Home Appliance Control System using Hand Gesture with ML

Mr. D. A. Birari[1], Pratik Jadhav[2], Priyanka Pagar[3], Harshada Pachpute[4], Mahendra Pawar[5]

[2,3,4,5]Student, Dept. of Information Technology, KBT College of Engineering, Maharashtra, India
[1]Assistant Professor, Dept. of Information Technology, KBT College of Engineering, Maharashtra, India

---

Abstract - *Modern science and technology is all about making our lives more comfortable. As the study of various technologies is progressing, the definition of comfort in our lifestyle is evolving. Tasks that involved so much complexities and hard work a few years ago can now be done by just pushing a button. But by being possessed with thoughts of our own comfort, we often forget the people who need it the most; the physically challenged people. They face so many limitations in everyday life that what may seem luxury to us is ironically necessity to them. To provide a satisfactory resolution to this, we have built smart Home Appliance Control System for Physically Disabled People Using Kinetic. Our systems goal is to make their lives easy and comfortable by providing them with a self-dependable environment. To validate the systems acceptability we developed a prototype in our laboratory environment and performed real user study.*

Key Words: **Smart Home, Hand Gestures, Body Movement, Machine Learning, Kinect Sensor**

1. INTRODUCTION

Science and Technology has become an essential part of our lives. Computer Science is now reaching every field we can imagine. It is now an integral part of human life serving in various ways. It provides us with better alternatives and has simple solutions to complex problems. Now days Human Computer Interaction (HCI) is one of the major concerns in every aspects of our technology. As technologies are developing, the ways of interaction between Human and Computer are also evolving. Technology has taken us to the stars, made great depths of sea accessible and filled our lives with uncountable wonderful devices to provide us with smart and new ways of doing things. Yet, for some people even a normal daily life is still a luxury. People with physical disabilities still have to rely on other people's help to even do the simple things like turning a light on. Existing smart home frameworks are usually built for normal people in mind and are quite expensive.

This system deals with identifying the hand gestures by using Kinect sensor and using these actions system will control various gadgets around us. The main motive is to create more interaction with human and machine in a natural, intuitive and seamless manner. To build this system there is need of different types of algorithm like Skeleton-tracking, Normalization and 3D Euclidean distance algorithm. Researchers have successfully implemented this system in various fields with high accuracy and impact using a highly feasible system. But our system is based on Kinect sensor and more useful to physically challenged peoples.

It includes the basic overview of the proposed system which gives the short description about the seminar. Introduction Today world is a global hub due to advancements in technology Inventions and evolution in technology has made this possible. Home automation has an important role in people's life when it comes to their standard of living as it provides convenient and hassle free environment. Human-machine interaction technology defines the way human use machines and it plays crucial role for developing an ecosystem between user and machine. Speech based Human machine interaction has potential to be future of Human-machine interaction and recent advancement in technology have made it possible to operate computers and electronic devices through speech. While there are lots of research going on for the development of better speech recognition systems but very less efforts are made to develop speech based Human-machine interaction system for everyday use Primary purpose of any Machine or device is to provide some functionality for productivity, comfort and convince of user.

Technology aims to solve many of the problems people have in their everyday lives. For people with disabilities, a home automation system can be an invaluable tool for making simple tasks easier.

1.1 Related Works

There are many previous works related to this system. Some of them are mentioned here such as, in year 2011, the authors Piyare, Tazil et al. [3] proposed a system named as "Bluetooth Based Home Automation System Using Cellphone". This system used Bluetooth for communication and implemented using various Bluetooth protocols. It was applicable for short range distances. In year 2009, the authors Gill, S.H, Yao et al. [2] proposed a system named as "A Zigbee Based Home Automation System". In this system the authors used new communication technology like Zigbee. Zigbee is nothing but a wireless standard network used for Remote Control Products. In year 2012, the author Guesgen et al. [6] proposed a system named as Gestural Control of Household Appliances for the Physically Impaired. This system proposed that contribution in Gestural control of household appliances for the physically impaired. In year 2013, the authors Asma Ben Hadj Mohamed and Juan A. Holgado-Terriza [13] proposed a system named as "Assisting

people with disabilities through Kinect sensors into a smart house". In this paper, we propose a monitoring system based on Kinect sensors to control and monitor elderly people into a smart house. The system recognizes gestures and communicates them through a network. We test some hand gestures and how it can be recognized with this sensor. In year 2014, the author Y. Kim et al. [7] proposed a system named as Non-Contact Gesture Recognition Using the Electric Field Disturbance for Smart Device Application. This system proposed that contribution in to the Total Antioxidant Capacity. In year 2015, the authors Amrutha et al. [8] proposed a system named as "Voice Controlled Smart Home", which uses MATLAB programming solution for voice recognition part in their system and the system needs to be trained for the speech pattern of each user in order to maintain its high recognition accuracy. In year 2016, the authors Juan J. Ojeda-Castelo Jose, A. Piedra-Fernandez and Cesar Bernal-Bravo [16] proposed a system named as "Sign Communication for People with Disabilities Using Kinect Technology at Home". In this paper authors present a communicator which interprets a series of commands by means of corporal expressions. These body expressions are learned by a gesture recognition system according to the requirements and disability of the user. Each of the commands adapts themselves to daily tasks. The system learns gestures and associates them with concrete actions that the user wants to do or needs at that moment.
In year 2018, the authors Huang-Chia Shih and Chang-Hsian Ma [21] proposed a system named as "Hand Gesture Recognition Using Color-Depth Association for Smart Home". This system proposed that a robust hand gesture segmentation method which associates the depth and color information with online training. Different existing methods, when the hands close to the body part or in cluttered background, our system remains valid.

1.2 Problem Definition

Home Automation is nothing but to control home appliances using smart technologies. This system aims to enhance the home automation experience by collecting usage data from the user and applying prediction algorithms on it to predict the next step the user may take. Machine learning is a data analytics technique that teaches computers to do what comes naturally to humans: learn from experience. Hand gestures are one of the natural ways to understand the human language. To provide satisfactory results we are going to introduce a system "Smart Home Appliance Control System for Blind People Using Machine Learning". This system is based on machine learning. Our system's goal is to make blind people's lives easy and comfortable by providing them with a self-dependable environment. This is a monitoring system based on Kinect sensors to control the appliances. This system will be built to control the electric appliances like light and fan. It will follow and use various machine algorithms. The purpose of the system is to help blind people on their own and live without disturbing others. The input given to the system will be a physical action (various gestures). It consists of both hardware and software. It is different than IoT. There will be no use of any smart phone.

2. SYSTEM DESIGN

2.1 System Architecture

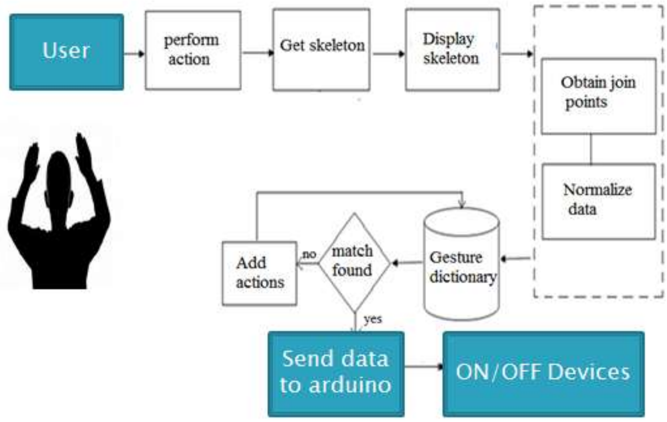

Fig -1: Architecture of Proposed System

The above block diagram shows the architecture of the proposed system. The user is in front of the camera doing a sign or getting ready to do so. A new frame is obtained and the video stream is updated with the skeleton of the user overlapped onto it. Three main blocks are executed: the first block consists of obtaining the data of the joints of interest (JoI) required for the frame descriptor, the second block consists of normalizing these data, and the third one consists of gesture dictionary.

Then, if the working mode is set to TRAINING (meaning that the user is adding a new sign to the training set), the frame descriptor is added to the correspondent file of the dictionary. Otherwise, if the mode is set to TESTING (meaning that the user wants to translate the sign that is been done), the frame descriptor is added to the current test sample. Then, the system checks if the current frame is the last frame of the sign. After a sign is finished and if the working mode is TESTING, the test sign is compared using a classifier with the signs from the dictionary and according to that arduino control the appliances.

2.2 Flowchart

This chart shows the working flow of the system.

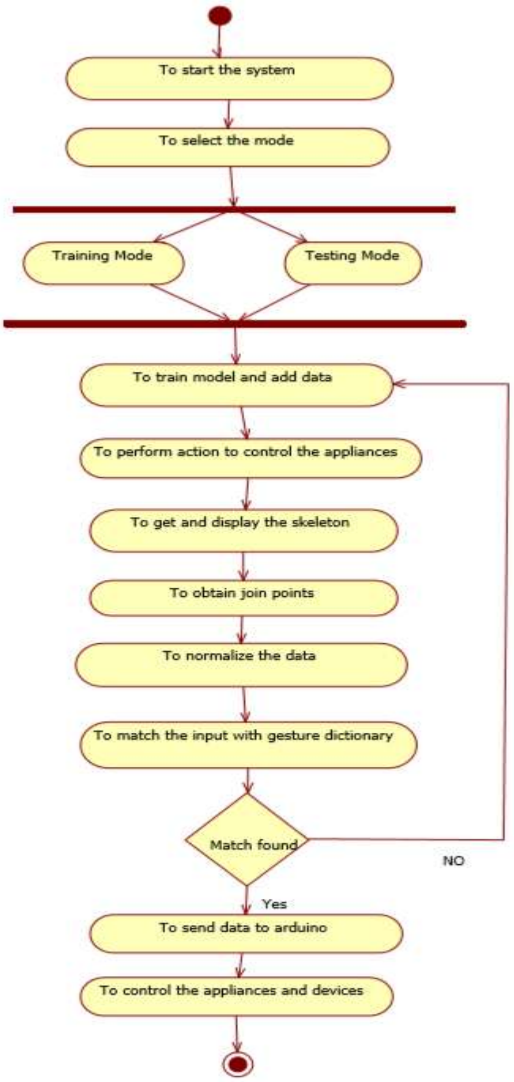

Chart -1: Flowchart

3. KINECT SENSOR

Kinect is the official name of Xbox 360 Console. It is produced by Prime Sense Company accompanying with Microsoft on June, 2010. A Windows version has been released on February, 2012. While in 2009, its announcement caused a great impact in the Computer Vision and Computer Graphics communities. This Microsoft product deciphers a versatile way for interaction in gaming, depending completely on voice and gestures. Gradually Kinect has become a popular device used in gaming, theater performances, robotics, natural interfacing etc.

The word "Kinect" originates from the word "Kinematics" because it senses object skeletal in running condition. Kinect possesses both hardware and software-based technology that deciphers the RGB and depth signals. The hardware portion of this sensor consists of a RGB camera, a three dimensional depth sensor (an infrared camera and a projector), an accelerometer, an array of microphone and a tilt motor which can produce RGB images, depth images, and acceleration forces due to gravity, audio signals and rotation of head position respectively. As far as the software tools are concerned, they are able to catch human motions in three dimensional (3D) spaces.

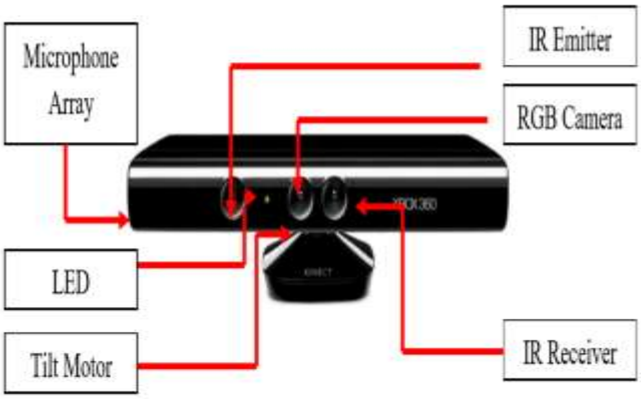

Fig -2: Kinect sensor depicting the hardware parts

3.1 Kinect Hardware Tools:

The Kinect is a black horizontal bar sensor and it looks like a webcam. Figure depicts the arrangement of a Kinect sensor, consisting of a color camera, an infrared (IR) laser projector, an IR camera. The IR projector projects known light speckle pattern into the 3-D images whereas the IR camera captures the reflected IR speckles. The speckle is invisible to the color (RGB) camera but it is viewed by the IR camera.

Each component of the Kinect hardware is described below.

a. RGB Camera: It provides three fundamental color components (Red-Green Blue) of a video. The RGB camera can capture images at a resolution of 640×480 pixels, with a channel rate of 8 bits, operating at 30 Hz. It also has the ability to switch the camera to a higher resolution of 1280×1024 pixels, working at 10 frames per second. The features of RGB camera are automatic white balancing, flicker avoidance, black reference, defect correction and color saturation. The RGB camera has been chosen as the origin of the reference frame. The x direction is along the length of the Kinect sensor, while the y-direction points vertically up and down and the z-direction is the depth measured by the sensor.

b. Depth Sensor: The three dimensional (3D) depth sensor possesses an infrared (IR) emitter and an IR camera. The IR emitter produces a noisy pattern of builded IR light. The projector along with the camera jointly creates a depth image, which allows the distance between the camera and an object. The sensing range of the Kinect device (depth sensor) can be adjusted to a specified limit. The IR camera can seize images at a resolution of 1200×960 pixels, operating at a frequency of 30 Hz. The images are further sampled to 640×480 pixels with a channel rate of 11 bits. Thus it has the sensitivity of 2048 levels.

c. Accelerometer: An electro-mechanical device that can measure acceleration forces. The forces can be of static or dynamic type. The accelerometer detects the path of acceleration forces due to gravity. This incorporates the system fixing its head at an exact level and calibrating the value such that the head can be able to move at specific angles.

d. Microphone Array: The array of microphone consists of four channels. Each channel can process an audio of 16 bit. The sampling rate of microphone is 16 kHz.

e. Motor Section: The Kinect sensor system consists of two sub-systems operated by motor with some gearing. One is to tilt the Kinect sensor up and down up to 27 degrees. The other is the accelerometer, used for determination of the head position of Kinect sensor.

The distance range within which the object should be placed in front of the Kinect sensor is nearly 1.2 to 3.5 m or 3.2 to 11 ft. The depth resolution diminishes, with an increment, at an amount of 1 cm at 2 m distance away from the Kinect.
The sensor system has an angular field of view of 45 degrees vertically, 58 degrees horizontally and 70 degrees diagonally.

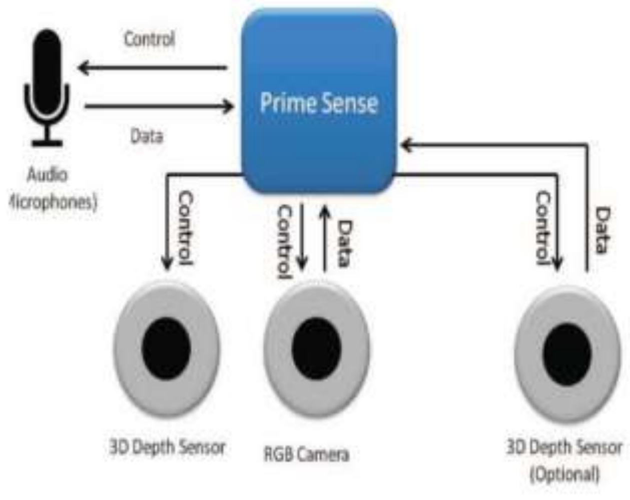

Fig -3: Kinect sensor structure diagram

4. ALGORITHMS

4.1 Skeleton-tracking

Prior to the application of the proposed action recognition approach, the depth maps captured by the kinect sensor are processed by a skeleton-tracking algorithm. The depth maps of the utilized dataset were acquired using the OpenNI API2.

To this end, the OpenNI high-level skeleton-tracking module is also used for detecting the performing subject and tracking a set of joints of his/her body.

More specifically, the OpenNI tracker detects the position of the following set of joints in the 3D space G = {gi , i ∈ [1, I]} ≡ {T orso, Neck, Head, Lef t shoulder, Lef t elbow, Lef t wrist, Right shoulder, Right elbow, Right wrist, Lef t hip, Lef t knee, Lef t foot, Right hip, Right knee, Right foot}. The position of joint gi is implied by vector pi(t) = [x y z] T , where t denotes the frame for which the joint position is located and the origin of the orthogonal XY Z co-ordinate system is placed at the center of the kinect sensor. An indicative example of a captured depth map and the tracked joints is given in Fig. 1.
The OpenNI skeleton-tracking module requires user calibration in order to estimate several body characteristics of the subject. In recent versions of OpenNI, the 'auto-calibration' mode enables user calibration without requiring the subject to undergo any particular calibration pose. Since no calibration pose was captured for the employed dataset, the OpenNI's (v. 1.5.2.23) 'auto-calibration' mode is used in this work. The experimental evaluation showed that the employed skeleton-tracking algorithm is relatively robust for the utilized dataset. In particular, the position of the joints is usually detected accurately, although there are some cases where the tracking is not correct. Characteristic examples of the latter are the inaccurate detection of the joint positions when very sudden and intense movements occur (e.g. arm movements when performing actions like 'punching') or when self-occlusions are present (e.g. occlusion of the knees when extensive body movements are observed during actions like 'golf drive'). [20]

4.2 Normalization Algorithm:

The problem is that the skeleton size grows as the person moves towards the Kinect sensor. We need to build our system in a way to be independent of the position of the person.

The normalizationmust take into account the position of the user. The deaf user can be at different positions of the room and consequently the data must be stored accordingly to that position. As shown in Fig 2, a slight variation in depth can cause a considerable variation of the X and Y values. The distances between one joint and another one can drastically vary depending on the position of the user.

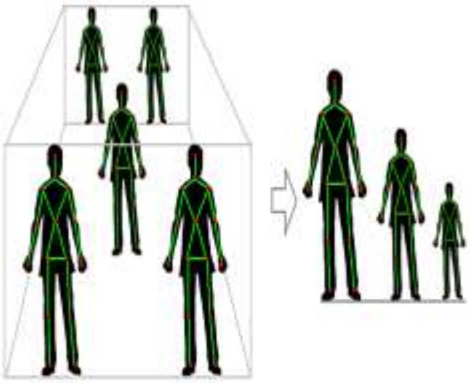

Fig -4: Normalization required for the position of the user.

4.3 Euclidean distance algorithm

According to the natural postures from the real human motion, researchers have proposed the hypothesis that each posture has different physical pattern as Range-based Algorithm. The algorithm concept is defined the relation of body parts and excerpting the human postures form Range between body parts. They represent the ranges between body parts will increase or decrease in y-axis depending on the postures as example of different range between shoulder and Knee in standing, sitting and lie-down postures. In this study, the proposed different of each 3D Euclidean distance between 5 joints pattern and analyzed the change rates of each relation as hypothesis. 3D Euclidean distance between joints of skeleton is:

$$d(p,q) = \sqrt{(p_1 - q_1)^2 + (p_2 - q_2)^2 + (p_3 - q_3)^2} \text{ .... (1)}$$

Where d(p,q) is distant value of the result between p as initial position and q as end position skeletal joints such as head left foot. Whereas p_1, $p_2$and $p_3$are instance of x, y and z axis respectively of initial skeletal position and q_1, q_2 and $q_3$are instance of x, y and z axis respectively of end skeletal position.

5. ADVANTAGES AND APPLICATIONS

5.1 Advantages

- Hand gestures are used to control the home appliances such as fans, lights, TV, etc. The platform used for the recognition of the gesture is the Kinect Sensor.
- The future advancement will be based on the Machine Learning basis; we can control the home appliances in and around the world by the help of Machine learning.
- Not only the appliances but also these gestures are used to control volumes tuning, TV channels, speed controls, the regulator of a fan can be controlled.
- By the future advancement technologies gestures can be used to control cars and even software applications. It can provide high quality images, high performance, high accuracy and high reliable way to control the devices.

5.2 Applications

- Low cost and expandable allowing a variety of devices to be controlled
- Saves money and energy
- Low power requirement.
- Simple circuitry as it does not require special hardware.
- Devices can controlled more comfortably
- It requires small space.

6. CONCLUSIONS

In this paper, we have presented the design and implementation of a home automation system controlled by hand gestures. Experimental results showed the feasibility of using the gesture as a remote control interface for controlling home automation devices. The set of gestures used in the gestural interface was selected after a thorough study of comfort ability with a group of participants. However, the gestures may also be easily transferable to other application domains due to their character of universality. The use of HGCS system was very intuitive, because it just required a brief explanation to the user to learn it, being functional and adaptable. The construction of activities by the controller (HGC) using the hierarchical tree TGC helps the execution of actions on multiple home devices. It seems to be a significant step toward future explorations in this field of research. On conclusion, the gestural interface allows a complete freedom of movement and breaks the barrier of using touch user interface.

As future work we want to increase the gestures library to customize the system according to a user profile. The comfort ability is not the same for every user, and the system can accommodate different usage. We want to expand the applications to industrial control where the gestural interface should be more precise.

REFERENCES

[1] Gonzalo Pomboza-Junez and Juan A. Holgado-Terriza. "Control of Home Devices based on Hand Gestures". US Patent 7,161,489 B2. Sep. 9, 2004.

[2] Gill, K., Yang, S. H., Yao, F., & Lu, X. (2009). A zigbee-based home automation system. IEEE Transactions on Consumer Electronics, 422-430.
[3] Piyare, R., & Tazil, M. (2011). Bluetooth based home system using cell phone. 2011 IEEE 15th International Symposium on Consumer Electronics (ISCE) (pp. 192195). Singapore: IEEE.
[4] D.Martinez,"MSc Thesis - Sign Language Translator using Microsoft Kinect XBOX 360TM", VIBOT 5., Department of Electrical Engineering and Computer Science, Computer Vision Lab, University of Tennessee. Yang Quan and Peng Jinye, "Application of improved sign language recognition and synthesis technology in IB", Industrial Electronics and Applications, 2008. ICIEA 2008. 3rd IEEE Conference on.
[5] Computing and Communications, Galveston, TX, 2009, pp. 1-9 H. Guesgen and D. Kessell, "Gestural Control of Household Appliances.
[6] Guesgen, H. W., & Kessell, D. (2012). Gestural Control of Household Appliances for the Physically Impaired. Twenty-Fifth International Florida Artificial Intelligence Research Society Conference.
[7] Y. Kim and C. Moon, "Non-Contact Gesture Recognition Using the Electric Field Disturbance for Smart Device Application", in International Journal of Multimedia and Ubiquitous Engineering, Vol 9, Issue 2, 2014.
[8] Amrutha, S., Aravind, S., Ansu, M., Swathy, S., Rajasree, R., & Priyalakshmi, S. (2015). Voice Controlled Smart Home. International Journal of Emerging Technology and Advanced Engineering
[9] Deval G. Patel (2015), 'Point Pattern Matching Algorithm for Recognition of Gestures', International Journal of Recent Technology and Engineering (IJRTE), Vol.31, No.2, pp.20-33.
[10] Gupta, P. (2015). Human Voice Controlled Home Appliances. International Journal of Advanced Research in Computer and Communication Engineering.
[11] Rathnayake, K. A., Diddeniya, S. I., Wanniarachchi, W. K., Nanayakkara, W. H., & Gunasinghe, H. N. (2016). Voice operated home automation system based on Kinect sensor.
[12] Aditya, Arun Ramamurthy et al (2014), 'Recognition of dynamic hand gestures',International Journal of Recent Technology and Engineering (IJRTE), Vol.11, No.10,pp.14-34
[13] Asma Ben Hadj Mohamed and Juan A. Holgado-Terriza"Assisting people with disabilities through Kinect sensors into a smart house".Journal of Visual Communication and Image Representation 17, no. 6 (2009): 1190-1208.
[14] Ahmed M. Elshafee and Karim Alaa Hamed. "Design and Implementation of a WiFi Based Home Automation System." IEEE Transactions on Pattern Analysis and Machine Intelligence (IEEE) 25, no. 6 (2003): 564-577
[15] Ankita N. Chadha, Jagannath H. Nirmal, and Pramod Kachare. "A Comparative Performance of Various Speech Analysis-Synthesis Techniques." Jornal of Electronic Imaging, January 2014: 231-240.
[16] Juan J. Ojeda-Castelo Jose, A. Piedra-Fernandez and Cesar Bernal-Bravo. "Sign Communication for People with Disabilities Using Kinect Technology at Home." ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences: 133-138.
[17] Shihab Shahriar Hazaroi and Lamia Alam. "Designing a Sign Language Translation System using Kinect Motion Sensor Device." Robotic and Sensors Environments. Ottawa: IEEE, 2017. 150-155.
[18] Héctor J. Carlo, Yeiram Martínez, Cristina Pomales-García. "Real-Time Dock Door Monitoring System Using a Kinect Sensor." Computer Vision Workshops. Barcelona: IEEE, 2017. 1154-1160
[19] Dutta, Tilak. "Evaluation of the Kinect sensor for 3-D kinematic measurement in the workplace." Applied Ergonomics 43, 2012: 645-649.
[20] Https://books.google.com/books?id=cRS6BQAAQBAJ
[21] Huang-Chia Shih, Chang-Hsian Ma "Hand Gesture Recognition Using Color-Depth Association for Smart Home". (2018)
[22] Martin, Chris C., et al. "A Real-time Ergonomic Monitoring System using the Microsoft Kinect." Systems and Information Egineering Design Symposium. IEEE, 2012. 50-55.