

# Image Processing Based Customized Image Editor and Gesture Controlled Embedded Robot Coupled with Voice Control Features

Somnath Kar<sup>1</sup>

Department of Electrical Engineering  
Maulana Abul Kalam Azad University of Technology  
Kolkata, West Bengal, India

Ankit Jana<sup>2</sup>

Department of Electrical Engineering  
Maulana Abul Kalam Azad University of Technology  
Kolkata, West Bengal, India

Debarshi Chatterjee<sup>3</sup>

Department of Electrical Engineering  
Maulana Abul Kalam Azad University of Technology  
Kolkata, West Bengal, India

Dipayan Mitra<sup>4</sup>

Department of Electrical Engineering  
Maulana Abul Kalam Azad University of Technology  
Kolkata, West Bengal, India

Soumit Banerjee<sup>5</sup>

Department of Electrical Engineering  
Maulana Abul Kalam Azad University of Technology  
Kolkata, West Bengal, India

Debasish Kundu<sup>6</sup>

Gobindapur Saphali Memorial Polytechnic Gobindapur,  
Guskara, Burdwan,  
West Bengal, India

Sudipta Ghosh<sup>7</sup>

Bengal Institute of Technology and Management,  
Santiniketan, Gopalnagar, Birbhum,  
West Bengal, India

Sauvik Das Gupta<sup>8</sup>

Electrical and Computer Engineering  
Oklahoma State University  
Stillwater, OK, USA

**Abstract**—In modern sciences and technologies, images gain much broader scopes due to the ever growing importance of scientific visualization (of often large-scale complex scientific/experimental data) like microarray data in genetic research, or real-time multi-asset portfolio trading in finance etc. In this paper, a proposal has been presented to implement a Graphical User Interface (GUI) consisting of various MATLAB functions related to image processing and using the same to create a basic image processing editor having different features like, viewing the red, green and blue components of a color image separately, color detection and various other features like noise addition and removal, edge detection, cropping, resizing, rotation, histogram adjust, brightness control that is used in a basic image editor along with object detection and tracking. This has been further extended to provide reliable and a more natural technique for the user to navigate a robot in the natural environment using gestures based on color tracking. Additionally, Voice control technique has been employed to navigate the robot in various directions in the Cartesian plane employing normal Speech recognition techniques available in Microsoft Visual Basic.

**Keywords**—Image Processing; Image Editor; Gesture Control; Embedded Robot; Voice Control

## I. INTRODUCTION

In imaging science, image processing is processing of images using mathematical operations by using any form of signal processing for which the input is an image, such as a photograph or video frame; the output of image processing may be either an image or a set of characteristics or parameters related to the image.

MATLAB based image processing [1] is a very convenient platform and suitable for programming. An image is a matrix of pixel values. MATLAB considers every input as a matrix. For this reason, MATLAB provides an easy tool for image processing as a user can easily access each and every pixel value from the image matrices and edit it. Moreover, there is an 'image processing tool box' built in MATLAB for this purpose. Guerrero, J. [2] has demonstrated the use of Matlab and GUI for image processing and implemented in a deep vein thrombosis screening system.

Object tracking is a mature discipline aiming to define techniques and systems for processing videos from cameras placed in a specific environment. The need for high power computers, the availability of high quality and inexpensive

video cameras, and the increasing need for automated video analysis has generated a great deal of interest in object tracking algorithms. Tracking an object in video has a variety of real world applications; these include autonomous aerial reconnaissance, remote surveillance, and advanced real time collision avoidance systems. There are three key steps in video analysis viz. detection of interesting moving objects, tracking of such objects from frame to frame, and analysis of frames to recognize their behavior. In its simplest form, tracking can be defined as the problem of estimating trajectory of an object in the image plane as it moves around a scene. The main aim is to track the real-time moving objects in different video frames with the help of a proposed algorithm. P.R.V. Chowdary [3] showed us an example to implement image processing algorithm for gesture recognition. The work of T. Mahalingam [4] in vision based moving object tracking through enhanced color image segmentation using Haar classifiers is very helpful to understand the underlying principle of color tracking which is also this project's backbone.

On the other hand, hand gestures can be interpreted as described by Luigi Lamberti<sup>1</sup> and Francesco Camastra in their paper [5] where they have modeled a color classifier performed by Learning Vector Quantization. J.S. Kim [6] et al. developed a pattern recognizing algorithm to study the features of the hand. Some gesture recognition systems involve adaptive color segmentation [7], hand finding and labeling with blocking, morphological filtering, and then gestures are found by template matching. These processes do not provide dynamicity for the gesture inputs.

Several other gesture controlled robotic system uses various ways to recognize the gesture commands such as hand finding and labelling with blocking, showing specific number of fingers [8] for specific command, hand position and orientation which are measured by ultrasonic for gesture recognition. M. Mahalakshmi [9] used CAMSHIFT algorithm for real time vision based object tracking. T. Said [10] proposed a different way of controlling multi robots through multi-object color tracking.

Some other technologies use Microsoft Xbox 360 Kinect(C) [11] for gesture recognition. Kinect gathers the color information using an RGB and depth information and Infra-Red camera respectively. This system though is not very cost effective.

Keeping all previous works in mind, here a system has been developed where the main objective is to provide reliable and a more natural technique for the user to navigate a robot in the environment using gestures. Certainly, building a less costly and robust system was the motive behind this work. The primary focus on building this gesture controlled robot is on the type of gestures. The gesture, this work mostly concentrated on was color tracking. A system is proposed where the robot will track the movement of the particular color and will move along that direction. In absence of any command the robot will stop.

The system implementation involves the design of an advanced User Interface that controls the robot's movement with the help of either a specific color tracking through image processing or specific voice commands. This process has been

tested for two separate colors. The color detection and tracking algorithm has been evaluated on a self-developed embedded prototype built on an open source AVR microcontroller-based platform (*ARDUINO*). To understand the underlying principles of image processing, at first an image editor has been developed with the help of Matlab Graphical User Interface (GUI). This image editor demonstrates the functions that has been used in the main color detecting and tracking algorithm. The next section describes the hardware components required, followed by section III, describing the methodology to design the system. Then comes the experimental evaluation section followed by section V with conclusion and section VI with future works.

## II. HARDWARE PLATFORM

The hardware part mainly consists of a digital computer, an Arduino Uno board, Dual H-Bridge Motor Driver, Integrated laptop webcam, DC Geared Motors, Metal chassis, two wheels and a free wheel which are being discussed along with their specific functions.

### A. Arduino Uno with Arduino Cable

An Arduino Uno is an Atmel 8-bit, 16-bit or 32-bit AVR microcontroller based board, with complementary components which helps in circuit incorporation. The Arduino Uno can be programmed with the Arduino Software. The ATmega328 on the Arduino Uno comes preprogrammed with a bootloader that allows users to upload new code to it without the use of an external hardware programmer. This board has a 5V linear regulator and a 16 MHz crystal oscillator.

### B. Dual H-Bridge Motor Driver Circuit

The dual H-bridge motor driver circuit used here consists L293D IC as main driver. The Device is a monolithic integrated high voltage, high current four channel driver designed to accept standard DTL or TTL logic levels and drive inductive loads (such as relays, solenoids, DC and stepper motors) and switching power transistors. It uses two bridges; each pair of channels is equipped with an enable input. A separate supply input is provided for the logic, allowing operation at a lower voltage and internal clamp diodes are included. It takes digital signal as an input from the Arduino and gives digital output to the DC motors of the robot. It also amplifies voltage (from 5V to 12V) and current (from 40-50 mA to 250mA) per pin. So, overall power is amplified. Though the current amplification could be done by transistors in Darlington pair connection, which is less costly, the main purposes of using this IC are

- Unlike transistors it accepts bidirectional current and
- It can control two motors simultaneously.

There's a PWM input per driver so one can control motor speed. It runs at 5V logic. This holds good for motor voltages from 4.5V up to 36V. This works well for the 12V motors which have been used here.

### C. Integrated laptop Webcam

The Webcam used here is a basic webcam associated with the HP 430 notebook. This webcam provided necessary color tracking facility for this project.

#### D. DC Geared Motors

The DC geared motors serve the main function of this project. Two DC motors with metallic gear head were used to run the two back wheels of the robot. The motors used were of 12 Volt, 100 rpm rating. It generates 1.5 kg-cm torque, which is enough to drive the wheels of the robot. Its no-load current is 60 mA (maximum) and full-load current is 300 mA (maximum). By adjusting motors rotations, the robot was being able to move forward, backward, right or left.

#### E. DC adapter with header

A DC adapter of 12V output is used to give power to the Dual H-Bridge Motor Driver Circuit.

#### F. Metal Chassis

The metal chassis worked as the main building block of the motor. It holds the circuit components on one back and the motors on the other side.

#### G. Two Wheels and One Free Wheel

The two motors drive two wheels differently at the same time. A free wheel is introduced for proper balancing of the motor.

### III. METHODOLOGY

Gesture is a movement of a body part especially the hand to express an idea. Here, in this project this gesture phenomenon is for giving command.

The whole system implementation could be divided into several steps. The possible functions of image processing were realized and tested in a Matlab GUI based image editor.

Image processing operations can be roughly divided into three major categories, Image Compression, Image Enhancement and Restoration, and Measurement Extraction. Image defects which could be caused by the digitization process or by faults in the imaging set-up can be corrected using Image Enhancement techniques. Once the image is in good condition, the Measurement Extraction operations can be used to obtain useful information from the image like color tracking. The use of MATLAB as a Digital Image Processing Tool has made the development of many applications which incorporate different Image Enhancement function very easy. Unlike coding programs the user of a GUI need not to understand the details of how the tasks are performed. GUI components can include menus, toolbars, push buttons, radio buttons, list boxes, and sliders, just to name a few. GUIs created using MATLAB tools can also perform any type of computation, read and write data files, communicate with other GUIs, and display data as tables or as plots. The fig. 1 illustrates the work of the GUI created for image enhancement. Here auto histogram adjustment process shows the resultant enhanced image.

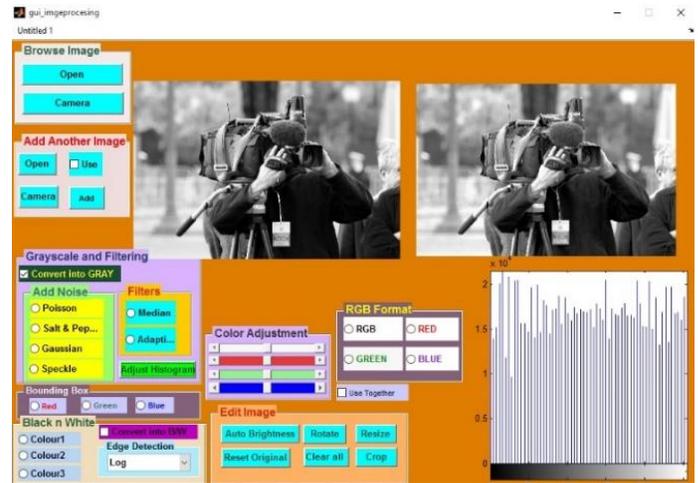


Fig. 1. GUI of Image Processing

This proposed system consists of two main hardware components: the computer which runs the Matlab, Visual Basic, and the Arduino Uno microcontroller board which is flexible, inexpensive, offers a variety of digital and analog inputs, serial interface and digital outputs. Arduino Uno controls the robot by controlling the action of motor driver circuit. Arduino Uno also enables the user to control the motor through voice commands. The computer communicates with the Arduino Uno microcontroller board through USB data transfer cable.

#### A. The Color Tracking Through Image Processing

Movement of Image frame with particular color is taken as an input and processed using Image Processing. This image processing is the fundamental part of tracking. The tracking system uses the Matlab coding. Matlab tracks the command direction and send the direction to the Arduino software which directly gives command to the Arduino Uno board. With the help of motor driver circuit, the Arduino controls the movement of the robot in the required direction as given by the command.

The Matlab program first opens the video capturing frame. This frame captures video through laptop associated webcam in RGB format for infinite time until user commands to stop. From this captured video continuous snapshots are taken out. This RGB snapshots are the main tools for the process done by Matlab.

The RGB images are flipped in both rows and columns to correct the effect of flipped image taken by webcam. Then the particular color (viz. Red, Blue, Green) upon which the Matlab color tracking program is based on, is extracted from it; the color is extracted from the RGB picture. Generally, this image contains dusty noise. So a median filter is used to filter out the noise from this image. This monochromatic image is

then converted to black and white. From this black and white image, the area, co-ordinates of the centroid and the bounding box containing the color used for giving command can be easily found out. The centroid contains the  $x$  and  $y$  co-ordinates, which are printed over the bounding box. The changes of  $x$  co-ordinate represent the movement of hand along  $x$ -axis i.e. in right or left direction and similarly the changes in  $y$  co-ordinate represent the movement of hand in upward or downward direction.

The tracking program functions this way: the program tracks the centroid and so the bounding box and according to which direction the centroid moves, the program sends commands to the robot to move in that direction via Arduino software. Our hand movement is never absolute in one direction. For example, if we move our hand in the right, some upward or downward movement occurs along with. But here the maximum change occurs in  $x$  co-ordinate and slight change take place in  $y$  co-ordinate. So a random threshold value of  $y$  co-ordinate (say 20) has been taken to omit the changes of  $y$  co-ordinate. So the Matlab will send command (exact direction to move horizontally along the right only) to Arduino. Similar pattern is used for other directions.

### B. Working of Arduino

The Arduino software programming takes the directions, which came from Matlab, as input. According to the case it runs the required function and sends accurate instruction to the motor.

With the Arduino software the Arduino Uno sends necessary information to the motor driver control circuit, which then controls the motions of the motors and thus the movement of the robot.

### C. Voice control

At the end of the work, a special feature of giving command to the motor has been introduced. This is voice command. The robot can accept voice commands through Microsoft Visual Basic 2008 Express Edition. Microsoft Visual Basic accepts voice commands and sends instructions to the Arduino software. This Arduino software then recognizes the input and sends the required instruction to the Arduino Uno board, which with the help of motor control circuit runs the motor in the direction accordingly.

## IV. EXPERIMENTAL EVALUATION

The whole system that has been developed is given in fig. 2. It shows us different parts and connections of the circuit. The input pins of dual H-bridge motor driver circuit are connected to Arduino's digital pins whereas the output pins are fed to the two motors. This can be seen in the bottom part of the system which is given in the fig 3.



Fig. 2. Top view of the whole robot

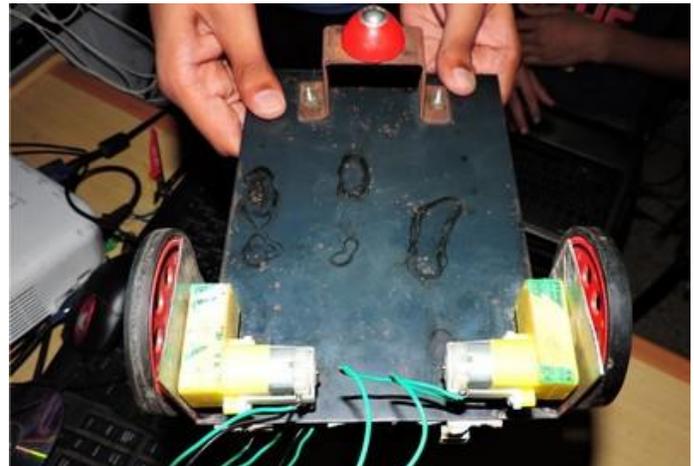


Fig. 3. The view of the bottom of the robot

Fig. 4 shows the input given to the Matlab for tracking of red color. The rectangular box in the picture is the bounding box and in it its area and the co-ordinates of the centroid is printed.

In the additional part of controlling the robot with voice command, the Microsoft Visual Basic has been used and it is shown in fig. 5, where the robot was given command to move forward.

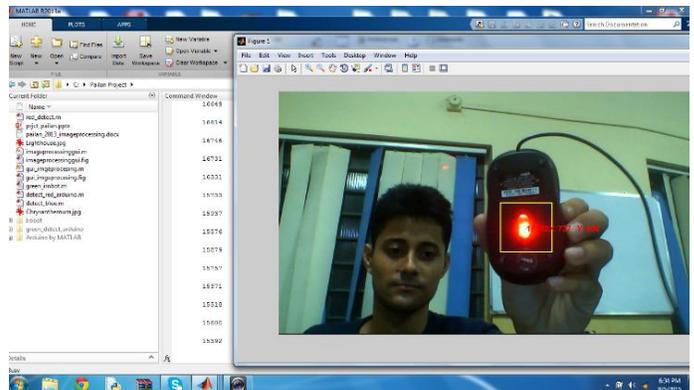


Fig. 4. Color tracking using Matlab

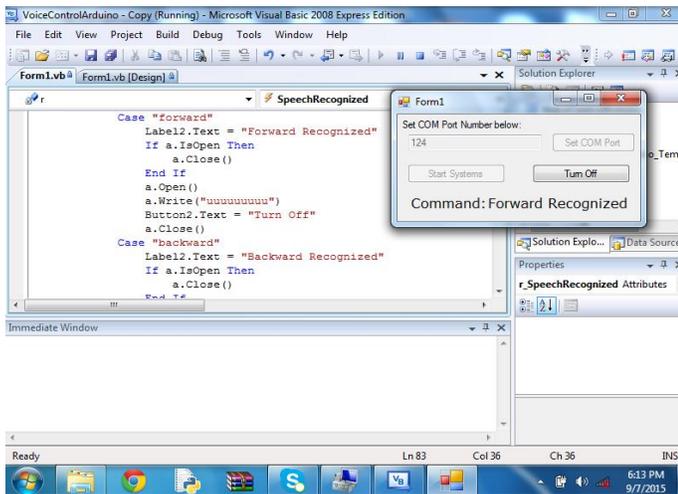


Fig. 5. Voice command using Microsoft Visual Basic

## V. DISCUSSION AND CONCLUSION

A low cost computer vision system that can be executed with low power in-built laptop web cam was one of the main objectives of this work, which has been implemented successfully. The system has been experimented with around 6 colors and the results achieved higher average precision.

From the results obtained, it can be concluded that the tracking algorithm is quite efficient. The main advantage of this algorithm is that even tilted colored objects can easily be recognized and analyzed. The colored object used to give the command for movement direction, if it is multicolored, only one specific color is detected. This overrules the possibility of interfering of other moving objects nearby, provided it is of different color.

However, there are several factors which limit the efficiency of the detection rate. Some of them are discussed below.

- Ambient lighting intensity vastly effects the outcome.
- the colored objects should be properly placed in front of the webcam so that the entire region is captured.
- Gesture made in this method involves only one color at a time and this reduces the number of gestures that can be sent within a definite time interval.
- Monochromatic color dependency affects the subsequent detection rates.

Whereas, the voice command recognition based implementation is much more simple and direct. It is independent of ambient conditions. Unless it gets the exact moving commands it does not send signals to the robot. So, the ambient noise leaves no interference.

This system has many potential application in many fields, from home applications to big industrial applications. Being at one particular position, many robotic applications at a time can be controlled to carry out many jobs. The type of jobs are somewhat specific like lifting things up, moving and putting them down, pushing things, stacking things,

positioning things and vehicles. In industries, in harsh environments, instead of sending people, the robots can be sent to carry out jobs like pulling the hot items and put them outside to be cooled down.

With a commercial hardware package and wireless facility, this system could be used in homes and in some small industries to control robots to move appliances.

## VI. FUTURE WORKS AND APPLICATIONS

Controlling a robot, in real time, through the gestures is a novel approach and its applications are myriad. The use of service robot to domestic users and industries in the upcoming years would need such methods extensively. The approach has huge potential once it gets further optimized, as its time complexity is higher, with the help of hardware having better specifications.

There are several possibilities for future renovation. The tracking algorithm could be implemented through GUI. This would make it more user-friendly and simultaneously, a robust interface for interaction. The speech recognition system can be used in the voice command algorithm to make it user dependent for more security. Implementing this concept on an i-Robot create would be more efficient and versatile.

Coupling this system with a wheelchair can be very effective for physically challenged people, whose legs are not working but the limbs, to become independent for moving from one place to other.

On the other hand, with camera facilities and more intelligent algorithm to work independently, this system can be very useful in big industries for carrying out some jobs and to carry things and place it at specific place. This would require the wireless communication technology to free the robots from direct wired connection with the control system.

Use of more efficient wireless communication technique and a camera on the robot unit would improve the performance of the system to a great extent and can be incorporated in future applications like investigating lives of wild animals, investigating through narrow tunnels and recovering necessary items from a place under devastating fire etc. where human beings can not be present.

As a whole it can be concluded that the system has a huge scope of further research and application which can prove to be effective in various fields.

## ACKNOWLEDGMENT

We would like to thank e-school learning (ESL), Kolkata for assisting us with all necessary technical details and working platform to build this project successfully.

## AUTHORS' CONTRIBUTIONS

Somnath, Ankit, Debarshi contributed in designing the image editor in MATLAB, implementing it in the gesture controlled robot, in testing and writing the report. Soumit and Dipayan assisted in assembling the hardware parts of the robot and in data collection. Debasish and Sudipta designed the Voice control part in Visual Basic. Sauvik conceptualized the

problem, designed it and helped in making the final coding in MATLAB.

REFERENCES

- [1] Justyna Inglot, 'Advanced Image Processing with Matlab', in Bachelor's Thesis Information Technology, May 2012, Date of the bachelor's thesis 07.05.2010, Mikkeli University of Applied Sciences.
- [2] J. Guerrero, "Image Processing and Analysis with Matlab", International Conference on Electrical, Communications, and Computers, 2009. CONIELECOMP 2009, Cholula, Puebla, pp. xvi, 26-28 Feb. 2009.
- [3] P.R.V. Chowdary, M.N. Babu, T. V. Subbareddy, B.M. Reddy, V. Elamaran, "Image processing algorithms for gesture recognition using Matlab", 2014 International conference on Advanced Communication Control and computing Technologies (ICACCCT), Ramanathapuram, pp. 1511-1514, 8-10 May 2014.
- [4] T. Mahalingam, M. Mahalakshmi, "Vision based moving object tracking through enhanced color image segmentation using Haar classifiers", Conference on Trendz in Information Sciences & Computing (TISC), 2010, Chennai, pp. 253-260, 17-19 Dec 2010.
- [5] Luigi Lamberti and Francesco Camastra, "Real-Time Hand Gesture Recognition using a Color Glove," *Department of Applied Science, University of Naples Parthenope*, 2010, pp.321-344.
- [6] J.S. Kim, C.S. Lee, K.J. Song, B. Min, Z. Bien, "Real-time hand gesture recognition for avatar motion control," *Proceedings of HCI'97*, pp. 96-101, February 1997.
- [7] Chao Hy Xiang Wang, Mrinal K. Mandal, Max Meng, and Donglin Li, "Efficient Face and Gesture Recognition Techniques for Robot Control," CCECE, 1757-1762, 2003.
- [8] Jagdish Lal Raheja, Radhey Shyam, Umesh Kumar and P Bhanu Prasad, "Real-Time Robotic Hand Control using Hand Gestures", Second International Conference on Machine Learning and Computing, 2010.
- [9] M. Mahalakshmi, "Real time vision based object tracking using CAMSHIFT algorithm with enhanced color image segmentation", 2010 International Conference on Computing Communication and Networking Technologies (ICCCNT), Karur, pp. 1-8, 29-31 July 2010.
- [10] T. Said, S. Ghoniemy, G. Ismail, I. A. EL-Ghafar, 13th International Conference on Aerospace Sciences & Aviation Technology, ASAT- 13, May 26 – 28, 2009.
- [11] Gesture controlled robot using Kinect <http://www.e-yantra.org/home/projects-wiki/item/180-gesture-controlled-robot-using-firebirdv-and-kinect>