Developing Natural User Interface for Windows Operating System Using Kinect

İsmail Öztel, Hüseyin Eski, Cemil Öz
Faculty of Computer and Information Sciences, Department of Computer Engineering, Sakarya University, Turkey

Abstract

Computers have become an integral part of life. Many projects on computers have been developed to make life easier. Human beings physically connect to the world through their limbs, especially their hands. They perform most of their everyday tasks with them; however, along with their hands, they also rely on devices such as mice, keyboards or joysticks to work with computers and computer applications. Virtual reality input devices like data glove, motion tracker and Kinect could overcome the limitations of these devices. In this study, we developed a Natural User Interface for Windows operating system using Kinect. In this way, a person can use a computer with gestures. Kinect Sensor was used to monitor users' gestures. This application can also be used in many other areas including education, health, sport, etc.

Key words: Virtual Reality, Kinect for Windows, Gestures, Natural User Interface

1. Introduction

In parallel to the advanced technology, natural user interfaces using inputs like sound, human gestures are developed for computer systems to replace input devices such as keyboards and mice [1]. The use of this type of input devices to operate computer systems is highlighted by its ergonomic benefits in gaming, graphic design as well as in fulfilling control and command functions.

Today, human-machine interaction is one of the popular study areas. Virtual reality, augmented reality, simulations and gaming technologies use various devices including NUIs (Natural User Interface), data gloves, haptic device, flock of bird, etc. instead of fixed devices limiting the user's motions. Most of the studies focusing on this area are often based on the analysis of human motions.

Machine-human interaction has become an integral part of our era, and many scientific and business studies have been made to make this interaction more meaningful and natural. Significant part of these studies focus on the analysis of human motions. Natural motions of human are tracked by sensors and systems are controlled via application software directives.

Analysis of human motions is crucial in various areas including video games, virtual reality applications, scientific visualizations and motion based control systems. These analyses require

*Corresponding author: Address: Faculty of Computer and Information Sciences, Department of Computer Engineering Sakarya University, 54187, Sakarya TURKEY. **E-mail address: ioztel@sakarya.edu.tr, Phone: +902642953240**

sensor based systems to detect 3D human motions [2]. Kinect device is one of these sensors. Kinect became very popular after its introduction for gaming purposes in Xbox gaming console. Following Microsoft's publication of software development kit (SDK) [3], Kinect went beyond its intended use and started to be included in industrial and academic studies with the mission of making the life of humans easier.

We can discuss studies to detect real-time human motions under 3 headings. First heading is the computer based vision. The second heading includes motion detection using special tools like data gloves, and the final heading covers studies made on both [4]. Among these methods, motion detection using special tools (e.g. electromagnetic sensor systems) may be limited as the tools have to be placed on the body which results in disrupted naturalness [5].

The recent developments in the sensor technologies have brought a new dimension to the human-computer interaction. These developments enabled functioning in negative conditions such as environments without illumination or inadequate background [5]. Gestures are used often in the process of interaction with systems. Studies conducted in this area have provided solutions to different sub-problems, enabling system control. Examples of system-human interaction include robot control, computer games, software controls, electronic device controls, sign language detection, etc. [6].

Today, the natural user interface concept is becoming more and more widespread in system-user interactions. Particularly, the studies on this subject are further advanced with the use of Kinect device and other similar sensors [1].

The literature features many Kinect supported studies, mostly on natural user interface. In a study conducted in 2013 by Andrea Sanna et al., a quadrotor was controlled by detecting user motions with Kinect and analyzing them in the software developed [7]. Han JiaQi et al. presented a study in 2013 in which they detected user motions via Kinect. They used these motions to control a mini vehicle they developed. They also enabled users to achieve steering motion by creating a graphic interface [8]. Unseok Lee and Jiro Tanaka made a study in 2013 and detected user hand motions supported by Kinect and the shapes drawn in the free space were transferred on a monitor by tracing finger in 3D space. They also changed and rotated dimensions on pictures and videos in the free space [5]. I-Tsun Chiang et al. conducted a study in 2012 which intended to develop the visual skill performance of elderly living in a wheelchair. Elderly users who live in a wheelchair stood facing the monitor and interacted with it via Kinect device. Reaction times and hand-eye coordination of the users were evaluated to conclude that their visual skill performance were developed and improved through the system [9]. Binnur Gören developed a robot to help elderly in their daily exercise program in her thesis study (2011). The system used Kinect for the analysis of human motion. Robot learned how motions should be performed based on these analyses and showed these motions to elderly people asking them repeat the same. It provided feedback on the performance of the motions made [10]. In his master's thesis (2012) Ahmet Ali Süzen developed a house automation system for disabled individuals to deal with stuff at their homes in a safe and practical manner. The system developed opens and closes objects in the house such as doors, television and lights with Kinect device detecting, and analyzing in the system, the user's motions [11]. In a study (2012) by Jing

Tong et al. multiple Kinect devices were used to model 3D human body in virtual environment. The multistage project included an initial modeling attempt with a single Kinect followed by a test with two Kinect devices until finally the model that was observed to demonstrate the highest performance was delivered using three Kinect devices [12].

2. System hardware

Special wearable systems, data gloves, electromagnetic sensors used to detect human motions limit movement and result in discomfort for the user. Vision based techniques were developed to overcome these handicaps. Vision based techniques do not require contact with a device; several sensors and motion identifying algorithms are sufficient in order for motions to be detected. Therefore, this technique offers a natural and impressive way of using these systems [7]. Recently, Kinect device which are very popular both in the industry and academy has been proved to work based on vision.

Communications with the Kinect device is performed by Kinect Software Development Kit (SDK). With the declaration of Microsoft to support SDK in the long term and periodically, it can be suggested that Kinect will maintain its popularity [1].

Other reasons for preference include the price of Kinect at around $100 and its relative affordability compared to other vision based systems as well as its capability to deliver large scale projects [2].

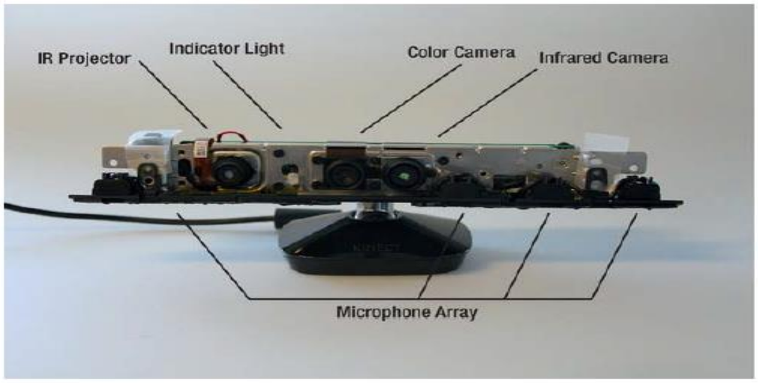

Figure 1. Kinect [13]

Kinect device connects to computer via USB connection. Data from USD is evaluated by application software and outputs are obtained. Kinect also requires a power supply and uses an electrical cable for supply as an alternative to USB.

With Kinect systems establish communications by simply detecting the users' motions and sounds without requiring any other manually-controlled device. It performs these actions with the built-in cameras, infrared heat emitter and microphones (Figure 1). Kinect's algorithm to detect human motions is shown in Figure 2.

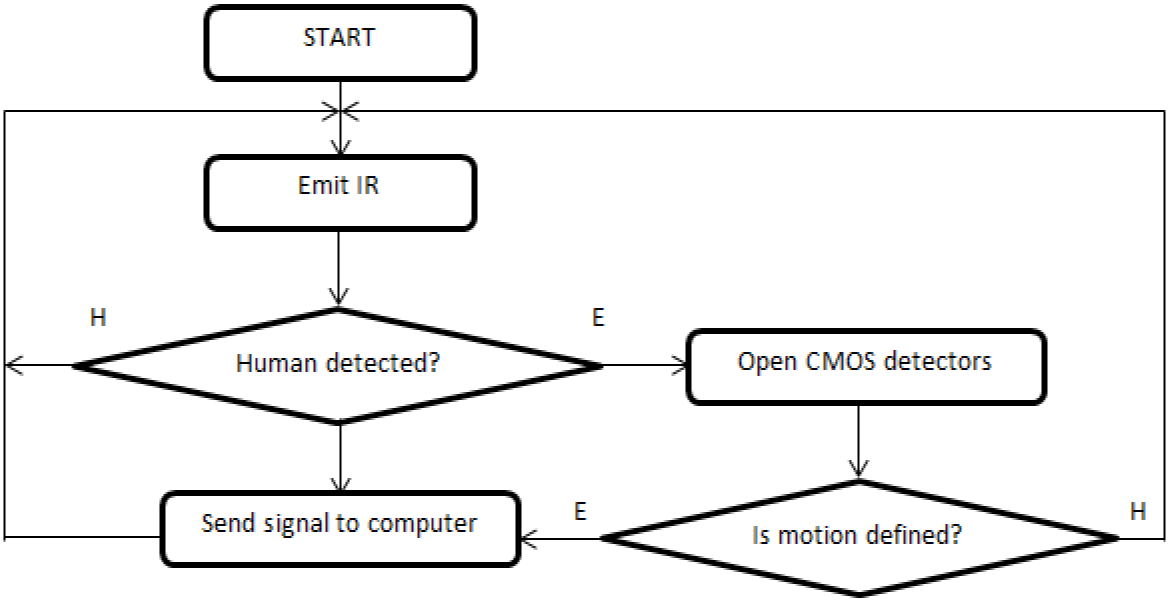

Figure 2. Kinect motion detection algorithm [11]

Kinect device can see an area of 53 degrees horizontally and 43 degrees vertically. Depth sensor works effectively within 1.2m - 3.5m. Kinect can detect up to 6 persons at the same time while it can only track the motions of 2 people [11]. The device can create a skeleton model of a person by tracing 20 joint points of the human (Table 1)

Table 1. Kinect skeleton joints

1.	Head	**6.**	Right wrist	**11.**	Right knee	**16.**	Right foot
2.	Left shoulder	**7.**	Right hand	**12.**	Left foot	**17.**	Left hip
3.	Right shoulder	**8.**	Left hand	**13.**	Left ankle	**18.**	Right hip
4.	Left elbow	**9.**	Left knee	**14.**	Right ankle	**19.**	Center hip
5.	Left wrist	**10.**	Right elbow	**15.**	Neck	**20.**	Chest

2.1. The application

This study developed a natural user interface for the use of operation systems and applications. Study equipment included Kinect device; Microsoft Kinect SDK for the programming of Kinect; and Windows Graphic API for the programming of operating system control.

X, y, z coordinate plane positions of 20 joint joints of the user obtained from the Kinect sensor and the depth image data were used to classify the user's motions. In the application program, data from Kinect were used to obtain information about the position of the person. In the application, user's gesture position was taken as reference for each Windows control function. In order to set a standard, Kinect executes monitoring within a certain area as required by the algorithm (Figure 4). This area is proportional to the monitor area. If suitable positions are captured after user data from the device of the user are matched with identified positions, the relevant action is realized on the operating system. Table 2 provides the operating system functions and the corresponding position the user must take together with details of the position.

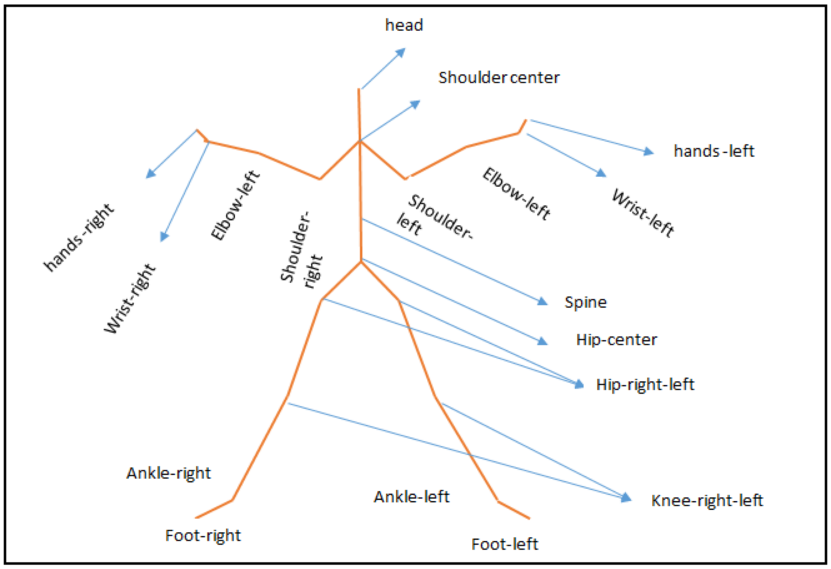

Figure 3. Kinect skeleton model

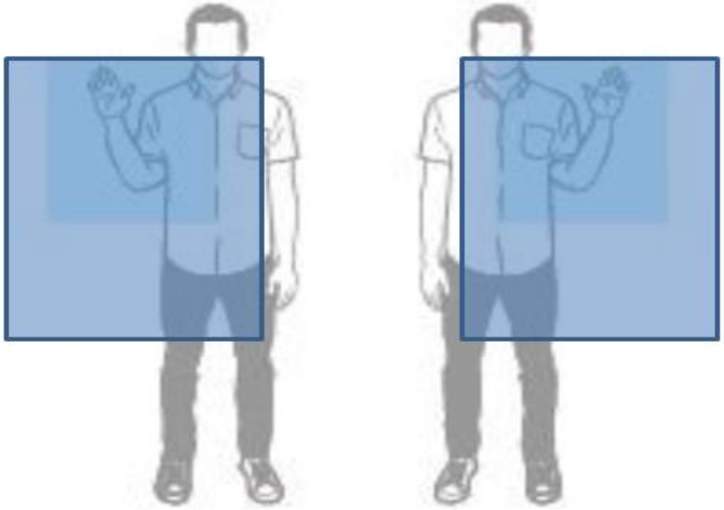

Figure 4. User movement area

Table 2. Identified motions and their descriptions

FUNCTION	MOTION	DESCRIPTION
Activate mouse control	Extend right hand straight forward	Distance between the right hand and right shoulder must be above 85% of the distance between shoulder and elbow, and distance between shoulder and hand must be below 1.8 times the distance between shoulder and elbow.
Deactivate mouse control	Pull right hand towards shoulder	Control is released when distance between right hand and right shoulder is below 85% of distance between shoulder and elbow.
Single click	Raise left hand above the shoulder	Distance between left hand and left shoulder when left hand is above left shoulder must not exceed the distance between shoulder and elbow.
Double click	Extend left hand straight forward	Distance between left hand and left shoulder will active when it is above 66% of the distance between shoulder and elbow.
Right click	Extend left hand sideways	Distance between left hand and left shoulder will activate when it is above the 1.2 times the distance between shoulder and elbow.
Scroll function	Extend right hand forward and move downwards or upwards	Distance between right hand and right shoulder will activate when it is above the 1.8 times the distance between shoulder and elbow.
Close active program (ALT+F4)	Extend left hand upwards above the shoulder when mouse control is deactivated	Left hand must not exceed the distance between shoulder and elbow when left hand is above left shoulder.
Shift between active windows (WINDOWS + TAB)	Extend left hand sideways when mouse control is deactivated	Distance between left hand and left shoulder will activate when it is above the 1.2 times the distance between shoulder and elbow.

For example, the user stands facing Kinect to take over the Mouse control and extends his right hand straight forward as seen in Figure 5. The distance of right hand between shoulder and elbow was taken as reference for the user to take over the control. When the distance between right hand

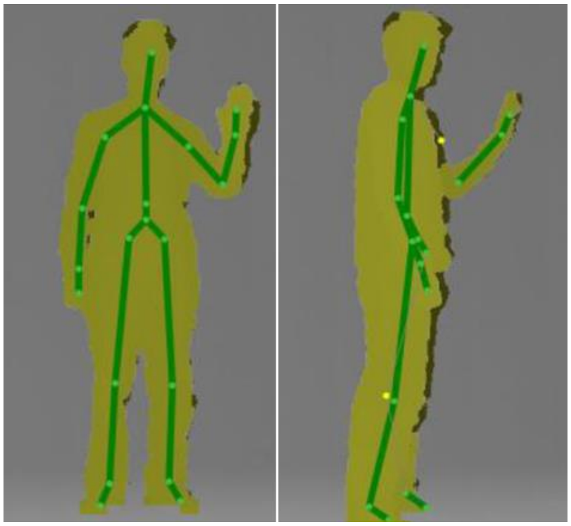

Figure 5. The depth map showing the moment when user takes over the mouse control

and shoulder reaches or exceeds this reference point, the control is taken over by the user. As seen in the table, distances identified for functions are not based on a fixed length. The reason for this is to ensure users with different body sizes to use the system without experiencing any problems.

Conclusions

This study delivered a natural user interface which uses human gestures and body motions. The user interface software was developed using object oriented programming technology and C# programming language with Microsoft Kinect SDK and Microsoft Graphic API (gdi32.dll) libraries. X, y, z position data for human joints obtained from the Kinect sensor and the depth image detected with a depth camera were filtered and classified through a rule-based algorithm. The human motions classified were matched with the basic input devices mouse and keyboard control keys.

The system developed works successfully with Windows operating system and programs. However, the performance of the program varies depending on the device where the application works. It may take 3-4 seconds for the program to start on a computer with average features. As the gestures of the user can only be monitored within the area identified in the algorithm, the motion will not be monitored out of the borders of this area and the application may experience perceivable hesitancy at the application during returns to the area.

Acknowledgements

This work was supported by Sakarya University Scientific Research Foundation (Project number: 2012-12-10-0002)

References

[1] Schmitt C, Zierold M, Pönisch F. Development and Usage of a NUI Framework. The 3rd IEEE International Conference on Consumer Electronics. 2013.
[2] Oz C, Leu M C. Human Motion Analysis with Artificial Neural Network Using the Kinect Sensor. 3rd International Conference on Science Engineering and Technology. 2012.
[3] Kıvrak H. İşaret Dili Hareketlerinin Bir İnsansı Robot Tarafından Taklit Yoluyla Gerçeklenmesi. M.S. thesis. İstanbul Teknik Üniversitesi.2013.
[4] Oz C, Leu M C. American Sign Language Word Recognition with a Sensory Glove Using Artificial Neural Networks. Engineering Applications of Artificial Intelligence. 2011.
[5] Lee U, Tanaka J. Finger controller: Natural User İnteraction Using Finger Gestures. Human-Computer Interaction. 2013.
[6] Memiş A. Kinect RGB Görüntülerinde ve Derinlik Haritalarında Uzam-Zamansal Özellikleri Kullanarak İşaret Dili Tanıma. M.S. thesis. Yıldız Teknik Üniveritesi.2013.

[7] Sanna A, Lamberti F, Paravati G, Manuri F. A Kinect-Based Natural Interface for Quadrator Control. Entertainment Computing. 2013.
[8] JiaQi h, Santoso M, Gook L B. Immersive Driving Car Simulation for Children Using Natural User Interface Controller. 2013 International Symposium on Ubiquitous Virtual Reality. 2013.
[9] Chaing I, Isai J, Chen S. Using Xbox 360 Kinect Games on Enhancing Visual Performance Skills on Institutionalized Older Adults with Wheelchairs. Fourth IEEE International Conference On Digital Game and Intelligent Toy Enhanced Learning. 2013.
[10] Görer B. Developing a Fitness Coach Robot for Elderly People in Assisted Living Environments. M.S. thesis. Boğaziçi Üniversitesi. 2011.
[11] Süzen A A. Kinect Teknolojisi Kullanılarak Engelliler için Ev Otomasyonu. M.S: thesis. Süleyman Demirl Üniversitesi. 2012
[12] Tong J, Zhou J, Liu L, Pan Z, Yan H. Scanning 2D Full Human Bodies Using Kinects. IEEE Transactions on Visualization and Computer Graphics. 2012.
[13] Webb J, Ashley J. Beginning Kinect Programming with the Microsoft Kinect SDK. APRESS. 2012