

Evaluación del proceso de Reconstrucción 3D mediante el uso del Sensor de Profundidad Kinect

Chacón Jackson Karen Ibeth¹, Cruz Martínez Yessica², Pedraza Ortega Jesús Carlos³, Delgado Rosas Manuel³, Aceves Fernández Marco Antonio³, Sotomayor Olmedo Artemio⁴

Instituto Tecnológico de Cerro Azul¹, Instituto Tecnológico Superior de Alvarado², Facultad de Informática de la Universidad Autónoma de Querétaro³, Facultad de Ingeniería de la Universidad Autónoma de Querétaro³

Resumen

Este artículo presenta la evaluación del procesos para llevar a cabo la reconstrucción 3D de objetos y escenarios, utilizando el sensor de profundidad Kinect.

Se propone el uso del Processing para realizar la reconstrucción de manera manual; realizando las capturas del objeto como nubes de puntos con el Kinect, mediante la librería OpenNI, para posteriormente trasladar los mapas de profundidad obtenidos a un software conocido como MeshLab, en el cual se lleva a cabo la unión de mallas mediante la aplicación del algoritmo ICP y obtener al final la reconstrucción tridimensional del objeto. Con el propósito validar el método propuesto, se lleva a cabo la reconstrucción de objetos utilizando un software comercial y se compara con la metodología propuesta.

Palabras clave: Nube de puntos, Algoritmo ICP, Reconstrucción 3D.

1. Introducción y objetivos

En los últimos tres años, el uso del sensor de profundidad Kinect, ha tomado importancia como una herramienta de desarrollo de proyectos de investigación, a partir de que fue lanzado originalmente como una extensión de una consola de videojuegos, en la actualidad se ha extendido más allá de la industria del videojuego.

Esta tecnología ha dado paso al desarrollo de nuevas formas creativas de interactuar con computadoras y tarjetas de propósito particular basadas en microcontroladores.

Entre el desarrollo de proyectos de investigación que han ido surgiendo a partir de la aparición del Kinect han sido principalmente en Estados Unidos y Europa:

En los primeros trabajos, mediante el uso del sensor Kinect para la construcción de mapas tridimensionales y la aplicación en un sistema teleoperado, se desarrollan los algoritmos de mapeo tridimensional y finalmente se aplican para realizar la teleoperación de un robot móvil [1].

Posteriormente, otras investigaciones se enfocaron en la realización de un sistema de tracking usando el sensor Kinect, el SDK no oficial para Kinect OpenNI, utilizando y la librería de visión artificial PointCloudsLibrary (PCL), este sistema es implementado para realizar tracking a imágenes médicas, lo cual requiere que el seguimiento de movimiento sea preciso y fiable, finalmente realizan pruebas y obtienen muy buenos resultados [2].

Una compañía rusa llamada ARDoor ha creado una tienda en el "espejo" llamada Puerta AR Kinect Fitting Room que sobrepone la ropa en los compradores. Ellos pueden rápidamente ver cómo les va quedar la ropa que quieren comprar. Este trabajo es de un investigador de la Universidad Técnica de Munich creó algo que él llama el "espejo mágico". Diseñado para la enseñanza de la anatomía, el sistema crea la ilusión de un espejo que muestra las entrañas del usuario a través de un "agujero" en el cuerpo [3].

Los investigadores del MIT construyeron un dispositivo que controla la computadora a través de movimientos de la mano y los dedos Investigadores de Fraunhofer Institute de Alemania crearon un elegante sistema de control de PCs con Windows que se basa en gestos de las manos.

En la mayoría de los desarrollos presentados con anterioridad [7-10], no se había considerado el

proceso de reconstrucción total (360° grados) utilizando únicamente un sensor Kinect.

El objetivo de éste artículo es dar a conocer las diferentes formas de llevar a cabo la reconstrucción 3D utilizando únicamente un Sensor Kinect. Este método propone la captura de múltiples mallas de un mismo objeto con un ángulo de rotación, y por medio de la unión de mallas mediante el algoritmo ICP se obtiene la reconstrucción total del objeto (360°).

2. Kinect

La cámara Kinect, desarrollada por el fabricante PrimeSense para Microsoft, está orientada para su utilización únicamente con la consola XBOX 360. Sin embargo, al estar equipada con un conector USB, es posible conectarla a un PC. Aunque Microsoft no ha lanzado drivers oficiales para su uso en PC, en poco tiempo algunos desarrolladores han propuesto potentes librerías para acceder a la funcionalidad de Kinect.

Estas librerías serán presumiblemente compatibles con nuevas cámaras de las mismas características que están por llegar al mercado, como la cámara wavi-xtion presentada en el CES 2011, que llegará de la mano de PrimeSense y ASUS [4].

El sensor Kinect es un dispositivo que cuenta con un sensor CMOS de imágenes en RGB (CMOS Image Sensor RGB) y un sensor CMOS de imágenes infrarrojas (CMOS Image Sensor For IR), y un micrófono multi-array, su frecuencia de video es de 30 Hz y cuenta con una resolución VGA de 640 x 480 pixeles, su límite de rango visual es aproximado de 1.2 a 1.5 m. aunque en las pruebas que se han realizado, los resultados han arrojado que la distancia mínima es de 50 cm aproximadamente en el cual se obtienen buenas tomas del objeto, por lo que se ha optado por colocar los objetos a ésta distancia obteniendo buenos resultados.

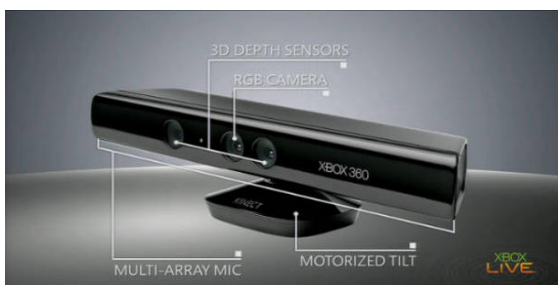


Fig. 1. Sensor Kinect

Las librerías no oficiales que se han desarrollado para trabajar con el Sensor Kinect en la PC son:

- PrimeSense NITE 1.5.2.2
- Sensor 5.1.2.1
- Sensor Kinect mod 5.1.2.1

2.1 Processing con Kinect

Processing es conocido como un lenguaje de programación orientado principalmente para diseñadores. Está basado en el lenguaje java y es de código abierto, compatible para computadoras con sistemas operativos Windows, Linux y Mac. Para poder trabajar con Processing en conjunto con el Kinect debemos realizar algunas instalaciones de librerías las cuales son:

- OpenNI 1.5.4.0
- SimpleOpenNI

Utilizando Processing podemos obtener una nube de puntos que son convertidas a una malla tridimensional del contorno del objeto por medio de algoritmos de triangulación. Mediante un grupo de mallas tridimensionales se puede llevar a cabo la reconstrucción tridimensional completa (360°) de los objetos.

2.2 Reconstructme

Este software emplea tecnologías de vanguardia para capturar modelos 3D en tiempo real, es similar a la de una cámara de vídeo convencional, simplemente moverse por el objeto a ser capturado. Sin embargo, en lugar de una secuencia de vídeo se obtiene un modelo 3D completo en tiempo real [5].

2.3 MeshLab

Es un software para la unión de mallas, en el cual se pueden realizar filtros y suavizados a las mallas.

Para poder trabajar las mallas obtenidas con Processing en MeshLab se utilizará la biblioteca modelbuilder para éste propósito, ya que proporciona herramientas que ayudan a que sea más fácil la creación, manipulación y exportación de las mallas 3D en MeshLab.

3 Metodología

Para el desarrollo de este proyecto la metodología propuesta se muestra en la figura 2.

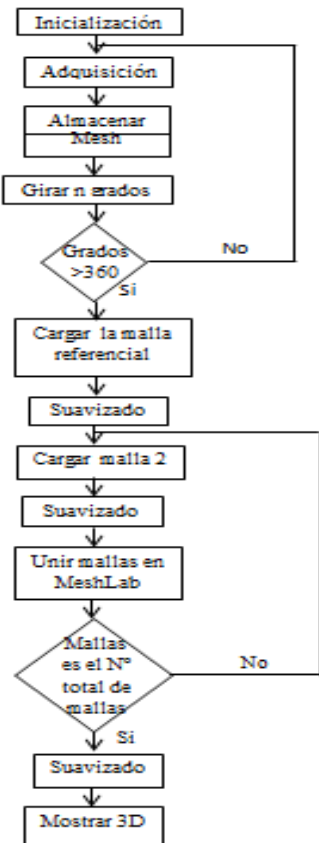


Fig. 2. Diagrama de Metodología

3.3 Adquisición

Primero se tomaron las imágenes de los objetos a reconstruir, se coloca la base giratoria la cual está dividida en 8 secciones de 22.5 grados cada una y al finalizar las 8 secciones, obtener los 360° del objeto; el objeto se coloca encima de la base a una distancia de 50 cm aproximadamente del sensor Kinect, como se muestra en la Fig. 3.

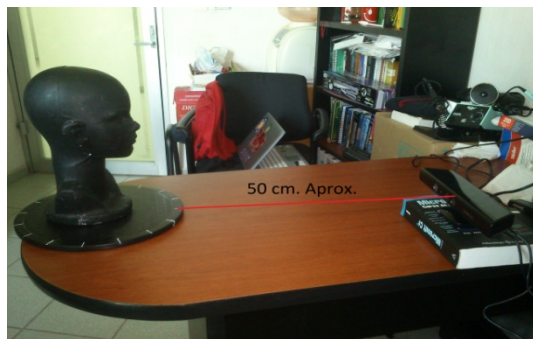


Fig. 3. Distancia aproximada 50 cm.

3.4 Almacenamiento de Mesh

Una vez que se tiene la primera toma del objeto, se guarda en el formato deseado (*.ply, *.obj, *.stl, *.3ds, *.wrl entre otros), posteriormente se gira la base y se toma la siguiente imagen y así sucesivamente hasta obtener las 8 capturas del objeto figura 3.1. y figura 3.2.



Fig. 3.1. Captura de imagen a 45°



Fig. 3.2. Captura de imagen a 180°

3.3 Carga de Mallas y suavizado en MeshLab

Al terminar de obtener todas las capturas del objeto, se carga en MeshLab la primera nube de puntos tomada, la cuál será nuestra referencia para comenzar con la reconstrucción; lo primero que se realizara será la limpieza de basura que haya logrado captar el kinect durante la toma del objeto.

Después de tener la malla limpia se realizará un filtrado, para de ésta manera realizarle un suavizado, esto se puede lograr mediante algoritmos basados en b-splines.

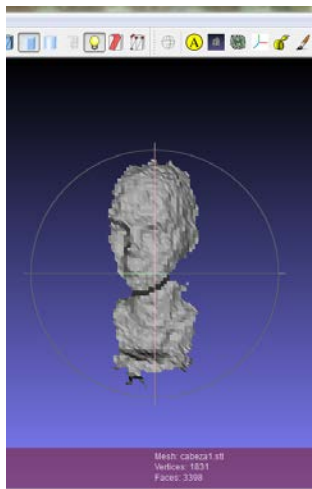


Fig. 3.3. Malla sin suavizar

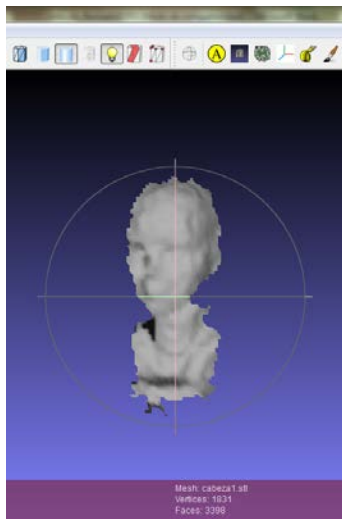


Fig. 3.4. Malla suavizada con el filtro

Una vez de tener la primera malla de referencia limpia y suavizada se procede a guardarla y realizar el mismo procedimiento con la malla dos y así sucesivamente con las demás hasta tener todas listas.

3.4 Unión de mallas con algoritmo Iterative Closest Point (ICP)

Una vez que se tienen todas las mallas listas se procede a cargar la primera malla de referencia y posteriormente la segunda y se utiliza el algoritmo ICP en el cual la primera fase consiste en establecer los pares de puntos más cercanos entre las 2 mallas a unir.

3.5 Ultimo suavizado de Mallas

Después de tener todas las mallas unidas como corresponde se les aplica un último suavizado a toda la imagen terminada con el filtro “Compute Normals for point sets”.

Al terminar éste último procedimiento se puede apreciar la Reconstrucción 3D completa del objeto con un acabado más suave y muy parecido al objeto en tiempo real.

4.- Resultados del Trabajo

Se realizaron varias tomas de nubes de puntos de diferentes objetos, una vez que se obtuvieron cada una de las tomas en sus 360°, fueron importadas a MeshLab, donde se llevó a cabo la unión de mallas.

Se importaron dos diferentes mallas del mismo objeto, en las cuales la idea es alinear las dos nubes de puntos marcando cuatro puntos homólogos de manera manual, para saber dónde colocar cada uno de los puntos, te guías por las capturas de la imagen, y los colocas donde se cree que quedarán alineadas correctamente como se muestra en la figura 4, una vez que se tienen marcados, se aplicara el algoritmo ICP y al ser procesado este alineara las dos mallas como se muestra en la figura 4.1.

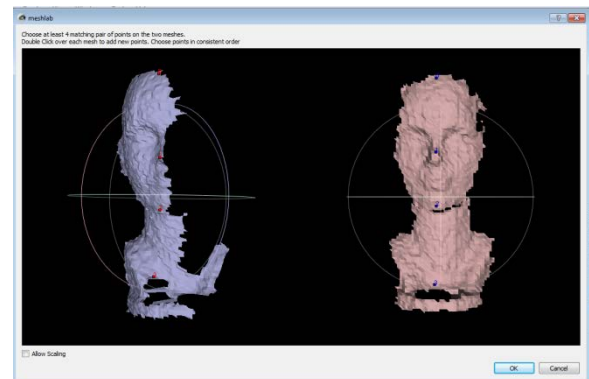


Fig. 4. Marcar Puntos Homólogos

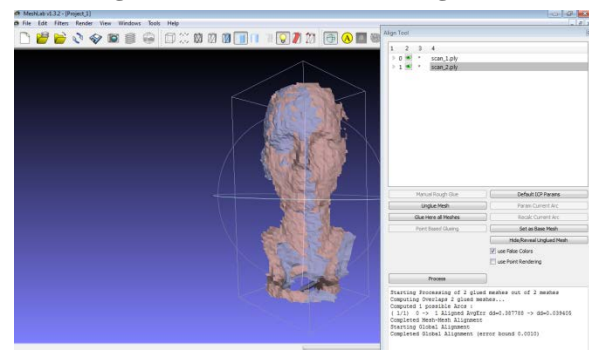


Fig. 4.1. Alineamiento de dos mallas.

Al seguir este procedimiento de unión de mallas de nubes de puntos, se formó la reconstrucción del objeto en sus 360° como se observa en la figura 4.2. Y por último se realizó un filtrado y suavizado a la figura que se obtuvo del objeto ya reconstruido completamente, como se muestra en la figura 4.3

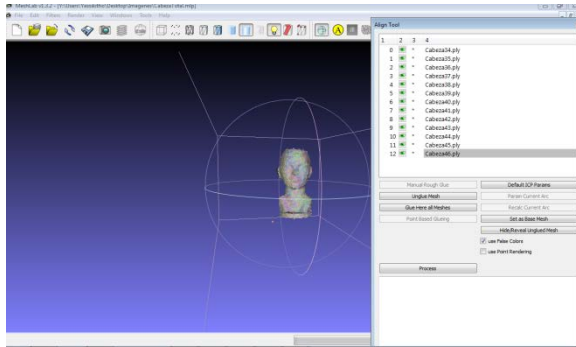


Fig. 4.2 Figura 360°

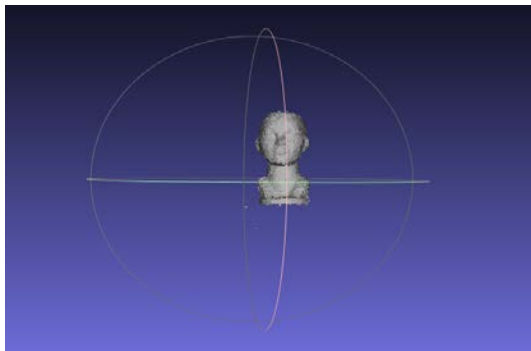


Fig. 4.3. Resultado Final

4 Conclusiones y Trabajo a futuro

Se propuso una metodología para llevar a cabo el proceso de reconstrucción 3D de un objeto en 360°. En cada una de las etapas realizadas en el trabajo, arrojaron resultados favorables debido a que se obtuvo la forma de los objetos a digitalizar presentando un error muy pequeño (menor al 5%). El sensor Kinect obtiene una nube de puntos, las cuales después se importan en MeshLab para realizar una reconstrucción completa de un objeto, no obstante se pueden realizar diferentes pruebas donde se utilicen objetos de un tamaño más grande, como por ejemplo a un ser humano, para la realización de

dichas pruebas se contó con una base giratoria y una cabeza de unicel mostrada en la figura 4.5.



Fig. 4.5 Base giratoria y Cabeza de Unicel

Uno de los trabajos a futuro que se pueden realizar es que por medio de un escaneo por medio de una línea láser la toma del objeto tenga una mejor resolución, y por lo tanto se pueda elaborar una reconstrucción tridimensional completa del objeto donde este muestre un filtrado y un suavizado más parecido al objeto, con la finalidad de establecer una comparación con los resultados.

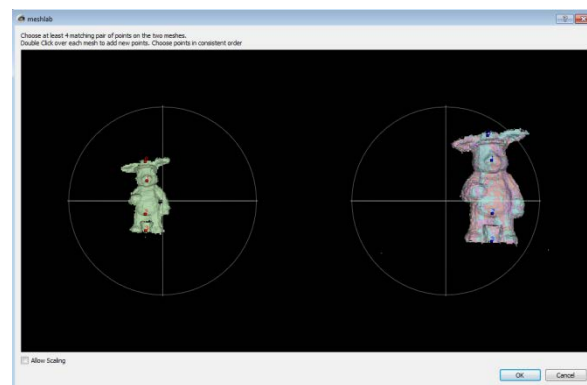


Fig. 4.6. Captura con Processing

En la figura 4.6 se puede mostrar la imagen de un objeto capturado con Processing, en la cual se está realizando la unión de mallas, y en la siguiente figura 4.7 se observa el objeto reconstruido completamente.

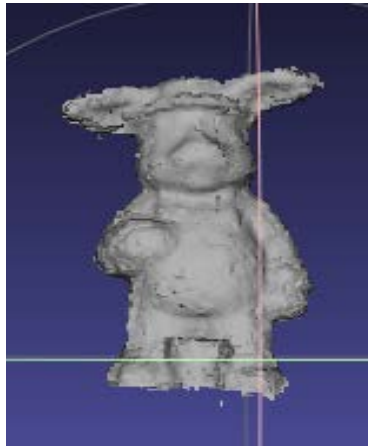


Fig. 4.7 Reconstrucción completa con Processing.

Ahora se presenta una toma de un objeto con escaneo de línea láser, se puede observar que la calidad de la imagen tiene una mejor resolución por lo cual se puede ver claramente que la figura tiene un mejor acabado que la toma de la cámara con el sensor Kinect véase figura 4.8.

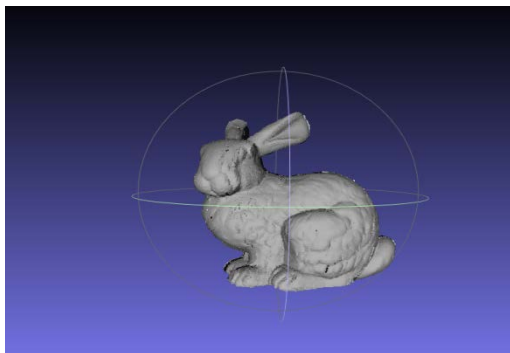


Fig. 4.8. Reconstrucción de un objeto en MeshLab de un objeto con escaneo por línea láser

Por otro lado uno de los trabajos a futuro que se pretende realizar es una aplicación basada en MeshLab y Processing, la cual contenga una interfaz en la que solo se gire el objeto y automáticamente se esté llevando a cabo la reconstrucción tridimensional en de forma automática.

Referencias

[1] Woodall W., Bevil D., "Using the Microsoft Kinect for 3D Map Building and Teleoperation",

Position Location and Navigation Symposium (PLANS), 2012 IEEE/ION, Páginas:1054 – 1061, Myrtle Beach, SC, 23 a 26 abril, 2012

[2] Philip J. Noonan; Tim F. Cootes; William A. Hallett; Rainer Hinz, "The Design and Initial Calibration of an Optical Tracking System Using the Microsoft Kinect, Institute of Electrical and Electronics, Engineers, 2011 IEEE Nuclear science symposium conference record, 2011.

[3] Villaseñor C., "SDK de Kinect: historia de éxito accidental de Microsoft", PCWorld México, Revista en línea, 201, <http://www.pcworld.com.mx/Articulos/13303.htm>

[4] Asus Wavi Xtion, Solución de tecnología PrimeSense para ordenadores, <http://www.wavixtion.es/>,

[5] <http://reconstructme.net/projects/reconstructmeqt/>, visitada el 7 de agosto de 2013

[6] "El Repositorio de escaneo 3D Stanford" <http://graphics.stanford.edu/data/3Dscanrep/>, visitada el 29 de Julio de 2013

[7] Izadi, Shahram, David Kim, Otmar Hilliges, David Molyneaux, Richard Newcombe, Pushmeet Kohli, Jamie Shotton et al. "KinectFusion: 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, pp. 559-568. ACM, 2011.

[8] Smisek, Jan, Michal Jancosek, and Tomas Pajdla. "3D with Kinect." In Consumer Depth Cameras for Computer Vision, pp. 3-25. Springer London, 2013.

[9] Henry, Peter, Michael Krainin, Evan Herbst, Xiaofeng Ren, and Dieter Fox. "RGB-D mapping: Using depth cameras for dense 3D modeling of indoor environments." In the 12th International Symposium on Experimental Robotics (ISER), vol. 20, pp. 22-25. 2010.

[10] Tong, Jing, Jin Zhou, Ligang Liu, Zhigeng Pan, and Hao Yan. "Scanning 3d full human bodies using kinects." Visualization and Computer Graphics, IEEE Transactions on 18, no. 4 (2012): 643-650.