Gestures based Wireless Robotic Control using Image Processing

Gaurav Chauhan
Dept. of Electronics & Telecommunication Engineering
MIT Academy of Engineering, Alandi (D)
Pune, India
gaurav_chauhan15@outlook.com

Prasad Chaudhari
Dept. of Electronics & Telecommunication Engineering
MIT Academy of Engineering, Alandi (D)
Pune, India
prasad21.nc@gmail.com

***Abstract*— many people in this world are known to gestures, a powerful communication medium amongst humans. They say actions speak louder than words. Communicating with gestures is a powerful method. Industrial as well as robots used for home purpose are often controlled by remote controllers. The use of gestures as a way to control can help a lot providing ease and can be more useful for the handicapped. Gestures provide a separate complementary modality to speech for expressing ones ideas. Furthermore, Image processing is one of the most effective method used to process image signals. The main purpose of gesture recognition research is to identify a particular human gesture and convey information to the user pertaining to individual gesture. Sign language recognition from hand motions or postures is an active area in gesture recognition research for Human Computer Interaction (HCI). This paper aims to brief the practical approach of robotics through an effective method of image processing using gestures as a mode of control.**

Keywords—Image processing, HCI, Gestures, Recognition, RGB,

I. INTRODUCTION

A gesture is a form of non-verbal communication or non-vocal communication in which visible bodily actions communicate specific messages, either in place of, or in unification with, speech. Gestures are a powerful means of communication among humans. Face gesturing is so deeply embedded in our communication system that people often continue gesturing when speaking on the telephone, mobiles or with one another. Hand gestures provide a distinct corresponding modality to speech for expressing ones notions. Information associated hand gestures in a conversation is degree, dissertation structure, spatial and time-based structure. The term Gesture is defined as ``movement to convey meaning'' or ``the use of motions of the limbs or body as a means of expression; a movement usually of the body or limbs that expresses or emphasizes an idea''. [1] The main purpose of gesture recognition research is to identify a particular human gesture and convey information to the user pertaining to individual gesture. Modelling and recognizing a gesture is a difficult challenge since gestures occur dynamically both in shape and in duration. Sign language recognition from hand motion or hand posture is an active area in gesture recognition research for Human Computer Interaction (HCI). Computer recognition of hand gestures may provide a more natural-computer technology interface, allowing persons to point, or rotate a CAD structured model by rotating their hands. [2] Hand gestures can be classified in two types:

1. Static Gestures
2. Dynamic Gestures

A static gesture is a specific hand configuration and pose, represented by a single image while a dynamic gesture is gesture motion, represented by a sequence of images. We will focus on the recognizing of a static gesture.

A robot is an electro-mechanical or may be virtually artificial host, mostly a motion based machine that is influenced by a computer program and an electronic circuitry. Furthermore, it can also be used by handicapped people for their betterment. Gestures could be hand, fingers, or eyes, the idea behind this is the same. The goal of this project is to build a Human Computer Interaction (HCI) system that has the ability to recognize hand gestures of humans and convert them into appropriate commands. These commands are used to drive a mobile robot and accordingly the required tasks are performed. The whole process should be done in real-time, the thing that necessities employing fast and efficient algorithms. In addition, the actual implementation of our system should take practicality and ease of use into account. This is because a system as ours is intended to be used in daily life and hence needs to be both simple and efficient to use. [3]

II. LITERATURE SURVEY

Gesture recognition enables humans to interface with the machine (HMI) and interact naturally without any mechanical or electro-mechanical devices. Using the idea of gesture recognition, we can point a finger at the computer screen so that the cursor will move accordingly. This could in theory make conventional input devices such as mouse, keyboards and even touch screens out of work. Gestures recognition can be conducted with techniques from computer vision and image processing. Most of the complete hand interactive systems can be considered to be comprised of three layers: detecting, tracking and recognizing. The detection layer is responsible for defining and extracting visual features that can be attributed to the presence of hands in the field of view of the camera(s). The tracking layer is responsible for accomplishing temporal data association between successive image frames, so that, at each instant in time, the system must be aware of ``what is where''. Moreover, in modal-based methods, tracking also provides a way to maintain estimates of typical parameters, variables and features that are not directly observable at a certain moment in time. Lastly, the recognizing layer is responsible for grouping the spatiotemporal data extracted in the previous layers and assigning the resulting groups with labels associated ta particular classes of gestures.

We can use DSP processors other than the ordinary microcontroller, but to make the project cost- effective, it's not a favorable option. DSP processors are highly qualified for signal processing but on the other hand, the starter kits are high cost. As a result, we're using a PC with a computing software such as matlab. Matlab is appropriate for such use with the availability of image processing toolbox, which has all the tools for acquiring an image and its parameters. We can also use methods for processing images. The ATMega microcontroller has a RISC architecture, better than the conventional CISC microcontrollers. It has reduced instruction set and the programming tools are free, readily available.

Over a decades ago, robots used to be wired. Then came the wireless RF modules, one of which is ZigBee. Zigbee is low-cost, low-power, wireless mesh network standard targeted at wide development of long battery life devices in wireless control and monitoring applications. ZigBee devices can transmit data over long distances by passing data through a mesh network of intermediate devices to reach more distant ones. ZigBee module has a well-defined data rate of 250 kbit/sec, best suited for intermittent data communications within a sensor or input device.

III. WORKING

The working of the whole system can be divided into two parts namely –

(I) Controlling Unit (II) Assembly Unit

1. *Controlling Unit*

At first, an input of image is taken through the webcam on the computer/laptop. This input is then processed through matlab using the Image Acquisition Toolbox. [2] A programming code was written to assign a command to the taken input image signal. For example, we assigned an alphabet to the command. These signals which have been assigned commands are then exported from the PC to a wireless RF module (in this case a Zigbee module) using a RS 232 to TTL converting IC (MAX 232), as the RS232 signals aren't compatible with the RF module i.e. the ZigBee. The signals that are in analog nature are converted to digital nature so that they're compatible with the RF module. This RF module is used to have a wireless control; containing a transmitter on one end and a receiver on the other. The main function of this unit is to assign commands and have a control over the Robot Assembly Unit. The Controlling unit is as shown in block diagram in fig 1.

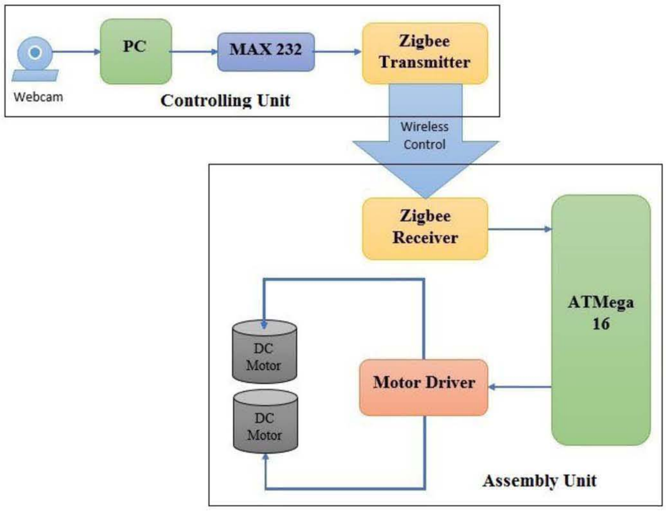


Figure 1: Block Diagram

2. *Assembly Unit*

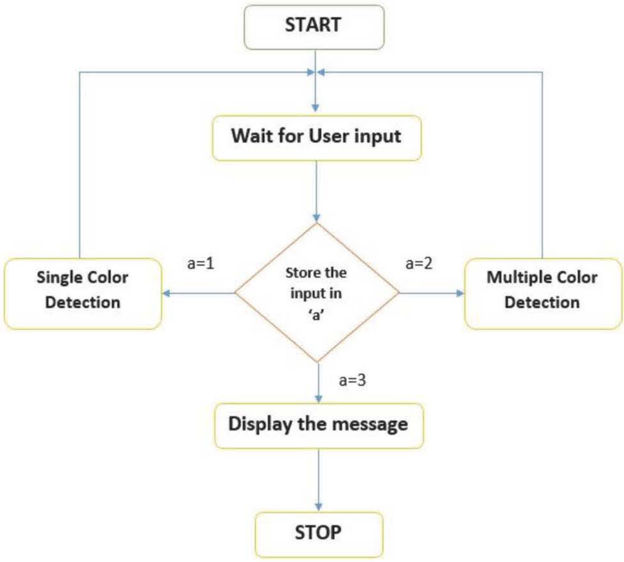
The Assembly Unit is mainly of two parts – The microcontroller (Here, we're using ATMega 16) as shown in fig 1. The important function of this unit is to receive the command signals and as per them, drive the robot assembly.

We're using the RF receiver in this unit to receive control commands from the controlling unit. The commands received by the RF receiver will be given to the microcontroller As discussed earlier, the commands received are in alphabets which are recognized and as per it the robot assembly will move. The microcontroller will then decode these commands and send it to the motor driver. The motor driver will drive the DC motor accordingly. A code in embedded C is written to recognize these commands and drive the motor driver. The Robot assembly consists of two DC motors acting as wheels and a motor driver (L293D). The motor driver is a standard IC used to drive DC motors. [4]

IV. IMAGE PROCESSING

Image processing is any form of information processing for which the input is an image, such as photos or frames of video; the output is not necessarily an image, but can be for case in point a set of features of the image. Many of the image-processing techniques involve treating the image as s two-dimensional signal and applying standard signal processing techniques to it. It is clear that image processing do have some kind of manipulation towards what it is fed to do. LabVIEW, their concept manual describes Image Processing as functions to study, filter and process images in NI Vision. A pixel, short for picture element, can be thought of as a tiny dot containing information about the picture. When you snap a picture, these tiny bits of information are gathered by the camera's sensor. The information is being stored in a plane of information. Each plane represents three colors that are red, green and yellow plane. Each plane has the intensity from 0 up to 255 or 8-bit of information per plane. These three color combination makes up all the colors we could see in RGB images. A flowchart for image processing is shown in fig. Simple calculation of these 8-bit information is as follows:-

$$22^{n} = m\,Bits\ ;\ 2^{8} = 256$$



Flowchart 1: Image Processing

1. Webcam

Webcam is a video camera that feeds or streams its image in real time to or through a computer or computer network. When ``captured" by the computer, the video streamed or image may be saved, viewed or sent on to other networks vie systems such as internet, email as an attachment. Webcams are known for their low manufacturing cost and flexibility. Now-a-days, these are frequently used for video calling, live video conferences etc.

2. Image Acquisition

For the signals to be send to the microcontroller, first, we have acquire images from the webcam of the laptop/computer. As these images are acquired, we can convert them into suitable formats and assign a code to them. Image Acquisition Toolbox lets us acquire images and videos directly into MATLAB from PC-compatible imaging hardware. We can detect hardware automatically, configure its properties, preview an acquirement, and get images and videos. We can use a range of imaging devices from inexpensive Web cameras or industrial frame grabbers to high-end scientific cameras that meet low-light, high-speed, and other perplexing requirements. Collectively, MATLAB, Image Acquisition Toolbox, and Image Processing Toolbox provide a complete environment for developing customized imaging applications. We can acquire images and videos, visualize data and develop processing algorithms and analyzing techniques, and create graphical user interfaces. Some Key Features are:

- By design, it automatically detects image and video acquisition devices.
- Manages device configurations.
- Provides live video previewing.
- Acquires static images and continuous videos.

Image acquisition Toolbox helps us connect to and configure our hardware, preview the acquisition, and acquire and visualize image data. With this us control your image acquisition parameters and incorporate them into M-scripts, applications built within MATLAB. Within the toolbox, the Image Acquisition Tool is a graphical user interface for working with image and video acquisition devices in MATLAB. With this tool, we can see all hardware available an your PC, change device settings, preview an acquisition, control acquisition parameters, and acquire image or video data. The flowchart for finger detection is shown in fig. To use the image Acquisition Toolbox to obtain image data, you must perform the following basic steps:

1. Install and configure the image acquisition device.
2. Create a video input object.
3. Preview the video stream.
4. Configure image acquisition object properties.
5. Acquire image data. [5]

3. Image Acquisition Toolbox

The Image Acquisition Toolbox, as talked about earlier, enables us to acquire images or videos and display their parameters using commands. The hardware can be detected automatically. It provides a programmatic interface to help us work with the acquisition hardware in Matlab. It enables us to customize the acquisition process to include integrating image processing functionality to identify objects or some things, enhance the imagery or also construct mosaics and paranormal views as the data is acquired. This can be used on Windows, Linux and Macintosh systems. The system is designed to acquire imagery as fast as the camera and PC can support, enabling analysis and processing of high-speed imaging applications.

Some of the code used and its meaning is as follows:

- v=videoinput('winvideo', 1, 'YUY2_640x480'):
  creating a "videoinput" object 'v'.
- im=getsnapshot(v):
  storing the image in 'im' variable.
- imshow(im):
  displaying the captured image.
- imwrite(im,'c:\image.jpg'):
  writing/storing the captured image in the hard-disk.
- v.FramesPerTrigger= 1:
  Setting the "FramesPerTrigger" value to '1'.
- delete(v):
  deleting the "videoinput" object to avoid filling up of memory.
- IM2 = imdilate(IM,SE):
  Enlarges the grayscale, binary or a packed binary image i.e. IM, returning the dilated image, IM2.
- input=rgb2gray(im):

Converts the image in variable im to rgb

- imhist(input):
  Plots the histogram for the image input. The image should be converted to binary for that.

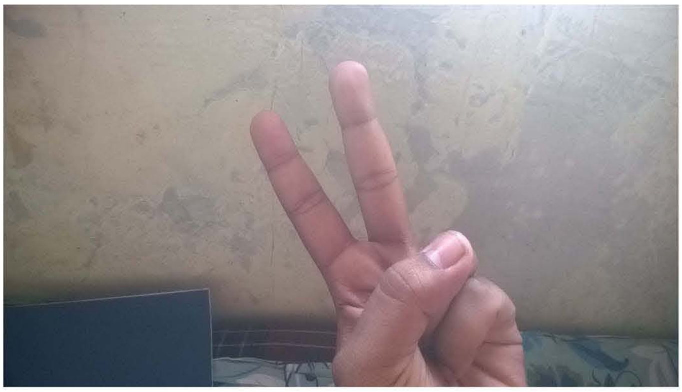


Figure 2: Captured image by webcam (two fingers)

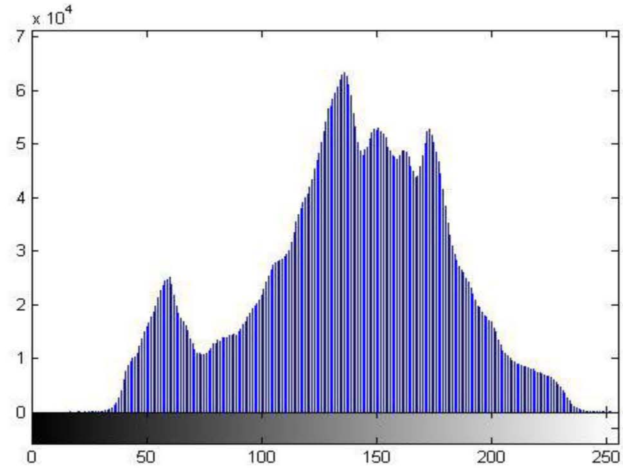
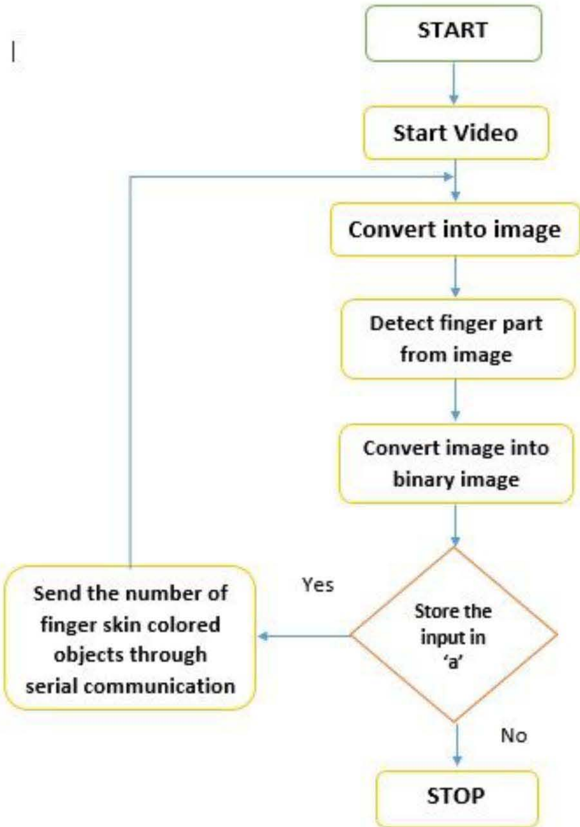


Figure 3: Histogram for fig. 2

Figure 4: Binary image produced of fig. 2



Flowchart 2: Finger Detection

Figure 5: Captured image by webcam (five fingers)

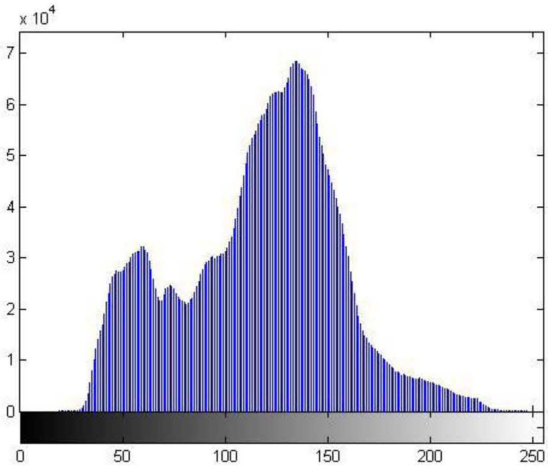


Figure 6: Histogram for fig. 5

Figure 7: Binary image produced of fig 4

4. *Processing Algorithm*

- First the camera shall capture the image of the finger gesture the user is making as shown in fig 1 & 3.
- Then the captured image is processed using the matlab.
- In the processing stage, first, all the red, green and blue parts are detected and are separated from each other.
- The Background noise is removed i.e. anything apart from the color black.
- The threshold value is calculated for that particular surrounding.
- After this, we initialize the image and convert it into binary image; white denoting upper finger skin color dots and black denoting absence of finger color skin color dots.
- The logical image is produced and number of dots are counted.
- This counted value is then compared with the database.
- Finally, the commands are sent to the robot using the RF module to the microcontroller and to motor driver. [6]

V. RESULTS

Fingers detection and tracking, testing results are tabulated below:

Number of Tests	Number of Successful Detections	Number of Defective Detections	Percentage of Successful Detections
10	08	2	80

Table 1: Analysis of tests

The above table shows that out of 10 tests, our system was able to detect and track the fingers 80% of the time. Failure can be accounted to different lighting conditions that the fingers are susceptible to. The brightness of the room can greatly affect the tracking and detection process. Lighting conditions affect accurate detection because the detection algorithm used for capturing and tracking the fingers is highly sensitive to color variations.

VI. ADVANTAGES

- Hand gesture recognition system is very useful for the physically impaired persons. The system can be trained to help these people to communicate with each other or to carry out tasks.
- Human machine interface can be made easy as Image processing is one of the revolutionary topics in electronic world.
- Due to finger detection method, hand free movements can be detected and this can be used in advanced 3D gaming, taking the game technology to a whole new level.
- It is fully automatic, so there is no need of any man-power. It also doesn't have frequent problems.
- The system is self-sufficient and uses very simple methodology. It increases system reliability.
- Also, it is cost effective with use of cheap but powerful components.

VII. FUTURE SCOPE

1. *Sign Language Recognition*

Just as speech recognition can write out speech to text, certain types of gesture recognition software can transcribe the symbols represented through sign language into text.

2. *For Socially Assistive robotics*

By using proper sensors (accelerometers and gyros) worn on the body of a patient and by reading the values from those sensors, robots can assist in patient and by reading those values from the sensors, robots can assist in patient rehabilitation. One of the best example can be Stroke Rehabilitation.

3. *Directional indication through Pointing*

Pointing has a very explicit and specific purpose in our society; to reference an object or location based on its position relative to ourselves. The use of gestures recognition to decide where a person is pointing. Is useful for identifying the perspective of statements or instructions. This application is of precise attention in the field of robotics and image processing.

4. *Control through Facial Gestures*

Controlling a computer through facial gestures is a useful application of gesture recognition for users who may not physically be able to use a mouse or keyboard. Eye tracking in particular can be used for controlling cursor motions or focusing on elements of a display.

5. *Alternative Computer Interfaces*

Foregoing to the old-styled keyboard and mouse setup used to interrelate with a computer, strong gestures recognition system could allow users to achieve success in doing frequent or common tasks using hand or face gestures in front of a camera.

6. *Immersive Game Technology*

Gestures can be used to control actions and interfaces within video games to try and make game player's experience more interactive or immersive.

7. *Virtual Controllers*

For systems where a lot of finding or acquiring a physical controller could require not much time, gestures can be used as an alternative controlling tool. Controlling subordinate devices in an automobile, or controlling a television set are cases of such usage.

8. *Effective Computing*

An effective computing, gesture recognition is used in the process of identifying emotional expressions through computer systems.

9. *Remote Control*

Through the use of gesture recognition, remote controlling has been possible with the movement of a hand or finger of various devices. The signal must not only show the desired response, but as well as device to be controlled.

VIII. Conclusion

This paper successfully briefed about a system to recognize the fingers from image sequences by the motion trajectory of a single hand which is suitable for real time application. There are many approaches to gesture recognition and each approach has its strengths and weaknesses. The strong point of the proposed method in this paper is its simplicity of the system and cost effectiveness. The system with the proposed algorithm achieved 80% average recognition rate which was very good as compared to the obstacles faced like inadequate light conditions and approximate threshold values. The proposed method was having some problems recognizing the fingers, but this is the limitations of the technology, inevitable problems. The system has an advantage that the actions performed by the robot are controlled from gestures and some additional gestures can be designed for human Computer Interaction (HCI) such as sign language translation.

IX. Acknowledgement

The successful completion of our research work within the stipulated time frame is a result of collective efforts of our group as well as all the people who provided the continuous support. Here, we would like to thank all those people for their timely guidance. Our Head of Department for his encouragement and providing the required resources. We would also like to thank our Project Guide for her constant guidance without which our research paper was impossible. Also we would like to thank our college, MIT AOE, Alandi (D), Pune for providing us the chance to present our knowledge in terms of this research paper.

X. REFERENCES

1] Pragati Garg, Naveen Aggarwal and Sanjeev Sofat, “Vision Based Hand Gesture Recognition”, World Academy of Science, Engineering and Technology, Vol.3, 2009-01-28.

2] Vladimir I. Pavlovic, Rajeev Sharma, and Thomas S. Huang, "Visual interpretation of hand gestures for human-computer interaction: A review", IEEE Trans. Pattern Analysis and Machine Intelligence., July, 1997, Vol. 19 (7), pp. 677 - 695.

3] Cristina Manresa, Javier Varona, Ramon Mas and Francisco J. Perales, “Real –Time Hand Tracking and Gesture Recognition for Human-Computer Interaction”, Electronic Letters on Computer Vision and Image Analysis, 0(0):1-7, 2000, Received 1 January 2000; revised 1 January 2000; accepted 1 January 2000.

4] Harish kumar kaura, Vipul Honrao, Sayali Patil, Pravish Shetty, “Gesture Controlled Robot using Image processing”, International Journal of Advanced Research in Artificial Intelligence, Vol. 2, No. 5, 2013.

5] Lars Bretzner, Ivan Laptev, Tony Lindeberg, “Hand Gesture Recognition using Multi-Scale Colour Features, Hierarchical Models and Particle Filtering”, Proceedings of the Fifth IEEE International Conference on Automatic Face and Gesture Recognition (FGRí02), 2002.

6] Apostolos Tsagaris1, Sotiris Manitsaris, “Color Space Comparison For Skin Detection in Finger Gesture Recognition”, International Journal of Advances in Engineering & Technology, Sept. 2013.

7] Rafiqul Zaman Khan and Noor Adnan Ibraheem, “Hand Gesture Recognition: A Review”, International Journal of Artificial Intelligence & Applications (IJAIA), Vol.3, No.4, July 2012.