ISSN: 1692-7257 - Volumen 3 – Número Especial - 2020

Revista Colombiana de
Tecnologías de Avanzada

Recibido: 1 de abril 2020
Aceptado: 3 agosto 2020

KINECT INTERACTION SYSTEM APPLIED TO PROCESS MANIPULATION

SISTEMA DE INTERACCIÓN CON KINECT APLICADO A MANIPULACIÓN DE PROCESOS

Camilo Pardo-Beainy*, Edgar Gutiérrez-Cáceres*,
Daniel Pardo*, Milton Medina*
Fabián Jiménez**

***Universidad Santo Tomás**
Electronic Engineering Faculty,
Tunja, Boyacá Colombia.
Tel: 57-8-7440404, Ext. 5612
E-mail: camilo.pardo@usantoto.edu.co, eagutierrezca@gmail.com,
pardo_1104@hotmail.com, mer1156@hotmail.com

****Universidad Pedagógica y Tecnológica de Colombia**
Electronic Engineering School,
Tunja, Boyacá, Colombia.
Tel: 57-8-7405626 Ext: 2565
E-mail: fabian.jimenez02@uptc.edu.co

Abstract: This article presents the aspects related to the development of an interface that uses a depth and RGB image of a Kinect to analyze the movements of a user and the manipulation of an environment with specific characteristics. This system works as a manipulation interface, where a user controls different applications from the analysis of the information captured from the acquired images. The Kinect system captures the information, and the program performs filtering, selection and evaluation tasks to carry out an action.

Keywords: Kinect, Digital Image Processing, Depth Map, Process Manipulation, Tracking Process.

Resumen: En este artículo se presentan los aspectos relacionados con el desarrollo de una interfaz que emplea una imagen de profundidad y RGB de un Kinect para analizar los movimientos de un usuario y la manipulación de un entorno con características específicas. Este sistema funciona como una interfaz de manipulación, en donde un usuario controla diferentes aplicaciones a partir del análisis de la información captada de las imágenes adquiridas. El sistema Kinect capta la información y el programa desarrolla tareas de filtrado, selección y evaluación de esta, para realizar una acción.

Palabras clave: Kinect, procesamiento digital de imágenes, Mapa de Profundidad, Manipulación de Procesos, Proceso de seguimiento.

1. INTRODUCTION

User-machine interfaces in automation systems represent one of the main focuses of interest at an industrial level, related to control and process handling systems; communication with the user in these processes requires easy interaction with the machine (López et al., 2013; Niño et al., 2020), where the interface offers the necessary possibilities to execute the actions or changes required on the system (Pardo-Beainy et al., 2014).

In this paper, the procedures developed with the Kinect sensor that allowed to recognize the operating capabilities of the device, as well as its technical characteristics, are related with the purpose of developing an interface with the ability to track the movements of some objects or extremities of the user's body in order to present a precise, attractive and simple interface that allows the user to carry out manipulation commands associated with a specific system.

Two fundamental phases were considered for this development; for the first phase, a Kinect device recognition and verification program was used, which was performed in LabVIEW software, using the subVIs to capture the RGB images and depth maps delivered by the device. For the second phase, filtering procedures were contemplated, detection of coordinates related to the depth maps of the captured scene and monitoring of circular-shaped sections in an established selection range.

2. KINECT DEVICE CHARACTERISTICS

The Kinect Sensor has received a lot of interest from the academic community, due to its rapid recognition capabilities of a person's position, becoming a significant 3D sensor. It has advantages such as its low cost, high capture speed and generation of measurement parameters, which make it a suitable device for object recognition, 3D reconstruction and applications in Robotics.(Vivas et al., 2017).

2.1 Kinect as 3D Sensor

This sensor has an elongated shape and is connected to a base, with the purpose of locating it in a horizontal position. The kinect has:

- Infrared projector and camera (which function as a depth detection system).
- RGB camera.
- Two-way multi-array microphone.

These elements allow the acquisition of images of users and their movements, additionally, facial and voice recognition tasks can be developed (Ivorra Martínez, 2015).

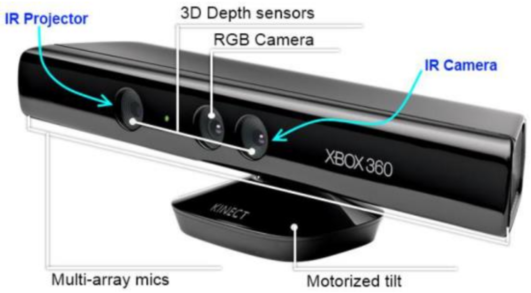


Fig. 1. Kinect Sensor Features
(Eduardo et al., 2014).

2.2 Infrared Image

The Infrared camera is used to acquire and decode the patterns of points emitted by the infrared projector; in this way a triangulation is carried out in the 3D scene. This camera can also present tangential and radial distortions(Smisek et al., 2011).

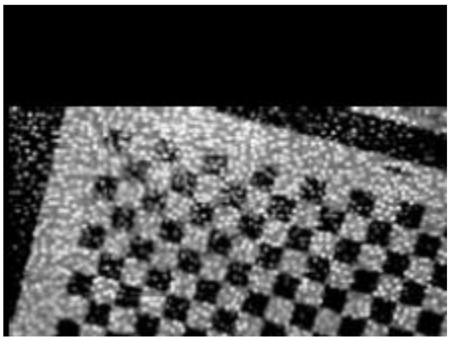


Fig. 2. Infrared image captured by the Kinect.
(Štrbac, Markovi, et al., 2012)

2.3 RGB Image

The Kinect sensor's RGB camera has the ability to make 30 acquisitions per second. The above represents 1 frame acquired every 30msec using the standard resolution mode provided by the sensor: 640x480 in RGB format.

Fig. 3. Kinect Sensor RGB Image.
(Štrbac, Markovi, et al., 2012)

2.4 Depth Map

One of the most relevant features of the Kinect device is the ability to generate a depth map of the scene. This depth map is built through a triangulation process between a beam projected in infrared and an image captured by the IR camera. Figures 3 and 4 present an example of an RGB and depth sensing system respectively using the Kinect sensor. (Štrbac, Markoviü, et al., 2012).

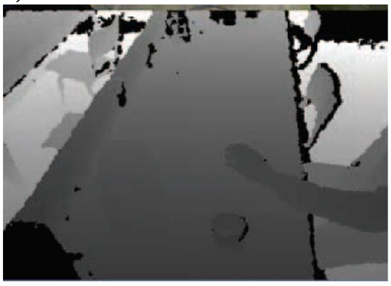


Fig. 4. Depth map delivered by the Kinect.
(Štrbac, Markovi, et al., 2012)

The Kinect captures images from a CMOS color sensor operating at a frequency of 30Hz, encoded in 32-bit data with a resolution of 640x480 pixels in VGA resolution. The monochrome CMOS system has a QVGA resolution of 320x240 pixels and is encoded in 16-bit data, with a sensitivity of 65536 levels. (Štrbac, Markoviü, et al., 2012).

To develop the calculation of distances between the sensor and an object, the Kinect device projector emits an infrared beam that projects a cloud of points on the objects in the scene; through the IR camera the pattern is captured, and the device hardware calculates the depth by triangulation for each emitted point. The Kinect sensor has a measurement range of 40cm to 4m (Mathe et al., 2012).

Figure 5 shows the Kinect camera system and sensor components, which are associated with the image recognition process.

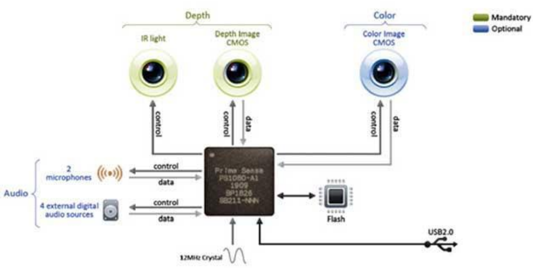


Fig. 5. Sensor components.
(Mathe et al., 2012)

The camera system has two main purposes:

- Generate the 3D depth map of the captured scene.
- Recognize users who are in movement among other objects in a captured image, starting from the recognition of the segments that represent the joints of a person's body and a scheme represented in gray scale (Mathe et al., 2012).

The Kinect tries to recognize the distance at which the objects in a scene are, also acquiring movements in real time, thus distinguishing the depth at which each object is, presenting a resolution of 1cm and generating an estimate of height and width of approximately 3mm.. Figure 6 shows the hardware system composed of the IR projector and the infrared camera, added to the firmware of the device and a processor system that uses algorithms oriented to the processing of three-dimensional images. (Mathe et al., 2012).

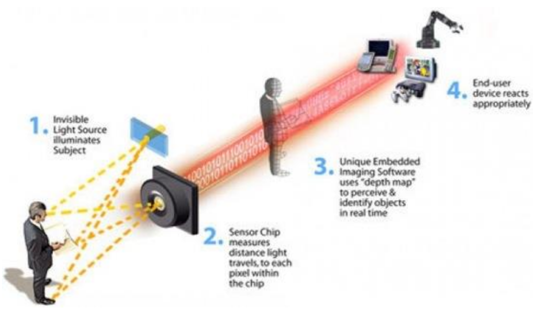
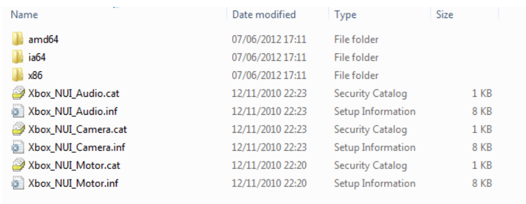


Fig. 6. IR imaging.
(Mathe et al., 2012)

3. PROPOSED ALGORITHM

3.1 Kinect Sensor Recognition and Verification in LabVIEW Software.

In order to operate the Kinect device with LabVIEW software, it is necessary to consider the device drivers so that it can be properly recognized, in table I and figure 7, the required drivers for device recognition are shown:



Name	Date modified	Type	Size
amd64	07/06/2012 17:11	File folder	
ia64	07/06/2012 17:11	File folder	
x86	07/06/2012 17:11	File folder	
Xbox_NUI_Audio.cat	12/11/2010 22:23	Security Catalog	1 KB
Xbox_NUI_Audio.inf	12/11/2010 22:23	Setup Information	8 KB
Xbox_NUI_Camera.cat	12/11/2010 22:23	Security Catalog	1 KB
Xbox_NUI_Camera.inf	12/11/2010 22:23	Setup Information	8 KB
Xbox_NUI_Motor.cat	12/11/2010 22:20	Security Catalog	1 KB
Xbox_NUI_Motor.inf	12/11/2010 22:20	Setup Information	8 KB

Fig. 7. Kinect drivers.

Tabla 1: Drivers for Installation

Device	Associated Drivers
Audio	Xbox_NUI_Audio
Video	Xbox_NUI_Camera
Kinect Engine	Xbox_NUI_Motor

Additionally, to acquire the data provided by the Kinect, it is necessary to consider some subVIs for handling the device. The most relevant ones are listed below:

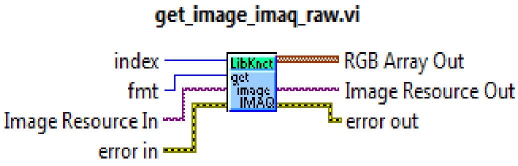


Fig. 8. RGB image acquisition subVI.

Get_image_imaq. This subVI supplies the RGB image delivered by the Kinect, as well as a set of data associated with the acquired image.

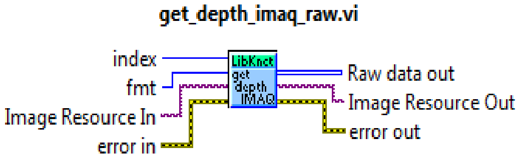


Fig. 9. Depth Map Acquisition SubVI.

Get_depth_imaq, returns the depth image of the scene, where the objects farthest and closest to the device are shown.

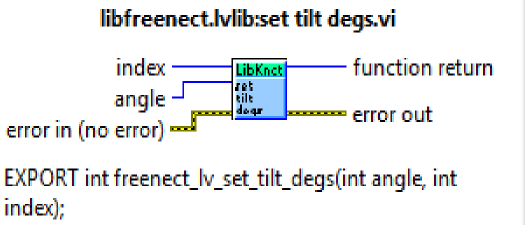


Fig. 10. Sensor View Angle SubVI.

The subVI Set tilt degs, provides a tool to adjust the desired Angle for the servomotor to position the Kinect as directed by the user.

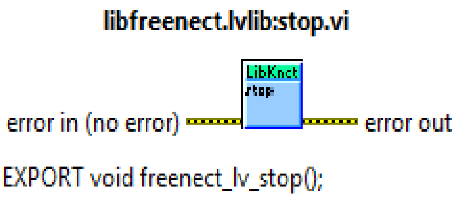


Fig. 11. Data stop subVI.

Stop, this subVI stops the execution of the previous virtual instruments.

In order to corroborate the proper functioning of the virtual instruments previously presented, a program called "Point Cloud" was designed, responsible for acquiring and viewing the different data delivered by the Kinect sensor. This program can be seen in figure 12.

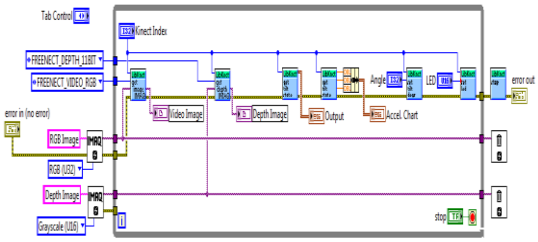


Fig. 12. Kinect Device Test Program.

With the program presented in the previous figure, it is possible to obtain the depth map and the RGB image of the scene, in addition to providing the Angle of the sensor, characteristics of the accelerometer, among other parameters. Figure 13 presents an example of the RGB Image and the acquired depth map:

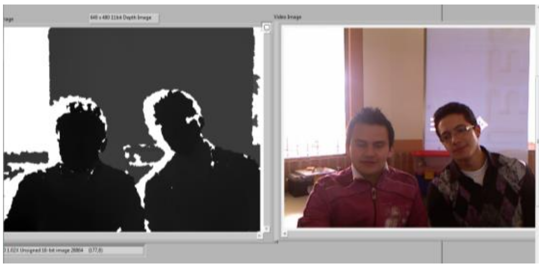


Fig. 13. Images (RGB and depth) delivered by the Kinect sensor.

3.2 Coordinate detection and filtering through depth map

With the depth map, the most distant and closest elements of the image are determined, which allows defining ranges of working depths and in this way generating a depth filter, considering only the data that are at the required depths. (Bevly, 2012).

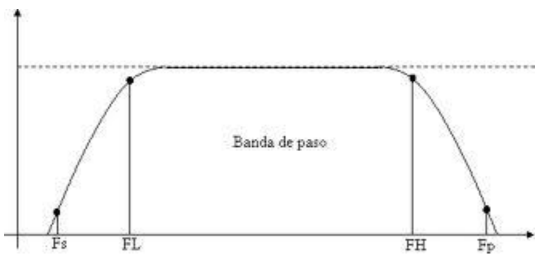


Fig. 14. Image depth working range.

Figure 14 presents the aforementioned depth range; only objects detected in this range will be considered. The range was chosen by selecting the best sensor detection points; For this, numerical values are taken that are proportional to the distance between the objects and the device.

The range of depths delivered by the Kinect varied between 1.2 and 3.5 meters; with an equivalent numerical value that varies between 10,000 and 70000. Table 2 presents the selected range, taking a depth between 1.5 and 2.0 meters.

Tabla 2. Numerical values, associated with the depth in the selected range.

Minimum Range Distance	Maximum Range Distance	Minimum Value of selected depth	Maximum Value of selected depth
1,5m.	2,0m.	20000	25000

Objects that are outside the mentioned range will be removed from the image and replaced with a white color; likewise, the objects within the range will be presented in black, as shown in figure 15.

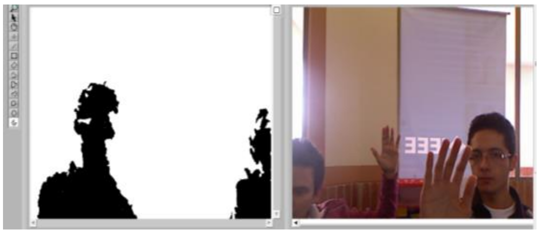


Fig. 15. Functional working range

Having the definition of the desired depths, the next step consists of recognizing a particular object in the scene within the allowed range. In this case, filtering and detection procedures applied to the black and white image were generated. To perform edge smoothing, a Laplacian filter is used, shown in Figure 16; and subsequently a detection of objects in the shape of a circle is carried out to give a specific detection coordinate (Štrbac, Markoviü, et al., 2012; Um et al., 2011).

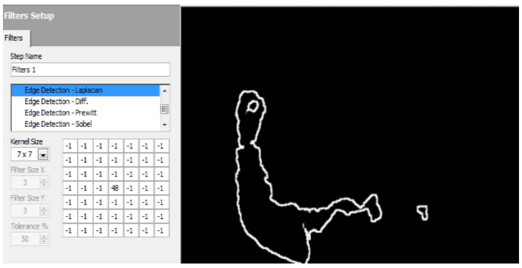


Fig. 16. Laplacian filter.

In figure 17, the coordinate of the circle detected in a plane is observed; With this coordinate, a numerical value is established and gives the user the possibility of locating at a specific point of the interface, to interact with the process to be worked on.

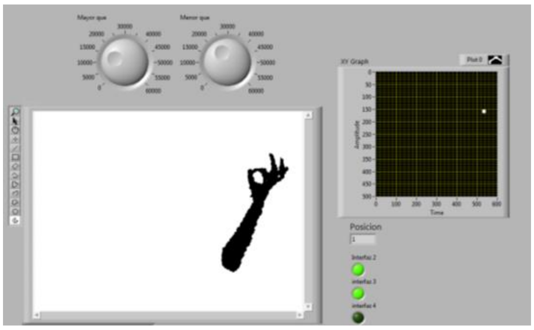


Fig. 17. Coordinate detection on the environment.

The value of the coordinate detected at the required depth is located on a plane previously divided into four quadrants. If the value of the coordinate complies with the values of the subdivision, an indicator led will light that points to the selected process within the system with which the user is interacting, as shown in figure 18.

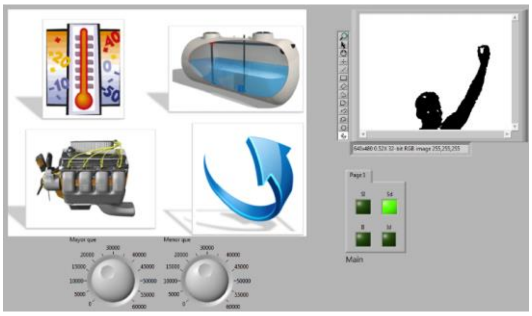


Fig. 18. Division of work quadrants in the system

In figure 19, an image with 4 LEDs is presented, each one of them is associated with one of the quadrants of the plane (upper right, upper left, lower right, lower left); if the coordinate of the detected circle is within the mentioned quadrant, this region will be activated; that is, when the user locates a circle in one of the aforementioned quadrants, the system interface will generate an action on the process associated with the quadrant, for example activating an alarm, a solenoid valve, etc.

Fig. 19. Indicators associated with the quadrants

3.3 Applications

The developed system allows the manipulation of different applications and / or processes; however, within the initial tests manipulations of processes such as level, speed and temperature were developed. These variables are present in the control of an industrial fluid simulation plant; having the following advantages:

- Ability to manipulate applications through artificial vision, replacing traditional interfaces.
- Incorporate new devices and sensors with the use of new technologies converging to industrial processes.
- Operate multiple processes through a simple, dynamic and comfortable interface for the user.

The interface developed can be linked to new applications such as Control Interfaces in Home Automation, Robotics, Medicine, Psychology, Security, among others.

4. CONCLUSIONS

At present, the development of applications using computer vision is opening new perspectives in areas such as industrial instrumentation, medicine, etc. The interface developed in this project emulates a cursor to select and interact with processes within a system, giving the possibility of non-contact interaction.

The use of subVIs and the work with LabVIEW software, allowed to facilitate the data acquisition processes of the Kinect sensor, and to link its operation with industrial processes and SCADA systems.

The characteristics and properties of the kinect facilitated the development of the program and the processing of the information, as well as the acquisition of depth maps in reduced times and the manipulation of each of the system tasks.

REFERENCES

Bevly, D. (2012). Using the Microsoft Kinect for 3D Map Building and Teleoperation. *Transform*, 1054–1061.

Chiang, I., & Chen, S. (2012). Institutionalized Older Adults with Wheelchairs. *Time*. https://doi.org/10.1109/DIGITEL.2012.69

Eduardo, J., Salvatore, J., Osio, M., & Morales. (2014). Detección de objetos utilizando el sensor Kinect. *12 Th Latin American and Caribbean Conference for Engineering and Technology*, 1–10.

Ivorra Martínez, E. (2015). *Desarrollo de técnicas de visión hiperespectral y tridimensional para el sector agroalimentario* [Universidad Politécnica de Valencia]. https://riunet.upv.es/bitstream/handle/10251/48541/Ivorra - Desarrollo de técnicas de visión hiperespectral y tridimensional para el sector agroali....pdf?sequence=1

López, D. A., Hernández, J. E., & Peña Cortes, C. A. (2013). Advances in the control of bipedal platforms using the system Kinect. *Revista Colombiana de Tecnologías de Avanzada*, *2*, 117–124. https://doi.org/https://doi.org/10.24054/16927257.v22.n22.2013.419

Mathe, L., Samban, D., & Gómez, G. (2012). Estudio del funcionamiento del sensor Kinect y aplicaciones para bioingeniería. *Interface*, 1–6.

Niño, C., Castro, S., & Medina, B. (2020). Characterization for location in the capture of video applied to artificial vision techniques in the people detection. *Revista Colombiana De Tecnologias De Avanzada (Rcta)*, *38*(38), 121–143.

Pardo-Beainy, C., Gutiérrez, E., Jiménez López, F. R., & Sosa, L. (2014). Drawing robot controlled through the kinect sensor. *Revista Colombiana de Tecnologıas de Avanzada*, *2*(24), 43–48.

Smisek, J., Jancosek, M., & Pajdla, T. (2011). 3D with Kinect. *Image (Rochester, N.Y.)*, 1154–1160.

Štrbac, M., Markovi, M., & Popovi, D. B. (2012). Kinect in Neurorehabilitation : Computer Vision System for Real Time Hand and Object Detection and Distance Estimation. *11th Symposium on Neural Network Applications in Electrical Engineering, NEUREL-2012*, 127–132.

Štrbac, M., Markoviü, M., & Popoviü, D. B. (2012). Kinect in Neurorehabilitation : Computer Vision. *Image (Rochester, N.Y.)*.

Um, D., Ryu, D., & Kal, M. (2011). *Multiple Intensity Differentiation for 3-D Surface Reconstruction With Mono-Vision Infrared*. *11*(12), 3352–3358.

Vivas, J. E., Imbachi, S. C., & Florez, J. F. (2017). Assessment of estimation techniques optics depth measures. *Revista Colombiana De Tecnologias De Avanzada (Rcta)*, *1*(29), 138–143.