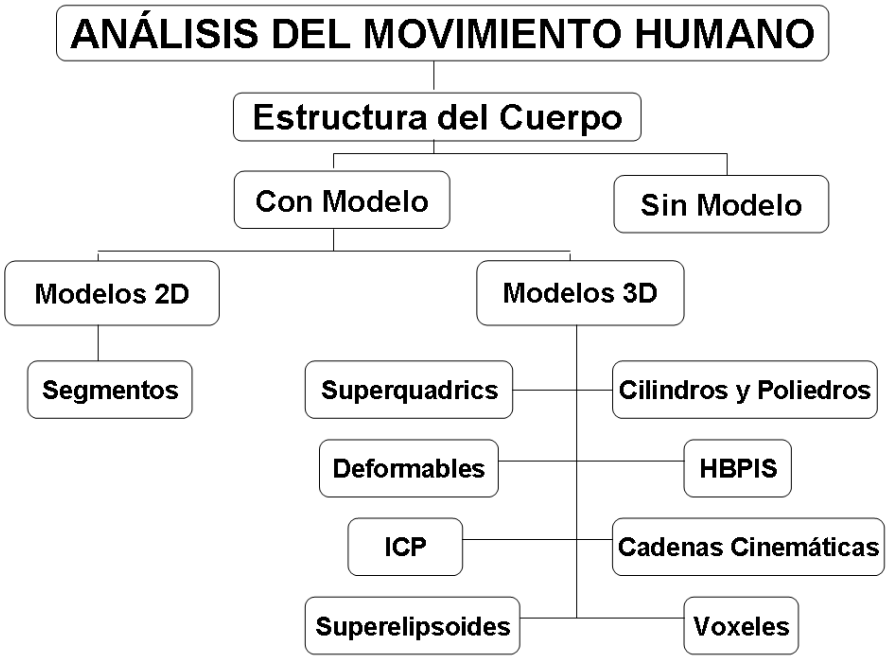


Figura 2.1. Diagrama estado del arte.

La mayoría de estas investigaciones utilizan modelos del cuerpo humano en dos (2D) o tres (3D) dimensiones, en las cuales se hace seguimiento de la información anatómica y cinemática relevante o se realiza un ajuste de éste a planos 2D de la imagen o a representaciones 3D [31]. Los modelos 3D normalmente constan de dos características: la estructura del esqueleto (figura de segmentos) y el músculo que lo rodea (basada en superficie). La primera característica consiste en una colección de segmentos y ángulos articulares, con varios grados de libertad en el lugar de la articulación; la segunda utiliza polígonos, cilindros y volúmenes en general para representar las diferentes formas de las partes del cuerpo [21]. Se han propuesto diferentes tipos de modelos: segmentos [35, 47, 69], *superquadrics* [1] [9, 20, 21, 55], cilindros, poliedros [7, 14], deformables [64], estrategia de identificación de las partes del cuerpo humano (HBPIS - *Human Body Part Identificaction Strategy*) [28], algoritmo iterativo del punto más cercano (ICP - *Iterative Closest*

[1] *Superquadrics*: familia de objetos tridimensionales paramétricos, útiles para el modelamiento geométrico y de sencilla representación matemática.

Point) [31], cadenas cinemáticas [51, 53], superelipsoides [29], voxeles [2] [50], entre otros, que serán descritos más adelante.

2.1. Modelos 2D

2.1.1. Segmentos

Haciendo captura de movimiento con una cámara desde el plano sagital, Persson [47] desarrolló un método 2D libre de marcadores, con el cual acopló un sencillo modelo articulado (por ejemplo, ver imagen 2.2(d)) a los contornos de uno de los miembros inferiores a través de transformaciones de distancia. Hizo seguimiento de los segmentos del objeto en cada cuadro y realizó un trabajo experimental sin aplicación en casos reales, con lo cual obtuvo la ubicación de las articulaciones del tobillo, la rodilla y la cadera, las posiciones de los segmentos y el ángulo de orientación de la rodilla. El inconveniente que se encontró con este método fue que durante la marcha los contornos cambian su forma y se mueven con relación a los huesos debido a la contracción del músculo, problema similar al método con marcadores, cuando éstos no se encuentran adecuadamente adheridos.

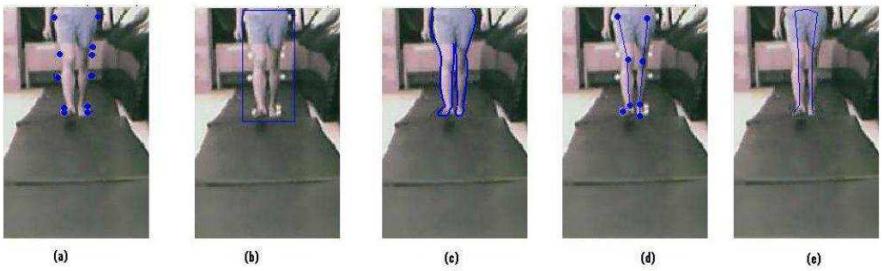


Figura 2.2. Representaciones del cuerpo humano. (a) Múltiples puntos, (b) pieza rectangular, (c) contorno, (d) modelo articulado, (e) esqueleto.

Con el propósito de obtener información útil relacionada con la posición y dinámica de las extremidades inferiores, Zhou *et al.* [69] propusieron una estructura bayesiana para introducir conocimiento a priori. Este conocimiento fue construido a partir de un sencillo modelo articulado y de parámetros

[2] Voxel: unidad cúbica que compone un objeto tridimensional. Constituye la unidad mínima procesable de una matriz tridimensional y es el equivalente del pixel en un objeto 2D.

tanto estáticos como dinámicos. La aplicación de este estudio se orientó hacia el reconocimiento de personas. Verificaron la aproximación sobre siluetas extraídas de imágenes de personas, caminando en ambientes tanto interiores como exteriores capturados a partir de una cámara.

Por otra parte, Goffredo *et al.* [35] realizaron captura de movimiento con una cámara de amplio rango de direcciones, aplicando un método estándar de sustracción de fondo para imágenes RGB, estimando la pose de los miembros inferiores, siguiendo resultados de estudios anatómicos e implementando un algoritmo de reconstrucción para normalizar y corregir las inclinaciones de los miembros inferiores. Obtuvieron una aproximación 2D de la estimación de la postura cuadro a cuadro y de los ángulos de los miembros inferiores para identificación de la marcha en aplicaciones de vigilancia.

2.2. Modelos 3D

2.2.1. *Superquadrics*

Algunos estudios se han orientado a la reconstrucción 3D a partir de múltiples cámaras. En el trabajo de Gavrila *et al.* [21] se implementó un modelo 3D basado en *tapered superquadrics*, estimando el movimiento, reconstruyendo y haciendo correspondencias, obtuvieron un sistema de visión para seguimiento de movimientos humanos sin restricciones, basado en modelos 3D y desde múltiples vistas para aplicaciones de realidad virtual y vigilancia.

Marzani *et al.* [20] describieron una aproximación del análisis del movimiento del cuerpo humano en un espacio 3D utilizando tres cámaras. Extrajeron puntos característicos de las fronteras del cuerpo en las imágenes iniciales, ajustaron *superquadrics* 2D a estos puntos mediante un proceso de agrupamiento difuso y finalmente reconstruyeron un modelo 3D basado en *superquadrics* articulados, cada uno de los cuales representa diferentes partes de la pierna. Con el sistema que desarrollaron se pueden obtener parámetros dinámicos del movimiento, características utilizadas en el análisis de discapacidades de la marcha.

Calais *et al.* [9] encontraron que se hace necesario mejorar diferentes etapas de todo el proceso para obtener un sistema totalmente confiable. Estas etapas incluyen la inicialización, con el fin de asegurar una adecuada predicción de la ubicación de la pierna humana en la escena, lo cual conduce a un mejor seguimiento de la persona en la secuencia.

Sundaresan *et al.* [55] presentaron un método basado en visión por computador para realizar captura de movimiento sin marcadores. Modelaron el cuerpo humano como un conjunto de *superquadrics* conectados a una estructura articulada, calcularon la representación de los datos de volumen (voxeles) a partir de imágenes capturadas desde múltiples cámaras y propusieron un algoritmo de seguimiento similar al filtro *Kalman* extendido, para estimar el movimiento articulado utilizando múltiples entradas que combinan tanto la información espacial como temporal. Finalmente obtuvieron un modelo articulado y una estimación exacta de la postura, siendo su contribución más notable la realización del seguimiento para múltiples señales, combinando las señales espaciales de manera intuitiva.

2.2.2. Cilindros y poliedros

Utilizando otro tipo de modelo, Meyer *et al.* [14] describieron una aproximación para la extracción de objetos articulados, orientada al análisis de marcha. Realizaron detección automática de las partes del cuerpo, seguimiento de cada parte, basándose en contornos y estimación de la forma y posición de sencillos bloques 3D, relacionando el contorno extraído de la imagen con la proyección del bloque 3D dentro del plano de ésta. Obtuvieron como resultado las trayectorias de las partes en movimiento, las cuales hacen posible la clasificación entre diferentes tipos de desórdenes de la marcha. Wan *et al.* [7] propusieron un método que no requiere segmentación de primer plano como paso introductorio esencial. Combinaron el modelo del cuerpo humano con los parámetros de postura y las imágenes de múltiples cámaras, además realizaron la segmentación y captura de movimiento simultáneamente mediante contornos activos y *level sets*[3]. Con este método obtuvieron resultados alentadores sobre imágenes reales, lo cual demuestra su potencial.

2.2.3. Modelos deformables

Utilizando otro tipo de modelo, Lu *et al.* [64] establecieron una estructura de captura del movimiento humano sin marcadores. Hicieron una mejora de un modelo de deformación de piel, segmentaron las siluetas de imágenes multivista, utilizando una mezcla de Gaussianas para sustracción de fondo estático y finalmente incorporaron la postura humana y las siluetas con un filtro de partículas (*Annealed Particle Filtering*). Los experimentos obtenidos mostraron que la aproximación propuesta fue adecuada para resolver el problema de captura de movimientos humanos complejos.

[3] *Level set*: técnica numérica utilizada para el seguimiento de superficies y formas.

2.2.4. Estrategia de identificación de las partes del cuerpo humano (HBPIS)

Kakadiaris *et al.* [28] desarrollaron un sistema de captura de los miembros superiores de varios sujetos, realizando movimientos 3D complejos en presencia de oclusión significativa, utilizando múltiples cámaras. Construyeron un banco de pruebas para captar parámetros geométricos y cinemáticos, obtuvieron un modelo a priori del sujeto e hicieron predicción de oclusiones entre las partes del cuerpo, utilizando un algoritmo de identificación de las partes del cuerpo (HBPIS) para extraer la forma tridimensional de los miembros superiores del sujeto. Luego hicieron estimación de movimiento y por último, predijeron el movimiento del modelo con un filtro *Kalman*. Los datos cinemáticos y geométricos que adquirieron con el sistema proporcionan medidas importantes en el diseño de dispositivos de rehabilitación. Los autores afirman que no existe restricción para usar este sistema en los miembros inferiores.

2.2.5. Algoritmo iterativo del punto más cercano (ICP)

Mundermann *et al.* [31] presentan una novedosa aproximación para captura exacta del movimiento sin marcadores, en la cual combinaron un preciso algoritmo de seguimiento con una base de datos de modelos articulados que se compone de formas humanas y datos antropométricos. En la etapa de seguimiento de los segmentos del cuerpo en secuencias de *visual hulls* [4] emplearon un algoritmo iterativo de puntos articulados más cercanos (ICP). La base de datos proporciona modelos articulados que relacionan muy bien la apariencia exterior de los *visual hulls*. Acoplaron esta información con una cadena cinemática mejorada por ecuaciones de regresión antropométrica. Las desviaciones en las cadenas cinemáticas de la ubicación de centros articulares verdaderos, fueron compensadas por un método de restricción de articulaciones suaves. Finalmente obtuvieron una salida exacta y más correcta anatómicamente hablando, apropiada para aplicaciones biomecánicas y clínicas.

2.2.6. Cadenas cinemáticas

Capturando el movimiento con una cámara y haciendo uso de un modelo compuesto de cadenas cinemáticas, Saboune *et al.* [53] extrajeron posiciones 3D de puntos característicos del cuerpo humano, los cuales pueden

[4] *Visual hull*: representación 3D estándar de secuencias dinámicas a partir de múltiples imágenes.

ser útiles para el análisis de marcha. Con el argumento de que de acuerdo a estudios recientes, las caídas se están convirtiendo en la primera causa de accidentes mortales entre personas de la tercera edad, propusieron una metodología para detectar la tendencia de estas caídas en adultos mayores, observando sus actividades en el hogar. Obtuvieron la silueta de la persona aplicando sustracción de fondo pixel a pixel, ajustaron un modelo sencillo y genérico a la imagen segmentada usando función de verosimilitud e hicieron seguimiento aplicando un filtro de partículas a intervalos. Finalmente extrajeron parámetros de marcha, los cuales fueron comparados con sistemas basados en marcadores, encontrándolos alentadores. El uso de varias cámaras podría producir mejores resultados, sin embargo, no lo hicieron porque esto complicaría el método.

El propósito del trabajo presentado por Corazza *et al.* [51] consistió en desarrollar y probar un método que proporciona un modelo de formas superficiales, libre de fisuras, de las formas del cuerpo humano de un sujeto, junto con las ubicaciones de los centros articulares. Una de las contribuciones de este trabajo fue expandir el espacio de formas humanas [13], introduciendo información de la ubicación de diez centros articulares de un sujeto específico. Las ubicaciones óptimas de los centros articulares en las mallas 3D se lograron a través de regresión lineal. El algoritmo de generación automática del modelo fue aplicado a mallas 3D de diferentes calidades y resolución como escáneres láser y *visual hulls*. Utilizando un espacio continuo, desarrollaron un algoritmo iterativo de registro para posturas y formas, siguiendo la generación automática de un modelo del sujeto específico a partir de una postura estática capturada. El algoritmo de registro se construyó utilizando un espacio de formas humanas, teniendo en cuenta un modelo articulado de segmentos rígidos, segmentación de las diferentes partes del cuerpo en la malla de datos y transformación global inversa para cada segmento en la malla de datos. El resultado más importante de este trabajo fue la generación automática de un modelo a partir de un *visual hull* sin necesidad de utilizar *hardware* dedicado y costoso. Aportaron además valores biomecánicos cuantitativos a la base de datos del espacio de formas humanas utilizada en este trabajo, lo cual representa una herramienta disponible para la generación de modelos de mallas sin fisuras, con información exacta de modelos cinemáticos.

2.2.7. Superelipsoides

Kehl *et al.* [29] propusieron un método, con el cual integran múltiples señales de la imagen como bordes, información de color y reconstrucción volumétrica. La combinación de esta información ayuda al seguidor a manejar ambigüedades como extremidades en contacto o fuertes oclusiones de las partes del cuerpo. Acoplaron un modelo articulado, construido de superelipsoides. Aplicaron el método de optimización SMD (*Stochastic Meta Descent*), el cual se basa en un gradiente descendente y es utilizado para encontrar la mejor postura que se ajusta a las imágenes. El resultado final fue un seguimiento robusto de movimientos complejos de todo el cuerpo humano.

2.2.8. Voxeles

A diferencia de las investigaciones presentadas hasta ahora, el estudio de Beriault *et al.* [50] se centró en mejorar los aspectos de diseño y calibración de sistemas multicámara y reconstrucción de la forma de la silueta. Propusieron un análisis cuidadoso con el objetivo de lograr un óptimo sistema sincronizado multicámara. Después hicieron una exacta calibración utilizando un método flexible, el cual permite un posicionamiento libre de la cámara. Para la extracción de la silueta, formularon un novedoso método basado en regiones, que no requiere de fondos estáticos. Finalmente obtuvieron una reconstrucción de los datos en voxeles a color en escenas complejas y sin restricciones. En este trabajo se interesaron particularmente en monitorear el movimiento humano, utilizando visión pasiva para la evaluación del desempeño de un pianista, así como para la prevención de lesiones. En este estudio encontraron además que el diseño de sistemas para la captura de movimiento sin marcadores es pasado por alto en muchas ocasiones, sin embargo cuando se escoge un equipo de captura inadecuado o un diseño del sistema incorrecto, se puede atenuar la calidad espacial y temporal de la entrada de video, con lo cual se puede afectar el análisis de la postura.

2.3. Métodos sin modelo

Existen otros trabajos dirigidos al análisis del movimiento, que no utilizan modelo alguno. Chu *et al.* [12] desarrollaron una aproximación para captura de movimiento de estructuras cinemáticas articuladas, libre de marcadores. Utilizaron múltiples cámaras y no usaron un a priori de modelo

cinemático, razón por la cual no realizaron el paso de inicialización. Desarrollaron un método libre de modelo (NSS - *Nonlinear Spherical Shells*) para extraer los puntos característicos del esqueleto, los cuales fueron unidos en una estructura de árbol para un frame particular en movimiento. Con un método de segmentación de primer plano, tomaron todas las siluetas y formaron un volumen. Valiéndose de la información de los puntos característicos y el volumen obtenido, determinaron la postura. Aunque no especificaron una aplicación clínica del método, encontraron que la velocidad relativa del procedimiento de captura hace el procesamiento de cada cuadro manejable, la implementación no requiere un modelo explícito de cinemática humana, ni proceso de inicialización y tampoco optimización de los parámetros con respecto al volumen. Finalmente, la curva del esqueleto puede ser útil para la representación de la postura.

Haciendo uso de codificación rápida y robusta de formas, lo cual produce un espacio vectorial sencillo y aplicando clasificación de datos vectoriales mediante reconocimiento de patrones y técnicas de aprendizaje de máquina, Bauckhage *et al.* [6] obtuvieron una serie de vectores característicos, después de mapear una cuadrícula sobre una silueta un poco rotada, que son un fuerte indicador de una marcha anormal.

Realizando una sustracción de fondo mediante umbralización de las tres componentes de color en el espacio YUV y estimando la postura de los miembros inferiores con base en proporciones de los segmentos corporales, Goffredo *et al.* [34] desarrollaron un método para la detección de los puntos articulares de los miembros inferiores, incluso en situaciones de oclusión. Validaron los resultados obtenidos contra secuencias de imágenes reales y sintéticas, encontrando buena exactitud en la estimación de la postura de los miembros inferiores. Adicionalmente compararon sus resultados con datos encontrados en estudios biomecánicos y concluyeron que el método propuesto puede utilizarse en el contexto médico.

Capítulo 3

Conceptos Preliminares

3.1. Análisis de marcha

La marcha se define como una forma de locomoción caracterizada por periodos de apoyo y balanceo de los miembros inferiores. El análisis de marcha estima de forma objetiva parámetros cinemáticos, cinéticos y electromiográficos del movimiento, los cuales en conjunto permiten determinar la acción muscular predominante durante el movimiento realizado y si la contracción muscular produce o consume energía [60]. El análisis cinemático de la marcha es la medición de los desplazamientos angulares de cada segmento articulado, de la traslación de los segmentos y de toda la masa corporal durante el movimiento del cuerpo al caminar; lo que lleva a realizar comparaciones necesarias para distinguir una marcha patológica de una marcha normal [57]. En dicho análisis es necesario conocer parámetros del movimiento como aceleración angular, velocidades y desplazamientos de diferentes partes del cuerpo, especialmente las piernas, además de las excursiones [1] de cadera y rodilla [14]. Algunas aplicaciones clínicas del análisis de marcha incluyen [9, 14, 17, 31, 56]:

- Diagnóstico de patologías mecánicas asociadas a enfermedades del sistema musculoesquelético.
- Planificación prequirúrgica.
- Seguimiento postquirúrgico.
- Desarrollo y evaluación de tratamientos de rehabilitación.

[1] Excursión articular: rango de flexión y extensión dentro del cual se mueve una articulación.

- Análisis del consumo de oxígeno.
- Evaluación de resultados terapéuticos con medicación oral.
- Técnicas de neurofacilitación.
- Evaluación de la locomoción en casos de parálisis cerebral para determinar una cirugía o fijar una órtesis.
- Observación del progreso de desórdenes neuromusculares como enfermedad de *Parkinson* o distrofia muscular.
- Estimación de la efectividad del reemplazo de una prótesis articular.
- Revisión de las mejoras en el diseño de órtesis.
- Cuantificación de los cambios en el diseño de prótesis.

La marcha necesita de gran integración sensorial y motriz para coordinar de forma correcta los movimientos de las distintas articulaciones.

Los orígenes del análisis de marcha se remontan al siglo XVII en Europa y continúan hasta principios del siglo XX. Descubrimientos tan notables como los de Borelli, Galvani, Newton, Descartes, Marey, Carlet, los hermanos Weber, Scherb, Duchenne, Muybridge, Braune y Fischer proporcionaron un sólido fundamento científico del entendimiento actual de la marcha humana. Los hermanos Weber reportaron uno de los primeros estudios cuantitativos de los parámetros de tiempo y distancia durante la locomoción humana [56]. Su trabajo constituyó un modelo para estudios cuantitativos posteriores del movimiento humano. Marey y Muybridge fueron los primeros en cuantificar patrones del movimiento humano utilizando técnicas fotográficas. Muybridge utilizó una serie de cámaras para tomar múltiples fotografías consecutivas de animales y seres humanos en movimiento [3, 4]. Braun y Fischer [3, 4] utilizaron los principios de la mecánica clásica newtoniana, la geometría de Descartes y conceptos matemáticos de Borelli para realizar mediciones de los movimientos de los segmentos corporales y calcular las fuerzas y energía invertida por las articulaciones. Aunque los principios de la investigación realizada por Braun y Fisher son válidos hoy en día, sus métodos de estudio eran bastante intensivos para permitir aplicaciones clínicas en pacientes.

Después de la segunda guerra mundial, la necesidad de comprender el movimiento humano produjo desarrollos de nuevas técnicas para el tratamiento de veteranos heridos e inválidos. Para tal fin, la Universidad de California proporcionó una gran fuente de conocimiento de la mecánica del movimiento humano y su trabajo fue la base de técnicas fundamentales empleadas actualmente para el estudio de la locomoción humana [3].

Las primeras técnicas de análisis de marcha incluyeron luz estroboscópica, cintas reflectivas y goniómetros manuales, electrogoniometría, película cinematográfica y sistemas de marcadores pasivos con entrada manual de las posiciones del marcador, mediciones automatizadas de movimiento, métodos de cálculo de ángulo articular, entre otros [57]. Hoy en día, este estudio se realiza en laboratorios de análisis de marcha. Las técnicas utilizadas comprenden captura del movimiento a través de múltiples cámaras, uso de módulos de electromiografía (EMG) y plataformas de fuerza [44].

Ciclo de marcha

La marcha es una actividad cíclica, en la cual se realizan movimientos secuenciales que se presentan entre dos contactos iniciales consecutivos del mismo pie (figura 3.1) y consta de dos fases para cada miembro, la fase de apoyo y la fase de balanceo. La fase de apoyo constituye el 60 % del ciclo, se presenta cuando el pie está en el suelo y soporta peso. Se divide en cinco etapas: contacto inicial (contacto de talón), respuesta de carga, posición media (apoyo unipodal), fase de apoyo final (apoyo bipodal final) y prebalanceo.

La fase de balanceo corresponde al 40 % del ciclo de marcha, se presenta cuando el pie no está sosteniendo peso y se está moviendo hacia adelante. Comprende tres etapas: balanceo inicial (aceleración), balanceo medio y balanceo final que termina en el momento del contacto inicial [66].

3.1.1. Análisis dinámico del movimiento

El análisis del movimiento se lleva a cabo desde el punto de vista dinámico: cinemático o cinético [22]. El análisis cinemático se refiere a las características del movimiento sin tener el cuenta las fuerzas que lo ocasionan. Este análisis implica la descripción del movimiento para determinar la posición, velocidad y aceleración de un objeto. El análisis cinemático puede aplicar al movimiento lineal o al angular. Mediante el análisis del movimiento lineal o angular es posible tener una mayor comprensión del movimiento

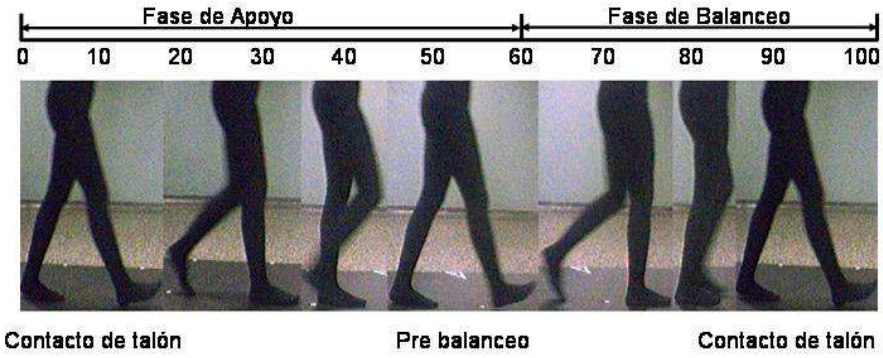


Figura 3.1. Fases del ciclo de marcha.

humano.

El análisis cinético estudia las fuerzas que producen, impiden o modifican el movimiento. Mediante este análisis, es posible conocer cuantitativamente la acción muscular, las restricciones óseas, capsulares y ligamentosas, además de las fuerzas de gravedad, inercia y de reacción del piso [33]. El análisis cinético en la marcha normal es útil para evidenciar patologías del sistema locomotor [60].

Aunque en el análisis dinámico intervienen tanto la cinemática como la cinética, en algunos casos, es suficiente realizar únicamente el análisis cinemático debido a que permite identificar las causas de alteraciones y anomalías en la marcha [33, 60].

3.1.2. Captura de movimiento

Como se mencionó anteriormente, la cinemática es el estudio del movimiento sin tener en cuenta las fuerzas que lo generan [22]. En el análisis cinemático del movimiento, es necesario realizar el seguimiento de los segmentos corporales cuando un paciente se mueve a través de un espacio [5]. Se han empleado varios métodos para estudiar la cinemática del cuerpo humano. En general, se hace captura de movimiento, tecnología utilizada para reunir información del desplazamiento de la persona; básicamente consiste en grabar y digitalizar un evento real, detectar y hacer seguimiento de un

número de puntos clave [2], regiones o segmentos, con lo cual finalmente se obtiene una representación bi o tridimensional de la caminata [46].

La captura del movimiento es un proceso importante para estudios en biomecánica. Se ha utilizado tradicionalmente para el apoyo al diagnóstico de patologías mecánicas asociadas a enfermedades musculoesqueléticas. Recientemente, ésta se ha utilizado también en el desarrollo y evaluación de tratamientos de rehabilitación e intervenciones preventivas para enfermedades musculoesqueléticas. Aunque el análisis del movimiento es clínicamente útil, el uso de rutinas para el análisis de marcha ha tenido un crecimiento limitado. El problema de su valor clínico, está asociado a diversos factores, incluyendo la aplicabilidad de las tecnologías existentes en problemas clínicos y la cantidad de tiempo y costos requeridos para la recopilación, procesamiento e interpretación de datos [32].

Existen diferentes técnicas para medir el movimiento absoluto de los segmentos corporales, las cuales varían en aspectos tales como exactitud, resolución, frecuencia de muestreo, alcance, espacio de trabajo, costo, limitación, conveniencia, susceptibilidad a cambios en la iluminación, facilidad de calibración, número de mediciones simultáneas, orientación y aplicación [18,61]. Las mediciones realizadas mediante sistemas mecánicos se basan en estructuras unidas que utilizan potenciómetros para determinar la rotación de cada articulación que esté disponible. Otros sistemas usan campos electromagnéticos o ultrasonido en el seguimiento de un grupo de sensores. Algunos emplean cámaras que digitalizan diferentes vistas del movimiento, las cuales se utilizan para ubicar la posición de los puntos clave o los marcadores. También se pueden diseñar combinaciones de dos o más de estas tecnologías, con el fin de reducir las limitaciones inherentes a cada una [46]. Las nuevas tecnologías buscan hacer seguimiento en tiempo real de un número ilimitado de puntos clave o de todos los segmentos de la persona, sin limitaciones de espacio y con un margen de error muy bajo. Se emplean sistemas no invasivos que usan procedimientos de visión por computador y pueden recuperar el movimiento a partir de datos ópticos utilizando modelos de movimiento [42]. Estos sistemas tienen aplicaciones en biomecánica, biometría, interacción hombre-máquina, animación y realidad virtual.

En aplicaciones clínicas, los métodos mecánicos son inadecuados porque

[2]Puntos clave: áreas que mejor representan el movimiento de las diferentes partes del cuerpo de la persona.

son incómodos, su ubicación varía con el movimiento de ciertas articulaciones [30] y limitan el movimiento del paciente al espacio permitido por el instrumento [18, 61]. Los sistemas electromagnéticos son sensibles a campos magnéticos en el ambiente y materiales ferromagnéticos en el espacio de movimiento [18], lo cual puede afectar la seguridad de pacientes que utilizan marcapasos o dispositivos similares. Los sistemas ultrasónicos por su parte, se han utilizado en robótica para calibración y en biomecánica para el seguimiento del movimiento. Para el seguimiento de puntos con una precisión y velocidad moderadas, los seguidores ultrasónicos son una alternativa razonable, sin embargo, debido a que se necesitan al menos tres puntos para inferir la orientación del cuerpo, es difícil medir toda la postura del cuerpo a tasas adecuadas [18]. De todos estos, el más popular y difundido para el análisis clínico de la marcha es el basado en visión por computador (video), de hecho los principales productos comercialmente disponibles para tal fin se basan en este método [30]. Estos métodos permiten que el usuario realice su movimiento en un espacio lo suficientemente amplio (6 m aproximadamente) para que el movimiento no sea limitado, aspecto indispensable para evitar alteraciones del gesto durante el análisis.

Sistemas de visión artificial

Los sistemas basados en visión artificial emplean cámaras que digitalizan diferentes vistas del movimiento, las cuales se utilizan para ubicar la posición de los puntos clave o los marcadores adheridos al objeto a ser seguido. Estos sistemas se agrupan en tres grandes categorías [28]: aquellos que usan marcadores activos (luminosos), aquellos que utilizan marcadores pasivos (reflectivos) o aquellos que no usan marcadores.

Los marcadores son dispositivos que al ubicarse sobre el objeto, son visibles. En los sistemas basados en visión artificial, generalmente se utilizan los dispositivos de acoplamiento de carga (cámaras CCD) para grabar el objeto a ser seguido y detectar el movimiento de los marcadores sobre éste. Normalmente se emplean múltiples cámaras con el fin que el objeto pueda ser seguido en tres dimensiones en lugar de dos. Mientras que sólo dos cámaras son necesarias para lograr un seguimiento tridimensional, es recomendable utilizar más para proporcionar múltiples resultados y de esta manera evitar oclusión de los marcadores.

Cada cámara se ubica de forma tal que se tenga una sola perspectiva de los marcadores. Posteriormente se hace triangulación de los marcadores para

seguirlos en tres dimensiones. Esta técnica revela el primer gran problema de los sistemas basados en visión por computador: determinar la correspondencia de los marcadores en cada vista de las cámaras. Con el fin de utilizar diversas vistas de un mismo marcador para triangular su posición, un marcador debe ser distinguido de los otros marcadores alrededor de éste.

- Marcadores activos.Los sistemas de marcadores activos utilizan marcadores que generan su propia luz; habitualmente hacen uso de diodos emisores de luz (LEDs) para definir los segmentos articulares de interés. El detector del marcador es una extensa área de detectores de posición (PSDs), análogo a una película fotográfica. Haciendo uso de las cámaras adecuadas, el PSD produce cuatro señales eléctricas derivadas de la imagen proyectada de un LED sobre el PSD. Los cuatro contactos del PSD, dos coordenadas en x y dos en y, registran las corrientes de salida. El efecto de la intensidad de la luz se invierte y se calcula el promedio de la posición bidimensional del punto de luz de la imagen en la superficie del fotodiodo. Con el fin de observar la posición de más de un LED o marcador activo, se utiliza multiplexación por división de tiempo. Esto significa que el sistema no puede hacer seguimiento de múltiples puntos simultáneamente. Esta situación se soluciona activando un LED a la vez. Por lo tanto, la identificación automática del marcador la realiza de por sí el sistema, debido a que cada cuadro contiene sólo un marcador y la orden de activación del mismo es conocida.

  En los sistemas de marcadores activos también se hace uso de diodos emisores de luz infrarroja (IREDs), los cuales emiten luz visible al sistema pero no al usuario. Como en los sistemas ultrasónicos, los detectores o cámaras pueden ser ubicados en el objeto a ser seguido o en puntos estacionarios alrededor del objeto. Claro está que poner cámaras sobre el cuerpo humano sería demasiado incómodo, por esta razón ubicar marcadores en el cuerpo y cámaras en puntos estacionarios del lugar es generalmente preferible. Para la detección del marcador, el enfoque frecuentemente utilizado es PSDs (fotodiodos de efecto lateral), los cuales están disponibles en cuadrículas de 1 cm^2. La luz incidente induce una corriente medida en cada borde de la cuadrícula para producir la posición $x - y$ del centroide de la luz incidente [18].

La ventaja de este tipo de sistema es la directa monitorización en línea de múltiples fuentes de luz, aspecto que disminuye la cantidad de memoria necesitada e incrementa la tasa de adquisición y procesamiento. Aunque el número de marcadores empleados tiene efecto sobre la frecuencia de muestreo del sistema, algunos sistemas permiten más de cincuenta marcadores con una frecuencia de muestreo por encima de 150 Hz. Usando menos marcadores, la frecuencia de muestreo se incrementa a más de 1KHz. La resolución espacial alta es además un beneficio de los sistemas de marcadores activos basados en la tecnología PSD. Sin embargo, utilizar LEDs como marcadores tiene diversas desventajas, por ejemplo, en algunas ocasiones se emplean cables para proporcionar potencia y fuente de sincronización a los LEDs, lo cual es incómodo para algunos pacientes [5].

- Marcadores pasivos. Los sistemas de marcadores pasivos tienen en común que utilizan marcadores reflectivos. Generalmente estos marcadores son esferas livianas cubiertas con cinta reflectiva. Esta cinta es diseñada para reflejar directamente la luz a lo largo de su línea de incidencia. Los sistemas de marcadores pasivos toman ventaja de esta propiedad llenando el volumen de captura de datos con luz generada por una fuente cercana a la lente de la cámara. En consecuencia, los pulsos de luz de una cámara en el volumen de captura, iluminan el marcador y lo reflejan de nuevo a la cámara. La luz pulsada es utilizada para detener el movimiento en el volumen de captura y proporcionar una ubicación instantánea del marcador. La frecuencia del pulso generalmente corresponde a la tasa de muestreo del sistema, cuyo rango oscila entre 50 y 250 Hz para sistemas de marcadores pasivos. Una sola cámara proporciona datos planos acerca de la ubicación del marcador. Al adicionar una segunda cámara y utilizando una transformación lineal directa se tiene la ubicación tridimensional del marcador a ser calculado. Los sistemas de marcadores pasivos utilizan como mínimo dos cámaras. Durante movimientos complejos el uso de más cantidad de cámaras aumenta la habilidad del sistema de seguir completamente un marcador a través de un recorrido de movimiento en el espacio tridimensional.

El software para diferentes sistemas de marcadores pasivos comprende rutinas de linealización y calibración. Esta cantidad de características internas y externas de la cámara incluyen distorsión de la lente y otras no linealidades introducidas por la operación del sistema. Aunque

algunos sistemas realizan en tiempo real el análisis de los datos, se requiere la intervención del usuario para asegurar la precisión del conjunto de datos.

- Sistemas sin marcadores. Sistemas no invasivos para la estimación de la cinemática del cuerpo humano que no requieren marcadores o accesorios ubicados en el cuerpo. Eliminar la necesidad de marcadores reduce considerablemente el tiempo de preparación del paciente, permite la simplicidad y evaluaciones más significativas del movimiento humano en investigación y práctica clínica [32].

  El enfoque principal en el desarrollo de sistemas de visión para captura de movimiento sin marcadores se basa en el uso de sistemas pasivos. La ventaja de los sistemas pasivos es que sólo se basan en la captura de imágenes y por ende proporcionan una estructura ideal para captura de sujetos en su ambiente natural.

3.2. Estimación de movimiento

La estimación de movimiento consiste en extraer el objeto de interés a partir de una secuencia video y posteriormente detectar y hacer seguimiento de puntos clave dentro de éste a lo largo de los cuadros del video. La figura 3.2 ilustra este proceso.

3.2.1. Detección del objeto

La detección del objeto en movimiento (segmentación) consiste en separar cada cuadro de la secuencia de video en regiones disyuntas. Formalmente la segmentación se define como:

Sea F el conjunto de todos los píxeles y $P()$ un predicado de uniformidad (homogeneidad) definido en grupos de píxeles conexos, la segmentación es una partición del conjunto F en un conjunto de regiones $(S_1, S_2, ..., S_n)$ tales que:

$$\bigcup_{i=1}^{n} S_i = F \text{ con } S_i \bigcap S_j = \oslash,\ i \neq j$$

El predicado de uniformidad $P(S_j) = verdadero$ para todas las regiones S_j y $P(S_i \bigcup S_j) = falso$, cuando S_i es adyacente a S_j.

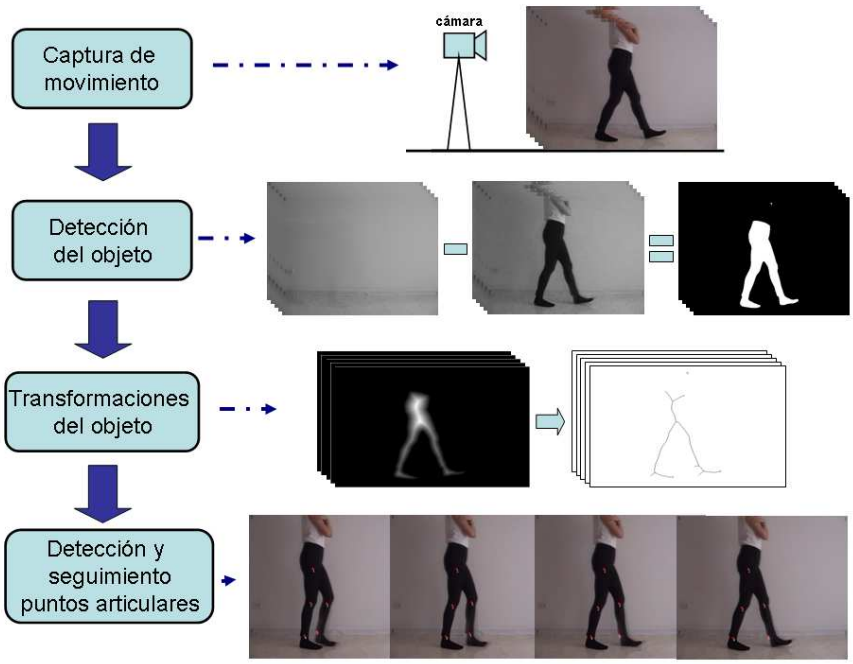


Figura 3.2. Etapas de la estimación de movimiento.

La finalidad de la segmentación es identificar los componentes semánticamente significativos de la secuencia de video y agrupar los píxeles que pertenecen a dichos componentes [70]. Algunos métodos para alcanzar tal objetivo incluyen sustracción de cuadros consecutivos, máscara de detección de cambio (CDM), sustracción de fondo y estimación de fondo, entre otros [27, 43, 59, 65].

Sustracción de cuadros consecutivos

La sustracción de cuadros consecutivos es una sencilla operación que calcula la diferencia de las características de un mismo pixel entre cuadros sucesivos de la escena en una secuencia de imágenes.

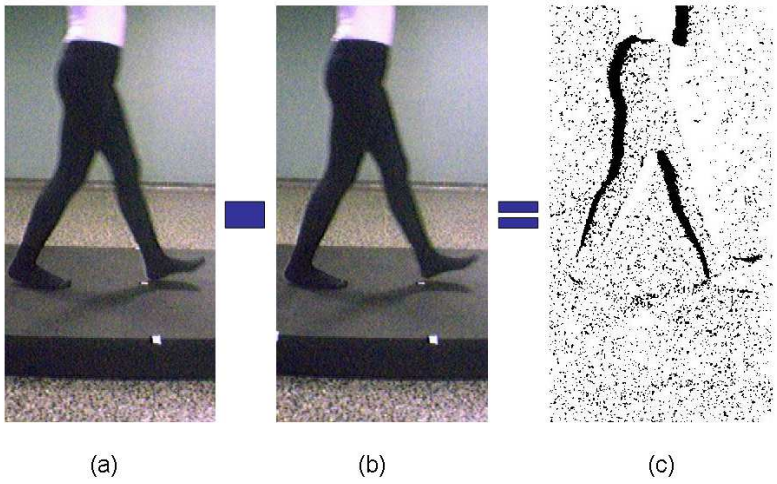


Figura 3.3. Diferencia de cuadros consecutivos. (a)$imagen_i$, (b)$imagen_{i-1}$, (c)resultado de la diferencia de cuadros consecutivos.

Este método se basa en la ecuación:

$$|imagen_i - imagen_{i-1}| < umbral \tag{3.1}$$

donde $imagen_i$ es el cuadro actual e $imagen_{i-1}$ es el cuadro anterior de la secuencia de imágenes. Este algoritmo omite objetos de primer plano con movimientos casi o totalmente nulos y cambios de iluminación sutiles,

además que identifica como objetos de primer plano a objetos con movimiento oscilatorio que forman parte del fondo, es vulnerable a cambios repentinos de luz y sólo puede recuperar de forma parcial el borde de los objetos en movimiento, (ver figura 3.3).

Máscara de detección de cambio

El método conocido como máscara de detección de cambio (CDM) utiliza una estimación estadística y trabaja con el grado de cambios de la reja espacial. Este método garantiza resultados de detección robusta en presencia de ruido, sin embargo el valor de diferencia entre cuadros depende de la velocidad del objeto en movimiento, por consiguiente la calidad de la segmentación no puede ser sistemáticamente conservada si la velocidad del objeto cambia significativamente en la secuencia (ver figura 3.4) y adicionalmente la sombra del objeto en la región del fondo puede generar problemas en la aproximación basada en detección de cambio [2, 52].

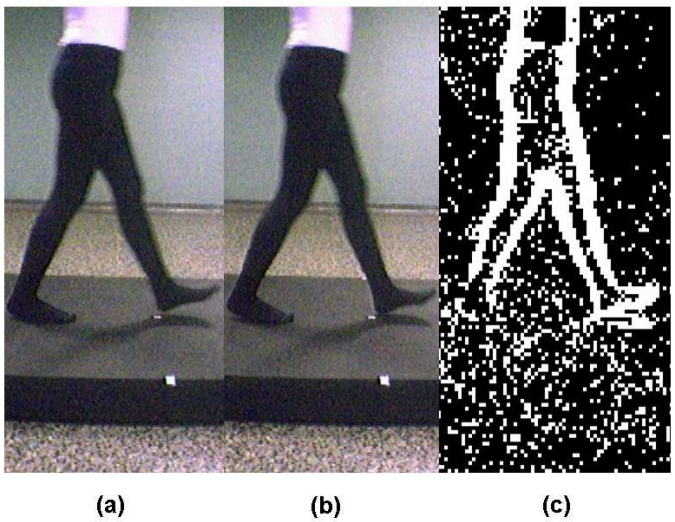


Figura 3.4. Máscara de detección de cambio.a)$imagen_{i-1}$, (b)$imagen_i$, (c)resultado de máscara de detección de cambio. Tiempo de procesamiento= 582,36 s

Sustracción de fondo

Identificar objetos en movimiento desde una secuencia de video capturada con una cámara fija es una tarea fundamental y crítica en muchas aplicaciones de visión por computador. Una manera frecuente de llevar a cabo este proceso es realizando sustraccción de fondo, técnica efectiva para extraer objetos móviles de interés en escenas estáticas. Un algoritmo tradicional de sustracción de fondo asume un objeto (u objetos) en movimiento con respecto a un fondo estático y segmenta el objeto (u objetos) en movimiento clasificando los píxeles como primer plano o como fondo de acuerdo a modelos estadísticos entrenados (ver figura 3.5). La imagen del fondo es básicamente una representación de la escena sin objetos en movimiento y es necesario que sea actualizada regularmente con el fin de adaptar las condiciones de variación en la iluminación y ajustes en la geometría [15, 59, 65, 68].

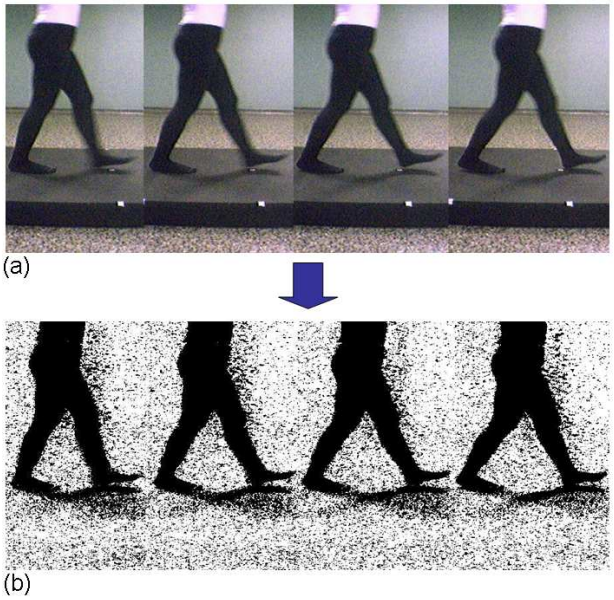


Figura 3.5. Sustracción de fondo utilizando filtro exponencial recursivo clásico $M_t = \alpha I_t + (1 - \alpha) M_{t-1}$. Número de cuadros= 77. Valor de $\alpha = 0$. Umbral= 0,1. Tiempo de procesamiento= 251,21 s. (a)Secuencia original. (b)Resultado de la sustracción de fondo.

El desempeño de la sustracción de fondo depende principalmente de la técnica usada para modelar el fondo de la escena. Un algoritmo robusto para

modelar el fondo debería manejar situaciones en las cuales nuevos objetos sean introducidos, o aquellos que ya estaban sean removidos de la escena. Incluso en una escena estática los cambios pueden ocurrir cuadro a cuadro debido al ruido y a oscilaciones de la cámara [71]. Algunas técnicas de sustracción de fondo se basan en modelos que utilizan las características de textura [15], otras se apoyan en uno o múltiples modelos de probabilidad, como la Gaussiana o mezcla de Gaussianas [27].

Estimación de fondo estático *Sigma-Delta* **($\Sigma - \Delta$)**

La estimación de fondo $\Sigma - \Delta$ es un método sencillo y eficiente para detectar aquellos píxeles que han cambiado significativamente en la escena estática, relacionados a una constante de tiempo que depende del número de píxeles en nivel de gris y de la velocidad de los cuadros del video.

Modulación $\Sigma - \Delta$ En la modulación $\Sigma - \Delta$ se sobremuestrea una señal de mensaje entrante a una tasa mayor que la de Nyquist [3] para aumentar la correlación entre muestras adyacentes de la señal, permitiendo de esta manera construir la señal codificada con cuantificación simple [23]. Los conversores $\Sigma - \Delta$ tienen aplicaciones en sistemas de comunicaciones, audio, básculas industriales y dispositivos de medidas de precisión. Esta popular técnica se utiliza para fabricar conversores análogo-digital y digital-análogo de alta resolución [10]. La característica fundamental de estos conversores es que son el único método de conversión de bajo costo que suministran rango dinámico alto y flexibilidad en conversión de señales de entrada con bajo ancho de banda [25].

La figura 3.6 muestra el diagrama de bloques de un sistema de modulación $\Sigma - \Delta$. La entrada es la diferencia entre la señal análoga y la de retroalimentación (llamada también señal de referencia). Esta diferencia pasa por la etapa Σ (*Sigma*) del conversor en la cual se realiza la integración. El modulador Δ (*Delta*) es un comparador seguido por un multiplicador, el cual se alimenta a partir de un generador de pulsos externos (reloj) para producir una señal codificada de 1 bit. El uso de integración a la entrada del transmisor requiere de una diferenciación en el receptor. Esta necesidad se elimina debido a que la diferenciación se cancela por la integración en el

[3] Según el teorema de muestreo de *Nyquist-Shannon*, para poder replicar con exactitud la forma de una onda es necesario que la frecuencia de muestreo sea superior al doble de la máxima frecuencia a muestrear.

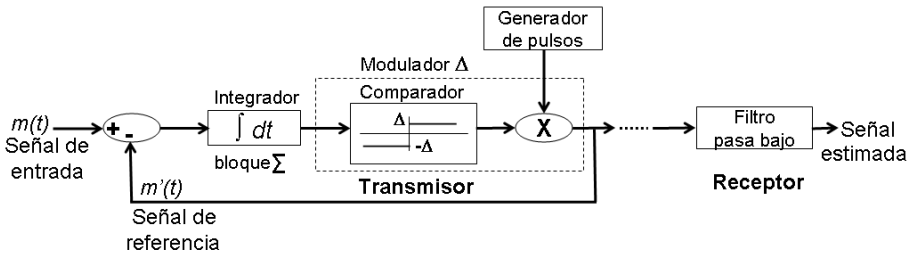


Figura 3.6. Sistema de modulación $\Sigma - \Delta$.

receptor Δ convencional, de modo que el receptor de un sistema de modulación $\Sigma - \Delta$ consiste en un filtro pasa bajo. Si la muestra de entrada $m(t)$ es mayor que la aproximación [4] más reciente, se aplica un incremento positivo $+\Delta$ a la aproximación. Por el contrario, si la muestra de entrada es más pequeña, se aplica un incremento negativo $-\Delta$ a la aproximación [23]. En síntesis, el modulador $\Sigma - \Delta$ es un seguidor cuyo objetivo es hacer que la señal estimada sea lo más cercana posible a la señal de entrada, evitando que la señal cambie rápidamente de una muestra a la siguiente. En el análisis de imágenes, este operador se utiliza para estimar el fondo, de forma tal que su varianza sea menor a la del fondo original.

Estimación $\Sigma - \Delta$. La estimación de fondo $\Sigma - \Delta$, es un método en el cual se estiman los cambios de cada pixel de la imagen mediante operaciones sencillas como comparaciones, incrementos y decrementos [36–38]. A pesar de su simplicidad computacional, esta técnica tiene robustas propiedades debido a que está estadísticamente fundamentada con respecto a la estimación de distribuciones monomodales [37]. Este filtro tiene interesantes propiedades relacionadas con la detección de cambio de señales que varían con el tiempo. Una interpretación de esta estimación de fondo es considerarla como la simulación de una conversión digital de señales análogas que varían con el tiempo usando modulación $\Sigma - \Delta$ [38]. El filtro $\Sigma - \Delta$ trabaja mejor sobre aplicaciones cuyo tiempo derivativo absoluto es menor que la unidad (el cambio de la señal es mínimo), lo cual implica que el error de modulación es proporcional a la tasa de variación de la señal, correspondiente a la medida de verosimilitud del movimiento del pixel.

[4] Aproximación de escalera a la versión sobremuestreada de la señal de mensaje.

El filtro $\Sigma - \Delta$ es empleado además para calcular la varianza temporal del pixel, de modo que sea posible decidir si los píxeles cambian hacia los estados "en movimiento" o "estacionario". El interés se centra en aquellos píxeles cuya tasa de variación es mayor que su actividad temporal. El enfoque de este método es la estimación recursiva de distribuciones monomodales, asumiendo que las series de tiempo correspondientes a los diferentes valores posibles tomados por el fondo a lo largo del tiempo presentan un único valor con la frecuencia más grande en el proceso [36].

Sabiendo que I_t es la secuencia de entrada de este método y M_t es el valor estimado del fondo, si I_t es una señal aleatoria discreta, la proporción entre el número de índices $\tau < t$ tal que $I_\tau < M_t$ y el número de índices $\tau < t$ tal que $I_\tau > M_t$, converge en una media de 1. Por lo tanto M_t es una aproximación de la mediana de I_t [39].

La no linealidad y la propiedad de convergencia de la media de la estimación $\Sigma - \Delta$ se dan por el hecho de que la estadística M_t es actualizada con un incremento constante ± 1 [38], independiente de la diferencia entre el cuadro y la media actual $(I_t - M_{t-1})$. Si el incremento dependiera linealmente sobre la diferencia se obtendría $M_t = M_{t-1} + \alpha(I_t - M_{t-1}) = \alpha I_t + (1-\alpha)M_{t-1}$, que es un filtro exponencial recursivo clásico (o promedio de movimiento) usado normalmente en métodos recursivos de ajuste Gaussiano [8, 48, 54]. En su expresión más sencilla α es una constante en el intervalo [0,1]. En la expresión más robusta de la estimación de ajuste Gaussiano, α ya no es una constante y se calcula a partir de la probabilidad del cuadro actual $\alpha(t) = N_{V_t}^{M_t}(I_t)$, donde $N_{\sigma^2}^{\mu}(x)$ es la función densidad normal de la media μ y la varianza σ^2. La varianza V_t es estimada como el movimiento promedio de las diferencias cuadradas $(I_t - M_t)^2$. El cálculo del cuadrado y la probabilidad Gaussiana del cuadro actual hacen que estos métodos sean sensibles a aproximaciones numéricas, mientras que la estimación $\Sigma - \Delta$ sólo utiliza cálculos exactos con enteros. La ventaja de los métodos de ajuste Gaussiano es su extensión a la estimación multimodal de distribuciones Gaussianas [48, 54], a través de las cuales es posible trabajar con escenas complejas. El único parámetro visible del método $\Sigma - \Delta$ es N, el número usado en el cálculo de la varianza V_t. El rango del valor de N es pequeño (entre 1 y 4) y generalmente se escoge una potencia de 2 para propósitos de optimización.

3.2.2. Seguimiento de puntos clave en un objeto

El seguimiento se define como el problema de estimar la trayectoria e información del centro de un objeto (de acuerdo a su orientación, área o forma) en el plano de la imagen a medida que se mueve alrededor de la escena. El seguimiento de objetos puede hacerse complejo debido a la pérdida de información causada por la proyección del mundo 3D en una imagen 2D, ruido en las imágenes, movimientos complejos, naturaleza no rígida o articulada, oclusiones parciales o completas y formas complejas de los objetos, cambios en la iluminación de la escena y exigencia de procesamiento en tiempo real.

El seguimiento de objetos es una tarea importante en el campo de visión por computador, especialmente en el análisis automático de video. Sus aplicaciones incluyen biomecánica, biometría, vigilancia, indexación de video, interacción hombre-máquinas, supervisión de tráfico y navegación de vehículos, entre otros.

El seguimiento se puede facilitar poniendo unas restricciones al movimiento. Por ejemplo, casi todos los algoritmos de seguimiento asumen que el movimiento del objeto es suave, sin cambios abruptos. El objeto en movimiento se puede limitar a tener una velocidad o aceleración constantes, basándose en información *a priori*. De otra parte, el conocimiento *a priori* sobre el tamaño, apariencia y forma del objeto puede ser utilizado también para simplificar el problema [1]. La tarea del seguimiento se realiza teniendo en cuenta la representación del objeto, selección de características, detección y seguimiento.

Las tareas de detectar el objeto y establecer correspondencia entre la representación del objeto, a través de los cuadros, se pueden realizar de forma separada o conjunta. En el primer caso, las regiones de objetos posibles en cada cuadro se obtienen por medio de un algoritmo de detección y después se hace rastreo de los objetos a través de los cuadros. En el otro caso, la detección del objeto y la correspondencia son estimadas conjuntamente actualizando iterativamente la ubicación del objeto y la información de la región obtenida en cuadros anteriores. En cualquier caso, los objetos se representan utilizando modelos de puntos, núcleo o silueta [1]. En el seguimiento de puntos, los objetos detectados en cuadros consecutivos son representados por puntos y su asociación se basa en el estado anterior del objeto, el cual puede incluir su movimiento y su posición. Esta aproximación requiere de un mecanismo externo como el filtro *Kalman* para detectar los

objetos en cada cuadro.

Filtro *Kalman***.**

El filtro *Kalman* es un conjunto de ecuaciones matemáticas que proporcionan un método computacional eficiente para la estimación de los estados de un sistema lineal minimizando el error cuadrático medio. Este filtro tiene la ventaja de realizar estimaciones de estados anteriores, actuales y futuros, incluso en situaciones en las que se desconoce la naturaleza precisa del sistema modelado [62].

El filtro *Kalman* se enfoca en el problema de tratar de estimar el estado $x \in \Re^n$ de un proceso controlado de tiempo discreto regido por la ecuación lineal estocástica de diferencias:

$$x_k = Ax_{k-1} + Bu_{k-1} + w_{k-1} \tag{3.2}$$

con una medición $z \in \Re^m$:

$$z_k = Hx_k + v_k \tag{3.3}$$

El ruido del proceso y de la medición está representado por las variables aleatorias w_k y v_k repectivamente, las cuales se asumen independientes entre sí y con distribución normal de probabilidad:

$$p(w) \sim N(0, Q) \tag{3.4}$$

$$p(v) \sim N(0, R) \tag{3.5}$$

En aplicaciones prácticas, las matrices de covarianza de ruido del proceso Q y de la medición R pueden cambiar con cada paso del tiempo o de la medición, sin embargo en este caso se asumen constantes.

La matriz A de transición de estados, de tamaño n x n (ecuación 3.2) relaciona el estado en un tiempo anterior $k-1$ con el estado en un tiempo k, en ausencia de alguna acción o ruido del proceso. En la práctica A podría cambiar con cada paso de tiempo, sin embargo en esta situación se asume constante:

$$A = \begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

La matriz B de n x l asocia la entrada de control opcional $u \in \Re^l$ al estado x:

$$B = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

La matriz de mediciones H, de tamaño m x n (ecuación de medición 3.3) relaciona el estado a la medición z_k. Esta matriz también se asume constante aunque en la práctica podría cambiar con cada paso del tiempo o de la medición:

$$H = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}$$

Algoritmo del filtro *Kalman* discreto El filtro *Kalman* estima un proceso utilizando una forma de control realimentado, en el cual hace la estimación del estado del proceso en un tiempo cualquiera y posteriormente obtiene una realimentación por medio de datos observados. De acuerdo a este criterio, las ecuaciones del filtro *Kalman* se clasifican como ecuaciones de actualización temporal (predicción) y de actualización de la medición (estimación). Las ecuaciones de actualización temporal se encargan de hacer una proyección futura del estado actual (k) y del error de covarianza estimados, con base al estado en el momento $k-1$. Las ecuaciones de actualización de la medición son responsables de la etapa de realimentación, incorporando una nueva medición en la estimación anterior con el objeto de obtener una estimación mejorada en el momento k. El filtro *Kalman* trabaja mediante un mecanismo de predicción-corrección, al suponer el nuevo estado y su incertidumbre y corregir la predicción con la nueva medida. Este proceso se muestra en la figura 3.7.

Las ecuaciones específicas en las que se basa la actualización temporal son:

$$\hat{x}_{\bar{k}} = A\hat{x}_{k-1} + Bu_{k-1} \tag{3.6}$$

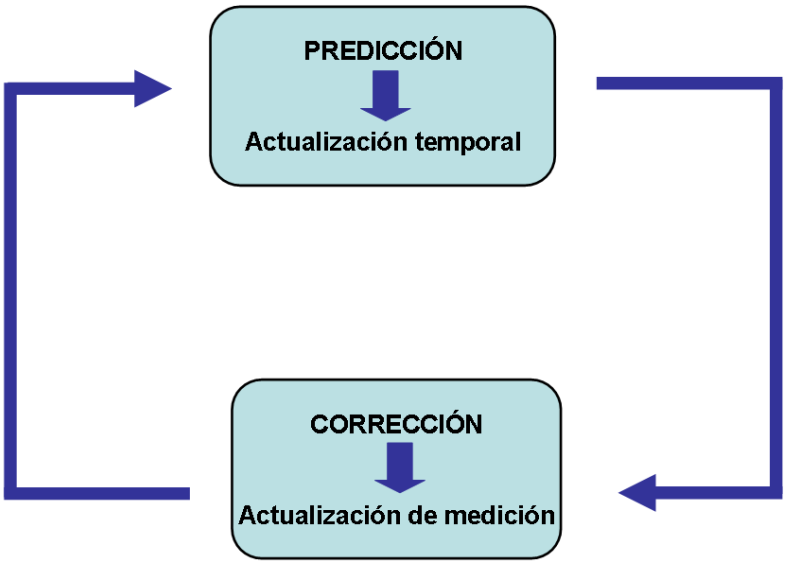


Figura 3.7. Diagrama cíclico del filtro *Kalman*. La actualización temporal proyecta el estado actual estimado hacia adelante en el tiempo. La actualización de medición ajusta la estimación proyectada mediante una medición actual en ese tiempo.

$$P_{\bar{k}} = AP_{k-1}A^T + Q \tag{3.7}$$

Las ecuaciones de actualización temporal 3.6 y 3.7 proyectan el estado y covarianza estimadas de forma adelantada en un tiempo $k-1$ a un tiempo k. A y B se conocen de la ecuación 3.2 y Q se obtiene de la ecuación 3.4.

Las ecuaciones mediante las cuales se obtiene la actualización de las mediciones son:

$$K_k = P_{\bar{k}}H^T \left(HP_{\bar{k}}H^T + R\right)^{-1} \tag{3.8}$$

$$\hat{x}_k = \hat{x}_{\bar{k}} + K_k \left(z_k - H\hat{x}_{\bar{k}}\right) \tag{3.9}$$

$$P_k = \left(I - K_k H\right) P_{\bar{k}} \tag{3.10}$$

La tarea primordial durante la actualización de las mediciones consiste en calcular la ganancia del filtro *Kalman*, K_k. Posteriormente es necesario actualizar la medición del proceso para obtener z_k y después generar el estado estimado *a posteriori*, incorporando la medición como se muestra en la ecuación 3.9. Finalmente se calcula el error de covarianza estimado *a posteriori*, mediante la ecuación 3.10.

Después de aplicar la actualización temporal y de la medición, el proceso se repite tomando como referencia las nuevas estimaciones del estado y de la covarianza del error.

Capítulo 4

Extracción y Seguimiento de los Miembros Inferiores

El principal problema tratado en esta investigación es la extracción de la silueta de los miembros inferiores a partir de diferentes secuencias de video de marcha, de manera que se pueda obtener información del movimiento sin utilizar marcadores. En la literatura existen diferentes investigaciones encaminadas al análisis de marcha sin marcadores, en las cuales utilizan generalmente modelos del cuerpo humano en 2D ó 3D para obtener información cinemática relevante (ver capítulo 2). En el presente trabajo se obtuvo la silueta de los miembros inferiores mediante el método de estimación de fondo estático Σ - Δ explicado en la sección 3.2.1 y posteriormente se realizó la detección y seguimiento de los centros articulares articulares mediante el uso de transformaciones de distancia, esqueletización (sección 4.3) y filtro *Kalman* (sección 3.2.2).

La figura 4.1 resume los diferentes componentes del método utilizado. La primera parte (sección 4.2) es la segmentación del objeto en movimiento, cuyo propósito es obtener la silueta de los miembros inferiores. En segundo lugar (secciones 4.3 y 4.4) está el parámetro de detección, el cual consiste en hacer corresponder la observación (silueta) con un conjunto de puntos representativos de las articulaciones de cadera (coxofemoral), rodillas (tibiofemoral) y tobillos (tibiotarsiana), para obtener la información del movimiento de los miembros inferiores a lo largo de la secuencia de video.

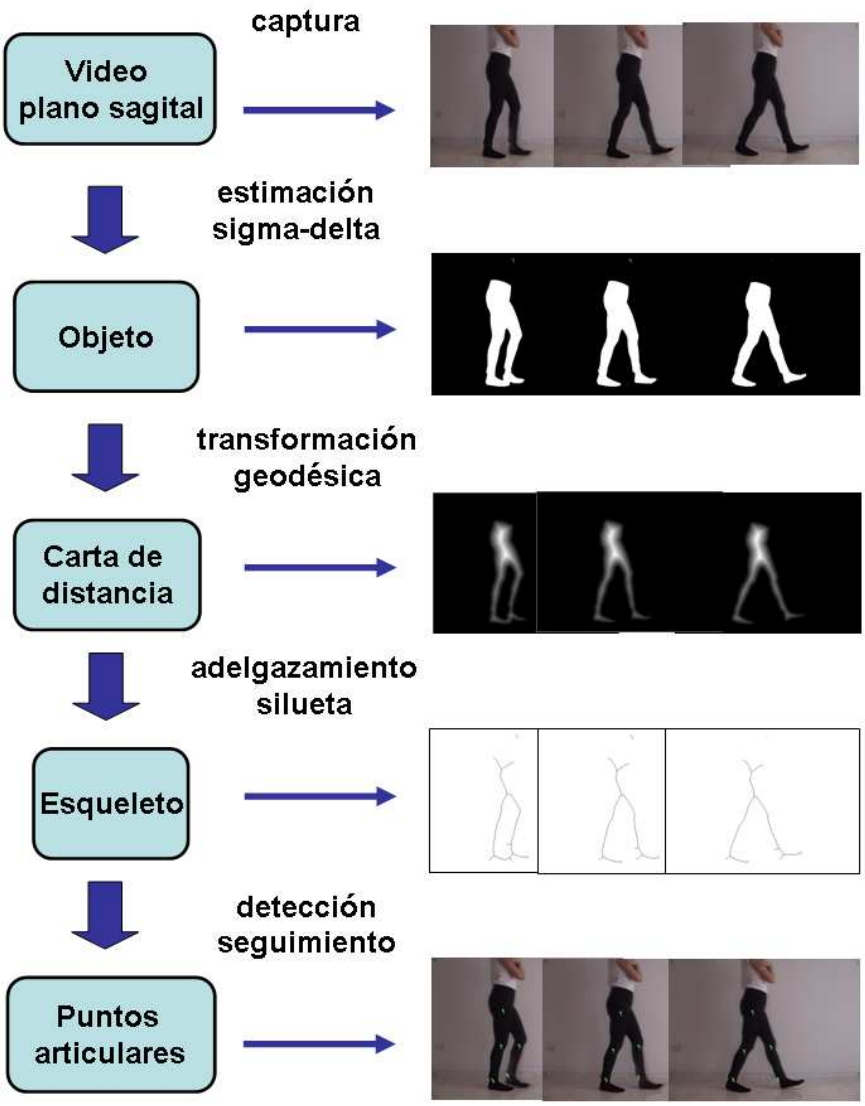


Figura 4.1. Visión general de la metodología aplicada.

4.1. Captura de video

El proceso inicial consistió en realizar la captura del movimiento. Para tal fin se grabaron seis videos de 49 ciclos de marcha; actividad que se llevó a cabo en el laboratorio de análisis de movimiento de la Universidad Nacional de Colombia sede Bogotá. Los videos tienen resolución de 640×480 pixeles, a 30 cuadros por segundo, bajo condiciones de iluminación semicontrolada con una cámara fija *firewire Unibrain Fire-i* previamente calibrada. Adicionalmente se realizaron 22 videos con ciclos de marcha completos de 12 inviduos sanos, seis hombres entre los 9 y los 73 años y, seis mujeres entre los 21 y los 69 años. Doce de los videos fueron realizados bajo condiciones de iluminación semicontrolada y los diez restantes con iluminación natural. Estos 22 videos fueron grabados con una cámara *Sony Handycam* DCR-HC38, a 30 cuadros por segundo, con una resolución de 720×480 pixeles. Todos los videos fueron capturados en el plano sagital, debido a que es en éste donde se aprecian la totalidad de los movimientos articulares más relevantes. Adicionalmente, en este plano es en el cual los especialistas centran su mayor interés en el análisis de marcha [30, 33, 60].

4.2. Segmentación

El siguiente paso después de las capturas consistió en obtener la silueta de los miembros inferiores. Debido a que la adquisición del video se realizó con una cámara fija, la segmentación fue manejada como un problema de sustracción de fondo. Inicialmente se exploraron algunos métodos como diferencia de cuadros consecutivos, máscara de detección de cambio y media adaptativa, cuyos resultados se observan en la sección 3.2. Dado que las segmentaciones obtenidas con las técnicas mencionadas anteriormente no fueron satisfactorias, se implementó el método de estimación de fondo estático Σ-Δ (presentado en la sección 3.2.1), donde dada una secuencia de imágenes I_t, se extrae una imagen binaria S_t, en la cual los píxeles x tales que $S(x) = 0$ forman la escena estática o fondo, mientras que los píxeles x tales que $S(x) = 1$ constituyen los objetos en movimiento o primer plano.

La estimación Σ-Δ sólo proporciona una clasificación de primer plano a nivel de pixel y no es suficiente para obtener las siluetas deseadas, razón por la cual esta puede combinarse con regularización espacio-temporal. En nuestro sistema, esto se hizo aplicando un filtro morfológico, seguido por una reconstrucción geodésica [36]. Como resultado se obtuvieron secuencias

suficientemente largas de los objetos binarios de interés.

4.3. Transformación geodésica y esqueletización

Una vez la silueta de los miembros inferiores fue extraída utilizando la estimación de fondo estático Σ-Δ, se determinaron las relaciones articulares fundamentales, realizando una detección de los puntos representativos de las articulaciones de cadera, rodillas y tobillos. Para hacer esto, una transformación geodésica euclidiana se aplicó a la silueta segmentada $D(p) = min\{dist_M(p,q), q \in O\}$ [11], generando un mapa de distancias cuyo objeto de interés es la silueta en sí, por ejemplo la silueta es el conjunto de distancia cero y la información de distancia es propagada hacia el centro de la misma (ver apéndice B).

Sean S_t la imagen binaria y d la distancia sobre el plano discreto $\mathbb{Z}^2$. La transformada de distancia de S_t se define como:

$$\begin{array}{rccl} F^d_{S_t}: & \mathbb{Z}^2 & \longrightarrow & \mathbb{N} \\ & x & \mapsto & \min\limits_y\{d(x,y); y \in \mathbb{Z}^2, S_t(y) = 0\} \end{array}$$

El mapa de distancias obtenido a partir de la transformación geodésica se redujo a un grafo con el fin de obtener el esqueleto de los miembros inferiores. Este proceso se realizó aplicando un algoritmo de adelgazamiento (ver apéndice C), donde dada la imagen obtenida de la transformación geodésica D_t y un elemento estructural compuesto $B = (B_1, B_2)$, el esqueleto se calcula a partir de [63]:

$$D_t \triangleright B = D_t - (D_t * B)$$

La esqueletización encuentra los ejes medios de la silueta y aunque en algunos cuadros del video estos ejes no corresponden a los centros articulares, permite tener una estimación de estos puntos donde se toman las medidas cinemáticas del movimiento.

4.4. Detección y seguimiento de centros articulares

Después de obtenerse el esqueleto E_t de la silueta, se hizo la detección de cinco puntos correspondientes a la cadera (la que se aprecia en el video), las rodillas y los tobillos. Este proceso consistió en realizar una marcación

inicial sobre la silueta binaria de los puntos correspondientes a dichas articulaciones, con base en los puntos de referencia antropométrica (trocánter mayor, cóndilos femorales y maléolos) [33, 49]. Posteriormente, estos puntos se ajustaron al esqueleto calculado de forma tal que se ubicaran sobre los ejes longitudinales de éste, los cuales tienen un valor máximo de intensidad (1). A cada máximo del esqueleto le corresponde una posición (x, y), indicando que se tiene un conjunto de puntos sobre los que se realizó el ajuste de los puntos anteriormente marcados. El proceso de ajuste consistió en determinar la distancia euclidiana entre cada punto marcado y cada una de las posiciones del esqueleto. Aquél punto del esqueleto cuya distancia con respecto al punto marcado fuera mínima, se convertía entonces en el punto ajustado.

Sea $X = \{x_c, x_r^1, x_r^2, x_t^1, x_t^2\}$ el conjunto de posiciones en $\mathbb{Z}^2$ de las cinco articulaciones del esqueleto. Sea μ_t^X el conjunto de las cuatro distancias calculadas para ubicar los puntos a la imagen de acuerdo a X. La función de ajuste de los puntos con respecto a E_t se define como:

$$\Phi(\mu_t^X, S_t) = \sum_{z \in \mu_t^X} F_{E_t}^d(z)$$

Una vez obtenidos los cinco puntos, fueron corregidos de manera que se conservaran en cierto margen las distancias antropométricas entre los segmentos cadera-rodillas y rodillas-tobillos. Este proceso se denota como el siguiente problema de optimización:

$$X_{puntos}^t = \arg \max_{X \in \mathbb{Z}^2} \Phi(\mu_t^X, E_t)$$

Los puntos articulares detectados corresponden a la medición, utilizada por el filtro *Kalman* para determinar la ubicación de las articulaciones en un tiempo k con base en la información que se tiene en un tiempo $k-1$. El filtro obtiene una predicción y una estimación, necesarias para hacer una proyección futura del estado actual e introducir una nueva medición en la estimación anterior, con el fin de obtener una estimación mejorada en el momento actual. En este problema en particular, la estimación realizada por el filtro *Kalman* fue muy valiosa en los casos en que una pierna ocluía la otra. En esta situación el esqueleto no corresponde al eje longitudinal de cada miembro y los puntos detectados no se aproximan mucho a la verdadera ubicación de las rodillas y los tobillos, por lo tanto la estimación se acerca un poco más a la posición de rodillas y tobillos.

4.5. Resultados generales y análisis

La estimación $\Sigma - \Delta$ se llevó a cabo tomando una escena de fondo inicial, con base en la cual se segmentó la silueta de los miembros inferiores. Algunos experimentos se realizaron en el espacio de color YUV sobre la componente de luminancia. El primer paso fue estimar el fondo estático M_t comparando el cuadro de entrada I_t con un fondo de referencia M_{t-1} e incrementando o decrementando un valor correspondiente a 1. Después se calculó el valor absoluto de la diferencia entre el cuadro de entrada I_t y el fondo estimado M_t. La estimación $\Sigma - \Delta$ se implementó en más de 30 cuadros para obtener resultados robustos.

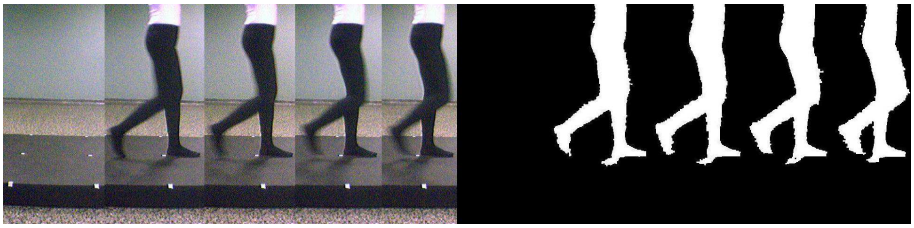


Figura 4.2. (izq.) Secuencia de marcha en laboratorio de marcha, (der.) resultados de la estimación de fondo $\Sigma - \Delta$.

Los resultados obtenidos con esta estimación proporcionaron una silueta de los miembros inferiores bien definida. Se presentaron inconvenientes relacionados con sombras proyectadas debido a que algunas capturas se realizaron bajo condiciones de iluminación semicontrolada. Las figuras 4.2(izq.), 4.3(a)(c), 4.4(a) y 4.5(a) muestran diferentes secuencias de video capturadas en ambientes con iluminación no uniforme. En las figuras 4.6(a) y 4.7(a) se observan dos secuencias con iluminación natural. Los resultados de la estimación $\Sigma - \Delta$ se muestran en las figuras 4.2(der.), 4.3(b)(d), 4.4(b), 4.5(b), 4.6(b) y 4.7(b).

Las secuencias binarias obtenidas después de la estimación $\Sigma - \Delta$ se utilizaron para calcular una carta de distancias cuyo principal objetivo fue proporcionar un pseudo-esqueleto a partir de estas imágenes. Este mapa de distancias se obtuvo utilizando una métrica euclidiana, cuyos resultados se observan en las figuras 4.4(c), 4.5(c), 4.6(c) y 4.7(c). Posteriormente se llevó a cabo un proceso de adelgazamiento (esqueletización) con el cual se obtuvo la estructura del objeto. En las figuras 4.4(d), 4.5(d), 4.6(d) y 4.7(d) se muestra el esqueleto de los miembros inferiores.

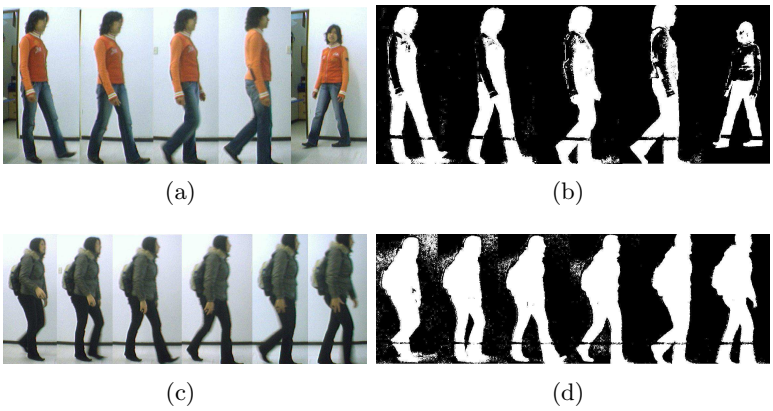


Figura 4.3. Diferentes secuencias de marcha. (a)Secuencia de caminata capturada en un ambiente con iluminación artificial, (b)resultado de la estimación de fondo $\Sigma - \Delta$, 769 cuadros. (c)Secuencia de caminata usando morral, capturada en un ambiente con iluminación artificial, (d)resultado de la estimación de fondo $\Sigma - \Delta$, 197 cuadros.

El cálculo del esqueleto de la silueta de los miembros inferiores fue la base para el proceso de detección de los cinco centros articulares de cadera, rodillas y tobillos. Estos puntos se encontraron ajustando sobre el esqueleto unas posiciones iniciales, las cuales fueron corregidas durante la secuencia de marcha con ayuda de la predicción de cada cuadro anterior de video realizada por el filtro *Kalman*. En las figuras 4.4(e), 4.5(e), 4.6(e) y 4.7(e) se observan diferentes secuencias de un ciclo de marcha con la medición de estos puntos articulares. En las figuras 4.4(f), 4.5(f), 4.6(f) y 4.7(f) se muestra el resultado de la predicción de las posiciones articulares.

El principal inconveniente encontrado en la etapa de seguimiento fue la oclusión de los puntos detectados en rodillas y tobillos en las fases de balanceo del ciclo de marcha (ver figura 3.1). Este problema se solucionó separando las posiciones de los puntos articulares al momento de la oclusión, de acuerdo al criterio de adelantar cierta cantidad de posiciones el punto de rodilla o tobillo del miembro que estuviera por delante de la cadera y atrasar aquellas posiciones de los puntos ocluidos que se encontraran por detrás de la cadera.

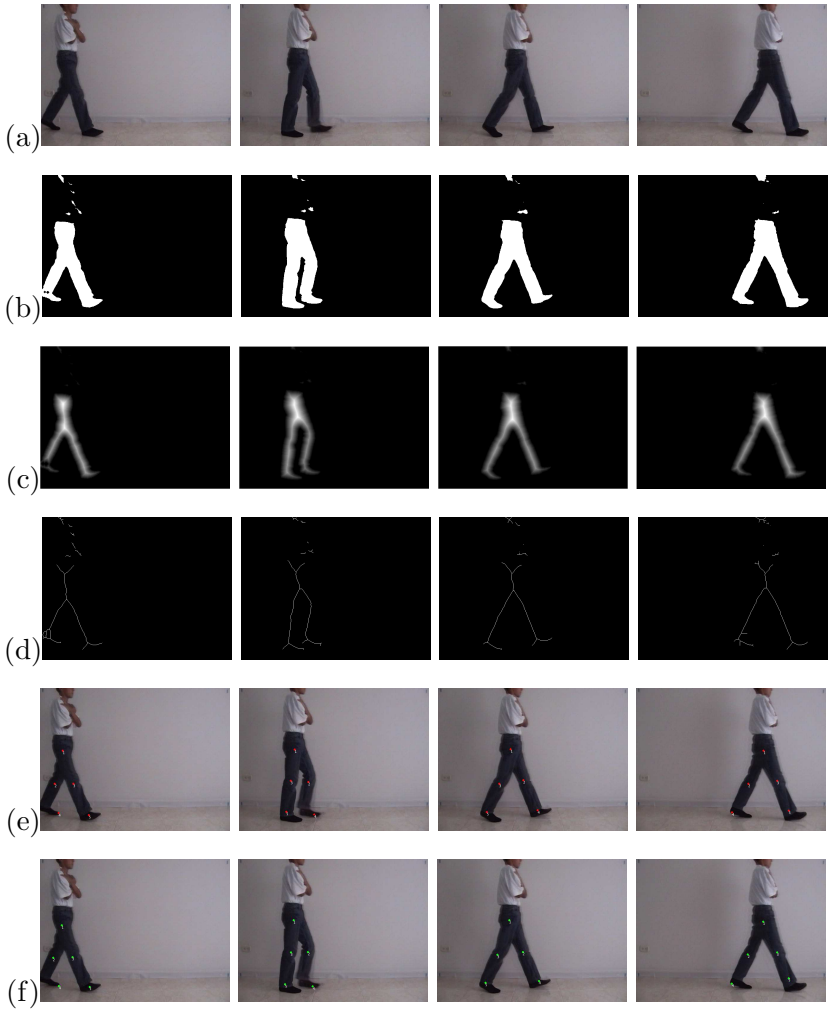


Figura 4.4. Proceso y resultados de la aplicación en un niño de 9 años. Ambiente con iluminación semicontrolada, 32 cuadros de video. Las imágenes muestran los cuadros 132, 143, 148 y 163 de la secuencia, así: (a)video original, (b)segmentaciones, (c)carta de distancias, (d)esqueleto, (e)mediciones y (f)estimaciones.

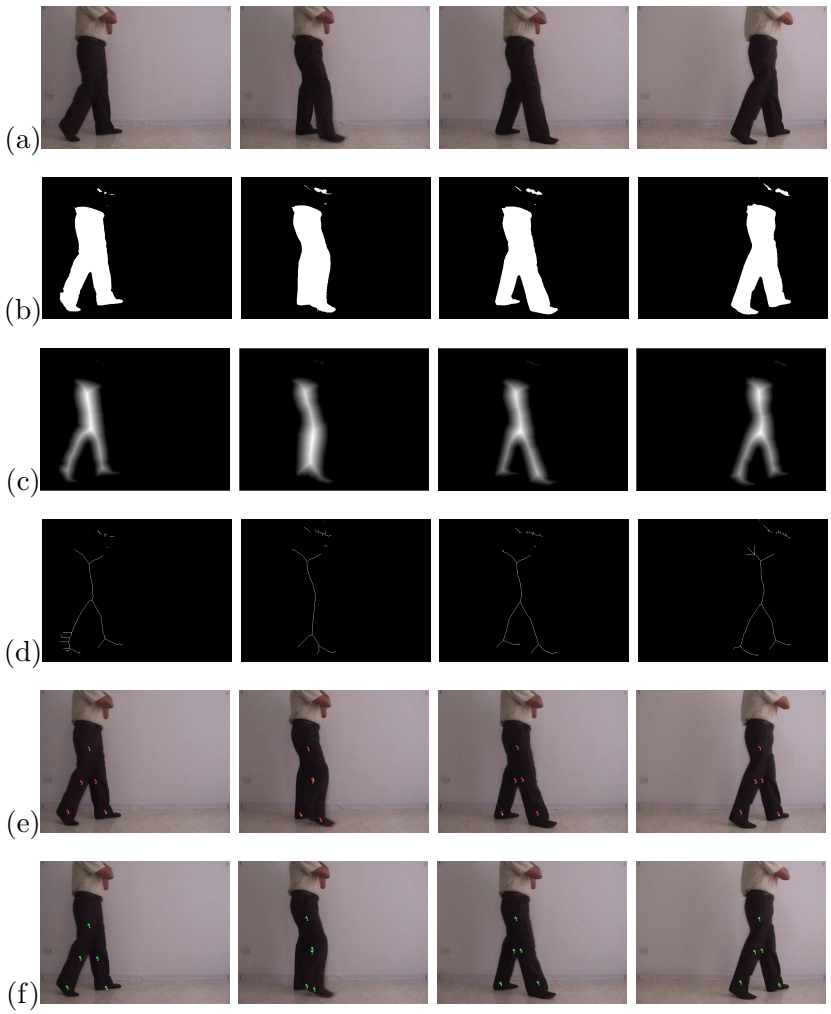


Figura 4.5. Proceso y resultados de la aplicación en hombre adulto de 73 años. Ambiente con iluminación semicontrolada, 39 cuadros de video. Las imágenes muestran los cuadros 183, 192, 197 y 219 de la secuencia, así: (a)video original, (b)segmentaciones, (c)carta de distancias, (d)esqueleto, (e)mediciones y (f)estimaciones.

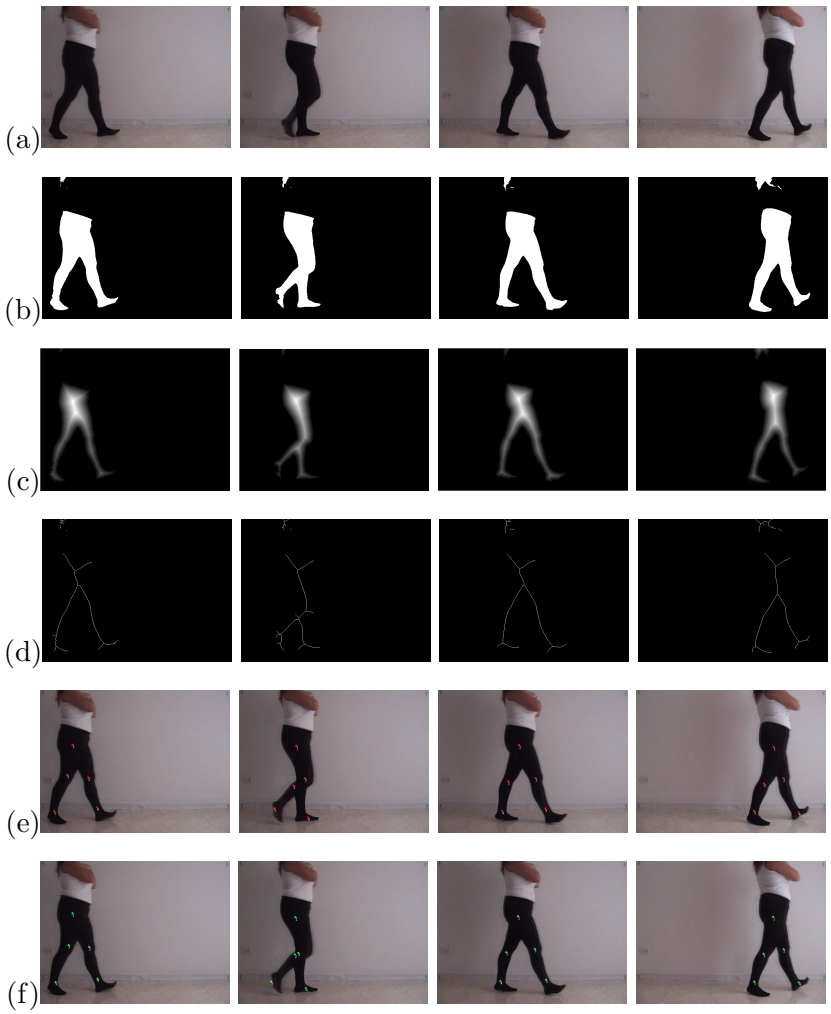


Figura 4.6. Proceso y resultados de la aplicación en mujer adulta de 21 años. Ambiente con iluminación natural, 34 cuadros de video. Las imágenes muestran los cuadros 163, 170, 178 y 196 de la secuencia, así: (a)video original, (b)segmentaciones, (c)carta de distancias, (d)esqueleto, (e)mediciones y (f)estimaciones.

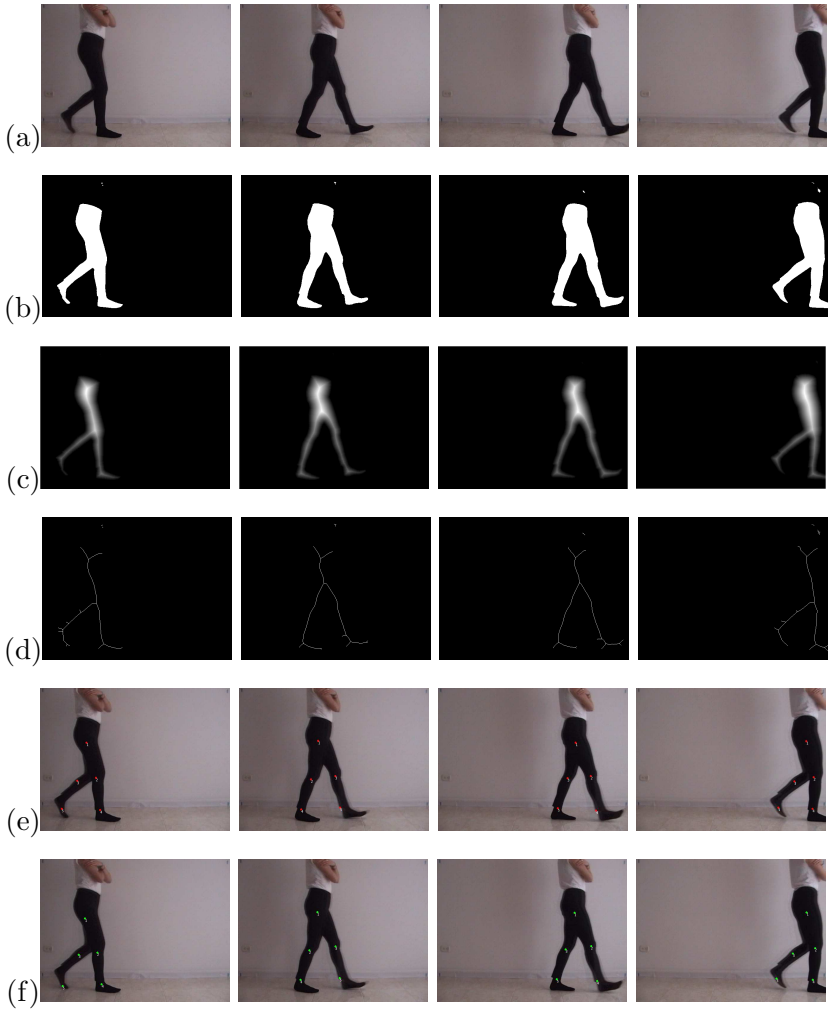


Figura 4.7. Proceso y resultados de la aplicación en mujer adulta de 32 años. Ambiente con iluminación natural, 36 cuadros de video. Las imágenes muestran los cuadros 160, 170, 186 y 195 de la secuencia, así: (a)video original, (b)segmentaciones, (c)carta de distancias, (d)esqueleto, (e)mediciones y (f)estimaciones.

4.6. Validación de resultados

En la figura 4.8 se observan los resultados de aplicar algunos métodos de detección de objetos explicados en la sección 3.2. Cada resultado se contrastó con una imagen de referencia (*foreground truth*), segmentada manualmente con el fin de demostrar que el método de estimación de fondo Σ-Δ tiene un mejor desempeño. El análisis se hizo utilizando el coeficiente de similitud de *Jaccard*[1]. Los rangos de este coeficiente se encuentran entre 0 y 1, donde el 1 significa en este caso que las imágenes son completamente semejantes entre sí [41]. Dados dos objetos binarios A y B, su coeficiente de similitud de *Jaccard* está dado por [45]:

$$J = \frac{M_{11}}{M_{01} + M_{10} + M_{11}}$$

donde

M_{11}: número total de características donde A y B tienen valor de 1.

M_{01}: número total de características donde A es 0 y B es 1.

M_{10}: número total de características donde A es 1 y B es 0.

Tabla 4.1. Coeficientes de similitud de *Jaccard*.

Método	Coeficiente
Media adaptativa	0.9077
Diferencia de cuadros	0.9393
Sigma-Delta	0.9920

En la tabla 4.1 se observan los coeficientes de similitud de *Jaccard* calculados para la media adaptativa, diferencia de cuadros y estimación $\Sigma - \Delta$. Es evidente por los resultados cualitativos y cuantitativos que el método de estimación $\Sigma - \Delta$ tiene mayor semejanza con la imagen de referencia y por lo tanto permite obtener una silueta de los miembros inferiores bien definida.

La postura de los miembros inferiores se determinó manualmente en cada cuadro del video de acuerdo a las características de un ciclo de marcha,

[1] Coeficiente de similitud de *Jaccard*: estadístico utilizado para comparar la similitud y diversidad de los grupos de una muestra.

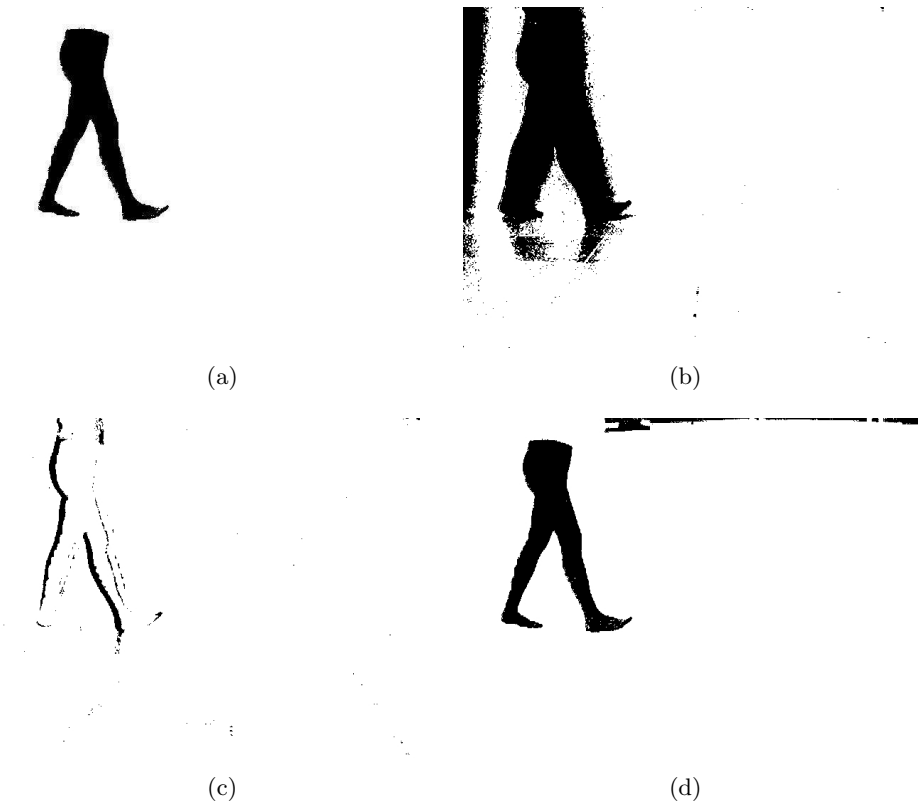


Figura 4.8. Comparación entre algunos métodos de detección de objeto. (a)*foreground truth* o imagen de referencia, (b)resultado de la detección utilizando media adaptativa. (c)Método de diferencia de cuadros, (d)resultado de la estimación de fondo $\Sigma - \Delta$.

en el cual el pie inicia con un evento específico hasta que este se repite en el mismo miembro; posteriormente se obtuvieron los desplazamientos angulares alrededor de la cadera y la rodilla para el método implementado, para el método que utiliza marcadores y para las marcaciones de un experto. El cálculo de los desplazamientos angulares se llevó a cabo, estimando los ángulos absolutos de tronco, muslo y pierna, además de los ángulos relativos de cadera y rodilla [22]. El ángulo absoluto de tronco, muslo y pierna (ver figura 4.9) describe la orientación de cada uno de estos segmentos corporales en el espacio. Dado que durante la marcha el tronco no presenta movimientos significativos, se asumió su ángulo absoluto igual a 90^o, mientras que en el caso del muslo y la pierna, se tuvieron en cuenta las siguientes relaciones:

$$tan\ \theta_{muslo} = \frac{y_c - y_r}{x_c - x_r} \tag{4.1}$$

$$tan\ \theta_{pierna} = \frac{y_r - y_t}{x_r - x_t} \tag{4.2}$$

donde (x_c, y_c), (x_r, y_r) y (x_t, y_t) son las parejas de puntos que representan respectivamente los centros articulares de cadera, rodilla y tobillo. La combinación de los puntos x y y en las ecuaciones 4.1 y 4.2 puede arrojar resultados en cualquiera de los cuatro cuadrantes del plano cartesiano, los cuales normalmente se ubican bien sea en el primer o cuarto cuadrante. En los otros casos, se hace necesaria una corrección sumando 180^o al ángulo absoluto obtenido.

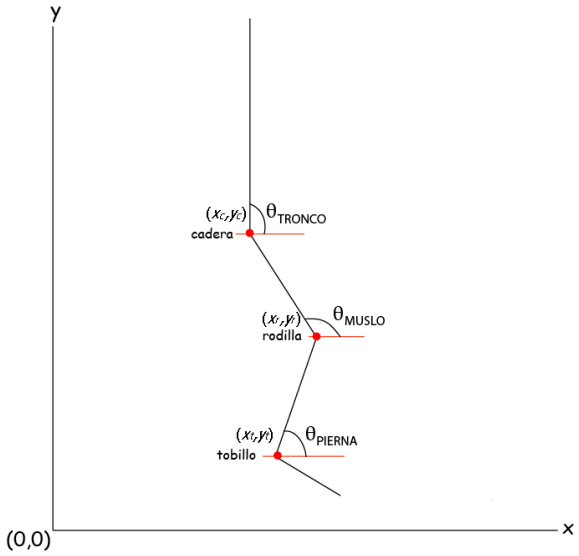


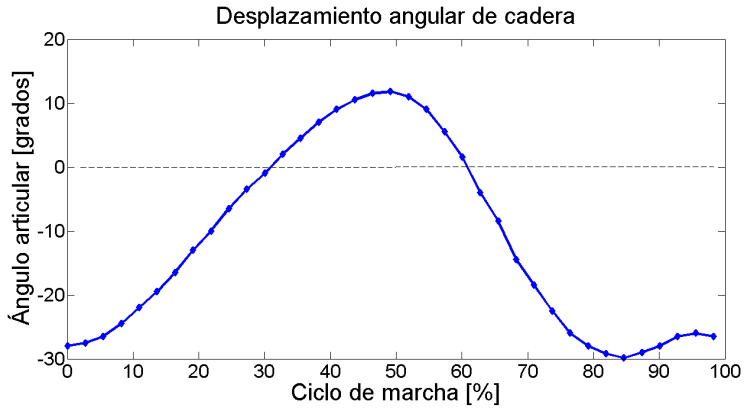
Figura 4.9. Modelo articular del miembro inferior.

Los ángulos relativos de cadera y rodilla, conocidos también como ángulos articulares, describen la cantidad de flexión-extensión en la articulación. Estos ángulos se obtuvieron mediante las siguientes expresiones:

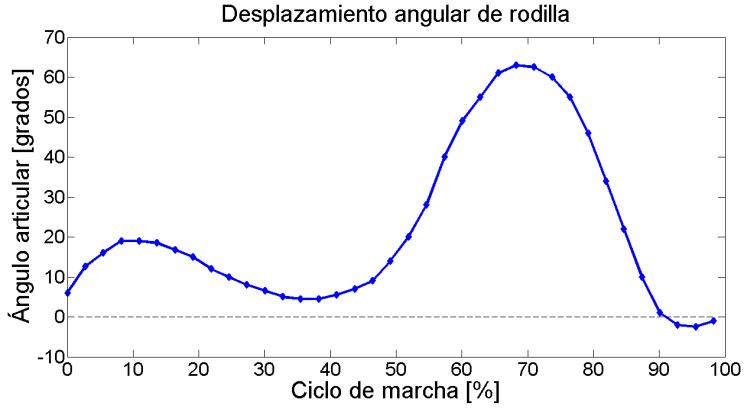
$$\theta_{cadera} = \theta_{muslo} - \theta_{tronco} \tag{4.3}$$

$$\theta_{rodilla} = \theta_{muslo} - \theta_{pierna} \tag{4.4}$$

Los ángulos relativos obtenidos fueron la base para conocer los desplazamientos articulares angulares alrededor de los ejes de movimiento del miembro inferior principal durante un ciclo de marcha en terreno plano. Las figuras 4.10(a) y (b) muestran las curvas teóricas de los desplazamientos angulares de cadera y rodilla [66], las cuales fueron utilizadas como base de comparación cualitativa de las curvas encontradas con el método desarrollado.



(a)



(b)

Figura 4.10. Desplazamientos angulares teóricos. (a)Cadera, (b)Rodilla.

Las figuras 4.11, 4.12, 4.13 y 4.14 muestran cuatro diferentes curvas de desplazamientos angulares de cadera y rodilla en un ciclo de marcha, de cuatro individuos (dos hombres y dos mujeres), obtenidos con el método que utiliza marcadores, con el método implementado en esta investigación

y con las marcaciones de un experto. Cada grupo de curvas pertenecen al mismo individuo con el fin de minimizar los errores que se pueden introducir al comparar los resultados del método desarrollado con aquellos generados mediante el método que utiliza marcadores cuando estos últimos corresponden a diferentes personas. Cualitativamente, en estas curvas se observa un comportamiento global muy similar a las tendencias mostradas en las curvas teóricas (ver figura 4.10). Los resultados experimentales evidencian que en la fase de apoyo, durante el prebalanceo y balanceo izquierdo (aproximadamente 20 %-40 % del ciclo) y en la fase oscilante durante el prebalanceo y balanceo derecho (aproximadamente 65 %-85 % del ciclo) las curvas presentan un comportamiento diferente a los resultados con marcadores, debido a que en esos periodos se presenta oclusión de las rodillas y de los tobillos y a que durante estas fases los centros articulares de interés no corresponden con los ejes laterales de la silueta obtenidos mediante el proceso de esqueletización.

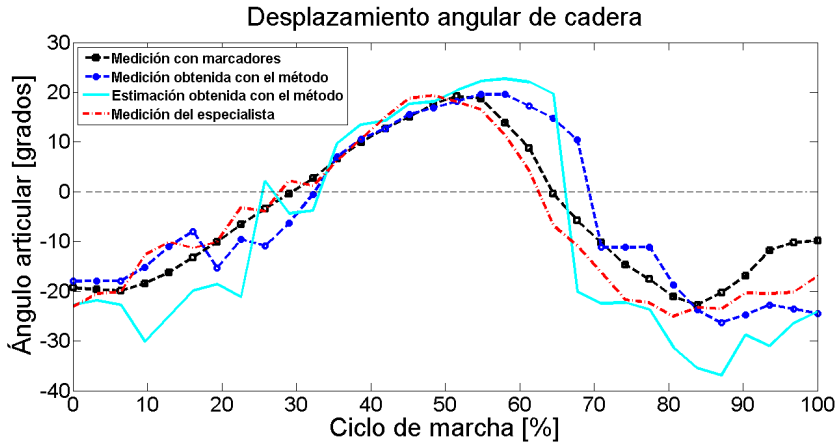
Las curvas 4.11 son las más similares a las teóricas 4.10 porque el ciclo de marcha empezó en contacto inicial derecho. En el caso de las curvas 4.12 y 4.13 hay variabilidad porque el ciclo de marcha de cada individuo empezó en contacto inicial izquierdo, además en el caso de las curvas 4.12 la persona lleva pantalones holgados, lo cual introduce ciertas incertidumbres al resultado. Lo mismo sucede con las curvas 4.14 cuyo ciclo de marcha correspondiente empieza en prebalanceo izquierdo. En general estos dos últimos grupos de curvas presentan la tendencia de las curvas teóricas (figura 4.10). En todos los casos, la mayor variabilidad se presenta en las estimaciones, debido a que en ocasiones los puntos articulares se ubican muy alejados del esqueleto, e incluso algunos puntos se ubican por fuera de la silueta. Esta situación no sucede con las mediciones, las cuales están ligadas al esqueleto y por tanto tienden a ser más fieles a la realidad.

Las curvas de los desplazamientos angulares de cadera y rodilla obtenidas con el método desarrollado y con la marcación de un experto, se contrastaron con aquellas obtenidas mediante el método que utiliza marcadores (referencia). Para tal fin se implementó un estimador llamado distancia de *Fréchet*. La distancia de *Fréchet* es una medida de similitud entre dos curvas, la cual tiene en cuenta la ubicación y el orden de los puntos a lo largo de éstas. Una curva se define como un mapeo continuo $f : [a, b] \to V$, donde $a, b \in \Re$, $a \leq b$ y (V, d) es un espacio métrico. Dadas dos curvas $f : [a, b] \to V$ y $g : [a', b'] \to V$, su distancia de *Fréchet* se define como [19]:

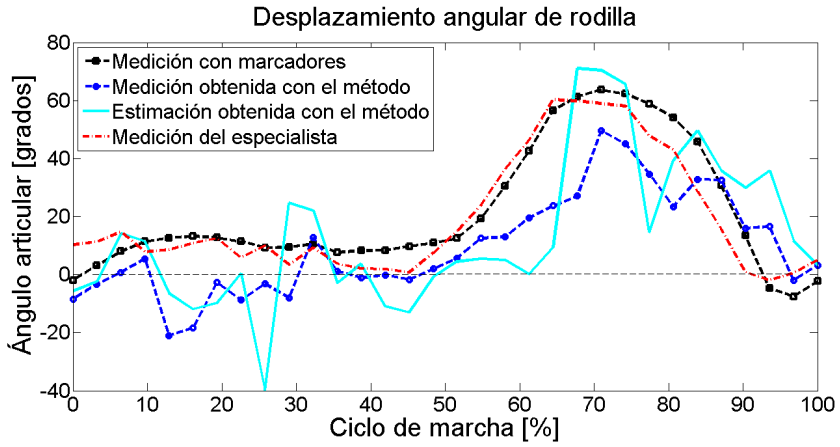
$$\delta_F(f,g) = \underset{\alpha,\beta}{inf}\, \underset{t\in[0,1]}{max}\, d\left(f\left(\alpha\left(t\right)\right), g\left(\beta\left(t\right)\right)\right)$$

donde α y β son funciones continuas arbitrarias no decrecientes de $[0,1]$ sobre $[a,b]$ y $[a',b']$ respectivamente. En las tablas 4.2 y 4.3 se observan los resultados de este estimador sobre las curvas de desplazamientos articulares angulares de cadera y rodilla. La información mostrada en las tablas se encuentra discriminada de acuerdo a las condiciones de iluminación. En la tabla 4.2 se observan los resultados obtenidos bajo condiciones semicontroladas, mientras que en la tabla 4.3 se aprecian los resultados con iluminación natural. Ambas tablas están organizadas por individuos, de acuerdo al sexo y a la edad.

Los resultados muestran la distancia máxima inferior entre dos puntos diferentes de dos curvas con longitudes distintas tal que no se permitan retrocesos de los puntos de las curvas, como sucede en la marcha, proceso durante el cual la velocidad es constante o variable, se repiten ciertos patrones y durante ésta, las personas siempre están moviéndose hacia adelante. En estos resultados se evidencia que hay similitud entre los desplazamientos angulares de cadera y rodilla obtenidos con el método desarrollado, con el que utiliza marcadores y con las marcaciones realizadas por un experto. Al comparar en cada una de las tablas los datos obtenidos por sexo, se puede evidenciar que en todos los casos los resultados son muy similares, se presentan algunas diferencias significativas en los desplazamientos angulares de rodilla, en el caso de la tabla 4.2, donde la referencia (método con marcadores) y las estimaciones de hombres y mujeres tienen valores de media y desviación estándar distintas. Al analizar los resultados en cuanto a grupos de edades, por ejemplo, en el caso de los hombres, aquellos cuyas edades están comnprendidas entre 30 y 38 años y las mujeres cuyas edades están entre 21 y 48 años, se presentan disparidadess notorias en el caso de mediciones y estimaciones de rodilla frente al método que utiliza marcadores. Las diferencias que se aprecian en cada uno de estos casos se deben en primer lugar a que en algunos videos, mientras el sistema se estabiliza, los puntos estimados se ubican muy alejados del esqueleto e incluso por fuera de éste; en segunda instancia, esto se debe a la oclusión que se presenta en las fases de prebalanceo y balanceo derecho e izquierdo y por último, porque si bien los desplazamientos angulares se estimaron sobre el mismo individuo, éstos fueron calculados sobre videos diferentes, en los cuales se presenta variabilidad en la velocidad y el ancho del paso de la marcha, situación que introduce ruido a los resultados.

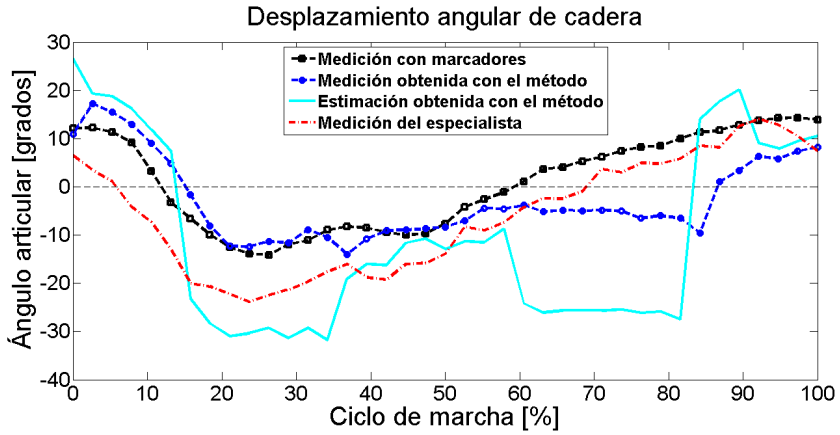


(a)

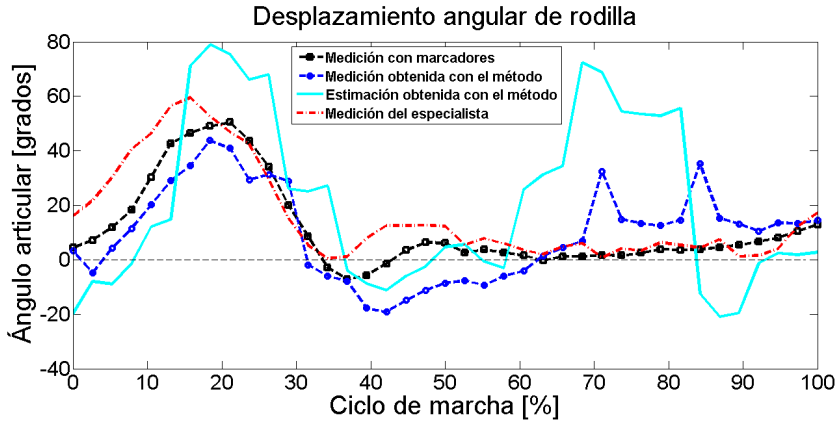


(b)

Figura 4.11. Desplazamientos angulares en un niño de 9 años, correspondientes a la secuencia mostrada en la figura 4.4. (a)Cadera, (b)Rodilla.

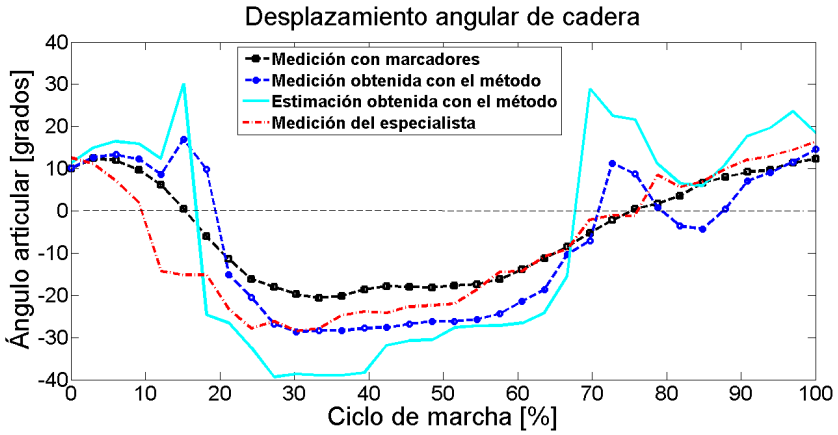


(a)

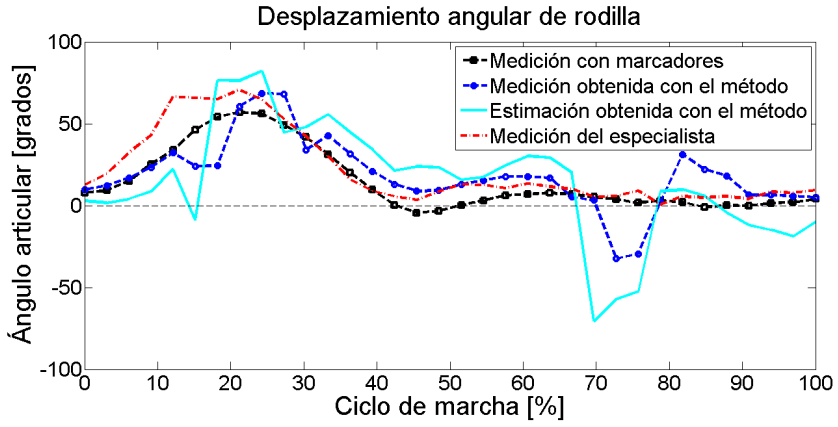


(b)

Figura 4.12. Desplazamientos angulares en hombre adulto de 73 años, correspondientes a la secuencia mostrada en la figura 4.5. (a)Cadera, (b)Rodilla.

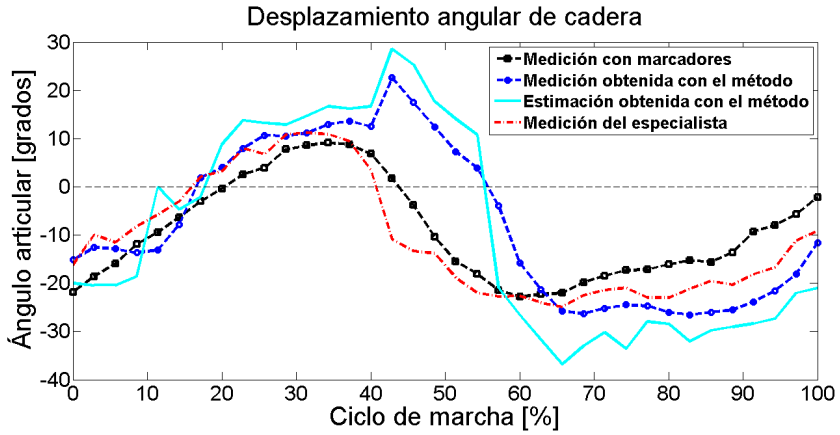


(a)

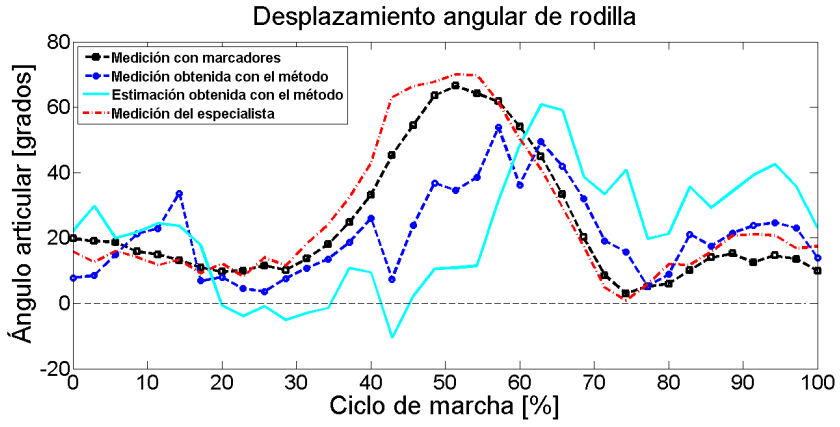


(b)

Figura 4.13. Desplazamientos angulares en mujer adulta de 21 años, correspondientes a la secuencia mostrada en la figura 4.6. (a)Cadera, (b)Rodilla.



(a)



(b)

Figura 4.14. Desplazamientos angulares en mujer adulta de 32 años, correspondientes a la secuencia mostrada en la figura 4.7. (a)Cadera, (b)Rodilla.

Tabla 4.2. Distancia de *Fréchet. Iluminación semicontrolada.*

Iluminacion Semicontrolada				
		Mediciones	Estimaciones	Especialista
Niño	Cadera	0.713	0.698	0.754
9 años	Rodilla	0.644	0.625	0.631
Hombre adulto 2	Cadera	0.765	0.775	0.882
32 años	Rodilla	0.711	0.725	0.662
Hombre adulto 3	Cadera	0.809	0.770	0.824
32 años	Rodilla	0.659	0.677	0.661
Mujer adulta 1	Cadera	0.769	0.759	0.829
21 años	Rodilla	0.621	0.630	0.663
Mujer adulta 2	Cadera	0.602	0.612	0.616
32 años	Rodilla	0.636	0.625	0.680
Mujer adulta 3	Cadera	0.610	0.616	0.613
34 años	Rodilla	0.601	0.602	0.620
Mujer adulta 4	Cadera	0.649	0.602	0.585
36 años	Rodilla	0.673	0.596	0.572
Mujer adulta 5	Cadera	0.785	0.770	0.772
48 años	Rodilla	0.642	0.642	0.673

4.6.1. Discusión

En este capítulo se presentó el desarrollo del método de extracción y seguimiento de los miembros inferiores durante la marcha sin marcadores. El principal problema abordado fue la extracción de los miembros inferiores en cada secuencia de video. Este proceso se llevó a cabo mediante el método de estimación de fondo estático Σ - Δ abordado en la sección 3.2.1. Al objeto binario obtenido se le aplicó una transformación de distancia y una posterior esqueletización (sección 4.3), la cual permite tener información de la forma del objeto. Posteriormente se realizó la detección, la cual consiste en hacer corresponder la observación (silueta) con un conjunto de puntos representativos de la articulación de cadera (la que se aprecia en el video), las rodillas

Tabla 4.3. Distancia de *Fréchet. Iluminación natural*

Iluminacion Natural				
		Mediciones	Estimaciones	Especialista
Niño	Cadera	0.669	0.725	0.782
9 años	Rodilla	0.732	0.718	0.723
Hombre adulto 3	Cadera	0.761	0.740	0.774
32 años	Rodilla	0.696	0.664	0.669
Mujer adulta 1	Cadera	0.784	0.747	0.808
21 años	Rodilla	0.637	0.636	0.694
Mujer adulta 2	Cadera	0.612	0.612	0.633
32 años	Rodilla	0.658	0.607	0.645
Mujer adulta 3	Cadera	0.746	0.764	0.696
34 años	Rodilla	0.733	0.708	0.727
Mujer adulta 4	Cadera	0.634	0.665	0.634
36 años	Rodilla	0.608	0.627	0.687

y los tobillos, para obtener la información del movimiento de los miembros inferiores a lo largo de la secuencia de video, mediante la implementación del filtro *Kalman* (sección 3.2.2). El método se probó sobre 22 secuencias de video capturadas bajo condiciones de iluminación natural y semicontrolada. Los resultados obtenidos se validaron mediante la distancia de *Fréchet*, encontrándose un comportamiento similar frente al método convencional que utiliza marcadores (referencia).

En la literatura (capítulo 2) se encuentran diferentes métodos de análisis de movimiento sin marcadores, los cuales varían de acuerdo al número de cámaras utilizadas, la representación de los datos capturados, tipos de algoritmos y la aplicación a regiones específicas o a todo el cuerpo humano. En el método mostrado por Persson [47] el análisis se hizo para una sola pierna, razón por la cual se pierde información del movimiento, además al trabajar con contornos, estos cambian su forma durante la marcha y se presenta un inconveniente similar a cuando se usan marcadores y no están bien adheridos al cuerpo. En el caso de Zhou *et al.* [69] y Goffredo *et al.* [35] aplicaron

sus métodos respectivamente a reconocimiento de la marcha (biometría) y a vigilancia, mas no a la detección de los centros articulares requeridos en el análisis de marcha. En cuanto a las investigaciones encaminadas a modelos 3D, basan fundamentalmente este proceso al modelado de superficies como es el caso de Gavrila *et al.* [21], Marzani *et al.* [20], Meyer *et al.* [14], Wan *et al.* [7], Kehl *et al.* [29] y Lu *et al.* [64] quienes tampoco enfocaron sus estudios al cálculo de los centros articulares de los miembros inferiores. Otros inconvenientes se presetan en el trabajo de Kakadiaris *et al.* [28], el cual se centró en los miembros superiores, Beriault *et al.* [50] enfocaron su análisis en el desempeño de un pianista y Chu *et al.* [12] trabajaron un método sin modelo en el cual el objetivo central es determinar la postura de la persona. Una situación diferente se presente en la investigación de Corazza *et al.* [51], quienes obtuvieron una estimación de centros articulares, sin embargo encontraron algunos inconvenientes en el caso de la cadera.

El método desarrollado en esta investigación para la extracción y seguimiento de los miembros inferiores sin utilizar marcadores tiene un funcionamiento adecuado. La captura del movimiento se realizó con los brazos recogidos, debido a que el análisis se centra en los miembros inferiores y los brazos generan inconvenientes para detectar el punto articular de cadera. Posterior al proceso de captura, el método extrae la silueta de los miembros inferiores con una definición del 99 %. Asimismo estima los centros articulares de cadera, rodillas y tobillos en cada secuencia de marcha. Este método fue evaluado en 22 videos, capturados bajo condiciones de iluminación natural y semicontrolada. La validación se llevó a cabo comparando los desplazamientos angulares de rodilla y tobillo obtenidas con el método desarrollado, con el método que utiliza marcadores y con la marcación de un experto. La inicialización del método, la ubicación de los marcadores y la marcación manual se realizaron con base en los puntos de referencia antropométrica (trocánter mayor, cóndilos femorales y maléolos) [33, 49]. El seguimiento del método con marcadores y el de la marcación de un experto se realizó manualmente. Esto fue ventajoso en el caso en que la persona usaba pantalones holgados, dado que si el marcador se desplazaba, manualmente se estimaba el punto articular de interés.

En general el método presenta buenos resultados y es útil para evitar alteraciones del gesto natural de la marcha, sin embargo cuando se presenta oclusión de rodillas y tobillos (fases de prebalanceo y balanceo izquierdo y derecho del ciclo de marcha), los ejes longitudinales del esqueleto calculado no coinciden con los centros articulares de rodillas y tobillos, razón por

la cual se presentan ciertas alteraciones en las curvas de desplazamientos angulares y por ende algunas diferencias entre este método y el convencional que utiliza marcadores. Este inconveniente puede ser resuelto en trabajos futuros implementando otro tipo de modelos que permitan conservar en todas las fases del ciclo de marcha, la relación de los segmentos articulares.

El seguimiento de los miembros inferiores a lo largo de la secuencia de video, permitió obtener los desplazamientos angulares de la cadera y la rodilla. La variabilidad en los resultados también se ve influenciada por las características de la marcha propias de cada individuo —como la longitud del paso, la velocidad de marcha, las medidas antropométricas, etc.—, por las marcaciones manuales y por el grado de conocimiento del experto.

A pesar de que el proceso de captura de movimiento se realizó con los brazos recogidos y a que en algunos casos la persona utilizaba pantalones holgados, el método arroja curvas de desplazamientos angulares con un comportamiento muy similar a aquellas obtenidas con marcadores o mediante la marcación de un experto. Estos resultados nos llevan a concluir que a futuro este método podría ser utilizado en aplicaciones clínicas de análisis de marcha.

Capítulo 5

Productos de la Tesis

La investigación desarrollada durante la tesis fue presentada en los siguientes eventos:

- **IV Seminario Internacional de Procesamiento y Análisis de Imágenes Médicas (SIPAIM)**.

  Ponencia: Extracción y seguimiento de los miembros inferiores para el análisis de marcha sin uso de marcadores.

  Autores: Lorenza Henao, Antoine Manzanera, Eduardo Romero.

  Año: 2008.

- **Latinamerican Conference on Networked and Electronic Media (LACNEM)**.

  Poster: *Capture and extraction of lower limbs for markerless gait analysis*.

  Autores: Lorenza Henao, Antoine Manzanera, Eduardo Romero.

  Año: 2009.

- **Quinto Seminario Internacional de Ingeniería Biomédica y Quinto Seminario Internacional de Procesamiento y Análisis de Imágenes Médicas (SIB-SIPAIM)**.

  Poster: Extracción y seguimiento de los miembros inferiores para el análisis de marcha sin el uso de marcadores.

  Autores: Lorenza Henao, Antoine Manzanera, Eduardo Romero.

Publicado en: Memorias del Quinto Seminario Internacional de Ingeniería Biomédica y Quinto Seminario Internacional de Procesamiento y Análisis de Imágenes Médicas.

Año: 2009.

- **Sexto Seminario Internacional de Procesamiento y Análisis de Imágenes Médicas (SIPAIM2010)**.

  Presentación oral: Extracción y seguimiento de los miembros inferiores para el análisis de marcha sin el uso de marcadores.

  Autores: Lorenza Henao, Antoine Manzanera, Eduardo Romero.

  Publicado en: Memorias del Sexto Seminario Internacional de Procesamiento y Análisis de Imágenes Médicas.

  Año: 2010.

Capítulo 6

Conclusiones y Perspectivas Futuras

La extracción de una silueta bien definida de los miembros inferiores es uno de los aspectos más importantes en la aplicación del análisis de marcha sin marcadores. Durante esta investigación se probaron diferentes métodos de segmentación como la máscara de detección de cambio (CDM), diferencia de cuadros, estimación de fondo con la media adaptativa y estimación recursiva de fondo estático *Sigma-Delta* ($\Sigma - \Delta$), encontrando que este último permite realizar una detección precisa de los objetos en movimiento. Las principales ventajas de este método son su bajo consumo en memoria y poco costo computacional. Este método fue comparado con otras técnicas relacionadas, encontrando un mejor desempeño en la estimación $\Sigma - \Delta$ en cuanto a la rapidez en el procesamiento y la definición de la silueta obtenida.

El cálculo de la transformada geodésica sobre la silueta de los miembros inferiores y el adelgazamiento de la carta de distancias mediante técnicas de esqueletización permitió obtener la estructura (esqueleto) de la silueta, la cual fue necesaria para realizar la detección de los puntos articulares de cadera, rodillas y tobillos y, su posterior seguimiento a lo largo de toda la secuencia de video.

El seguimiento de los miembros inferiores a lo largo de la secuencia de video, permitió obtener los desplazamientos angulares de la cadera y la rodilla. Las curvas que describen estos desplazamientos se compararon con aquellas obtenidas utilizando el método de marcadores, con las marcaciones realizadas por un especialista y con curvas teóricas, encontrando que

en el caso de la cadera, el método desarrollado se asemeja a las referencias establecidas, mientras que en el caso de las rodillas se presentan algunos comportamientos diferentes en las fases de prebalanceo y balanceo del ciclo de marcha, debido a la oclusión de las rodillas y los tobillos. La variabilidad en los resultados también se ve influenciada por las características de la marcha propias de cada individuo —como la longitud del paso, la velocidad de marcha, las medidas antropométricas, etc.—, por las marcaciones manuales y por el grado de conocimiento del experto. En general, este método arroja resultados muy alentadores, los cuales en el caso de la cadera se asemejan en un 70 % a aquellos obtenidos con el método que utiliza marcadores y en el de la rodilla su silimitud es del 64 %. Dados estos datos, se concluye que el método podría ser utilizado a futuro en aplicaciones médicas del análisis de marcha.

El método desarrollado conduce a futuras investigaciones en las cuales se integre este análisis monocular a un sistema multi-vista, de forma tal que se pueda obtener una completa representación 3D del cuerpo humano, lo que permite evitar las dificultades presentadas en las fases donde ocurre la oclusión, para obtener información mucho más certera de las posiciones articulares de caderas, rodillas y tobillos.

Esta aplicación también se encamina como trabajo futuro a realizar el análisis de marcha *in situ, donde el paciente se encuentre en un ambiente conocido y propio. También busca generalizar el no uso de los marcadores en este tipo de análisis, lo que generaría un importante logro al evitar alterar completamente el gesto de marcha.*

Apéndice A

Algoritmos Implementados

A.1. Estimación de Fondo $\Sigma - \Delta$

Cálculo de la media $\Sigma - \Delta$
Inicialización
  para cada pixel x*:*
    $M_0(x) = I_0(x)$
Para cada cuadro t
  para cada pixel x*:*
    si $I_t(x) > M_{t-1}(x) \Rightarrow M_t(x) = M_{t-1}(x) + 1$
    si $I_t(x) < M_{t-1}(x) \Rightarrow M_t(x) = M_{t-1}(x) - 1$

Cálculo de la diferencia entre la imagen y la media $\Sigma - \Delta$
(medida de verosimilitud del movimiento)
Para cada cuadro t
  para cada pixel x*:*
    $Diff_t(x) = |M_t(x) - I_t(x)|$

Cálculo de la varianza $\Sigma - \Delta$ *definida como la media* $\Sigma - \Delta$
de N *veces las diferencias diferentes a cero*
Para cada cuadro t
  para cada pixel x *tal que* $Diff_t(x) \neq 0$*:*
    si $V_{t-1}(x) < NxDiff_t(x) \Rightarrow V_t(x) = V_{t-1}(x) + 1$
    si $V_{t-1}(x) > NxDiff_t(x) \Rightarrow V_t(x) = V_{t-1}(x) - 1$

Cálculo de la etiqueta de movimiento haciendo comparaciones entre la diferencia y la varianza
Para cada cuadro t
  para cada pixel x*:*
    si $Diff_t(x) < V_t(x)$
      entonces $D_t(x) = 0$
      sino $D_t(x) = 1$

A.2. Filtro *Kalman*

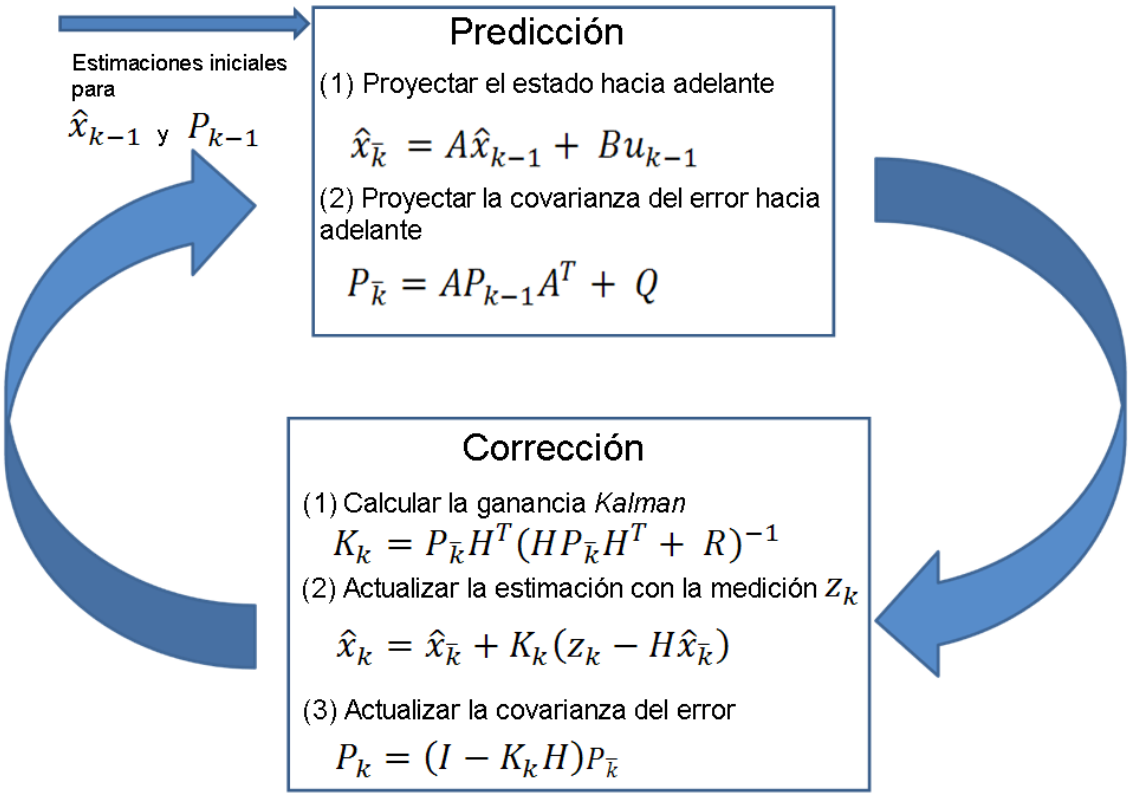


Figura A.1. Diagrama de bloques filtro *Kalman*.

A.3. Flujograma Método Completo

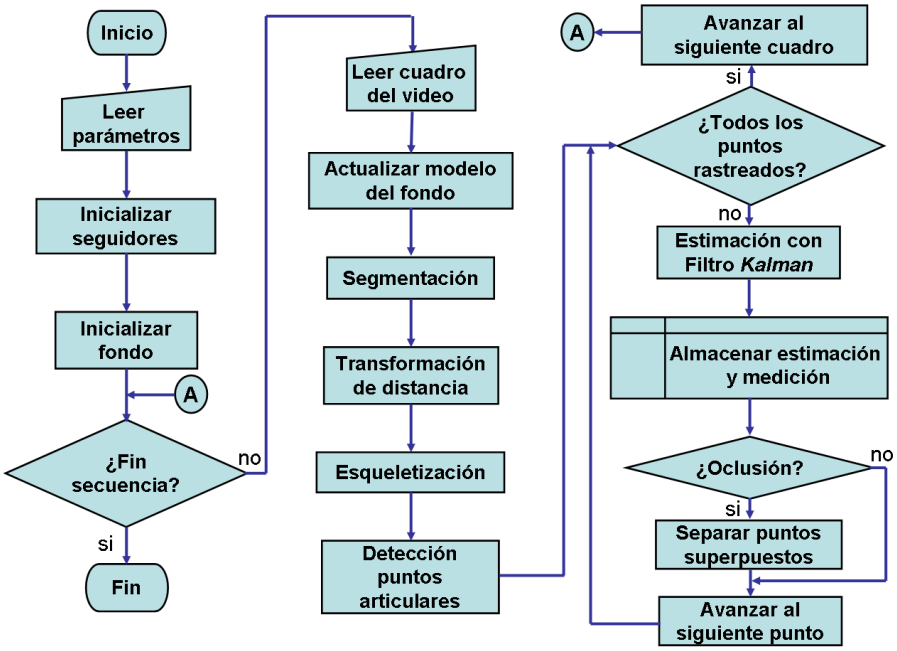


Figura A.2. Diagrama de flujo método completo.

Apéndice B

Transformación de Distancia

La transformación de distancia es la operación que calcula el mapa de distancias de una imagen binaria que representa al objeto, de forma tal que algún punto en el plano almacena la distancia al objeto de interés más cercano en la escena. La figura B.1(a)(izq.) muestra la distancia desde un punto p al punto q más cercano que pertenece al objeto [11].

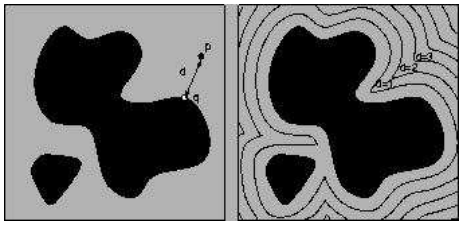


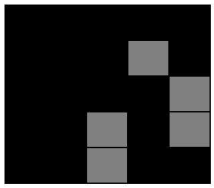
Figura B.1. (a) Distancia desde un punto al objeto (b) Mapa de distancias [11]

Partiendo de una imagen binaria I que consta de un objeto O y su fondo B, una transformación de distancia hace una imagen, el mapa de distancias D, en el cual el valor de algún pixel es la distancia desde este pixel al objeto O. La complejidad computacional es proporcional al producto del número de píxeles en O y en B.

$$D(p) = min\{dist_M(p,q), q \in O\} \tag{B.1}$$

Existen diferentes clases de métricas de distancia en el procesamiento de imágenes:

- *Chessboard.*
- *City-block.*



(a)

31.623	22.361	14.142	1	14.142
28.284	2	1	0	1
22.361	14.142	1	1	0
2	1	0	1	0
2	1	0	1	1

(b)

Figura B.2. Ejemplo carta de distancias (a) Objeto binario (b) Mapa de distancias

- *Chamfer.*
- *Euclidiana.*

Las distancias chessboard y city-block son fáciles de estimar y pueden ser acumuladas recursivamente considerando solamente un vecindario local a la vez. Estas distancias son diferentes de la distancia euclidiana y aunque pueden ser utilizadas en algunas aplicaciones, la más común es la métrica euclidiana [67]. Al igual que las distancias chessboard y city-block, la distancia chamfer también puede ser deducida a partir del valor de sus vecinos, sin embargo esta métrica es sólo una gruesa aproximación de la distancia euclidiana.

La transformada de distancia euclidiana (EDT) fue sugerida en 1980 por Danielsson [16]. Esta transformada produce un mapa de distancias en el cual cada pixel es un vector de dos componentes enteros positivos. La transformada de distancia euclidiana es apropiada porque tiene rotación invariante y el esqueleto del objeto en una imagen no cambia si la orientación del objeto se modifica. La desventaja de este método es la dificultad de descomponerlo en operaciones dentro de vecindarios locales porque este involucra una operación no lineal de raíz cuadrada [40].

Con el fin de definir un algoritmo de transformación de distancia euclidiana local, Danielsson [16] propone asumir que el pixel más cercano al objeto (NOP) es una propiedad local. El NOP de un pixel de fondo es el mismo que el de uno de sus vecinos. Con la métrica euclidiana, asumir esto es apropiado en un plano continuo, sin embargo puede ser incorrecto en una cuadrícula discreta, con configuraciones particulares de los píxeles de objetos. Esto produce una aproximación de transformada de distancia euclidiana, en la cual los errores no sistemáticos ocurren únicamente en algunas ubicaciones de la imagen.

La expresión general de la transformada de distancia euclidiana en $\mathbb{Z}^2$ *está dada por:*

$$d_E = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2},$$

donde (x_1, y_1) *y* (x_2, y_2) *son enteros.*

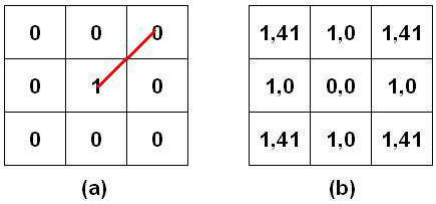


Figura B.3. Métrica Euclidiana. (a) Imagen binaria (b) Transformada de distancia

B.0.1. Distancia Geodésica

La distancia geodésica es una variación de la transformada de distancia, la cual es únicamente definida y calculada sobre una parte de la imagen que puede ser convexa o no. La distancia geodésica entre dos píxeles p *y* q *se define como la longitud del camino más corto entre* p *y* q*. Dado un camino* $P = \{p_1, p_2, ..., p_n\}$ *en un dominio conexo entre los píxeles* p_1 *y* p_n*, la longitud del camino* $l(P)$ *se define como la suma de las distancias vecinas* d_N *entre puntos adyacentes en el camino [11]:*

$$l(P) = \sum_{i=1}^{n-1} d_N(p_i + p_{i+1})$$

Apéndice C

Representación del Cuerpo Humano

En el proceso de seguimiento, los objetos se pueden representar por sus formas y apariencias. De acuerdo a la forma, las representaciones involucran puntos, formas geométricas primitivas, silueta y contorno del objeto, modelos articulados y modelos del esqueleto [1].

- *Puntos. El objeto se representa por un punto que es el centroide o por un conjunto de puntos. En general, la representación de puntos es adecuada para realizar seguimiento de objetos que ocupan pequeñas regiones en la imagen.*
- *Formas geométricas primitivas. Se representa la forma de un objeto por un rectángulo, una elipse, etc. Aunque las formas geométricas primitivas son más apropiadas para representar objetos rígidos sencillos, estas también son utilizadas para seguimiento de objetos no rígidos.*
- *Silueta y contorno del objeto. La representación del contorno define las fronteras de un objeto. La región dentro del contorno es la silueta del objeto. Este tipo de representación es pertinente para seguimiento de formas no rígidas complejas.*
- *Modelos articulados. El cuerpo humano es un objeto articulado con cabeza, torso y miembros superiores e inferiores conectados por articulaciones. Un modelo articulado es usualmente deducido a partir de la descripción morfológica de la anatomía del cuerpo humano, además del conjunto de información relacionada con las cadenas cinemáticas*

y centros articulares, (ver figura C.1). Con el fin de representar un objeto articulado, se pueden modelar sus partes constitutivas empleando una aproximación general como segmentos, cilindros, elipses, superquadrics, etc.

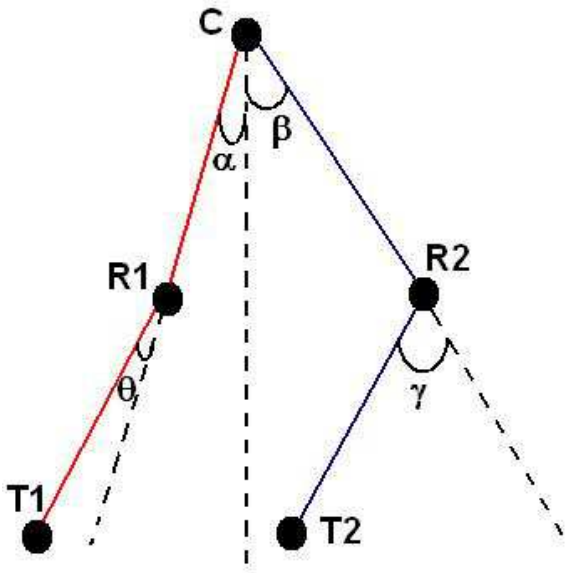


Figura C.1. Modelo articulado. α, β, θ y γ representan los movimientos de flexión y extensión de las articulaciones. C, T1, R1, T2 y R2 constituyen las articulaciones de cadera, rodilla y tobillo de la pierna 1 y 2 respectivamente

- *Modelos de esqueleto. En ocasiones es conveniente modificar un objeto en un modelo tipo esqueleto, proceso que consiste fundamentalmente en obtener un conjunto de líneas que corresponden al adelgazamiento total de la región sin que se pierda su forma básica [63]. El esqueleto de un objeto se puede calcular aplicando transformada de eje medial a la silueta del objeto o mediante operaciones morfológicas. Este tipo de modelo es utilizado frecuentemente como una representación de forma para reconocimiento de objetos. La representación del esqueleto se puede utilizar para representar objetos articulados y rígidos.*

En cuanto a las características de apariencia de los objetos, algunas representaciones de estas son [1]:

- *Densidades de probabilidad de la apariencia del objeto. Las densidades*

de probabilidad de las características de apariencia del objeto como color o textura se pueden calcular a partir de las regiones de la imagen especificadas por los modelos de forma. La densidad de probabilidad puede ser paramétrica(Gaussiana - mezcla de Gaussianas) o no paramétrica (ventana de Parzen - histograma).

- *Plantillas. Las plantillas se constituyen usando formas geométricas sencillas o siluetas. La ventaja de las plantillas es que utilizan la información espacial y de apariencia. Sin embargo, éstas sólo codifican la apariencia del objeto generada desde una sola vista, por lo tanto son apropiadas para el seguimiento de objetos cuya postura no varíe considerablemente durante este proceso.*

- *Modelos de apariencia activa. Estos modelos se generan para modelar simultáneamente la forma y apariencia del objeto. La forma del objeto se define por un conjunto de marcas, las cuales se establecen en las fronteras o dentro de la región del objeto. Para cada marca se almacena un vector de apariencia, el cual está en la forma de color, textura o magnitud del gradiente. Los modelos de apariencia activa requieren de una etapa de entrenamiento en la cual tanto la forma como su apariencia asociada son aprendidas de un conjunto de muestras.*

- *Modelos de apariencia multivista. Estos modelos codifican diferentes vistas de un objeto. Un enfoque para representar las diferentes vistas del objeto es generar un subespacio a partir de unas vistas dadas. Algunas aproximaciones para representación de forma y apariencia son análisis por componentes principales (PCA) y análisis de componentes independientes (ICA). Otras aproximaciones para aprender diferentes vistas de un objeto es realizando entrenamiento de clasificadores como máquinas de soporte vectorial o redes Bayesianas. Una limitación de los modelos de apariencia multivista es que las apariencias en todas las vistas son requeridas con anticipación.*

Bibliografía

[1] *O. Javed A. Yilmaz and M. Shah. Object tracking: A survey.* ACM Computing Surveys*, 38(4):1–45, 2006.*

[2] *T. Aach and A. Kaup. Statistical model-based change detection in moving video.* Signal Processing*, 31:165–180, 1993.*

[3] *T. P. Andriacchi and E. J. Alexander. Studies of human locomotion: Past, present and future.* Journal of Biomechanics*, 33:1217–1224, 2000.*

[4] *R. Baker. The history of gait analysis before the advent of modern computers.* Gait and Posture*, doi:10.1016/j.gaitpost.2006.10.014, 2007.*

[5] *J. Van Bogart. Motion analysis technologies.* IEEE*, pages 166–172, 2000.*

[6] *J. K. Tsotsos C. Bauckhage and F. E. Bunn. Automatic detection of abnormal gait.* Image and Vision Computing*, 27(1,2):108–115, 2007.*

[7] *L. Wang C. Wan, B. Yuan and Z. Miao, editors.* Model-Based Markerless Human Body Motion Capture Using Active Contour. Signal Processing, 2008. ICSP 2008. 9th International Conference on. *IEEE, 2008.*

[8] *T. Darrel C. Wren, A. Azarbayejani and A. Pentland. Pfinder: Real-time tracking of the human body.* IEEE Trans. on PAMI*, 1997.*

[9] *E. F. Desseree Calais and L. R. Legrand, editors.* First Results of a Complete Marker-Free Methodology for Human Gait Analysis. Proceedings of the 2005 IEEE Engineering in Medicine and Biology 27th Annual Conference Shanghai, China. *IEEE, 2005.*

[10] *J. C. Candy and G. C. Temes. Oversampling methods for a/d and d/a conversion.* Eds. New York IEEE*, 1991.*

[11] *O. Cuisenaire.* Distance Transformations: Fast Algorithms and Applications to Medical Image Processing. *PhD thesis, Universite catholique de Louvain (UCL), 1999.*

[12] *O.C. Jenkins C.W. Chu and M.J. Mataric, editors.* Markerless Kinematic Model and Motion Capture from Volume Sequences. Proceedings of the 2003 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR03). *IEEE, 2003.*

[13] *et al. D. Anguelov, editor.* SCAPE: Shape Completion and Animation of People, in SIGGRAPH Conference*, 2005.*

[14] *J. Denzler D. Meyer and H. Niemann, editors.* Model Based Extraction of Articulated Objects in Image Sequences for Gait Analysis. *IEEE, 1997.*

[15] *H. Zhang D. Zhou and N. Ray, editors.* Texture Based Background Subtraction. Proceedings of the 2008 IEEE International Conference on Information and Automation June 20 -23, 2008, Zhangjiajie, China. *IEEE, 2008.*

[16] *P.E. Danielsson. Euclidean distance mapping.* Computer Graphics and Image Processing*, 14:227–248, 1980.*

[17] *R.B. Davis. Clinical gait analysis.* IEEE ENGINEERING IN MEDICINE AND BIOLOGY MAGAZINE*, pages 35–40, 1988.*

[18] *N.I. Durlach and A.S. Mavor.* Virtual Reality: Scientific and Technological Challenges. *Committee on Virtual Reality Research and Development, National Research Council. National Academy Press, 1995.*

[19] *T. Eiter and H. Mannila. Computing discrete fréchet distance. Technical report, CD-TR 94/64, Technische Universitat Wien, 1994.*

[20] *E. Calais F. Marzani and L. Legrand. A 3d marker free system for the analysis of movement disabilities an application to the legs.* IEEE Transactions on Information Technology in Biomedicine*, 5(1):18–26, 2001.*

[21] *D. M. Gavrila and L. S. Davis, editors.* 3D Model Based Tracking of Humans in Action: a Multi View Approach. *IEEE, 1996.*

[22] *J. Hamill and K.M. Knutzen.* Biomechanical Basis of Human Movement. *Lippincott Williams and Wilkins, 2009.*

[23] *S. Haykin.* Sistemas de Comunicación. *Limusa Wiley, 2002.*

[24] *P. Horain and M. Bomb, editors.* 3D Model Based Gesture Acquisition Using a Single Camera. Proceedings ot the Sixth IEEE Workshop on Applications of Computer Vision (WACV02). *IEEE Computer Society, 2002.*

[25] *D. Jarman. A brief introduction to sigma delta conversion.* Intersil. Application Note*, (AN9504):1–7, 1995.*

[26] *S. Ji and H. W. Park, editors.* Region-Based Video Segmentation Using DCT Coefficients. *IEEE, 1999.*

[27] *Ch. Jungik K. Changgeol and L. Yillbyung, editors.* The Relational Properties among Results of Background Subtraction. Feb. ICACT 2008. *IEEE, 2008.*

[28] *I. Kakadiaris and D. Metaxas. Model based estimation of 3d human motion.* IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22(12):1453–1459, 2000.*

[29] *R. Kehl and L. Van Gool. Markerless tracking of complex human motions from multiple views.* Computer Vision and Image Understanding*, 104:190–209, 2006.*

[30] *C. Kirtley.* Clinical Gait Analysis. Theory and Practice. *Elsevier, 2006.*

[31] *S. Corazza L. Mundermann and T. P. Andriacchi, editors.* Accurately Measuring Human Movement Using Articulated ICP with Soft-joint Constraints and a Repository of Articulated Models. *Computer Vision and Pattern Recognition, 2007. CVPR07. IEEE Conference on, 2007.*

[32] *S. Corazza L. Mundermann and T.P. Andriacchi. The evolution of methods for the capture of human movement leading to markerless motion capture for biomechanical applications.* Journal of NeuroEngineering and Rehabilitation*, (doi:10.1186/1743-0003-3-6), 2006.*

[33] *J. Daza Lesmes.* Evaluación Clínico Funcional del Movimiento Corporal Humano. *Editorial Médica Panamericana, 2007.*

[34] *J.N. Carter M. Goffredo and M.S. Nixon, editors.* 2D Markerless Gait Analysis. IFMBE Proceedings. *IEEE, 2008.*

[35] *J.N. Carter M. Goffredo, R.D. Seely and M.S. Nixon. Markerless view independent gait analysis with self-camera calibration.* IEEE*, (978-1-4244-2154-1/08/), 2008.*

[36] *A. Manzanera. Sigma-delta background subtraction and the zipf law.* Pattern Recognition Letters*, 28(3):320–328, 2007.*

[37] *A . Manzanera and J. C. Richefeu. A robust and computationally eficient motion detection algorithm based on sigma-delta background estimation.* Indian Conference on Computer Vision, Graphics and Image Processing (ICVGIP'04). Kolkata-India*, pages 46–51, 2004.*

[38] *A. Manzanera and J. Richefeu. A new motion detection algorithm based on sigma-delta background estimation.* Pattern Recognition Letters*, 28:320–328, 2007.*

[39] *N. McFarlane and C. Schofield. Segmentation and tracking of piglets in images.* Machine Vision and Applications*, 8:187–193, 1995.*

[40] *P.K. Bora N. Sudha, S. Nandi and K. Sridharan. Efficient computation of euclidean distance transform for applications in image processing. pages 49–52, 1998.*

[41] *R. Nazirah and M. Daud, editors.* A Function Principle Approach to Jaccard Ranking Fuzzy Numbers. *2009 International Conference of Soft Computing and Pattern Recognition. IEEE Computer Society, 2009.*

[42] *C.B. Phillips N.I. Badler and B.L. Webber, editors.* Simulating Humans. *Oxford University Press, Oxford, UK, 1992.*

[43] *L. Niu and N. Jiang, editors.* A Moving Objects Detection Algorithm Based on Improved Background Subtraction. *IEEE, 2008.*

[44] *M. G. Pandy. Computer modeling and simulation of human movement.* Annual Reviews*, 3:245–273, 2001.*

[45] *M. Steinbach Pang-Ning Tan and V. Kumar.* Introduction to Data Mining. *Addison–Wesley, 2005.*

[46] *D. F.J. Perales. Human motion analysis and synthesis using computer vision and graphics techniques. Technical report, Computer Graphics and Vision Group Department of Computer Science Universitat de les Illes Balears (UIB).*

[47] *T. Persson. A marker-free method for tracking human lower limb segments based on model matching.* International Journal of Bio-Medical Computing*, 41:87–97, 1996.*

[48] *P. Power and J. Schoonees. Understanding background mixture models for foreground segmentation.* Imaging and Vision Computing New Zealand, Auckland, NZ*, 2002.*

[49] *D. Tyburski R.C. Davis, S. Õunpuu and J.R. Gage. A gait analysis data collection and reduction technique.* Human Movement Science*, 10:575–587, 1991.*

[50] *M. Cote S. Beriault and P. Payeur, editors.* Volumetric Modeling with Multiple Cameras for Markerless Motion Capture in Complex Scenes. Instrumentation and Measurement Technology Conference Proceedings, 2008. IMTC 2008. *IEEE, 2008.*

[51] *L. Mundermann S. Corazza, E. Gambaretto and T.P. Andriacchi. Automatic generation of a subject specific model for accurate markerless motion capture and biomechanical applications.* Biomedical Engineering, IEEE Transactions on: Accepted for future publication, *2008.*

[52] *S. Y. Ma S. Y. Chien and L. G. Chen. Efficient moving object segmentation algorithm using background registration technique.* IEEE Trans. Circuits and Systems for Video Technology, *12(7):577–586, 2002.*

[53] *J. Saboune and F. Charpillet, editors.* Using Interval Particle Filtering for Marker less 3D Human Motion Capture. Proceedings of the 17th IEEE International Conference on Tools with Artificial Intelligence (ICTAI05). *IEEE Computer Society, 2005.*

[54] *C. Stauffer and E. Grimson. Learning patterns of activity using real-time tracking.* IEEE Trans. on PAMI, *22(8):747–757, 2000.*

[55] *A. Sundaresan and R. Chellappa, editors.* Markerless Motion Capture using Multiple Cameras. Proceedings of the Computer Vision for Interactive and Intelligent Environment (CVIIE05). *IEEE Computer Society, 2005.*

[56] *D. H. Sutherland. The evolution of clinical gait analysis part l: Kinesiological emg.* Gait and Posture, *14:61–70, 2001.*

[57] *D. H. Sutherland. The evolution of clinical gait analysis. part ii kinematics.* Gait and Posture, *16:159–179, 2002.*

[58] *Vicon Motion Capture System. www.vicon.com, 2007.*

[59] *Du-Ming Tsai and Shia-Chih Lai. Independent component analysis-based background subtraction for indoor surveillance.* IEEE Transactions on Image Processing, *18(1):158–167, 2009.*

[60] *E. Viel.* La Marcha Humana, la Carrera y el Salto. *Masson, S.A., 2002.*

[61] *R. McGhee W. Frey, M. Zyda and W. Cockayne. Off-the-shelf, real-time, human body motion capture for synthetic environments. Technical report, NPSCS-96-003, Computer Science Dept., Naval Postgraduate School, Monterey, Calif., 1996.*

[62] *G. Welch and G. Bishop. An introduction to the kalman filter. Technical report, TR 95-041, Department of Computer Science, University of North Carolina at Chapell Hill, 2006.*

[63] *G. Pajares y J.M. de la Cruz.* Visión por Computador. *Alfaomega, 2002.*

[64] *R. Hartley H. Li Y. Lu, L. Wang and C. Shen. Multi-view human motion capture with an improved deformation skin model.* IEEE Computer Society, *(doi:10.1109/DICTA.2008.14), 2008.*

[65] *R. Aguilera Z. y R.M. Aguilar P. Análisis de desempeño de algoritmos para detección de movimiento para vigilancia por video. Technical report, Instituto Tecnológico y de Estudios Superiores de Monterrey campus Saltillo; Universidad Autonoma de San Luis Potosi.*

[66] *M. Nordin y V.H. Frankel.* Biomecánica Básica del Sistema Musculoesquelético. *McGraw Hill, 2004.*

[67] *Q. Z. Ye. The signed euclidean distance transform and its applications.* IEEE, *1988.*

[68] *S. Jiang Z. Wei and Q. Huang, editors.* A Pixel-Wise Local Information-Based Background Subtraction Approach. *IEEE, ICME 2008.*

[69] *A. Prugel-Bennett Z. Zhou and R.I. Damper. A bayesian framework for extracting human gait using strong prior knowledge.* IEEE Transactions On Pattern Analysis And Machine Intelligence*, 28(11):1738–1752, 2006.*

[70] *D. Zhang and G. Lu. Segmentation of moving objects in image sequence: A review. Technical report, Gippsland School of Computing and Information Technology Monash University, Churchill, Vic 3842, Australia.*

[71] *Hongxun zhang and De xu, editors.* Robust Estimation for Background Subtraction. Proceedings of the First International Conference on Innovative Computing, Information and Control (ICICIC 06)*. IEEE, 2006.*