



INAOE

Calificación de Gestos Terapéuticos del Brazo Humano con Modelos Ocultos de Markov

por

Gerardo Eliezer Quintana Torres

Tesis

Sometida al programa de ciencias de la computación,
del departamento de Ciencias Computacionales,
como requisito parcial para obtener el grado de

Maestro en Ciencias de la Computación

en el

Instituto Nacional de Astrofísica, Óptica y Electrónica

Octubre 2007

Tonantzintla, Puebla

Asesores:

Dr. Enrique Sucar Succar, INAOE

Dr. Leopoldo Altamirano Robles, INAOE

©INAOE, 2007

Derechos reservados

El autor otorga al INAOE el permiso de reproducir y
distribuir copias de esta tesis en su totalidad o en partes



Agradecimientos

Al Consejo Nacional de Ciencia y Tecnología (CONACYT) por apoyarme con la beca de maestría y al Instituto Nacional de Astrofísica Óptica y Electrónica (INAOE) por la formación académica.

A todos los pacientes del Instituto Nacional de Neurología y Neurocirugía (INNN), así como a Nayelly e Israel terapeutas del INNN.

A mis asesores Dr. Luis Enrique Sucar Succar y el Dr. Leopoldo Altamirano Robles por dirigirme en el transcurso de mi tesis. Así como a Gildardo Azcárate y Héctor Avilés por su valiosa ayuda.

A la Red GFU de Puebla.

A mi tío Arturo.

A Olmo, Félix, Victor y a la alberca del INAOE.

Mi mayor agradecimiento a Cecilia, mis padres, mis hermanas y hermano así como a mi mami Ceci.

Ya que sin todos ellos esta Tesis no se hubiera logrado.

Dedicatoria

A Cecilia.

A Jesús y Guille.

A Ligia, Zoe, Frida y Andrés.

A mi abuela Reyna y mami Ceci.

A Olmo, Félix y Victor.

A Pablo, Memo y Lulú.

Resumen

La enfermedad vascular cerebral afecta a millones de personas cada año en todo el mundo. Una de las secuelas de esta enfermedad es la hemiplejia. La hemiplejia es devastadora en países subdesarrollados o en personas de escasos recursos económicos, ya que para contrarrestar los efectos de la hemiplejia se necesita mucho tiempo de terapia y esta es muy costosa. Actualmente se necesitan nuevas opciones terapéuticas que eviten la necesidad de que el paciente se traslade de su hogar y pague demasiado por la terapia. Para esto se necesitan métodos automáticos que valoren la evolución de la rehabilitación.

En el presente proyecto se plantea el uso de los modelos ocultos de Markov (MOM) para evaluar los movimientos del paciente y valorar su progreso. En particular, utilizamos las métricas Levinson, Kullback - Leibler y Porikli, para comparar MOMs, generar una calificación paramétrica y automática, y así poder valorar ciertos gestos (en este caso los gestos son movimientos terapéuticos). Para esto, se desarrolló un sistema de seguimiento del brazo de una persona, basado en un modelo simplificado, para después calificar 3 gestos: flexión, circular y abducción. En base a esto, se desarrolló un MOM que representa el movimiento de una persona sana para cada gesto, el cual se compara con los MOMs obtenidos de cada paciente. Los resultados son comparados con las escalas que se utilizan en terapia, en particular la de índice motor y la de Fugl - Meyer. Del análisis de los resulta-

dos de las pruebas realizadas, se determina que la métrica para comparar MOMs que mejor se ajusta a calificar los tres gestos, es la de Porikli, en particular con respecto al índice motor.

Esto es la base de un sistema de terapia de bajo costo, enfocado a la rehabilitación de personas con discapacidad motora a causa de una enfermedad vascular cerebral y que puede ser utilizado en su hogar.

Abstract

Stroke or cerebrovascular accident (CVA) affects millions of people every year in all the world. One of the sequels of this disease is the hemiplegy. Because the effects of hemiplegy take a long time of therapy to cure and this is very expensive, it is devastating in underdeveloped countries and for people with limited economic resources. Today, new therapeutics options are required to avoid the need of the patient being in the hospital paying too much for the therapy. For this reason, automatic methods are needed that can evaluate the evolution of the rehabilitation.

In this work we propose the use of hidden Markov models (HMM) to qualify the movements of the patient, so the system can automatically evaluate his progress in the therapy. In particular we use the Levinson, Kullback - Leibler and Porikli metrics to compare HMMs and to generate a parametric and automatic qualification. This qualification is used to evaluate certain gestures (in this case, the gestures are therapeutics exercises). For this a visual tracking system was developed to track the patient's arm, and it serves as input to the models that qualify three gestures: flex, circular and abduction. Based on this, a HMM is learned to represent the movement of a healthy person for each gesture, which is compared with the HMMs obtained for each patient. The results are compared with the scales that are used in therapy, in particular the motricity index and the

Fugl - Meyer scale. From the analysis of several experiments, the Porikli metric was the best to qualify the three gestures, in particular in terms of the motricity index.

This work provides the bases for the development of a low-cost visual therapy system for rehabilitation after stroke.

Índice general

1. Introducción	3
1.1. Antecedentes	3
1.2. Objetivo general	5
1.3. Objetivos específicos	5
1.4. Descripción del sistema/resultados	6
1.5. Organización de la tesis	7
2. Rehabilitación motriz de las extremidades superiores	9
2.1. La enfermedad vascular cerebral	10
2.1.1. Enfermedad vascular cerebral isquémica	11
2.1.2. Enfermedad vascular cerebral hemorrágica	12
2.2. Secuelas de la enfermedad vascular cerebral	13
2.3. Terapia a pacientes hemipléjicos	15
2.3.1. Etapa flácida inicial	16
2.3.2. Etapa de espasticidad	16
2.3.3. Etapa de rehabilitación definitiva	17
2.4. Trabajos previos	17
2.4.1. <i>Java Therapy</i>	18
2.4.2. <i>T-Wrex</i>	20

2.4.3.	<i>Gesture Therapy</i>	20
2.5.	Escalas de la valoración de la discapacidad motriz	22
2.5.1.	Escala de Fugl - Meyer	23
2.5.2.	Índice de motricidad	24
2.5.3.	Motor Activity Log	24
2.6.	Resumen	25
3.	Análisis del movimiento humano	27
3.1.	Detección de movimiento basado en sensores	29
3.1.1.	Inerciales	30
3.1.2.	Infrarrojos	31
3.1.3.	Mixtos	32
3.1.4.	Otros	33
3.2.	Basados en visión con marcas	34
3.2.1.	<i>Motion Capture System (Mocap System)</i>	35
3.2.2.	<i>Webcams</i> - Visión con marcas	37
3.3.	Basados en visión sin marcas	38
3.4.	Comparación de métodos de seguimiento humano	42
3.5.	Métricas para trayectorias	43
3.5.1.	Métricas basadas en distancias	45
3.5.2.	Métricas basadas en características	47
3.6.	Resumen	47
4.	Modelos Ocultos de Markov	49
4.1.	Descripción de los MOM	50
4.2.	Definición	51
4.3.	Los 3 problemas básicos de los MOM	53
4.3.1.	Solución al problema 1: <i>evaluación</i>	54

4.3.2.	Solución al problema 2: <i>decodificación</i>	55
4.3.3.	Solución al problema 3: <i>estimación</i>	57
4.4.	Topologías de los MOM	59
4.5.	Métricas para comparar MOM	60
4.5.1.	Levinson	62
4.5.2.	Kullback-Leibler	63
4.5.3.	Porikli	64
4.6.	Resumen	65
5.	Descripción del sistema	67
5.1.	Introducción	67
5.2.	1a Etapa: Sistema visual	69
5.2.1.	Calibración	70
5.2.2.	Segmentación	73
5.2.3.	Seguimiento 2D	77
5.2.4.	Reconstrucción 3D	80
5.3.	2da. Etapa: Modelo de movimiento	83
5.3.1.	Selección de los gestos terapéuticos	83
5.3.2.	Características del movimiento	85
5.4.	3a. Etapa: Calificación	89
5.4.1.	Modelos de Markov	89
5.4.2.	Métricas	92
5.5.	Resumen	94
6.	Experimentos y resultados	97
6.1.	Seguimiento	98
6.2.	Movimiento humano	100
6.3.	Calificación	102

6.3.1. Levinson	102
6.3.2. Kullback - Leibler	106
6.3.3. Porikli	112
6.4. Comparación	112
6.4.1. Resultados	116
6.5. Persona sana	120
6.6. Análisis de resultados	121
6.7. Resumen	123
7. Conclusiones y trabajo futuro	125
7.1. Aportaciones de la tesis	127
7.2. Trabajo futuro	128

Índice de figuras

2.1. <i>Enfermedad vascular cerebral: a)Hemorrágica: ruptura de una arteria, b)Isquémica: bloqueo del flujo sanguíneo</i>	11
2.2. <i>Prototipo de “Java Therapy”, tele - rehabilitador de regeneración de la fuerza [38].</i>	18
2.3. <i>T-Wrex. a), b) y c) Utilización del brazo robótico, d)Juego de basket ball, e)Juego para limpiar las ventanas, f)Juego de manejo [23].</i>	19
2.4. <i>Gesture Therapy. Las imágenes a), b) y d) muestran el funcionamiento de Gesture Therapy en los juegos de pintar cuarto, de manejo y el de limpiar ventanas, respectivamente, c)Imagen que muestra la localización de la mano por Gesture Therapy y e) y f)Funcionamiento de Gesture Therapy en el juego de basket ball, cambiando de lugar la ventana que muestra el seguimiento de la mano. [46].</i>	21
3.1. <i>Gráfica de la trayectoria del movimiento de la marca de color rojo, que puede representar una articulación del cuerpo.</i>	28
3.2. <i>Xsens Motion System: a)Sensor MTi, b)Colocación en el cuerpo humano, c)Diagrama MTx, muestra las conexiones físicas del sistema, los sensores colocados a un bus, que a su vez, esta conectado a la PC [58].</i>	30

3.3. <i>Phase Space Motion Capture System: a)Led infrarrojo, b)Sensor infrarrojo, c)Calibración de las cámaras infrarrojas, d) y g)Sensores colocados en el cuerpo humano, e)Obtención de la posición 3D, f)Renderización .</i>	31
3.4. <i>Wii mote: a) Sensor inercial, b) Sensor infrarrojo, c) Leds infrarrojos[24]</i>	32
3.5. <i>Johansson: Moving Light Display, a) y b)Muestran las marcas de luz detectadas por el sistema, y c)Renderización por el sistema del cuerpo humano.</i>	34
3.6. <i>Optical Motion Capture, Qualisys [55]. a)Reconstrucción 2D, b)Reconstrucción 3D, c) Seguimiento del movimiento humano en deportes, d) Calibración, e) Seguimiento de movimiento de un barco, en un ambiente simulado. .</i>	36
3.7. <i>Seguimiento de objetos utilizando sustracción de imágenes, a), c) y e) Imágenes de la secuencia de video, b), d) y f) Resultados de la sustracción de imágenes que detectan objetos en movimiento.</i>	38
3.8. <i>Seguimiento del cuerpo humano[53]. a)Imagen original, b)Segmentación, c) Blobs.</i>	39
3.9. <i>Seguimiento del cuerpo humano[4]. Las imágenes muestran la localización del cuerpo humano en movimiento en cada una de las imágenes.</i>	40
3.10. <i>Seguimiento del rostro, Camshift[7]. a) y b) Localización de la posición del rostro, d) y e) Obtención de la distribución de probabilidad de píxeles color piel en la imagen, c) y f) Localización del centro de masa del objeto color piel.</i>	41
3.11. <i>Trayectoria del movimiento de un objeto. a)Detección del centro de masa del objeto en los frames 1 al 8, b)Coordenadas cartesianas del centro de masa del objeto en los frames 1 al 8, c)Trayectoria T_n del objeto donde $n = 8$</i>	44

4.1.	<i>a)Diagrama generalizado de un Modelo Oculto de Markov, b)Representación de una red bayesiana dinámica simple por medio de su red base y su red de transición</i>	50
4.2.	<i>Grafos que muestran las topologías de los MOM: a)Ergódica, b)Bakis, c)Cross - coupled</i>	59
4.3.	<i>Ambigüedades para las métricas convencionales[35]. a)Duraciones diferentes, b)Igual duración, pero diferentes instantes de tiempo, c)Efecto desigual en los instantes de tiempo</i>	60
4.4.	<i>Imagen que muestra la distancia de cruce entre dos MOM[35].</i>	64
5.1.	<i>Diagrama general del sistema. 1a. Etapa: Seguimiento visual, 2da. Etapa Modelo de Movimiento. 3a. Etapa Calificación.</i>	68
5.2.	<i>1a. Etapa: Sistema Visual.</i>	69
5.3.	<i>Parámetros intrínsecos. a)Patrón de calibrado, b) y c)Detección de las esquinas, d) Modelo pinhole.</i>	71
5.4.	<i>Parámetros extrínsecos. a)Configuración de las cámaras, b)Configuración original de las cámaras, c)Rotación de las cámaras.</i>	72
5.5.	<i>Segmentación: marca color rojo. a)Imagen original, b)Imagen que muestra el grado de pertenencia de cada pixel al color rojo deseado, c)Imagen binaria que muestra la pertenencia de cada pixel al color rojo deseado, d)Histograma binario que muestra la pertenencia de cada pixel al color rojo deseado, e)Histograma de la distribución de probabilidades del color rojo deseado, canal H y S, f)Histograma que muestra el grado de pertenencia de cada pixel al color rojo deseado, g)Imagen Segmentada.</i>	76
5.6.	<i>Diagrama de flujo del algoritmo camshift [7].</i>	77
5.7.	<i>Líneas de proyección.</i>	80

5.8. Selección de gestos. a)Flexión - extensión, b)Circular, c)Abducción - aducción.	84
5.9. Coordenadas esféricas, a)Obtención de las coordenadas esféricas en un ambiente 3D, b)Obtención de las coordenadas esféricas en el cuerpo humano, c)La imagen ilustra la postura del brazo en un instante de tiempo. α indica el ángulo entre el brazo y el antebrazo, los ángulos θ y ϕ representan las coordenadas esféricas del codo tomando como referencia al hombro. h , m y c representan las coordenadas tridimensionales (X , Y , Z) del hombro, codo y muñeca, respectivamente.	85
5.10. Obtención de las coordenadas esféricas por el sistema. En las dos imágenes se muestra el seguimiento de las marcas de color en las dos secuencias de video, capturadas por las dos webcams. En la gráfica de abajo se muestran las trayectorias de los tres ángulos: α , θ y ϕ	88
5.11. Topología Bakis Continua de los MOM utilizados en este proyecto de tesis.	91
6.1. Ambiente. Forma de colocar las cámaras y el paciente.	98
6.2. Seguimiento, a), b) y c) Seguimiento de las tres marcas de color en la webcam1 y d), e) y f) Seguimiento de las tres marcas de color en la webcam2.	99
6.3. Gráficas que muestran el movimiento del brazo a partir de la obtención de los ángulos generados por el movimiento del brazo, a)Flexión(3 ángulos), b)Flexión(3d), c)Abducción(3 ángulos), d)Abducción(3d), e)Circular(3 ángulos), f)Circular(3d).	100
6.4. Gráficas que muestran secuencias de observaciones discretas, a)16 clases, b)32 clases y c)64 clases.	101

6.5.	<i>Gráficas comparativas de las calificaciones al gesto flexión, a) y b) Resultados de Levinson en índice motor y Fugl - Meyer, c) y d) Resultados de Kullback - Leibler en índice motor y Fugl - Meyer y e) y f) Resultados de Porikli en índice motor y Fugl - Meyer.</i>	117
6.6.	<i>Gráficas comparativas de las calificaciones al gesto circular, a) y b) Resultados de Levinson en índice motor y Fugl - Meyer, c) y d) Resultados de Kullback - Leibler en índice motor y Fugl - Meyer y e) y f) Resultados de Porikli en índice motor y Fugl - Meyer.</i>	118
6.7.	<i>Gráficas comparativas de las calificaciones al gesto abducción, a) y b) Resultados de Levinson en índice motor y Fugl - Meyer, c) y d) Resultados de Kullback - Leibler en índice motor y Fugl - Meyer y e) y f) Resultados de Porikli en índice motor y Fugl - Meyer.</i>	119

Índice de Tablas

3.1. Comparación de técnicas de seguimiento. VCM = Visión con marcas. VCMI = Visión con marcas infrarrojas.	43
6.1. Muestra los resultados obtenidos al aplicar la métrica Levinson a los modelos del gesto flexión de cada uno de los pacientes y la persona sana, con el modelo del gesto flexión realizado correctamente. Utilizando ob- servaciones discretizadas en 16, 32, 64 128 y 256 clases.	103
6.2. Muestra los resultados obtenidos al aplicar la métrica Levinson a los modelos del gesto circular de cada uno de los pacientes y de la persona sana, con el modelo del gesto circular realizado correctamente. Utilizan- do observaciones discretizadas en 16, 32, 64 128 y 256 clases.	104
6.3. Muestra los resultados obtenidos al aplicar la métrica Levinson a los modelos del gesto abducción de cada uno de los pacientes y de la per- sona sana, con el modelo del gesto abducción realizado correctamente. Utilizando observaciones discretizadas en 16, 32, 64 128 y 256 clases. .	105
6.4. Muestra los resultados obtenidos al aplicar la métrica Kullback - Leibler a los modelos del gesto flexión de cada uno de los pacientes y de la persona sana, con el modelo del gesto flexión realizado correctamente. Utilizando observaciones discretizadas en 16, 32, 64 128 y 256 clases. .	106

6.5.	<i>Muestra los resultados obtenidos al aplicar la métrica Kullback - Leibler a los modelos del gesto circular de cada uno de los pacientes y de la persona sana, con el modelo del gesto circular realizado correctamente. Utilizando observaciones discretizadas en 16, 32, 64 128 y 256 clases. .</i>	107
6.6.	<i>Muestra los resultados obtenidos al aplicar la métrica Kullback - Leibler a los modelos del gesto abducción de cada uno de los pacientes y de la persona sana, con el modelo del gesto abducción realizado correctamente. Utilizando observaciones discretizadas en 16, 32, 64 128 y 256 clases. .</i>	108
6.7.	<i>Muestra los resultados obtenidos al aplicar la métrica Porikli a los modelos del gesto flexión de cada uno de los pacientes y de la persona sana, con el modelo del gesto flexión realizado correctamente. Utilizando observaciones discretizadas en 16, 32, 64 128 y 256 clases.</i>	109
6.8.	<i>Muestra los resultados obtenidos al aplicar la métrica Porikli a los modelos del gesto circular de cada uno de los pacientes y de la persona sana, con el modelo del gesto circular realizado correctamente. Utilizando observaciones discretizadas en 16, 32, 64 128 y 256 clases.</i>	110
6.9.	<i>Muestra los resultados obtenidos al aplicar la métrica Porikli a los modelos del gesto abducción de cada uno de los pacientes y de la persona sana, con el modelo del gesto abducción realizado correctamente. Utilizando observaciones discretizadas en 16, 32, 64 128 y 256 clases. . .</i>	111
6.10.	<i>Tabla que muestra la calificación obtenida por Levinson de la realización del ejercicio flexión de cada uno de los pacientes y de la persona sana.</i>	113
6.11.	<i>Tabla que muestra la calificación obtenida por las escalas, índice motor y Fugl - Meyer.</i>	114
6.12.	<i>Tabla que muestra la distancia euclidiana entre los resultados de las métricas y los resultados de las escalas.</i>	116

6.13. *Tabla que muestra la varianza entre los resultados de las métricas y los resultados de las escalas.* 116

6.14. *Tabla que muestra el promedio de la distancia euclidiana de los tres gestos.*120

6.15. *Tabla que muestra el promedio de la varianza de los tres gestos.* 120

Capítulo 1

Introducción

“El principio es la mitad del todo”.
Pitágoras de Samos (582 AC-497 AC)

1.1. Antecedentes

La detección y entendimiento de gestos en secuencias visuales tiene un alto valor para muchas aplicaciones, tales como, interacción humana con computadoras o robots, vigilancia, entrenamiento y sistemas de soporte médico. Es en esta última en la que el presente trabajo tendrá aplicación. Este interés es motivado gracias a los avances en el campo de visión computacional que generan cambios novedosos y radicales en la forma de interactuar con la computadora.

Cada año 200,000 personas en México sobreviven a un derrame cerebral y cerca de 700,000 en Estados Unidos [56]. Aproximadamente el 80% de ellos pierden la habilidad del movimiento. La deficiencia del movimiento después de un derrame cerebral es generalmente tratada con terapia física intensa por varias semanas después de la lesión, que generalmente es dirigida por un fisioterapeuta que diagnostica si se está llevando la rehabilitación apropiadamente. Desafortunadamente,

la terapia de los pacientes que han sufrido un derrame cerebral es intensa mientras se tiene esta ayuda profesional, así como la economía suficiente para cubrir los gastos que esto implica. La rehabilitación subsecuente en casa, es a menudo autodirigida con información poco profesional. Con lo anterior las metas de la ingeniería en rehabilitación están encaminadas en el desarrollo tecnológico que permita terapias sin la necesidad de que un fisioterapeuta este siempre presente.

Por tal motivo, ha surgido la inquietud de desarrollar sistemas que detecten el progreso de rehabilitación en pacientes con atrofias en la neuronas motoras que impiden al paciente tener una movilidad con libertad. En la actualidad, el experto en rehabilitación, utiliza diversas escalas para valorar la movilidad de sus pacientes. El problema de estas escalas es que la mayoría son ordinales y no paramétricas [32].

En los últimos años se ha realizado un esfuerzo por crear sistemas que ayuden en el proceso de rehabilitación para personas que han sufrido un derrame cerebral, utilizando *joysticks* [38] y algunos con sistemas visuales [60, 49, 48]. El objetivo de la rehabilitación es permitir a las personas que han sufrido un derrame, obtener el nivel más alto posible de independencia para que puedan ser productivos [57].

Los sistemas de seguimiento visual se basan en obtener la información del movimiento humano. Para obtener dicha información se pueden utilizar dos técnicas, utilizando marcas para indicar la posición de las articulaciones, simplificando parcialmente el problema del seguimiento del cuerpo humano y por otro lado, no utilizando marcas. Actualmente la técnica utilizando marcas es la mas utilizada, por su bajo costo y por ser mas precisa que la técnica libre de marcas.

En la presente propuesta se propone utilizar los modelos ocultos de Markov, para que automatice la calificación de la realización de gestos terapéuticos, utilizando información de movimiento. La motivación de utilizar los modelos ocultos de Markov, es por que tienen la habilidad de representar datos en estructuras

secuenciales con presencia de ruido, cuando solo se tiene el conocimiento de las observaciones, encontrando así la secuencia de estados más probable que producen dicha secuencia de observaciones [10].

1.2. Objetivo general

Desarrollar un algoritmo que califique movimientos del brazo humano utilizando modelos ocultos de Markov, y aplicarlo a la evaluación de gestos terapéuticos.

1.3. Objetivos específicos

- Seleccionar 3 gestos (movimientos con el brazo humano) que sean terapéuticos y que puedan ser calificables.
- Seleccionar características del movimiento humano a modelar.
- Obtener estas características del movimiento humano a partir de secuencias de video.
- Modelar los 3 gestos utilizando los modelos ocultos de Markov.
- Obtener por medio de comparar los modelos ocultos de Markov, una calificación clínica de la ejecución del gesto basada en un valor de probabilidad.
- Seleccionar escalas para valorar la movilidad utilizadas en rehabilitación.
- Comparar los resultados obtenidos con las escalas de valoración.

1.4. Descripción del sistema/resultados

En esta tesis se propone utilizar los modelos ocultos de Markov para generar una calificación automática de la realización de gestos con el brazo humano utilizados en rehabilitación. Para lograr esta calificación, se creó un sistema con tres etapas. La primera etapa está compuesta por los módulos capaces de obtener las características del movimiento humano: el segmentador, el seguimiento 2D y la reconstrucción 3D. La segunda etapa, está compuesta por el módulo que procesa las características del movimiento humano para que puedan ser útiles para el modelado. La última etapa, contiene las herramientas para modelar los MOM y compararlos.

Del análisis de los resultados de las pruebas, se determinó que la métrica para comparar modelos ocultos de Markov que mejor se ajusta a calificar los tres gestos, es *Porikli*. En particular, el índice motor es la escala que mejor se ajusta a las métricas que comparan modelos ocultos de Markov que representan trayectorias de los gestos terapéuticos. Lo que se tiene que mejorar y resolver como trabajo futuro es la correlación entre las métricas y las escalas, para esto se pretende utilizar sólo los incisos de la escala que califican el movimiento de la extremidad superior. También se necesita mejorar la precisión de la calificación de las métricas, para esto, se necesita de un método de seguimiento más robusto, o algún método que elimine el ruido de las características del movimiento humano.

En este documento se presentan cuatro aportaciones principales. La principal aportación de este proyecto de tesis, es la obtención de una calificación automática y paramétrica que resuelva el problema de las escalas utilizadas en rehabilitación por ser demasiado subjetivas. La segunda aportación es que esta escala puede ser capaz de ayudar a la motivación y rehabilitación de pacientes con secuelas de enfermedad vascular cerebral como la hemiplejía, sin contar con la presencia de

un terapeuta. La tercera aportación es la utilización de los modelos ocultos de Markov para la calificación de gestos. Esto abre una gran posibilidad de seguir investigando en crear modelos para la calificación de gestos, ya sea para la rehabilitación, entrenamiento, etc. Por último, este proyecto ayuda a la creación de una herramienta de bajo costo que puede ser incluida en un sistema para la rehabilitación de pacientes con hemiplejía capaz de ser utilizada por más personas que no cuenten con los recursos para pagar una rehabilitación o los medios para asistir a un centro de rehabilitación.

1.5. Organización de la tesis

La organización de la tesis continúa con los siguientes capítulos:

- Capítulo 2. En este capítulo se presenta una descripción de la enfermedad vascular cerebral, así como sus efectos, el proceso de rehabilitación y las escalas de valoración de la movilidad
- Capítulo 3. Se presenta una breve introducción de las técnicas y métodos para obtener las características del movimiento humano, así como la comparación entre ellos. También se presentan métricas basadas en distancia para comparar trayectorias.
- Capítulo 4. Se describen los modelos ocultos de Markov, sus tres problemas así como su solución y métricas que comparan modelos ocultos de Markov.
- Capítulo 5. Se describe de manera detallada la estructura general del sistema, así como las tres etapas que los constituyen, seguimiento visual, modelo de movimiento y calificación.

- Capítulo 6. En este capítulo se analizan los resultados de aplicar las métricas *Levinson*, *Kullback - Leibler* y *Porikli* y se comparan los resultados con la valoración realizada a pacientes por un especialista en rehabilitación del Instituto Nacional de Neurología y Neurocirugía.
- Capítulo 7. En este capítulo se redactan las conclusiones de la tesis y se dan posibles caminos que se pueden seguir sobre esta línea de investigación.

Capítulo 2

Rehabilitación motriz de las extremidades superiores

*“Para llamar a alguien con el dedo índice,
no piensen que solo contraen el flexor propio del dedo índice;
inhiben la flexión de todo el brazo”
Dr. Denis Williams, Neurólogo*

Las enfermedades vasculares cerebrales son las causas principales de discapacidad y muerte en muchos países. Un daño de las neuronas motoras puede producir hemiplejía. Inmediatamente después del inicio de la hemiplejía, se pierde el movimiento voluntario en las extremidades afectadas y se disminuye la resistencia al movimiento pasivo, ocasionando flacidez. Esta flacidez puede terminar en pocos días o convertirse en una condición permanente. En muchos casos esta flacidez es gradualmente remplazada por espasticidad¹.

Esta disfunción motriz afecta en gran medida la calidad de vida de las personas que sobreviven a un derrame cerebral. De todas las personas que sufren un derrame cerebral cada año, dos tercios sobreviven y requieren rehabilitación. Los objetivos

¹Tono muscular exagerado de origen cerebral que se manifiesta por contracciones involuntarias de los músculos

de la rehabilitación son ayudar a los sobrevivientes a ser lo más independientes posible y a lograr la mejor calidad de vida posible. Aún cuando la rehabilitación no “cura” la hemiplejia pues no revierte el daño cerebral, la rehabilitación puede ayudar sustancialmente a que las personas logren mejorar su calidad de vida a largo plazo.

Muchas personas que sufren de hemiplejia no logran recuperarse por la falta de un proceso de rehabilitación. Debido a esto, se ve la necesidad de llevar a más personas con secuelas de un derrame cerebral, a tratar alternativas de rehabilitación, que puedan ser utilizadas por ellos mismos. Para esto, es necesario contar con un sistema de seguimiento de ejercicios terapéuticos que sea fácil de usar, de bajo costo y que pueda calificar los avances de la rehabilitación, sin la necesidad de la presencia de un experto en rehabilitación.

En este capítulo se presenta una descripción de la enfermedad vascular cerebral, sus secuelas, algunos trabajos previos sobre rehabilitación y algunas escalas que miden la movilidad corporal.

2.1. La enfermedad vascular cerebral

La *enfermedad vascular cerebral*² (e. v. c.), sucede cuando se interrumpe repentinamente el suministro de sangre a una parte del cerebro o cuando un vaso sanguíneo en el cerebro se rompe derramando sangre en los espacios que rodean a las neuronas. Cuando una persona sufre una pérdida de flujo sanguíneo en el corazón, se dice que ha sufrido un ataque cardíaco, de la misma forma puede decirse que una persona con pérdida de flujo sanguíneo en el cerebro tiene un “ataque cerebral” o sufre un accidente cerebrovascular.

²El término en inglés para describir la enfermedad vascular cerebral es “*stroke*”. En español, se utilizan comúnmente los términos “ataque cerebrovascular” o “derrame cerebral”.

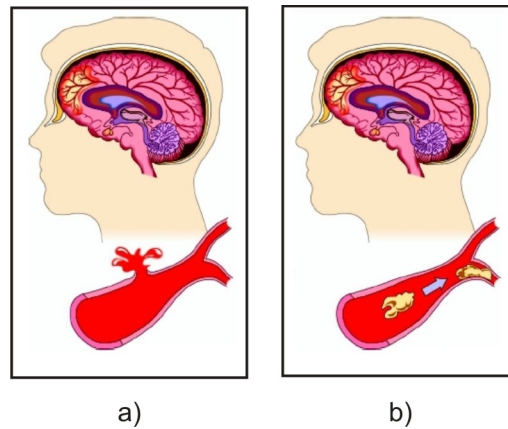


Figura 2.1: *Enfermedad vascular cerebral: a) Hemorrágica: ruptura de una arteria, b) Isquémica: bloqueo del flujo sanguíneo*

Las neuronas mueren cuando dejan de recibir oxígeno y nutrientes de la sangre o cuando son dañadas por una hemorragia repentina en el cerebro. Isquemia es el término utilizado para describir la pérdida de oxígeno y nutrientes en las neuronas cuando no existe un flujo adecuado de sangre. La isquemia conduce finalmente a un infarto y a la muerte de neuronas que con el tiempo son sustituidas por una cavidad llena de fluido en el cerebro lesionado.

Cuando se interrumpe el flujo de sangre al cerebro, algunas neuronas mueren inmediatamente, mientras que otras permanecen con riesgo de morir. Estas células dañadas constituyen la penumbra isquémica y pueden permanecer en un estado de riesgo por varias horas. Con tratamiento oportuno, estas células pueden salvarse.

Hay dos tipos de enfermedad vascular cerebral: la enfermedad vascular cerebral isquémica y la enfermedad vascular cerebral hemorrágica [56].

2.1.1. Enfermedad vascular cerebral isquémica

Una enfermedad vascular cerebral *isquémica* (e. v. c. i.) *Figura 4.1b*, sucede cuando una arteria que suministra sangre al cerebro queda bloqueada, reduciendo o

interrumpiendo repentinamente el flujo de sangre, y con el tiempo, ocasionando un infarto en el cerebro. Aproximadamente un 80 por ciento de todos los accidentes cerebrovasculares son de tipo isquémico [56]. Los coágulos de sangre son la causa más común de bloqueo arterial y de infarto cerebral. Cuando los coágulos de sangre se forman dentro de una arteria, ocasionan una lesión devastadora al interferir con el flujo normal de la sangre.

Los coágulos de sangre pueden ocasionar isquemia e infarto de dos formas: enfermedad vascular cerebral embólica y enfermedad vascular cerebral trombótica. La enfermedad vascular cerebral embólica, sucede cuando un coágulo que se forma en una parte del cuerpo fuera del cerebro puede trasladarse a través de los vasos sanguíneos y quedar atrapado en una arteria cerebral. Este coágulo libre se denomina émbolo y a menudo se forma en el corazón. La enfermedad vascular cerebral trombótica, es ocasionada por una trombosis. Una trombosis es la formación de un coágulo de sangre en una de las arterias cerebrales que permanece fijo a la pared arterial hasta que aumenta de tamaño, lo suficiente para bloquear el flujo de sangre al cerebro. Las enfermedades vasculares cerebrales isquémicas, también pueden ser ocasionados por estenosis³ [5].

2.1.2. Enfermedad vascular cerebral hemorrágica

En un cerebro saludable, las neuronas no entran en contacto directo con la sangre. El oxígeno vital y los nutrientes que las neuronas necesitan de la sangre llegan a las neuronas a través de paredes delgadas de los capilares cerebrales. Las glías⁴ forman una barrera sanguínea-cerebral o hemoencefálica⁵.

³Estrechamiento de una arteria.

⁴Celdas del sistema nervioso que soportan y protegen a las neuronas

⁵Trama compleja que rodea a los vasos sanguíneos y capilares y controla qué elementos de la sangre pueden pasar a través a las neuronas.

Cuando se rompe una arteria en el cerebro, la sangre pasa al tejido circundante y perturba no sólo el suministro de sangre sino el equilibrio químico delicado que las neuronas requieren para funcionar. A este tipo de accidente cerebrovascular se le llama enfermedad vascular cerebral *hemorrágica* (e. v. c. h.) *Figura 4.1a*. Estos accidentes hemorrágicos representan aproximadamente un 20 por ciento de todos los ataques cerebrovasculares [56]. La hemorragia ocurre de varias formas. Una causa común es una aneurisma⁶ sangrante. Con el tiempo, estos lugares débiles se dilatan o se hinchan en forma de globo bajo una presión arterial elevada. Las paredes delgadas de estas aneurismas en forma de globo pueden romperse y derramar sangre en el espacio que rodea a las neuronas.

La hemorragia también ocurre cuando las paredes arteriales se rompen. Las paredes arteriales incrustadas con placa⁷ pierden con el tiempo su elasticidad y se tornan quebradizas y delgadas, propensas a romperse [5].

2.2. Secuelas de la enfermedad vascular cerebral

La *hemiplejia* es una de las secuelas mas importantes de la enfermedad vascular cerebral que existen en la actualidad. La hemiplejia, es consecuencia en la mayoría de las ocasiones de la e. v. c.

La hemiplejia es un trastorno motor, en el que la mitad vertical del cuerpo está paralizada; por ejemplo, un brazo y su pierna correspondiente, así como el lado contrario del rostro tienen deficiencia de movilidad voluntaria. La hemiplejia es provocada por la muerte de neuronas encargadas de la motricidad del cuerpo. La muerte de las neuronas es normalmente ocasionada por una e. v. c. Dependiendo de las neuronas afectadas la e. v. c., puede afectar otras funciones del cuerpo

⁶Lugar débil o delgado en una pared arterial.

⁷Mezcla de sustancias grasas, incluyendo el colesterol y otros lípidos.

humano, como la autonomía de los esfínteres o cualquier otra músculo del cuerpo [56].

La Organización Mundial de la Salud (O. M. S.), reporta 200 casos nuevos de personas con e. v. c. por cada 100 mil habitantes por año. Cada año 200,000 personas en México sobreviven a un derrame cerebral y cerca de 700,000 en Estados Unidos [56]. Aproximadamente el 80 % de ellos sufren secuelas de e. v. c., es decir, pierden la habilidad del movimiento. Aproximadamente un 10 % no logra beneficio alguno de ningún tratamiento. Un 10 % de los enfermos consiguen una recuperación espontánea completa de 8 a 12 semanas, mientras que el 80 % de los pacientes con secuelas de e. v. c., puede beneficiarse de una rehabilitación específica [41].

Los costos de la e. v. c. son altos, los gastos directos están en relación con la prevención, el diagnóstico, el tratamiento y la rehabilitación. No existen en nuestro país datos claros en relación con el tema, pero en E. U. A., por un poco más de 3 millones de pacientes que sobreviven a un ictus⁸ (*stroke*), se requieren de unos 16.8 billones de dólares, distribuidos en gastos de hospital, enfermería, servicios médicos, medicamentos y falta de productividad [17].

Los pacientes con e. v. c. presentan inicialmente flacidez, la cual evoluciona hacia un aumento de espasticidad y posteriormente recuperan su función motora voluntaria. Sin embargo, este patrón de recuperación puede interrumpir su evolución en cualquier fase, por lo cual no se puede considerar que la espasticidad anuncia el regreso de la función motora. Los pacientes que se encuentran con recuperación obtendrán el 90 % en los primeros 3 meses posteriores al ataque [5], mientras que los pacientes que no presentan una adecuada recuperación cursan con cuadros de hemiplejía caracterizados por un síndrome de neuronas motoras

⁸Cuadro morboso que se presenta de un modo súbito y violento, como producido por un golpe

superiores dañadas, el cual propicia déficit neurológico que afecta en gran medida las actividades de su vida diaria. Un factor importante de estos síndromes es, sin duda, la espasticidad, la cual genera una gran disfunción en el movimiento de los pacientes provocando, en muchas ocasiones, posiciones viciosas y contracturas (limitación en la movilidad, no reductible, debido al acortamiento muscular o capsular) de algunos músculos tendones, ligamentos y articulaciones; además, la espasticidad complica en forma muy importante la rehabilitación de los pacientes impidiendo que ésta se lleve a cabo de manera más eficaz y rápida. Sólo una pequeña minoría (aproximadamente el 5 %) de estos pacientes recobra funcionalidad del miembro superior paralizado. Después de los 3 meses los prospectos de recuperación son muy pocos [3]

2.3. Terapia a pacientes hemipléjicos

El primer objetivo dentro la terapia, es la rehabilitación a corto plazo del paciente hemipléjico durante la etapa temprana y aguda, esto es, hacerlo más independiente en las actividades primarias de la vida diaria, como asearse, alimentarse y trasladarse. El tratamiento pone énfasis en el lado sano que ayudará a compensar al lado afectado. Este tipo de rehabilitación precoz lleva, lamentablemente, al desarrollo de espasticidad. Esto es producto de las reacciones del esfuerzo requerido, por la utilización solo del lado sano, pérdida del equilibrio y el miedo a las caídas [15].

La principal etapa de la terapia es mejorar el tono y la coordinación, obteniendo reacciones normales del lado afectado, en respuesta a la movilización. El objetivo de la terapia debe de inhibir los patrones de movimiento anormales del paciente y desarrollar paralelamente los patrones normales.

En una etapa posterior, el tratamiento se centra en la rehabilitación funcional

del individuo, esto es, ayudarlo a desarrollar técnicas y estrategias para llevar a cabo las actividades diarias con mayor independencia. Esta rehabilitación es gradual, partiendo por cosas simples como tomar una cuchara hasta realizar su aseo personal.

Las técnicas de rehabilitación empleadas en la hemiplejía dependen de la etapa de recuperación que haya alcanzado el paciente o en que etapa se haya detenido la mejoría. Es claro que para saber en que etapa se encuentra el paciente, se debe de contar con un método para su valoración. Las etapas de la rehabilitación son: etapa flácida inicial, etapa de espasticidad y etapa de recuperación relativa [15].

2.3.1. Etapa flácida inicial

La etapa flácida inicial se descubre poco después del inicio de la hemiplejía durando de días a meses. El paciente no puede mover el lado afectado por la pérdida de sus patrones anteriores de movimiento y los patrones del lado sano son inadecuados para compensar su pérdida.

En esta etapa no existe ninguna restricción del rango articular a los movimientos pasivos del lado afectado lo que ayuda a la rehabilitación.

2.3.2. Etapa de espasticidad

El desarrollo de la espasticidad es gradual y comienza ya en la etapa flácida inicial, por lo tanto el tratamiento durante esta etapa se superpone con el anterior. La espasticidad se desarrolla lentamente con predilección por los músculos flexores de la extremidad superior y los extensores del miembro inferior. Se debe evaluar inicialmente el grado de espasticidad mediante examen de la resistencia al estiramiento pasivo de músculos y extremidades.

2.3.3. Etapa de rehabilitación definitiva

Los pacientes que alcanzan esta etapa son en su gran mayoría los que no estaban gravemente afectados al inicio y han logrado una buena recuperación, o han andado bien con el tratamiento. El paciente ahora camina sin ayuda externa, puede utilizar su brazo afectado y sostener un objeto en la mano que le sea colocado. La espasticidad es ligera y no impide el movimiento, aunque aumenta transitoriamente al realizar un esfuerzo.

2.4. Trabajos previos

La principal causa por la cual las personas hemipléjicas no se rehabilitan, es por la falta de terapia, ya sea por que no la hayan tomado nunca o por que se haya truncado una vez iniciada. Las causas pueden ser muchas, pero las principales son: la falta de recursos económicos, la dependencia a trasladarse, y en algunos casos la ausencia de terapeutas cercanos [15], es decir, la persona hemipléjica no obtiene terapia si no puede llegar a un centro de rehabilitación, sea cual fuera.

Debido a esto, se han buscado alternativas que faciliten y mejoren la rehabilitación. La mayoría de estos trabajos se basan en asistentes robóticos que eliminan en gran parte el trabajo del terapeuta. Algunos ejemplos son: *Massachusetts Institute of Technology (MIT) - MANUS* [29], *Mirror Image Motion Enabler (MIME)* [8], *Assited Rehabilitation and Measurement (ARM) Guide* [39], *Bi-Manu-Track* [22] y *GENTLE/S* [12]. Estos proyectos proveen una rehabilitación menos asistida por el especialista, pero siguen siendo costosos y su ubicación sigue siendo el centro de rehabilitación.

Pero si el problema es poder llegar a un centro de rehabilitación, es necesario buscar alternativas que lleven la terapia a donde la persona hemipléjica se encuen-



Figura 2.2: Prototipo de “Java Therapy”, tele - rehabilitador de regeneración de la fuerza [38].

tre. Algunos de estos ejemplos son: *Java therapy*: Sistema que utiliza un brazo robótico para la rehabilitación, con comunicación vía - *Web* [38], *T-Wrex (Therapy Wilmington Robotic Exoskeleton)*: Sistema robótico pasivo de entrenamiento de bajo costo [23] y *Markerless video arm tracking*: Sistema de visión libre de marcas, para la rehabilitación [46]. A continuación se describirán estos tres ejemplos, los cuales se encuentran directamente relacionados con esta tesis.

2.4.1. *Java Therapy*

La lesión del cerebro como resultado de una enfermedad vascular cerebral, es una de las causas principales de la discapacidad en los Estados Unidos [38]. Actualmente, por desgracia, para los personas con alguna lesión en el cerebro, existe poca tecnología terapéutica que puedan practicar desde sus hogares y que al mismo tiempo se les pueda monitorear.

Java therapy, Figura 2.2, presenta una alternativa para resolver este problema. *Java Therapy*, es un sistema robótico barato de tele - rehabilitación para la terapia del brazo y de la mano después de alguna lesión en el cerebro que haya producido hemiplejía.

“*Java Therapy*”, es un sitio *Web* con un conjunto de bibliotecas de *actividades*

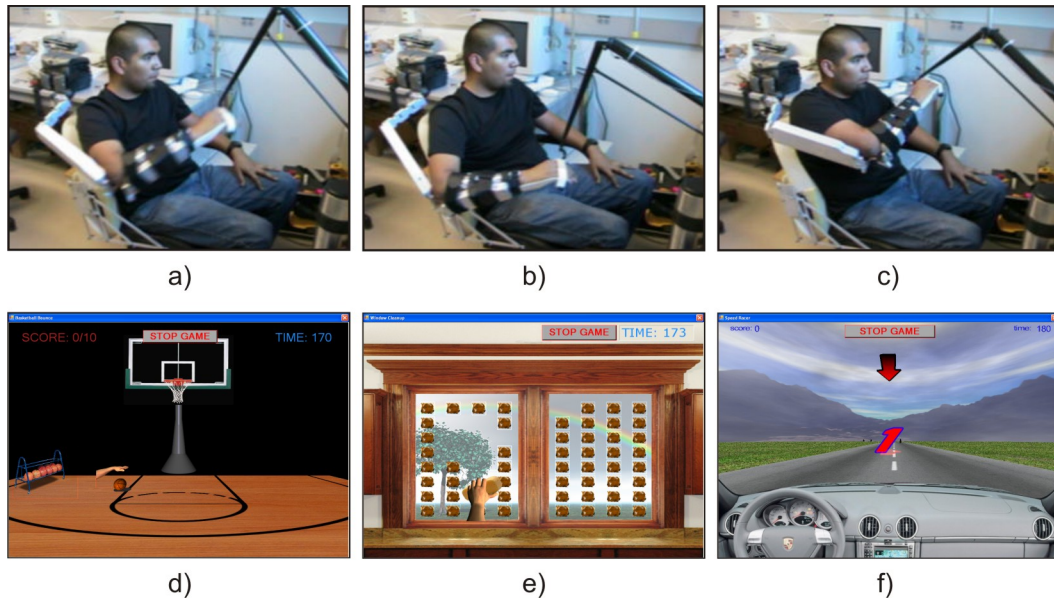


Figura 2.3: *T-Wrex*. a), b) y c) Utilización del brazo robótico, d) Juego de basket ball, e) Juego para limpiar las ventanas, f) Juego de manejo [23].

terapéuticas y evaluaciones de las mismas. Las *actividades terapéuticas* se pueden realizar con un “*Joystick*”, una palanca terapéutica comercial, que se utiliza para la regeneración de la fuerza. El *Joystick*, no sólo es *pasivo*, sino también *activo*, es decir, no sólo puede seguir el movimiento del brazo, sino que también puede resistir físicamente en el movimiento mientras el usuario realiza la terapia. Los movimientos de rehabilitación solicitados por *Java Therapy*, pueden ser generados por patrones de rehabilitación terapéuticos, que se establecen por un experto en rehabilitación. El problema *Java therapy*, es su necesidad de conexión a internet, ya que algunos hogares no cuentan con banda ancha o múltiples líneas telefónicas.

2.4.2. *T-Wrex*

El *Hardware* de *T-Wrex*, es una ortosis⁹ (brazo robótico) con antigraedad para el brazo humano, diseñada para rehabilitar deficiencia de movimiento en el brazo a través de entrenamiento sin la presencia de un terapeuta. Es pasivo con 5 grados de libertad, provee un espacio en 3D y también obtiene la fuerza de presión de la mano. Como se vio en la *sección* 2.3, uno de los objetivos de la terapia, es la rehabilitación funcional del individuo, esto es, ayudarlo a desarrollar técnicas y estrategias para llevar a cabo las actividades diarias. Las actividades diarias pueden ser limpiar la mesa, sostener una taza etc. Los patrones de movimientos solicitados por el *software* de *T-Wrex*, son a menudo juegos que imitan estas actividades diarias (ver *Figura* 2.3). El *software* de *T-Wrex*, obtiene el movimiento del paciente por medio del brazo robótico, que es calibrado antes de comenzar con la terapia, para obtener la amplitud de movimiento y fuerza de presión del paciente. Además, el sistema proporciona la regeneración cuantitativa del funcionamiento del movimiento, permitiendo que los usuarios y sus cuidadores, determinen el progreso de la rehabilitación. *T-Wrex*, elimina el problema de no tener banda ancha para la conexión a internet, pero el precio del brazo robótico sigue siendo demasiado elevado.

2.4.3. *Gesture Therapy*

Gesture Therapy [46], es la continuación del trabajo de *T-Wrex*. *Gesture Therapy*, es necesidad de un *Joystick* o brazo robótico y sin la utilización de marcas. Las motivaciones son las mismas a *Java therapy* y *T-Wrex*, sólo que *Gesture Therapy*, busca crear un sistema de mucho menor costo al utilizar 2 *webcams* en lugar de

⁹Aparato ortopédico aplicado, se aplica externamente a una parte del cuerpo para corregir deformidad o para mejorar una función.

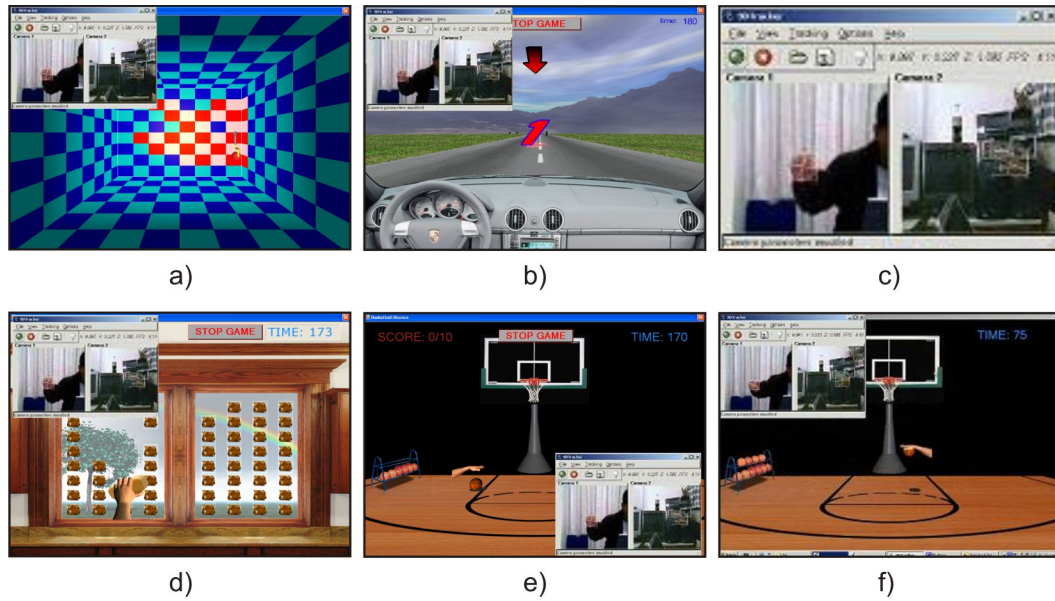


Figura 2.4: *Gesture Therapy*. Las imágenes a), b) y d) muestran el funcionamiento de *Gesture Therapy* en los juegos de pintar cuarto, de manejo y el de limpiar ventanas, respectivamente, c) Imagen que muestra la localización de la mano por *Gesture Therapy* y e) y f) Funcionamiento de *Gesture Therapy* en el juego de basket ball, cambiando de lugar la ventana que muestra el seguimiento de la mano. [46].

un *Joystick*, y sin la necesidad de estar conectado a internet.

El sistema de visión cuenta con un segmentador de piel, un seguidor del movimiento y un proceso de obtención de la posición espacial de la mano. El resultado es la obtención de las coordenadas x, y, z en el tiempo t_n . Estas coordenadas son enviadas a un sistema que lee la posición de la mano, que por medio de juegos, el paciente puede rehabilitar sus movimientos (ver *Figura 2.4*).

2.5. Escalas de la valoración de la discapacidad motriz

El mejorar la función motora y disminuir la espasticidad en el paciente con parálisis cerebral es una tarea desafiante, por lo que no se puede subestimar su complejidad. En la mayoría de los casos, esta condición que persistirá de por vida, se modifica, por lo que en las diferentes etapas de su desarrollo, se estarán buscando opciones de tratamiento de acuerdo a las necesidades, edad y expectativas del paciente. Es primordial que el profesional de la salud haga su mejor esfuerzo, para efectuar una evaluación adecuada.

Los distintos métodos de rehabilitación se han enfrentado al desafío de presentar evidencia de su efectividad; desafortunadamente, los principios subyacentes a la mayoría de los métodos de rehabilitación física para el tratamiento de la parálisis cerebral adolecen de pruebas científicas irrefutables de su utilidad. A pesar de la falta de pruebas suficientes de su eficacia, los programas de terapia física generalmente se consideran útiles en la rehabilitación. En medicina clínica, la metodología clásica para definir la efectividad de un procedimiento es el estudio controlado y aleatorio.

Para determinar la utilidad de un procedimiento terapéutico, ya sean terapias

de rehabilitación física o psicológica, la farmacoterapia oral, cirugía ortopédica y neurocirugía, sobre la función motora y la espasticidad en pacientes con parálisis cerebral; se requiere, por un lado, de escalas de medición objetivas y, por otro, escalas apropiadas de acuerdo con lo que deseamos medir, como: impedimento, dishabilidad, discapacidad, calidad de vida, etc. [9].

El desarrollo de mediciones confiables y válidas ha sido particularmente difícil en el campo de la rehabilitación, virtualmente todas las mediciones usadas para determinar la función y la mayoría de las escalas que cuantifican movimiento y espasticidad son ordinales y por lo tanto no paramétricas [32].

2.5.1. Escala de Fugl - Meyer

La escala *Fugl-Meyer*, evalúa las extremidades en cuanto a la actividad reflejada, sinergia flexora, sinergia extensora, colocación de la mano a la región lumbar, abducción pura de hombro de 0 grados a 90 grados, movimientos de muñeca y mano, coordinación - velocidad, movimiento de rodilla y tobillo. La calificación es de 88 puntos en una persona sin alteración [11]. La forma de calificar es valorando la movilidad así como el dolor de las articulaciones del brazo y de la pierna. A la tabla de movilidad se le asigna un puntaje de 0 si no hay movilidad, 1 movilidad disminuida y 2 movilidad casi normal. A la tabla de dolor se le asigna 0 si hay dolor en todo el movimiento, 1 si solo hay poco dolor al final del movimiento, y 2 si no hay dolor. La escala *Fugl-Meyer*, consta de 22 valoraciones. La escala *Fugl-Meyer* agrupa las escalas dependiendo del grado de discapacidad; menor a 40 puntos discapacidad severa, 40 a 64 puntos discapacidad, 65 a 75 discapacidad moderada, 76 a 87 ligera discapacidad.

2.5.2. Índice de motricidad

El *índice de motricidad* es una evaluación simple de la función motora en los pacientes con hemiplejía, toma en cuenta tres rangos de movimientos en el miembro torácico que son: la abducción del hombro, flexión de codo y la prensión. En el miembro pélvico evalúa la flexión de la cadera, extensión de rodilla y de tobillo. Todos los rangos de movimientos se evalúan indiferentes a la gravedad y contra resistencia cuando sea posible, otorgándole diversas calificaciones [11]. Los valores se basan en la observación. La forma de valorar a un paciente utilizando la escala de *índice de motricidad* para la prensión es el siguiente: 0 sin movimiento, 11 inicio de prensión (cualquier movimiento de pulgar o dedos), 19 agarra el cubo, pero es incapaz de mantenerlo contra gravedad, 22 agarra el cubo y lo mantiene en el aire, pero se le desaloja fácilmente, 26 agarra el cubo con fuerza pero más débil que el lado sano y 33 prensión normal. Para la movilidad de flexión y abducción: 0 sin movimiento, 9 contracción palpable en el músculo pero sin movimiento, 14 se aprecia movimiento pero no del arco completo ni contra gravedad, 19 movimiento arco completo contra gravedad, no contra resistencia, 25 movimiento contra resistencia pero más débil que el otro lado y 33 fuerza normal. Dando un total máximo para una persona sana de 100.

2.5.3. Motor Activity Log

La escala *Motor Activity Log - 14 (MAL)* para extremidades superiores, es una entrevista estructurada que obtiene información acerca de 14 actividades de la vida diaria (AVD). A los pacientes se les pregunta que tan bien (escala de calidad de movimiento) y que tanto (escala de cantidad de uso) utilizan su brazo deteriorado para lograr cada AVD, las cuales se califican en un rango de 0 a 5. 0 para la menor calificación y 5 para la mejor calificación. El rango puede ir de 0 para una total

discapacidad a 140 puntos para una persona sana [51].

2.6. Resumen

La enfermedad vascular cerebral afecta a millones de personas cada año en todo el mundo. Para las personas que han sufrido la enfermedad vascular cerebral, la hemiplejia es una de sus más “dolorosas” secuelas. Uno de los efectos de la hemiplejia es la incapacidad de tener una vida normal, ya que en la mayoría de los casos, las personas se vuelven dependientes. La hemiplejia es devastadora en países subdesarrollados o en personas de escasos recursos económicos, ya que para contrarrestar los efectos de la hemiplejia se necesita mucho tiempo de terapia y es muy costosa.

Para resolver este problema, actualmente existen algunas propuestas para ofrecer más opciones de rehabilitación, como *Java Therapy*, *T-Wrex* y *Markerless video arm tracking*, en donde el paciente puede realizar su terapia en cualquier lugar, sin la necesidad de acudir a un centro de rehabilitación.

En estas propuestas, no es necesario que un terapeuta de la rehabilitación motriz se encuentre presente. Por tal motivo, se necesita de una escala *automática* que mida el progreso de la rehabilitación del paciente. Actualmente existen varias escalas de la valoración motriz, como *Fugl Meyer*, *Motor Activity Log* e *índice de motricidad*, pero todas estas se necesita la presencia de un especialista, además de que estas escalas no son *paramétricas*.

La finalidad de este proyecto es generar una métrica *paramétrica*, que califique ciertos movimientos terapéuticos y proponerla como escala que valore la movilidad del paciente.

Capítulo 3

Análisis del movimiento humano

*“Todo lo que se mueve es movido por otro acto”
Aristóteles*

El estudio del *movimiento humano* es un tema tratado desde hace mucho tiempo y por múltiples culturas y civilizaciones. Obtener características dinámicas sobre las acciones de ciertas partes del cuerpo, es de suma importancia para los terapeutas y los investigadores de la bioinformática, ya que con dicha información se pueden crear tratamientos y/o tecnologías para la recuperación de los pacientes después de alguna afección o enfermedad.

Hoy en día, y con la ayuda de las nuevas tecnologías, este estudio es posible realizarlo con los medios audiovisuales y con computadoras como principal herramienta de procesamiento masivo de datos. La capacidad del hardware actual y las herramientas de software disponibles hacen perfectamente viable la visualización en tiempo real de objetos complejos, como evidentemente es la estructura del cuerpo humano.

Todos estos ingredientes científicos - tecnológicos, permiten modelar una persona como un conjunto de sólidos articulados (los huesos) con elementos deformables (los músculos), si es preciso, y visualizarlo sobre un dispositivo bidi-

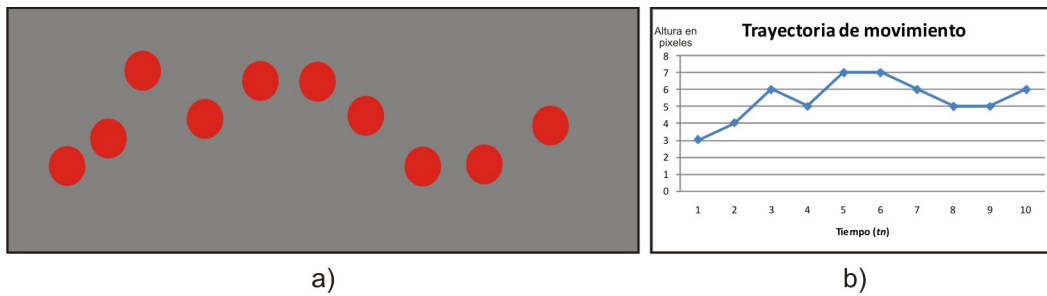


Figura 3.1: *Gráfica de la trayectoria del movimiento de la marca de color rojo, que puede representar una articulación del cuerpo.*

mensional como es el monitor o bien un entorno inmersivo tridimensional. Al mismo tiempo, la mecánica actual permite realizar un estudio biomecánico del sistema articular propuesto como si fuese un mecanismo físico convencional. El problema de todos estos procesos es la generación precisa de los modelos que permitan simular el comportamiento real de la persona en movimiento.

Por ello, una de las principales fuentes de información para el estudio del movimiento del cuerpo humano ha sido siempre la obtención directa de datos captados a partir del movimiento de un sujeto objeto de estudio. Esta recolección de datos se realiza mediante diferentes técnicas, ya sea a partir de informaciones proporcionadas por sensores electrónicos situados en puntos estratégicos del cuerpo o mediante la utilización de marcadores reflectantes cuyas posiciones son extraídas a partir del análisis digital de imágenes obtenidas estereoscópicamente, o sin marcas. Podemos clasificar el análisis del movimiento humano en tres categorías:

- Sistemas basados en sensores
- Sistemas basados en visión con marcas
- Sistemas basados en visión sin marcas

Una vez que se obtiene el movimiento humano, puede ser graficado en el tiempo, ver *Figura 3.1b*. El movimiento puede ser visto como una trayectoria que cambia en el tiempo. Al trabajar con trayectorias es fácil preguntarse si dos trayectorias son iguales o que tan semejantes son. Existen diferentes métodos para comparar trayectorias, algunos se basan en distancias y otros en características.

En la primera parte de este capítulo se presentan técnicas para detectar el movimiento humano utilizando sensores, marcas y sin marcas. En la segunda, se describen algunas métricas para comparar trayectorias: basadas en distancias y basadas en características.

3.1. Detección de movimiento basado en sensores

Un sensor es un dispositivo que detecta manifestaciones de cualidades o fenómenos físicos, como la energía, velocidad, aceleración, tamaño, cantidad, etc. Podemos decir también que es un dispositivo que aprovecha una de sus propiedades con el fin de adaptar la señal que mide para que la pueda interpretar otro elemento.

Los sistemas basados en sensores que se utilizan para la obtención del movimiento humano se pueden clasificar en:

- *Inerciales*: Acelerómetros, giroscopios.
- *Infrarrojos*: Leds, cámaras infrarrojas.
- *Mixtos*: Inerciales e infrarrojos.
- *Otros*: Magnéticos y acústicos.

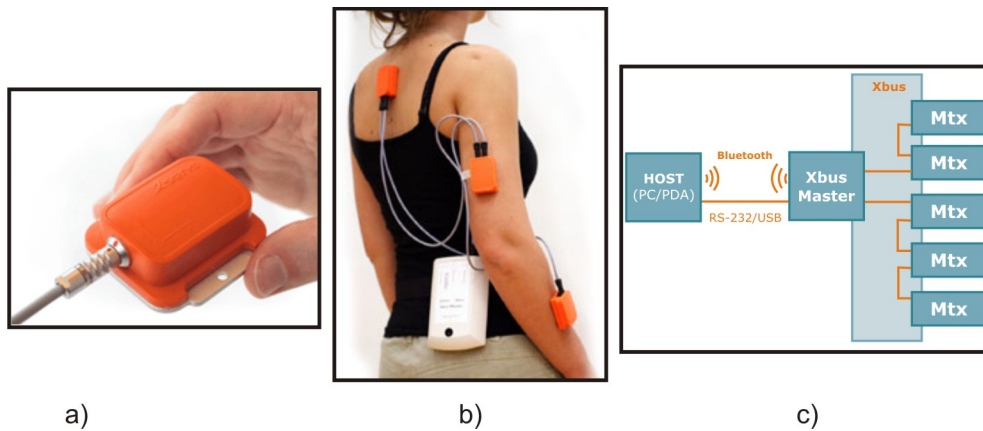


Figura 3.2: *Xsens Motion System*: a) Sensor *MTi*, b) Colocación en el cuerpo humano, c) Diagrama *MTx*, muestra las conexiones físicas del sistema, los sensores colocados a un bus, que a su vez, está conectado a la PC [58].

3.1.1. Inerciales

Los sensores *inerciales* son aquellos que utilizan la propiedad de los cuerpos de no modificar su estado de reposo o movimiento si no es por la acción de una fuerza. Los sensores inerciales pueden estar formados por giroscopios, acelerómetros o ambos. Se denomina *acelerómetro* a un instrumento destinado a medir aceleraciones, está contruido por un eje que cambia su dirección en sentido contrario a la aceleración. Los *giroscopios* son dispositivos mecánicos formados por un cuerpo que gira alrededor de su eje de simetría. Cuando se somete el giroscopio a un torque que tiende a cambiar la orientación de su eje de rotación, éste cambia de orientación en una dirección perpendicular a la dirección del movimiento.

Un ejemplo de un sistema inercial [59, 27] es el *Xsens* [58], de la compañía *Xsens Motion Tech*. El sensor que utiliza es un acelerómetro *MTi*, Figura 3.2a, que mide la aceleración y campo magnético para obtener medidas 3D. Combinado con el software *MT*, éste mecanismo provee información 3D en tiempo real en forma

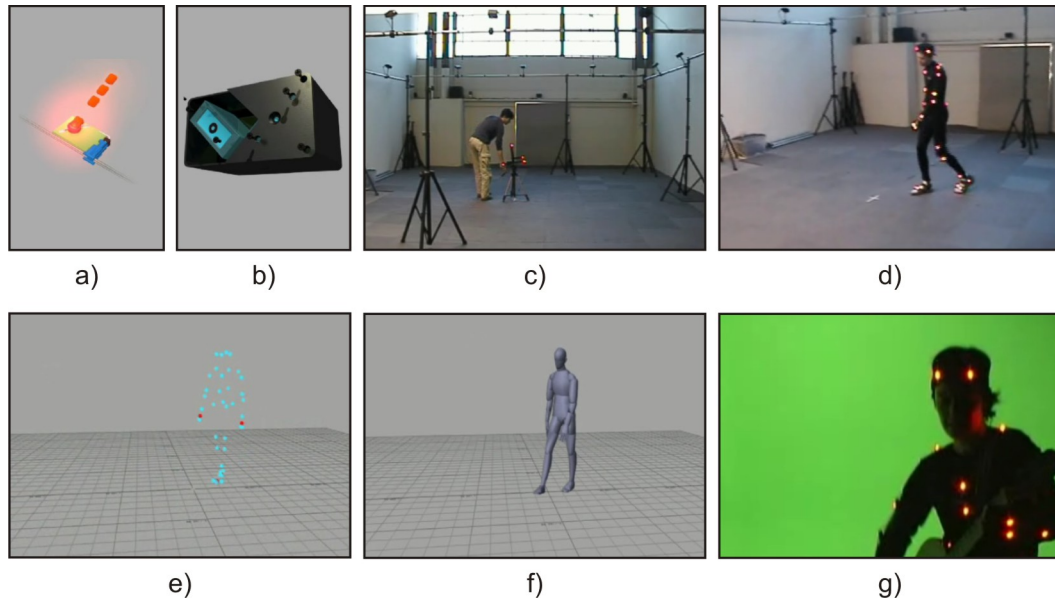


Figura 3.3: *Phase Space Motion Capture System*: a) *Led infrarrojo*, b) *Sensor infrarrojo*, c) *Calibración de las cámaras infrarrojas*, d) y g) *Sensores colocados en el cuerpo humano*, e) *Obtención de la posición 3D*, f) *Renderización*

de ángulos de *Euler*¹, a una frecuencia de 512 Hz. Debido a su tamaño compacto y a su alto rendimiento, el dispositivo MTi ha sido fácilmente integrado en el campo de la bioinformática, robótica, animación, realidad virtual, etc. Sin embargo, el dispositivo (con seis unidades MTi, *Figura 3.2b*) y el software de seguimiento MT que tiene un costo muy elevado, aproximadamente \$135,000 pesos.

3.1.2. Infrarrojos

Los sistemas que utilizan *sensores infrarrojos* están compuestos por un conjunto de *leds* infrarrojos colocados en diferentes partes del cuerpo, que por lo general

¹Constituyen un conjunto de tres coordenadas angulares que sirven para especificar la orientación de un sistema de referencia de ejes ortogonales, normalmente móvil, respecto a otro sistema de referencia de ejes ortogonales normalmente fijos.

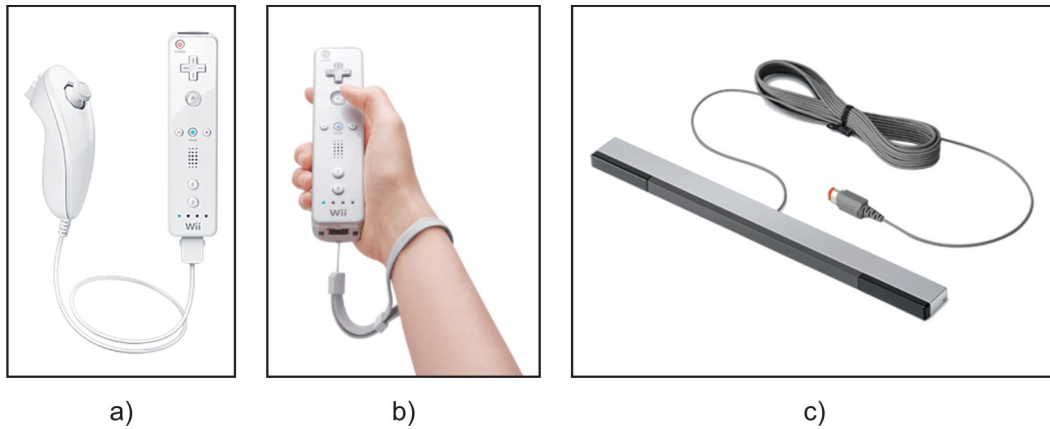


Figura 3.4: *Wii mote*: a) *Sensor inercial*, b) *Sensor infrarrojo*, c) *Leds infrarrojos*[24]

son las articulaciones del cuerpo. La luz infrarroja emitida por los *leds*, es captada por unas cámaras, que se encuentran colocadas de tal forma que cubra un área de 360 grados de un espacio controlado.

Un ejemplo de sistemas infrarrojos es el PhaseSpace [54]. Su funcionamiento es a partir de señales de luz infrarroja emitidas por leds, *Figura 3.3a*, colocados en las articulaciones del cuerpo humano, *Figuras 3.3d y 3.3g* . Estas son recibidas por unas cámaras colocadas alrededor de la persona, *Figuras 3.3b y 3.3c*, 8 en total. Puede leer hasta 128 leds infrarrojos. Por su gran cantidad de cámaras no tiene problemas de oclusión, el problema al igual de *Xsens*, es su precio industrial, aproximadamente \$250,000 pesos, que es alto para proyectos pequeños.

3.1.3. Mixtos

Los sensores infrarrojos tienen la desventaja de no poder medir la aceleración y cambios de dirección por si solos. Y los inerciales pueden fallar al no tener cambios bruscos de velocidad y de dirección y no detectar la posición en 3D por si solos. Al juntar las dos tecnologías se obtiene un sistema mucho más robusto.

Un ejemplo de sistemas mixtos es el utilizado por la compañía Nintendo en

su sistema *wiimote*[24], *Figura 3.4*. El control contiene dos sensores inerciales un acelerómetro, para detectar la fuerza y velocidad de movimiento y un giroscopio que detecta los giros de movimiento, *Figura 3.4a*. El sistema también cuenta con un sensor infrarrojo, éste percibe la luz infrarroja de 8 leds colocados en un panel arriba del televisor, 4 en cada extremo, *Figuras 3.4b y 3.4c*. Dándole así la información de posición tridimensional. Combina estos tres tipos de señales: luz infrarroja, acelerómetro y giroscopio para obtener la información de movimiento, posición y aceleración. Sus problemas son las distancias requeridas para la utilización del control. A distancias muy alejadas, mayores a 10 metros, la posición muestra errores. Algunas señales de teléfono y el zoom infrarrojo de cámaras pueden dar mal funcionamiento al sistema, ya que se interfiere con las señales infrarrojas del sistema.

3.1.4. Otros

Existen otros tipos de sensores que no entran en ninguna clasificación anterior, como son los acústicos y los magnéticos. Los sensores acústicos son dispositivos que utilizan una señal ultrasónica de baja frecuencia producida por un transmisor fijo para determinar la posición y orientación en tiempo real de un receptor móvil. Su forma de funcionar es la siguiente: Una fuente triangular (3 altavoces) produce sonidos ultrasónicos, el receptor (con tres micrófonos), mide la distancia de cada altavoz a todos los micrófonos (9 medidas en total). La distancia se calcula multiplicando el “tiempo de vuelo” por la velocidad del sonido en el aire. Estos sensores son los menos usados, ya que captan mucho ruido.

Los sensores magnéticos son dispositivos que utilizan un campo magnético producido por un transmisor fijo para determinar la posición en tiempo real de un receptor móvil. Su forma de funcionar es la siguiente: Una fuente fija produce

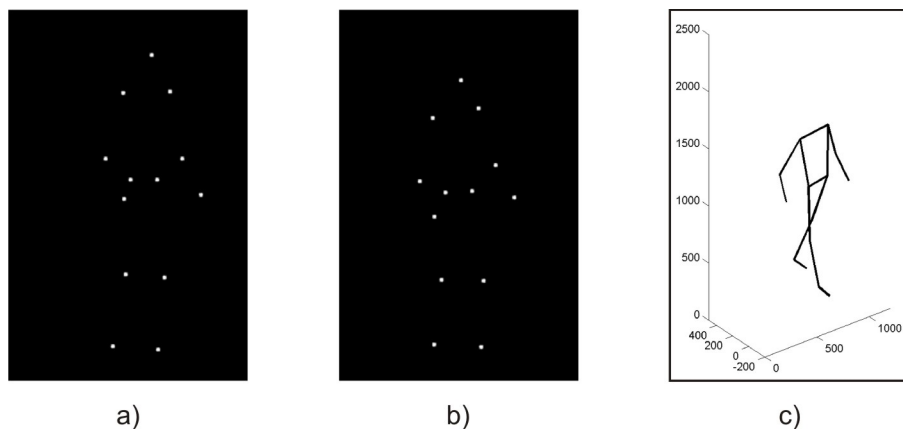


Figura 3.5: *Johansson: Moving Light Display*, a) y b) Muestran las marcas de luz detectadas por el sistema, y c) Renderización por el sistema del cuerpo humano.

campos magnéticos de baja frecuencia para medir la posición. El tamaño de la fuente está relacionado con el volumen de trabajo. El receptor se une al objeto y tiene tres antenas perpendiculares. La distancia se deduce de los voltajes inducidos en las antenas. Se necesita calibración. Los sensores son pequeños y ligeros y no tienen problemas de oclusión. Un ejemplo de sistemas que utilizan sensores magnéticos es el creado por *Ascension Technology Corporation*. Este sistema se le conoce como Parvada de pájaros (*flock of birds*) [13]. El problema de este tipo de sistemas es el costo, alrededor de \$30,000 pesos.

3.2. Basados en visión con marcas

En 1975 el psicólogo Gunnar Johansson realizó un trabajo relacionado con *motion perception*. La *percepción del movimiento*, es una rama de la ciencia actual, que intenta resolver el problema de obtener conocimiento del movimiento. Johansson, en su trabajo de investigación: *Moving Light Display* (MLD) [26], grabó videos en un ambiente oscuro. Lo que grabó fueron secuencias de movimiento del cuerpo hu-

mano, al cual le colocó luces en sus articulaciones. Las secuencias que grabó fueron actividades humanas como, caminar, saltar, bailar, correr, etc. Los resultados que obtuvo fueron muy interesantes, pues con solo el análisis de las luces, el podía determinar que actividad estaba realizando la persona y no sólo eso, sino que podía determinar su edad y su género.

Los resultados de Johanson abrieron una puerta a la investigación de Visión Computacional relacionada con *motion perception*. La mayoría de los trabajos de reconocimiento de movimiento humano en el área de visión computacional se basan en el trabajo de Johansson. Son los que actualmente se conocen como *sistemas de visión con marcas*. El utilizar marcas para la detección del movimiento humano, reduce muchos de los problemas de la visión computacional, como es la segmentación. Algunos ejemplos de sistemas que utilizan marcas para la detección del movimiento son: Mocap [55], Coda [44] y Vicon [45]. Ya que estos tres sistemas tienen un funcionamiento muy similar, en la siguiente sección sólo se analizará uno de ellos, Mocap.

3.2.1. *Motion Capture System (Mocap System)*

Optical Motion Capture [55] es un método eficiente que permite seguimiento correcto de movimientos de cualquier tipo. Los sistemas de *Qualisys* utilizan las cámaras fotográficas digitales de alta velocidad, para capturar el movimiento exacto de un objeto, con los *marcadores* pasivos, activos² o mixtos. La tecnología es exacta y robusta y entrega datos de alta calidad al observador en tiempo real. El software hace fácil realizar *cálculos básicos* del movimiento, tales como velocidad, aceleración, rotaciones y ángulos. La captura óptica del movimiento permite

²Al contrario de los marcadores pasivos, los marcadores activos, emiten señales, y son más usados como sensores.

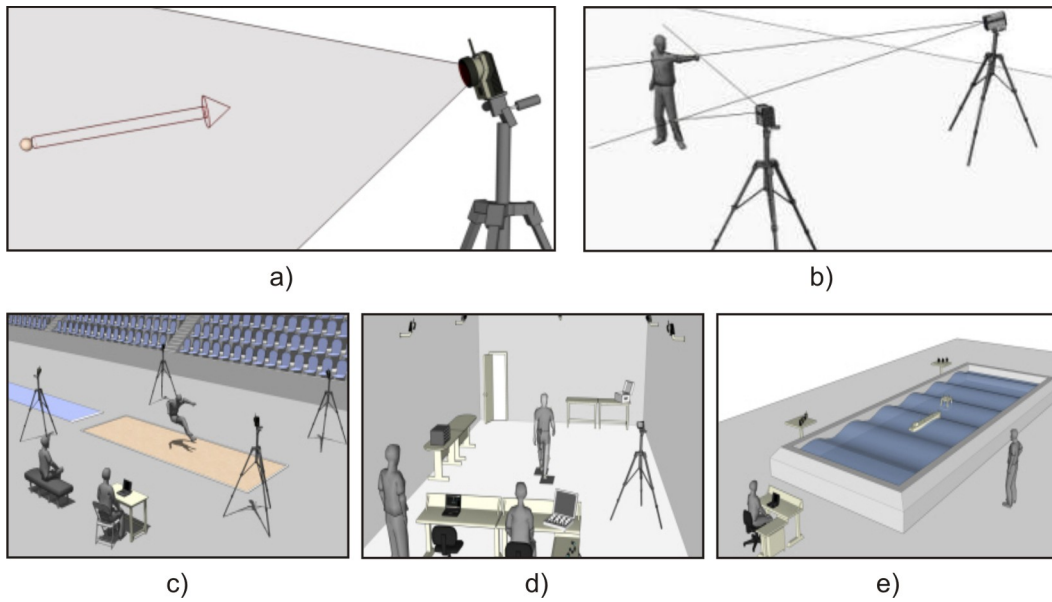


Figura 3.6: *Optical Motion Capture, Qualisys [55]. a)Reconstrucción 2D, b)Reconstrucción 3D, c) Seguimiento del movimiento humano en deportes, d) Calibración, e) Seguimiento de movimiento de un barco, en un ambiente simulado.*

obtener los movimientos que serían difíciles de medir de otras maneras. El *Mocap*, realiza el seguimiento de objetos de la siguiente manera:

- *Reconstrucción 2D*. El sistema modular de medición consiste en una o más cámaras fotográficas. Cada una emite un haz de luz infrarroja. Los marcadores reflexivos pequeños se colocan en el objeto o la persona que desea seguir. La cámara emite luz infrarroja y los marcadores la reflejan de nuevo a la cámara fotográfica. La cámara mide la posición en 2D del objeto (ver *Figura 3.6a*).
- *Reconstrucción 3d*. El software combina la información 2D de las diferentes cámaras para calcular la posición 3D del objeto (ver *Figura 3.6b*).
- *Calibración*. Para obtener la posición en 3D, el sistema necesita ser calibrado. Simplemente se mueve una varita con un sensor alrededor del espacio de seguimiento, mientras que un objeto inmóvil sirve de referencia en el espacio. La calibración se acaba aproximadamente después de 10-20 segundos (ver *Figura 3.6d*).

3.2.2. *Webcams* - Visión con marcas

Este modelo es el utilizado en este proyecto de tesis, el cual será analizado detalladamente en el capítulo 5. Se basa principalmente en seguir el movimiento humano utilizando el modelo de visión con marcas, el cual es tomado del trabajo de Cañas y González [19]. Cañas y González realizan seguimiento de un objeto en una secuencia de video utilizando 2 *webcams*, para obtener su posición en tres coordenadas (x, y, z) . Este modelo resulta ser muy económico, pues el costo es de alrededor de \$1,000 pesos (el costo de las *webcams*). No es muy preciso, pues su

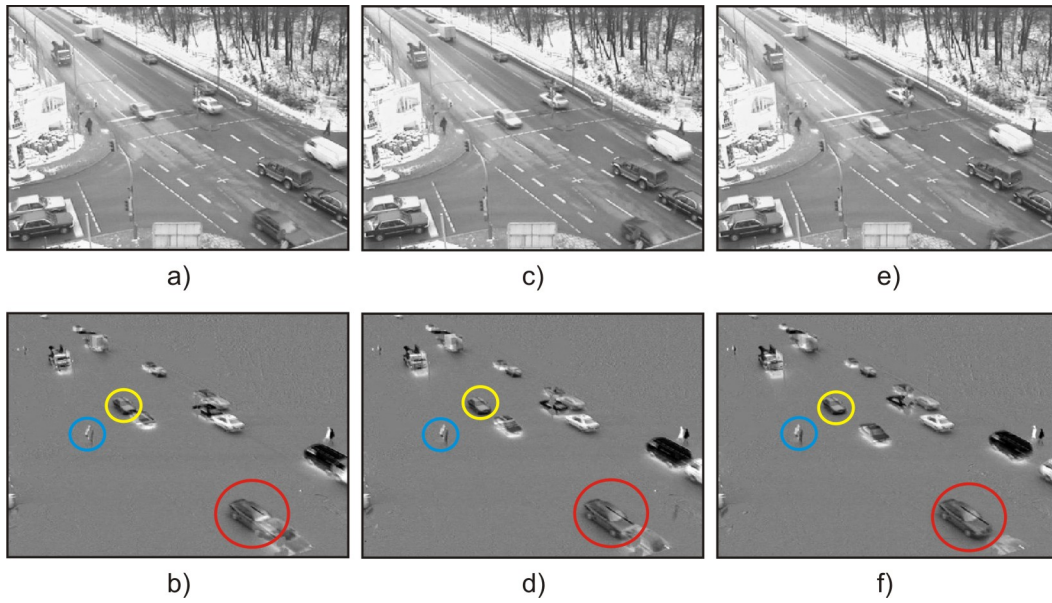


Figura 3.7: *Seguimiento de objetos utilizando substracción de imágenes, a), c) y e) Imágenes de la secuencia de video, b), d) y f) Resultados de la substracción de imágenes que detectan objetos en movimiento.*

precisión tiene un error de -5 cm. a 5 cm. (dependiendo de la colocación de las cámaras y de las condiciones de iluminación).

3.3. Basados en visión sin marcas

Una técnica que busca superar las restricciones de seguir objetos utilizando marcas, como son el grado de movilidad y la oclusión de las marcas, son las técnicas de visión libres de marcas. La *Figura 3.7* muestra en ejemplo sencillo que obtiene las secuencias de movimiento de varios objetos en un video. La técnica utilizada en este ejemplo, se basa en la substracción de imágenes. La imagen actual,

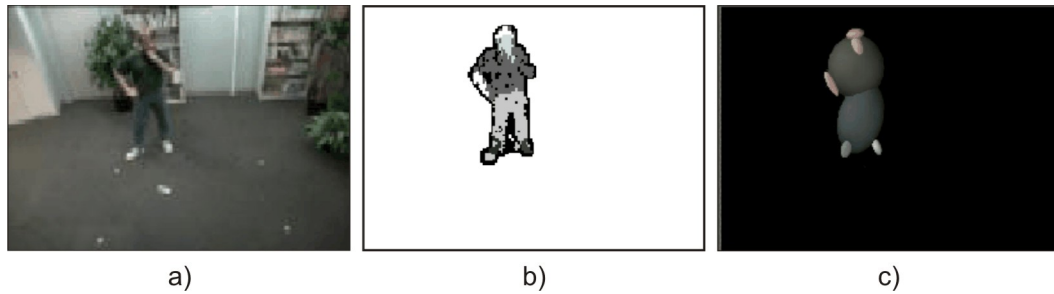


Figura 3.8: *Seguimiento del cuerpo humano*[53]. a) *Imagen original*, b) *Segmentación*, c) *Blobs*.

*Foreground*³, se resta de la imagen anterior o fondo, *Background*,⁴, obteniendo así, sólo los objetos que cambiaron de posición (ver Figuras 3.7b, 3.7d y 3.7f). Una vez detectado el objeto en movimiento, se pueden aplicar técnicas de detección de bordes y contornos para definir el objeto. Muchas técnicas en seguimiento del movimiento humano libre de marcas, se basan en detectar el cuerpo humano utilizando características del mismo.

Al no utilizar marcas, nos enfrentamos a varios problemas, ya que existen varias dificultades, incluyendo ambigüedades en la profundidad, deformidades en la apariencia, complejidad en los modelos cinemáticos y oclusiones [43, 42].

Normalmente se usa conocimiento del movimiento y forma del cuerpo humano para detectar el cuerpo humano y evitar los problemas de oclusión de ciertas partes del cuerpo. Por ejemplo, Wren [53] presenta un enfoque basado en regiones, considerando al cuerpo humano como un conjunto de *blobs*⁵, que pueden ser identificadas por información espacial y una distribución *Gaussiana* del color. En el

³En visión computacional, es el término utilizado para hacer mención del *primer* plano que se desea segmentar.

⁴En visión computacional, es el término utilizado para hacer mención del plano fijo que se desea eliminar al segmentar el objeto en movimiento.

⁵Cristopher Wren, en su trabajo[53] menciona que el cuerpo humano esta formado por regiones: La cabeza, tronco, piernas, brazos y pies, a los cuales el les llama *blobs*, manchas.

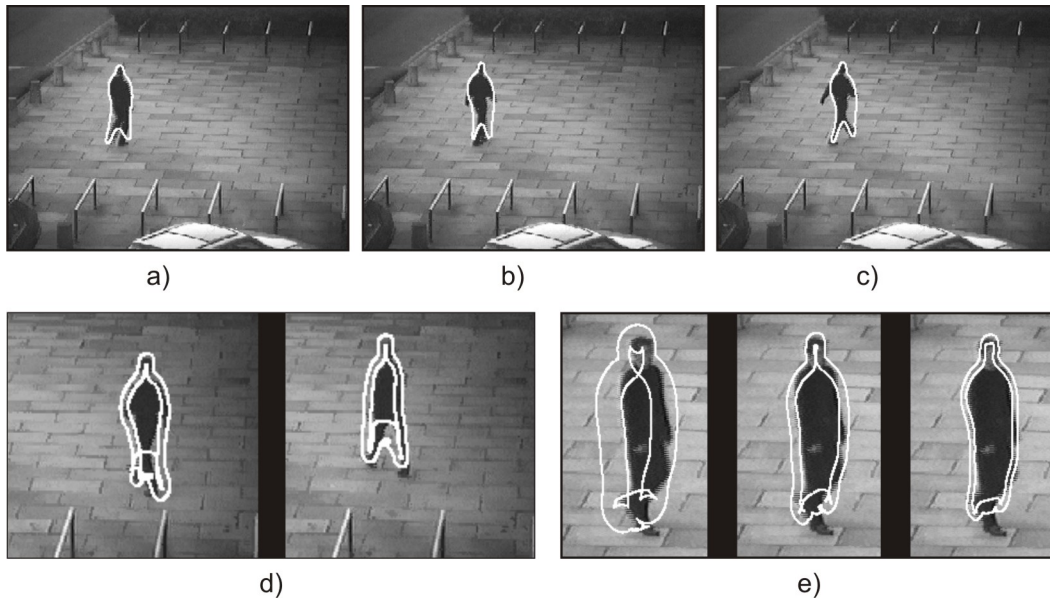


Figura 3.9: *Seguimiento del cuerpo humano*[4]. Las imágenes muestran la localización del cuerpo humano en movimiento en cada una de las imágenes.

proceso de inicialización, el foreground puede ser extraído del background previamente creado. Las Figuras 3.8a, 3.8b y 3.8c, muestran los resultados obtenidos por Wren, sin embargo, no se pueden identificar los puntos de unión en las extremidades.

Baumberg y Hogg [4] utilizaron los modelos de contornos activos (Active Shape Models-ASM) para el seguimiento de peatones (ver Figura 3.9) y utilizaron *B-Splines* para representar los diferentes contornos; las regiones del primer plano de la imagen son extraídas por medio de la substracción del fondo. Ellos aplicaron el filtro de Kalman para lograr las operaciones espacio-temporales. El sistema de Baumberg y Hogg genera automáticamente un modelo físico de la persona usando conjuntos de datos de entrenamiento. Los modelos del movimiento del cuerpo humano son limitados a simples contornos, donde las características descriptivas del movimiento no se pueden obtener fácilmente.

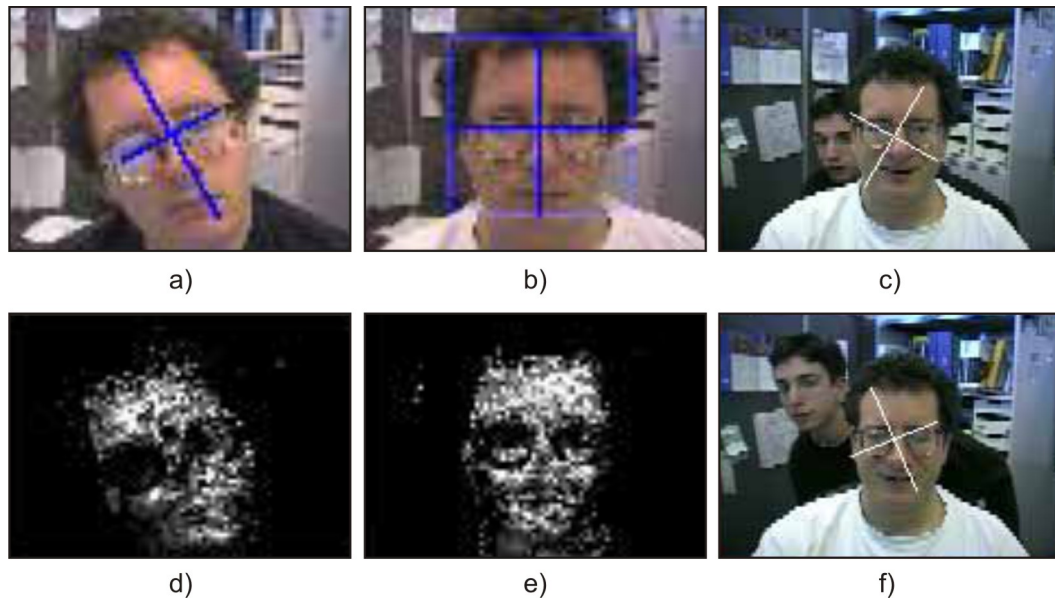


Figura 3.10: *Seguimiento del rostro, Camshift[7]. a) y b) Localización de la posición del rostro, d) y e) Obtención de la distribución de probabilidad de píxeles color piel en la imagen, c) y f) Localización del centro de masa del objeto color piel.*

Bradski[7] propone un algoritmo llamado *CamShift* orientado al seguimiento del rostro humano para su uso en interfaces humano-computadora. Utiliza una ventana de búsqueda en donde identifica la mayor densidad de puntos pertenecientes al rostro. La ventana de seguimiento se posiciona en el centro de masa del objeto y se modifican sus dimensiones si éste cambia de tamaño (ver *Figura 3.10*). Este algoritmo ofrece una forma eficiente de seguimiento en tiempo real, sin embargo, tiene ciertos problemas en la fase de segmentación.

3.4. Comparación de métodos de seguimiento humano

El *cuadro 3.1*, es una tabla que compara los diferentes sistemas mencionados en este capítulo. La tabla compara los sistemas y métodos por tecnología, precisión, ambiente, precio y dificultad:

- *Tecnología*: Determina el tipo de tecnología en la que fueron desarrollados (visión con marcas (VCM), visión con marcas infrarrojas (VCMI) o sensores).
- *Precisión*: Determina la precisión al detectar el movimiento o si le afecta la oclusión (alta, media o baja).
- *Ambiente*: Determina en que espacio o ambiente puede funcionar el sistema (*cerrado*: colocación fija del sistema - espacio controlado, *abierto*: colocación dinámica del sistema y *semiabierto*: facilidad de cambiar de espacio del sistema bajo previa configuración).
- *Precio*: Determina el costo de la adquisición del sistema (*bajo*: menor a \$1,000 pesos; *medio*: entre \$1,000 pesos y \$50,000 pesos; *alto*: entre \$50,000 pesos y \$100,000 pesos y *muy alto*: más de \$100,000 pesos).

Sistema	Tecnología	Precisión	Ambiente	Precio	Dificultad
Mocap	VCMi	Alta	Cerrado	M.Alto	Alta
Wiimote	Sensor	Media	Abierto	Alto	Baja
WebCam	VCM	Media	S. Abierto	Bajo	Baja
PhaseSpace	Sensor	Alta	Cerrado	M.Alto	Alta
Xsens	Sensor	Alta	Abierto	M.Alto	Alta
Flock of birds	Sensor	Alta	Cerrado	Alto	Alta

Tabla 3.1: Comparación de técnicas de seguimiento. VCM = Visión con marcas. VCMi = Visión con marcas infrarrojas.

- *Dificultad*: Determina la dificultad de instalación y uso del sistema (baja, media y alta).

En el presente trabajo se desea realizar un *seguimiento del movimiento humano* utilizando una técnica que nos permita cambiar de ambiente, ya que el sistema, tiene como meta en un futuro poder ser utilizado en los hogares de los pacientes en rehabilitación. Además se necesita que sea fácil de implementar. No se necesita que sea de precisión milimétrica. Y que sea muy económico. Por los requisitos deseados, como se puede apreciar en el *cuadro 3.1*, la técnica que más se adecúa a las necesidades del proyecto es la utilización de seguimiento visual con marcas del movimiento utilizando *webcams*.

3.5. Métricas para trayectorias

En visión computacional, cuando hablamos de seguir objetos, una *trayectoria* es una secuencia temporal de coordenadas representando la ruta de movimiento de un objeto sobre su tiempo de vida, es decir, el número de cuadros de video en que el objeto existe. Estas coordenadas corresponden a las marcas de posición

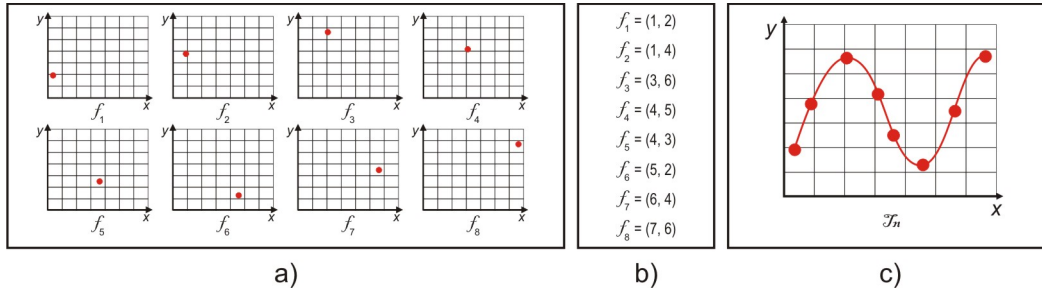


Figura 3.11: Trayectoria del movimiento de un objeto. a) Detección del centro de masa del objeto en los frames 1 al 8, b) Coordenadas cartesianas del centro de masa del objeto en los frames 1 al 8, c) Trayectoria T_n del objeto donde $n = 8$

del objeto en cuadros consecutivos (ver Figuras 3.11a y 3.11b). Por lo tanto, es posible ver la trayectoria como una colección de valores que puedan abstraer la forma de movimiento del objeto (ver Figura 3.11c). En notación, una trayectoria en 2D se puede escribir como [35]:

$$T = \{p_n\} = \{(x_1, y_1, t_1), (x_2, y_2, t_2), \dots, (x_N, y_N, t_N)\} \quad (3.1)$$

donde x_i y y_i son las coordenadas del centro de masa del objeto, t_i es el índice del tiempo, y N es la duración de la trayectoria.

Al obtener la ruta del movimiento de un objeto, se puede obtener mucha información de esta trayectoria, como por ejemplo:

- *Coordenadas cartesianas.* Los datos obtenidos por las marcas de posición, que generalmente indican el centro de masa del objeto, son las coordenadas cartesianas del objeto.
- *Coordenadas polares o esféricas.* Al obtener coordenadas cartesianas de varios objetos articulados, se pueden obtener ángulos de las líneas de intersección de los objetos, con los datos que se pueden obtener las coordenadas polares o esféricas.

- *Información contextual.* Información obtenida por la relación espacial del objeto que se quiere seguir, con otros: esta más lejos de..., esta más cerca de..., etc.

Al hacer el seguimiento de 2 o más objetos, y obtener sus trayectorias, es fácil preguntarse: ¿qué similitud hay entre ellas?. Para hacer una comparación entre trayectorias, es decir, entre el movimiento de dos o más objetos, se debe definir que tipo de información se va a comparar, si es sólo información de las distancias entre ellas o también sus características. Por lo tanto, existen métricas basadas en distancias y métricas basadas en características

3.5.1. Métricas basadas en distancias

Existen métricas sencillas basadas en técnicas estadísticas para comparar dos trayectorias de igual duración. La métrica más simple para calcular la distancia entre dos trayectorias, T_1 y T_2 , es la distancia media de las coordenadas, definida como [35]:

$$D_m(T_1, T_2) = \frac{1}{N} \sum_{n=1}^N d_n^2 \quad (3.2)$$

donde d_n^2 indica el desplazamiento entre las posiciones, y es calculado usando la distancia Cartesiana:

$$d_n^2 = [(x_n^1 - x_n^2)^2 - (y_n^1 - y_n^2)^2]^{\frac{1}{2}} \quad (3.3)$$

o usando $L - normal$:

$$d_n^L = [(x_n^1 - x_n^2)^L - (y_n^1 - y_n^2)^L]^{\frac{1}{L}} \quad (3.4)$$

donde x_j y y_j son las coordenadas del centro de masa del objeto que representan la posición del objeto j en el tiempo j de la trayectoria T_i . Se puede notar que la métrica de la distancia media hace tres suposiciones críticas:

- La duración de ambas trayectorias son iguales: $N^1 = N^2 = N$.
- Las coordenadas están sincronizadas: $t_n^1 = t_n^2$, esto es, las posiciones en ambas trayectorias son tomadas en el mismo instante de tiempo.
- La velocidad de captura es constante.

Por consiguiente, es evidente que la media de las distancias es muy sensitiva a las correspondencias parciales, y no toma en cuenta las distorsiones en el tiempo.

Para proveer mayor información descriptiva, se pueden incorporar estadísticas de segundo orden como la distancia basadas en la mediana, la varianza, el mínimo o el máximo [35]. La *varianza* está definida como:

$$D_v(T_1, T_2) = \frac{1}{N} \sum_{n=1}^N (d_n - D_m(T_1, T_2))^2 \quad (3.5)$$

donde d_n es el desplazamiento en la ecuación 3.3, y $D_m(T_1, T_2)$ es la media de la distancia en la ecuación 3.2.

Para calcular la *mediana*, los desplazamientos son ordenados respecto a sus magnitudes como $d_n \rightarrow d_m$, para que el valor del componente intermedio de la lista sea asignado:

$$D_{med} = \begin{cases} d_{\frac{N+1}{2}} & N \text{ impar} \\ \frac{1}{2}(d_{\frac{N}{2}} + d_{\frac{N+1}{2}}) & N \text{ par} \end{cases} \quad (3.6)$$

El *mínimo* y el *máximo* de dos trayectorias se definen como:

$$D_{min}(T_1, T_2) = \min\{d_n\}, \forall_n \in \{1 \cdots N\} \quad (3.7)$$

$$D_{max}(T_1, T_2) = \max\{d_n\}, \forall_n \in \{1 \cdots N\} \quad (3.8)$$

Si bien, estas medidas estadísticas proveen información extra, heredan el defecto de la métrica ordinaria basada en la media de la distancia $D_m(T_1, T_2)$ de la ecuación 3.2. Además, ninguna de las métricas es suficiente para hacer una evaluación precisa de la similitud [35]. Ya que se basan sólo en obtener la distancia

entre las trayectorias que cumplen los requisitos de tener la misma longitud, la misma velocidad de captura, etc. Además de que no pueden medir la diferencia de velocidad, aceleración y la dirección o cambios de dirección.

3.5.2. Métricas basadas en características

Además del sistema de coordenadas, podemos encontrar otras características en las trayectorias. A pesar de su simplicidad, la duración (lapso de vida) es una característica distintiva. Por ejemplo, la longitud de desplazamiento de una mosca al tener un comportamiento de alimentación, puede ser la misma al aparearse, más no sucede así con la longitud de su trayectoria. El tamaño total de una trayectoria es definido como:

$$L_T = \sum_{n=2}^N |p_n - p_{n-1}|, \quad (3.9)$$

donde p_i es un punto de la trayectoria T . Esto es diferente del desplazamiento total del objeto:

$$D_T = |p_N - p_1| \quad (3.10)$$

donde p_i es un punto de la trayectoria T .

Así como la duración de las trayectorias, podemos encontrar otras características, como: la aceleración, orientación, etc., es decir, propiedades dinámicas. Por lo tanto, las métricas basadas en características, son aquellas que utilizan estas propiedades dinámicas, para comparar trayectorias. Algunas métricas que se basan en características se detallarán posteriormente (*sección 4.5*).

3.6. Resumen

En este capítulo se revisaron diferentes técnicas para la obtención del movimiento humano. Los sistemas basados en sensores, se dividen en inerciales, ópticos y mix-

tos. Los inerciales obtienen la posición del cuerpo humano en el tiempo utilizando acelerómetros y giroscopios. Los ópticos, emiten luz infrarroja que es detectada por unas cámaras para determinar su posición. Los ópticos, no pueden medir la aceleración y los cambios bruscos de dirección y los inerciales tienen deficiencias al obtener la posición en el espacio. Existen sistemas que utilizan sensores inerciales y ópticos para evitar las deficiencias de ambos. Los sistemas de visión con marcas, se benefician de las articulaciones del cuerpo humano para poder detectar su forma al colocarle marcas en dichas articulaciones, estas marcas pueden ser activas, pasivas o mixtas. Las marcas eliminan el problema de la segmentación, pero ocasionan que se pierda a la persona o alguna de sus extremidades por el problema de oclusión. Los sistemas de visión sin marcas, no tienen el problema de oclusión, ya que la mayoría se basa en detectar al cuerpo humano por medio de definir su estructura antes de la detección, como lo son los *blobs* o manchas y que definen la cabeza, las extremidades, el tronco y las manos y los pies, y los *stick* que determinan la estructura en forma de alambre del cuerpo humano.

Una vez obtenido el movimiento humano, este puede ser representado como una trayectoria, en donde las coordenadas cartesianas del objeto obtenidas en cada cuadro de video forman una curva temporal. El objetivo de este proyecto es obtener calificaciones de ejercicios terapéuticos. Al ver los movimientos como trayectorias abre la posibilidad de comparar los ejercicios realizados, con los ejercicios ya definidos como correctos. En este capítulo se revisaron métricas de distancia, las cuales tienen varias restricciones, como son igual duración e igual número de muestra de cuadros por unidad de tiempo. Por lo tanto, se necesita de una técnica que no tenga estas restricciones. Como se verá en el capítulo siguiente, los Modelos Ocultos de Markov pueden resolver estos problemas.

Capítulo 4

Modelos Ocultos de Markov

*“Existe, pues, una continuidad predecible,
en la medida que una serie de caracteres
permite anticipar la probabilidad
de otra sucesión de caracteres”
Andrei Andreivich Markov*

Para obtener una calificación objetiva de los movimientos del brazo, se necesita de una escala paramétrica. Una forma de lograr una calificación paramétrica de un movimiento dado, es comparando este movimiento con un movimiento bien realizado. Los movimientos pueden ser considerados como trayectorias si obtenemos la sucesión de posiciones en el espacio del brazo. Una solución para poder comparar dos trayectorias es utilizando las técnicas vistas en la sección 3.5.1. Pero como ya vimos, estas tienen algunas restricciones. Una forma de resolverlo sin tener estas restricciones es utilizando los modelos ocultos de Markov (MOM).

Los MOM fueron introducidos en una serie de artículos estadísticos por *Leonard E. Baum* en la segunda mitad de la década de los 60's, pero fué hasta finales de los años 70's cuando los MOM fueron utilizados en reconocimiento de voz; y más tarde en la década de los 80's y 90's, en otras aplicaciones como al análisis de secuencias biológicas y en el reconocimiento de secuencias visuales [37].

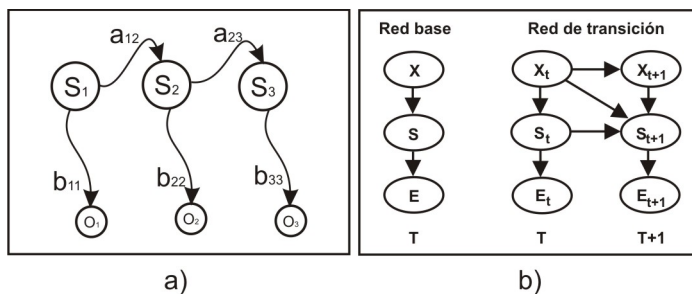


Figura 4.1: a) *Diagrama generalizado de un Modelo Oculto de Markov*, b) *Representación de una red bayesiana dinámica simple por medio de su red base y su red de transición*

Los modelos ocultos de Markov describen un proceso de probabilidad el cual produce una secuencia de eventos o símbolos observables. Son llamados ocultos porque hay un proceso de probabilidad subyacente que no es observable, pero afecta la secuencia de los eventos observados. Un modelo oculto de Markov, puede ser pensado como una máquina de estados finitos donde las transiciones entre estados dependen de la ocurrencia de algún símbolo. A esto se encuentra asociado una distribución de probabilidad que describe la probabilidad con la cual un símbolo ocurrirá durante la transición.

En este capítulo se presenta una descripción de los modelos ocultos de Markov, su definición, sus tres problemas básicos, las diferentes topologías de los modelos ocultos de Markov y algunas métricas para comparar modelos ocultos de Markov.

4.1. Descripción de los MOM

Los modelos ocultos de Markov o por sus siglas en español, MOM¹, son un modelo estadístico en el que se asume que el sistema a modelar es un proceso de Markov de parámetros desconocidos. El objetivo es determinar, cuáles son los parámetros desconocidos (u ocultos, de ahí que se llamen modelos ocultos) de

¹En inglés, el término para referirse es HMM, de *Hidden Markov Models*.

dicha cadena a partir de los parámetros que si se pueden observar. Un MOM puede ser considerado como la red bayesiana dinámica más² simple. Un diagrama general de un MOM se puede apreciar en la *Figura 4.1*.

En un modelo de Markov (normal), el estado se puede observar directamente por el observador, por lo que las probabilidades de transición entre estados son los únicos parámetros. En un modelo oculto de Markov, el estado no es visible directamente, sino que sólo lo son las variables o símbolos influidos por el estado, llamadas observaciones. Cada estado tiene una distribución de probabilidad sobre los posibles símbolos de salida u observaciones. De tal manera, la secuencia de observaciones generada por un MOM proporciona cierta información acerca de la secuencia de estados.

4.2. Definición

Un MOM, denotado como λ , es un proceso doblemente *estocástico*³. La primera capa estocástica es un proceso de Markov de *primer orden*, representado por el diagrama de transición de estados de la *Figura 4.1*. La probabilidad de transición del estado S_1 al estado S_2 es $P(q_t = S_2 | q_{t-1} = S_1)$. La probabilidad de ir al estado S_2 en el tiempo t dado que el estado en el tiempo $t - 1$ es S_1 . La segunda capa *estocástica* del MOM es el conjunto de probabilidades de salida para cada estado.

Formalmente, un Modelo Oculto de Markov Discreto⁴ de primer orden se caracteriza por los siguientes elementos [37]:

1. N , el número de estados en el modelo. El conjunto de estados individuales

²Consisten en una representación de los estados del proceso en un tiempo (red base) y las relaciones temporales entre dichos procesos (red de transición) ver *Figura 4.1*.

³Teoría estadística de los procesos cuya evolución en el tiempo es aleatoria.

⁴MOM compuesto de observaciones caracterizadas por símbolos discretos de un alfabeto finito.

se denota por $S = \{S_1, S_2, \dots, S_N\}$, y un estado en el tiempo t como q_t .

2. M , el número de símbolos de observación por estado. El conjunto de símbolos de observación $O = \{O_1, \dots, O_N\}$, se interpretan como la salida física del sistema modelado. Los símbolos individuales se denotan como $V = \{v_1, v_2, \dots, v_M\}$.

3. La distribución de probabilidad de transición entre estados $A = \{a_{ij}\}$, donde:

$$a_{ij} = P(q_t = S_j | q_{t-1} = S_i), \quad 1 \leq i, j \leq N$$

para:

$$a_{ij} \geq 0, \quad 1 \leq i, j \leq N \quad y \quad \sum_{j=1}^N a_{ij} = 1, \quad 1 \leq i \leq N$$

$$A = \begin{pmatrix} a_{1,1} & \cdots & a_{1,N} \\ \vdots & & \vdots \\ a_{N,1} & \cdots & a_{N,N} \end{pmatrix} \quad (4.1)$$

4. La distribución de probabilidad de símbolos de observación en el estado j , $B = \{b_j(k)\}$, donde:

$$b_j(k) = P(O_t = v_k | q_t = S_j), \quad 1 \leq j \leq N, \quad 1 \leq k \leq M$$

para:

$$b_j(k) \geq 0, \quad 1 \leq j \leq N, \quad 1 \leq k \leq M \quad y \quad \sum_{k=1}^M b_j(k) = 1, \quad 1 \leq j \leq N$$

$$B = \begin{pmatrix} b_{1,1} & \cdots & b_{1,N} \\ \vdots & & \vdots \\ b_{M,1} & \cdots & b_{M,N} \end{pmatrix} \quad (4.2)$$

5. La distribución de probabilidades iniciales para los estados $\pi = \{\pi_i\}$, donde:

$$\pi_i = P(q_1 = S_i), \quad 1 \leq i \leq N$$

para:

$$\pi_i \geq 0, \quad 1 \leq i \leq N \quad y \quad \sum_{i=1}^N \pi_i = 1$$

$$\pi = \left(\pi_1 \quad \cdots \quad \pi_N \right) \tag{4.3}$$

Por conveniencia un Modelo Oculto de Markov se puede especificar de forma compacta por medio de la siguiente notación:

$$\lambda = (A, B, \pi)$$

Donde A representa la distribución de probabilidad de transición de estados (*matriz* 4.1), B representa la distribución de probabilidad de símbolos de observación (*matriz* 4.2) y, π representa la distribución de probabilidades iniciales (*vector* 4.3).

4.3. Los 3 problemas básicos de los MOM

Existen tres problemas básicos de interés que deben ser resueltos para que un modelo (λ) pueda ser usado en aplicaciones reales. Estos problemas son descritos en [37], de la siguiente forma:

1. **Problema 1. Evaluación.** Dada una secuencia de observaciones $O = O_1, O_2, \dots, O_T$, y el modelo $\lambda = \{A, B, \pi\}$, ¿cómo podemos calcular eficientemente

$P(O|\lambda)$, la probabilidad de la secuencia de observaciones dado el modelo?. Es decir, que la probabilidad de la secuencia observada, sea generada por el modelo.

2. **Problema 2.** *Decodificación.* Dada una secuencia de observaciones $O = O_1, O_2, \dots, O_T$, y el modelo λ , ¿cómo podemos encontrar una secuencia de estados correspondiente $Q = q_1, q_2, \dots, q_T$ que haya generado dicha secuencia de observaciones? Se trata de descubrir la parte oculta del modelo, esto es, encontrar la secuencia más probable de estados.
3. **Problema 3.** *Estimación.* ¿Cómo podemos ajustar los parámetros del modelo $\lambda = \{A, B, \pi\}$ que maximice $P(O|\lambda)$?

4.3.1. Solución al problema 1: *evaluación*

Existen dos formas de resolver este problema por el método directo y el método *forward - backward*. El cálculo de $P(O|\lambda)$, de acuerdo a la forma *directa*, es de un orden de complejidad $2T \cdot N^T$ operaciones [37]. Para propósitos prácticos este cálculo es computacionalmente infactible. Por tal motivo sólo se detallará el método *forward - backward* [37].

Considérese la variable $\alpha_t(i)$ denominada *variable forward*, definida como:

$$\alpha_t(i) = P((O_1, O_2, \dots, O_t), q_t = S_i | \lambda) \quad (4.4)$$

que corresponde a la probabilidad de una secuencia parcial de observación, O_1, O_2, \dots, O_t , (hasta el tiempo t) y el estado S_i en el tiempo t , dado el modelo λ . Se puede resolver inductivamente a la variable $\alpha_t(i)$ como sigue:

1. *Inicialización:*

$$\alpha_1(i) = \pi_i b_i(O_1), \quad 1 \leq i \leq N \quad (4.5)$$

Se inicializa la *variable forward* como la probabilidad conjunta del estado S_i y la observación inicial O_1 .

2. *Inducción:*

$$\alpha_{t+1}(j) = \left[\sum_{i=1}^N \alpha_t(i) a_{ij} \right] b_j(O_{t+1}), \quad 1 \leq t \leq T-1, \quad 1 \leq j \leq N \quad (4.6)$$

Se calcula la *variable forward* inductivamente y a través del tiempo.

3. *Terminación:*

$$P(O|\lambda) = \sum_{i=1}^N \alpha_T(i) \quad (4.7)$$

Se calcula la probabilidad deseada $P(O|\lambda)$ como la suma de las variables terminales $\alpha_T(i)$.

El problema de evaluación también se puede resolver con el uso de la *variable backward* $\beta_t(i)$, definida como[37]:

$$\beta_t(i) = P(O_{t+1}, O_{t+2}, \dots, O_T | q_t = S_i, \lambda) \quad (4.8)$$

que corresponde a la probabilidad de la secuencia parcial del tiempo $t+1$ hasta T , dado el estado S_i en el tiempo t y el modelo λ . De la misma forma, se puede resolver el problema 1 utilizando la variable $\beta_t(i)$ inductivamente [37].

4.3.2. Solución al problema 2: *decodificación*

Se desea encontrar la secuencia de estados más probable Q que haya generado una secuencia de observaciones dada O . Existe una técnica formal para encontrar esta secuencia de estados, basada en métodos de programación dinámica, y es llamada el algoritmo de *Viterbi* [37]. Para hallar la mejor secuencia de estados, $Q =$

$\{q_1, q_2, \dots, q_T\}$, para una secuencia de observaciones dada $O = \{O_1, O_2, \dots, O_T\}$, se define la siguiente variable:

$$\delta_t(i) = \max_{q_1, q_2, \dots, q_{t-1}} P[(q_1, q_2, \dots, q_t) = i, (O_1, O_2, \dots, O_t) | \lambda] \quad (4.9)$$

es decir, $\delta_t(i)$ es la probabilidad más alta a través de una ruta única, en el tiempo t , que cuenta para las primeras t observaciones y termina en el estado S_i . Por inducción se tiene:

$$\delta_{t+1}(i) = [\max_j \delta_t(i) a_{ij}] \cdot b_j(O_{t+1}) \quad (4.10)$$

Para mantener la secuencia de estados, es necesario mantener un seguimiento del argumento que maximice $\delta_{t+1}(i)$, para cada t y j por medio de un arreglo $\psi_t(j)$. El proceso completo para hallar la mejor secuencia de estados es como sigue:

1. *Inicialización:*

$$\delta_1(i) = \pi_i b_i(O_1), \quad 1 \leq i \leq N \quad (4.11)$$

$$\psi_1(i) = 0 \quad (4.12)$$

2. *Recursión:*

$$\delta_t(j) = \max_{1 \leq i \leq N} [\delta_{t-1}(i) a_{ij}] b_j(O_t), \quad 2 \leq t \leq T, \quad 1 \leq j \leq N \quad (4.13)$$

$$\psi_t(j) = \arg \max_{1 \leq i \leq N} [\delta_{t-1}(i) a_{ij}], \quad 2 \leq t \leq T, \quad 1 \leq j \leq N \quad (4.14)$$

3. *Terminación:*

$$P^* = \max_{1 \leq i \leq N} [\delta_T(i)] \quad (4.15)$$

$$q_T^* = \arg \max_{1 \leq i \leq N} [\delta_T(i)] \quad (4.16)$$

4. *Recuperar la secuencia de estados:*

$$q_t^* = \psi_{t+1}(q_{t+1}^*), \quad t = T-1, T-2, \dots, 1 \quad (4.17)$$

4.3.3. Solución al problema 3: *estimación*

Se desea ajustar los parámetros del modelo (A, B, π) para maximizar una secuencia de observaciones O dado el modelo λ , esto es, $P(O|\lambda)$. Una forma de ajustar los parámetros del modelo λ es mediante el algoritmo de *Baum-Welch* [37].

Para describir el procedimiento de reestimación (actualización y mejoramiento iterativo) de los parámetros del MOM, se necesita definir primero la variable auxiliar $\xi_t(i, j)$ [37], la cual define la probabilidad de estar en el estado S_i en el tiempo t , y en el estado S_j en el tiempo $t + 1$, dado el modelo y la secuencia de observaciones. Esto es:

$$\xi_t(i, j) = P(q_t = S_i, q_{t+1} = S_j | O, \lambda) \quad (4.18)$$

La variable $\xi_t(i, j)$ se puede definir en términos de las variables *Forward - Backward*), de la siguiente manera:

$$\begin{aligned} \xi_t(i, j) &= \frac{\alpha_t(i) a_{ij} b_j(O_{t+1}) \beta_{t+1}(j)}{P(O|\lambda)} \\ &= \frac{\alpha_t(i) a_{ij} b_j(O_{t+1}) \beta_{t+1}(j)}{\sum_{i=1}^N \sum_{j=1}^N \alpha_t(i) a_{ij} b_j(O_{t+1}) \beta_{t+1}(j)} \end{aligned} \quad (4.19)$$

donde el numerador es justamente $P(q_t = S_i, q_{t+1} = S_j, O|\lambda)$ y la división por $P(O|\lambda)$ nos da la medida de la probabilidad deseada.

También es necesario definir la variable $\gamma_t(i)$, que corresponde a la probabilidad de estar en el estado S_i en el tiempo t , dada la secuencia de observaciones y el modelo. Esto es:

$$\gamma_t(i) = P(q_t = S_i | O, \lambda) \quad (4.20)$$

Se puede relacionar a la variable $\gamma_t(i)$ con la variable $\xi_t(i, j)$ si se suma sobre j , esto es:

$$\gamma_t(i) = \sum_{j=1}^N \xi_t(i, j) \quad (4.21)$$

Si se suma $\gamma_t(i)$ sobre el tiempo t , se obtiene una cantidad que puede ser interpretada como el número de veces (sobre el tiempo) que el estado S_i es visitado. De manera similar, la suma de $\xi_t(i, j)$ sobre t (desde $t = 1$ hasta $t = T - 1$) puede ser interpretada como el número esperado de transiciones del estado S_i al estado S_j . Esto es:

$$\sum_{t=1}^{T-1} \gamma_t(i) = \text{numero esperado de transiciones del estado } S_i \quad (4.22)$$

$$\sum_{t=1}^{T-1} \xi_t(i, j) = \begin{array}{l} \text{numero esperado de transiciones} \\ \text{del estado } S_i \text{ al estado } S_j \end{array} \quad (4.23)$$

Usando las fórmulas anteriores, se puede obtener un método para reestimar los parámetros de un MOM. Las fórmulas de reestimación para π , A , y B son:

$$\bar{\pi}_i = \begin{array}{l} \text{frecuencia esperada (numero de veces)} \\ \text{en el estado } S_i \text{ en el tiempo } (t = 1) = \gamma_1(i) \end{array} \quad (4.24)$$

$$\bar{a}_{ij} = \frac{\begin{array}{l} \text{numero esperado de transiciones} \\ \text{del estado } S_i \text{ al estado } S_j \end{array}}{\begin{array}{l} \text{numero esperado de veces en el estado } j \\ \text{y haber observado el simbolo } v_k \end{array}} \quad (4.25)$$

$$= \frac{\sum_{t=1}^{T-1} \xi_t(i, j)}{\sum_{t=1}^{T-1} \gamma_t(i)}$$

$$\bar{b}_j(k) = \frac{\begin{array}{l} \text{numero esperado de veces en el estado } j \\ \text{y haber observado el simbolo } v_k \end{array}}{\begin{array}{l} \text{numero esperado de veces} \\ \text{en el estado } j \end{array}} \quad (4.26)$$

$$= \frac{\sum_{t=1}^T \gamma_t(j), O_t = v_k}{\sum_{t=1}^T \gamma_t(j)} \quad (4.27)$$

Si iterativamente se utiliza el modelo $\bar{\lambda} = (\bar{A}, \bar{B}, \bar{\pi})$ para repetir la reestimación, entonces se puede mejorar la probabilidad $P(O|\lambda)$ hasta alcanzar algún

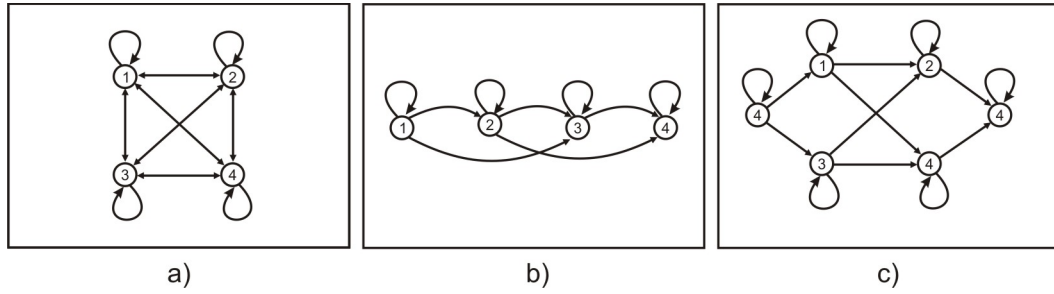


Figura 4.2: Grafos que muestran las topologías de los MOM: a)Ergódica, b)Bakis, c)Cross - coupled

óptimo local [37].

4.4. Topologías de los MOM

Un MOM puede ser representado como un grafo dirigido, donde los nodos representan a los estados y las flechas representan las transiciones entre los estados. Por lo tanto podemos encontrar diferentes topologías dependiendo de las propiedades del MOM. La topología específica que permita modelar lo mejor posible las propiedades observadas depende en gran medida de las características del problema. Las topologías más usadas son [37]:

1. *Ergódicas*: Grafos completamente conectados, en los cuales cada estado del modelo puede ser alcanzado desde cualquier otro estado en un número finito de pasos (ver *Figura 4.2a*).
2. *Bakis*: Grafos conectados de Izquierda-derecha o hacia adelante, los cuales tienen la propiedad de que en la medida que el tiempo crece, se avanza en la secuencia de observaciones asociada O y en esa misma medida el índice que señala el estado del modelo permanece o crece, es decir, los estados del sistema van de izquierda a derecha (ver *Figura 4.2b*). Estas topologías

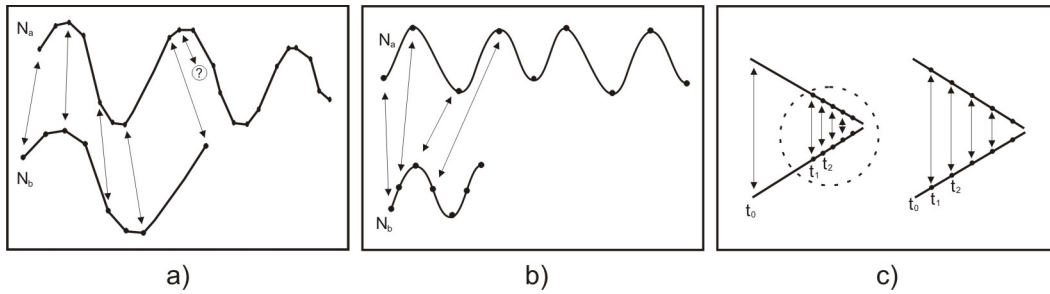


Figura 4.3: *Ambigüedades para las métricas convencionales*[35]. a) *Duraciones diferentes*, b) *Igual duración, pero diferentes instantes de tiempo*, c) *Efecto desigual en los instantes de tiempo*

modelan bien los aspectos lineales de las secuencias, un ejemplo de su uso es en secuencias biológicas, en reconocimiento de la voz, y en el reconocimiento de gestos visuales.

3. *Cross - coupled*: Grafos de Izquierda-derecha y paralelos, es como si fueran dos topologías izquierda-derecha conectadas entre sí (ver *Figura 4.2c*).

4.5. Métricas para comparar MOM

En visión computacional, cuando hablamos de seguir objetos, un aspecto importante en la evaluación de los resultados del seguimiento es una métrica de la distancia que determine la similitud entre las trayectorias [35]. La mayoría de las medidas existentes [25, 40], calculan la distancia media entre los puntos correspondientes de dos trayectorias de igual duración. Como se vió en la *sección 3.5.1*, las medidas estadísticas como las distancias basadas en la varianza, la mediana, el mínimo y el máximo, son también propuestas para extender la descripción de la similitud. Sin embargo, (como se mencionó en la *sección 3.5.1*) existen ciertas limitaciones cuando se tratan de comparar las trayectorias:

- *Duraciones iguales.* Las trayectorias están limitadas a igual duración. Por *duración* se entiende, como al número de puntos coordinados que constituyen la trayectoria (ver *Figura 4.3a*).
- *Muestras de cuadros con el mismo intervalo de tiempo.* Las coordenadas de las trayectorias son tomadas en diferentes instantes de tiempo (diferencias constantes) (ver *Figura 4.3b*).
- *Normalización.* Si se normaliza o parametriza la duración de la trayectoria para que pueda ser comparada, esta normalización destruye las propiedades temporales de la trayectoria.
- *Igualdad en los instantes de tiempo.* Las coordenadas de las trayectorias son tomadas en diferentes instantes de tiempo (diferencias desiguales) (ver *Figura 4.3c*). Este es un problema común especialmente en el seguimiento de objetos en *tiempo real* que procesan el flujo de datos de las cámaras de video.

Como la mayoría de las trayectorias cumplen con estas restricciones, es necesario utilizar una técnica que pueda medir efectivamente las diferencias entre trayectorias sin estas restricciones. Una alternativa es utilizar los MOM que, como se verá más adelante, pueden capturar las propiedades dinámicas de las trayectorias (ver *sección 3.5.2*). Existen varias técnicas para obtener una métrica que pueda medir la similitud entre trayectorias basadas en los MOM, a continuación se presentan las más relevantes.

En general, las métricas basadas en MOM tratan de proyectar la trayectoria T (o un conjunto de trayectorias), dentro del espacio del modelo λ , que es caracterizado por los parámetros del MOM, esto es, $\lambda = \{A, B, \pi\}$. De esta forma, se pretende reemplazar la información de la trayectoria con los símbolos de ob-

servación, y la matriz de transición de estados. Por lo tanto, para comparar dos trayectorias, T_1 y T_2 , se necesita comparar sus modelos λ_1 y λ_2 , que representan a ambas trayectorias, respectivamente. En un caso más general, se puede entrenar un modelo λ que represente a una clase de trayectorias, de esta forma, se puede determinar la similitud de una nueva trayectoria respecto al modelo entrenado. En las siguientes secciones se presentan algunas métricas para comparar pares de MOM, que pueden ser utilizadas como una alternativa en la comparación de trayectorias.

4.5.1. Levinson

La propuesta de Levinson [31], fué la primera medida que apareció en la literatura para comparar pares de MOM con observaciones discretas. La medida, que es una *distancia Euclidiana* sobre la matriz de las probabilidades de observación de estados, esta definida como:

$$D_e(\lambda_1, \lambda_2) = \sqrt{\frac{1}{N} \sum_{j=1}^N \sum_{k=1}^M \|b_{jk}^1 - b_{jk}^2\|^2} \quad (4.28)$$

donde b_{jk}^i es la matriz de probabilidad de observación de estados en el modelo λ y el número de estados de ambos modelos debe de ser idéntico, $N^1 = N^2$. Un caso más general, es utilizando la mínima distancia euclidiana:

$$D_{mec}(\lambda_1, \lambda_2) = \sqrt{\frac{1}{N} \sum_{j=1}^N \min_j \sum_{k=1}^M \|b_{jk}^1 - b_{jk}^2\|^2}$$

donde b_{jk}^i es la matriz de probabilidad de observación de estados en el modelo λ . Sin embargo, esta medida de la distancia no toma en cuenta la estructura temporal representada en la *cadena de Markov*, y es fácil encontrar algún par de modelos λ_1 y λ_2 donde la distancia tienda a cero pero las medidas de probabilidad generadas $P(O_1|\lambda_1)$ y $P(O_2|\lambda_2)$ sean completamente diferentes [16].

4.5.2. Kullback-Leibler

Juang y Rabiner [36], proponen una medida de la distancia para comparar pares de MOM. Esta medida está basada en la distancia de Kullback-Leibler [30] entre dos funciones de distribución de probabilidad, que sigue los conceptos de la divergencia, entropía cruzada, o información de discriminación. La medida de la distancia entre dos MOM, λ_1 y λ_2 , está definida como:

$$D_k(\lambda_1, \lambda_2) = \frac{1}{T} [\log(P[O^1|\lambda_1]) - \log(P[O^1|\lambda_2])] \quad (4.29)$$

donde O^1 es la secuencia de observaciones generada por el modelo λ_1 y T es la longitud de la secuencia de observaciones. Esta medida se puede interpretar como que tan bien el modelo λ_2 coincide con las observaciones generadas por el modelo λ_1 , relativo a que tan bien el modelo λ_1 coincide con las observaciones generadas por si mismo[37].

Se puede notar que la medida de la distancia de la ecuación 4.29 *no es simétrica*. Una extensión de esta medida es su versión *simétrica*, definida como[36]:

$$D_{sk}(\lambda_1, \lambda_2) = \frac{D(\lambda_1, \lambda_2) + D(\lambda_2, \lambda_1)}{2} \quad (4.30)$$

que es el promedio de las dos distancias *no-simétricas*. La medida $D_{sk}(\lambda_1, \lambda_2)$ es simétrica respecto a λ_1 y λ_2 y representa la medida de la dificultad de discriminar $P(O|\lambda_1)$ y $P(O|\lambda_2)$.

Un problema en la medida de la distancia de la ecuación 4.29 es su convergencia. En [36] se realizan una serie de pruebas experimentales y concluyen que la distancia converge en medida que el tamaño de la secuencia de observaciones T incrementa. En general, la medida $D(\lambda_1, \lambda_2)$ es sensitiva a los parámetros del modelo λ : M (el número de símbolos de observación por estado) y N (el número de estados).

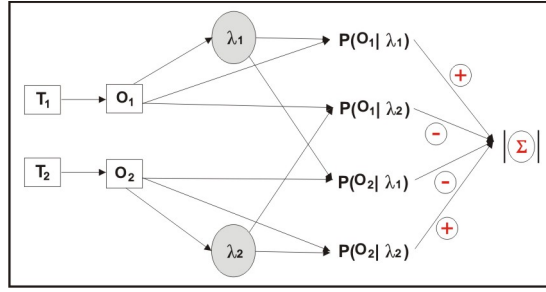


Figura 4.4: Imagen que muestra la distancia de cruce entre dos MOM[35].

4.5.3. Porikli

Porikli [35] propone una medida de la distancia para comparar trayectorias de un objeto basadas en la representación de MOM. La información de la trayectoria se parametriza en un modelo λ , donde los símbolos de observación sustituyen las posiciones del objeto, y la matriz de transición de estados captura las propiedades dinámicas de la trayectoria. La medida, llamada *distancia de cruce*, esta definida como:

$$D_{porikli}(T_1, T_2) = |P(O_1|\lambda_1) + P(O_2|\lambda_2) - P(O_1|\lambda_2) - P(O_2|\lambda_1)| \quad (4.31)$$

que corresponde al cruce de las trayectorias con el otro modelo, como se ilustra en la Figura 4.4. Donde O_i es la secuencia de observaciones (trayectoria) y $P(O_i|\lambda_i)$ es la probabilidad de la secuencia de observaciones dado el modelo. Los términos $P(O_1|\lambda_1)$ y $P(O_2|\lambda_2)$ indican la similitud de las trayectorias de pertenecer a su propio modelo. Los términos de cruce $P(O_1|\lambda_2)$, $P(O_2|\lambda_1)$ revelan la similitud de la trayectoria generada por los otros modelos de trayectorias. En otras palabras, si dos trayectorias son idénticas, los términos de cruce tendrán un valor máximo, así la ecuación 4.31 será igual a cero. Por otro lado, si las trayectorias son diferentes, su similitud de ser generadas por el otro modelo será pequeña, esto es, la distancia será grande.

La medida de la distancia de la ecuación 4.31 está pensada para modelar una sola trayectoria en un modelo λ . Por lo tanto, para comparar dos trayectorias, la información de cada trayectoria es utilizada para entrenar el modelo. Así, se tendría un modelo por cada trayectoria que se pretenda comparar. Además, se utilizan MOM de densidad continua, esto es, se utilizan *mezclas de gaussianas* para modelar los símbolos de observación en el modelo.

4.6. Resumen

Un MOM es un modelo estadístico en el que se asume que el sistema a modelar es un proceso de Markov de parámetros desconocidos. El objetivo es determinar los parámetros desconocidos (u ocultos) de dicha cadena a partir de los parámetros observables. Los MOM son herramientas útiles para modelar series de datos en el tiempo. Existen tres problemas básicos para que los MOM puedan ser aplicados en el mundo real: a) evaluación, que determina la probabilidad de la secuencia de observaciones dado el modelo, b) decodificación, que consiste en encontrar la secuencia de estados más probable que haya generado una secuencia de observaciones dada, y c) entrenamiento, que consiste en ajustar los parámetros del MOM para maximizar la probabilidad dada una secuencia de observaciones y el modelo.

Si los MOM son vistos como grafos dirigidos, entonces se pueden definir diferentes topologías: ergódicas, que representan a los grafos completamente conectados; Bakis, grafos conectados de izquierda a derecha e izquierda-derecha paralelas, grafos conectados de izquierda a derecha en paralelo.

Como los MOM pueden representar la dinámica de un proceso, es fácil representar una trayectoria como un MOM. Actualmente existen diferentes métodos para comparar MOM's. El primer método es el propuesto por Levinson, que se basa principalmente en comparar la matriz de observaciones, de esta forma se

comparan los modelos por la similitud de sus probabilidades por generar observaciones en cada uno de los estados. La distancia de Kullback-Leibler, es una buena aproximación para comparar dos MOM, ya que considera todos los parámetros de un MOM; esta basada en los conceptos de la divergencia, entropía cruzada, o información de discriminación. Un problema de esta distancia es la convergencia, debido a que depende de la longitud de la secuencia de observaciones generada por el modelo. Porikli, diseña una distancia para comparar trayectorias de movimiento de objetos, donde la información de la trayectoria se reemplaza por los símbolos de observación del modelo, y los estados capturan las propiedades espacio-temporales de la trayectoria. Sin embargo, se requiere un modelo por cada trayectoria individual, lo que hace complicado modelar varias trayectorias de una misma clase.

Los modelos ocultos de Markov, como se verá en el capítulo 5, constituyen la tercera etapa del sistema general que se desarrolló en este proyecto de tesis para obtener la calificación automática de gestos terapéuticos. En el siguiente capítulo se describirá el sistema desarrollado en este trabajo de tesis, el cual esta conformado por tres etapas: seguimiento visual, modelo de movimiento y calificación.

Capítulo 5

Descripción del sistema

5.1. Introducción

Antes de poder obtener una calificación para un ejercicio terapéutico, es necesario contar con una herramienta capaz de detectar las características del ejercicio. Para esto se necesita saber que características son necesarias para poder calificar el ejercicio. En nuestro caso, necesitamos los ángulos en el tiempo del movimiento del brazo humano al realizar el ejercicio. Para obtener los ángulos, se necesita de las posiciones en el espacio de las articulaciones del brazo humano en el tiempo. Un seguidor visual con marcas utilizando dos *Webcams*, puede resolver el problema de obtener las posiciones de las articulaciones. Para poder seguir las marcas, se necesita de un segmentador y de un seguidor 2D que ayude a predecir la posición de las marcas en el siguiente cuadro. Una vez obtenidas en cada una de las dos *Webcams*, las posiciones en dos dimensiones de las marcas, es decir de las articulaciones, es necesario construir una representación en tres dimensiones. Para la reconstrucción 3D, se necesitan los parámetros intrínsecos y extrínsecos de las cámaras.

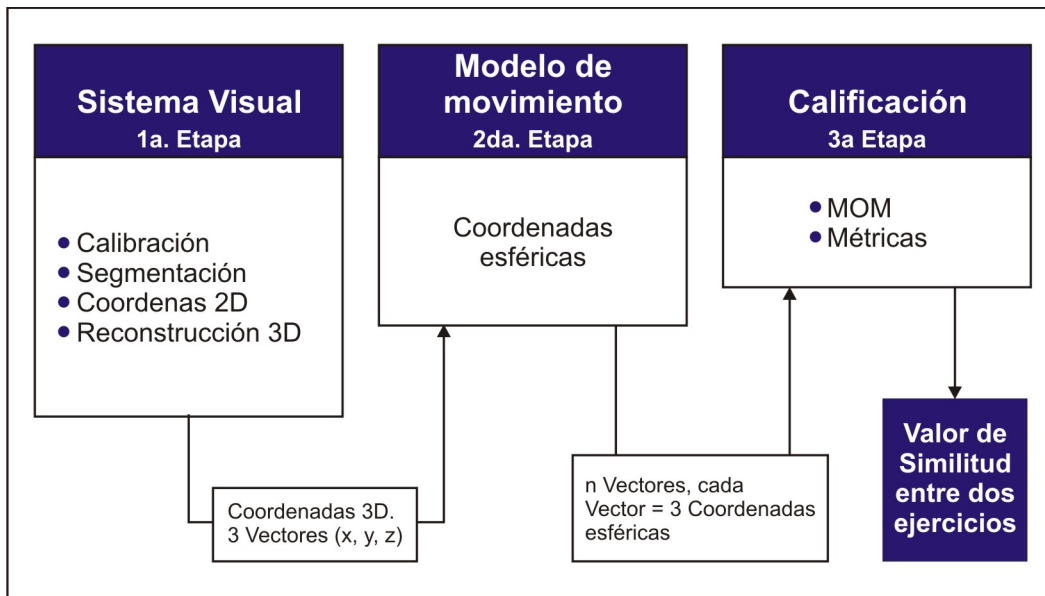


Figura 5.1: Diagrama general del sistema. 1a. Etapa: Seguimiento visual, 2da. Etapa Modelo de Movimiento. 3a. Etapa Calificación.

El paso siguiente es entrenar los MOM que se encargarán de generar valores de similitud entre la realización de los ejercicios y los ejercicios bien hechos. Para entrenarlos se necesitan secuencias de ejercicios bien realizados. Como al realizar el ejercicio los datos de las posiciones de las articulaciones del brazo son continuos, se necesita realizar un proceso de *discretización* de los datos.

En este capítulo se analizará cada una de las partes del sistema que obtiene una calificación automática de los ejercicios terapéuticos. La primera parte describe el sistema visual: *calibración, segmentador, seguimiento 2D y reconstrucción 3D*. La segunda parte, describe el modelo de movimiento para la obtención de características del ejercicio, así como la descripción de los mismos. Y la última parte, describe la utilización de los MOM: *discretización y entrenamiento* para la utilización de las *métricas* para la calificación. El diagrama general del sistema se puede apreciar en la *Figura 5.1*

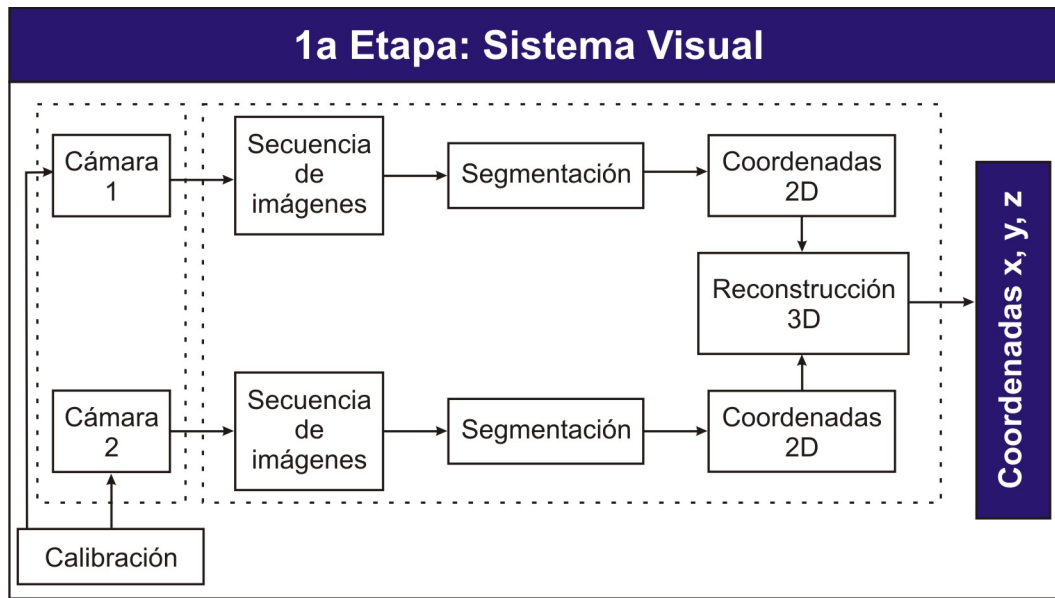


Figura 5.2: 1a. Etapa: Sistema Visual.

5.2. 1a Etapa: Sistema visual

La 1a. Etapa es el sistema visual (ver Figura 5.2), que se encarga de obtener las coordenadas x, y, z , localizando y siguiendo el brazo a partir del procesamiento de las secuencias de imágenes capturadas por dos *webcams*. Para seguir el brazo se utilizan marcas (cintas de color) en cada articulación (hombro, codo y muñeca). El sistema visual esta compuesto por los siguientes módulos:

- *Calibración*. Este módulo se encarga del calibrado de las cámaras. Se obtienen los parámetros intrínsecos que son utilizados para obtener la posición 3D de las marcas de color¹.

¹Aún cuando los parámetros extrínsecos no se obtienen por medio de la calibración, si son necesarios para la reconstrucción 3D, por lo tanto en esta etapa se obtienen estos parámetros manualmente.

- *Segmentación*. Este módulo se encarga de la localización de las marcas de color dentro de una imagen. En el proceso de segmentación de las marcas, se aplica información de color (histogramas de color).
- *Seguimiento 2D*. Este módulo se encarga del seguimiento de un objeto a través de una secuencia de imágenes de video. El seguimiento se aplica en las secuencias de imágenes en cada una de las dos *webcams*.
- *Reconstrucción 3D*. En este módulo se obtienen las coordenadas 3D del objeto en seguimiento. Se aplica un algoritmo de intersección de líneas de visión al objeto en seguimiento de ambas cámaras para obtener una estimación de la posición 3D.

5.2.1. Calibración

Para obtener las coordenadas x, y, z de las articulaciones del brazo, se necesita saber tanto el modelo de proyección como la posición en el espacio de las cámaras. La calibración de la cámara es un proceso que relaciona un *modelo ideal* con el dispositivo físico y determina la posición y orientación de ésta respecto al sistema de referencia del mundo [21]. A esto se les llama parámetros de la cámara *intrínsecos*² y *extrínsecos*³, respectivamente [34]. Por lo tanto, la calibración consiste en obtener los parámetros intrínsecos y extrínsecos de las cámaras.

Los parámetros intrínsecos son obtenidos por un algoritmo de calibrado que utiliza un patrón de calibrado que es colocado frente a la cámara (ver *Figura 5.3a*). El algoritmo de calibración detecta las esquinas de las cuadrículas del tablero, almacenando una secuencia de imágenes con diferentes vistas del tablero (ver *Figuras 5.3b y 5.3c*).

²Longitud focal, centro óptico y tamaño del pixel.

³Posición y orientación.

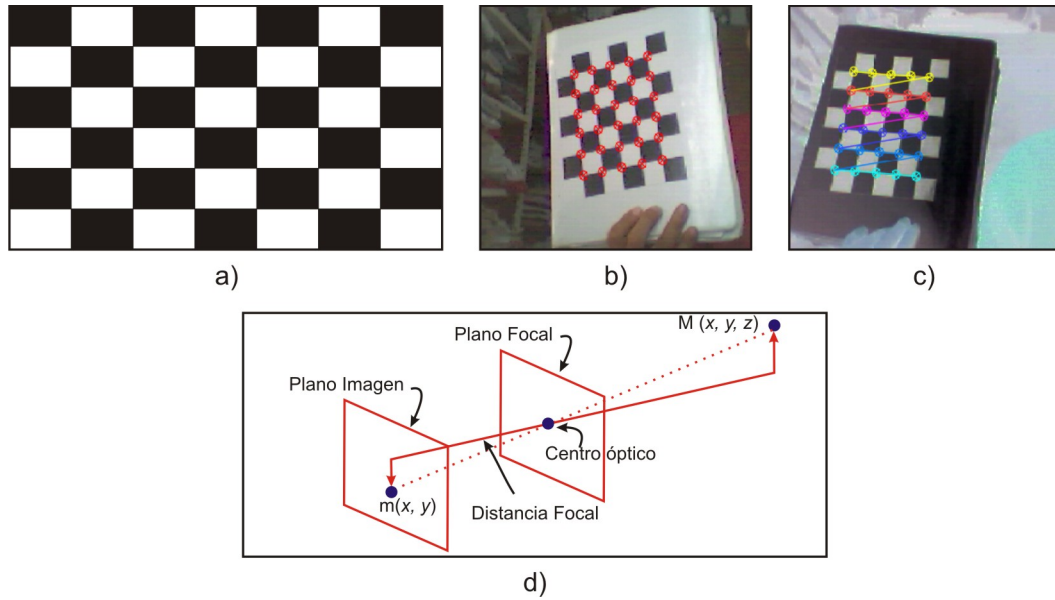


Figura 5.3: *Parámetros intrínsecos. a) Patrón de calibrado, b) y c) Detección de las esquinas, d) Modelo pinhole.*

El modelo de proyección utilizado es el modelo *pin-hole* (ver Figura 5.3d), que está basado en el principio de colinealidad, donde cada punto en el espacio del objeto es proyectado en una línea recta a través del centro óptico (centro de proyección) del plano de la imagen [21].

La relación entre un punto M en 3D y su proyección en la imagen m esta dada por la fórmula:

$$m = A [Rt] M \quad (5.1)$$

donde A es la matriz intrínseca de la cámara, definida como:

$$A = \begin{bmatrix} f_x & 0 & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{bmatrix} \quad (5.2)$$

y (c_x, c_y) son las coordenadas del centro óptico. (f_x, f_y) son las longitudes focales sobre los ejes (x, y) . Los parámetros intrínsecos describen la geometría interna y

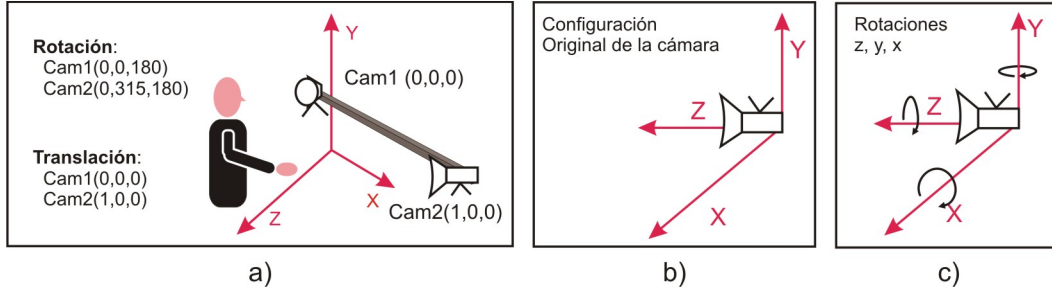


Figura 5.4: *Parámetros extrínsecos. a) Configuración de las cámaras, b) Configuración original de las cámaras, c) Rotación de las cámaras.*

características ópticas de los lentes y del dispositivo de proyección de la imagen. La longitud focal es la distancia entre el lente y el plano de la imagen. El centro óptico es el centro de la imagen en coordenadas de píxeles. (R, t) son los parámetros extrínsecos que describen la posición y orientación de la cámara en el sistema de referencia del mundo. Relacionan el sistema de coordenadas del mundo con el sistema de coordenadas de la cámara (ver *Figura 5.4*). R es la matriz de rotación y t el vector de translación, definidos como:

$$R = \begin{bmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{bmatrix}, \quad t = \begin{bmatrix} t_1 \\ t_2 \\ t_3 \end{bmatrix} \quad (5.3)$$

Un punto en el marco de referencia de la cámara $[X^C, Y^C, Z^C]^T$ está relacionado con su correspondiente punto en el marco de referencia del mundo $[X^W, Y^W, Z^W]^T$ por:

$$P^C = R(P^W - t) \quad (5.4)$$

Para obtener los parámetros intrínsecos de las cámaras, se utilizan las librerías de visión computacional *OpenCV* (*Open Source Computer Vision Library*) desarrolladas por *Intel* [50]. Los parámetros extrínsecos se obtienen manualmente por medio de medir la posición y orientación del sistema de referencia de la cámara

(centro del lente) respecto al sistema de referencia del mundo. La *Figura 5.4a* muestra el sistema de referencia del mundo y la colocación de las cámaras. La cámara 1 se encuentra en el centro de referencia del mundo en la posición $(0, 0, 0)$ y rotada 180 grados en el eje Z , orientada hacia el plano XY . La cámara 2 se encuentra en la posición $(0, 1, 1)$ y rotada 180 grados en el eje Z , y 315 grados en el eje Y , igualmente orientada hacia el plano XY ⁴. La rotación de las cámaras, depende de los grados en que gira la cámara, en los diferentes ejes (en el orden z, y, x , ver *Figura 5.4c*), respecto a la configuración original de la cámara (ver *Figura 5.4b*).

5.2.2. Segmentación

Para poder seguir el brazo, es necesario encontrarlo dentro de la imagen. La segmentación tiene la finalidad de encontrar, dentro de la imagen, las marcas de color que representan las articulaciones del brazo. Por lo tanto, la segmentación es el proceso mediante el cual se calcula qué píxeles de la imagen pertenecen al objeto de interés. Para segmentar un objeto de una imagen se pueden utilizar diferentes técnicas dependiendo que primitiva se utilice como discriminante. Básicamente se pueden establecer tres clases de segmentación [33]:

- *Basados en Pixel* (Por elementos): Se utilizan los valores de color de cada elemento de la imagen.
- *Basados en bordes* (Por discontinuidad): Se detectan cambios de intensidad en la frontera de los objetos y se extraen los bordes del mismo.
- *Basados en regiones* (Por continuidad): Se analizan los valores de color en grandes áreas de la imagen.

⁴Esta es la configuración utilizada en este trabajo de tesis, pero puede ser cualquier configuración, dependiendo de las características del lugar.

La segmentación basada en pixel, es el método utilizado en este proyecto de tesis. Para lograr localizar las marcas de color se realiza una segmentación empleando la diferencia de color. Aunque los métodos de segmentación basados en color han probado ser eficientes en una variedad de aplicaciones de visión, existen varios problemas asociados con estos métodos en donde la consistencia del color en diferentes ambientes es uno de los más importantes [28]. Para solventar este problema, es preciso considerar un modelo de color que sea robusto ante los cambios de iluminación. El modelo de color utilizado es el *HSV* [20], el cual permite separar el brillo o luminosidad del color, utilizando las componentes *H-S* (croma y saturación), e ignorando la componente *V* del modelo.

El método para localizar las marcas de color adheridas al brazo, se basa en el método propuesto por Swain y Ballard, llamado *Retroproyección de Histograma* [47], el cual sugiere la representación de los objetos por medio de histogramas de color, los cuales son almacenados en una base de datos. El objeto que se busca debe estar representado por un histograma M . También, se obtiene el histograma de la imagen I que se está analizando, y se calcula un tercer histograma R , resultado de dividir el histograma del modelo entre el histograma de la imagen. Es este histograma R , el que se mapea sobre la imagen, esto es, los valores de la imagen son sustituidos por los valores indexados por R . Así, el proceso para segmentar es el siguiente:

- Para cada objeto (marca de color = articulación) en la escena, se toman muestras del color y se crea un histograma bidimensional M en el espacio de color ($H-S$), donde $M(H, S)$ especifica el número de pixeles con el valor de la dupla (H, S) .
- Calcular el histograma I de la imagen a analizar.

- Obtener el histograma R .

$$R = \frac{M}{I} \quad (5.5)$$

- Para cada pixel de la imagen $I(x, y)$ se toma el valor del color (H, S) , y el pixel de la imagen resultante se sustituye por:

$$I^*(x, y) = \begin{cases} 255 & \text{si } R(H, S) > 1 \\ 128 \cdot \text{mín}(R(H, S), 1) & \text{de otro modo} \end{cases} \quad (5.6)$$

De esta forma, podemos transformar una imagen de color a una imagen en escala de grises, donde el valor de gris de cada pixel, muestra su grado de pertenencia con el modelo de color. Esta imagen se transforma en una imagen binaria si se utiliza un umbral como índice de pertenencia, donde tiene valor uno si pertenece al objeto y en otro caso, su valor es cero. Este proceso se aplica en cada una de las imágenes capturadas por las dos cámaras de video, buscando un objeto (marca de color) a la vez. La *Figura 5.5e*, muestra el histograma 2D en el espacio de color $(H - S)$, que representa la distribución de color de la marca roja (articulación = codo), y las *Figuras 5.5b* y *5.5c* muestran las imágenes resultantes al aplicar el método *Retroproyección de Histograma*, en escala de grises y binaria, respectivamente.

Después de segmentar las marcas de color, el siguiente paso es obtener las posiciones iniciales de cada una, a partir de las imágenes segmentadas. Para ello se utilizan los momentos de orden cero y uno [18]. El momento de orden cero coincide con el área del objeto, definido como:

$$m_{00} = \sum_x \sum_y I(x, y) \quad (5.7)$$

$$m_{10} = \sum_x \sum_y xI(x, y) \quad (5.8)$$

$$m_{01} = \sum_x \sum_y yI(x, y) \quad (5.9)$$

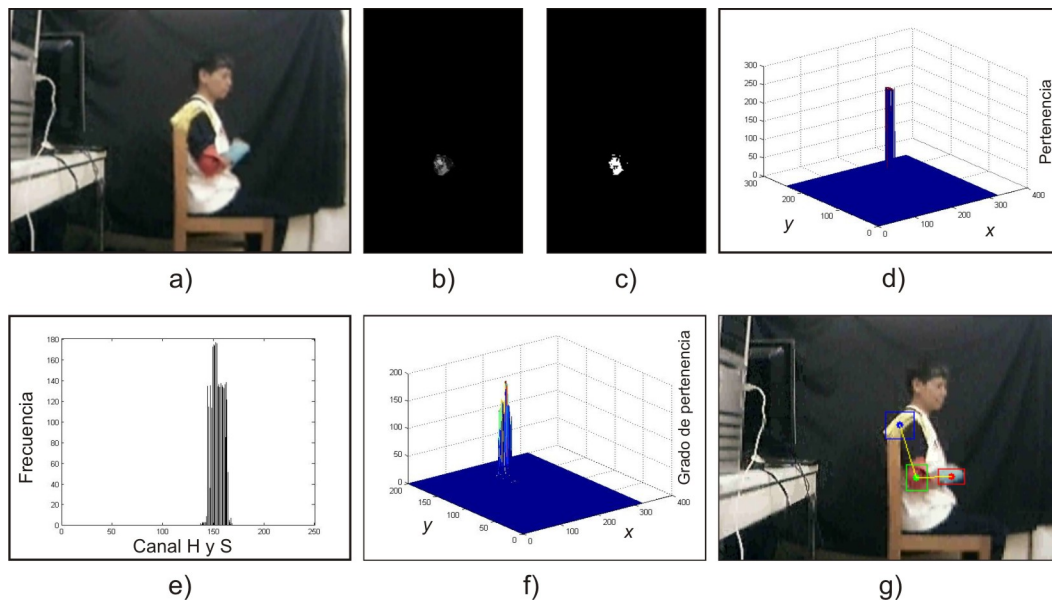


Figura 5.5: *Segmentación: marca color rojo. a)Imagen original, b)Imagen que muestra el grado de pertenencia de cada pixel al color rojo deseado, c)Imagen binaria que muestra la pertenencia de cada pixel al color rojo deseado, d)Histograma binario que muestra la pertenencia de cada pixel al color rojo deseado, e)Histograma de la distribución de probabilidades del color rojo deseado, canal H y S, f)Histograma que muestra el grado de pertenencia de cada pixel al color rojo deseado, g)Imagen Segmentada.*

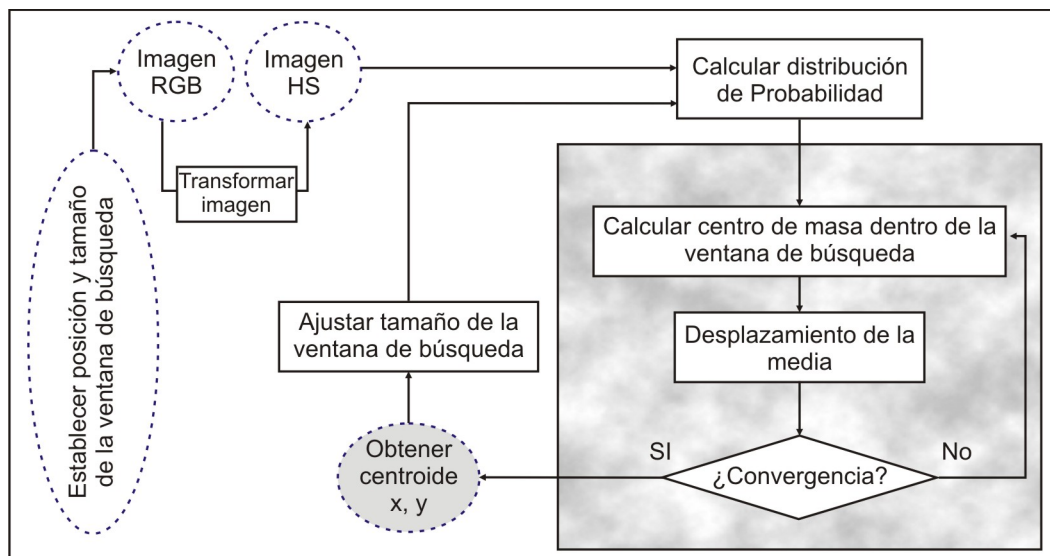


Figura 5.6: Diagrama de flujo del algoritmo camshift [7].

Los momentos de orden uno, m_{01} y m_{10} , junto con m_{00} , determinan el llamado centro de gravedad del objeto:

$$X_c = \frac{m_{10}}{m_{00}}, Y_c = \frac{m_{01}}{m_{00}} \quad (5.10)$$

Este procedimiento, nos proporciona las coordenadas de la ventana que será utilizada en la fase de seguimiento, con el fin de reducir el cálculo, explicado en la siguiente sección. La Figura 5.5g muestra las posiciones iniciales de las articulaciones del brazo, encerradas en rectángulos. El sistema utiliza los primeros 10 cuadros de la secuencia de imágenes para detectar las marcas.

5.2.3. Seguimiento 2D

Después de la localización de las marcas de color, el siguiente paso es su seguimiento a través de una secuencia de imágenes. El problema de seguimiento puede definirse como sigue: dada la localización de un objeto en un tiempo t , determinar la posición más probable de ese objeto en un tiempo $t + \Delta T$ [14].

Existen dos categorías de sistemas para resolver el problema de seguimiento: *basado en modelos* y *basado en características* [52]. El seguimiento basado en modelos se basa en la idea de aproximar un objeto conocido a priori -o modelo- con el objeto a seguir presente en la imagen. La desventaja de este enfoque es su alto costo computacional. El seguimiento basado en características se enfoca en detectar alguna característica del objeto a seguir (textura, borde, color, etc), siendo de bajo costo computacional, pero menos robusto ante problemas de ruido como la oclusión o la distracción de objetos con características semejantes. En este trabajo se utiliza el algoritmo *Camshift*, propuesto por Bradski [7], que se basa en el color como característica discriminante.

El algoritmo *Camshift* [7] se basa en ideas de estadística robusta y distribuciones de probabilidad, que ascendiendo en gradientes de densidad, permite encontrar la moda de distribuciones de probabilidad. Los estadísticos robustos son aquellos que tienden a ignorar los datos atípicos en la muestra (puntos que caen lejos de la región de interés), por lo tanto ayudan a compensar el ruido y las distracciones en los datos de visión (objetos que no son el seguido). De esta forma podemos encontrar la zona con mayor densidad de puntos pertenecientes a los datos, y dado que los datos son los pixeles que pertenecen a una marca de color, encontrar la mayor mancha en la imagen del color del objeto y sus dimensiones.

Camshift ajusta el tamaño de la ventana de búsqueda en el transcurso de sus operaciones. El tamaño inicial de la ventana de búsqueda se puede establecer en cualquier valor razonable. Para distribuciones discretas (datos digitales), el tamaño mínimo de la ventana es 3. En lugar de un conjunto de tamaños o una adaptación externa del tamaño de la ventana, *Camshift* confía en la información del momento de orden 0 para adaptar continuamente el tamaño de la ventana en cada cuadro de video. El momento de orden 0 representa el área de la distribución bajo la ventana de búsqueda. El tamaño de la ventana se establece como una

función del momento de orden 0. El algoritmo *Camshift* calcula entonces utilizando cualquier tamaño de ventana que no sea nulo (mayor o igual que 3 si la distribución es discreta). El algoritmo *Camshift* realiza los siguientes pasos (Figura 5.6):

1. Elegir el tamaño de la ventana de búsqueda. Este tamaño es determinado de forma heurística en relación de la distancia de las marcas con la cámara de video.
2. Elegir la posición inicial de la ventana de búsqueda.
3. Desplazar la media (una o varias iteraciones), guardar el momento de orden 0. La media se aplica a una imagen en escala de grises, en nuestro caso, obtenida por medio del método *Retroproyección de Histograma* [47]. Los momentos de orden 0 y 1 se obtienen por medio de las ecuaciones 5.7 y 5.10.
4. Establecer el tamaño de la ventana de búsqueda igual a una función del momento de orden 0 encontrado en el paso 3.
5. Repetir los pasos 3 a 4 hasta la convergencia (la posición de la media se mueve menos que un umbral predeterminado). En este caso, no deberá sobrepasar 10 iteraciones, o en caso contrario, el desplazamiento de los centroides en los ejes x y y , no exceda el tamaño de la ventana de búsqueda.

Cuando sigue un objeto *Camshift*, trabaja con una imagen de la distribución de probabilidad del color. *Camshift* calcula el centroide de la distribución de probabilidad del color dentro de la ventana de búsqueda, vuelve a centrar la ventana, y calcula el área para obtener el siguiente tamaño de ventana. De esta forma no necesitamos calcular la distribución de probabilidad del color en la imagen completa, sino que basta con restringir el cálculo de la distribución a una pequeña

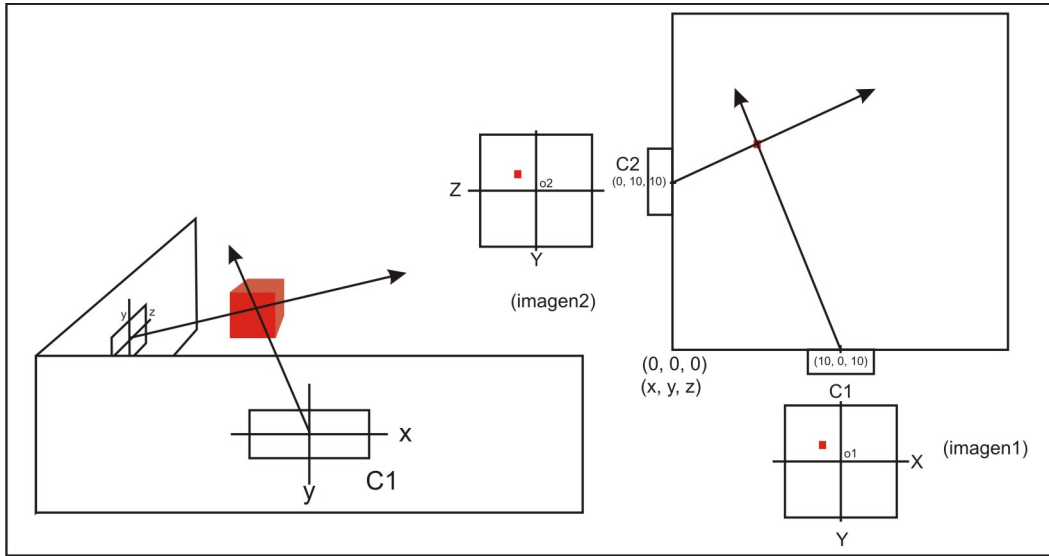


Figura 5.7: Líneas de proyección.

región de la imagen que envuelve a la ventana de búsqueda actual de *Camshift*. Esto evita gran cantidad de cálculos cuando el color del objeto no domina la imagen. De esta forma, podemos obtener las posiciones de las marcas de color de cada articulación del brazo. El proceso se aplica en las secuencias de imágenes capturadas por las dos cámaras de video.

5.2.4. Reconstrucción 3D

Después de obtener las posiciones de cada marca en el plano 2D, el siguiente paso es la reconstrucción 3D para obtener las coordenadas x, y, z de cada una de las articulaciones del brazo. Al tener dos cámaras se puede realizar un proceso de triangulación detectando en ambas imágenes algún punto de interés de los objetos buscados.

El método utilizado en este proyecto de tesis para estimar las posiciones de las articulaciones en 3D, es el enfoque basado en *múltiples cámaras* [1], en nuestro

caso dos *Webcams*, donde cada cámara obtiene la posición 2D del objeto en el plano de la imagen; es decir, el centroide de la ventana de seguimiento de las marcas de color, y los combina con los datos de las demás cámaras.

Este método, esta basado en el de Gonzalez y Cañas [19], el cual consiste en proyectar las líneas de visión de dos cámaras al objeto buscado, según un modelo geométrico. La *Figura 5.7*, ilustra la posición tridimensional del objeto en las que se intersectan las rectas de proyección, o en su defecto su punto de cruce más cercano. El método asume que las cámaras se encuentran calibradas, y se conoce su posición y orientación en el sistema de referencia del mundo.

Una vez obtenidas las coordenadas de cada marca de color en cada imagen, para obtener su posición 3D deshacemos la proyección de cada una de las cámaras, reconstruyendo la línea de visión⁵. Esta línea contiene todos los puntos tridimensionales que se proyectan en esas coordenadas en la imagen.

Reconstruir la línea de visión buscada consistirá en invertir el proceso de proyección y la rotación/traslación, para esto, es necesario conocer los parámetros *intrínsecos* y *extrínsecos*, vistos en la *sección 5.2.1*. Para reconstruir la línea de visión es necesario conocer dos puntos por los que pase. El primero será el *centro óptico* de la cámara, que es igual al vector t de la *ecuación 5.3*. Para el otro punto tomamos el centro de masa del objeto (marca de color), obtenido en la fase de seguimiento 2D (u, v) . Esto es, tomando en cuenta la relación que existe entre las coordenadas de la imagen en píxeles, y las espaciales:

$$\frac{1}{Z} \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = A \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} f_x & 0 & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} \quad (5.11)$$

donde A es la matriz que define los parámetros intrínsecos, tomada de la *ecuación*

⁵La *línea de visión* es la proyección de una línea a partir del centro de la imagen al centro de masa del objeto

5.2. Queda la expresión:

$$\begin{aligned} u &= \frac{X}{Z} f_x + c_x \\ v &= \frac{Y}{Z} f_y + c_y \end{aligned} \quad (5.12)$$

despejamos (u, v) (la posición 2D del objeto), y para simplificar puede tomarse $z = 1$. De este modo se obtiene la expresión:

$$\begin{aligned} x &= \frac{(u - c_x)}{f_x} \\ y &= \frac{(v - c_y)}{f_y} \\ z &= 1 \end{aligned} \quad (5.13)$$

Este nuevo punto define la posición del objeto visto desde la cámara (P^C), por lo que se debe considerar la nueva posición y orientación de la cámara. Tomando los *parámetros extrínsecos* de la ecuación 5.3, y la ecuación 5.4, que relaciona un punto en el marco de referencia de la cámara con respecto al marco de referencia del mundo, el nuevo punto 3D del objeto queda definido por la siguiente expresión:

$$P^W = P^C R^{-1} + t \quad (5.14)$$

Una vez calculadas las rectas de proyección de cada cámara con respecto al objeto, el corte de las dos líneas de visión, o en su defecto, el punto más cercano de cruce [6], marcará la estimación de la posición del objeto en tres dimensiones.

El procedimiento se aplica para cada marca de color, esto es, se proyectan tres líneas de visión de cada cámara a las marcas de color adheridas al brazo. De esta forma, se obtiene la posición tridimensional de cada marca respecto al sistema de referencia del mundo, lo cual nos permite tener una representación del brazo por medio de una estructura de alambre (*stick*), al trazar líneas de unión entre cada articulación.

5.3. 2da. Etapa: Modelo de movimiento

Una vez que se determinó el proceso de segmentación y seguimiento del brazo en un espacio tridimensional, el siguiente paso es determinar qué características son convenientes para analizar el movimiento del brazo. Pero para saber que características utilizar hay que definir antes los ejercicios terapéuticos a calificar.

5.3.1. Selección de los gestos terapéuticos

El término *gesto* se puede definir como un movimiento o conjunto de movimientos ya sea del rostro, de las manos o de cualquier parte del cuerpo, con los cuales, se pueden expresar señales, instrucciones, estados de ánimo, etc. Los gestos se pueden clasificar en dos clases debido a su número de movimientos: gestos estáticos y gestos dinámicos. Los *gestos estáticos* solo requieren de alguna postura del cuerpo. Por otro lado, los *gestos dinámicos* son aquellos que involucran varios movimientos en su realización. Aunque el gesto es visto como un símbolo o un lenguaje, también puede ser visto como una sucesión de movimientos corporales de una rutina de ejercicios. De tal forma, algunas personas o sistemas inteligentes pueden observar el gesto y analizarlo con la finalidad de calificar su correcta ejecución.

En este proyecto de tesis, el gesto es considerado como un conjunto de movimientos del brazo del paciente, que en su conjunto forman un ejercicio terapéutico en el proceso de rehabilitación. En este proyecto de tesis se propuso como objetivo particular la definición de 3 gestos terapéuticos. Estos gestos fueron definidos por un especialista en rehabilitación del INNN⁶. Los gestos son:

1. *Flexión-extensión*: Inicialmente el brazo está extendido y se flexiona en dirección paralela al torso, y se regresa a la posición inicial (ver *Figura 5.8a*).

⁶Instituto Nacional de Neurología y Neurocirugía

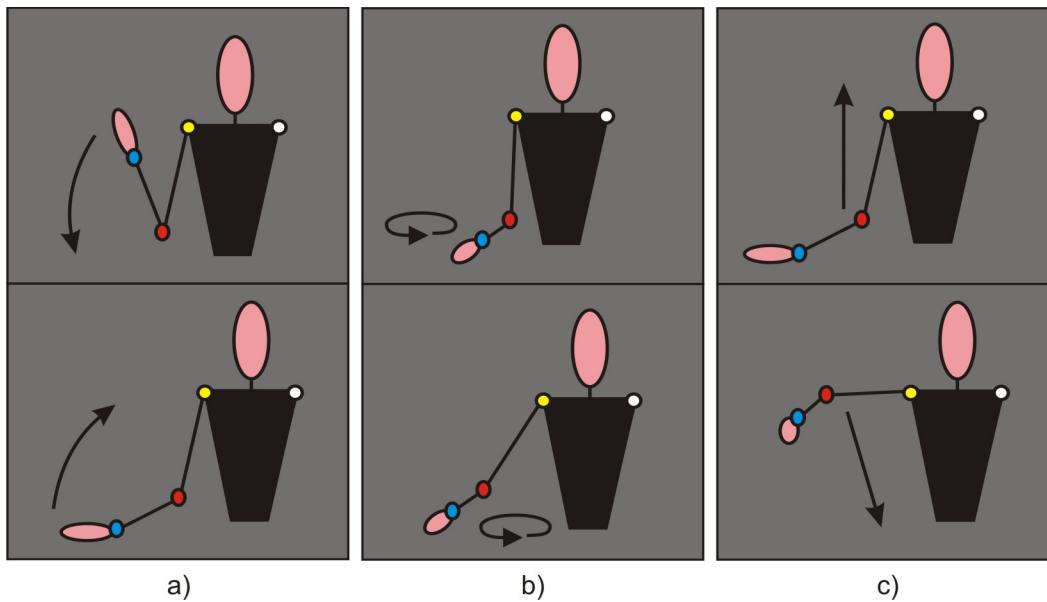


Figura 5.8: Selección de gestos. a) Flexión - extensión, b) Circular, c) Abducción - aducción.

Este ejercicio terapéutico fortalece el bíceps y aumenta la movilidad de los músculos flexores y extensores del brazo. Beneficia el poder de levantar objetos.

2. *Circular*: En el inicio el brazo está extendido frontalmente con la palma de la mano empuñada, y se mueve la mano en forma circular en sentido inverso a las manecillas del reloj, regresando a su posición inicial (ver *Figura 5.8b*). Este ejercicio terapéutico aumenta la coordinación del movimiento de los músculos flexores, extensores, abductores y aductores. Beneficia el realizar la limpieza de objetos.
3. *Abducción*: Inicialmente el brazo está flexionado hacia arriba paralelo a las piernas, y se mueve el brazo hacia arriba, hasta el punto máximo que pueda alcanzar, regresando a su posición inicial (ver *Figura 5.8c*). Este ejercicio

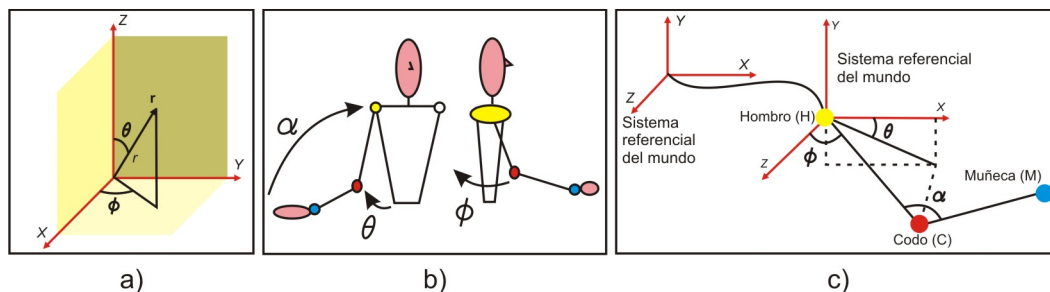


Figura 5.9: *Coordenadas esféricas, a) Obtención de las coordenadas esféricas en un ambiente 3D, b) Obtención de las coordenadas esféricas en el cuerpo humano, c) La imagen ilustra la postura del brazo en un instante de tiempo. α indica el ángulo entre el brazo y el antebrazo, los ángulos θ y ϕ representan las coordenadas esféricas del codo tomando como referencia al hombro. h , m y c representan las coordenadas tridimensionales (X, Y, Z) del hombro, codo y muñeca, respectivamente.*

terapéutico aumenta la movilidad de los músculos abductores y aductores del brazo. Beneficia al paciente de poder vestirse por si solo.

5.3.2. Características del movimiento

El movimiento del brazo humano, como todo nuestro cuerpo, depende de las articulaciones por las cuales esta formado. El brazo se encuentra constituido por varias articulaciones; el hombro, el codo y la muñeca son las principales. Los gestos utilizados en esta tesis se basan en el movimiento de estas tres articulaciones. Para diferenciarlos visualmente, se colocaron tres bandas de colores distintos. Se seleccionaron tres colores que fueran fáciles de segmentar. Los colores que se seleccionaron son: el amarillo para el hombro, el rojo para el codo y el azul para la muñeca.

Las métodos más comunes para representar las características de movimiento están basados en la trayectoria descrita por el objeto a seguir. Sin embargo, debido

a que el brazo está representado por un conjunto de tres objetos, articulaciones (marcas de color), el uso de estas trayectorias para representar el movimiento provocaría errores al colocar al paciente en diferente posición, así como al calificar pacientes con características físicas diferentes, brazos mas largos, estatura, etc.

Una alternativa para modelar el movimiento del brazo, es la representación de la postura en cada instante de tiempo, por medio de ángulos entre las articulaciones y el torzo del usuario. Los ángulos son invariantes a translaciones y escalado, y son independientes de las características físicas de la persona.

En este proyecto de tesis se proponen tres grados de libertad para representar la postura del brazo (ver *Figura 5.9*). El primer ángulo (α) esta formado entre el brazo y el antebrazo. Para el segundo y tercer ángulo (θ, ϕ), se considera la posición del brazo respecto al torzo. Para esto, se toma como nuevo punto de referencia al hombro y se obtienen las coordenadas esféricas del codo.

Por lo tanto, el conjunto de características que describe la trayectoria del brazo al realizar los gestos, esta formado por $V = [\alpha, \theta, \phi]$. Las posiciones tridimensionales de cada articulación estan definidas como:

$$h = \begin{bmatrix} X_h \\ Y_h \\ Z_h \end{bmatrix}, c = \begin{bmatrix} X_c \\ Y_c \\ Z_c \end{bmatrix}, m = \begin{bmatrix} X_m \\ Y_m \\ Z_m \end{bmatrix} \quad (5.15)$$

donde h, m y c representan las posiciones 3D del hombro, codo y muñeca, respectivamente.

El ángulo entre el brazo y el antebrazo (α) se calcula a partir del ángulo que forman dos líneas en el espacio [2], dadas por las siguientes ecuaciones:

$$Q = M - C$$

$$R = C - H$$

donde H, C y M son las coordenadas del hombro, codo y muñeca (ver *Figura*

5.9c).

$$\begin{aligned}\alpha &= \cos^{-1} \left(\frac{\|R \cdot Q\|}{\|R\| \cdot \|Q\|} \right) \\ &= \cos^{-1} \left(\frac{X_R X_Q + Y_R Y_Q + Z_R Z_Q}{\sqrt{X_R^2 + Y_R^2 + Z_R^2} \sqrt{X_Q^2 + Y_Q^2 + Z_Q^2}} \right)\end{aligned}\quad (5.16)$$

Los ángulos entre el brazo y el torzo (θ, ϕ) , se obtienen por medio de las coordenadas esféricas del codo tomando como referencia el hombro. Para esto, se realiza una translación del origen del sistema de referencia del mundo a la posición del hombro, esto es:

$$d = C - H \quad (5.17)$$

donde el vector d representa la nueva posición del codo después de aplicar la translación. De esta forma, para convertir las coordenadas cartesianas a esféricas se aplican las siguientes ecuaciones [2]:

$$\begin{aligned}\rho &= \sqrt{X_d^2 + Y_d^2 + Z_d^2} \\ \theta &= \tan^{-1} \left(\frac{Y_d}{X_d} \right) \\ \phi &= \cos^{-1} \left(\frac{Z_d}{\rho} \right)\end{aligned}\quad (5.18)$$

donde θ es el ángulo que se forma al proyectarse el segmento de línea en el plano (X, Y) . Y ϕ es el ángulo que indica la apertura del segmento de línea respecto al eje Z (ver *Figura 5.9*). Finalmente, los tres ángulos son normalizados entre $[0.0 - 1.0]$, al dividirse entre 2π radianes:

$$\begin{aligned}\alpha^* &= \frac{\alpha}{2\pi} \\ \theta^* &= \frac{\theta}{2\pi} \\ \phi^* &= \frac{\phi}{2\pi}\end{aligned}\quad (5.19)$$

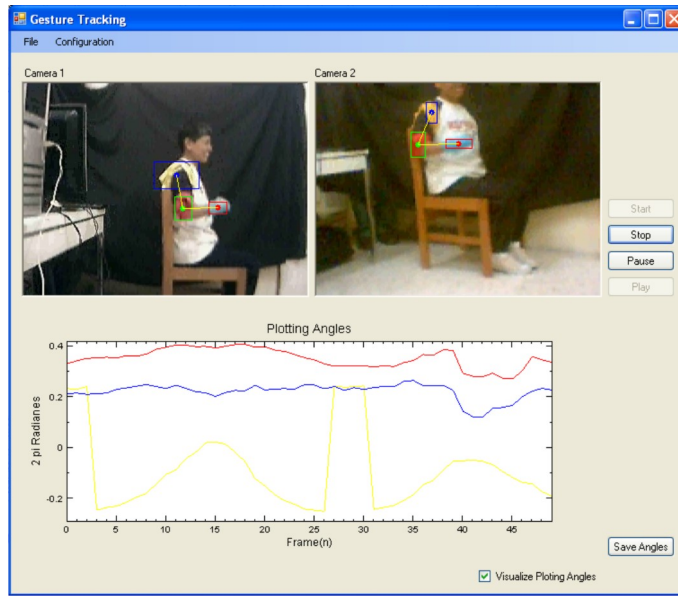


Figura 5.10: Obtención de las coordenadas esféricas por el sistema. En las dos imágenes se muestra el seguimiento de las marcas de color en las dos secuencias de video, capturadas por las dos webcams. En la gráfica de abajo se muestran las trayectorias de los tres ángulos: α , θ y ϕ .

De esta forma, se puede representar la trayectoria descrita por el movimiento del brazo al realizar un gesto, bajo el sistema de coordenadas α, θ, ϕ , como (ver Figura 5.10):

$$F = \{(\alpha_1^*, \theta_1^*, \phi_1^*), \dots, (\alpha_n^*, \theta_n^*, \phi_n^*)\} \quad (5.20)$$

donde n es la longitud de la trayectoria.

5.4. 3a. Etapa: Calificación

5.4.1. Modelos de Markov

Una vez logrado el seguimiento del brazo de una persona y obtener su estructura 3D, así como su extracción de características de movimiento, en nuestro caso sus coordenadas esféricas, el siguiente paso es modelar la trayectoria del brazo que describe un gesto. En este proyecto de tesis, el objetivo del análisis del movimiento es obtener la calificación de la realización de un ejercicio terapéutico. Para lograr esto se necesita comparar un gesto ejecutado *correctamente* por una persona sana, con otro ejecutado por un paciente que haya sufrido un derrame cerebral y que sufra hemiplejía. Para esto, se propone una métrica basada en modelos probabilísticos, por medio del uso de los Modelos Ocultos de Markov (MOM) [37], que como se expuso en el capítulo 4, es una técnica que permite una representación apropiada a secuencias de datos *espacio-temporales*. Al representar nuestros videos de la realización de los gestos como secuencias de datos dentro de un MOM, los datos pueden ser continuos o discretos [37]. En este trabajo de tesis se utilizan datos discretos, ya que es mucho más fácil representar las secuencias con datos discretos [37]. Por lo tanto el siguiente paso es discretizar nuestros datos con el fin de obtener símbolos discretos que serán utilizados por los MOM [37], en la fase del análisis del movimiento.

Discretización

Para la discretización, se utiliza el algoritmo *K-means* [18], el cual permite clasificar los datos en K clases o *clusters*, dentro de nuestro espacio de características (α, θ, ϕ) . En otras palabras, el algoritmo permite pasar del conjunto de datos continuos (que representan la trayectoria del brazo) a un conjunto de datos discretos

más compacto representado por K clases.

El algoritmo *K-means*, es un método sencillo, pero muy eficiente para realizar agrupamiento. Este algoritmo se basa en la distancia mínima entre el punto medio de cada clase (K) y los puntos de las características. A continuación se describe el funcionamiento del algoritmo *K-means*:

1. *Inicialización*. Establecido previamente el número de clases (K), se escogen al azar entre los elementos a agrupar K vectores, que representan los centroides de las K clases ($Z_i, 1 \leq i \leq K$).
2. *Clasificación*. Se distribuyen todas las muestras $\{X\}_{1 \leq i \leq n}$ entre las K clases C_i usando el clasificador de mínima distancia ($X \in C_i$ *si* $\|X - Z_i\| \leq \|X - Z_j\|, \forall i = 1, 2, \dots, K \ i \neq j$).
3. *Actualización*. Una vez redistribuidos los elementos a agrupar entre las diferentes clases, se actualizan los centroides de las clases por medio de calcular el punto medio de los vectores de entrenamiento en cada grupo.
4. *Finalizar*. Se termina el agrupamiento cuando no cambian los centroides, de lo contrario se continúa en el paso 2.

A este proceso de agrupamiento también se le conoce como *Cuantización Vectorial* (VQ), el cual nos da como resultado un *Libro de códigos* (*Codebook*), que contiene un número de símbolos y las coordenadas del centroide de cada clase. El proceso del reconocimiento del VQ es el siguiente. Una vez obtenido el libro de códigos por medio del algoritmo *K-means*, se introduce un vector (α, θ, ϕ) representando los datos de un gesto, se calcula la distancia *Euclídea* con cada una de las clases⁷, y se determina su pertenencia a alguna de las clases utilizando la

⁷En este trabajo de tesis se utilizaron *codebooks* de 16, 32, 64 y 128 clases.

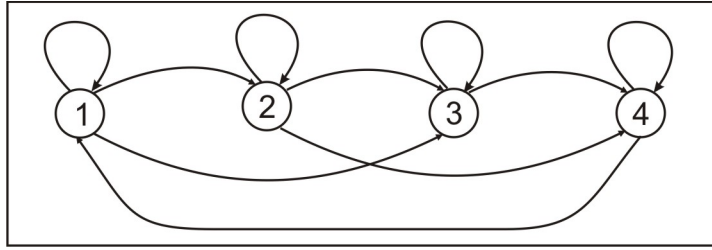


Figura 5.11: *Topología Bakis Continua de los MOM utilizados en este proyecto de tesis.*

distancia mínima. Así, obtenemos los símbolos de observación utilizados para entrenar los Modelos Ocultos de Markov [37] ($F = \{(\alpha_1, \theta_1, \phi_1), \dots, (\alpha_n, \theta_n, \phi_n)\} \xrightarrow{VQ} O = \{o_1, \dots, o_n\}$).

Topología de los MOM empleada

El siguiente paso, es escoger una topología para los MOM que sea adecuada para el análisis de gestos dinámicos. La topología escogida es la estructura *Bakis* (Izquierda-Derecha Continua) (ver sección 4.4), la cual puede modelar fácilmente secuencias de observaciones cuya estructura cambia en el tiempo (ver Figura 5.11). En estos modelos, la transición entre estados ocurre de izquierda a derecha, regresando al estado inicial, conforme pasa el tiempo.

Entrenamiento de los MOM

Cada uno de los gestos a analizar (*flexión-extensión*, *abducción-aducción* y *circular*), está representado por un MOM. El método de entrenamiento se aplica a dos grupos: *normal* y *paciente*. El método para entrenar a cada MOM es el siguiente:

1. Se obtuvieron 2 grupos de distintas secuencias de observaciones de movimiento correspondientes a cada gesto: *normal* y *paciente*.

- *Normal*. Se realizaron varias secuencias de los 3 gestos bien realizados por personas sanas, que fueron acreditadas por un especialista en rehabilitación del INNN, de las cuales se obtuvieron secuencias de video. Las cuales se utilizaron para los ejercicios corrector y la persona sana.
 - *Paciente*. Se realizaron varias secuencias de los 3 gestos realizados por 10 diferentes pacientes del INNN, que sufren hemiplejia por causa de un e.v.c.h y e.v.c.i (ver sección 2.1), de las cuales se obtuvieron secuencias de video. El grupo *paciente*, esta conformado por los gestos realizados por el paciente a analizar, comparándose posteriormente cada uno de los MOM con los del grupo *normal*
2. Se estimaron los parámetros correspondientes a cada uno de los tres MOM, para cada uno de los gestos bien realizados y para cada uno de los tres gestos realizados por los 10 pacientes hemiplejicos, así como los correctos y además los realizados por una persona sana, 36 MOMs en total. Se empleó el algoritmo *Baum-Welch*, detallado en la sección 4.3.3. Guardándose en un base de datos los 36 MOM, 1 para cada gesto y sus correspondientes secuencias de observaciones. La definición de los MOM es la siguiente:

- *Normal*. Flexión-extensión = λ_0^f , circular = λ_0^c , abducción-aducción = λ_0^a .
- *Paciente*. Flexión-extensión = $\lambda_1^f, \lambda_2^f, \dots, \lambda_{10}^f$, circular = $\lambda_1^c, \lambda_2^c, \dots, \lambda_{10}^c$, abducción-aducción = $\lambda_1^a, \lambda_2^a, \dots, \lambda_{10}^a$.

5.4.2. Métricas

Una vez obtenidos los modelos de cada uno de los gestos bien realizados así como los modelos de los gestos realizados por los pacientes, lo siguiente es obtener una

calificación de cada uno de los gestos realizados por los pacientes. El primer paso es generar una calificación automática para que en el siguiente paso se comparen estos resultados con métricas utilizadas por los especialistas en rehabilitación.

Comparación de gestos empleando MOM

Una vez entrenados los MOM para cada gesto del grupo *normal* para entrenamiento y del grupo *paciente*, as como el de la persona sana (36 en total), se emplean cada una de las métricas basada en la distancia detalladas en la *sección* 4.5. Para esto, para cada una de las secuencias de los gestos realizados por los pacientes, se toma su MOM así como el MOM del gesto bien realizado correspondiente al mismo gesto, y se obtiene su distancia.

- *Levinson*. Utilizando la distancia de la *ecuación* 4.28:

$$D^e(\lambda_1, \lambda_2) = \sqrt{\frac{1}{N} \sum_{j=1}^N \sum_{k=1}^M \|b_{jk}^1 - b_{jk}^2\|^2}$$

se obtiene la distancia de cada uno de los 30 MOM de los 10 pacientes entre los 3 MOM de los gestos bien realizados.

- *Porikli*. Utilizando la distancia de la *ecuación* 4.31:

$$D^{porikli}(T_1, T_2) = |P(O_1|\lambda_1) + P(O_2|\lambda_2) - P(O_1|\lambda_2) - P(O_2|\lambda_1)|$$

se obtiene la distancia de cada uno de los 30 MOM de los 10 pacientes entre los 3 MOM de los gestos bien realizados.

- *Kullback-Leibler*. Utilizando la distancia de la *ecuación* 4.29:

$$D^k(\lambda_1, \lambda_2) = \frac{1}{T} [\log(P[O^1|\lambda_1]) - \log(P[O^1|\lambda_2])]$$

se obtiene la distancia de cada uno de los 30 MOM de los 10 pacientes entre los 3 MOM de los gestos bien realizados.

Valoración del paciente

Al obtener una calificación automática, es necesario probar su validez con métricas utilizadas en el proceso de rehabilitación. En la *sección 2.5*, se mencionaron 3 escalas de la valoración de la discapacidad motriz ampliamente utilizadas, de las cuales, en este proyecto de tesis se utilizarán la *Fugl-meyer* e *índice de motricidad*.

5.5. Resumen

En este capítulo se presentó el desarrollo del sistema para la calificación automática de gestos terapéuticos utilizados en rehabilitación. El sistema general se divide en tres módulos principales: *a) Sistema visual*, *b) Modelo de movimiento* y *c) Calificación del movimiento*. El sistema visual localiza y sigue las articulaciones del brazo de una persona (hombro, codo y muñeca), empleando cintas de color (amarillo, rojo, azul) como marcas distintivas. El seguimiento se aplica en las secuencias de imágenes capturadas por dos *webcams* de video, obteniéndose las posiciones 2D de cada articulación en el plano de la imagen en cada *webcam*. Posteriormente, se estima la posición 3D de cada marca por medio de la intersección de líneas de visión (de cada cámara a las marcas), reconstruyendo el brazo en una estructura de alambre (*stick*) 3D. El Modelo de movimiento obtiene las trayectorias del movimiento del brazo. Dichas trayectorias se modelan por medio de ángulos entre el brazo y el antebrazo, y entre el brazo y el torzo del usuario, extraídas de la estructura de alambre 3D del brazo. Por último se califica el movimiento utilizando MOM. Para esto se discretizan los datos de las trayectorias a ser usados por los MOM. Se entrenan los MOM correspondientes a los tres gestos, empleando las secuencias de observaciones de movimiento obtenidas. Se crean dos grupos de datos de entrenamiento, uno para los gestos *correctos*

realizados por personas sanas, y otro para los gestos realizados por el paciente. Posteriormente, se obtiene una métrica entre dos MOM empleando la distancia de *Levinson*, *Kullback-Leiblar* y *Porikli*. Finalmente, el especialista en rehabilitación obtiene una valoración del paciente utilizando las métricas de *Fugl-meyer* e *índice de motricidad*. Los resultados experimentales y su análisis, se detallan en el siguiente capítulo.

Capítulo 6

Experimentos y resultados

En este capítulo se muestran los resultados de comparar movimientos del brazo humano realizados por pacientes hemipléjicos, con movimientos realizados por una persona sana. Estos movimientos, tanto el realizado por la persona con hemiplejia como el realizado por una persona sana, son convertidos en modelos a partir de la extracción de características de las secuencias de movimiento. La comparación se realiza entre dos modelos, el de la persona con hemiplejia, y el de una persona sana. Esto es con el fin de obtener la semejanza entre los dos modelos, y así poder generar una calificación. La comparación de los modelos se realiza utilizando las métricas para comparar modelos ocultos de Markov: *Levinson*, *Porikli*, y *KullbackLeibler*.

Los movimientos del brazo humano fueron definidos por un especialista en rehabilitación del INNN. Los gestos son: *Flexión-extensión*: Inicialmente el brazo extendido y flexionarlo en dirección paralela al torzo, y regresarlo a la posición inicial. *Circular*: El brazo extendido frontalmente con la palma de la mano empuñada, y mueve la mano en forma circular en sentido inverso a las manecillas del reloj, regresando a su posición inicial. *Abducción*: Inicialmente el brazo flexionado hacia arriba paralelo a las piernas, y mueve el brazo hacia arriba, hasta el punto

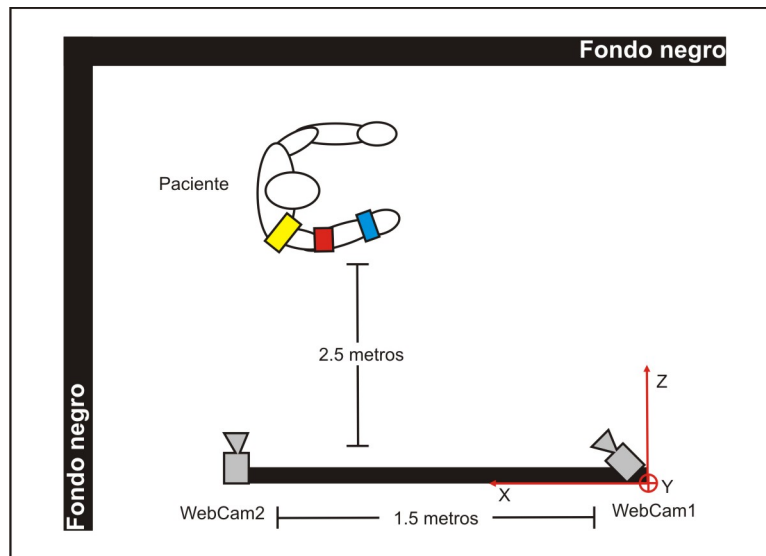


Figura 6.1: *Ambiente. Forma de colocar las cámaras y el paciente.*

máximo que pueda alcanzar, regresando a su posición inicial.

Una vez obtenidas las distancias de los modelos realizados por los pacientes hemipléjicos respecto al modelo del movimiento realizado por una persona sana, se prosigue a comparar estos resultados con las valoraciones de movilidad realizadas por un especialista en rehabilitación.

En la primera sección se muestran los resultados del seguimiento, en la segunda sección se muestran los resultados de la obtención de las características del movimiento, en las siguientes secciones se muestran los resultados obtenidos de comparar los modelos con las diferentes métricas, así como su comparación con la valoración realizada por un especialista de la rehabilitación del INNN.

6.1. Seguimiento

El primer paso de la experimentación es obtener la posición de las articulaciones del brazo. Este primer paso está definido por el seguimiento. Como se mencionó en

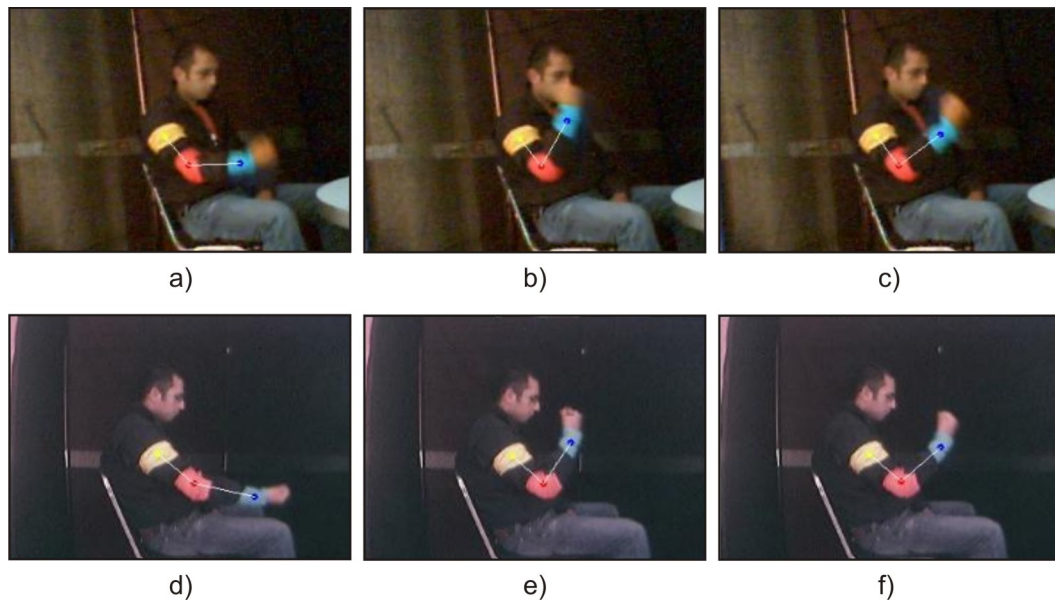


Figura 6.2: *Seguimiento, a), b) y c) Seguimiento de las tres marcas de color en la webcam1 y d), e) y f) Seguimiento de las tres marcas de color en la webcam2.*

el capítulo 3, el método de seguimiento utilizado en este proyecto de tesis se basa en la técnica de visión con marcas usando *webcams*. El proceso de seguimiento o sistema visual esta formado por tres etapas: segmentación, seguimiento 2D y reconstrucción 3D. La segmentación se encarga de la localización de las marcas de color dentro de una imagen. Las articulaciones están marcadas con cintas de color: amarillo - hombro, rojo - codo y azul - muñeca, para poder realizar una segmentación mas exacta. Para la segmentación de las marcas, se aplica información de color (histogramas de color). El seguimiento 2D se encarga del seguimiento de un objeto a través de una secuencia de imágenes de video. Y la reconstrucción 3D obtiene las coordenadas 3D del objeto en seguimiento. Se aplica un algoritmo de intersección de líneas de visión al objeto en seguimiento de ambas cámaras para obtener una estimación de la posición 3D. Para esto es necesario realizar una calibración de las cámaras para obtener los parámetros intrínsecos y los parámetros

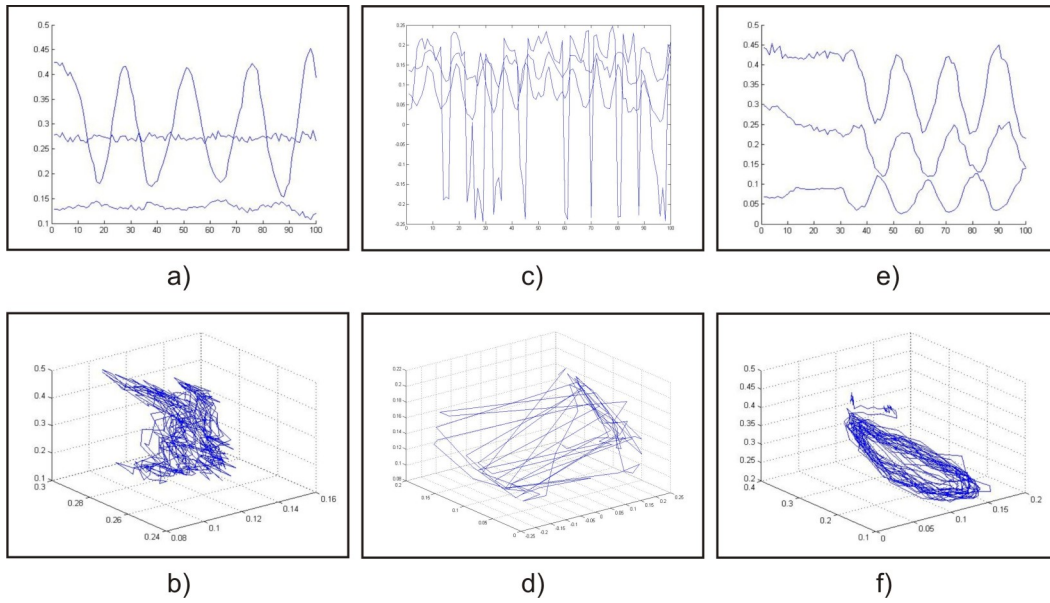


Figura 6.3: Gráficas que muestran el movimiento del brazo a partir de la obtención de los ángulos generados por el movimiento del brazo, a) Flexión(3 ángulos), b) Flexión(3d), c) Abducción(3 ángulos), d) Abducción(3d), e) Circular(3 ángulos), f) Circular(3d).

extrínsecos que se obtienen manualmente. La Figura 6.1, muestra la colocación de las cámaras y del paciente al momento de tomar los videos.

Las Figuras 6.2a, 6.2b y 6.2c, muestran el seguimiento de las tres marcas de color en la *webcam1* y las Figuras 6.2d, 6.2e y 6.2f, muestran el seguimiento de las tres marcas de color en la *webcam2*.

6.2. Movimiento humano

Al obtener los ángulos que se forman entre las articulaciones del brazo al realizar los ejercicios terapéuticos se obtienen 3 trayectorias, que corresponden a cada uno de los 3 ángulos, que se pueden apreciar en las Figuras 6.3a, 6.3c y 6.3e. En las Figuras 6.3b, 6.3d y 6.3f, se pueden apreciar las gráficas generadas en un espacio

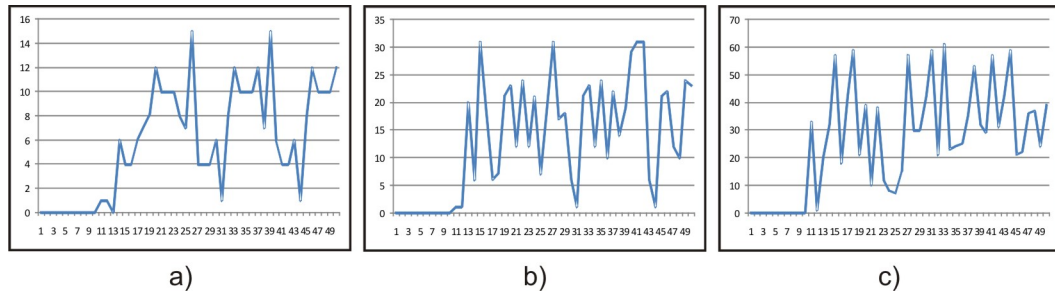


Figura 6.4: Gráficas que muestran secuencias de observaciones discretas, a)16 clases, b)32 clases y c)64 clases.

3D por una sola trayectoria formada por la obtención de los 3 ángulos.

En la *Figura 6.3a* que muestra el ejercicio de flexión, se puede apreciar como el ángulo formado entre el codo y la muñeca oscila aumentando y disminuyendo, mientras que los otros dos ángulos permanecen casi sin cambio alguno. En la *Figura 6.3c* que muestra el ejercicio de abducción, se puede apreciar como el ángulo formado entre el codo y el cuerpo va aumentando y disminuyendo, mientras que los otros dos ángulos permanecen casi sin cambio alguno. El que sólo un ángulo tenga cambios, no sucede en el ejercicio circular, ya que como muestra la *Figura 6.3e*, los tres ángulos oscilan aumentando y disminuyendo.

Los 3 ángulos (α, θ, ϕ) son las características necesarias para generar los modelos, pero para esto es necesario discretizarlos. La discretización se realizó con el algoritmo *K - means*. En este proyecto de tesis se discretizaron las características en 16, 32, 64, 128 y 256 clases, las *Figuras 6.4a*, *6.4b* y *6.4c*, muestran solamente las observaciones discretas con 16, 32 y 64 clases, respectivamente, del ejercicio de flexión.

6.3. Calificación

La calificación de los gestos terapéuticos se basa en obtener la distancia entre dos modelos, uno realizado correctamente y el otro por una persona con hemiplejía o persona sana. Cada paciente, 10 en total, así como la persona que realiza el gesto correctamente y la persona sana, realizan los tres gestos, flexión, circular y abducción, los cuales fueron grabados en secuencias de video. Fueron 100 repeticiones de cada gesto que realizaron los pacientes, la persona sana y la persona que realizó el ejercicio correcto, un total de 3,600, los cuales fueron grabados en videos. Al mismo tiempo en que fueron tomados los videos, un especialista en rehabilitación del INNN aplicó las escalas *Fugl-meyer* e *índice de motricidad*. De cada gesto realizado se genera un MOM. Los MOM de los pacientes se comparan con el MOM de la persona sana, utilizando las métricas para comparar modelos ocultos de Markov: *Levinson, Porikli, y Kullback Leibler*. Posteriormente, estos resultados serán comparados con las escalas *Fugl-meyer* e *índice de motricidad*. A continuación se muestran los resultados de aplicar *Levinson, Porikli, y Kullback Leibler* a los MOM.

6.3.1. Levinson

Levinson se basa en obtener la distancia euclidiana de dos modelos, basándose en sus matrices de probabilidad de observación de los dos MOM que se desean comparar. Las *Tablas* 6.1, 6.2 y 6.3, muestran los resultados de aplicar la métrica *Levinson* (ecuación: 4.28). Las columnas *Paciente(n)*, muestra el resultado de comparar la trayectoria del *Paciente(n)* con la trayectoria de la persona sana (correcto) al realizar el ejercicio de flexión, circular y abducción, respectivamente. Las columnas 16, 32, 64, 128 y 256 indican el número de clases al momento de discretizar las características del movimiento.

Flexión Levinson					
	16	32	64	128	256
Correcto	0	0	0	0	0
Sano	0.0151	0.0341	0.03554	0.0251	0.0232
Paciente 1	0.0353	0.0601	0.0329	0.0353	0.1814
Paciente 2	0.0229	0.0551	0.0353	0.0601	0.0551
Paciente 3	0.0908	0.1814	0.1360	0.0501	0.0353
Paciente 4	0.0242	0.0551	0.0651	0.0353	0.1716
Paciente 5	0.1814	0.1813	0.2569	0.0360	0.0908
Paciente 6	0.1585	0.0551	0.0601	0.0601	0.0501
Paciente 7	0.0945	0.0329	0.0601	0.0353	0.1814
Paciente 8	0.0569	0.0229	0.0908	0.0601	0.0329
Paciente 9	0.0360	0.0146	0.3908	0.0551	0.0272
Paciente 10	0.0260	0.0945	0.0014	0.0908	0.0242

Tabla 6.1: Muestra los resultados obtenidos al aplicar la métrica Levinson a los modelos del gesto flexión de cada uno de los pacientes y la persona sana, con el modelo del gesto flexión realizado correctamente. Utilizando observaciones discretizadas en 16, 32, 64, 128 y 256 clases.

Circular Levinson					
	16	32	64	128	256
Correcto	0	0	0	0	0
Sano	0.0253	0.0542	0.0096	0.0632	0.0674
Paciente 1	0.0887	0.0178	0.0029	0.0195	0.0029
Paciente 2	0.0824	0.0491	0.0153	0.0817	0.0449
Paciente 3	0.0695	0.0465	0.2360	0.0166	0.0504
Paciente 4	0.0520	0.0157	0.0566	0.0743	0.0680
Paciente 5	0.0187	0.0878	0.2746	0.0745	0.0846
Paciente 6	0.0344	0.0273	0.0698	0.0423	0.0529
Paciente 7	0.0344	0.0536	0.1285	0.0863	0.0209
Paciente 8	0.0820	0.0936	0.0907	0.0089	0.0283
Paciente 9	0.0509	0.0187	0.1908	0.4790	0.0358
Paciente 10	0.0003	0.0327	0.6801	0.0335	0.0235

Tabla 6.2: Muestra los resultados obtenidos al aplicar la métrica Levinson a los modelos del gesto circular de cada uno de los pacientes y de la persona sana, con el modelo del gesto circular realizado correctamente. Utilizando observaciones discretizadas en 16, 32, 64 128 y 256 clases.

Abducción Levinson					
	16	32	64	128	256
Correcto	0	0	0	0	0
Sano	0.0353	0.0601	0.0090	0.0421	0.1814
Paciente 1	0.0618	0.0592	0.0010	0.0618	0.0219
Paciente 2	0.0189	0.0667	0.0150	0.0818	0.0700
Paciente 3	0.0268	0.0570	0.4309	0.0936	0.0828
Paciente 4	0.0297	0.0085	0.1094	0.0454	0.0694
Paciente 5	0.0337	0.0073	0.2987	0.0252	0.0436
Paciente 6	0.0310	0.0193	0.0601	0.0603	0.0986
Paciente 7	0.0331	0.0592	0.0883	0.0445	0.0346
Paciente 8	0.0444	0.0352	0.0613	0.0226	0.0074
Paciente 9	0.0823	0.0816	0.2837	0.0797	0.0452
Paciente 10	5.72E-05	0.0257	0.7653	0.0268	0.0570

Tabla 6.3: *Muestra los resultados obtenidos al aplicar la métrica Levinson a los modelos del gesto abducción de cada uno de los pacientes y de la persona sana, con el modelo del gesto abducción realizado correctamente. Utilizando observaciones discretizadas en 16, 32, 64 128 y 256 clases.*

Flexión Kullback - Leibler					
	16	32	64	128	256
Correcto	0	0	0	0	0
Sano	0	0	0	0	0
Paciente 1	0	0	0	0	0
Paciente 2	0	0	0	0	0
Paciente 3	Infinito	Infinito	Infinito	0.00023	Infinito
Paciente 4	Infinito	Infinito	Infinito	0.00012	Infinito
Paciente 5	Infinito	Infinito	0.00084	0.0003	Infinito
Paciente 6	Infinito	Infinito	Infinito	0.00051	Infinito
Paciente 7	Infinito	Infinito	Infinito	0.0003	Infinito
Paciente 8	Infinito	Infinito	0.00073	0.00242	Infinito
Paciente 9	Infinito	Infinito	Infinito	0.00439	Infinito
Paciente 10	Infinito	Infinito	Infinito	0.002808	Infinito

Tabla 6.4: Muestra los resultados obtenidos al aplicar la métrica Kullback - Leibler a los modelos del gesto flexión de cada uno de los pacientes y de la persona sana, con el modelo del gesto flexión realizado correctamente. Utilizando observaciones discretizadas en 16, 32, 64 128 y 256 clases.

6.3.2. Kullback - Leibler

Kullback - Leibler se basa en obtener la distancia de dos modelos, basándose en que tan bien el modelo λ_2 coincide con las observaciones generadas por el modelo λ_1 , y relativo a que tan bien el modelo λ_1 coincide con las observaciones generadas por si mismo. Las Tablas 6.4, 6.5 y 6.6, muestran los resultados de aplicar la métrica *Kullback - Leibler* (ecuación: 4.29). Las columnas Paciente(n), muestra el resultado de comparar la trayectoria del Paciente(n) con la trayectoria

Circular Kullback - Leibler					
	16	32	64	128	256
Correcto	0	0	0	0	0
Sano	0	0	0	0	0
Paciente 1	0	0	0	0	0
Paciente 2	0	0	0	0	0
Paciente 3	Infinito	Infinito	Infinito	0.00023	Infinito
Paciente 4	Infinito	Infinito	Infinito	0.00012	Infinito
Paciente 5	Infinito	Infinito	0.00084	0.0003	Infinito
Paciente 6	Infinito	Infinito	Infinito	0.00051	Infinito
Paciente 7	Infinito	Infinito	Infinito	0.0003	Infinito
Paciente 8	Infinito	Infinito	0.00073	0.00242	Infinito
Paciente 9	Infinito	Infinito	Infinito	0.00439	Infinito
Paciente 10	Infinito	Infinito	Infinito	0.002808	Infinito

Tabla 6.5: Muestra los resultados obtenidos al aplicar la métrica Kullback - Leibler a los modelos del gesto circular de cada uno de los pacientes y de la persona sana, con el modelo del gesto circular realizado correctamente. Utilizando observaciones discretizadas en 16, 32, 64 128 y 256 clases.

Abducción Kullback - Leibler					
	16	32	64	128	256
Correcto	0	0	0	0	0
Sano	0	0	0	0	0
Paciente 1	0	0	0	0	0
Paciente 2	0	0	0	0	0
Paciente 3	Infinito	Infinito	Infinito	0.00017	Infinito
Paciente 4	Infinito	Infinito	Infinito	0.00290	Infinito
Paciente 5	Infinito	Infinito	Infinito	0.00384	Infinito
Paciente 6	Infinito	Infinito	Infinito	0.00748	Infinito
Paciente 7	Infinito	Infinito	0.00077	0.00549	Infinito
Paciente 8	Infinito	Infinito	Infinito	0.00890	Infinito
Paciente 9	Infinito	Infinito	0.00017	0.00456	Infinito
Paciente 10	Infinito	Infinito	Infinito	0.00064	Infinito

Tabla 6.6: Muestra los resultados obtenidos al aplicar la métrica Kullback - Leibler a los modelos del gesto abducción de cada uno de los pacientes y de la persona sana, con el modelo del gesto abducción realizado correctamente. Utilizando observaciones discretizadas en 16, 32, 64 128 y 256 clases.

Flexión Porikli					
	16	32	64	128	256
Correcto	0	0	0	0	0
Sano	0.0273	0.0426	0.0361	0.0902	0.0142
Paciente 1	0.0342	0.0473	0.0483	0.0975	0.0228
Paciente 2	0.0459	0.0605	0.0637	0.3490	0.0049
Paciente 3	0.0790	0.0889	0.0177	0.1360	0.0778
Paciente 4	0.0374	0.0072	0.0499	0.2834	0.0012
Paciente 5	0.0765	0.0890	0.0864	0.2569	0.0708
Paciente 6	0.0238	0.0054	0.0808	0.5323	0.0405
Paciente 7	0.0927	0.0123	0.0514	0.2132	0.0122
Paciente 8	0.0247	0.0694	0.0724	0.4909	0.0063
Paciente 9	0.0982	0.0863	0.0774	0.5346	0.0471
Paciente 10	0.0181	0.0964	0.0456	0.5103	0.0564

Tabla 6.7: Muestra los resultados obtenidos al aplicar la métrica Porikli a los modelos del gesto flexión de cada uno de los pacientes y de la persona sana, con el modelo del gesto flexión realizado correctamente. Utilizando observaciones discretizadas en 16, 32, 64 128 y 256 clases.

de la persona sana(correcto) al realizar el ejercicio de flexión, circular y abducción, respectivamente. Las columnas 16, 32, 64, 128 y 256 indican el número de clases al momento de discretizar las características del movimiento. El resultado *infinito* se debe a que algunas veces la probabilidad de que λ_1 genere la secuencia de observaciones O_2 es 0, por lo tanto $\log(P(O_2|\lambda_1)) = -\text{infinito}$.

Circular Porikli					
	16	32	64	128	256
Correcto	0	0	0	0	0
Sano	0.0362	0.0847	0.0172	0.0932	0.0287
Paciente 1	0.0482	0.0646	0.0282	0.1210	0.0936
Paciente 2	0.0518	0.0683	0.0573	0.3239	0.0992
Paciente 3	0.0785	0.0481	0.0458	0.1413	0.0578
Paciente 4	0.0628	0.0361	0.0019	0.3702	0.0383
Paciente 5	0.0510	0.0610	0.0198	0.3728	0.0393
Paciente 6	0.0516	0.0411	0.0283	0.5423	0.0431
Paciente 7	0.0217	0.0350	0.0720	0.5982	0.0418
Paciente 8	0.0529	0.0669	0.0880	0.5380	0.0702
Paciente 9	0.0746	0.0810	0.0569	0.7654	0.0094
Paciente 10	0.0390	0.0109	0.0795	0.9098	0.0135

Tabla 6.8: Muestra los resultados obtenidos al aplicar la métrica Porikli a los modelos del gesto circular de cada uno de los pacientes y de la persona sana, con el modelo del gesto circular realizado correctamente. Utilizando observaciones discretizadas en 16, 32, 64 128 y 256 clases.

Abducción Porikli					
	16	32	64	128	256
Correcto	0	0	0	0	0
Sano	0.0823	0.0273	0.0182	0.0752	0.0263
Paciente 1	0.0060	0.0739	0.0010	0.0898	0.0585
Paciente 2	0.0616	0.0068	0.0150	0.2978	0.0284
Paciente 3	0.0858	0.0012	0.4309	0.2329	0.0842
Paciente 4	0.0048	0.0551	0.1094	0.2590	0.0928
Paciente 5	0.0030	0.0759	0.2987	0.2180	0.0940
Paciente 6	0.0263	0.0555	0.0601	0.4312	0.0835
Paciente 7	0.0494	0.0616	0.0883	0.3983	0.0542
Paciente 8	0.0523	0.0552	0.0613	0.3943	0.0769
Paciente 9	0.0476	0.0909	0.2837	0.4998	0.0135
Paciente 10	0.0682	0.0200	0.7653	0.6131	0.0626

Tabla 6.9: Muestra los resultados obtenidos al aplicar la métrica Porikli a los modelos del gesto abducción de cada uno de los pacientes y de la persona sana, con el modelo del gesto abducción realizado correctamente. Utilizando observaciones discretizadas en 16, 32, 64 128 y 256 clases.

6.3.3. Porikli

Porikli se basa en obtener la distancia de dos modelos, basándose en que tan bien los modelo λ_1 y λ_2 coinciden con las observaciones generadas por el otro modelo (en forma de cruce), y relativo a que tan bien los modelos λ_1 y λ_2 coinciden con las observaciones generadas por si mismos. Las *Tablas* 6.7, 6.8 y 6.9, muestran los resultados de aplicar la métrica *Porikli* (ecuación: 4.31). Las columnas *Paciente(n)*, muestran el resultado de comparar la trayectoria del *Paciente(n)* con la trayectoria de la persona sana(correcto) al realizar el ejercicio de flexión, circular y abducción, respectivamente. Las columnas 16, 32, 64, 128 y 256 indican el número de clases al momento de discretizar las características del movimiento.

6.4. Comparación

Una vez obtenidas las distancia de los modelos realizados por los pacientes hemipléjicos respecto al modelo del movimiento realizado por una persona sana, se continúa entonces a comparar estos resultados con las valoraciones de movilidad realizadas por un especialista en rehabilitación.

Se le tomaron videos a 10 pacientes del INNN con secuelas de una enfermedad vascular cerebral. Los 10 pacientes, nombrados $p1, p2, \dots, p10$, realizaron los tres gestos a calificar: flexión, abducción y circular. Así mismo a estos pacientes se les valoró con las escalas *activity Log*, índice motor y *Fugl - Meyer*, de las cuales las dos últimas se utilizaron para comparar nuestros resultados de este proyecto de tesis. La escala *Fugl - Meyer* califica la capacidad de movimiento y prensión de la extremidad superior en un rango de 0 a 44, 0 para nada de movilidad - prensión y 44 para movilidad/prensión normal. La escala índice motor califica la capacidad de movimiento y prensión de la extremidad superior en un rango de 0% a 100%,

Gesto Flexión		
Relación Levinson - Índice Motor		
	Levinson	Índice Motor
Correcto	1.0000000000	1.0000000000
Sano	0.9644543437	1.0000000000
Paciente 1	0.9670287527	1.0000000000
Paciente 2	0.9646096583	1.0000000000
Paciente 3	0.8639281085	0.9200000000
Paciente 4	0.9448207378	0.9000000000
Paciente 5	0.7430860200	0.7600000000
Paciente 6	0.9398154985	0.6200000000
Paciente 7	0.9398154985	0.5900000000
Paciente 8	0.9091866101	0.5600000000
Paciente 9	0.6091012396	0.5000000000
Paciente 10	0.9985600666	0.3000000000

Tabla 6.10: *Tabla que muestra la calificación obtenida por Levinson de la realización del ejercicio flexión de cada uno de los pacientes y de la persona sana.*

Gesto Flexión		
Relación Levinson - Fugl - Meyer		
	Levinson	Índice Motor
Correcto	1.000000	1.000000
Sano	0.962698	1.000000
Paciente 1	0.967029	1.000000
Paciente 2	0.96461	1.000000
Paciente 3	0.863928	1.000000
Paciente 4	0.944821	0.977273
Paciente 5	0.743086	0.333333
Paciente 6	0.939815	0.863636
Paciente 7	0.939815	0.863636
Paciente 8	0.909187	0.977273
Paciente 9	0.609101	0.931818
Paciente 10	0.99856	0.818182

Tabla 6.11: *Tabla que muestra la calificación obtenida por las escalas, índice motor y Fugl - Meyer.*

0 % para nada de movilidad/prensión y 100 % para movilidad/prensión normal.

Los resultados obtenidos de la comparación entre los modelos ocultos de Markov utilizando *Levinson*, *Porikli*, y *Kullback - Leibler*, se normalizaron entre 0 y 1 al igual que la valoración obtenida por el especialista. La *Tabla 6.10* muestra estos resultados al aplicar *Levinson*, así mismo la *Tabla 6.11* muestra la calificación al mismo paciente comparándola con índice motor y *Fugl - Meyer*. Las *Figuras 6.5a* y *6.5b*, muestran los resultados al graficar estas tablas.

Para comparar los resultados obtenidos por las métricas *Levinson*, *Porikli*, y *KullbackLeibler* con las escalas índice motor y *Fugl - Meyer*, se aplicó la distancia euclidiana:

$$d_e^2 = \sqrt{\frac{1}{N} \sum_{i=1}^N \|m_i - e_i\|^2} \quad (6.1)$$

donde m_i es el resultado obtenido por la métrica aplicada al paciente i , e_i es la valoración utilizando las escalas del paciente i . N es el número de pacientes, en nuestro caso $N = 10$.

Al aplicar la *ecuación 6.1* se corre el riesgo de no medir la variación de los datos. Por lo tanto, también es necesario calcular la varianza

$$v^2 = \frac{1}{N} \sum_{i=1}^N \|d_e - (m_i - e_i)\|^2 \quad (6.2)$$

donde $d_e = \frac{1}{N} \sum_{i=1}^N \|m_i - e_i\|$.

En la siguiente sección se muestran los resultados obtenidos en cada uno de los gestos al aplicar las métricas *Levinson*, *Porikli*, y *KullbackLeibler* y las escalas índice motor y *Fugl - Meyer*. Únicamente se muestran los resultados obtenidos con las secuencias de observaciones discretizadas por la clase (16, 32, 64, 128 y 256) que tuvo el mejor resultado al aplicar las *ecuaciones: 6.1* y *6.2*. La clase que mejor resultado tuvo al aplicar *Levinson* fue 16 y *Porikli* y *Kullback - Leibler* 128.

Métricas						
	Índice motor			Fugl - Meyer		
Gesto	Levinson	K - L	Porikli	Levinson	K - L	Porikli
Flexión	0.29227	0.22872	0.16433	0.13224	0.16207	0.30420
Circular	0.20517	0.13503	0.16433	0.18172	0.40648	0.43993
Abducción	0.22959	0.28736	0.19999	0.24340	0.39415	0.29562

Tabla 6.12: Tabla que muestra la distancia euclidiana entre los resultados de las métricas y los resultados de las escalas.

Métricas						
	Índice motor			Fugl - Meyer		
Gesto	Levinson	K - L	Porikli	Levinson	K - L	Porikli
Flexión	0.04491	0.02220	0.00938	0.00813	0.01347	0.02564
Circular	0.01892	0.00997	0.00638	0.02143	0.08188	0.04413
Abducción	0.02254	0.04113	0.00742	0.03538	0.07586	0.01555

Tabla 6.13: Tabla que muestra la varianza entre los resultados de las métricas y los resultados de las escalas.

6.4.1. Resultados

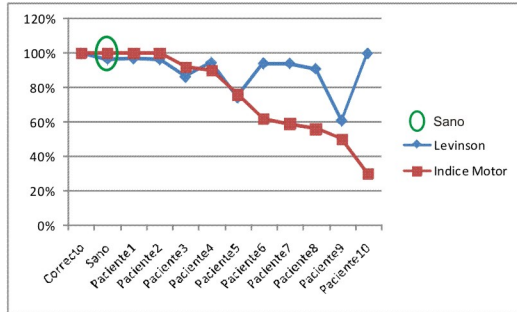
En esta sección, se muestran los resultados de comparar la calificación obtenida al aplicar las métricas a los gestos y los resultados de las escalas aplicadas a los pacientes. Las Figuras 6.5, 6.6 y 6.7, muestran estas comparaciones para el gesto flexión, circular y abducción, respectivamente.

En la Tabla 6.12 se muestran los resultados de aplicar la ecuación 6.1. Como se puede apreciar los mejores resultados en el gesto circular y abducción son los obtenidos al aplicar *Porikli* y comparándolos con la escala índice motor. Así para el gesto flexión es *Kullback - Leibler* igualmente comparándolo con el índice motor.

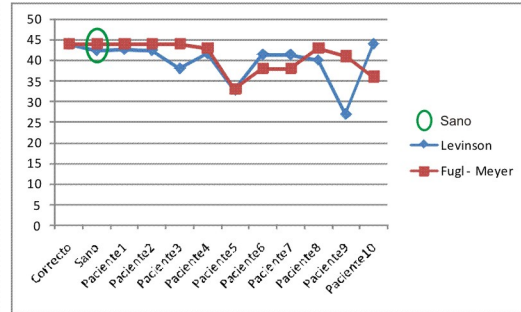
Flexión

Índice motor

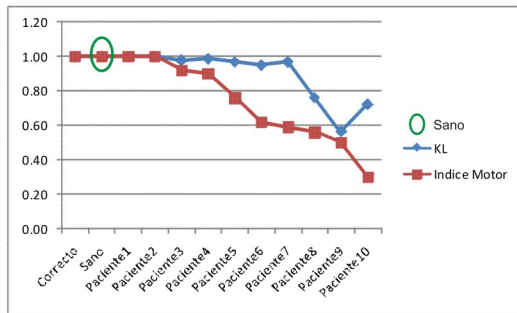
Fugl - Meyer



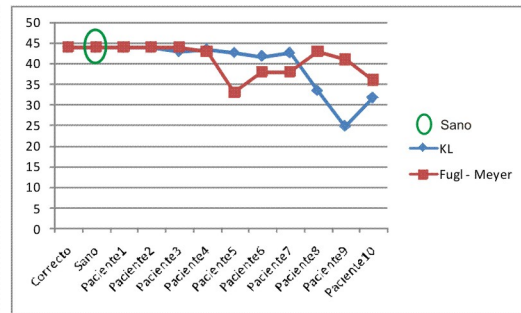
a) Levinson



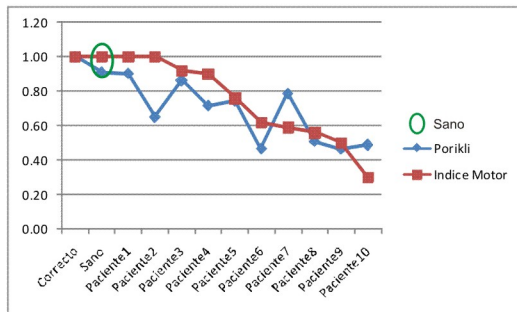
b) Levinson



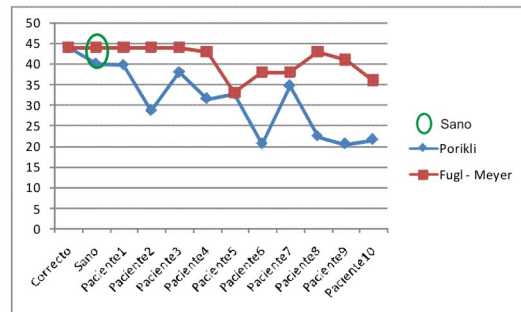
c) Kullback - Leibler



d) Kullback - Leibler



e) Porikli



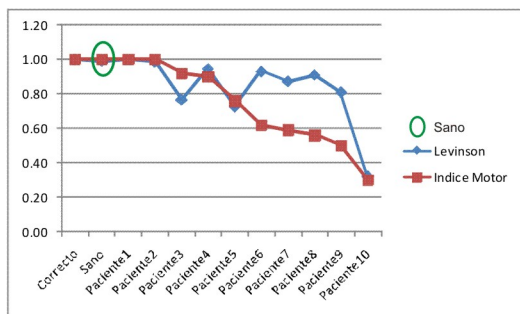
f) Porikli

Figura 6.5: Gráficas comparativas de las calificaciones al gesto flexión, a) y b) Resultados de Levinson en índice motor y Fugl - Meyer, c) y d) Resultados de Kullback - Leibler en índice motor y Fugl - Meyer y e) y f) Resultados de Porikli en índice motor y Fugl - Meyer.

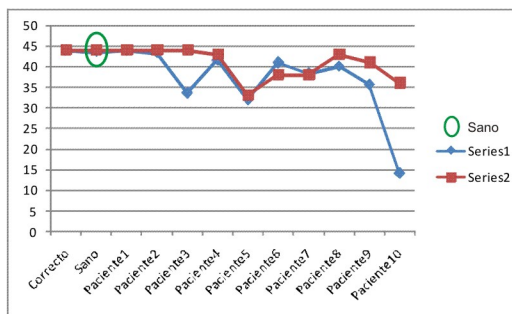
Circular

Índice motor

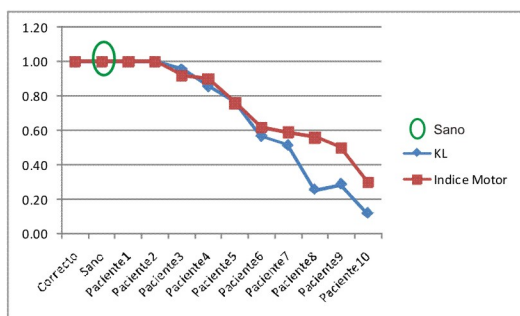
Fugl - Meyer



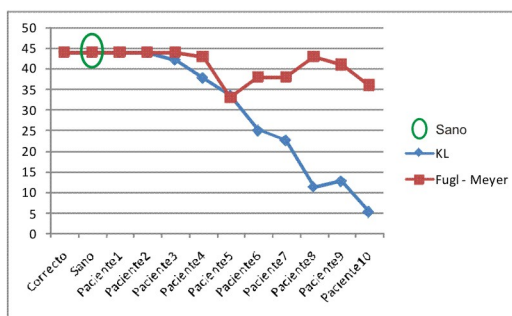
a) Levinson



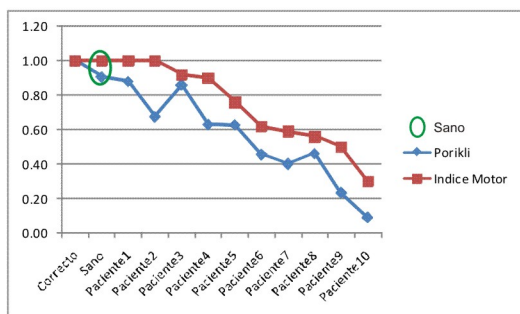
b) Levinson



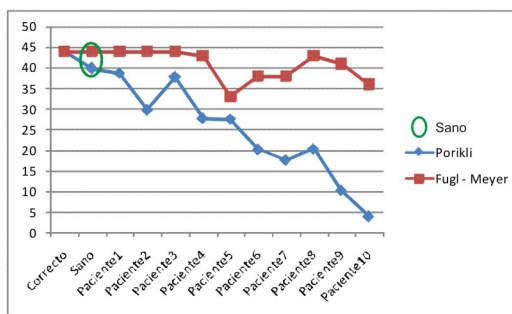
c) Kullback - Leibler



d) Kullback - Leibler



e) Porikli



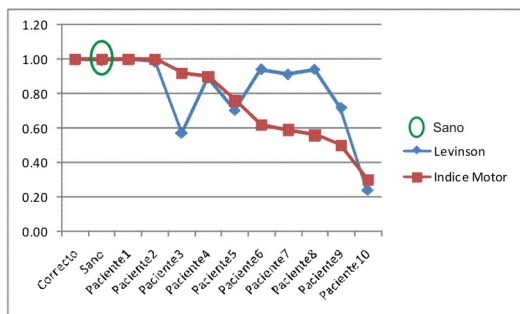
f) Porikli

Figura 6.6: Gráficas comparativas de las calificaciones al gesto circular, a) y b) Resultados de Levinson en índice motor y Fugl - Meyer, c) y d) Resultados de Kullback - Leibler en índice motor y Fugl - Meyer y e) y f) Resultados de Porikli en índice motor y Fugl - Meyer.

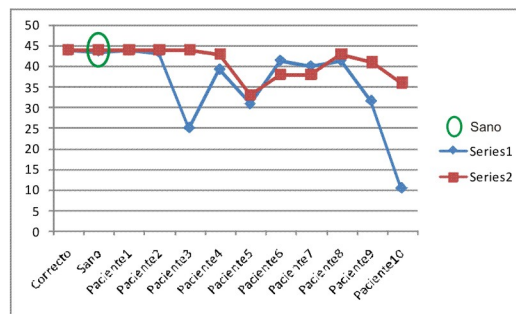
Abducción

Índice motor

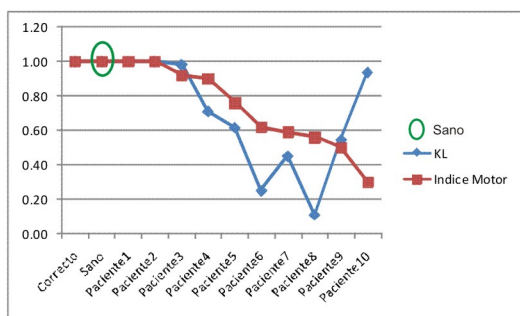
Fugl - Meyer



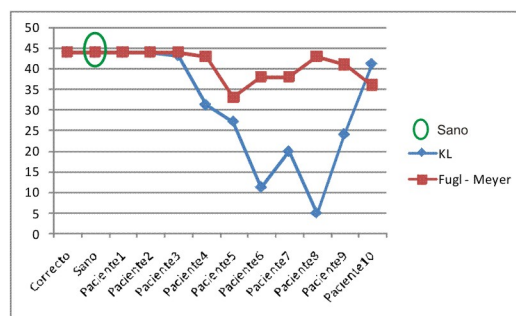
a) Levinson



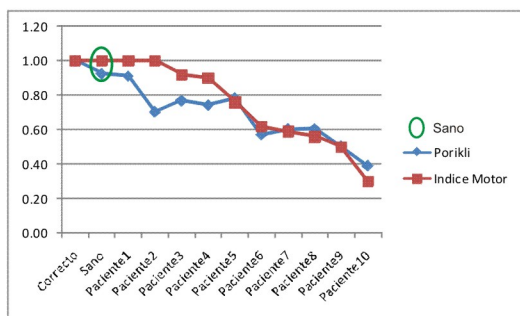
b) Levinson



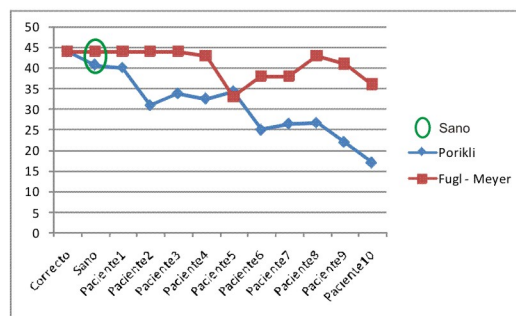
c) Kullback - Leibler



d) Kullback - Leibler



e) Porikli



f) Porikli

Figura 6.7: Gráficas comparativas de las calificaciones al gesto abducción, a) y b) Resultados de Levinson en índice motor y Fugl - Meyer, c) y d) Resultados de Kullback - Leibler en índice motor y Fugl - Meyer y e) y f) Resultados de Porikli en índice motor y Fugl - Meyer.

Métricas						
	Índice motor			Fugl - Meyer		
	Levinson	K - L	Porikli	Levinson	K - L	Porikli
Promedio	0.24235	0.21704	0.17622	0.18578	0.32090	0.34658

Tabla 6.14: Tabla que muestra el promedio de la distancia euclidiana de los tres gestos.

Métricas						
	Índice motor			Fugl - Meyer		
	Levinson	K - L	Porikli	Levinson	K - L	Porikli
Promedio	0.02879	0.02443	0.00773	0.02165	0.05703	0.02844

Tabla 6.15: Tabla que muestra el promedio de la varianza de los tres gestos.

En la *Tabla 6.13* se muestran los resultados, de determinar que tan espaciados están los resultados, dando como mejor resultado *Porikli* en los tres gestos.

Estos resultados determinan que la métrica para comparar modelos ocultos de Markov que mejor se ajusta a calificar nuestros tres gestos, es *Porikli*. En particular, el índice motor es la escala que mejor se ajusta a las métricas que comparan modelos ocultos de Markov que representan trayectorias de los gestos terapéuticos.

6.5. Persona sana

Se realizó una comparación entre los gestos realizados por una persona sana y el gesto realizado correctamente. Como se puede apreciar en las tablas de las secciones: 6.3.1, 6.3.2 y 6.3.3 y en la gráficas de la figuras: 6.5, 6.6 y 6.7 las calificaciones obtenidas con las métricas *Levinson*, *Kullback - Leibler* y *Porikli* fueron las mas altas, dando una calificación alrededor del 90%. Es decir, que

una calificación utilizando la métrica de *Porikli* a un gesto de flexión, circular y abducción bien realizado es del 90 %.

6.6. Análisis de resultados

Del análisis de los resultados de estas pruebas, se tienen las siguientes conclusiones:

- La similitud obtenida al aplicar *Levinson*, puede determinar distancias “cero”, es decir, determinar que dos trayectorias de movimiento son iguales, cuando son totalmente diferentes.
- La similitud obtenida al aplicar *Levinson*, determina que tan probable es que las dos trayectorias pasen por cierto punto en el espacio, pero no determina el tiempo en que pasan por cierto punto.
- La similitud obtenida por *Kullback Leibler*, a menudo es infinita, es decir, totalmente diferentes, se tienen que seleccionar secciones de la trayectoria para que esto no suceda. El número de observaciones para que converja es impreciso, ya que se basa en determinar la similitud de los logaritmos de la probabilidad de los modelos.
- Estos resultados determinan que la métrica para comparar modelos ocultos de Markov que mejor se ajusta a calificar nuestros tres gestos, es *Porikli*. En particular, el índice motor es la escala que mejor se ajusta a las métricas que comparan modelos ocultos de Markov que representan trayectorias de los gestos terapéuticos.
- Los mejores resultados se obtuvieron con *Porikli*, al comparar el gesto circular con la escala índice motor, ya que esta escala se basa principalmente en valorar en conjunto los movimientos flexor y abductor y la presión de

la mano, y el gesto circular tiene cambios en el ángulo flexor, y abductor, dando así una mejor correlación. Esto no sucede en los gestos de sólo flexión y abductor, ya que el paciente puede tener movimiento en el músculo flexor, más no en el abductor o viceversa.

- También se pudo apreciar que la calificación a una persona sana realizada por las métricas *Levinson*, *Kullback - Leibler* y *Porikli* es de las más altas, un poco más de 90%. Es decir, que una calificación utilizando la métrica de *Porikli* a un gesto de flexión, circular y abducción bien realizado es del 90%.
- La comparación con la escala *Fugl - Meyer*, dio los peores resultados, ya que la escala *Fugl - Meyer*, así como valora movimiento, valora presión y el dolor al realizar el movimiento, una persona con dolor, puede realizar el movimiento o por el contrario una persona no puede realizar el movimiento aún sin dolor.
- Una forma de mejorar la correlación entre las métricas y las escalas es utilizar sólo los incisos de la escala que califican el movimiento de la extremidad superior.
- Para mejorar la precisión de la calificación de las métricas se necesita de algún método que elimine el ruido de la segmentación, ya que el centro de masa varía aún estando el paciente sin realizar ningún movimiento. Tal vez se pueda solucionar seleccionando otro método para la obtención de las características del movimiento humano.

6.7. Resumen

En este capítulo se mostraron los resultados de la segmentación y seguimiento realizadas por el sistema, para poder obtener las características del movimiento de los pacientes al realizar el ejercicio terapéutico. También se mostraron los resultados de la obtención de los tres ángulos: (α, θ, ϕ) , que son las características obtenidas del movimiento. También se mostraron ejemplos de las gráficas que muestran la discretización de las características en clases de 16, 32 y 64. En las últimas secciones de este capítulo se mostraron los resultados de la comparación entre las calificaciones obtenidas al aplicar *Levinson*, *Kullback - Leibler* y *Porikli* a cada uno de los gestos y las escalas aplicadas a los pacientes: índice motor y *Fugl - Meyer*. En estos resultados se pudo apreciar que la métrica para comparar modelos ocultos de Markov que mejor se ajusta a calificar nuestros tres gestos, es *Porikli*. En particular, el índice motor es la escala que mejor se ajusta a las métricas que comparan modelos ocultos de Markov que representan trayectorias de los gestos terapéuticos.

Estos resultados también muestran una posible clasificación de pacientes con hemiplegia, que puede ser utilizada para calificar su movilidad al realizar tres gestos terapéuticos.

Capítulo 7

Conclusiones y trabajo futuro

Cada año 200,000 personas en México sobreviven a un derrame cerebral y cerca de 700,000 en Estados Unidos. Aproximadamente el 80 % de ellos pierden la habilidad del movimiento. La deficiencia del movimiento después de un derrame cerebral es generalmente tratada con terapia física intensa por varias semanas después de la lesión, que generalmente es dirigida por un fisioterapeuta que diagnostica si se está llevando la rehabilitación apropiadamente. Desafortunadamente, la terapia de los pacientes que han sufrido un derrame cerebral es intensa mientras se tiene esta ayuda profesional, así como la economía suficiente para cubrir los gastos que esto implica. La rehabilitación subsecuente en casa, es a menudo autodirigida con información poco profesional.

Por tal motivo, ha surgido la inquietud de desarrollar sistemas que detecten el progreso de rehabilitación en pacientes con atrofias en la neuronas motoras que impiden al paciente tener una movilidad con libertad. En la actualidad, el experto en rehabilitación, utiliza diversas escalas para valorar la movilidad de sus pacientes. El problema de estas escalas es que la mayoría son ordinales y no paramétricas.

En esta tesis se propone utilizar los Modelos Ocultos de Markov para generar una calificación automática de la realización de gestos con el brazo humano utilizados en rehabilitación. Para lograr esta calificación, se creó un sistema visual capaz de obtener las características del movimiento humano para poder modelarlo. Este sistema visual esta formado por tres etapas: segmentación, seguimiento 2D y Reconstrucción 3D. La segmentación se encarga de la localización de las marcas de color dentro de una imagen. Las articulaciones están marcadas con cintas de color: amarillo - hombro, rojo - codo y azul - muñeca, para poder realizar una segmentación mas exacta. Para la segmentación de las marcas, se aplica información de color (histogramas de color). El seguimiento 2D se encarga del seguimiento de un objeto a través de una secuencia de imágenes de video. Y la reconstrucción 3D obtiene las coordenadas 3D del objeto en seguimiento. Se aplica un algoritmo de intersección de líneas de visión al objeto en seguimiento de ambas cámaras para obtener una estimación de la posición 3D. Para esto es necesario realizar una calibración de las cámaras para obtener los parámetros intrínsecos y los parámetros extrínsecos.

Los Modelos Ocultos de Markov, pueden modelar la trayectoria del movimiento de un objeto a partir de asignar dos distribuciones de probabilidades. Así, el Modelo Oculto de Markov, puede ser visto como dos distribuciones de probabilidades. De esto, es fácil suponer que al tener dos Modelos Ocultos de Markov, se puede obtener la distancia entre los modelos a partir de la distribuciones de sus probabilidades. *Levinson*, *Kullback - Leibler* y *Porikli*, son métricas utilizadas para obtener la semejanza entre Modelos Ocultos de Markov. *Levinson*, compara la semejanza de las probabilidades entre los dos modelos y *Kullback - Leibler* y *Porikli*, comparan dos modelos, basándose en determinar que tan probable es que un modelo genere los mismos cambios de posición del objeto que el otro modelo.

Estas métricas fueron aplicadas a los movimientos de diez pacientes con hemi-

plejia del INNN. Los diez pacientes realizaron tres gestos terapéuticos, flexión, circular y abducción. Las métricas fueron aplicadas comparando el movimiento del paciente con el movimiento correcto realizado por una persona sana. Así mismo, estos diez pacientes y la persona sana, fueron valorados de su movilidad por un especialista en rehabilitación del INNN utilizando las escalas: índice motor y *Fugl - Meyer*. Estas valoraciones fueron comparadas con los resultados de aplicar *Levinson*, *Kullback - Leibler* y *Porikli* a los modelos Ocultos de Markov de los movimientos.

Del análisis de los resultados de éstas pruebas, se determina que la métrica para comparar modelos ocultos de Markov que mejor se ajusta a calificar nuestros tres gestos, es *Porikli*. En particular, el índice motor es la escala que mejor se ajusta a las métricas que comparan modelos ocultos de Markov que representan trayectorias de los gestos terapéuticos. Una forma de mejorar la correlación entre las métricas y las escalas es utilizar sólo los incisos de la escala que califican el movimiento de la extremidad superior. Para mejorar la precisión de la calificación de las métricas se necesita de un método de seguimiento más robusto, o algún método que elimine el ruido de las características del movimiento humano.

7.1. Aportaciones de la tesis

La principal aportación de este proyecto de tesis, es la utilización de los Modelos Ocultos de Markov para la calificación de gestos. Otra aportación importante es la obtención de una calificación automática y paramétrica que resuelva el problema de las escalas utilizadas en rehabilitación por ser demasiado subjetivas. Esto se vuelve prometedor para el especialista en rehabilitación, ya que brinda una técnica objetiva al momento de valorar a sus pacientes. Esto abre una gran posibilidad de seguir investigando en crear modelos para la calificación de gestos, ya sea para la

rehabilitación, entrenamiento, etc. Así mismo este proyecto contribuye a crear una herramienta de bajo costo (alrededor de \$1,000 pesos) que pueda ser incluida en un sistema para la rehabilitación de pacientes con hemiplejía capaz de ser utilizada por más personas que no cuenten con los recursos para pagar una rehabilitación o los medios para asistir a un centro de rehabilitación. Una consecuencia es que esta escala puede ser capaz de ayudar a la motivación y rehabilitación de pacientes con secuelas de enfermedad vascular cerebral como la hemiplejía, sin contar con la presencia de un terapeuta.

7.2. Trabajo futuro

El trabajo futuro se va a separar en dos direcciones, la primera es en resolver los problemas presentados en este trabajo y la segunda, son las motivaciones para continuar la investigación en el área de calificación de gestos.

1. *Movimiento humano*: La segmentación, es uno de los principales problemas al obtener las características del movimiento humano, ya que aún no se cuenta con un método de visión que pueda segmentar objetos en cualquier ambiente con cambios de iluminación. La segmentación utilizada en este proyecto de tesis obtiene la posición de las articulaciones del brazo, a partir de encontrar los centros de masas de las marcas de color. Estos centros cambian constantemente, aún sin que la persona se mueva. Para este trabajo de tesis, estos cambios no son demasiado influyentes en nuestros resultados, pero es claro que si se quiere obtener una calificación más exacta del movimiento, se necesita encontrar un método que siga el movimiento humano sin estos cambios involuntarios del centro de masa del objeto. Así mismo también es necesario un método que resuelva el problema de oclusión, ya que la selección de gestos para este proyecto de tesis, estuvo restringida a aquellos

gestos que no ocluyeran las marcas de color.

2. *Calificación de gestos*: Sin duda esta es un área prometedora, sus aplicaciones son bastas. Un trabajo futuro es seguir con la calificación de gestos, no solo del brazo, sino de todas las articulaciones del cuerpo, así como movimiento facial. Ya que la hemiplejia se presenta en todo el cuerpo, no sólo en las extremidades superiores. Otra investigación prometedora, sería en obtener una correlación directa entre la calificación del gesto con los daños cerebrales del paciente. Esta área no sólo puede servir para rehabilitación, sino también para entrenamiento de movimientos, como por ejemplo a los cirujanos para mejorar sus movimientos al operar o los gimnastas al realizar sus movimientos.

Bibliografía

- [1] J. K. Aggarwal and Q. Cai. Human motion analysis: A review. Technical report, Department of Electrical and Computer Engineering, The University of Texas at Austin, 1999.
- [2] D. H. Ballard and C. M. Brown. *Computer Vision*. Prentice-Hall, 1982.
- [3] Barkheit, Thilmann, Ward, Poewe, and Wissel. A randomized, double-blind, placebo-controlled, dose-ranging study to compare the efficacy and safety of three doses of botulinum toxin type a (dysport) with placebo in upper limb spasticity after stroke. *Stroke*, 31:2402–2406, 2000.
- [4] A. Baumberg and D. Hogg. An efficient method for contour tracking using active shape models. *Proc. IEEE Workshop on Motion of Non-Rigid and Articulated Objects*, pages 194–199, 1994.
- [5] R. Bonita. Epidemiology of stroke. *Lancet*, 339:342–244, 1992.
- [6] P. Bourke. *The Shortest Line Between Two Lines In 3D*. disponible en: <http://astronomy.swin.edu.au/~pbourke/geometry/lineline3d/>, 2006.
- [7] G.R. Bradski. Computer vision face tracking as a component of a perceptual user interface. *Workshop on Applications of Computer Vision*, 1:214–219, 1998.

- [8] C. G. Burgar, P. S. Lum, P. C. Shor, and H. F. Machiel Van der Loos. Deleppment of robots for rehabilitation therapy: the palo alto va/sanford experience. *J. Rehabil. Re. Develop.*, 37(6):663–673, 2000.
- [9] Raúl Calderón-Sepúlveda. Escalas de medición de la función motora y la espasticidad en parálisis cerebral. *Mex Neuroci*, pages 285–289, 2002.
- [10] L. Campbell, D. Becker, A. Azarbayejani, A. Bobick, and A. Pentland. Invariant features for 3-d gesture recognition. 1996.
- [11] A. Cerrato-Cuadra and R. Carrillo-Mezo. Activación del área motora primaria mediante rmf posrehabilitación en pacientes con miembro torácico parético, como secuela de enfermedad vascular cerebral. *Neurocién*, 10:57–65, 2006.
- [12] S. Coote, E. Stokes, B. Murphy, and W. Harwin. The effect of gentle/s robot mediated therapy on upper extremity function post stroke. *Proceedings International Conference on Rehabilitation Robotics Korea*, pages 59–61, 2003.
- [13] Ascension Technology Corporation. Flock of birds, technical description of dc magnetic trackers. Technical report, Ascension Technology Corporation, 2008.
- [14] J. L. Crowley and J. Coutax. Vision for man machine interaction. *Robotics and Autonomous Systems*, 19:347–358, 1997.
- [15] Patricia M. Davies. *Pasos a seguir. Tratamiento integrado de paciente con hemiplejía*. Médica Panamericana, S.A., 2003.
- [16] M. Falkhausen, H. Reininger, and Wolf. Calculation of distance measures between hidden markov models. Technical report, Institut fr Angewandte Physik, Johann Wolfgang Goethe Universitt, FRG, 1995.

- [17] Kasper; Braunwald & Fauci. *HARRISON. Principios de medicina interna.*, volume 2. Mc Graw Hill, 16 edition, Junio 2005.
- [18] D. M. Gómez-Allende. *Reconocimiento de Formas y Visión Artificial*. Editorial Ra-Ma, 1993.
- [19] P. B. González and J. M. Caas. Seguimiento tridimensional usando dos cámaras. Technical report, Universidad Rey Juan Carlos, España, Octubre 2004. (In Spanish).
- [20] R. C. Gonzalez and R. E. Woods. *Digital Image Processing*. Adisson-Wesley, 1992.
- [21] R. Hartley and A. Zisserman. Multiple view geometry in computer vision. *Press Syndicate of the University of Cambridge*, 2000.
- [22] S. Hesse, G. Schulte-Tiggs, M. Konrad, A. Bardeleben, and C. Werner. Robot-assisted arm trainer for the passive and active practice of bilateral forearm and wrist movements in hemiparetic subjects. *Arch Phys Med. Rehabil*, 84(6):915–920, 2003.
- [23] S. J. Housman, V. Le, T. Rahman, R. J. Sanchez, and D. J. Reinkensmeyer. Arm-training with t-wrex after chronic stroke: Preliminary results of a randomized controlled trial. 2007.
- [24] <http://wii.nintendo.com>. Wiimote nintendo, 2007.
- [25] C. Jaynes, S. Webb, R. Steele, and Q. Xiong. An open development environment for evaluation of video surveillance systems. *Proc. of PETS*, 2002.
- [26] G. Johansson. Visual motion perception. *Scientific American*, pages 76–88, 1975.

- [27] A. Kourepenis, A. Petrovich, and M. Meinberg. Development of a monolithic quartz resonator accelerometer. 1989.
- [28] V. Kravtchenko. Tracking color objects in real time. Master's thesis, The University of British Columbia, 1999.
- [29] H. I. Krebs, N. Hogan, B. T. Volpe, M. L. Aisen, L. Edelstein, and C. Diels. Overview of clinical trials with mit-manus: a robot-aided neuro-rehabilitation facility. *Technol. Health Care*, 7(6):419–423, 1999.
- [30] S. Kullback. Information theory and statistic. *New York: Wiley*, 1958.
- [31] S. E. Levinson, L. R. Rabiner, and M. M. Sondhi. An introduction to the application of the theory of probabilistic function of a markov process to automatic speech recognition. *B. S. T. J.*, 62:1036–1074, 1983.
- [32] José Ismael Mariscal. Comparación de proporciones de la ecuperacin funcional en la muñeca y mano espásticas de pacientes con secuelas de enfermedad vascular cerebral, entre un grupo tratado con fisioterapia y toxina botulínica tipo a, y otro con fisiterapia y placebo. Master's thesis, Universidad de Colima, 2003.
- [33] X. Munoz, J. Freixenet, X. Cufí, and J. Martí. Strategies for image segmentation combining region and boundary information. *Pattern Recognition Letters*, 24:375 – 392, 2003.
- [34] G. Pajares and J. M. De La Cruz. *Visión por computador: Imágenes y aplicaciones*. Editorial Alfa-Omega, 2002.
- [35] F. Porikli. Trajectory distance metric using hidden markov model based representation. Technical report, Mitsubishi Electronic Research Laboratories, 2005.

- [36] B. H. Juang AND L. R. Rabiner. A probabilistic distance measure for hidden markov models. *AT&T Technical Journal*, 64:391–408, 1985.
- [37] L. R. Rabiner. A tutorial on hidden markov models an selected applications in speech recognition. *Proceedings of the IEEE*, 77:257–286, 1989.
- [38] Reinkensmeyer, Pang, Nessler, and Painter. Java therapy: Web-based robotic rehabilitation. 2002.
- [39] D. J. Reinkensmeyer, L. E. Kahn, M. Averbuch, A. McKenna-Cole, B. D. Schmit, and W. Z. Rymer. Understanding and treating arm movement impairment after chronic brain injury: progress with the arm guide. *J. Rehabil. Re. Develop.*, 37(6):653–662, 2000.
- [40] A. Senior, A. Hampapur, Y. Tian, L. Brown, S. Pankanti, and R. Bolle. Appearance models for occlusion handling. *Proc. of PETS*, December 2001.
- [41] Jang SH, Kim YH, Cho SH, Lee JP, Park JW, and Kwon YH. Cortical reorganization induced by task-oriented training in chronic hemiplegic stroke patients. *Neuroreport*, 14:137 – 141, 2003.
- [42] H. Sidenbladh. *Probabilistic Tracking and Reconstruction of 3D Human Motion in Monocular Video Sequences*. PhD thesis, Dept. of Numerical Analysis and Comp. Sci., Sweden, 2001.
- [43] C. Sminchisescu. *Estimation Algorithms for Ambiguous Visual Models Three-Dimensional Human Modelling and Motion Reconstruction in Monocular Video Sequences*. PhD thesis, Institute National Politechnique de Grenoble (INRIA), July 2002.
- [44] CODA Software. Disponible en: <http://www.charndyn.com/>, 2007.

- [45] VICON Software. Disponible en: <http://www.vicon.com/>, 2007.
- [46] L. E. Sucar, G. Azcárate, R. Leder, and D. Reinkensmeyer. Markerless video arm tracking as a tool for biomechanics/rehabilitation engineering.
- [47] M. J. Swain and D. H. Ballard. Color indexing. *International Journal of Computer Vision*, pages 11–32, 1991.
- [48] Y. Tao and H. Hu. Building a visual tracking system for home-based rehabilitation. *Proc. of the 9 Chinese Automation and Computing Society Conference in the UK, Luton, Eng.,*, pages 131–161, 2003.
- [49] Y Tao, H. Hu, and H. Zhou. Integration of vision and inertial sensors for home-based rehabilitation. *Proc. of IEEE International Conference on Robotics and Automation*, pages 106–112, 2005.
- [50] OpenCV *Open Source Computer Vision Library*. Disponible en: <http://www.sourceforge.net/projects/opencvlibrary>, 2007.
- [51] G. Uswatte, E. Taub, D. Morris, M. Vignolo, and K. McCulloch. Reliability and validity of the upper-extremity motor activity log-14 for measuring real-world arm use. *American Heart Association*, 2005.
- [52] J. J. Wang and S. Singh. Video analysis of human dynamics - a survey. Technical report, Department of Computer Science, University of Exeter, 2006.
- [53] C. Wren, A. Azarbayejani, T. Darrell, and A. Pentland. Pfinder: Real-time tracking of the human body. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, pages 780–785, 1997.
- [54] www.phasespace.com. Phase space motion capture, 2007.

- [55] www.qualisys.com. Optical motion capture, 2007.
- [56] www.strokeassociation.org. American stroke association, 2007.
- [57] www.xplain.com. Rehabilitacion despues de un derrame cerebral, 2007.
- [58] www.xsens.com. Xsens system, 2007.
- [59] H. Xie and G. Fedder. A cmos z-axis capacitive accelerometer with comb-finger sensing. Technical report, The Robotics Institute, Carnegie Mellon University, 2000.
- [60] H. Zhou and H. Hu. Inertial motion tracking of human arm movements in stroke rehabilitation. *Proc. of IEEE International Conference on Mechatronics and Automation*, pages 1306 – 1311, 2005.