

## Comparación de dos Sistemas de Captura de Movimiento por medio de las Trayectorias Articulares de Marcha

*D.A. Bravo M.<sup>1</sup>, C.F. Rengifo R.<sup>2</sup>, W. Agredo R.<sup>3</sup>*

<sup>1</sup>Departamento de Física, Universidad del Cauca, Calle 5 No. 4-70, Popayán, Colombia.

<sup>2</sup>Departamento de Electrónica, Instrumentación y Control, Universidad del Cauca, Calle 5 No. 4-70, Popayán, Colombia.

<sup>3</sup>Laboratorio de Biomecánica, Universidad Autónoma de Occidente, Calle 25 No. 115-85. Cali, Colombia.

### RESUMEN

En la actualidad, los métodos más comunes para una adecuada captura del movimiento humano en tres dimensiones requieren de un entorno de laboratorio y la fijación de marcadores, accesorios o sensores a los segmentos corporales. Sin embargo, el alto costo de estos equipos es un factor limitante en diversos entornos de trabajo. Sistemas de captura de movimiento como Microsoft Kinect <sup>TM</sup> presentan un enfoque alternativo a la tecnología de captura de movimiento.

En este trabajo se comparan dos sistemas de captura de movimiento por medio de las trayectorias articulares y las medidas antropométricas de una persona en un ciclo de marcha normal. El primero de ellos, es un sistema comercial de precisión que utiliza marcadores (Vicon <sup>TM</sup>) y el segundo, es la cámara Microsoft Kinect <sup>TM</sup>. Ambos sistemas se evaluaron con el propósito de comparar la diferencia geométrica y el error RMS entre las trayectorias articulares de la marcha humana obtenidas por cada uno de los sistemas. Los resultados muestran una varianza mayor en las medidas antropométricas y trayectorias articulares para el Kinect, aunque este sistema es de bajo costo y de fácil uso e instalación, no puede ser utilizado para un análisis preciso de la cinemática de la marcha humana.

**Palabras clave:** biomecánica, captura de movimiento, Microsoft Kinect <sup>TM</sup>, marcha humana, trayectorias articulares.

### Correspondencia:

Diego A Bravo M.

Universidad del Cauca, Calle 5 No. 4-70, Popayán, Colombia.

Correo electrónico: [dibravo@unicauca.edu.co](mailto:dibravo@unicauca.edu.co)

**Fecha de recepción:**

14 de agosto de 2015

**Fecha de aceptación:**

1 de marzo de 2016

## ABSTRACT

Currently, the most common methods for proper capture of human movement in three dimensions require a laboratory environment and setting markers, accessories or sensors to the body segments. However, the high cost of this equipment is a limiting factor in diverse environments. Motion capture systems such as Microsoft Kinect<sup>TM</sup> present an alternative approach to motion capture technology.

In this paper, two motion capture systems are compared by means of joint trajectories and anthropometric measurements of a person in a normal gait cycle. The first is accurate trading system that uses markers (Vicon<sup>TM</sup>), and the second is Microsoft Kinect<sup>TM</sup> camera. Both systems were evaluated in order to compare the geometric difference and the RMS error between the joint human walking paths obtained for each of the systems. The results show a greater variance in anthropometric measures and joint trajectories for Kinect, although this system is inexpensive and easy to use and install, cannot be used for precise kinematic analysis of human walking.

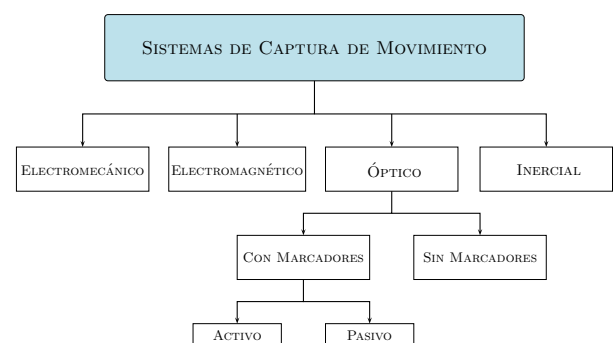
**Keywords:** biomechanics, Microsoft Kinect<sup>TM</sup>, motion capture, human gait, joint trajectories.

## INTRODUCCIÓN

En el mercado existen sistemas especializados para la captura de movimiento, por ejemplo: PeakMotus (Vicon<sup>TM</sup>), SkillSpector<sup>TM</sup>, DartFish<sup>TM</sup>, entre otros. Estos sistemas son utilizados para adquirir las características del movimiento de un ser humano con un grado de precisión relativamente alto. Pueden clasificarse según su tecnología en: electromecánicos, electromagnéticos, inerciales y ópticos. La clasificación se ilustra en la Figura 1.

En los sistemas de captura de movimiento electromecánicos (Gypsy7<sup>TM</sup>), la captura del movimiento se realiza usando sensores mecánicos. En el proceso de captura de movimiento, la persona viste unos trajes especiales, adaptables al cuerpo humano. Estos trajes son generalmente estructuras rígidas compuestas de barras metálicas o plásticas, unidas mediante potenciómetros colocados en las principales articulaciones. Básicamente, el actor coloca la estructura en su cuerpo y mientras se mueve, el traje se adapta a los movimientos que este realiza; los potenciómetros recogen toda la información del grado de rotación de las articulaciones. La

desventaja de estos sistemas con respecto a otros, es la incapacidad de medir traslaciones globales (miden las posiciones relativas de los miembros, pero no el desplazamiento del actor en el escenario). Por otro lado, dicho sistema supone que la mayoría de los huesos humanos están unidos por articulaciones de un grado de libertad con centro instantáneo de rotación fijo, pero no tiene en cuenta rotaciones complejas que se producen en las articulaciones humanas; las estructuras suelen ser rígidas y restringen el movimiento del actor.



**Figura 1.** Clasificación de sistemas de captura de movimiento.

Para los sistemas de captura de movimiento electromagnéticos (Polhemus<sup>TM</sup>) [18], se dispone de una colección de sensores electromagnéticos que miden la relación espacial con un transmisor cercano. Los sensores se colocan en el cuerpo y se conectan a una unidad electrónica central; están constituidos por tres espiras ortogonales que miden el flujo magnético, determinando posición y orientación del sensor. Un transmisor genera un campo electromagnético de baja frecuencia que los receptores detectan y transmiten a la unidad electrónica de control.

En los sistemas de captura de movimiento inerciales (Xsens<sup>TM</sup>) [27], se colocan sensores inerciales en distintas partes del cuerpo (acelerómetros triaxiales y giroscopios). Una ventaja es que se obtienen datos precisos de aceleración y orientación del individuo. Sin embargo, no es posible medir traslaciones globales y una desventaja es que estos sensores son muy sensibles a cambios en los campos magnéticos.

Los sistemas ópticos utilizan los datos recogidos por sensores de imagen para inferir la posición de un elemento en el espacio, utilizando una o más cámaras sincronizadas para proporcionar proyecciones simultáneas. Generalmente se usan marcadores pegados al actor, pero los sistemas más recientes permiten recoger datos confiables, rastreando superficies del sujeto identificadas dinámicamente. Estos sistemas entregan la posición cartesiana  $(x, y, z)$  de cada marcador en un marco de referencia inercial; la orientación de una superficie se calcula utilizando la posición relativa, de al menos, 3 marcadores. Los sistemas ópticos de captura de movimiento, permiten la grabación en tiempo real, con algunas limitaciones como son: el número de cámaras, marcadores y actores. Estos sistemas pueden capturar un gran número de marcadores a frecuencias del orden de hasta 2000 cuadros por segundo.

Los sistemas ópticos de captura de movimiento son los más utilizados en laboratorios de biomecánica [15] y se pueden clasificar como sistemas ópticos sin marcadores y con marcadores, estos a su vez se dividen en: activos y pasivos. En los sistemas ópticos con marcadores pasivos (Vicon<sup>TM</sup>), se colocan al sujeto marcadores reflectantes en todos los puntos de interés de captura de movimiento, a diferencia de los sistemas ópticos activos (Optotrak<sup>TM</sup>), donde se colocan marcadores activos (diodos emisores de luz [LED]). Una desventaja en los sistemas ópticos con marcadores es la oclusión; aquí los marcadores no aparecen en varias tomas de la cámara, debido a la obstrucción de la línea de visión de los objetos de la escena o por otras partes del cuerpo del sujeto. La mayoría de los paquetes de pos-procesamiento comerciales tienen la capacidad de tratar con los marcadores ocluidos; para ello se crean marcadores virtuales para sustituir la información de los ocluidos, o pueden usar marcadores redundantes (más que el mínimo requerido en el protocolo estándar), para compensar los marcadores ocluidos.

El movimiento de los marcadores se suele utilizar para deducir el movimiento relativo entre dos segmentos consecutivos, con el objetivo de definir con precisión el movimiento de una articulación. El movimiento de la piel (donde se coloca el marcador), en relación con el hueso subyacente, es el principal factor que limita la aplicación de algunos sensores [4], [14], [22].

Los sistemas de captura de movimiento sin marcadores, como la cámara Microsoft Kinect y Organic Motion<sup>TM</sup>, ofrecen una alternativa distinta a la tecnología de captura de movimiento [20]. El Kinect es una cámara capaz de estimar la geometría 3D de la escena adquirida, a 30 cuadros por segundo; está construida con un sensor de profundidad de resolución espacial de  $640 \times 480$  píxeles, una vídeo-cámara VGA de la misma resolución y un arreglo de micrófonos para reconocimiento

de voz. El Kinect fue desarrollado como un dispositivo periférico para uso con la consola de juegos Xbox 360, aunque se ha adaptado fácilmente para otros campos como: la robótica [12], seguimiento del esqueleto humano [24], reconstrucción 3D [16], terapia asistencial [8] y biomecánica [23].

Con excepción del Kinect, la característica común de los sistemas descritos anteriormente es su alto costo y dificultad de implementación. Surge entonces el interrogante: ¿Es posible construir un sistema de captura de movimiento con Kinect de bajo costo y fácil implementación? En nuestro caso particular, un sistema de captura de movimiento permitirá extraer las trayectorias articulares de la marcha humana, con el fin de adaptarlas como referencias para el ciclo de marcha de un robot bípedo. Algunos autores han comparado el Kinect con otros sistemas de captura de movimiento con marcadores; en [7, 25], se comparan las variaciones de las posiciones cartesianas de un sistema de captura de movimiento con marcadores y Kinect; en [23] se comparan las variaciones de los ángulos de un brazo robótico obtenidas por Kinect y un sistema con marcadores. Estos aportes en el área, han sido trabajos sobre soluciones específicas que no integran la captura de movimiento para el cálculo de las trayectorias articulares de la marcha humana, propósito de este trabajo. En este estudio, se comparan las diferencias entre las trayectorias articulares obtenidas por un sistema de captura de movimiento con marcadores y Kinect; también se evalúa la variación de las medidas antropométricas entre los dos sistemas. Para cumplir con este objetivo, las piernas del sujeto de estudio se modelaron como una estructura robótica según los parámetros geométricos de Denavit-Hartenberg [11]. Después se desarrolló una aplicación para transformar las coordenadas cartesianas que entregan los dos sistemas de captura de movimiento en coordenadas articulares; en esta parte se propuso un nuevo algoritmo, basado en los mínimos cuadrados

para el cálculo de la cinemática inversa, utilizando las coordenadas cartesianas del experimento y el modelo geométrico de las piernas. Posteriormente, estos datos fueron analizados estadísticamente, mostrando el error y la varianza asociada a las medidas antropométricas y articulares.

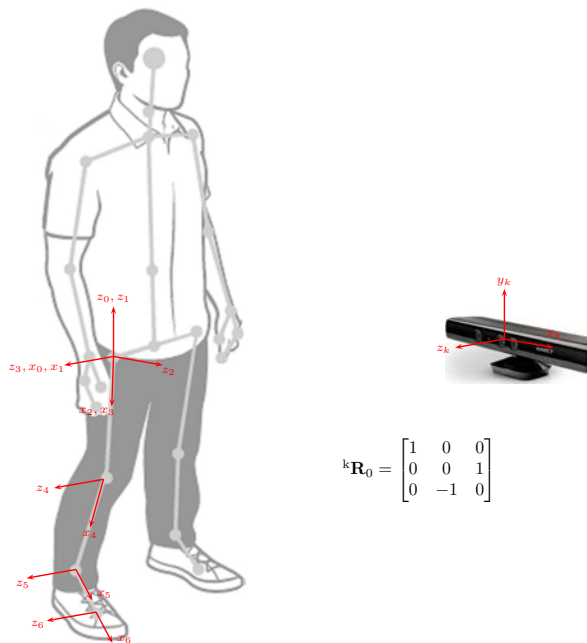
## DISEÑO DEL EXPERIMENTO

Para realizar el experimento de captura de movimiento, se dispuso de un sistema con 12 cámaras OptiTrack<sup>TM</sup> con resolución de  $1280 \times 1024$  píxeles, una caminadora con velocidad variable, un sensor Kinect y el software apropiado para la recolección de datos. El experimento se realizó en el Laboratorio de biomecánica de la Universidad Autónoma de Occidente. La idea básica de funcionamiento del sistema óptico utilizado, es que cada una de las cámaras envía luz infrarroja que se refleja de regreso cuando choca con una superficie reflectante. La posición cartesiana  $(x, y, z)$  del centro de un marcador reflectante (normalmente esférico) puede ser identificado por la intersección de las luces infrarrojas de tres cámaras. Kinect entrega también, la posición cartesiana  $(x, y, z)$  de 20 puntos articulares del esqueleto óptico del usuario; para miembros inferiores son 8 vectores que contienen las tres posiciones cartesianas de cada uno. Dichos vectores, corresponden a la cadera, la rodilla, el tobillo y la posición del metatarso del pie en cada una de las piernas.

La Figura 2 muestra al sujeto de estudio con 15 marcadores unidos a los miembros inferiores de su cuerpo. La distribución de marcadores en el cuerpo sigue un protocolo determinado. El protocolo Helen Hayes, implementado en este experimento, es un protocolo estándar muy utilizado para el análisis de la marcha [6]. El sujeto de prueba fue una mujer sana sin antecedentes de problemas músculo-esqueléticos. La altura fue de 154 *cm*, su peso fue 54 *kg* y la edad de 22 años; su participación en este estudio fue voluntaria. Este trabajo contó con el aval



**Figura 2.** Puesta de marcadores según el protocolo Helen Hayes.



**Figura 3.** Ubicación de ejes para cada articulación y el referente del sensor.

del Comité de Ética de la Vice-rectoría de Investigaciones de la Universidad del Cauca (Colombia).

Cada pierna se modeló con los parámetros geométricos de Denavit-Hartenberg [11]. En la Figura 3, se muestra la ubicación de ejes para la pierna derecha; en este diagrama nos interesa ver la evolución temporal de las articulaciones de la Tabla 1.

**Tabla 1.** Coordenadas articulares de la pierna derecha.

$\theta$ [°]	Coordenada articular
$q_1$	Cadera en el plano transversal
$q_2$	Cadera en el plano frontal
$q_3$	Cadera en el plano sagital
$q_4$	Rodilla en el plano sagital
$q_5$	Tobillo en el plano sagital

**Tabla 2.** Parámetros geométricos de la pierna.

$j$	$i$	$\sigma_j$	$\alpha_j$	$d_j$	$q_j$	$r_j$
1	0	0	0	0	$q_1$	0
2	1	0	$-\pi/2$	0	$q_2 + \pi/2$	0
3	2	0	$\pi/2$	0	$q_3$	0
4	3	0	0	$D3$	$q_4$	0
5	4	0	0	$D4$	$q_5 + \pi/2$	0
6	5	0	0	$D5$	0	0

La tabla de parámetros geométricos de la pierna derecha para la Figura 3 se muestra en la Tabla 2, donde  $D3$  es la longitud del fémur,  $D4$  es la longitud de la tibia y  $D5$  es la distancia que hay entre el tobillo y la posición del metatarso del pie. Las coordenadas articulares  $q_1, q_2, \dots, q_5$  son los ángulos en los distintos planos: transversal, sagital y frontal; la especificación de cada uno de ellos se puede leer en la Tabla 1.

La selección de los parámetros y los grados de libertad de la Tabla 1 se hizo con el fin de adaptar las trayectorias articulares de un ser humano a las trayectorias angulares que deben seguir los servomotores del robot bípedo Bioloid. Es por esta razón que las articulaciones de los miembros inferiores se modelan como si fueran las piernas de un robot. Para que las trayectorias articulares sean independientes del marco de referencia del sistema de captura de movimiento (Vicon o Kinect), se hace una transformación homogénea entre el referente del sistema de visión y el marco de referencia cero de la pierna, [9]. Los datos para ambos sistemas fueron adquiridos mediante archivos de texto en formato ASCII (\*.txt). Las coordenadas de cada marcador se registraron en metros, estos archivos fueron importados en Matlab

© y se transformaron los datos de ambos sistemas en alineación con el marco de referencia cero de la pierna mediante una matriz de transformación homogénea (2), como se explica en la siguiente sección.

### Modelo geométrico directo de la pierna

El modelo geométrico directo permite calcular la posición cartesiana y orientación de cada una de las articulaciones de la cadena cinemática de la Figura 3 en el referente del Kinect  $\langle x_k, y_k, z_k \rangle$ . El vector que denota la posición cartesiana del referente cero con respecto al marco de referencia inercial (Sistema de Captura de movimiento, Vicon o Kinect) es  ${}^s p_0$  y la matriz de rotación que denota la orientación del referente cero con respecto al marco de referencia inercial es  ${}^s \mathbf{R}_0$ . El sentido de la marcha en el eje  $x$ . La matriz de rotación se escribe en términos de los ángulos de euler  $\phi$ ,  $\theta$  y  $\psi$ .

$${}^s \mathbf{R}_0(\phi, \theta, \psi) = \begin{bmatrix} C\phi C\theta & C\phi S\theta S\psi - S\phi C\psi & C\phi S\theta C\psi - S\phi S\psi \\ S\phi C\theta & S\phi S\theta S\psi + C\phi C\psi & S\phi S\theta C\psi - C\phi S\psi \\ -S\theta & C\theta S\psi & C\theta C\psi \end{bmatrix} \quad (1)$$

Donde

$$\begin{aligned} S\phi &\triangleq \sin \phi & C\phi &\triangleq \cos \phi \\ S\theta &\triangleq \sin \theta & C\theta &\triangleq \cos \theta \\ S\psi &\triangleq \sin \psi & C\psi &\triangleq \cos \psi \end{aligned}$$

Para facilidad en la notación se define  $\sigma \triangleq \{\phi, \theta, \psi\}$ , entonces la matriz de transformación para pasar del marco de referencia cero al referente del Kinect en términos de  ${}^s \mathbf{R}_0(\sigma)$  y  ${}^s p_0$  es

$${}^s \mathbf{T}_0(\sigma) = \begin{bmatrix} {}^s \mathbf{R}_0(\sigma) & {}^s p_0 \\ \mathbf{0} & 1 \end{bmatrix} \quad (2)$$

La transformación  ${}^s \mathbf{T}_0(\sigma)$  permite expresar la posición de la cadera en términos del marco de referencia inercial. La matriz de transformación para pasar del marco de

referencia  $j$  a  $i$  está definida por un vector  ${}^i p_j$  que expresa la posición del marco  $j$  en  $i$  y una matriz de rotación  ${}^i \mathbf{R}_j$  que relaciona la orientación del marco  $j$  en  $i$ . La matriz de rotación depende, además, del vector de posiciones articulares  $q$ .

$${}^i \mathbf{T}_j(q) = \begin{bmatrix} {}^i \mathbf{R}_j(q) & {}^i p_j \\ \mathbf{0} & 1 \end{bmatrix}$$

Donde:

$$\begin{aligned} {}^i \mathbf{R}_j(q) &= \begin{bmatrix} Cq_j & -Sq_j & 0 \\ C\alpha_j Sq_j & C\alpha_j Cq_j & -S\alpha_j \\ S\alpha_j Sq_j & S\alpha_j Cq_j & C\alpha_j \end{bmatrix} \\ {}^i p_j &= \begin{bmatrix} d_j \\ -r_j S\alpha_j \\ r_j C\alpha_j \end{bmatrix} \end{aligned}$$

Los parámetros  $\alpha_j$ ,  $d_j$ ,  $q_j$  y  $r_j$  fueron presentados en la Tabla 2.

### Trayectorias articulares como un problema de optimización

La anatomía humana se puede representar mediante una secuencia de cuerpos rígidos conectados por articulaciones [1]. Kinect permite el seguimiento óptico del esqueleto humano y entrega las coordenadas cartesianas  $(x, y, z)$  de 20 puntos articulares. Por otra parte, el sistema de marcadores ofrece las coordenadas cartesianas  $(x, y, z)$  de cada uno de los marcadores adheridos al sujeto de estudio según el protocolo Helen Hayes. En este trabajo estamos interesados en las trayectorias articulares de los miembros inferiores en el plano sagital y frontal. Por lo tanto, fue necesario desarrollar una aplicación para transformar las coordenadas cartesianas  $(x, y, z)$  en coordenadas articulares  $(q_1, q_2, \dots, q_n)$ . Generalmente este problema se resuelve utilizando cinemática inversa [17]. Los intentos por resolver el problema de cinemática inversa en robótica, han utilizado numerosos enfoques que van desde iteraciones sucesivas del modelo geométrico, diversas técnicas de optimización [3] y métodos



basados en aprendizaje, [10], [26]. El problema de cinemática inversa consiste en mapear del espacio cartesiano al espacio articular, esto significa encontrar la función no lineal:

$$\mathbf{q}^{ref} = \mathbf{f}^{-1}(\mathbf{X}_h) \quad (3)$$

Donde  $\mathbf{q}^{ref} \in \mathbb{R}^n$  es el vector de coordenadas articulares,  $\mathbf{X}_h \in \mathbb{R}^{3m}$  es el vector con las coordenadas cartesianas obtenidas mediante el sistema de captura de movimiento. Siendo  $m$  el número de marcadores considerados y  $n$  el número de grados de libertad del robot. Las técnicas analíticas tiene la ventaja de un cálculo más rápido, pero están limitadas a robots con un pequeño número de articulaciones. Los métodos basados en optimización son más genéricos y tienen la ventaja de incorporar restricciones físicas y/o geométricas.

Con el fin de resolver el problema de cinemática inversa por optimización, la función de costo (4) llamada *error de marcador*, es minimizada. La función *error de marcador*, se define como la distancia entre un marcador (o centro de la articulación) experimental y el marcador virtual correspondiente. En cada intervalo de tiempo, el algoritmo de mínimos cuadrados calcula un vector de coordenadas generalizadas  $\mathbf{q}^{ref}(t_k)$ , que minimiza la suma de la norma del error de marcador.

$$\min_{\mathbf{q}^{ref}(t_k)} \left[ \sum_{i=1}^m \left\| \mathbf{X}_i^{exp}(t_k) - \hat{\mathbf{X}}_i(\mathbf{q}(t_k)) \right\|^2 \right] \quad (4)$$

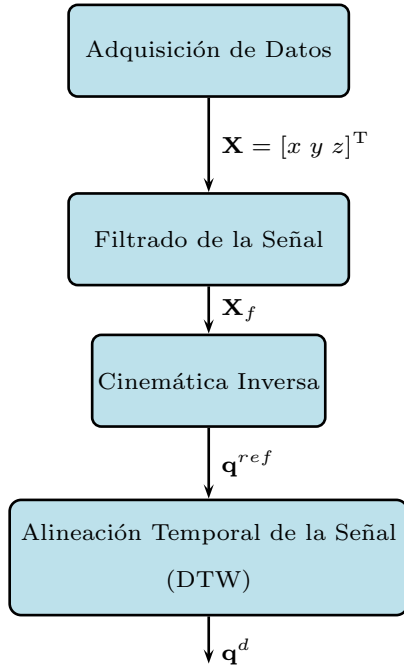
Donde  $\mathbf{X}_i^{exp}(t_k) \in \mathbb{R}^3$  denota la posición cartesiana del marcador experimental  $i$  medida en el instante  $t = t_k$  por el sistema de captura de movimiento y  $\hat{\mathbf{X}}_i(\mathbf{q}(t_k)) \in \mathbb{R}^3$  son las posiciones cartesianas estimadas, asociadas a un vector  $\mathbf{q}(t_k)$  dado;  $\hat{\mathbf{X}}_i(\cdot)$  es la función que contiene el modelo geométrico directo del robot. Una vez que cada una de las trayectorias articulares  $\mathbf{q}^{ref}$  se calculan, el robot podría realizar un seguimiento de las trayectorias articulares con

el controlador apropiado. Sin embargo, hay muchas diferencias entre un robot bípedo y un ser humano real. Esto significa que no podemos simplemente enviar estos datos al robot y esperar que se mueva correctamente.

## Procesamiento de las Señales de Captura de Movimiento

El procedimiento utilizado para calcular las trayectorias angulares de la pierna derecha en un ciclo normal del marcha se muestra en la Figura 4.

La Figura 4 muestra el diagrama en bloques del procesamiento de la señal de captura de movimiento. Primero, la *adquisición de datos* se hace simultáneamente para ambos sistemas de captura, Vicon a 240 Hz y Kinect a 30 Hz durante un periodo de 3 segundos, este tiempo es suficiente para obtener las trayectorias angulares de la pierna derecha durante un ciclo de marcha. Ambos sistemas entregan un vector  $\mathbf{X}$  que contiene las coordenadas cartesianas de todas las articulaciones, cada uno de sus componentes es un vector en  $\mathbb{R}^3$ . El bloque *filtrado de la señal* se encarga de eliminar frecuencias no deseadas, se implementó un filtro Butterworth con  $f_c = \frac{f_s}{2}$ , se utilizó la interpolación spline cúbica entre cada par de datos sucesivos para asegurar la continuidad de las trayectorias. Las trayectorias cartesianas filtradas  $\mathbf{X}_f$  se transforman en trayectorias articulares  $\mathbf{q}^{ref}$  mediante la *cinemática inversa*. Finalmente con el propósito de comparar cualitativamente las trayectorias medidas por ambos sistemas de captura de movimiento se deben graficar las trayectorias medidas por los dos sistemas de captura. Sin embargo como los sistemas se ejecutan a diferentes frecuencias de muestreo y el instante de inicio no es el mismo para los dos, para alinear en el tiempo las trayectorias se utilizó el algoritmo de procesamiento de señales *Dynamic Time Warping* que permite alinear en el dominio del tiempo dos señales capturadas a distintas frecuencias de muestreo [2].



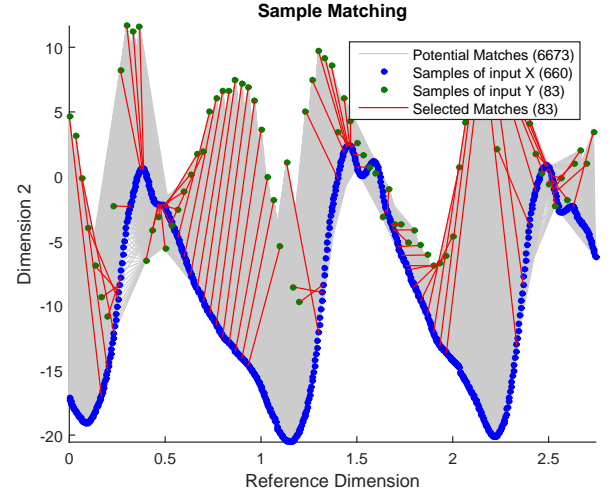
**Figura 4.** Diagrama en bloques del procesamiento de la señal.

## RESULTADOS Y DISCUSIÓN

El procesamiento de las señales de captura de movimiento es un factor muy importante para el análisis de resultados. En esta sección, se muestran los resultados de las trayectorias angulares de la pierna derecha en un ciclo normal de marcha, las diferencias obtenidas por medio de ambos sistemas de captura de movimiento se calculan en forma gráfica y cuantitativa.

### Dynamic Time Warping

DTW por sus siglas en inglés es un algoritmo para encontrar un alineamiento óptimo de dos secuencias (dependientes del tiempo). Esta técnica ha sido usada para aplicaciones del reconocimiento de voz, y clasificación del movimiento humano [28].



**Figura 5.** DTW para la articulación  $q_3$ .

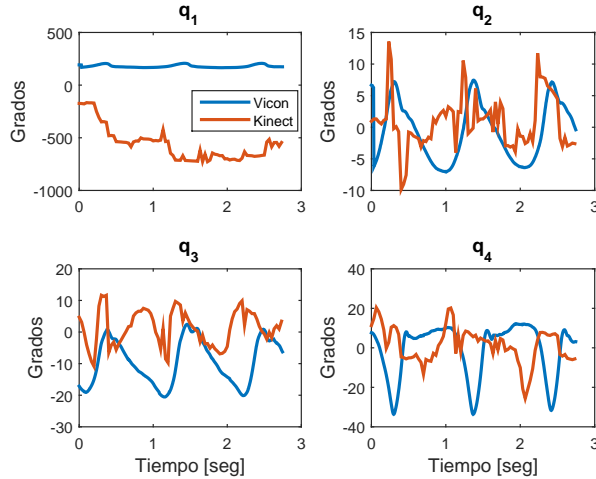
Dadas dos series temporales  $\mathbf{X} = [x_1, \dots, x_{n_x}]^T \in \mathbb{R}^{n_x}$  y  $\mathbf{Y} = [y_1, \dots, y_{n_y}]^T \in \mathbb{R}^{n_y}$ , DTW permite alinear  $\mathbf{X}$  y  $\mathbf{Y}$  tal que la siguiente norma es minimizada:

$$\|\mathbf{X}W_x - \mathbf{Y}W_y\|^2 \quad (5)$$

Donde  $W_x$  y  $W_y$  son matrices con la dimensión adecuada que solamente contienen los valores discretos  $[0, 1]$ , el algoritmo se puede estudiar en detalle en [21]. En la Figura 5 se muestra el resultado de alineamiento temporal para la articulación de la cadera en el plano sagital,  $q_3$ , aquí se puede inferir que al aplicar DTW para ambas señales (Vicon y Kinect) el algoritmo selecciona 83 puntos que minimizan la Ecuación (5).

Las principales articulaciones asociadas con la locomoción humana son: la cadera, la rodilla, tobillos y el metatarso del pie. El movimiento de la cadera combinada con la rotación de la pelvis permite alargar el paso, [5]. Durante un ciclo de marcha, el movimiento de la cadera en el plano sagital es esencialmente sinusoidal. Por lo tanto, el fémur se mueve de atrás hacia adelante y viceversa. La articulación de la rodilla permite los movimientos de flexión y extensión de la pierna durante la locomoción. La Figura 6 muestra la comparación de trayectorias angulares obtenidas por el procedimiento indicado en la Figura 4.





**Figura 6.** Trayectorias angulares pierna derecha.

**Tabla 3.** Errores RMS para cada articulación.

$\theta$ [°]	RMS
$q_1$	756.22
$q_2$	5.95
$q_3$	14.07
$q_4$	19.04

Para comparar de manera cuantitativa las trayectorias se estableció como métrica el error RMS, definido como:

$$\text{RMS} = \sqrt{\frac{1}{N} \sum_{i=1}^N (K_i - V_i)^2} \quad (6)$$

Donde  $N$  es el número total de datos utilizado en el intervalo de tiempo de duración de la trayectoria,  $K_i$  es la medida del Kinect y  $V_i$  es la medida de Vicon para una trayectoria articular dada. Los resultados obtenidos para cada articulación se presentan en la Tabla 3.

Existe una diferencia entre los dos conjuntos de datos debido a los diferentes métodos de captura de posición articular. El Kinect utiliza algoritmos de procesamiento de imágenes basado en la silueta del participante para las coordenadas  $X$  e  $Y$  y se basa en un sensor de profundidad infrarrojo para la coordenada  $Z$ .

**Tabla 4.** Medidas Antropométricas del Femur obtenidas por tres sistemas de medición.

Sistema	$\mu$ [m]	$\sigma$ [m]
Vicon	0.418	0.013
Kinect	0.413	0.003
Cinta Métrica	$0,418 \pm 0,001$ m	

El algoritmo de esqueletización del Kinect es en esencia, sólo una estimación de la posición de las articulaciones, debido al algoritmo de captura de movimiento sin marcadores [20]; mientras que Vicon triangula el punto central de los grupos de marcadores colocados en las posiciones articulares de interés. Este factor determina el que error RMS sea mayor para la articulación  $q_1$ , correspondiente a la rotación de la cadera en el plano transversal.

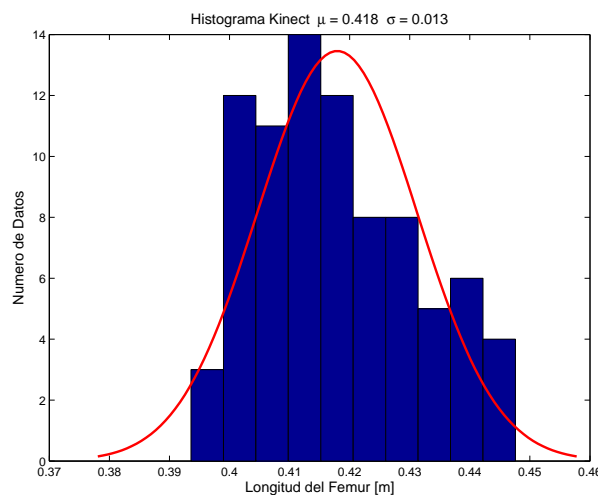
Para comparar las medidas antropométricas del fémur se midió directamente la longitud del fémur del sujeto de estudio con una cinta métrica, con el propósito de comparar las medidas obtenidas con los sistemas ópticos y un tercer sistema de referencia. Los resultados se presentan en la Tabla 4.

La Figura 7 muestra una distribución de los datos obtenidos por Kinect para la longitud del fémur, estos se aproximaron por una curva Gaussiana de media igual a 0.418 m, con una desviación estándar de 0.013 m, mientras que la Figura 7b muestra una distribución gaussiana para la longitud del fémur tomada con Vicon de media igual a 0.413 m, con una desviación estándar de 0.003 m. Aunque ambos sistemas coinciden con dos cifras significativas en la longitud del fémur, el sistema Vicon resulta ser  $0,013/0,03 \approx 12$  veces más preciso que el Kinect.

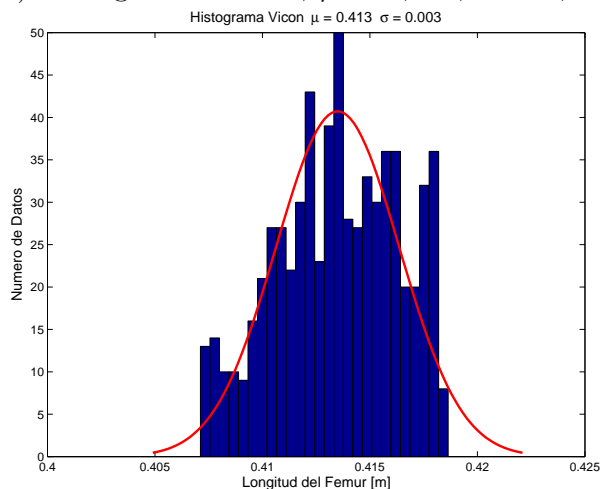
Los estudios para determinar la resolución espacial del Kinect en las coordenadas  $X, Y, Z$ , evidencian que el error asociado a la medición de las coordenadas cartesianas aumenta a medida que el sujeto de análisis se aleje del sensor (escala logarítmica), obteniendo valores mínimos y máximos de: 1 mm hasta 6.5 mm para las coordenadas  $X$  e  $Y$ ; y 2 mm hasta 5 cm en la coordenada  $Z$ , para valores de distancia del sujeto de estudio

al sensor de 0.8 m hasta 4 m respectivamente, [13]. Las desviaciones en la medidas de las coordenadas es considerable frente a la presencia de fuentes parásitas infrarrojas, por tal motivo es necesario realizar un adecuado control de estas.

Aunque este trabajo solamente contó con un participante, en la Ref. [19] donde se tuvieron 25 sujetos de estudio (para una mayor validez estadística), los autores llegaron a la misma conclusión principal presentada aquí.



a) Histograma Kinect,  $\mu = 0,418$ ,  $\sigma = 0,013$



b) Histograma Vicon,  $\mu = 0,413$ ,  $\sigma = 0,003$

**Figura 7.** Comparación de histogramas de la longitud del fémur. La distribución de datos en la longitud del fémur es aproximada por una curva Gaussiana para ambos sistemas de captura.

## CONCLUSIONES

En este trabajo se compararon las trayectorias de la marcha humana obtenidas mediante dos sistemas de captura de movimiento. Para ello se implementó un algoritmo para la estimación de las posiciones articulares humanas a partir de las posiciones cartesianas mediante ajuste por mínimos cuadrados. Los resultados antropométricos obtenidos con Kinect mostraron una mayor varianza en relación a los obtenidos con Vicon en una razón 1/12. Los movimientos a lo largo del plano  $Z$  afectan significativamente la precisión de las lecturas basadas en Kinect, en consecuencia las trayectorias articulares no coinciden con el sistema Vicon.

La sincronización de los sistemas de captura de movimiento es esencial, aunque ambos sistemas trabajan a frecuencias de muestreo distintas, el algoritmo DTW permite alinear en el dominio del tiempo ambas señales y permite comparar de manera cualitativa las dos formas de onda. El sistema de captura de movimiento con Kinect es de bajo costo y fácil instalación y se presenta como una opción para realizar un estudio preliminar de los parámetros cinemáticos de la marcha humana. Sin embargo, el análisis de los datos de validación indica que el sensor Kinect no se puede emplear para hacer un análisis de movimiento humano preciso. La mejora de los algoritmos y métodos experimentales para la captura de movimiento humano con Kinect sigue siendo una área activa e interesante para la investigación.

## Agradecimientos

Los autores expresan sus más sinceros agradecimientos a la Universidad del Cauca y a la Universidad Autónoma de Occidente en Colombia por todo el apoyo académico y financiero brindado en este proyecto.

## REFERENCIAS

- [1] K. Abdel-Malek and J. Arora, *Human Motion Simulation: Predictive Dynamics*. Elsevier Science, 2013.
- [2] K. Adistambha, C. Ritz, and I. Burnett, "Motion classification using dynamic time warping," in *Multimedia Signal Processing, 2008 IEEE 10th Workshop on*, Oct 2008, pp. 622–627.
- [3] K. Ayusawa and Y. Nakamura, "Fast inverse kinematics algorithm for large dof system with decomposed gradient computation based on recursive formulation of equilibrium," in *Intelligent Robots and Systems (IROS), 2012 IEEE/RSJ International Conference on*, Oct 2012, pp. 3447–3452.
- [4] A. Cappozzo, A. Cappello, U. d. Croce, and F. Pensalfini, "Surface-marker cluster design criteria for 3-d bone movement reconstruction," in *Biomedical Engineering*, 1997.
- [5] C. Chevallerau, G. Bessonnet, G. Abba, and Y. Aoustin, *Bipedal Robots. Modeling, design and building walking robots*, 1st ed. Wiley, 2009.
- [6] M. P. A. Chris Kirtley, *Clinical Gait Analysis. Theory and Practice*, 1st ed. Churchill Livingstone, 2006.
- [7] R. A. Clark, K. J. Bower, B. F. Mentiplay, K. Paterson, and Y.-H. Pua, "Concurrent validity of the microsoft kinect for assessment of spatiotemporal gait variables," *Journal of Biomechanics*, vol. 46, no. 15, pp. 2722 – 2725, 2013.
- [8] R. A. Clark, Y.-H. Pua, K. Fortin, C. Ritchie, K. E. Webster, L. Denehy, and A. L. Bryant, "Validity of the microsoft kinect for assessment of postural control," *Gait & Posture*, vol. 36, no. 3, pp. 372 – 377, 2012.
- [9] P. I. Corke, *Robotics, vision and control : fundamental algorithms in Matlab*, 1st ed., ser. Star, 73.; Springer tracts in advanced robotics, 73. Springer, 2011.
- [10] B. Damas and J. Santos-Victor, "An online algorithm for simultaneously learning forward and inverse kinematics," in *Intelligent Robots and Systems (IROS), 2012 IEEE/RSJ International Conference on*, Oct 2012, pp. 1499–1506.
- [11] J. Denavit and R. S. Hartenberg, "A kinematic notation for lower-pair mechanisms based on matrices," *Trans. ASME, J. Appl. Mech.*, vol. 22, no. 2, pp. 215 – 221, 1965.
- [12] G. Du, P. Zhang, J. Mai, and Z. Li., "Markerless kinect-based hand tracking for robot teleoperation." *International Journal of Advanced Robotic Systems*, 2012.
- [13] T. Dutta, "Evaluation of the kinect sensor for 3-d kinematic measurement in the workplace," *Applied Ergonomics*, vol. 43, no. 4, pp. 645 – 649, 2012.
- [14] J. P. Holden, J. A. Orsini, K. L. Siegel, T. M. Kepple, L. H. Gerber, and S. J. Stanhope, "Surface movement errors in shank kinematics and knee kinetics during gait," *Gait & Posture*, vol. 5, no. 3, pp. 217 – 227, 1997.
- [15] V. Ivancevic and T. Ivancevic, *Human-Like Biomechanics: A Unified Mathematical Approach to Human Biomechanics and Humanoid Robotics*, ser. Intelligent Systems, Control and Automation: Science and Engineering, v. 28. Springer, 2008.
- [16] S. Izadi, D. Kim, O. Hilliges, D. Molyneaux, R. Newcombe, P. Kohli, J. Shotton, S. Hodges, D. Freeman, A. Davison, and A. Fitzgibbon, "Kinect-fusion: Real-time 3d reconstruction

- and interaction using a moving depth camera,” in *Proceedings of the 24th Annual ACM Symposium on User Interface Software and Technology*, ser. UIST '11. New York, NY, USA: ACM, 2011, pp. 559–568.
- [17] W. Khalil and E. Dombre, *Modeling, Identification and Control of Robots*, 2nd ed., ser. Kogan Page Science. Paris, France: Butterworth - Heinemann, 2004.
- [18] S. J. Lee, Y. Motai, and H. Choi, “Tracking human motion with multichannel interacting multiple model,” *Industrial Informatics, IEEE Transactions on*, vol. 9, no. 3, pp. 1751–1763, Aug 2013.
- [19] M. J. Malinowski, E. Matsinos, and S. Roth, “On using the Microsoft Kinect™ sensors in the analysis of human motion,” *ArXiv e-prints*, Dec. 2014.
- [20] C. D. Mutto, P. Zanuttigh, and G. M. Cortelazzo, *Time-of-Flight Cameras and Microsoft Kinect(TM)*. Springer Publishing Company, Incorporated, 2012.
- [21] L. Rabiner and B.-H. Juang, *Fundamentals of Speech Recognition*. Upper Saddle River, NJ, USA: Prentice-Hall, Inc., 1993.
- [22] C. Reinschmidt, A. van den Bogert, B. Nigg, A. Lundberg, and N. Murphy, “Effect of skin movement on the analysis of skeletal knee joint motion during running,” *Journal of Biomechanics*, vol. 30, pp. 729–732, 1997.
- [23] A. Schmitz, M. Ye, R. Shapiro, R. Yang, and B. Noehren, “Accuracy and repeatability of joint angles measured using a single camera markerless motion capture system,” *Journal of Biomechanics*, vol. 47, no. 2, pp. 587 – 591, 2014.
- [24] L. A. Schwarz, A. Mkhitarian, D. Mateus, and N. Navab, “Human skeleton tracking from depth data using geodesic distances and optical flow,” *Image and Vision Computing*, vol. 30, no. 3, pp. 217 – 226, 2012.
- [25] E. Stone and M. Skubic, “Passive in-home measurement of stride-to-stride gait variability comparing vision and kinect sensing,” in *Engineering in Medicine and Biology Society, EMBC, 2011 Annual International Conference of the IEEE*, Aug 2011, pp. 6491–6494.
- [26] H. Toshani and M. Farrokhi, “Real-time inverse kinematics of redundant manipulators using neural networks and quadratic programming: A lyapunov-based approach,” *Robotics and Autonomous Systems*, vol. 62, no. 6, pp. 766 – 781, 2014.
- [27] J.-T. Zhang, A. C. Novak, B. Brouwer, and Q. Li, “Concurrent validation of xsens mvn measurement of lower limb joint angular kinematics,” *Physiological Measurement*, vol. 34, no. 8, p. N63, 2013.
- [28] F. Zhou and F. De la Torre, “Generalized time warping for multi-modal alignment of human motion,” in *Computer Vision and Pattern Recognition (CVPR), 2012 IEEE Conference on*, June 2012, pp. 1282–1289.

