

Control de Movimiento de un Robot Humanoide por medio de Visión de Máquina y Réplica de Movimientos Humanos*

Motion Control of a Humanoid Robot Through Machine Vision and Human Motion Replica

Artículo de Investigación Científica - Fecha de Recepción: 24 de Marzo de 2013 - Fecha de Aceptación: 2 de Septiembre de 2013

Robinson Jiménez Moreno

Ingeniero Electrónico. Magister en Automatización Industrial. Universidad Militar Nueva Granada. Bogotá D.C, Colombia. robinson.jimenez@unimilitar.edu.co.

Fabio Andrés Espinosa Valcárcel

Ingeniero Mecatrónico. Universidad Militar Nueva Granada. Bogotá D.C, Colombia. fabio.espinosa@unimilitar.edu.co

Darío Amaya Hurtado

Ingeniero Electrónico. Doctor en Mecánica. Universidad Militar Nueva Granada. Bogotá D.C, Colombia. dario.amaya@unimilitar.edu.co.

Para citar este artículo / To reference this article:

R. Jiménez, F. Espinosa, and D. Amaya, "Control de movimiento de un robot humanoide por medio de visión de máquina y réplica de movimientos humanos," *INGE CUC*, vol. 9, no. 2, pp. 44–51, 2013.

Resumen: En este artículo se presenta el desarrollo e implementación de un sistema de captura de movimiento antropomórfico mediante técnicas de visión de máquina basado en el dispositivo Kinect, con el fin de realizar el control de movimiento imitativo de un agente robótico Bioloid en el Grupo de Aplicaciones Virtuales (GAV) del Programa de Ingeniería en Mecatrónica de la Universidad Militar Nueva Granada (UMNG). Dados los múltiples grados de libertad de un brazo humano, se busca simplificar una interfaz de control que permita replicar los movimientos de este en un robot humanoide. En este artículo se presentan las técnicas usadas para mejorar el nivel de precisión de los datos entregados por el Kinect y los métodos personalizados de transmisión y codificación de las órdenes enviadas al robot. Los resultados obtenidos derivan en un sistema que cumple con las exigencias básicas de estabilidad, precisión y velocidad de respuesta en la imitación.

Palabras clave: visión de máquina, interacción humano-robot, Kinect, detección de movimiento, imitación.

Abstract: This paper presents the development and implementation of an anthropomorphic motion capture system through machine vision based on the Kinect device in order to achieve the imitative motion control of a Bioloid robotic agent of the GAV research group from the Mechatronics Engineering program at UMNG. We present the techniques used to improve data precision delivered by Kinect, as well as custom methods of transmission and coding of commands sent to the robot. Results derive in a system that meets the basic requirements of stability, accuracy and speed of imitation response.

Keywords: Machine vision, human-robot interaction, Kinect, motion detection, imitation.

* Este artículo fue desarrollado en el marco del proyecto de investigación "Plataforma de teleoperación para manipulación de materiales explosivos. FASE III", avalado y financiado por la Vicerrectoría de Investigaciones de la Universidad Militar Nueva Granada; se desarrolló del 15 de julio de 2012 al 15 de julio del 2013, con código interno de aprobación ING 1069 y asociado al Grupo de Investigación en Aplicaciones Virtuales (GAV) del Programa de Ingeniería en Mecatrónica de la Universidad Militar Nueva Granada.

I. INTRODUCCIÓN

Las interacciones humano-robot han sido motivo de estudio a lo largo de los últimos veinte años, especialmente desde que se visualizó el potencial de la inclusión de los robots en actividades de la vida cotidiana de los seres humanos. Estas interacciones siempre se vieron limitadas debido a una falta de naturalidad en la estructura y, en especial, en los movimientos de los robots, en parte a causa del material de construcción y la electrónica que involucraban; de igual forma, estos estaban sujetos a procesos de programación complejos y largos, todo lo cual limitaba el rango de posibilidades en las posiciones que podían lograr.

Además la interacción directa entre humano y robot se había realizado hasta ahora por medio de controles que no eran lo suficientemente intuitivos (*joysticks*, mandos, pulsadores) para generar una correspondencia entre los movimientos humanos y los del robot [1].

El interés en naturalizar los movimientos de los robots se debe en gran parte a la necesidad que se tiene en diferentes campos para que estos puedan realizar movimientos complejos; pero además para que en campos como la robótica de servicio o la enseñanza estos sean aceptados más rápidamente por los usuarios [2].

Por lo anterior, recientemente han aparecido tendencias de desarrollo para lograr naturalizar los movimientos de los robots a través de la imitación directa de los movimientos humanos, en una clara relación al método de aprendizaje humano, el cual consiste principalmente, al menos en las primeras etapas de desarrollo infantil, en la imitación. Estos intentos han hecho uso de *hardware* dedicado y personalizado, generalmente complejo y costoso, entre los cuales se encuentra el uso de acelerómetros, giroscopios y sistemas de rastreo infrarrojo. Entre estos se encuentran investigaciones como [3], en la que se hace uso de un sistema de rastreo con marcadores infrarrojos para la imitación del movimiento del brazo humano; también en [4] se hace uso de un complejo sistema de navegación inercial para lograr replicar los movimientos de un usuario por parte del robot humanoide NAO.

Los sistemas de visión de máquina ofrecen una alternativa para la réplica de movimientos humanos en robots; gracias al desarrollo de *hardware* relativamente económico y con buen nivel de precisión como lo es el Kinect de Microsoft [5], el número de investigaciones por medio de estos sistemas, que podrían contribuir a este campo, ha aumentado considerablemente. Algunas de las que más relevancia tienen en el desarrollo de este proyecto, tanto por su similitud como por su contribución indirecta, son: [6], en la que se plantea el uso del Kinect como dispositivo de control de un esqueleto

robótico de 8 articulaciones; [7], en la que se hace un estudio sobre la factibilidad del control de sistemas robóticos usando el Kinect y que dio como conclusión la factibilidad de esta propuesta pero con ciertas limitaciones; [8], en la que se usa el Kinect para determinar actividades no estructuradas, de tal forma que se puedan incrementar las capacidades de los futuros robots de servicio; [9] y [10], en las que se profundiza en la detección de movimientos y gestos humanos usando Kinect y algoritmos de aprendizaje y predicción, los cuales pueden ser usados en entornos de enseñanza robot-humano. Por último se hace notar que la detección de movimientos para el control de robots no cae únicamente dentro del campo de la robótica humanoide, ya que en [11] se hace uso del Kinect para controlar un robot móvil terrestre.

Como trabajo previo a esta investigación, en el Grupo de Aplicaciones Virtuales (GAV) del Programa de Ingeniería en Mecatrónica de la Universidad Militar Nueva Granada se ha venido trabajando con el Kinect en desarrollos de robótica y automatización [12], explorando sus ventajas en el procesamiento de información de profundidad y técnicas de visión de máquina. Es así como ante la necesidad de operar en tiempo real un grupo de brazos robóticos de corte industrial, de seis grados de libertad y de referencia Mitsubishi RV6SQL, se planteó el diseño de una interface de fácil operación que permitiese a un usuario manipular ágilmente ambos brazos por réplica del movimiento humano, tomando como algoritmo base las funciones del Kinect operando sobre un robot humanoide. Como aporte, esta investigación desarrolla la manipulación en tiempo real de ambos brazos, ya que trabajos relacionados han hecho avances en este tema con solo un brazo.

Este artículo consta de 5 secciones: en la sección 2 se presenta una explicación del desarrollo del sistema de adquisición de movimientos por medio del uso de Kinect y los métodos usados para mejorar los datos obtenidos; en la sección 3 se explica cómo se usan los datos obtenidos anteriormente para controlar el agente robótico humanoide Bio-loid y las técnicas usadas para lograr una operación satisfactoria; en la sección 4 se presentan los resultados obtenidos de la implementación del sistema, y finalmente, las conclusiones se presentan en la sección 5, junto con posibles trabajos futuros.

II. ADQUISICIÓN DE MOVIMIENTO

Como se explicó anteriormente, se utilizó el dispositivo de sensado de movimiento Kinect, el cual consta de una cámara de captura de colores y de un conjunto emisor-cámara infrarroja para el mapeo de la profundidad del ambiente. Gracias a este conjunto infrarrojo que pueden realizar tareas de detección de movimiento de estructuras esqueléticas

CONTROL DE MOVIMIENTO DE UN ROBOT HUMANOIDE POR MEDIO DE VISIÓN DE MÁQUINA Y RÉPLICA DE MOVIMIENTOS HUMANOS

cas como la humana [12]. Como tal, esta capacidad no es inherente al Kinect y depende de algoritmos externos de detección de patrones. Dentro de la funcionalidad provista por estos algoritmos se encuentra la posibilidad de acceder directamente a la información acerca de parámetros esenciales del cuerpo humano, como por ejemplo, ángulos y medidas de las extremidades, así como también la posibilidad de obtener indirectamente medidas de velocidad o aceleración de una articulación o extremidad; esto da lugar a funcionalidades más avanzadas, como por ejemplo, la detección de poses y gestos, los cuales amplían las capacidades de operación de un robot con múltiples grados de libertad.

Como se puede ver en la figura 1, los sistemas de detección del Kinect son capaces de detectar hasta 20 articulaciones corporales; además de ser capaces de predecir el movimiento de estas y de tener robustez frente a la oclusión de una o más de las articulaciones detectadas.

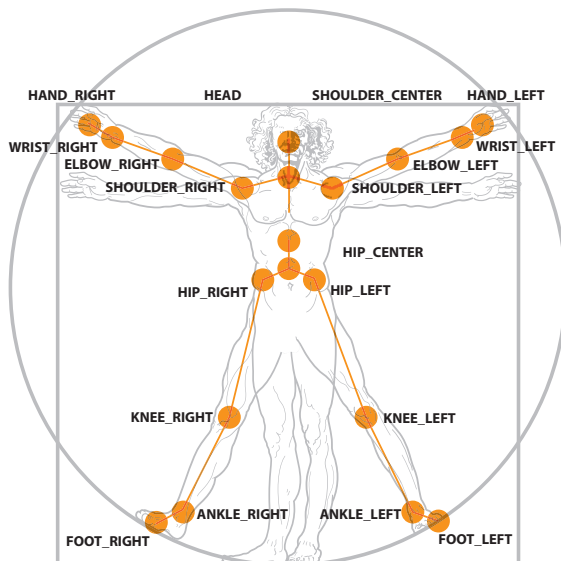


Figura 1. Articulaciones rastreadas por Kinect

Fuente: [www. http://msdn.microsoft.com/en-us/library/hh438998.aspx](http://msdn.microsoft.com/en-us/library/hh438998.aspx).

Como parte del desarrollo del proyecto se creó una aplicación capaz de medir diferentes parámetros del cuerpo, como lo son el ángulo de los brazos respecto a un plano horizontal, respecto a un plano vertical o respecto a una articulación previa (figura 2). Vale la pena aclarar que estas extremidades se escogieron solamente como demostración, ya que prácticamente todas son susceptibles de ser medidas; en la figura 3 se muestra un ejemplo de esta aplicación en la que se puede observar las articulaciones que se pueden rastrear y, además, los diferentes planos sobre los cuales se puede realizar este rastreo.

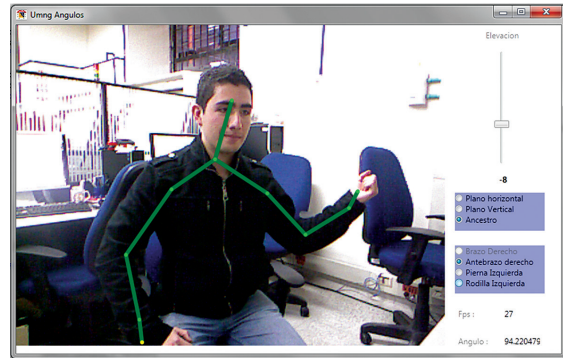


Figura 2. Rastreo de articulaciones respecto a articulaciones previas
Fuente: Elaboración propia.

Sobre esta aplicación se realizaron pruebas sobre los datos provistos por el rastreo esquelético, con el fin de minimizar el error de posición estimada de las articulaciones (fluctuación) (figura 4) a través de técnicas de suavizado basadas en filtros ARMA. Este suavizado está basado en ciertos parámetros que se mencionan a continuación:

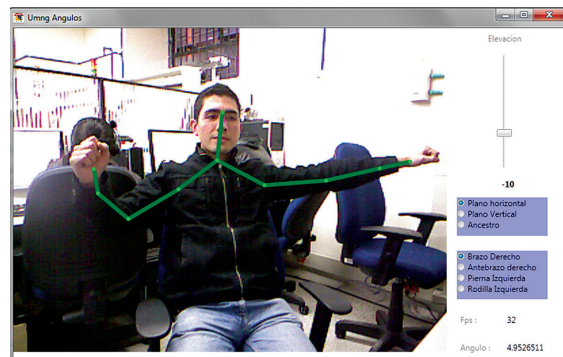


Figura 3. Rastreo de articulaciones respecto a un plano horizontal
Fuente: Elaboración propia.

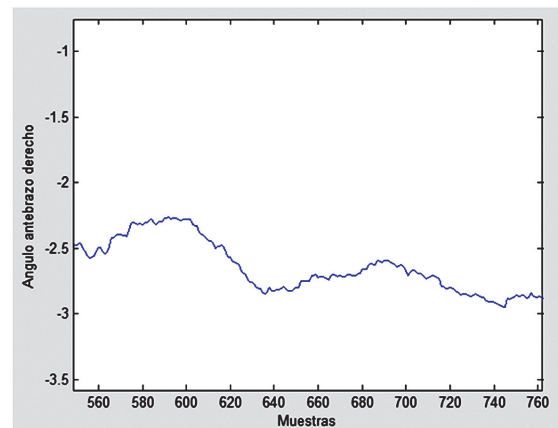


Figura 4. Fluctuaciones en los datos obtenidos
Fuente: Elaboración propia.

- **Suavizado:** indica el nivel de suavizado que se va a aplicar sobre los datos, de tal forma que no presente fluctuación (*jitter*). Incrementar el nivel de suavizado trae como consecuencia negativa un incremento en el tiempo de establecimiento entre los datos y la posición real de la extremidad o articulación, tal como se puede ver en la figura 5.

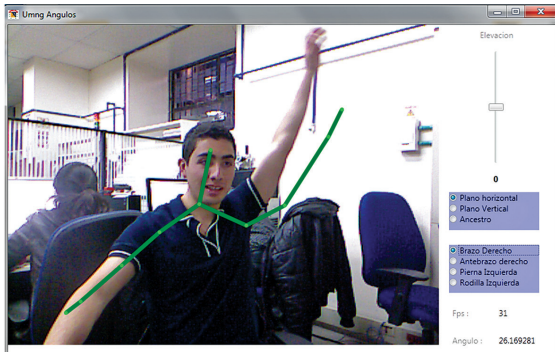


Figura 5. Desfase entre la articulación real y la rastreada
Fuente: Elaboración propia.

- **Corrección:** indica la velocidad a la cual se van a corregir los datos que presenten fluctuaciones o errores respecto a la posición real.
- **Predicción:** indica el nivel de predicción (número de intervalos que se desean predecir) que se va a usar. Valores altos podrían resultar en la aparición de sobreimpulsos (sobrepaso de la posición deseada) en la respuesta, como se puede ver en la figura 6.

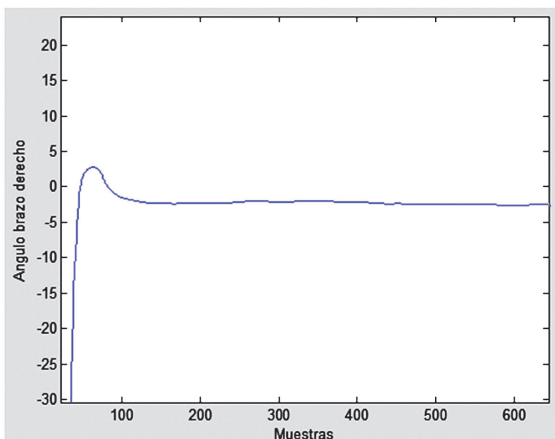


Figura 6. Sobreimpulso generado por el uso de datos predichos
Fuente: Elaboración propia.

- **Radio de fluctuación (*jitter radius*):** es el tamaño máximo de fluctuación que una articulación o extremidad puede tener.

- **Radio máximo de desviación:** indica el nivel de desviación máximo que los datos suavizados pueden tener sobre los datos sin procesar.

De estas pruebas se determinó que se debe hacer una concesión entre velocidad y precisión, por lo cual se obtuvieron unos rangos de valores de los parámetros que permiten establecer diferentes formas de operación según el tipo de aplicación que se desee construir.

III. CONTROL DE MOVIMIENTO

Una vez se implementó el sistema para la adquisición de los parámetros relevantes (ángulos) del movimiento de una persona, se propuso la implementación de tres formas de control del movimiento del agente robótico, las cuales se describen a continuación:

- Transmisión directa del ángulo actual

Usando esta metodología, una se vez se obtenía el ángulo actual de una articulación deseada, este se transmitía directamente al robot junto con un valor estimado de velocidad. Esta metodología presentó un gran problema debido a la sensibilidad tanto del Kinect (figura 7) como de los servomotores del robot, ya que ante cambios mínimos en la extremidad sensada, el robot presentaba oscilaciones.

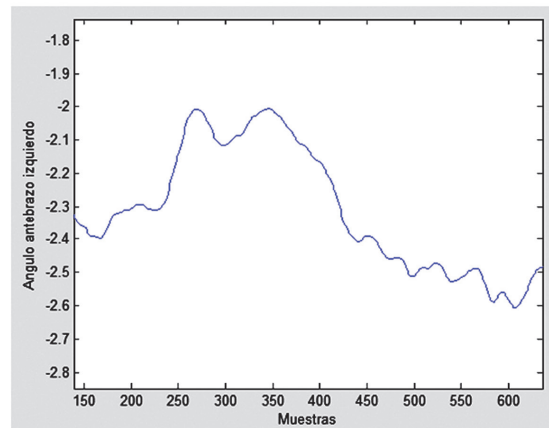


Figura 7. Fluctuaciones en los datos obtenidos
Fuente: Elaboración propia.

- Transmisión de órdenes en forma de impulso

Para evitar estas oscilaciones se implementó dentro de la aplicación un algoritmo de regresión lineal (ecuación 1) con el cual se obtenían tendencias de movimiento, las cuales eran interpretadas como comandos de acción, los cuales eran transmitidos al robot mientras, se mantuviera una tendencia.

Estas tendencias se pueden visualizar gráficamente en la figura 8. Aunque de esta forma se logró mejorar la estabilidad en los movimientos del robot, se perdió precisión de posición de los mismos.

$$\hat{B} = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^n (x_i - \bar{x})^2} \quad (1)$$

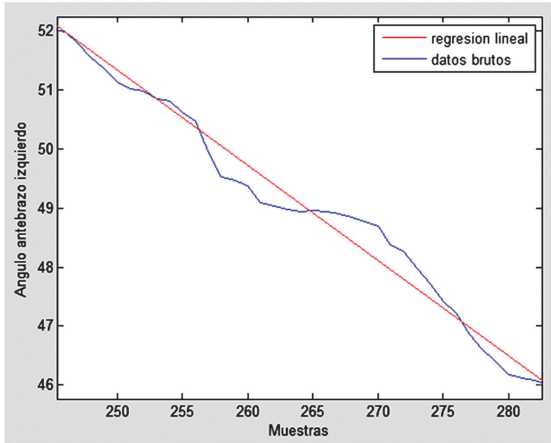


Figura 8. Regresión lineal de datos en bruto
Fuente: Elaboración propia.

- Transmisión del ángulo una vez se detecta una tendencia

Debido a que los métodos anteriores presentaban deficiencias tanto en precisión como en estabilidad, se implementó uno que hacía uso de las buenas características de ambos; este método consiste en determinar la tendencia de movimiento de un conjunto de datos mediante la ecuación 1, para de esta forma enviar el valor del ángulo actual solo cuando se presente una tendencia determinada por encima de un umbral predefinido.

El ángulo es luego transmitido desde la aplicación hasta el robot por medio de comunicación serial, usando una codificación personalizada en forma de cadena de caracteres, de tal forma que junto con el valor del ángulo también se encuentra el ID del servomotor que se desea mover. Aunque la transmisión ocurre en forma secuencial con la detección de movimiento para lograr un sistema más eficiente, las transmisiones individuales para cada servo ocurren en paralelo; esto quiere decir que el movimiento en el robot no es secuencial. Una vez los mensajes codificados llegan hasta el robot, este ejecuta una rutina de lectura e interpretación del mensaje, la cual da lugar al movimiento deseado, junto con una velocidad predefinida.

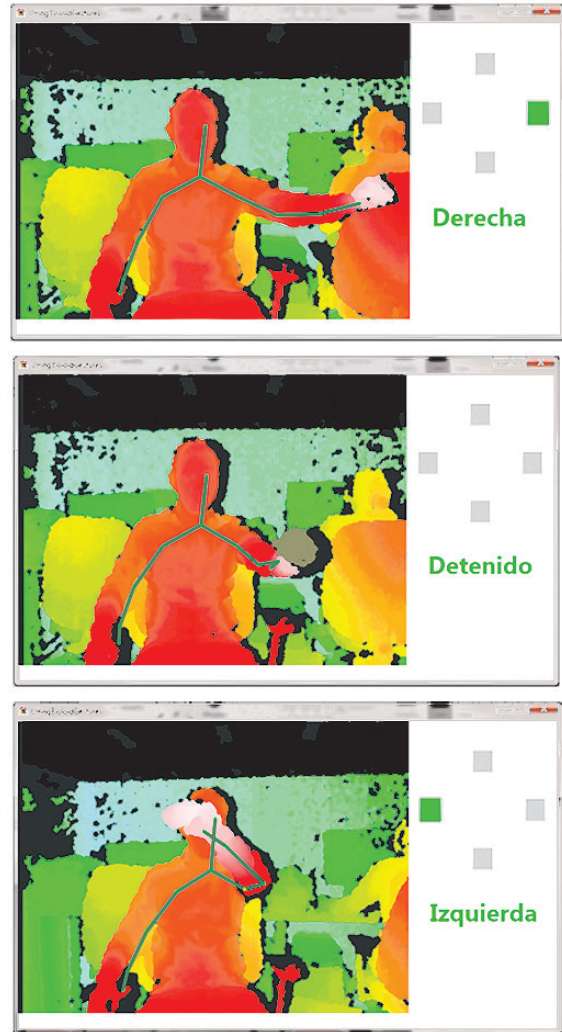


Figura 9. Interpretación visual de tendencias de movimientos
Fuente: Elaboración propia.

En la figura 9 se aprecia la percepción del movimiento obtenido respecto a una articulación, el cual puede ser determinado como izquierda, derecha, arriba o abajo.

IV. RESULTADOS

Este sistema se implementó mediante un computador con un procesador de 2 núcleos a 3.1 GHz y 4 GB de memoria RAM; se usó la distribución de Kinect para Xbox® 360 y como agente robótico un humanoide Bioloid de la compañía Robotis© de 18 grados de libertad, de los cuales por ahora solo se usaron 6, correspondientes a las extremidades superiores. Además, como plataforma de desarrollo para la adquisición de movimientos se usó la distribución Express de visual C#, junto con el



Figura 10. Esquema general del sistema
Fuente: Elaboración propia.

kit de desarrollo de software provisto por Microsoft® para el Kinect en su versión 1.5, y para la programación del robot se usó AVR Studio para la programación del microcontrolador Atmel y el software Roboplus de Robotis para la carga del programa en el microcontrolador y para la verificación de los movimientos. En la figura 10 se puede apreciar una estructura general del flujo de información en el sistema creado.

Respecto a la implementación del sistema de adquisición de movimientos, básicamente se presentaron inconvenientes en tres parámetros importantes: Fluctuaciones, Retardo y Precisión. Para solucionar los problemas de fluctuaciones pequeñas en los datos se usó los filtros ARMA proporcionados dentro del SDK provisto por Microsoft para el manejo del Kinect [13], [14]. Gracias a estos filtros se logró una reducción total de las fluctuaciones (figura 11), pero a costa de la aparición de un retardo temporal, lo cual causó un amortiguamiento en la salida del filtro (figura 12).

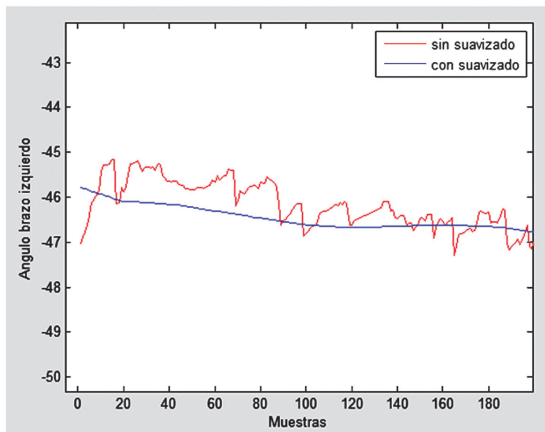


Figura 11. Suavizado de los datos con fluctuaciones
Fuente: Elaboración propia.

Cabe anotar que los datos brutos entregados por el sensor ya presentaban de por sí un pequeño desfase temporal respecto a los movimientos reales de una articulación, por lo tanto se hizo uso de las funciones de predicción, también presentes dentro del SDK mencionado anteriormente [13], [14], [15]. Gracias a esto se logró reducir este retardo temporal, que en ocasiones alcanzaba valores de hasta 50 (figura 12) muestras, y dado que la frecuencia de muestreo estaba establecida en 30 Hz, estas equivaldrían a 1.66 segundos de retardo. Sin embargo, el uso de predicciones para los datos también generó un efecto indeseado como lo es el sobreimpulso (figura 12), el cual alcanzaba un máximo de 17.5 %, lo cual no es deseable.

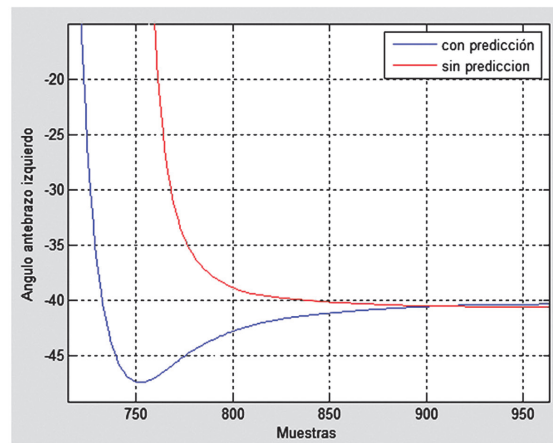


Figura 12. Retardo de los datos en bruto y sobreimpulso del sistema usando datos predichos
Fuente: Elaboración propia.

Para solucionar los inconvenientes presentados por usar los métodos de suavizado y predicción sobre los datos, se ejecutó un proceso iterativo de selección de los mejores valores para estos dos parámetros,

teniendo en cuenta que es preferible que exista un pequeño retardo pero no un sobreimpulso; de esta forma se obtuvo un conjunto de valores que permitieron obtener un sistema de detección de movimiento con error estacionario de posición igual a cero, críticamente amortiguado, con un retardo de máximo 20 muestras o 0.66 segundos y con un sobreimpulso extremadamente pequeño (figura 13).

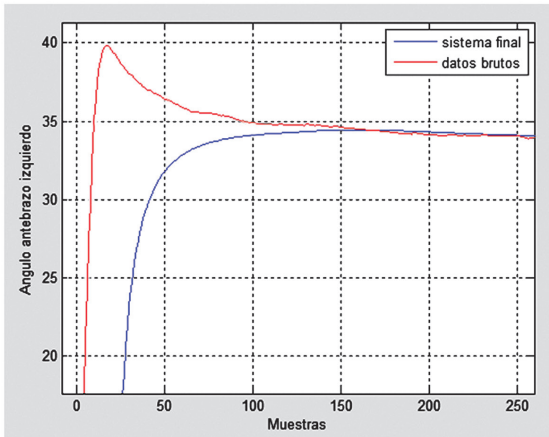


Figura 13. Respuesta del sistema obtenido versus datos en bruto
Fuente: Elaboración propia.

El sistema creado anteriormente, junto con la técnica de transmisión de valores de ángulo basada en tendencias (figura 14), permitieron una operación adecuada de las articulaciones seleccionadas del agente robótico (figuras 15 y 16). Estas articulaciones corresponden (aunque no completamente) al hombro (dos grados de libertad) y antebrazo humano (un grado de libertad).

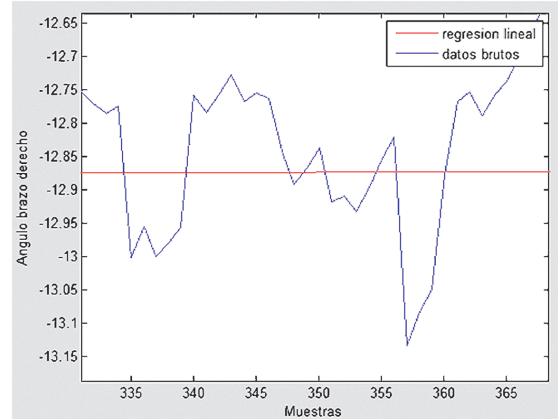


Figura 14. Regresión lineal de datos en bruto
Fuente: Elaboración propia.

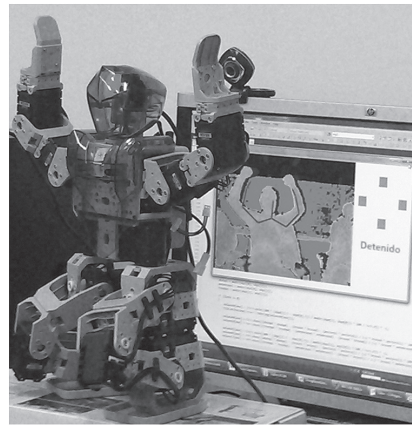


Figura 15. Ejemplos de funcionamiento del sistema
Fuente: Elaboración propia.

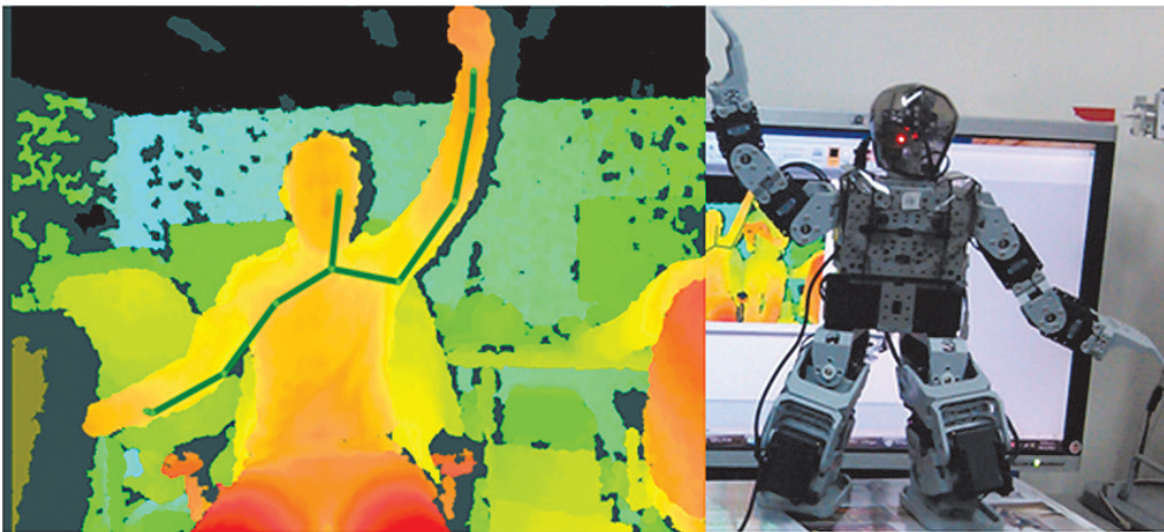


Figura 16. Ejemplos de funcionamiento del sistema
Fuente: Elaboración propia.

V. CONCLUSIONES

Fue necesario hacer una concesión entre las diferentes características del sistema de adquisición (precisión, retardos, sobreimpulso y velocidad de respuesta) con el fin de obtener una respuesta que fuera útil para las necesidades prácticas que se tenían, por lo cual se dio mayor prioridad a la precisión.

Se logró construir un sistema que presenta un buen tiempo de respuesta; de igual forma, presenta la capacidad de responder ante cambios bruscos en las posiciones de las articulaciones y en el cual no estaban presentes oscilaciones indeseables, lo cual genera una operación agradable e interactiva del sistema.

Se evidenció la necesidad de crear métodos de transmisión y control personalizados con el fin de lograr un mayor nivel de control sobre el agente robótico, dado que la naturaleza del sensor de movimiento generaba comportamientos indeseados ante métodos tradicionales de operación.

Se concluye que mediante un sensor como el Kinect es posible realizar el telecontrol de dispositivos robóticos de tipo antropomórfico de forma eficiente y en tiempo real, lo cual facilita y simplifica así la interfaz de operación.

Como trabajo futuro se sugiere el incremento del número de articulaciones por controlar, el afinamiento de los parámetros de control del sistema (suavizado y predicción) y su aplicación en el telecomandado de dos brazos robóticos industriales de forma natural, mediante el seguimiento de los movimientos del operador.

REFERENCIAS

- [1] C.I. Penaloza, Y. Mae, K. Ohara, and T. Arai, "Social human behavior modeling for robot imitation learning", *Mechatronics and Automation (ICMA)*, 2012 International Conference on , vol., no. , pp. 457-462, Aug. 5 8 2012.
- [2] S. Hak, N. Mansard, O. Ramos, L. Saab, and O. Stasse, "Capture, recognition and imitation of anthropomorphic motion", *Robotics and Automation (ICRA)*, 2012 IEEE International Conference on , vol., no., pp. 3539-3540, May. 14 18 2012.
- [3] Hsien-I Lin, Yu-Cheng Liu, and Chi-Li Chen, "Evaluation of human-robot arm movement imitation", *Control Conference (ASCC)*, 2011 8th Asian , vol., no., pp. 287-292, May. 15 18 2011.
- [4] J. Koenemann and M. Bennewitz, "Whole-body imitation of human motions with a Nao humanoid", *Human-Robot Interaction (HRI)*, 2012 7th ACM/IEEE International Conference on , vol., no., pp. 425, March. 5 8 2012.
- [5] J. Smisek, M. Jancosek, and T. Pajdla, "3D with Kinect", *Computer Vision Workshops (ICCV Workshops)*, 2011 IEEE International Conference on , vol., no., pp.1154-1160, Nov. 6 13 2011.
- [6] J. Ekelmann and B. Butka, "Kinect controlled electromechanical skeleton", *Southeastcon*, 2012 *Proceedings of IEEE* , vol., no., pp.1-5, March. 5 18 2012.
- [7] R. A. El-laithy, J. Huang, and M. Yeh, "Study on the use of Microsoft Kinect for robotics applications", *Position Location and Navigation Symposium (PLANS)*, 2012 IEEE/ION , vol., no., pp.1280-1288, April 23 26 2012.
- [8] J. Sung, C. Ponce, B. Selman, and A. Saxena, "Unstructured human activity detection from RGBD images", *Robotics and Automation (ICRA)*, 2012 IEEE International Conference on , vol., no., pp. 842-849, May. 14 18 2012.
- [9] A. Bigdelou, T. Benz, L. Schwarz, and N. Navab, "Simultaneous categorical and spatio-temporal 3D gestures using Kinect" *3D User Interfaces (3DUI)*, 2012 IEEE Symposium on , vol., no., pp.53-60, March. 4 5 2012.
- [10] I. I. Itauma, H. Kivrak, and H. Kose, "Gesture imitation using machine learning techniques", *Signal Processing and Communications Applications Conference (SIU)*, 2012 20th , vol., no., pp.1-4, April 18 20 2012.
- [11] E. Machida, Meifen Cao, T. Murao, and H. Hashimoto, "Human motion tracking of mobile robot with Kinect 3D sensor", *SICE Annual Conference (SICE)*, 2012 *Proceedings of* , vol., no., pp. 2207-2211, Aug. 20 23 2012.
- [12] R. Jimenez, O. Aviles, and F. Espinosa, "Level measurement comparison between 3D vision system based on Kinect and ultrasonic industrial sensor", in *Pakistan Asian Transactions On Engineering*, vol. 2, fasc.5, pp.10 - 19, 2012.
- [13] M. A. Livingston, J. Sebastian, Z. Ai, and J. W. Decker, "Performance measurements for the Microsoft Kinect skeleton", *Virtual Reality Short Papers and Posters (VRW)*, 2012 IEEE , vol., no., pp.119-120, March. 4 8 2012.
- [14] Skeletal Joint Smoothing White Paper. Available: <http://msdn.microsoft.com/en-us/library/jj131429.aspx#ID4EWPAAE>. [fecha de consulta: noviembre de 2012].
- [15] Joint Filtering. Available: <http://msdn.microsoft.com/en-us/library/jj131024.aspx>. [fecha de consulta: noviembre de 2012]