

MICROSOFT KINECT

Diana Beltran Guerrero, Luis Basañez Villaluenga

Universidad Politécnica de Catalunya, Av Diagonal, 647, Planta 11, 08028, Barcelona, España
Instituto de Organización y Control de Sistemas Industriales

{diana.beltran, luis.basanes}@upc.edu

Palabras clave: Microsoft Kinect, nube de puntos 3D, profundidad, disparidad.

ÍNDICE

1. INTRODUCCIÓN.....	2
2. MODO DE FUNCIONAMIENTO.....	3
2.1 Calibración.....	6
2.2 Integración de la profundidad con el color.....	7
2.3 Precisión y resolución de la profundidad.....	7
3. UTILIZACIÓN DEL SENSOR KINECT CON UN PC.....	7
3.1 SDK PARA WINDOWS 7.....	8
3.2 OPENNI SDK.....	8
4. EJEMPLOS DE LA INFORMACIÓN OBTENIDA DEL SENSOR KINECT UTILIZANDO ROS.....	9
5. ALGUNOS TRABAJOS REALIZADOS CON EL SESOR KINECT.....	11
5.1 “Visual Odometry and Mapping for Autonomous Flight Using an RGB-D Camera”... ..	11
5.2 “RGBD-HuDaAct: A color-Depth Video Database for Human Daily Activity Recognition”.....	11
5.3 “Unstructured Human Activity Detection from RGB Images”.....	11
5.4 “Real-time 3D visual SLAM with a hand-held RGB-D camera”.....	11
5.5 “RGB-D mapping: Using Kinect-Style depth cameras for dense 3D modeling of indoor environments”.....	12
5.6 “Towards a benchmark for RGB-D SLAM evaluation”.....	12
5.7 “Human Activity Detection from RGBD Images”.....	12
5.8 “Real-Time Plane Segmentation Using RGB-D Cameras”.....	12
5.9 “An evaluation of the RGB-D SLAM system” [23]	13
5.10 “Unsupervised Feature Learning for RGB-D Based Object Recognition”.....	13
6. CONCLUSIONES.....	13
7. BIBLIOGRAFIA.....	14

1 INTRODUCCIÓN

El mundo de los videojuegos ha ido avanzando día a día en el intento de que el usuario sienta cada vez con mayor realismo que está dentro del juego. El 4 de noviembre de 2010 fue presentado en Norteamérica uno de los controladores de juego que más ha llamado la atención en los últimos tiempos: el ya muy conocido Microsoft Kinect, un dispositivo diseñado para permitir al usuario controlar e interactuar con la consola sin necesidad de tener contacto físico con un control tradicional de videojuego. Kinect usa un interfaz de usuario que reconoce gestos, voz, objetos e imágenes y que, por tanto, permite controlar con el propio cuerpo el personaje que aparece en la pantalla.

El sensor Kinect conocido inicialmente como “Project Natal”[1] fue vendido inicialmente por Microsoft para controlar e interactuar con el Xbox 360 sin la necesidad de un control físico, y fue anunciado por primera vez el 1 de junio de 2009 durante “the Electronic Entertainment Expo (E3) 2009”. El 13 de junio de 2010 durante la E3 2010 Microsoft anunció que el nombre oficial del dispositivo sería “Kinect” y que la fecha de lanzamiento en Norteamérica sería el 4 de noviembre de 2010. El lanzamiento en Europa fue pospuesto una semana. Habiendo vendido 8 millones de unidades en sus primeros 60 días, Kinect ha conseguido el record mundial Guinness por ser el dispositivo electrónico más rápidamente vendido [2].

El dispositivo cuenta con un proyector laser, una cámara infrarroja, una cámara RGB, micrófonos y un procesador personalizado que ejecuta el software, diseñado por la compañía israelí “Prime Sense”, que reconoce rostros, voz y que proporciona captura de movimiento de todo el cuerpo en 3D. En la figura 1 se puede observar los elementos físicos del sensor Kinect.

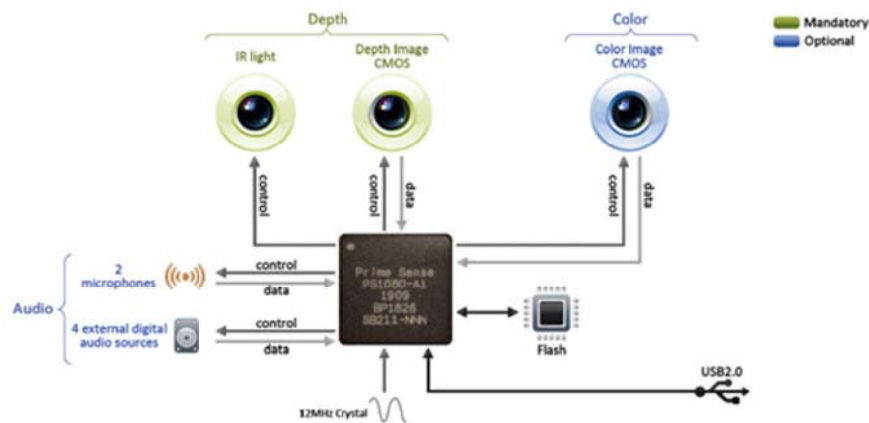


Fig. 1. Elementos internos del sensor Kinect [3]

Poco después de su lanzamiento, los investigadores de diferentes ámbitos se dieron cuenta de que este dispositivo no solo servía como un control de videojuego, sino que

tenía un gran potencial por las características de los datos capturados que podrían ser utilizados en diferentes aplicaciones incluyendo la correspondencia y el modelado 3D. Fue entonces cuando el español Héctor Martín, motivado por una recompensa que se ofrecía por el diseño de un controlador de código abierto para la Kinect, usó métodos de ingeniería inversa para el desarrollo de un controlador para GNU/Linux que permite el uso de la cámara RGB y las funciones de profundidad.

Investigaciones sobre calibración y geometría de sensores similares al Kinect, tales como SwissRanger[4] y PMD[5], han sido el tema de muchos trabajos [6-11], sin embargo estos sensores están basados en el principio de tiempo-de-vuelo para obtener sus datos de profundidad lo cual es fundamentalmente diferente del principio del sensor Kinect que se basa en la triangulación.

A partir de entonces y en muy poco tiempo se han desarrollado diferentes aplicaciones para el uso del sensor Kinect. Entre muchas otras se encuentran:

- Una aplicación para Windows KinEmote que permite controlar de forma gestual una aplicación de media center como Boxee.
- Unos investigadores del Instituto Tecnológico de Massachusetts (MIT), han conseguido un control de ordenadores mediante gestos con las manos más discretos y menos teatralizados que en otras soluciones.
- Unos estudiantes de la Escuela Politécnica de Lausana (Suiza) han desarrollado una aplicación que se basa en la utilización de dos dispositivos Kinect y que es capaz de detectar en tiempo real las posiciones de dos usuarios.
- Un grupo de científicos del Instituto de medicina forense de la Universidad de Berna ha desarrollado una aplicación para Kinect que permite interactuar con radiografías en 3D sin necesidad de tocar una pantalla o un teclado.
- Se ha desarrollado también una aplicación para utilizar Kinect como radar y sensor 3D para detectar objetos en un robot que vuela o bien utilizar una iPad para el control de los diversos parámetros de Kinect.

2 MODO DE FUNCIONAMIENTO

Como se ha mencionado anteriormente, el sensor Kinect posee un proyector laser, una cámara infrarroja con una resolución de 320x240 pixeles que funciona a 30 cuadros por segundo (fps) y una cámara RGB con una resolución de 640x480 pixeles que funciona a 30 fps. En la figura 2 se puede apreciar la ubicación del proyector y las cámaras en el sensor. El dispositivo también está equipado con un conjunto de micrófonos que permite recibir y reconocer comandos de voz. Posee un ángulo de visión de 57° horizontalmente y 43° verticalmente, mientras que con la base móvil se puede desplazar tanto para arriba como para abajo 27° permitiendo captar mayor espacio de su entorno. Algo importante a destacar del sensor Kinect son los requerimientos de alimentación. La Kinect requiere 12 voltios de corriente continua para operar y 12 vatios de potencia durante su funcionamiento. Estos requerimientos son superiores a los 5 voltios y 2.5 vatios del bus serial universal (USB) que incorpora la mayoría de

los computadores, lo que significa que la Kinect requiere un alimentador adicional como el que se muestra en la figura 3.

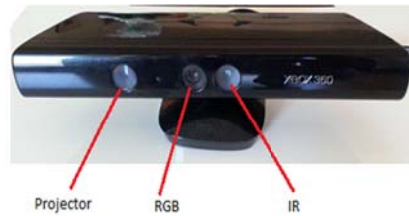


Fig. 2. Sensor Kinect [12].

Con todo esto el sensor captura profundidad e imágenes a color simultáneamente, la integración de los datos de profundidad y el color da como resultado una nube de puntos 3D en color que contiene alrededor de 300.000 puntos en cada cuadro. Registrando imágenes de profundidad consecutivas se puede obtener un incremento en la densidad de los puntos, además de poder crear una nube de puntos completa de entornos interiores en tiempo real.



Fig. 3. Dispositivo de alimentación para el sensor Kinect [3].

El láser del sensor Kinect proyecta un patrón como el que se muestra en la figura 4, que es utilizado para obtener un patrón de referencia que es almacenado dentro del sensor previamente. Este patrón de referencia es obtenido capturando un plano a una distancia conocida. Para una nueva imagen capturada de la escena a estudiar, la cámara infrarroja detenta un nuevo patrón, el cual se compara con el patrón de referencia por medio de una ventana de correlación. De esta manera se obtiene la disparidad.

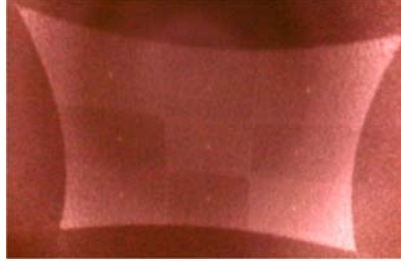


Fig. 4. Patrón proyectado por el láser [12].

En la figura 4 se tiene una representación del patrón de referencia almacenado en la memoria del sensor Kinect con respecto al nuevo patrón proyectado por el láser sobre un objeto en la escena.

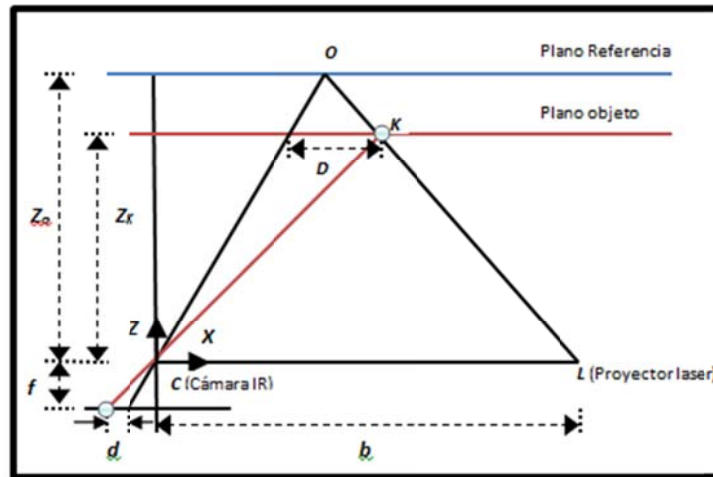


Fig. 5. Relación entre la profundidad relativa y la disparidad medida [13]

Como se aprecia en la figura 5, cuando la cámara IR detecta un nuevo objeto k se produce una disparidad d entre el patrón proyectado en dicho objeto y el almacenado en memoria. Por la separación que hay entre el proyector láser y la cámara IR (línea base b) y la profundidad del objeto (Z_k), se puede realizar una triangulación. Utilizando semejanza de triángulos a partir de la figura 5 se obtiene:

$$\frac{D}{b} = \frac{Z_0 - Z_k}{Z_0} \quad (1) \quad \text{y} \quad \frac{d}{f} = \frac{D}{Z_k} \quad (2)$$

Dónde:

D : Desplazamiento del punto k en el espacio del objeto.

b : Distancia entre el proyector y la cámara infrarroja conocida como longitud de la base.

Z_0 : Distancia desde el sensor al plano de referencia.
 Z_k : Profundidad del punto k en el espacio del objeto.
 d : Disparidad observada en el espacio de la imagen.
 f : Distancia focal de la cámara infrarroja.

Despejando D de la ecuación (2) y reemplazándola en la (1) obtenemos:

$$Z_k = \frac{Z_0}{1 + \frac{Z_0}{fb}d} \quad (3)$$

Con la ecuación (3) se puede obtener la profundidad de los objetos presentes en la escena. Los valores de Z_0 , f and b se determinan en la etapa de calibración del sensor Kinect.

2.1 Calibración

Los parámetros de calibración involucrados en el modelo matemático para la obtención de la coordenada 3D del sensor Kinect incluyen [13]:

- Distancia focal f .
- Punto principal inicial (x_0, y_0) .
- Coeficientes de distorsión δ_x, δ_y .
- Longitud de la base b .
- Distancia del patrón de referencia Z_0 .

Tres de los parámetros indicados se pueden obtener por medio de la calibración estándar de la cámara IR. El cálculo de la longitud de la base y la distancia del patrón de referencia no es sencillo por las siguientes razones. En la práctica no es posible transmitir la medida actual de disparidad, probablemente por limitaciones de ancho de banda. En lugar de esto, los valores de disparidad se normalizan entre 0 y 2047 y se transmiten como enteros de 11 bits. Por lo tanto, en la ecuación (3) d se debe reemplazar con $md'+n$ donde d' es la disparidad normalizada y m, n los parámetros de una normalización lineal, quedando la ecuación así [13]:

$$Z_k^{-1} = \left(\frac{m}{fb}\right) d' + \left(Z_0^{-1} + \frac{n}{fb}\right) \quad (4)$$

La anterior ecuación expresa la relación lineal entre la inversa de la profundidad de un punto con su correspondiente disparidad normalizada. Cuanto mayor sea la disparidad menor será la profundidad del objeto y viceversa.

2.2 Integración de la profundidad con el color

El sensor Kinect captura profundidad e imágenes a color simultáneamente. La integración de los datos de profundidad y los datos de color da como resultado una nube de puntos 3D en color que contiene alrededor de 300,000 puntos en cada cuadro. La integración de estos datos requiere una orientación de la cámara RGB en relación con el sistema de coordenadas de la profundidad [13].

En la práctica las imágenes de la cámara RGB son transmitidas en un tamaño reducido; por ello, es más conveniente realizar una calibración estéreo con las imágenes reducidas en vez de con las imágenes completas. Los parámetros resultantes describen la relación entre las coordenadas 3D de cada punto y su correspondiente pixel-coordenada en la imagen RGB reducida. Una vez estos parámetros se han estimado, se puede añadir color a la nube de puntos proyectando cada punto 3D a la imagen RGB e interpolando el color [13].

2.3 Precisión y resolución de la profundidad

Como se puede apreciar en el estudio realizado en [12] los factores que influyen en la precisión y resolución con los que el sensor Kinect proporciona información son los que se describen a continuación:

Condiciones del entorno en donde se realiza las medidas: El sensor Kinect solo funciona en zonas interiores, ya que se ve afectado por la interferencia de la luz solar, provocando falsas mediciones de profundidad.

Propiedades de los objetos a medir: el sensor Kinect no es apto para realizar mediciones en objetos reflectantes o transparentes, ya que no los detecta y por lo tanto no puede medir su profundidad.

Rango de trabajo: la resolución del sensor Kinect se ve afectada por la distancia a la que se encuentren los objetos a medir, a mayor distancia, menor es la resolución con la que se obtiene los datos de profundidad. El rango de trabajo donde mejor precisión tiene al obtener los datos de profundidad, es, según [12], de 1.4m a 5.0m, pero puede funcionar entre 0.5m y 5.0m.

3 UTILIZACIÓN DEL SENSOR KINECT CON UN PC

Existen diferentes herramientas para controlar el funcionamiento el sensor Kinect desde un PC, aunque este dispositivo fue creado inicialmente para ser usado con el Microsoft Xbox 360. A continuación se describen los dos kit de desarrollo de software (SDK) más usados.

3.1 SDK para Windows 7

El SDK para Windows es una kit de herramientas para desarrollar aplicaciones que permite a los investigadores y aficionados crear aplicaciones utilizando la tecnología del sensor Kinect en computadores bajo Windows 7. El kit incluye lo siguiente:

- Controladores, para usar el sensor Kinect en un computador con Windows 7.
- Interface de programación de aplicaciones (API's) e interfaces del dispositivo, junto con documentación técnica.
- Código fuente de ejemplos.

Los requerimientos del sistema son:

Hardware:

- Ordenador con un procesador dual core 2.66 GHz o más rápido.
- Tarjeta gráfica compatible con Windows 7 que soporte DirectX® 9.0c.
- 4 GB RAM recomendado aunque puede funcionar con 2 GB de RAM.

Software:

- Windows 7 (x86 or x64).
- Visual Studio 2010 Express (u otra edición 2010)
- Microsoft .NET Framework 4.0

3.2 OpenNi SDK

El kit de desarrollo de software (SDK) llamado OpenNi es un marco de desarrollo confiable que incluye una colección de herramientas y librerías diseñadas para ayudar a los programadores en proyectos de detección 3D.

Este SDK ofrece a los usuarios la posibilidad de iniciar y desarrollar soluciones 3D y middleware de visión por computador por igual.

Los requerimientos del sistema son:

- Un procesador Pentium 4 de 1.4 GHz, o un AMD Athlon 64/FX de 1 GHz.
- Por lo menos 64Mb de memoria disponible.
- 250Mb de espacio en disco duro disponible.
- Puerto USB 2.0.
- Microsoft Visual Studio 2008 y 2010

4 EJEMPLOS DE LA INFORMACIÓN OBTENIDA DEL SENSOR KINECT UTILIZANDO ROS.

Como se ha dicho anteriormente, existen diferentes formas de obtener la nube de puntos 3D generada por el sensor Kinect. Para ello basta con instalar los controladores que permiten tener acceso a la cámara, y esto se puede hacer tanto en Linux como en Windows. El sistema operativo de robótica diseñado por Willow Garage, conocido como ROS (Robot Operating System) [14], es una de las formas más fáciles de acceder a toda la información que este sensor puede proporcionar. Con el paquete OpenNi_camera de ROS se puede tener acceso a las imágenes generadas por la cámara RGB y la cámara IR. Utilizando el visualizador de ROS llamado *rviz* se pueden apreciar dichas imágenes como se muestra en la figura 6 además del mapa de profundidad y la nube de puntos a color que muestran la figura 7.



Fig. 6. (a) Imagen RGB, (b) Imagen IR [12].

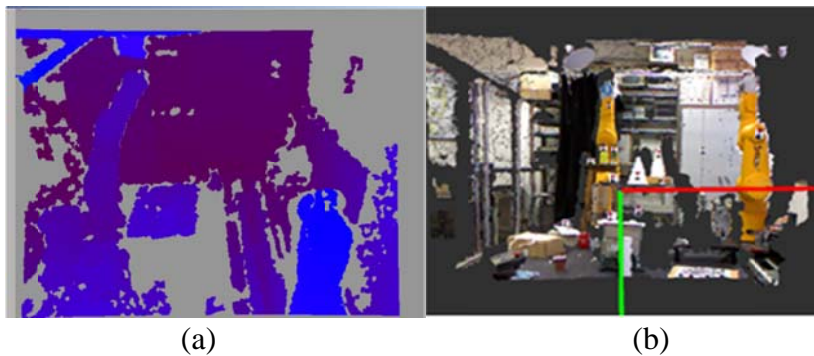


Fig. 7. (a) Mapa de disparidad, (b) Nube de puntos 3D [12].

En el estudio realizado en [12], se utiliza ROS para evaluar el funcionamiento y la precisión del sensor Kinect. Para ello se distribuye una cantidad de muestras a lo largo de una escena, con el fin de obtener la medida de la distancia a cada una de las muestras desde este sensor, a través de la nube de puntos tridimensional generada por el

mismo. Luego usando el medidor de distancia laser Bosch profesional DLE50 que tiene un rango de medición comprendido entre 0.05m y 50m con una precisión de $\pm 1.5\text{mm}$, se obtiene la distancia real, de cada una de las muestras, pudiendo así valorar la precisión del sensor. En la tabla 1 se ven reflejadas las medidas obtenidas. Los resultados de este estudio han sido muy positivos a favor del sensor Kinect, ya que se ha observado que, aunque sus cámaras poseen baja resolución, proporciona buenos resultados en un rango de trabajo de entre 1.4m y 5m, además de generar información en zonas de poca textura que otros sensores tridimensionales como las cámaras estéreo no pueden proporcionar de manera adecuada. Las principales desventajas que se han observado en éste estudio han sido que el sensor no funciona adecuadamente en entornos exteriores debido a la interferencia de la luz solar, que pierde información en objetos reflejantes o transparentes y que fuera de su rango de trabajo tiene baja precisión.

Table 1. Medidas de profundidad del sensor Kinect [12]

muestra	Mundo Real (cm)	Kinect (cm)	Error(cm)
1	145,8	146,1	0,300
2	164,0	164,9	0,900
3	169,0	170,6	1,600
4	175,0	Sin Info	Sin Info
5	184,6	183,3	1,300
6	193,4	196,9	3,500
7	206,9	207,6	0,700
8	207,9	208,8	0,900
9	233,7	235,9	2,200
10	241,4	249,6	8,200
11	247,1	251,4	4,300
12	252,8	229,7	23,100
13	260,1	266,9	6,800
14	299,1	301,9	2,800
15	480,8	490,7	9,900
16	490,4	483,8	6,600

5 ALGUNOS TRABAJOS REALIZADOS CON EL SENSOR KINECT

5.1 “Visual Odometry and Mapping for Autonomous Flight Using an RGB-D Camera” [15].

En este trabajo los autores describen un Sistema para realizar odometría y mapeo visuales, utilizando el sensor Kinect como cámara RGB-D, y su aplicación en vuelo autónomo. Este sistema permite vuelo 3D en cualquier entorno usando solo datos del sensor a bordo. Todos los cálculos y sensado requeridos para el posicionamiento local se lleva a cabo a bordo del vehículo. Los autores evalúan la efectividad de su sistema para estabilizar y controlar un micro-vehículo aéreo y demuestran su uso para construir mapas 3D detallados.

5.2 “RGBD-HuDaAct: A color-Depth Video Database for Human Daily Activity Recognition” [16].

En este trabajo los autores aportan dos contribuciones: (1) Han creado una base de datos con videos de actividades humanas llamada RGBD-HuDaAct que contiene un flujo de videos sincronizados de profundidades con color, para el reconocimiento de la actividad diaria de los humanos. (2) Han diseñado dos esquemas de fusión multi-modal que combinan de forma natural información de color y profundidad a partir de dos métodos de representación de características para el reconocimiento. A estos esquemas les han llamado puntos de interés espacio temporal (STIPS) e imágenes del historial del movimiento (MHIs).

5.3 “Unstructured Human Activity Detection from RGB Images” [17].

En este trabajo los autores realizan la detección y el reconocimiento de la actividad humana en entornos no estructurados. Para esto usan un sensor RGBD (el sensor Kinect) como sensor de entrada y calculan un conjunto de características basados en la posición y los movimientos del humano, como también basados en imágenes e información de nube de puntos. Prueban su algoritmo detectando y reconociendo 12 actividades diferentes, realizadas por 4 personas en distintos entornos y consiguen un buen funcionamiento incluso cuando la persona no ha sido vista en la etapa de entrenamiento.

5.4 “Real-time 3D visual SLAM with a hand-held RGB-D camera” [18].

Este artículo presenta un sistema de SLAM RGB-D para generar modelos 3D en color de objetos en entornos interiores utilizando el sensor Kinect. El sistema consta de cuatro importantes pasos: en primer lugar se extraen características SURF de las imágenes a color entrantes, luego se realizan correspondencias entre estas características con las características de imágenes previas. En base a estas correspondencias, se

estima la transformación relativa entre cuadros usando RANSAC. El tercer paso es mejorar la estimación inicial usando una variante del algoritmo ICP. Por último, se optimiza el grafico de posición resultante usando un solucionador de gráficos de posición [4]. La salida del algoritmo es un modelo 3D global del entorno percibido, representado como una nube de puntos a color.

5.5 “RGB-D mapping: Using Kinect-Style depth cameras for dense 3D modeling of indoor environments” [19].

En este artículo los autores investigan el uso de las cámaras para construir mapas 3D densos en entornos interiores. Presentan un sistema de mapeo RGB-D que utiliza un nuevo algoritmo de optimización combinando características visuales y alineación basado en la forma. Los autores evalúan el mapeo RGB-D en dos grandes entornos interiores y demuestran que combina eficientemente la información visual disponible en cámaras RGB-D como el sensor Kinect.

5.6 “Towards a benchmark for RGB-D SLAM evaluation” [20].

En este trabajo los autores presentan un conjunto de datos RGB-D con el objetivo de realizar una comparativa entre algoritmos de SLAM. Este conjunto de datos contiene imágenes en color, mapas de profundidad e información de la posición de la cámara. Además proponen una métrica de evaluación que puede ser usada para evaluar el rendimiento de un sistema SLAM. La comparación se lleva a cabo utilizando el sensor Kinect.

5.7 “Human Activity Detection from RGBD Images” [21].

Los autores llevan a cabo un reconocimiento de actividad utilizando el sensor Kinect. Muestran que pueden lograr un rendimiento muy fiable en la detección y reconocimiento de actividades comunes realizadas en entornos típicos humanos. Su método está basado en técnicas de aprendizaje automático. Evalúan su método en 12 actividades diferentes realizadas por 4 personas distintas en 5 entornos diferentes. Los resultados muestran un rendimiento global de 84.3% en la detección de la actividad correcta cuando la persona fue vista en la etapa de entrenamiento y de 64.2% cuando la persona no fue vista antes.

5.8 “Real-Time Plane Segmentation Using RGB-D Cameras” [22].

Los autores presentan un Sistema de obtención y procesado de información 3D, usando una cámara RGB-D, con velocidad de refresco superior a 30Hz que permite a un robot móvil detectar de forma fiable obstáculos y segmentar objetos manipulables y superficies de soporte como también la geometría de toda la escena. Usando imágenes integrales calculan las normales de la superficie local. Los puntos son luego agru-

pados, segmentados y clasificados en espacio normal y coordenadas esféricas. Los resultados demuestran que el sistema es capaz de detectar obstáculos de una manera fiable con altas velocidades de refresco, incluso en casos en que los obstáculos se mueven rápidamente o no se adhieren considerablemente al suelo.

5.9 “An evaluation of the RGB-D SLAM system” [23].

Los autores presentan un sistema para hacer SLAM utilizando una cámara RGB-D como el sensor Kinect. El sistema estima la trayectoria de un sensor Kinect y genera un modelo 3D denso del entorno. Los autores presentan la característica clave de su propuesta y evalúan su funcionamiento a través de un conjunto de datos recientemente publicado, incluyendo un conjunto de secuencias de diferentes escenas variando las condiciones de iluminación y las velocidades de la cámara. En particular, evalúan la precisión, robustez y tiempo de procesamiento para tres diferentes descriptores (SIFT, SURF, y ORB). Los experimentos demuestran que el sistema puede tratar con datos difíciles de una manera robusta en escenarios interiores y ser lo suficientemente rápido para operaciones online. El sistema está disponible como código abierto.

5.10 “Unsupervised Feature Learning for RGB-D Based Object Recognition” [24].

En este trabajo los autores presentan un algoritmo de búsqueda jerárquica de correspondencia llamado HMP. Proponen un aprendizaje no supervisado de características para el reconocimiento de objetos utilizando datos RGB-D haciendo una búsqueda adecuada de correspondencias en imágenes de color y de profundidad. Extensos experimentos utilizando varios conjuntos de datos revelan que el aprendizaje de características no supervisado de los datos sin procesar, puede producir reconocimiento con una precisión superior a la de los algoritmos de reconocimiento de objetos anteriores. Las innovaciones presentadas en este trabajo mejoran significativamente el funcionamiento de HMP aplicado a datos RGB-D.

6 CONCLUSIONES

Después de realizar este estudio sobre el sensor Kinect, se puede decir que es un dispositivo con notables características para el desarrollo de sistemas de visión artificial. Por su relación calidad-precio es un dispositivo altamente usado hoy en día por la comunidad científica y tiene las características necesarias para seguir siendo usado por mucho tiempo. Las posibilidades de trabajar con este sensor son extensas, ya que hay gran variedad de código abierto que permite su utilización tanto en Windows como en Linux, permitiendo así que muchas personas tengan acceso al dispositivo, no solo para actividades de investigación, sino que también para entretenimiento. Su sucesor, conocido como Kinect 2, ya ha salido a la venta y trae consigo más resolu-

ción, lo que permite obtener información con más precisión abriendo nuevas puertas para su uso en el mundo científico y de entretenimiento.

7 BIBLIOGRAFIA

1. Takahashi, D. (2009), «Microsoft games exec details how Project Natal was born», URL <http://venturebeat.com/2009/06/02/microsoft-games-executive->.
2. Microsoft (2011), «Kinect Sales Surpass Ten Million», URL: <http://www.xbox.com/en-us/press/archive/2011/0308-ten-million-kinects>, [Online; accessed 03-October-2011].
3. Avancini, M. Using Kinect to emulate an Interactive Whiteboard. Tesis 2011-2012.
4. MESA Imaging. Available online: <http://www.mesa-imaging.ch/> (accessed on 14 December 2011).
5. PMD [vision] CamCube 3.0. Available online: <http://www.pmdtec.com/products-services/pmdvisionr-cameras/pmdvisionr-camcube-30/> (accessed on 14 December 2011).
6. Breuer P., Eckes C., Muller S. Hand gesture recognition with a novel IR time-of-flight range camera—A pilot study. In: Gagalowicz A., Philips W., editors. Lecture Notes in Computer Science. Vol. 4418. Springer; Berlin, Germany: 2007. pp. 247–260.
7. Kahlmann T., Ingensand H. Calibration and development for increased accuracy of 3D range imaging cameras. *J. Appl. Geod.* 2008;2:1–11.
8. Kahlmann T., Remondino F., Ingensand H. Calibration for increased accuracy of the range imaging camera SwissRanger. Proceedings of ISPRS Commission V Symposium ‘Image Engineering and Vision Metrology’; Dresden, Germany. 25–27 September 2006; pp. 136–141.
9. Lichti D.D. Self-calibration of a 3D range camera. Proceedings of International Archives of the Photogrammetry Remote Sensing and Spatial Information Sciences; Beijing, China. 3–11 July 2008; pp. 927–932.
10. Lindner M., Schiller I., Kolb A., Koch R. Time-of-flight sensor calibration for accurate range sensing. *Comput. Vis. Image Underst.* 2010;114:1318–1328.
11. Shahbazi M., Homayouni S., Saadatseresht M., Sattari M. Range camera self-calibration based on integrated bundle adjustment via joint setup with a 2D digital camera. *Sensors*. 2011; 11:8721–8740. [PMC free article] [PubMed].
12. Beltran, D., Basañez, L. A Comparison between Active and Passive 3D Vision Sensors: BumblebeeXB3 and Microsoft Kinect. *Advances in Robotics Vol 1 y 2*. Springer; Madrid, Spain: 2013. pp. 725-734.
13. Khoshelham, K., Elberink, S. Accuracy and Resolution of Kinect Depth Data for Indoor Mapping Applications. *Sensors (Basel)*. 2012; 12(2): 1437–1454.
14. <http://www.willowgarage.com/pages/software/ros-platform,describes-origins-of-project-natal-game-controls/>, [Online; accessed 27-September-2011].
15. Huang, A., Bachrach, A., Henry, P., Krainin, M., Maturana, D., Fox, D., Roy, N. Visual Odometry and Mapping for Autonomous Flight Using an RGB-D Camera. *Int. Symposium on Robotics Research (ISRR)*, Flagstaff, Arizona, USA, Aug. 2011.
16. Bingbing Ni, Wang, G., Moulin, P. RGBD-HuDaAct: A Color-Depth Video Database for Human Daily Activity Recognition. *Consumer Depth Cameras for Computer Vision*. Springer, 2013. Pp 193-208.
17. Jaeyong S., Ponce, C., Selman, B., Saxena, A. Unstructured human activity detection from RGBD images. *Robotics and Automation (ICRA)*, 2012. Pp 842-849.

18. Engelharda, N., Endresa, F., Hessa, J., Sturm, J., Burgarda, W. Real-time 3D visual SLAM with a hand-held RGB-D camera. Proc. of the RGB-D Workshop on 3D Perception in Robotics at the European Robotics Forum 2011.
19. Henry, P., Krainin, M., Herbst, E., Ren, X., Fox, D. RGB-D mapping: Using Kinect-style depth cameras for dense 3D modeling of indoor environments. *The International Journal of Robotics Research* April 2012 31: 647-663, first published on February 10, 2012.
20. Sturm, J., Magnenat, S., Engelhard, N., Pomerleau, F., Colas, F., Cremers, D., Siegwart, R., Burgard, W. Towards a benchmark for RGB-D SLAM evaluation. In Proc. of the RGB-D Workshop on Advanced Reasoning with Depth Cameras at Robotics: Science and Systems Conf. (RSS), 2011.
21. Sung, J., Ponce, C., Selman, B., Saxena, A. Human Activity Detection from RGBD Images. Plan, Activity, and Intent Recognition: Papers from the 2011 AAAI Workshop (WS-11-16).
22. Holz, D., Holzer, S., Bogdan, R., Behnke, S. Real-Time Plane Segmentation Using RGB-D Cameras, *RoboCup 2011: Robot Soccer World Cup XV*. Springer. Pp 306-317.
23. Endres, F., Hess, J., Engelhard, N., Sturm, J., Cremers, D., Burgard, W. An evaluation of the RGB-D SLAM system. *Robotics and Automation (ICRA)*, 2012. Pp 1691 – 1696.
24. Bo, L., Ren, X., Fox, D. Unsupervised Feature Learning for RGB-D Based Object Recognition. In ISER, June 2012.