

Gesture-Based Robot Control with Kinect Sensor

Bhatt Meet, Joshi Hari, Vaghasiya Denil, Rohit R Parmar, Pradeep M Shah

Abstract: With the recent improvement in technology, controlling a robot wirelessly has become possible. Controlling a robot wirelessly increases its mobility. It also opens up the possibility of different kinds of robot designs. However, controlling a robot wirelessly using human computer interface bring more challenges to the researchers. This paper implements a human gesture-based robot control system. It uses a Kinect sensor which consists of a depth sensor, RGB camera. The Kinect sensor obtains the skeletal data from the subject. Using mathematical calculation, it calculates the angle deviation between the different parts of the arm based on its movement. The angle deviation data measured by the Kinect sensor will be transmitted by using a Wi-Fi module. The use of the Wi-Fi module will increase mobility. On the receiver, we use a microcontroller which will convert the received angle deviations values into equivalent PWM which will be used to control the servo motors. The combination of the different servo motors will result in the robotic movement

Index Terms – Kinect, Skeleton tracking, Raspberry Pi 3B, Servo motors, hardware, Human Computer Interface

I. INTRODUCTION

This paper implements a human gesture-based robot control system. It uses a Kinect sensor which consists of a depth sensor, RGB camera. The Kinect sensor obtains the skeletal data from the subject. Using mathematical calculation, it calculates the angle deviation between the different parts of the arm based on its movement. The receiver will receive the data and convert it to its equivalent PWM signal. The PWM signal will be used to control the servo motors present in the robot. The combination of the different servo motors will result in the robotic movement.

II. LITERATURE REVIEW

The classification of skeleton tracking algorithms can be done in several ways. First, we look at the pathways that follow the trajectory by using one-dimensional design methods [1, 2].

Subsequently, Shotton et al. [1] have suggested a new method with the help of which one can predict the 3D positions of physical interaction from a depth image and without temporal details and kinematic constraints.

Revised Manuscript Received on May 20, 2020.

Bhatt Meet, UG Student, Department of Electronics and Communication, G. H. Patel College of Engineering & Technology, Vallabh-Vidyanagar, Gujarat, India.

Joshi Hari, UG Student, Department of Electronics and Communication, G. H. Patel College of Engineering & Technology, Vallabh-Vidyanagar, Gujarat, India.

Vaghasiya Denil, UG Student, Department of Electronics and Communication, G. H. Patel College of Engineering & Technology, Vallabh-Vidyanagar, Gujarat, India.

Rohit R Parmar, Assistant Professor, Department of Electronics and Communication, G. H. Patel College of Engineering & Technology, Vallabh-Vidyanagar, Gujarat, India.

Pradeep M Shah, Assistant Professor, Department of Electronics and Communication, G. H. Patel College of Engineering & Technology, Vallabh-Vidyanagar, Gujarat, India.

That way, the average presentation of the organs is designed to obtain a problem mapping problem for each pixel separation. Finally, they have a 3D demonstration of the confidence of many body parts produced by incorporating the effects into local mode. As a result, it is able to predict the 3D positions of the body structure with high accuracy. Based on this algorithm they used skeleton trackers in this type of the Kinect SDK.

Zhang et al. [5] use images of individual depth in a composite point cloud and use the fine-grained particle filtering method and with their help make available estimates. Similarly, Liu et al. [6] presented a countless video navigation system with multiple views that reconstruct the skeletal movements and surface geometries of two highly interrelating data. The technique presented in this paper is based on the approximation posture of human skeleton which is shown by Kinect SDK 2.0 and varies from the methods used by the different studies.

Specifically, our goal was not to develop a process that calculates 3D body mass positions directly from raw depth images or RGB images, but instead to investigate how to generate real and accurate full-body computations in real-time by combining original-estimated multiple 3D bodies.

Brogardh in his work [4] discusses and introduces technical challenges that robot participants encounter. He concluded that model-controlled control is now a major technology for the control of robots and models. Designing industrial robots for a specific application is still very costly, costly and time-consuming. In such a situation, SMEs (small and medium enterprises) cannot benefit from robotic automation due to this time-consuming planning. [7] Today, in practical industrial applications, there are two main classes of robotic systems: online and offline.

In an online program, an instructional pendant is used to manually move the wrist or joints in the intended position as well as become familiar with each stage of the robot work. Appropriate positions and guidance are sent to the robotic control system and the robot's program is written. [4, 7]. In Paper [2] the authors provide a full review of recent studies on the methods of designing industrial robots, including the online system, the offline system, and the program using Augmented Reality (AR). [5, 7]

Human motion analysis using depth imagery was concerned in many scientific publications e.g. [5]. The topic was considered in terms of analysis of human behavior through visual information. Previously, analysis of human behavior was attained via images from a conventional camera, however, in recent times depth sensors have made a new kind of data available.

III. PROPOSED GESTURE RECOGNITION SYSTEM MODEL

A. Transmitter Side

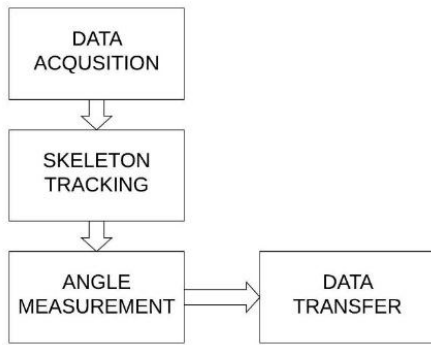


Figure 1: Flowchart for gesture recognition (Transmitter)

B. Data Acquisition

In the Data acquisition process, the input is captured in the form of an image or a video. The user has to stand in front of the device that is used to capture the input. Here Microsoft Kinect is used as an input capturing device. The Kinect can capture the RGB video stream, IR stream.

C. Skeleton Tracking

Microsoft Kinect sensor can sense 20 joints at a given time. The sensor returns the location of the joints of the user. Based on the coordinates of the joints a vector is drawn between two joints which represent the bone of a skeleton. The combination of the joints and vectors is represented as the skeleton of the user.

D. Angle Measurement

Once we have the skeleton of the user, we can use basic trigonometric functions to calculate the angle between the vectors. We measure 2 different angles for each arm like the angle between head-shoulder-elbow and angle between shoulder-elbow-wrist.

E. Data Transfer

Once we have the angle deviation values, we transfer the angle deviation on a webserver. The web server allows us to increase the mobility of the robots by eliminating the connection of the robot to the computer. On the receiver side, we use the data uploaded on the webserver.

F. Receiver Side

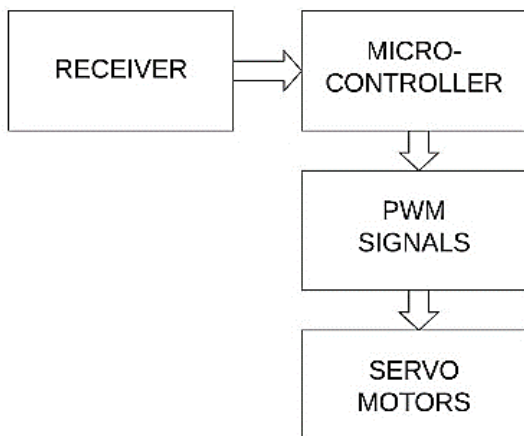


Figure 2: Flowchart for gesture recognition (Receiver)

G. Receiver

The receiver will fetch the data from the web-server. The receiver is linked to the robot’s control system.

H. Microcontroller

The microcontroller used here is Raspberry PI 3b. The microcontroller is used here to control the servo motors present in the robot. The microcontroller converts the angle deviation values into its equivalent PWM signals.

I. PWM Signals

The PWM signals are Pulse Width Modulated Signals. They are analog signals generated from a digital source. PWM signals then control the servo motors.

J. Servo Motors

Servo motors are the basic component which is used for movement in the robots. The servo motors consist of mainly 3 pins like +5V, Ground and PWM pin. The PWM pins control the rotation of the arm of the motor. Servo motor modules are used in robots which allows the user to control more than one servo motor by reducing the excess wiring. A servo motor module can control at least 6 motors.

IV. HARDWARE DESCRIPTION

A. Kinect



Figure 3: Xbox 360 Sensor

Kinect is a motion-sensing device produced by Microsoft. The Kinect was mainly used as an add-on to the existing gaming consoles. It consists of an RGB camera, a depth sensor and an array of microphones.

B. NodeMCU (ESP8266)



Figure 4: NodeMCU with ESP8266 Wi-Fi chip

NodeMCU is an open-source firmware and a development board for the ESP8266 Wi-Fi chip. ESP8266 is a low-cost Wi-Fi chip developed by Espressif Systems with TCP/IP protocol. It can be designed to connect to the Internet for the Internet of Things (IoT) and similar projects.

C. Raspberry Pi

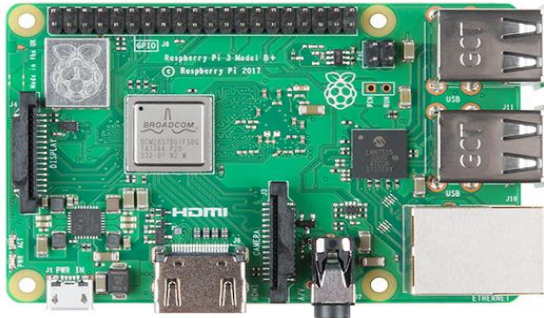


Figure 5: Raspberry Pi 3 board

The Raspberry Pi is used on the receiver end. Raspberry Pi 3 is an ARM-based small-sized computer that can be plugged into a computer monitor or a TV. It is used to convert the received angle deviation values into its equivalent PWM signals which are used to control the servo motor.

V. SOFTWARE DESCRIPTION

In Data acquisition the Kinect sensor uses the depth sensor and the RGB camera to capture the user.

In the angle measurement technique, we calculate the distance between the two joints. Based on the required angle, we need to select the specific parts of the body. The Kinect sensor keeps tracking the position of the defined joints. Based on the change in the position of the joints it measures the angle between those joints. The angle its measures will be in radians which will be converted into degrees. The different joint angles measured will be used to control the different servo motors present in the robot.

The measured angle values are then transmitted to the receiver using a Wi-Fi module. The Arduino IDE is a cross-platform application that is written in functions from C and C++. The ESP8266 module is programmed using this IDE. The measured angle values from the Kinect sensor uploaded on a web server using the Wi-Fi module.

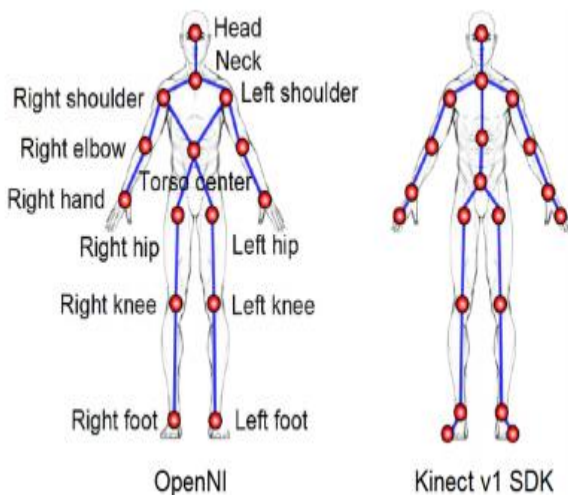
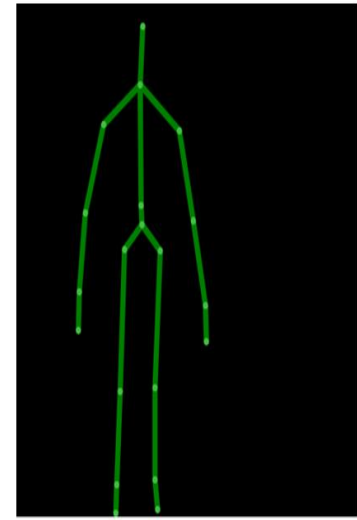


Figure 6: Different joint locations in a body

VI. RESULT & DISCUSSION

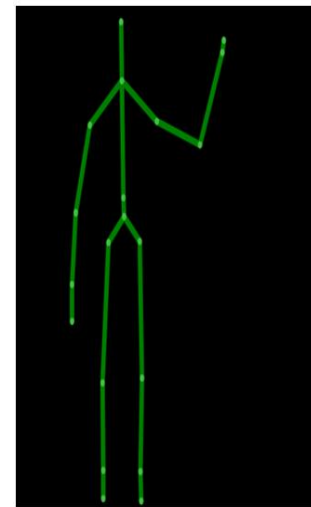
We are designing a control system for the given robot which allows the user to control the robot using his actions. The software implementation of the project consists of capturing the skeleton data stream using the Kinect Sensor. From the skeleton stream, we find out the angle required angles using

the trigonometric functions. Then the acquired output angle deviation values are uploaded on a webserver created using an ESP8266 Wi-Fi module. The receiver side fetches the data from the webserver and converts it into its equivalent PWM signal which controls the robot. The following figure below shows us the skeleton data stream and the measured angle for the various position of the user.



8.444 150.673 137.2

Figure 7: Skeleton frame when the user in the standing position



36.9 145.067 112.469

Figure 8: Skeleton frame when the user is lifting his right hand

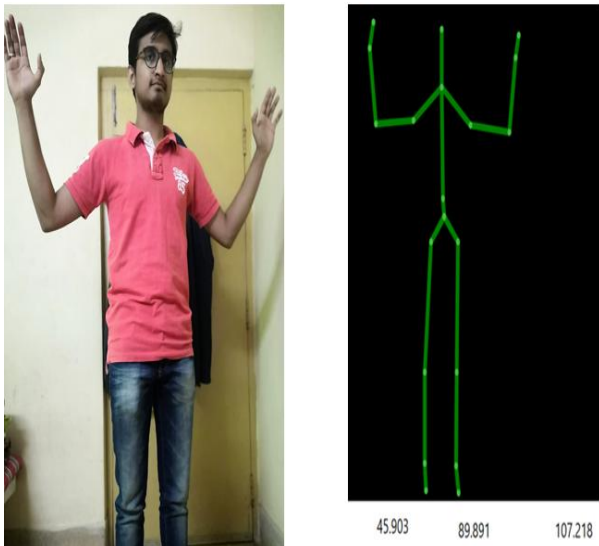


Figure 9: Skeleton frame when the user is lifting his both hands

VII. CONCLUSION

The software implementation of the project consists of capturing the skeleton data stream using the Kinect Sensor. From the skeleton stream, we find out the angle required angles using the trigonometric functions. Then the acquired output angle deviation values are uploaded on a webservice created using an ESP8266 Wi-Fi module.

REFERENCES

1. J. Shotton, A. Fitzgibbon, M. Cook, T. Sharp, M. Finocchio, R. Moore, A. Kipman, and A. Blake, "Real-time human pose recognition in parts from single depth images," International Conference on Computer Vision and Pattern Recognition (CVPR), 2011.
2. I. Mikic, M. M. Trivedi, E. Hunter, and P. Cosman, "Human body model acquisition and tracking using voxel data," *Int. J. Comput. Vis.*, vol. 53, no. 3, pp. 199-223, 2003.
3. F. Caillette, A. Galata, and T. Howard, "Real-time 3- D human body tracking using learned models of behavior," *Comput. Vis. Image Understand.*, vol. 109, pp. 112-125, 2008.
4. Brogardh, T.: Present and future robot control development - An industrial perspective. *Annual Reviews in Control* 31, 69-79 (2007)
5. Chen, L., Wei, H., Ferryman, J.: A survey of human motion analysis using depth imagery. *Pattern Recognition Letters* 34, 1995-2006 (2013)
6. Dutta, T.: Evaluation of the Kinect sensor for 3-D kinematic measurement in the workplace. *Applied Ergonomics* 43, 645-649 (2012)
7. Pan, Z., Polden, J., Larkin, N., VanDuin, S., Norrish, J.: Recent progress on programming methods for industrial robots. *Robotics and Computer-Integrated Manufacturing* 28, 87-94 (2012)
8. Madakam, S., Ramaswamy, R. and Tripathi, S. (2015) Internet of Things (IoT): A Literature Review. *Journal of Computer and Communications*, 3, 164-173
9. M. Van den Bergh, D. Carton, R. de Nijs, N. Mitsou, C. Landsiedel, K. Kuehnlenz, D. Wollherr, L. Van Gool, and M. Buss. Real-time 3D hand gesture interaction with a robot for understanding directions from humans. In *Proc. of 20th IEEE international symposium on robot and human interactive communication*, pages 357-362, 2011.
10. Biao Ma, Wensheng Xu, and Songlin Wang. A robot control system based on gesture recognition using Kinect. *TELKOMNIKA Indonesian Journal of Electrical Engineering*, 11(5):2605-2611, May 2013.
11. Ali, Ahmed & Elmisery, Fathy & Mostafa, Ramadan & Hussein, Mohammed. (2014). Motion Control of Robot by using Kinect Sensor. *Research Journal of Applied Sciences, Engineering and Technology*. 8. 1384-1388. 10.19026/rjaset.8.1111.
12. Kun Qian, JieNiu, and Hong Yang. Developing a gesture based remote human-robot interaction system using kinect. *International Journal of Smart Home*, 7(4), July 2013.

13. K. Lai, J. Konrad, and P. Ishwar. A gesture-driven computer interface using kinect. In *IEEE Southwest Symposium on Image Analysis and Interpretation (SSIAI)*, pages 185-188, 2012.

AUTHORS PROFILE



Meet Bhatt is a final year student pursuing Bachelor of Engineering in Electronics and communication field from G. H. Patel College of Engineering & Technology, Vallabh-Vidyanagar, Gujarat. He is planning to pursue Masters Degree abroad



Hari Joshi is a final year student pursuing Bachelor of Engineering in Electronics and communication field from G. H. Patel College of Engineering & Technology, Vallabh-Vidyanagar, Gujarat. At present, he claims the IT sector job at Capgemini, Pune



Denil Vaghasiya is a final year student pursuing Bachelor of Engineering in Electronics and communication field from G. H. Patel College of Engineering & Technology, Vallabh-Vidyanagar, Gujarat. He is presently preparing for Entrance exam GATE.



Prof. Rohit Parmar has post-graduated in Communication Engineering from Gujarat Technological University. He has more than eight years of academic experience in teaching at undergraduate level. He has taught different UG courses like Audio Video Systems, Communication Engineering, Basic Electronics, Digital Electronics, Electronics Measurement and Instrumentation, etc. His research interest lies in the domain of Biomedical Signal Processing and Image processing. He has guided 8 UG projects and 4 PG dissertations.



Prof. Pradeep M. Shah has completed his M.E. in Electronics & Communication from Dharmsinh Desai University in 2004. Presently, he is pursuing his Ph.D. from Shri Govindram Seksaria Institute of Technology and Science, Indore. He has 13.5 year of teaching experience at under graduate and post graduate level. His area of interests are wireless communication and digital signal processing.