

# Identificación de posturas del cuerpo humano utilizando un sistema de captura de movimiento

DAVID LÓPEZ<sup>1</sup>, ANTONIO BENITEZ<sup>1</sup>, MA. AUXILIO MEDINA<sup>1</sup>, JORGE DE LA CALLEJA<sup>1</sup>, AND MARCO A. ALONSO<sup>1</sup>

<sup>1</sup>Universidad Politécnica de Puebla, Departamento de Posgrado, 3er carril del Ejido Serrano S/N, San Mateo Cuanalá, Juan C. Bonilla, Puebla, México, 72640

Compiled 25 de septiembre de 2017

Este trabajo describe como identificar posturas humanas utilizando un dispositivo de captura de movimiento llamado Kinect. Al recuperar las configuraciones del cuerpo humano se obtienen comandos para ejecutar algunas tareas. La identificación de las posturas se hace utilizando una plantilla que almacena datos espaciales asociados a un conjunto de articulaciones en relación a un sistema de referencia global. Se presenta una propuesta para el análisis de error en algunos movimientos del cuerpo humano.

© 2017 Universidad Politécnica de Puebla

**Palabras clave:** Captura de Movimiento, identificación de posturas

<http://dx.doi.org/10.1364/ao.XX.XXXXXX>

## 1. INTRODUCCIÓN

El uso de las computadoras en un principio estaba restringido a profesionales del area debido primordialmente a la complejidad en el uso de las mismas; limite que poco a poco se ha reducido por los avances en sistemas de interaccion basados en el usuario que van desde entornos de texto hasta la creacion de avatares[1] que permiten al usuario ser representado como una entidad digital[2] en entornos que van desde la comunicacion hasta trabajo colaborativo pasando por plataformas dedicadas al entretenimiento como en la figura 1.

De igual forma, la interacción entre humanos y robots era en un inicio una situación que debía evitarse por los pocos avances y por los posibles errores que se presentaban en las distintas plataformas robóticas, lo cual generaba ambientes peligrosos para el ser humano.

Hoy en día debido a los grandes avances en informática, control y telecomunicaciones, se ha transformado en una condición de uso en ambientes como empresas, escuelas, hospitales (figura : 2)

Dentro de la clasificación de los robots, los humanoides son los que presentan el mayor potencial para ayudar a personas en el hogar, oficinas, hospitales y en particular sus áreas de rehabilitación según se demuestra en [3][4][5]. Sin embargo, si un robot se diseña con el objetivo de desempeñarse en estos entornos, es necesario que su interacción sea diseñada por humanos por medio de movimientos afines a la tarea que va a desempeñar.

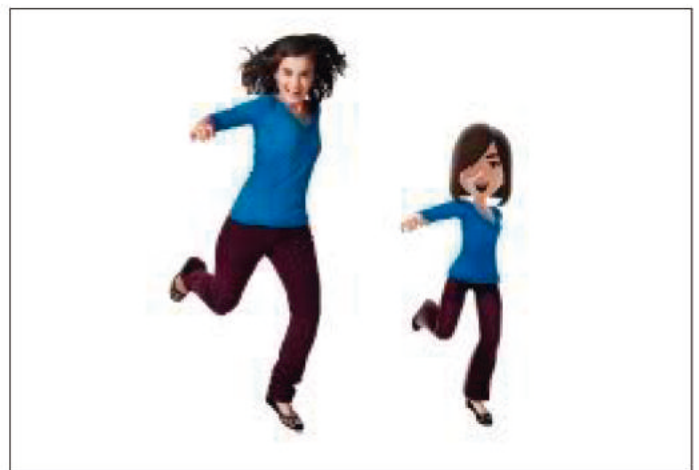


Fig. 1. Avatar digital de la plataforma de juegos de Microsoft.



Fig. 2. Darwin-OP en un entorno de trabajo con niños

Estas nuevas aplicaciones han derivado en la necesidad de sistemas que puedan capturar la información generada por el movimiento del cuerpo humano, el cual posee una estructura y manejo complejas, lo cual ha desencadenado una amplia gama de investigaciones concernientes al aprendizaje y simplificación de los datos obtenidos de la observación del movimiento para su posterior replicación en sistemas[6][7].

Como resultados obtenidos de estas investigaciones se permite que los robots incrementen sus capacidades autónomas y al mismo tiempo su capacidad de reacción a eventos en los ambientes donde han sido insertados.

Aunque existen diversos trabajos desarrollados por investigadores para imitar y replicar el movimiento desde un cuerpo humano hacia sistemas como robots[6][7], ésta no deja de ser un área de investigación abierta y desafiante para las nuevas arquitecturas que día a día se generan.

Actualmente, gracias al manejo, la expansión, la reducción de precios y el amplio desarrollo en centros de investigación y universidades con propósitos educativos o de entretenimiento con robots humanoides de todos tamaños y complejidad así como el uso cotidiano de avatares digitales en entornos educativos[8], los sistemas de captura e imitación de movimiento humano, son solicitados y utilizados por una amplia gama de aplicaciones [9][10], de manera particular en el área médica.

El trabajo que se describe en este artículo tiene como objetivo identificar posturas humanas utilizando un dispositivo de captura de movimiento llamado Kinect para la identificación de posturas y movimientos del cuerpo humano.

Este artículo está organizado de la siguiente manera: en la Sección 2, se describe el dispositivo de captura Kinect, valorando sus capacidades en hardware y los controladores disponibles en software. En la Sección 3, se discute el problema de la identificación de movimientos humanos por medio de posturas y su posterior procesamiento. En la Sección 4, se presentan los resultados y las limitaciones al utilizar plantillas como método de identificación. Por último, en la Sección 5 se presentan las conclusiones, el trabajo en proceso y algunos trabajos a futuro.

**2. CAPTURA DE MOVIMIENTO**

La captura de movimiento, también conocida como Motion Capture o simplemente MOCAP, es una técnica que sirve para almacenar movimientos de situaciones reales en formatos digitales[11].

Los dispositivos que se utilizan para esta tarea se apoyan en diversas técnicas que van desde el procesamiento de imágenes, hasta el uso de sensores electromecánicos para cumplir con la tarea asignada, lo cual ocasiona que el precio de estos sea dispositivos sea elevado, con precios en rangos de miles a decenas de miles de dólares[12].

Este trabajo utiliza el dispositivo Kinect que cubre las necesidades técnicas y además resulta económico puesto que su valor no sobrepasa los 100 dólares<sup>1</sup>.

**A. Kinect como dispositivo de captura**

Kinect es un dispositivo, inicialmente pensado como un simple controlador de juego, que gracias a los componentes que lo integran: sensor de profundidad, cámara RGB, arreglo de micrófonos y sensor de infrarrojos (emisor y receptor), es capaz de capturar el movimiento del cuerpo humano, reconocerlo y posicionarlo en un plano.

<sup>1</sup><https://www.microsoft.com/en-us/store/d/kinect-sensor-for-xbox-one/91hq5578vksc>

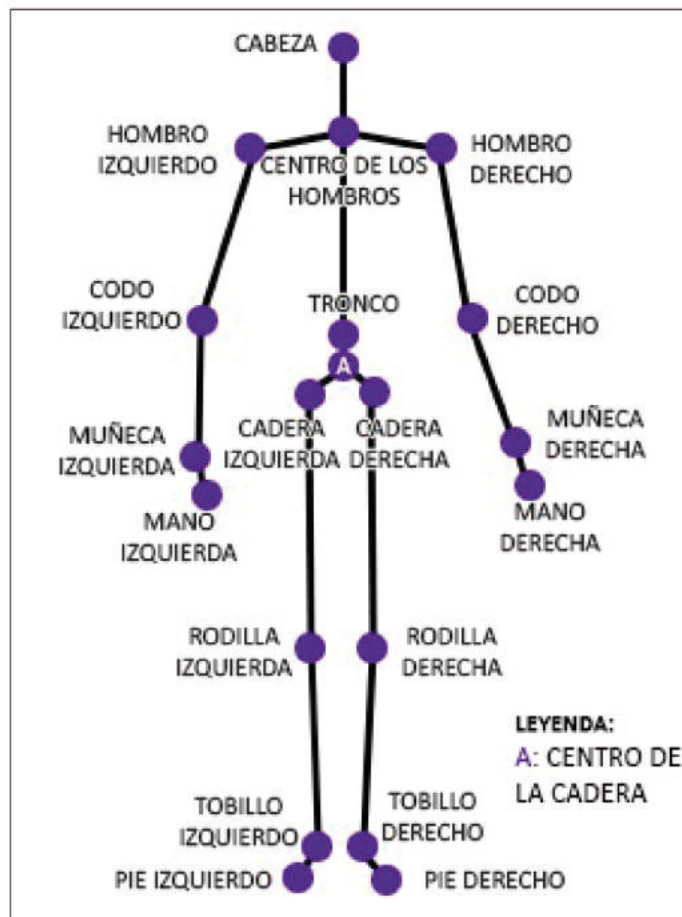
Gracias a toda la información que captura este dispositivo, los desarrolladores de software pueden hacer uso de él para programar toda una serie de aplicaciones cuya función principal es la interacción con elementos "virtuales" a través de los distintos movimientos del cuerpo humano[11].

Desde la salida de Kinect en 2010, se han publicado distintas librerías, controladores y APIs que han permitido desarrollar aplicaciones de escritorio con distintos fines. Algunas de esas plataformas son: el SDK de Windows para Kinect[13], OpenNI[14] y OpenKinect[15]. En este trabajo se utiliza el SDK de Microsoft, debido a sus continuas actualizaciones y el soporte para versiones actuales de Kinect.

**B. Windows SDK para Kinect**

Se utiliza el SDK de Windows como software de apoyo dadas las ventajas que ofrece para la creación de aplicaciones que detecten y formen un esqueleto o mapa de articulaciones simplificado a partir del procesamiento de la imagen infrarroja generada por un cuerpo humano posicionado frente al dispositivo[16].

El SDK en su versión 1.7 es capaz de reconocer de manera clara veinte puntos del cuerpo humano, cada uno de ellos con un identificador particular, según se puede ver en la figura 3<sup>2</sup>. A través del SDK también se pueden detectar las coordenadas 3D del conjunto de articulaciones antes mencionado, dicha tarea se denomina Seguimiento de Esqueleto (Skeletal Tracking)[17].



**Fig. 3.** Mapeo de articulaciones del esqueleto humano al esqueleto de Kinect.

<sup>2</sup><http://generacionxbox.com/kinect-para-xbox-one/>

Es muy importante tener en cuenta que todas las coordenadas que entrega Kinect están definidas en un sistema coordenado referenciado a él, tal y como se muestra en la figura 4<sup>3</sup>. A este sistema de coordenadas se le denominara con el subíndice 0.

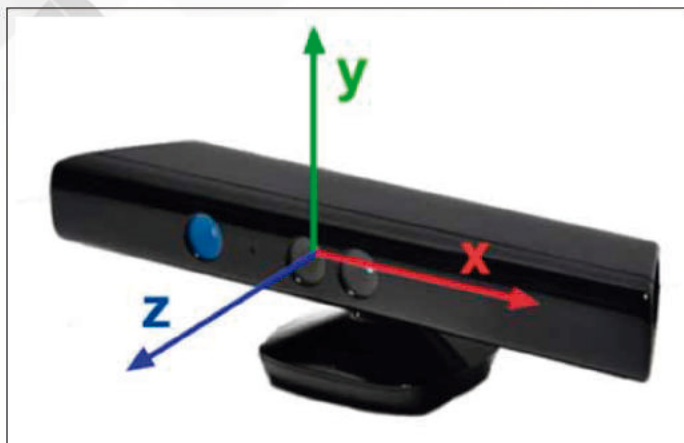


Fig. 4. Sistema de Coordenadas del Kinect.

Se debe considerar que cada articulación posee tres componentes en un espacio tridimensional(3D). Por ejemplo el punto que define la ubicación de la cabeza se denota  $0_p\text{CABEZA}$  y cada una de sus componentes se denota como sigue:

$$0_p\text{CABEZA} = \begin{bmatrix} 0_x\text{CABEZA} \\ 0_y\text{CABEZA} \\ 0_z\text{CABEZA} \end{bmatrix}$$

Las coordenadas 3D de cada una de las articulaciones se procesan para identificar posturas o sucesiones de posturas conocidas como gestos, que pueden ser interpretados como comandos que posteriormente se envíen a un agente como lo es un robot o un avatar virtual.

### 3. POSTURAS Y GESTOS

Como se mencionó con anterioridad, una postura se define como el estado integrado por un conjunto de coordenadas de un esqueleto en un instante de tiempo, y de esta forma un gesto se define como la sucesión de posturas a intervalos de tiempo definidos.

Este trabajo utiliza la técnica de generación de plantillas en formato XML para el almacenamiento de posturas identificables[18].

#### A. Selección de posturas

El sistema desarrollado en este proyecto se basa en el reconocimiento de posturas realizadas por personas de distinta morfología y diversas capacidades físicas, por lo cual las posturas deben de seleccionarse y analizarse de forma que se tomen en cuenta estos factores para su identificación.

Dados estos requerimientos, se seleccionaron posturas comunes en rutinas de ejercicio con diversas aplicaciones, que pueden ir desde mejorar el tono muscular hasta la rehabilitación.

La descripción y el análisis de las posturas se hicieron tomando en cuenta la anatomía del cuerpo humano y los alcances promedio de las articulaciones del mismo[19].

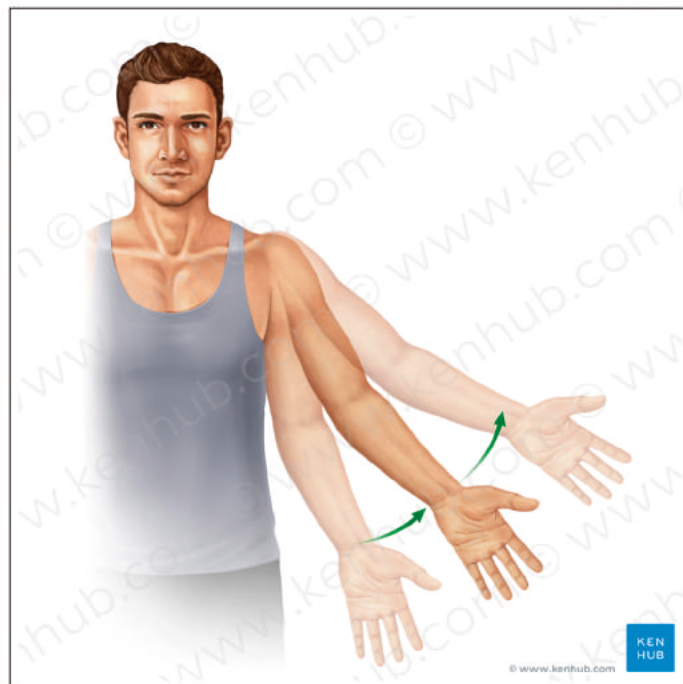


Fig. 5. Movimiento lateral del brazo en relación al hombro.

Como posturas de referencia se utilizaron la postura inicial y final ante los movimientos de abducción y aducción de los brazos extendidos a los lados en relación al torso, utilizando como articulación de referencia el hombro.

- Movimiento abducción y aducción:

Los hombros del cuerpo humano tienen la capacidad de moverse con un rango de movimiento lateral de  $0^\circ$  a  $180^\circ$  por los lados del cuerpo, tal como se define en la figura 5.

Para la representación del gesto de abducción del brazo izquierdo, se debe tomar como referencia  $0_p\text{HOMBRO\_IZQUIERDO}$  puesto que es la articulación que representa al hombro izquierdo en Kinect. A partir de esta información, se genera una plantilla en formato XML donde se vacía la información que representa las relaciones de las articulaciones que intervienen en las posturas relacionadas con el movimiento de abducción.

#### A.1. Posturas en piernas

También se utilizaron la postura inicial y final ante los movimientos de abducción y aducción de las piernas con relación al torso, utilizando como articulación de referencia la cadera como se muestra en la figura 6.

#### A.2. Posturas en brazos

- Movimiento abducción y aducción:

Las piernas del cuerpo humano tienen la capacidad de moverse con un rango de movimiento lateral hacia afuera con un rango de  $50^\circ$  y hacia adentro por  $35^\circ$ .

Para la representación del gesto de abducción de la pierna derecha, se debe tomar como referencia  $0_p\text{CADERA\_DRETERECHA}$  puesto que es la articulación que representa la sección derecha de la cadera en Kinect.

La formulación del archivo XML para la identificación de las posturas de brazos y piernas toma en cuenta una lista de parámetros descritos en la tabla 1.

<sup>3</sup><http://www.ergonautas.upv.es/lab/articulos/kinect.htm>



Fig. 6. Movimiento lateral de la pierna en relación a la cadera.

**Cuadro 1. Descripción del archivo XML**

Elemento	Atributo	Valor
Gestos	TiempoDeteccionGesto	Númérico en ms
	Descripcion	Texto
Gesto	TiempoMaximoEjecucion	Númérico en ms
	Permisividad	Decimal
	PrimerArticulacion	Nombre de articulación
Postura	SegundaArticulacion	Nombre de articulación
	RelacionInicial	Relación de cuadrante
	RelacionFinal	Relación de cuadrante

**B. Elementos del archivo XML**

**B.1. Cuadrantes**

El parámetro más importante en la detección de posturas esta dado por la relación de dos articulaciones, el proyecto ubica una articulación como el origen de un nuevo sistema de coordenadas y la otra se ubica en un cuadrante a partir de sus nuevas coordenadas en este sistema. La lista de relaciones que pueden tomar las articulaciones se expresa en la tabla 2.

**Cuadro 2. Relación de cuadrantes**

Coordenadas 2da Articulación	Relación
(x, y, z)	Derecha, Arriba y al Frente
(x, y, -z)	Derecha, Arriba y Atrás
(-x, y, z)	Izquierda, Arriba y al Frente
(-x, y, -z)	Izquierda, Arriba y Atrás
(x, -y, z)	Derecha, Abajo y al Frente
(x, -y, -z)	Derecha, Abajo y Atrás
(-x, -y, z)	Izquierda, Abajo y al Frente
(-x, -y, -z)	Izquierda, Abajo y Atrás
(x, y, 0)	Derecha, Arriba
(-x, y, 0)	Izquierda, Arriba
(x, -y, 0)	Derecha, Abajo
(-x, -y, 0)	Izquierda, Abajo

**B.2. Permisividad en posiciones**

Puesto que el sistema está diseñado para identificar posturas representadas por un ser humano, es idóneo mantener los parámetros de las posturas con cierta holgura puesto que la morfología y las capacidades físicas de las personas que pudiesen utilizar el sistema no son las mismas y el hecho de restringir las posiciones de las articulaciones a una medida angular exacta, ocasionaría que un alto porcentaje de posturas no fueran detectadas.

Para determinar el rango de permisividad en grados con el cual una articulación se puede detectar como presente en un cuadrante, se definio un conjunto de valores {0, 0,25, 0,5, 0,75, 1} que permite ampliar la identificación de posturas en cierto grado, como se observa en la tabla 3.

Esto permite refinar la detección de posturas dentro de un

**Cuadro 3. Permisividad para detectar una articulación dentro de un cuadrante**

Valor de permisividad	Valores de ángulo permitidos
0	45°
0.25	33.75° - 56.25°
0.5	22.5° - 67.5°
0.75	11.25° - 78.75°
1	0° - 90°

**Cuadro 4. Combinaciones utilizadas en el archivo XML para la detección de posturas**

No.	Tiempo Postura	Tiempo entre Posturas	Permisividad
1	100 ms	100 ms	0
2	100 ms	1000 ms	0
3	1000 ms	1000 ms	0
4	100 ms	100 ms	0.5
5	100 ms	1000 ms	0.5
6	1000 ms	1000 ms	0.5
7	100 ms	100 ms	1
8	100 ms	1000 ms	1
9	1000 ms	1000 ms	1

rango de grados para su posterior uso, respetando los movimientos generados por cualquier persona.

#### 4. RESULTADOS

En esta sección se discuten algunas características del Sistema, así como percepciones útiles que surgieron en el desarrollo del trabajo.

##### A. General

Las características del sensor Kinect, no permiten capturar todas las articulaciones del cuerpo humano, lo cual podría pensarse que limita la cantidad de posturas que pueden ser detectadas, punto que este trabajo salvaguarda analizando previamente las posturas por detectar. Existen posturas generables por el cuerpo humano que no pueden o no se desean replicar en agentes como avatares o robots.

Puesto que este trabajo tiene como fin la detección de posturas previamente delimitadas, no se requiere un alto grado de exactitud, por lo cual se aprovecha la velocidad de procesamiento en beneficio de la detección de posturas.

##### B. Porcentaje de identificación

La detección de la postura del brazo en un ángulo superior a los 90° en relación al hombro al igual que la detección de la pierna por encima de los 270° en relación a la cadera derecha, permitieron comprobar los valores que contiene la plantilla y seleccionar aquellos que fueron ideales para la activación de un comando.

Para probar los valores se tomaron 10 muestras con distintas combinaciones (tabla 4) y se obtuvieron los resultados enunciados según la tabla 6 para la postura en brazos y según la tabla 5 para piernas.

Los problemas más comunes fueron falsos positivos con tiempos pequeños entre detección de posturas y ejecuciones de las mismas; falsos positivos cuando la permisividad se encuentra con un valor de 1, y una mínima detección con permisividad en 0, lo cual limita la posición de la articulación a un ángulo exacto.

De este modo se recomienda utilizar tiempos de 1000 ms o superiores entre detección de posturas y tiempos entre 100 ms y 1000 ms en la ejecución de las mismas, trabajar con un valor de permisividad al 0.5 para evitar falsos positivos.

**Cuadro 5. Detección en postura de piernas**

Combinación	Detectados	No Detectados	Falsos Positivos
1	2	7	1
2	2	6	2
3	1	9	0
4	7	2	1
5	8	1	1
6	6	3	1
7	2	7	1
8	1	1	8
9	2	1	7

**Cuadro 6. Detección en postura de brazos**

Combinación	Detectados	No Detectados	Falsos Positivos
1	2	6	2
2	1	8	1
3	1	7	2
4	7	2	1
5	8	2	0
6	6	2	2
7	1	6	3
8	1	0	9
9	1	2	7

## 5. CONCLUSIONES

En este trabajo se demostro que es posible realizar la detección de posturas de manera simple utilizando plantillas XML con un formato entendible que permite ubicar articulaciones en tres dimensiones a intervalos de tiempo. La detección de la postura propuesta se puede hacer con los datos generados por cualquier persona sin distinción de morfología, dado que el conjunto de datos que genera la plataforma Kinect no toma en cuenta este factor.

El uso de este sistema permite al usuario común realizar interacciones con agentes de cualquier tipo de una manera simple.

## REFERENCIAS

1. E. Gadalla, I. Abosag, and K. Keeling, Qualitative Market Research: An International Journal **19**, 101 (2016).
2. E. Gadalla, I. Abosag, and K. Keeling, Qualitative Market Research: An International Journal **19**, 101 (2016).
3. C. D. Kidd and C. Breazeal, "Robots at home: Understanding long-term human-robot interaction," in "Intelligent Robots and Systems, 2008. IROS 2008. IEEE/RSJ International Conference on," (IEEE, 2008), pp. 3230–3235.
4. S. A. Setiawan, J. Yamaguchi, S. H. Hyon, and A. Takahashi, "Physical interaction between human and a bipedal humanoid robot-realization of human-follow walking," in "Robotics and Automation, 1999. Proceedings. 1999 IEEE International Conference on," , vol. 1 (IEEE, 1999), vol. 1, pp. 361–367.
5. K. Wada, T. Shibata, T. Saito, K. Sakamoto, and K. Tanie, "Psychological and social effects of one year robot assisted activity on elderly people at a health service facility for the aged," in "Robotics and Automation, 2005. ICRA 2005. Proceedings of the 2005 IEEE International Conference on," (IEEE, 2005), pp. 2785–2790.
6. J. B. Cole, D. B. Grimes, and R. P. Rao, "Learning full-body motions from monocular vision: Dynamic imitation in a humanoid robot," in "Intelligent Robots and Systems, 2007. IROS 2007. IEEE/RSJ International Conference on," (IEEE, 2007), pp. 240–246.
7. S. Cotton, M. Vanoncini, P. Fraise, N. Ramdani, E. Demircan, A. P. Murray, and T. Keller, *Applied Bionics and Biomechanics* **8**, 67 (2011).
8. T. Boellstorff, *Coming of age in Second Life: An anthropologist explores the virtually human* (Princeton University Press, 2015).
9. S. Kagami, K. Nishiwaki, J. J. Kuffner, Y. Kuniyoshi, M. Inaba, and H. Inoue, "Online 3d vision, motion planning and bipedal locomotion control coupling system of humanoid robot: H7," in "Intelligent Robots and Systems, 2002. IEEE/RSJ International Conference on," , vol. 3 (IEEE, 2002), vol. 3, pp. 2557–2562.
10. F. Zuher and R. Romero, "Recognition of human motions for imitation and control of a humanoid robot," in "Robotics Symposium and Latin American Robotics Symposium (SBR-LARS), 2012 Brazilian," (IEEE, 2012), pp. 190–195.
11. T. Cloete and C. Scheffer, "Benchmarking of a full-body inertial motion capture system for clinical gait analysis," in "Engineering in Medicine and Biology Society, 2008. EMBS 2008. 30th Annual International Conference of the IEEE," (IEEE, 2008), pp. 4579–4582.
12. M. Gleicher, *ACM SIGGRAPH Computer Graphics* **33**, 51 (1999).
13. J. Webb and J. Ashley, *Beginning Kinect Programming with the Microsoft Kinect SDK* (Apress, 2012).
14. S. Falahati, *OpenNI cookbook* (Packt Publishing Ltd, 2013).
15. F. Jurado, J. Asbusac et al., *Desarrollo de Videojuegos: Desarrollo de Componentes* (Cursos en Español, 2013).
16. J. Webb and J. Ashley, *Beginning Kinect Programming with the Microsoft Kinect SDK* (Apress, 2012).
17. Z. Zhang, *IEEE multimedia* **19**, 4 (2012).
18. V. Bloom, D. Makris, and V. Argyriou, "G3d: A gaming action dataset and real time action recognition evaluation framework," in "Computer Vision and Pattern Recognition Workshops (CVPRW), 2012 IEEE Computer Society Conference on," (IEEE, 2012), pp. 7–12.
19. J. L. Chicharro and A. F. Vaquero, *Fisiología del ejercicio* (Ed. Médica Panamericana, 2006).