Model for Object Detection using Computer Vision and Machine Learning for Decision Making

Aditya Raj
Information Technology
Rajkiya Engineering College
Bijnor, India.
nam2817120@gmail.com

Manish Kannaujiya
Information Technology
Rajkiya Engineering College
Bijnor, India

Ajeet Bharti
Information Technology
Rajkiya Engineering College
Bijnor, India

Rahul Prasad
Information Technology
Rajkiya Engineering College
Bijnor, India

Namrata Singh
Information Technology
Rajkiya Engineering College
Bijnor, India

Ishan Bhardwaj, PhD
Information Technology
Rajkiya Engineering College
Bijnor, India

ABSTRACT

This paper gives the efficient model of the process of detecting the object and analyzing the gesture of an object using the machine learning and computer vision for decision making after surveying various research in the field of pattern recognition.* It is based on the objects behavioural activities, Gesture recognition is very essential for making a decision. This research paper deals with Human-Computer Interaction (HCI). The object is may be hand, face, vehicle etc. In our research work we will extract 3D information about the captured image and calculate the distance of an object from the camera, detect the object gesture further we perform edge detection and then apply vision based and data based object gesture recognition. The implementation of the methodology will be done on the MATLAB tool.

Keywords
Object Gesture, Hand Gestures, Arduino, Ultrasonic Sensors, Human Computer Interaction (HCI), edge tracking.

1. INTRODUCTION
Smart work is essential now a days, it provides the efficient result with accuracy while saving the time. And especially when we talk about making a decision, and our decisions are based on the information that we have. Now, having the correct information is important. Now the question is from where we collect the information? Is it the correct source? So we decided to work on real-time images. We detect the object from the image with the help of algorithms, extract the 3D information from it, apply object gesture recognition techniques for analyzing the behavioural activities or changes of an object that will help us in making a decision. The object may be a hand, face, vehicle etc. Initially, we will take a hand as an object. Study its behavior or gestures through hand gesture recognition technique. To make a computer understand about the objects expressions and gesture is not that easy. Exchanging information through expressions and gestures are non verbally. The expression and gestures of an object are perceived through vision, we can study about it deeply in the research area of computer vision. There is a complex programming algorithm for analyzing the expression and gestures in machine language. The idea is to make a user-friendly interface. According to the need, we use the complex programming and techniques.

2. LITERATURE SURVEY
In paper [1], it works on a speech detection algorithm for speech activated hands-free application. Speech recognition provides a way to solve the problem for isolated word speech recognition. Many speech detection algorithms have been proposed for applications in car environments, a good speech detector should be noise robust, accurate and capable of real-time implementation. Normally for good condition moderate noise condition such as SNRs above 5db. However, for the car environment SNRs down to -10 DB. This algorithm contains three major parts: noise suppression, the robust endpoint detection and speech verification. In the first method, we have to, decrease SNRs to -10 DB for the car environment. They analyzed speech detection in several noise condition and consider these all that we have a need for the car environment extract information from that detected information. For this thing, we consider the subspace method and adaptive threshold method for speech detection. From these algorithms, we record the speech and after that step, the next step is speech verification for extract information from speech detection algorithm we compare the different noise suppression at different environment. Then after considering these all information we get a clear speech verification techniques.

In paper [2] they describe sensor ALIS. In 2002 they developed a new handheld landmine detection dual sensor ALIS (Advance Landmine Imaging System). ALIS is equipped with a metal detector and a GRP. It has a sensor tracking system which has record the GRP and metal detector signal with its location. ALIS uses two different systems, namely VNA (Vector Network Analyzer) based GRP and impulse GRP. They use a sensor tracking system. They record the image of recorded data in the graph and a considering information when they stored an image of different soil samples. Then extract the information from the metal detector data using 3D GPR image is reconstructed by the Kirchhoff migration algorithm. After applying Kirchoff algorithm on recorded data. Basically, they get two types of image: 1. Metal detector image 2. GPR image. These all are tested in CMAC (Cambodia Mine Action Center). There are 14 lanes in lanes in the test site: 6 lanes for handheld types and 6 for the vehicle-mounted type. The Lanes have three kinds of soils and then scanning all types of soil. They get different image and compare these all images data. We get the evaluation of landmine detection.

In paper [3], there are two certain techniques used for extracting the certain part of an image like as for mute people used mainly hand gestures to communicate with another person. This paper explains two different techniques of vision-based hand gesture and one based glove based technique. The vision-based techniques and second is a real-time algorithm. The data glove based technique is a flex sensor based glove. In static hand gesture recognition first, we have to click an image and save that image after that convert that image to grayscale and grayscale to a binary image. Then we apply morphological operation (for removing extra things or background) and to clear the image. Canny edge detection is also used for detection of figure edges and at last used Gaussian filter for clear detection. After edge detection, if we have to find or counting the number of fingers

then change the background pixel to the foreground pixel and foreground pixel to the background pixel to detect it. The number of changes recorded which gives a count no. Of Finger. So in vision-based techniques efficiency for detection is 100% and 86% in data glove based techniques.

In paper [4], they used different techniques for hand detection using arm segmentation from the depth data. This paper describes different techniques and algorithm for hand detection and palm gesture for detect hands and performs gesture recognition an RGD - D sensor is needed because it provides 3D data information like as distance between gesture sharing area and camera. First, they detect the face with RGB frame and then RGB pixel mapped on the depth map.

For extracting hand gesture we apply a convex hull algorithm which connects all the endpoints of hand and then by using area covers on a combination of starts points and end point they extract hand and palm gesture.

In paper [5], it builds upon a creative framework of the MARIACHI program at the stony book university. It integrates a research at the frontier of human knowledge with a broad discovery-based educational program. This project motivates teachers and students to experience cutting-edge technology and learn forefront science. The novel detection technique is radar-based and leads naturally to the implementation of strong educational and research programs in engineering and in particular signal processing. This research paper is mainly based on cosmic rays that collect the information about ultra-high energy cosmic rays. In MARIACHI experiment they use radar, scintillator ground detector, cyberinfrastructure. Overall these all are performed in an order and a number of students work on this experiment under the guidance of their teacher. In this research paper, MARIACHI is a high school project which has been completed by the students under the guidance of teachers.

In paper [6], they describe detecting invalid sample used in verifications and other techniques. In this section, it is given that if we scan the sample of hand or figure there are two type of categorized image:

1. Palm print feature
2. Hand geometry feature

This research paper is focused on second feature hand geometry which uses length, width, height, deviation and angles of a finger. Poor quality images are also responsible for invalid results. They describe three procedure for dealing with low-quality images and that is:

1. Requisitions
2. Quality based processing
3. Invoke human intervention

After these steps when we get object image then we detect the height of the finger and ratio between finger length. Because in general, the human body maintains some proportion (length of hand in approximately one-tenth of man's height. They measured the ratios between a various couple of finger lengths in a large collection of samples then storing these records. We check valid and invalid samples.

In paper [7], we focus on a w-band radar system, which is an excellent instrument for surveillance applications especially under challenging visibility conditions. The w-band radar system is capable to detect small drones. The sensor prevents in this work provides a wide bandwidth with high range revolution enabling not only the detection of objects in the perimeter but also a signature analysis. Using this technique radar is fully coherent. In this research paper, we recorded the drone location with four receiver channels of MORPS (Multichannel Radar for Perimeter

Surveillance). The whole paper is described UAS (Unmanned Aerial System). It is important for the protection of public and private areas. Overall further improvements for an operational UAS - detection radar system in the w-band. In future work, more efforts have to be targeting an enhancing the signal to noise ratio and by this increasing the detection range.

In paper [8], They use CMOS sensor technology for object detection and the main purpose of this research is to reduce the power consumption in hand gesture recognition system. There are two approaches used in gesture recognition.

Lower level for pose recognition and higher level for gesture recognition. Gesture recognition pipelines occur some steps:

1. Image acquisition
2. Hand segmentation
3. Feature extraction
4. Gesture classification
5. Gesture-based interaction
6. Appliance machine

After these steps images send to the image processor (ISP). These are some applications consumers appliance have the same characteristics:

The background is very less likely to change or changes very slowly in such applications except for the light condition changes. Even though the appliance has to continuously monitor for the gesture, the user doesn't control it continuously. This is rather done very occasionally and hand it more likely to move in and out of the field of view before and after almost every gesture. The gesture starts with a preparatory phase and ends with the retraction phase. Reducing the frame rate at the end of the gesture is a relatively simpler problem as part of gesture recognition process. So after the overall description, they used skin pixel percentage in the image for evaluating the performance, they captured the following statics: processed based percentage and false negative. If the percentage of a frame processed is high then lower power saving. After comparing this two static they get output for their program.

In paper [9], they are discussed about multiple hand gesture recognition systems under complex background. They also describe the gestures are of two types static and dynamic. Dynamic is also of three types:

1. Local motions
2. Global motions
3. Local and global motions

First, they used hand segmentation techniques which involves skin colour segmentation. In skin colour segmentation the input image is first taken and then the image in RGB colour space is converted into HCI colour space. After this process, they apply frame differencing. It simply means background subtraction. And then for multiple hands gesture, we use contour matching by applying contour matching algorithm we get multiple hands gesture clearly. Overall in this paper, they used an efficient algorithm for hand segmentation of dynamic hand gestures. They also tested the overall model under different background conditions.

In paper [10], they describe real-time hand gesture detection and recognition using a bag of features and support vector machine techniques. This paper is mainly based on symbol detection or motion detection using a hand. Like that, if we want to swipe left then swipe your hand from right to left. After using the algorithm for extraction of hand the hand gesture area only was saved in a small image, which will be used in extracting the key

points by the scale invariance feature transform (SHIFT) algorithm and then after applying vector quantization k- means clustering. We get a bag of word vectors and after storing these records if we match like stored data then they perform any action.

In paper [11], we read about adaptive skin colour model for hand segmentation. Hand segmentation is normally the first step in a gesture recognition or hand tracking. In this paper, they give a new technique for hand segmentation of the colour images using adaptive skin colour model. In this method, we capture the pixel value of a person's hand and convert them into YCbCr colour space. The technique will then map the YCbCr colour space to the cbcr plane to construct a clustered region of skin colour for space. Edge detection is applied to the cluster in order to create an adaptive skin colour boundaries for classification. The proposed algorithm introduces a new method for segmenting hand using adaptive skin colour model. After this algorithm, we find hand segmentation easily.

In paper [12], they discussed an algorithm for real-time text detection and recognition on handheld objects to assist blind people. First, we take an image and detect an object of interest detection to get the text pattern from the image. And text localization to obtain text regions and to recognize them. Text which is maximally stable is detected using Maximally Stable External Regions (MSER) features. A novel algorithm is evaluated on a variety of scenes. The detected text is compared with the template and convert into speech output. The text patterns are localized and binarized using Optical Character Recognition (OCR). The recognized text is converted to an audio output. The speech output is given to blind users.

In paper [13], they have discussed the algorithm for hand detection for the virtual touchpad. They used the normal method initially they draw a square or rectangular area which can be used for the touchpad and then apply action when the webcam detects any finger in that area. For complete this process for the next step and third step is touchpad detection. But before touchpad detection, we have to analyze edge detection, line segment detection, and quadrangle detection and then, at last, we have to apply hand detection but beforehand detection they also involving some process like as skin segmentation, fingerprint detection and state determination. After they identify the number and position of the fingertips. We determine the state of the system. The state is determined based only on fingertips counts. The state of this system can be illustrated as a finite state machine (FSM). There are a total of a six system states. The states depend on the no. And the type of finger detected. They detect no finger, move the pointer, left click, right click, middle click, gesture mode. In this type, they detect virtual notepad.

In paper [14], they have discussed none contact hand motion classification technique for application to human-machine interfaces. This paper presents an effective none contact technique for the measurement of human hand motion with wireless portable motion detection sensors, this technique allows the detection of the subject hand's movements. This method involves the measurement of the current generated due to the difference in capacitance between the subject's hand and the measurement electrodes in the motion detection sensors. The two motion detection sensors are capable of detecting the direction of subjects hand movement by measuring the time difference of peak of the electrostatic induction current. This paper we describe technique to recognize the direction of the hand movement with respect to the motion detectors sensors under perfect non-contact contains.

In paper [15], they were discussed to extract hand gesture from a complex because in previous use are seeing there is no complex background. So instead of using a simple background subtraction. the purposed system consists of four stages, detection of the appearance of hands, segmentation of hand regions, detection of full palm and hand gesture recognition. Detection of the appearance of hand is to find out when a hand appears in front of the camera. In HCI system, the segmentation of foreground objects such as hands, faces, and so on from the background is a major study for this study. Then detect the colour of the skin and after that segmentation of hand regions. After completing we apply the algorithm for detection of full palm and hand gesture recognition. And then the palm region is segmented from the background and filtered by the algorithm proposed. The palm is obtained and the system needs to know how many fingers have been raised. In order to detect no. Of fingers, the hand of the binary image will be transformed into a polar image for recognition. Polar hand image is a popular method. By this method, we get a proper recognition.

In paper [16], they have discussed convolution neural network based face detection. In this paper, they mentioned both the methods using the handcrafted method and entirely modern approach using a neural network with the help of a machine learning. Face detection is very hard to implement and there are no. Of ways and algorithm to find out the face. In this research paper, they used the handcrafted method with the help of HOG (Histogram Oriented Gradient) and second is LSB. And after this process, we get the face of an object using the neural network, it is implemented theoretically.

In paper [17], they have discussed real-time robotic hand control using a hand gesture. It means if we are doing the movement with our hand then the robot will have to perform the same action. The simple video camera is used for computer vision, which helps in maintaining gesture presentation. This approach consists of four module.

1. Real-time hand gesture formation monitor and gesture capture.

2. Feature extraction.

3. Pattern matching for gesture recognition and command determination corresponding to the shown gesture and performing an action by the robotic system.

4. The object of interest is extracted from the background and the portion of the hand representing the gesture is cropped out using the statistical property of hand. The extracted hand gesture is matched with the stored database of hand gesture using pattern matching.

In paper [18], they have discussed ensemble threshold segmentation for hand detection. This paper proposed an ensemble threshold segmentation (ETS), a method to segment the hand image for hand detection based on skin colour. Finally, the result was compared and analysed the accuracy among the existing colour model, with a threshold value and otsus method. At last ETS was implemented with five colour models using the optimized threshold range. The highest accuracy of segmentation method is ETS with the HSV colour model (98.51)% and lowest is ETS with YUU (71.18)% ETS and HSV colour model is optimized to segment the hand image for this dataset.

3. PROPOSED MODEL
ALGORITHM
Step: 1- Capture Image.

Step: 2- Subtract Background.

Step: 3- Object/Hand detection based on skin detection method.

Step: 4- Check whether hand is present or not.

(a)- if no then go to step 1.

(b)- if yes then go to the next step.

Step: 5- Discover and extract biggest contour.

Step: 6- Remove unwanted contour.

Step: 7- Locate hand position

Step: 8- Perform hand tracking.

Step: 9- Convexity defect.

Step: 10- Fingertips detection.

Step: 11- Binary conversion.

Step: 12- Noise removal apply Morphological transform and smoothing.

Step: 13- Edge detection.

Step: 14- Determine the total no. of fingers.

Step: 15- Find orientation about bounded area.

Step: 16- Extract the information from hand gesture/Modeling of gesture/Hand gesture recognition.

Step: 17- Interpret gesture and make decision.

Step: 18- Apply appropriate action.

Flow chart
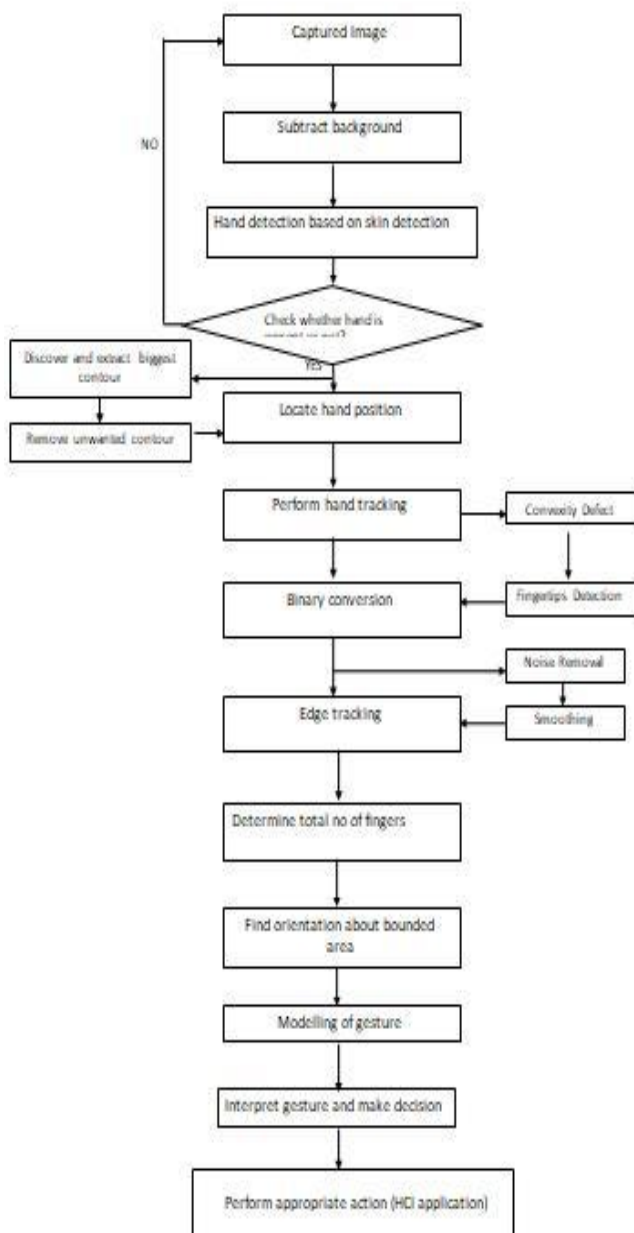
From the above process, we detect the object using machine learning and computer vision for decision making, basically, the aim of our paper is to recognize the object, object position, study the behaviour and gesture of an object for making an efficient and practical decision.

In our paper, we will extract 3D information about the object (image) and calculate the distance of hand (object) from the camera and subtract from the complete length using Depth Data and Static Palm Gesture Recognition. Then we get the image for Hand Gesture Recognition. After that Subtract the background obtained image. We use the Adaptive Skin Color method for detecting our hand gesture. For Edge detection, we are using Convex Hull Algorithm and for counting of a finger we use Vision Based and Data-Based Hand Gesture Recognition Technique then we can able to count the finger. For finding the edge detection of Image, we use Canny Edge detection the . After that overall, we select the contours part or moving part and perform the function for machine learning. Now we extract the meaningful information from observing the behaviour and gesture of an object through an algorithm for decision making. To read that particular gesture and take a decision is not that easy. Here we are using the concepts, techniques and algorithms of machine learning, computer vision and artificial intelligence.

4. CONCLUSION

In this research paper, we are proposing a model for object detection, it may be a hand, face or vehicle. This model is basically for hand detection using the concept of computer vision and machine learning based on the survey for the purpose of making a decision and HCI (Human-Computer Interaction). We have decided the various objects, techniques and algorithm to work on. This is only a proposed model. We are in between of implementation.

5. FUTURE SCOPE

Future scope: In this research paper, we have proposed a model based on the survey. We have decided the concept, techniques and algorithm to work on. We will implement this proposed model. Some modifications are possible.

6. REFERENCES

[1] D. Wu, M. Tanaku, R. Chen, L. Olorenshaw, M. Amador and X Menendez-Pidal, A ROBUST SPEECH DETECTION ALGORITHM FOR SPEECH ACTIVATED HANDS-FREE APPLICATIONS Spoken Language Technology, Sony US Research Laboratories 3300 Zanker Road, San Jose, CA 95 134, USA.

[2] Motoyuki Sato, Kazunori Takahashi, Jun Fujiwara, Hand held dual sensor ALIS and its evaluation test in Cambodia, Center for Northeast Asian Studies Tohoku Universit Sendai, Japan.

[3] Oinam Robita Chanu1, Anushree Pillai2, Spandan Sinha3, Piyanka Das4, Comparative Study for Vision Based and Data Based Hand Gesture Recognition Technique, Department of Biomedical Engineering SRM University, Kattankulathur, Chennai, India, 2017 International Conference on Intelligent Communication and Computational Techniques (ICCT) Manipal University Jaipur, Dec 22-23, 2017.

[4] Rytis Augustauskas, Arunas Lipnickas, Robust Hand Detection using Arm Segmentation from Depth Data and Static Palm Gesture Recognition , Kaunas University of Technology, Studentu 48, Kaunas 51367, Lithuania, The 9th IEEE International Conference on Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications 21-23 September, 2017, Bucharest, Romania.

[5] M´onica F. Bugallo†, Helio Takai‡, Michael Marx†, David Bynum and John Hover, Brook University, Stony Brook, NY 11794 Brookhaven National Laboratories, Upton, NY 11973.

[6] Javier Burgues, Julian Fierrez, Daniel Ramos, Maria Puertas, Javier Ortega-Garcia, Detecting Invalid Samples in Hand Geometry Verification Through Geometric Measurements, Biometric Recognition Group - ATVS, EPS - Univ. Autonoma de Madrid C/ Francisco Tomas y Valiente, 11 - Campus de Cantoblanco - 28049 Madrid, Spain {javier.burgues, julian.fierrez, daniel.ramos, maria.puertas.

[7] M. Caris, W. Johannes, S. Sieger, V. Port, S. Stanko, Detection of Small UAS with W-band Radar, Fraunhofer Institute for High Frequency Physics and Radar Techniques FHR Fraunhoferstr. 20, 53343 Wachtberg, GERMANY.

[8] Mahesh CHANDRA Consumer Products Division STMicroelectronics Pvt. Ltd. Greater Noida, Uttar Pradesh, India, Brijesh LALL Electrical Engineering Department Indian Institute of Technology Delhi, India, A Novel Method for Low Power Hand Gesture Recognition in Smart Consumer Applications.

[9] Ananya Choudhury Dept. of Electronics and Communication Engineering Gauhati University Guwahati-14, Assam, India, Anjan Kumar Talukdar Dept. of Electronics and Communication Engineering Gauhati University Guwahati-14, Assam, India , Kandarpa Kumar Sarma Dept. of Electronics and Communication Engineering Gauhati University Guwahati-14, Assam, India, A Novel Hand Segmentation Method for Multiple-Hand Gesture Recognition System under complex background.

[10] Nasser H. Dardas and Nicolas D. Georganas, Fellow, IEEE, Real-Time Hand Gesture Detection and Recognition Using Bag-of-Features and Support Vector Machine Techniques, IEEE TRANSACTIONS ON INSTRUMENTATION AND MEASUREMENT, VOL. 60, NO. 11, NOVEMBER 2011.

[11] Ahmad Yahya Dawod Faculty of Information Technology, Multimedia University, Cyberjaya, Malaysia, Junaidi Abdullah Faculty of Information Technology, Multimedia University, Cyberjaya, Malaysia, Md.Jahangir Alam Faculty of Information Technology, Multimedia University, Cyberjaya, Malaysia , Adaptive Skin Color Model for Hand Segmentation, 2010 International Conference on Computer Applications and Industrial Electronics (ICCAIE 2010), December 5-7, 2010, Kuala Lumpur, Malaysia.

[12] Samruddhi Deshpande Department of Instrumentation and Control Cummins College of Engineering For Women Karvenagar, Pune , Ms. Revati Shriram Department of Instrumentation and Control Cummins College of Engineering For Women Karvenagar, Pune, Real Time Text Detection and Recognition on Hand Held Objects to Assist Blind People., 2016 International Conference on Automatic Control and Dynamic Optimization Techniques (ICACDOT) International Institute of Information Technology (I²IT), Pune.

[13] Gerard Edwin#1, Iping Supriana#2 #School of Electrical and Informatics Engineering, Institut Teknologi Bandung Jl. Ganesha 10 Bandung 40132, Indonesia, Hand Detection for Virtual Touchpad, 2011 International Conference on Electrical Engineering and Informatics 17-19 July 2011, Bandung, Indonesia.

[14] Koichi Kurita, Noncontact Hand Motion Classification Technique for Application to Human–Machine Interfaces, IEEE TRANSACTIONS ON INDUSTRY APPLICATIONS, VOL. 50, NO. 3, MAY/JUNE 2014.

[15] Chung-Ju Liao, Shun-Feng Su, Ming-Chang Chen Department of Electrical Engineering National Taiwan University of Science and Technology Taipei, Taiwan, Vision-based Hand Gesture Recognition System for a Dynamic and Complicated Environment, 2015 IEEE International Conference on Systems, Man, and Cybernetics.

[16] Subham Mukherjee Institute of Engineering & Management, Kolkata ,Sumalya Saha Institute of Engineering & Management, Kolkata, Sounak Lahiri Institute of Engineering & Management, Kolkata, Ayan Das Institute of Engineering & Management, Kolkata, Ayan Kumar Bhunia Institute of Engineering & Management, Kolkata, Aishik Konwer Institute of Engineering & Management, Kolkata, Arindam Chakraborty Institute of Engineering & Management, Kolkata, Convolutional Neural Network based Face detection.

[17] Jagdish Lal Raheja, Radhey Shyam, Umesh Kumar, P Bhanu Prasad Digital Systems Group Central Electronics Engineering Research Institute (CEERI)/Council of Scientific & Industrial Research (CSIR) Pilani – 333031, Rajasthan India, Real-Time Robotic Hand Control using Hand Gestures, 2010 Second International Conference on Machine Learning and Computing.

[18] Sunthorn Rungruangbaiyok, Rakkrit Duangsoithong, Kanadit Chetpattananondh Department of Electrical Engineering, Faculty of Engineering Prince of Songkla University, Hat Yai, Songkhla, 90112, Thailand, Ensemble Threshold Segmentation for Hand Detection.

[19] Kjeldsen and Kendersi, devised a technique for doing skin-tone segmentation in HSV space, based on the premise that skin tone in images occupies a connected volume in HSV space.

[20] Etsuko Ueda and Yoshio Matsumoto presented a novel technique a hand-pose estimation that can be used Fig 1.2 VPL data glove 7 for vision-based human interfaces, in this method, the hand regions are extracted from multiple images obtained by a multiviewpoint camera system, and constructing the "voxel Model." Hand pose is estimated.