Gesture Recognition Based on Computer Vision on a Standalone System

Mrityunjay Sharma
&
Prabir Kumar Choudhury

Department of Electronics and Communication Engineering
National Institute of Technology Rourkela
Rourkela-769 008, Odisha, India

Gesture Recognition Based on Computer Vision on a Standalone System

Thesis submitted in

May 2015

to the department of

Electronics and Communication Engineering

of

National Institute of Technology Rourkela

in partial fulfillment of the requirements

for the degree of

B.Tech

in

Electronics and Instrumentation Engineering

by

Mrityunjay Sharma

[Roll: 111EI0261]

&

Electronics and Communication Engineering

by

Prabir Kumar Choudhury

[Roll: 111EC0169]

Under the guidance of

Prof. K.K. Mahapatra, ECE Dept.

NIT Rourkela

Department of Electronics and Communication Engineering
National Institute of Technology Rourkela
Rourkela-769 008, Odisha, India

Department of Electronics and Communication Engineering
National Institute of Technology Rourkela
Rourkela-769 008, Odisha, India

May 11, 2015

Certificate

This is to endorse that the work in the thesis entitled "**Gesture Recognition Based on Computer Vision on a Standalone System**" by **Mrityunjay Sharma** and **Prabir Kumar Choudhury** is a documentation of an authentic research work executed under my supervision and counsel in partial fulfillment of the requirements for the award of the degree of Bachelor of Technology in Electronics and Communication Engineering. After successful intrinsic plagiarism check, it is being declared that, this thesis work or any part of it has never been proposed for any academic degree or rewards elsewhere.

Prof. K.K. Mahapatra
Head of the Department
Department of Electronics and Communication Engineering
NIT Rourkela

Acknowledgment

We express our heartfelt and sincere regards towards our guide Prof. K.K. Mahapatra for his able and accomplished guidance and motivation throughout the course of the project which served as an incentive to keep our work within time. We would also like to mention the supervision and guidance of Asst. Prof. Ayas Kanta Swain and PhD. Scholar Sudeendra Kumar who act as stimulus to our work.

We would also like to thank our fellow friends and juniors who helped us in time of need with their illuminating ideas and played a role for successful completion of our thesis work. Finally, we would be fully indebted to our parents for helping us survive all the stress for the past year and not letting us give up.

Mrityunjay Sharma

Prabir Kumar Choudhury

Abstract

Our project uses computer vision methods gesture recognition in which a camera interfaced to a system captures real time images and after further processing able to recognize the gesture shown to be interpreted. Our project mainly aims at hand gestures and after extracting information we try to produce it as an audio or in some visual form. We have used adaptive background subtraction with Haar classifiers to implement segmentation then we used convex hull and convex defects along with other feature extraction algorithms to interpret the gesture. First, this is implemented on a PC or laptop and then to produce a standalone system, we have to perform all this steps on a system which is dedicated to perform only the given specified task. For this we have chosen Beaglebone Black as a platform to implement our idea. The development comes with ARM Cortex A8 processor supported by NEON processor for video and image processing. It works on a clock frequency of maximum 1 GHz. It is 32 bit processor but it can be used in thumb mode i.e. it can work in 16 bit mode. This board supports Ubuntu, Android with some modification. Our first task is to interface a camera to the board so that it can capture images and store those as matrixes followed by our steps to modify the installed Operating System to our purpose and implement all the above processes so that we can come up with a system which can perform gesture recognition.

Keywords: *Hand gestures, Haar classifiers, Feature extraction, ARM processor, OS*

Contents

List of Figures

Chapter 1

Introduction

1. Overview

Many people use computers in their free time. Special I/O devices have been invented with the aim of providing a communication between humans and Electronic Systems. Keyboard and Mouse are the two most common HID devices. With passing time and new technology coming up it can be seen as an attempt to provide the computer with more intelligence and providing humans the power to be able to perform more sophisticated tasks and communicate with the Electronic System. Successful human computer interfaces were created due to the result oriented efforts made by computer professionals.

The sole aim is to enable computers understand and interpret human language and cultivate an interface to interact with the humans. Enabling a computer to understand human speech, moodily expressions and human gestures are some instances in this regard. Gestures can be defined as the non-verbally exchanged information delivered through other parts of the body. A human being can perform uncountable gestures at a time. Since gestures shown by humans are sensed through vision, great interest has been put into this matter by computer vision researchers.

The project aims to recognize human gestures on a standalone system. A complex programming algorithm is required to code these gestures into machine language. An overview the recognition system is given for a better understanding. Hand gesture recognition is of utmost value for HCI, because of it widespread uses and applications in virtual realm and sign language interpreter. Previous work suggests that, old-fashioned vision-based hand gesture systems are still way behind many real-life application serving purpose.

The images which are captured have a quality which is sensitive to the present conditions of lighting and noisy backgrounds, because of the limitations of the optical transducers. Thus it is very hard detect as well as track the gesture shown robustly and with great precision. This limitation hugely affects the performance of hand gesture recognition based Human Machine

Interface. An efficient way to make hand gesture recognition more robust is to use multiple sensors for capturing moving features e.g. data glove method. But unlike optical transducers, such devices are more robust and are not affected by lighting conditions or noisy backgrounds. However, the inconvenience to user in wearing a data glove prevents its wide use. It also sometimes requires calibration, which may hinder the naturalness of the hand gesture. These data gloves are also quite expensive and not easy to handle. And therefore a result, it has not gained much popularity.

With the development of inexpensive and high depth cameras such as Xtion PRO LIVE sensor and Kinect devices, hand gesture recognition has found new opportunities. There still exists an open problem in using these types of high end cameras. Xtion Pro LIVE produces good results in tracking large objects, e.g., a human hand which occupies a very small portion of the image with complicated movements. In such cases, segmentation and may significantly affect the recognition process. Contour detection have also suffer from significant distortions that are local, along with large variations. Due to low resolution and inaccuracy of these cameras, if small objects are close to each other in the object they may be indistinguishable and hard to process.

An observation has been made in this regard proving that classic shape recognition methods, such as skeleton matching and shape contexts methods are not sufficient to perfectly recognize and interpret the contour with huge distortions. In a visible context, recognizing noisy shapes is hugely challenging especially in the case of a large number of gestures.

Gesture Recognition in a standalone system has been introduced recently and is a topic of research and huge applications. In this particular case, an AM335 SoC (System on Chip) is used, which has an ARM cortex A-8 series core and NEON co-processors, to develop the HID interface on the system. This system finds various applications in portable fields such as flying copters to small gadgets capable of gesture recognition.

2. Motivation

Humans, the gift of Nature, have an inimitable ability and that is “The power to speak, to communicate and connect through sounds. This speech or voice allows them to interact, communicate and understand each other. Thus spoken language becomes one of the primary qualities of humans. But unfortunately, everybody does not possesses this ability mainly due to the absence of a particular sense, i.e. hearing *(Daniel Capilla).*

In India, with a billion population there are around 6million deaf people reported according to NAD and sign language is the only alternative means of communication between deaf people and normal humans. It is very hard for normal people to communicate with others without an interpreter being involved. So here we thought of developing a system that have the power to interpret symbols used in sign languages into audio or plain text that can bring out a solution to this problem. Various interactive training programs can be conducted online or offline which will help to narrow down the gap between normal people and deaf people.

3. Objective

The major objectives of this project are:

- To develop a system capable of recognizing different types of hand gestures that aid human in their life activities.
- The developed system must be a standalone system consisting of a high speed processor, an optical transducer, and an output means of communication or display.
- The system must be fast responsive delivering output in real time with minimal latency and sufficient compensation.
- Finally, develop a product or a portable machine that will help the mankind to communicate or to improve their efficiency in their day-to-day activities.

4. Thesis Organization

The thesis is organized as follows.

Chapter 2, describes the hypothetical and methodological findings or contributions by various scholars on this kind of similar topic, or the literature survey.

Chapter 3, describes the details of the hardware used.

Chapter 4, deals with the technique and the code flow used in the project.

Chapter 5, describes the setting up and interfacing of the hardware.

Chapter 6, displays the result obtained after successful running of the hardware.

Chapter 7, sheds light on the various applications and uses of this project.

Chapter 8, concludes the project with the proposed future work.

Chapter 2

Literature

1. Literature

There has been imitations on the research on small scale systems which are able to recognize a full sign language. A glove-based gesture recognition system was designed by Christopher Lee and Yang sheng Xu that is able to recognize 14 letters from the hand alphabet that forms the basis of sign language, can learn new gestures and is able to update the present model of each gesture in the system on the fly, with a frequency 10Hz. Many devices for the detection of gesture such as advanced glove devices have been designed. They are the Sayre Glove, Power Glove and Dexterous Hand Master. The mostly used and successful glove based detection is the VPL Data Glove as shown in the figure 1.

Figure 1 VPL Data Glove

Developed by Zimmerman in 1970's this data glove is a prototype which is based on the patented technology of optical fibre sensors placed along the back of the hand of the fingers. Starner and Pentland also developed a gesture recognition system based on glove-environment system which is capable of recognizing around 40 signs derived from the American Sign Language (ASL) with a rate of 5Hz.

2. Sign Language

Nowadays, one can find a wide number of sign languages all over the world (more than 50) and almost every spoken language has its respective sign language. American Sign Language (ASL), Mexican Sign Language (LSM), French Sign Language (LSF), Italian Sign Language (LIS), Irish Sign Language (IRSL), British Sign Language (BSL), Australian Sign Language (Auslan), German Sign Language (DGS), Indian Sign Language (ISL), and Spanish Sign Language (LSE) are just a few of them. The American Sign Language is presently the most widely used and popular sign language among various others. The goal is to provide the reader with a basic knowledge about the sign languages. This section is not going to get into details of a single sign language because each one has its own rules. The following section will attempt to give a general description of the shared characteristics among the different sign languages: origin, phonology, and syntax (for the last two, contains a easy-to-understand description). By doing so, people who are not familiar with them will realize how complex it would be to design a whole Sign Language Translator and why the decision to simplify the system without taking into account these characteristics was made in the version of the system introduced here.

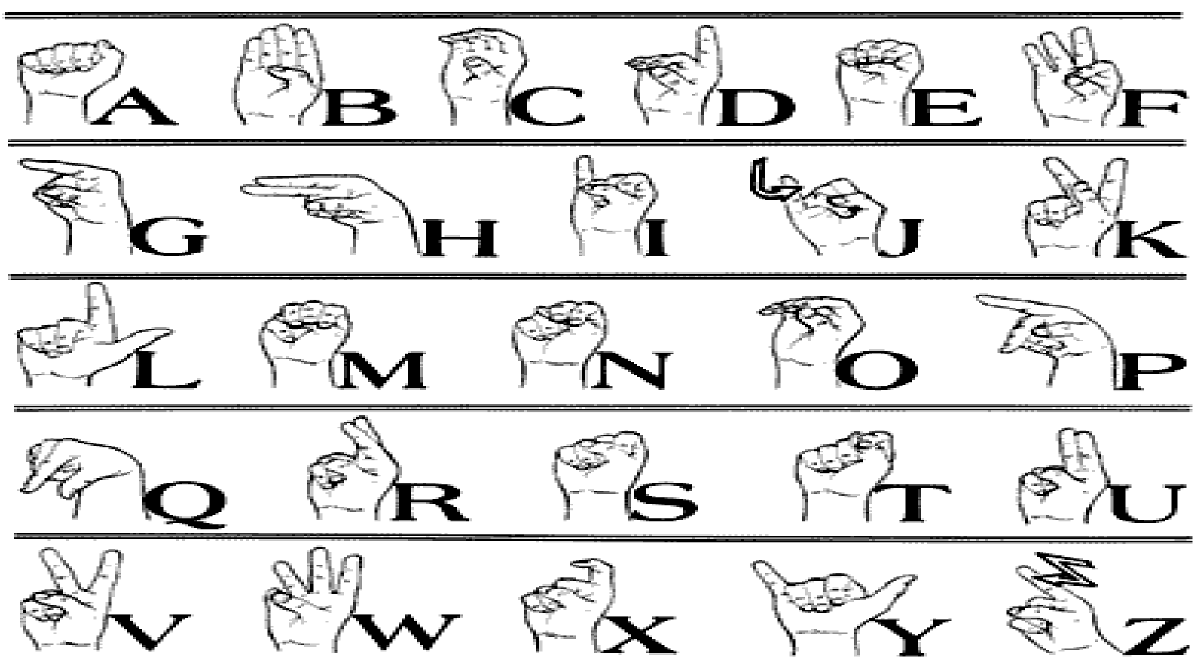

Figure 2 Sign Language

1. Origin of sign language

Sign language is mainly taught to deaf people, but its origin dates from the beginning history. In fact, gestures are the native way that kids have to express their feelings until they learn spoken language. Moreover, several hearing communities have used various sign languages to communicate with other ethnic groups that use entirely different phonologies (e.g. American Indians from the Great Plains). The starting real study of sign languages is relatively younger compared to spoken languages.

It dates from 1960, but today there is not an exact definition of their grammar. There is not yet a tradition in the use of a common transcription system that let us guess how young these disciplines are. There is an obvious quantitative and qualitative advance since the beginning of sign language linguistics, but there are still some methodological problems like the definition of a tool to transcript any sort of sign language.

2. Phonology

In spoken language, the phonology denotes the study of physical sounds present in human speech (called phonemes). Similarly, the phonology of sign language can be defined. Instead of sounds, the phonemes are considered as the different signs present in a row of hand signs. They are analyzed taking into account the following characteristics:

- Configuration: Hand shape while doing the sign.
- Orientation of the hand: Where the palm is pointing.
- Position: Where the sign is being done (mouth, forehead, chest, shoulder).
- Motion: Movement of the hand while doing the sign (straightly, swaying, circularly).
- Contact point: Dominant part of the hand that used to touch the body (palm, fingertip, back of the fingers).
- Plane: Where the sign is being done, depending on the distance with reference to the body (first plane is the one with contact to the body and fourth plane is the most remote one).

- Non-manual components: Refers to the information provided by the body (facial expression, movements of the shoulders or lip movements). For example, when the body leans front, it use to express future tense. When it is leaned back, expresses past tense. Also, non-manual signatures show grammatical in-formation such as question markers, negation or localization, conditional clauses, and relative clauses.

3. Morphology

Languages that are spoken have both in inflectional and derivational morphology. The former refers to the modification of words to express different grammatical categories such as tense, grammatical mood, aspect, person, number, grammatical voice, gender, and case. The latter is the process of forming a new word based on an existing word Sign languages have only one type of morphology which is derivational. Hereafter the main parameters those deal with the morphology are summarized:

- Degree: Known also as mouthing, it's the action to make what appear to be speech sounds to give emphasis on a word. For example, the sign "man tall" express "the man is tall". If this signature comes with a syllable "cha", the phrase becomes "that man was enormous".
- Reduplication: Repeating the exact sign several times. By doing so, the sign "chair" is prepared by repeating the verb "sit".
- Compounds: When a word is expressed as the combination of two different words. For example, the verb "agree" is composed by the verbs "think" and "alike" (one sign is executed just after the other).
- Verbal aspect: Verbs can be expressed in different methods so that we can say "to be sick", "to be often sick", "to get sick", "to be continuously sick", "to be sickly", "to be very sick", etc. Many of these include reduplication.
- Verbal number: To express singular or plural verbs. Reduplication is also used to express it.

Chapter 3

Hardware

1. Logitech HD (Optical Vision) C270 Camera

Figure 3 Logitech (optical Vision) Camera

The specifications of the camera used are:

Hardware and Operating System Specifications

The operating system must be Windows Vista or any version of Windows 7 (32-bit or 64-bit) or the different versions of Windows 8.

- The processor speed must be a minimum of 1 Giga Hz.
- An internet connection should be present for successful operation and updates.
- The ram required is 512 MB RAM or higher memory.

- An USB port (preferably 2.0) must be present.
- There should be a minimum of 200 MB hard drive space.

To enable the HD features in the above said camera the requirements are:

- The processor must be an Intel processor running at 2.4GHZ.
- The ram required is a minimum of 2 GB.
- Data speed must be 1MBps or higher.
- The minimum hard disk space required is200 MB.
- The minimum screen resolution should be 1280 x 720.
- A high speed USB port must be present.

Specifications Pertaining to Technical Details:

The specifications are:

- A system with High Definition facility of video calling is recommended preferably with resolution (1280 x 720) pixels.
- The maximum resolution of the photos is up to 3 Megabytes.
- The camera hardware is based on a liquid crystal based Logitech copyright technology.
- A USB is present to support high transfer rates.
- Noise is reduced with an inbuilt code for common mode rejection.
- The mechanical design is so made to fit to all portable devices.

The software provided by the Logitech support center has the following specs:

- It provided interface for various functions such as pan or zoom up to 300%.
- It can capture HD photos and can render videos.
- Any moving object can be detected.
- It can track human faces.

2. Beaglebone black

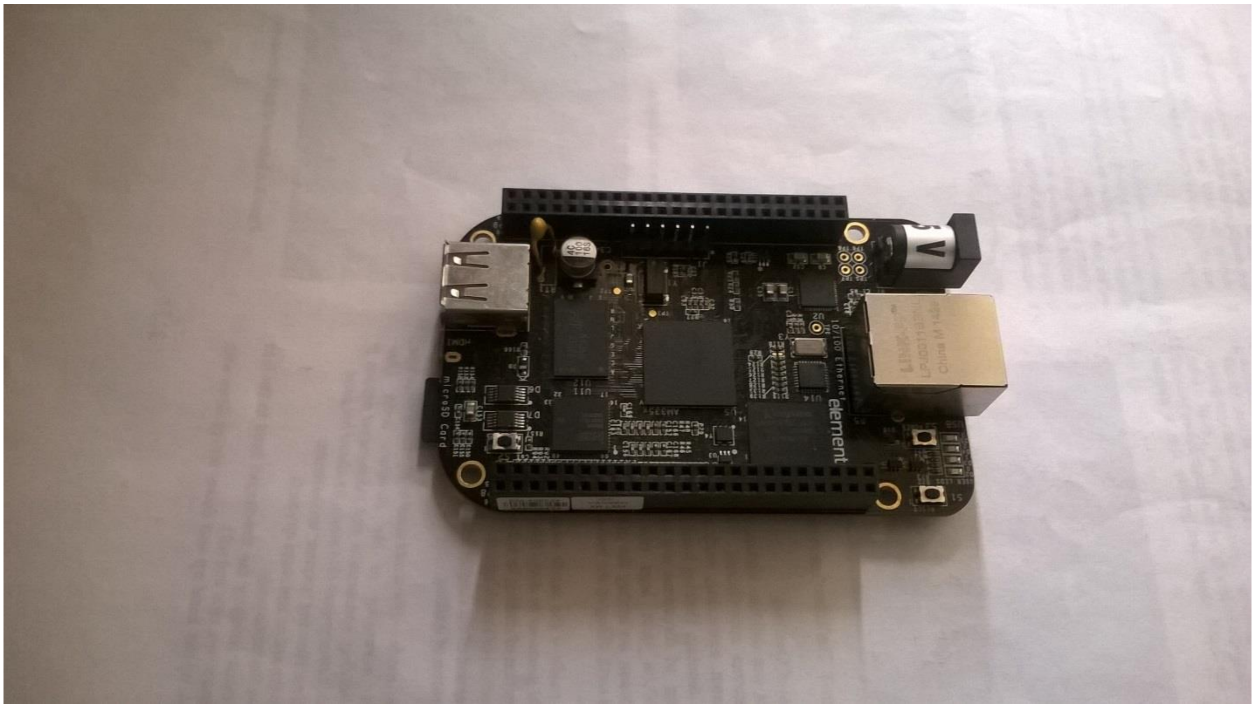

Figure 4 Beaglebone Black

Texas Instruments a high end corporation focusing on VLSI and embedded technology has come up with a very cheap single board computer commonly known as the SBC. It includes Panda boards, Beagle Boards but the one that has acquired mass attention is Beagle Bone. Its popularity is due to the fact that it is a SBC with very super performance and the cost is very less compared to other embedded boards thus can be used for students and individuals. At the heart of the SB lies a System on Chip or SOC. It has a higher performance RISC processor namely ARM architecture and the version is Cortex A-8 which is built for high end media processing applications. Compared to other open source boards such as Arduino it has very high performance and has the capability to support an OS owing to the fact that it has an MMU chip on its board.

Chapter 4

Technique and Code flow

1. Color Models

There are various color models used to describe a particular image. An image is a matrix of data or to say precisely pixels which contain a 3 dimensional value. Now in a 3 dimensional space a vector of points can be represented in 3 precise ways depending on the coordinate system. They are:

- Cartesian System
- Cylindrical System
- Spherical System

• **RGB:** It is the Cartesian coordinate's represented by 3 vector points in a 3d space. The point's represent the colors which are primary namely green, red, and blue. The best feature of this color space is its simplicity and the ease to process the data. But it has great disadvantages too in the field of image processing namely where luminance needs to be separated or chrominance.

• **HSV (Hue, Saturation, and Value):** It is the cylindrical coordinate and is represented by a radius and height and an angle. Hue is expressed with a color which is dominant and might be either red, green or purple in a particular area. The intensity of colorful in an area is given by saturation which is proportional to the brightness in that area. This model is very helpful for computer vision techniques and can discriminate between luminance and chrominance. This model fails to give satisfactory output in case of low illumination.

• **Y-Cb-Cr:** It is the spherical coordinate and is represented by the radius of the sphere and two angles that define the sphere in 2 planes. It is represented by Y, Cb and Cr. This model is also optical vision friendly and can easily separate segments in an image.

2. Segmentation

The first step towards recognizing gesture mainly hands or facial expressions is segmentation. After proper separation of the region of interest or in the case of gesture recognition hand other feature extraction techniques are adopted to understand the gesture shown by the person. To efficiently track a hand in a frame under any illumination efficient algorithm must be used to segment the Area of Interest.

The simplest form of segmentation is image thresholding where pixels below a certain level are excluded and thus the Region of Interest is separated from the frame. Presence of a skin like background also creates problems in skin color segmentation. Thus multiple feedback algorithm must be used to minimize background effects.

In this regard a background subtraction algorithm must be used which differentiates a moving object such as hand to a stationary background and is combined with skin segmentation to produce an efficient output.

3. Hand Segmentation in HSV Space

A innovative way to approach the algorithm for Image segmentation has been developed. In the below mentioned approach, segmentation based on color was done in the HSV color space.

The separation of H, S and V was done using these equations.

$V = \max\{R, G, B\}$

$\delta = V - \min\{R, G, B\}$

$S = \delta/V$

To obtain value for hue following are the cases

(i) If R=V ,then we can say $H=1/6(G-B)/\delta$

(ii) If G=V then we can say $H=1/6(2+ (B-R)/\delta)$

(iii) If B=V then we can say $H=1/6(4+(R-G)/\delta)$

The input image of green color samples was passed to the algorithm and from H-S histogram the H range = [0.4 0.55 0.6 0.6] and S range = [0.2 1.0] were experimented for segmentation. Algorithm could able to subtract dynamic background. Skin color samples needed to be passed to the algorithm for skin color detection. The drawback of this algorithm was training samples of the color need to be stored. It was sensitive to little variation in color brightness.

Figure 5 Segmented Image

4. Contour Detection

A fundamental step in image detection is edge or contour detections. The main aim in this regard is to detect any sharp change in pixel values which helps in determining the edge or the boundary of the frame.

There are 4 distinct approaches to the edge detection problem:

- Gradient and difference based operators
- Template matching
- Edge fitting
- Statistical Edge detection

Edge detection is only the first stage of the boundary based segmentation process. To aggregate these local edge elements, which are a relatively featureless representation, into structures better suited to the process of interpretation. This is normally achieved using process such as edge thinning, edge linking, gap filling and curve segment linking in order to generate a distinct, explicit and unambiguous representation of the boundary. There are several techniques for boundary detection and they varying the amount of knowledge or domain dependent information and that is used in the grouping process.

After segmentation contour detection is applied after which multiple small contours are detected and drawn using draw tool. Among these small contours the largest is found out and is assumed to contain the hand which is then smoothened out to best detect the hand features.

5. Convex Hull

In this type of feature extraction method after an efficient segmentation and contour detection the image is now converted to binary image and then this method is applied.

In this method, convexity and concavity defects are found using convex hull algorithms that find out these defects. These defects are the points where the image (binary) takes a sharp turn and is pointed out by placing green and red circles around these defects in the given frame. Counting the number of defects will give us information about the feature and the placement of these defects.

In hand gesture recognition convex hull is used to extract features from the hand. The convexity defects and the concave defects are determined by the convex hull process which gives the output as a set of points. These set of points are stored in a vector which is used to determine the feature. The red and blue dots mark the defects in the image shown below.

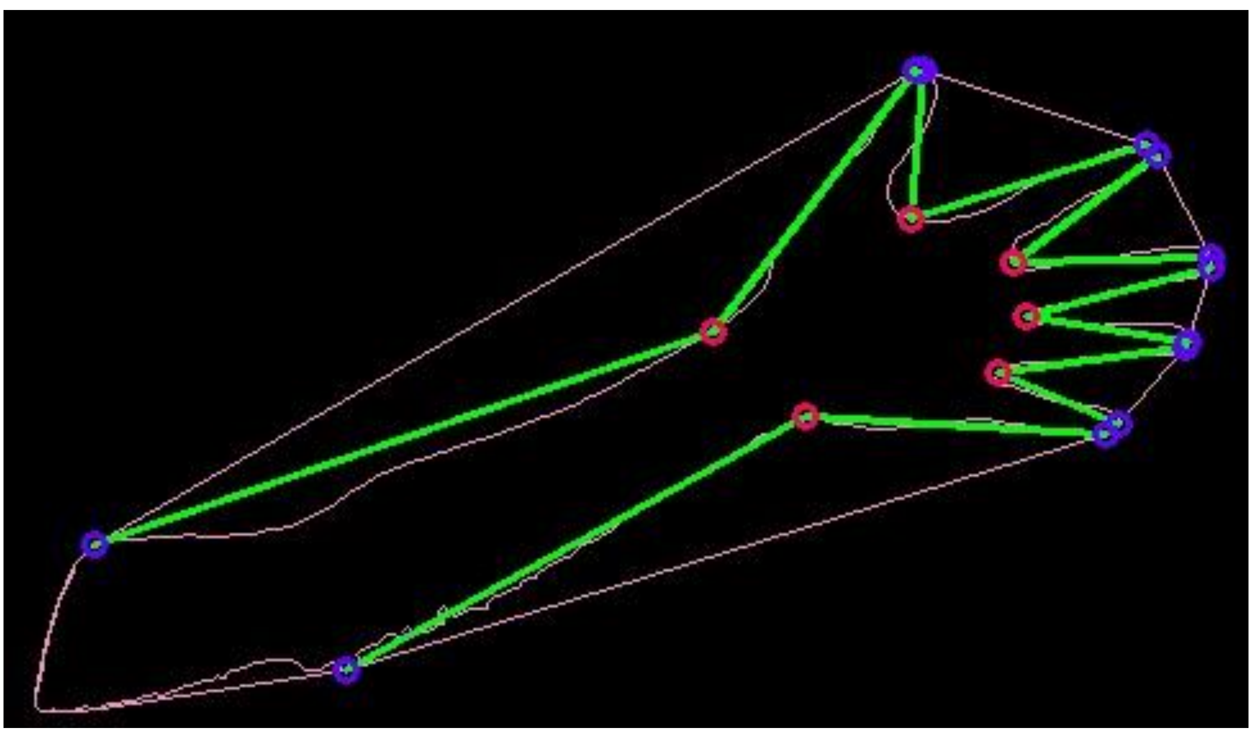

Figure 6 Defects

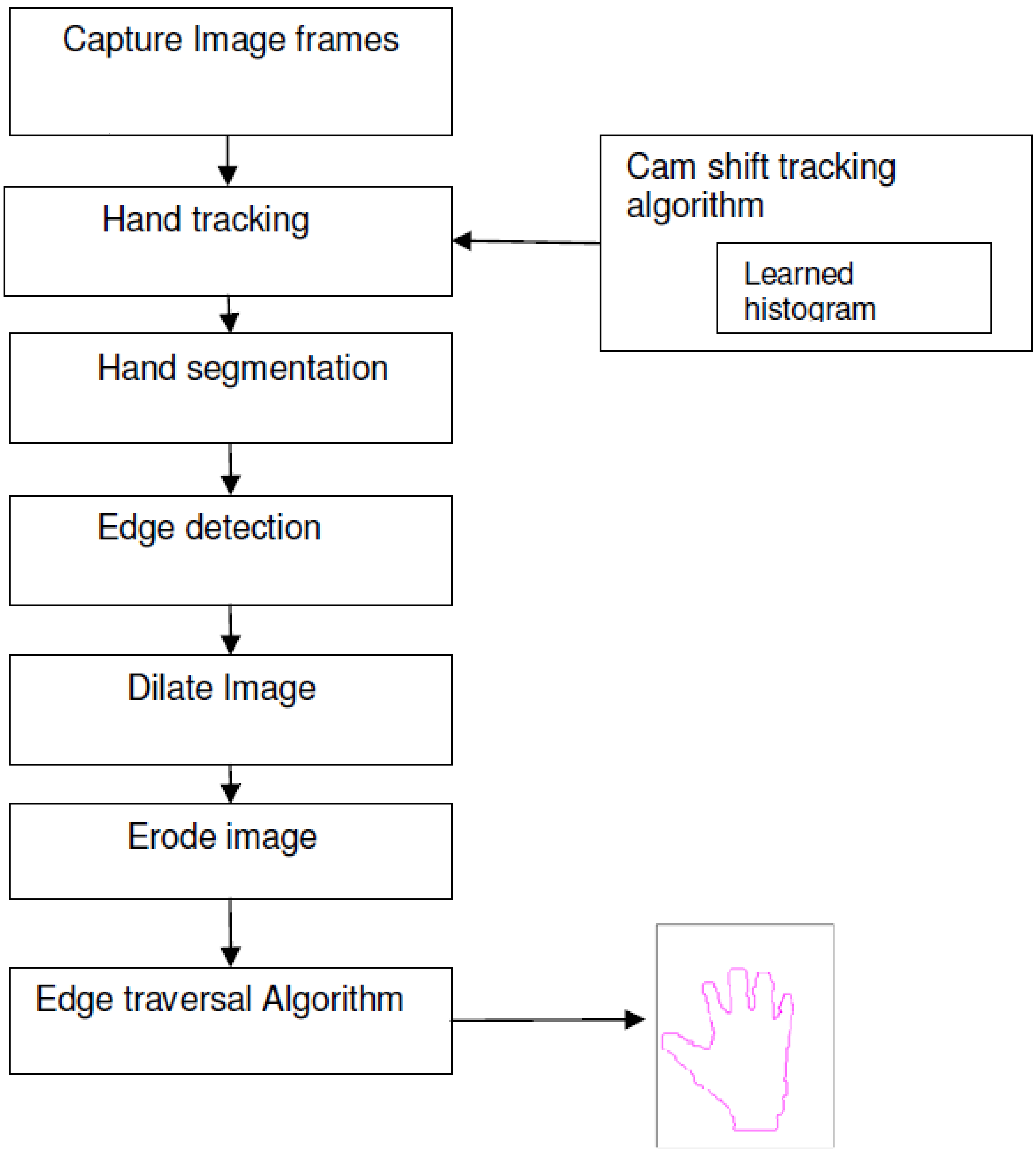

Figure 7 Architecture Design

6. Haar Classifiers

The image features obtained digitally used in recognition of objects are called **Haar (like) features**. It has acquired such a name because of its resemblance to Haar Wavelets and also because it was first time used in face detection.

Rectangular regions at a specific location adjacent to each other are considered in a detection window are summed up. The intensities of the pixel are summed up and the difference is calculated between the sums. This data is then used to find out the difference and then is categorized and subsections of an image is formed.

There are many types of haar windows such as 2 rectangle window or 3 rectangle window. The haar feature is a type of learner which is very weak in nature or can also be called as a weak classifier. Because of this very large number of features are required to describe a particular image with good accuracy. These features are then trained using multiple sample images to produce a trained classifier that now can be used to successfully used to detect an object in the particular frame.

The main advantage in using haar classifiers is its efficiency and its speed. The use of integral images reduces the time complexity and thus it can be calculate din a very short duration of time.

Chapter 5

Setup and Implementation

1. Setting up Beaglebone Black

The process involved to boot up beagle bone from an SD card (micro)

- The first step is to partition the SD card in the OS Ubuntu.
- Then the storage device is selected.
- The volume is unmounted.
- The partition is deleted.
- Two partitions are then created with names “boot” and “rootfs” that have the following file extensions “fat” and “ext3”.
- The boot partition is marked as bootable.

The next step in this regard is to cross compile the tool chain.

1. The steps involved in cross compiling are:

   Commands:
   ```
   #!/bin/sh
   export PATH=$PATH: /home/<jay(user)>/<path>
   export ARCH=arm
   export CROSS_COMPILE=arm-cortexa8-linux-gnueabihf-
   ```

2. In the first partition “boot” MLO, and uboot image is copied. Kernel image is also copied to the boot sector. Rootfs is also places in the second partition.

3. The booting process is started by pressing the switch on the board.

The hardware setup is described as follows:

Camera is connected to the USB 2.0 provided on it and Beaglebone Black is connected to PC using micro USB to USB cable.

To share internet over USB in beagle bone the steps are:

In BBB console

sudo su

ifconfig usb0 192.168.7.2

route add default gw 192.168.7.1

Chapter 6

Results

Various gestures were shown and recognized by the standalone system.

The gestures shown here are a few numeric examples that convey numeric message through gesture.

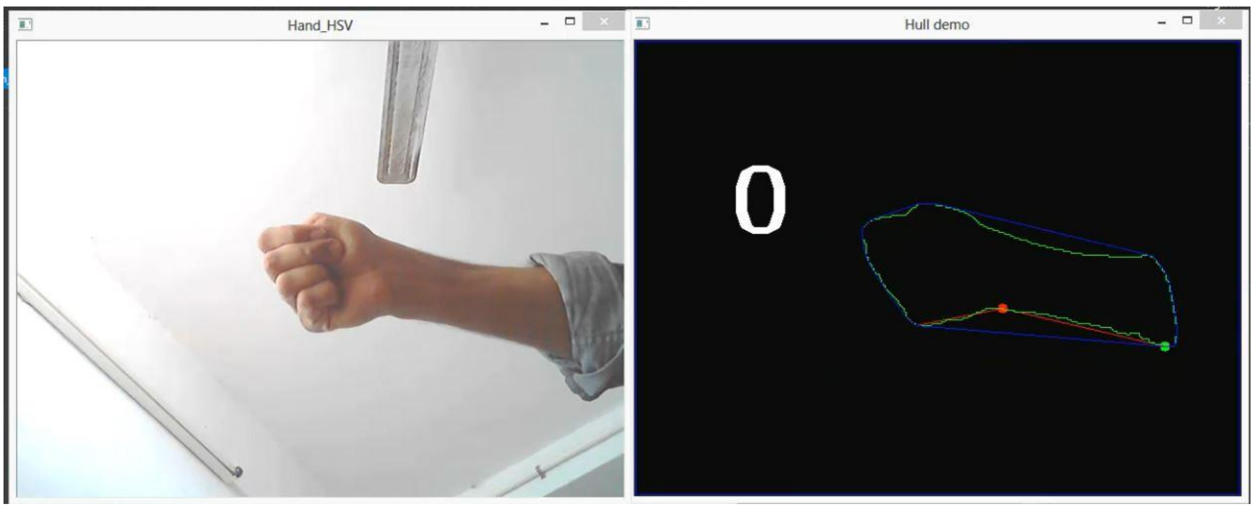

Figure 8 Result1

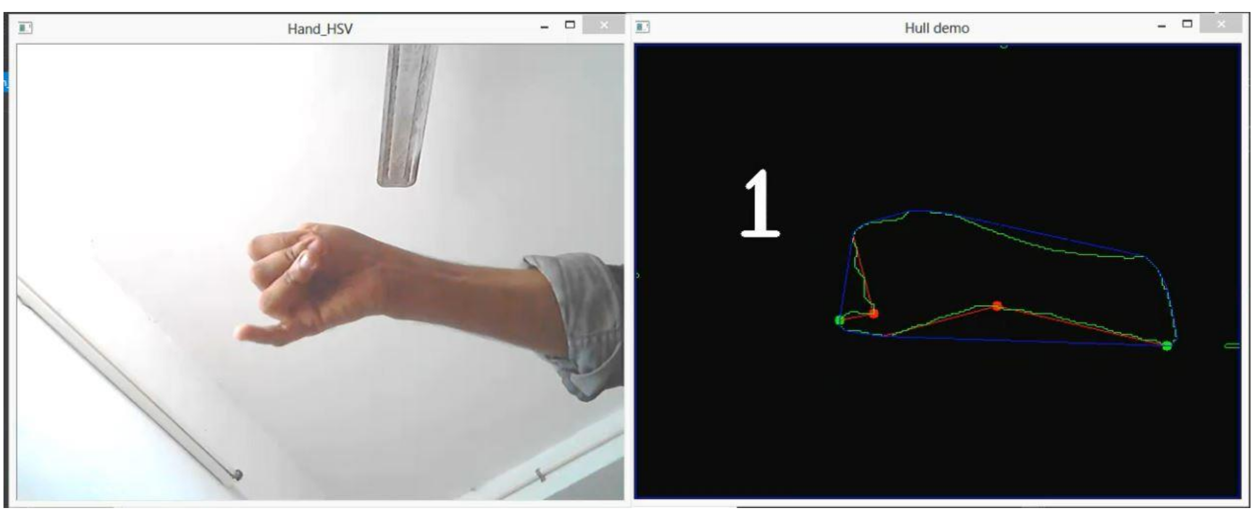

Figure 9 Result

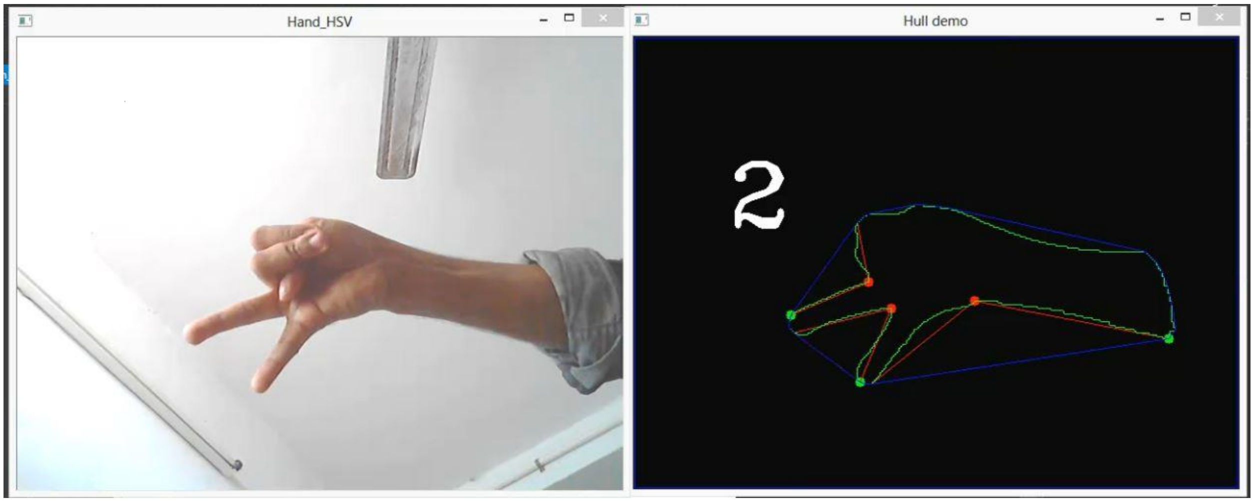

Figure 10 Result3

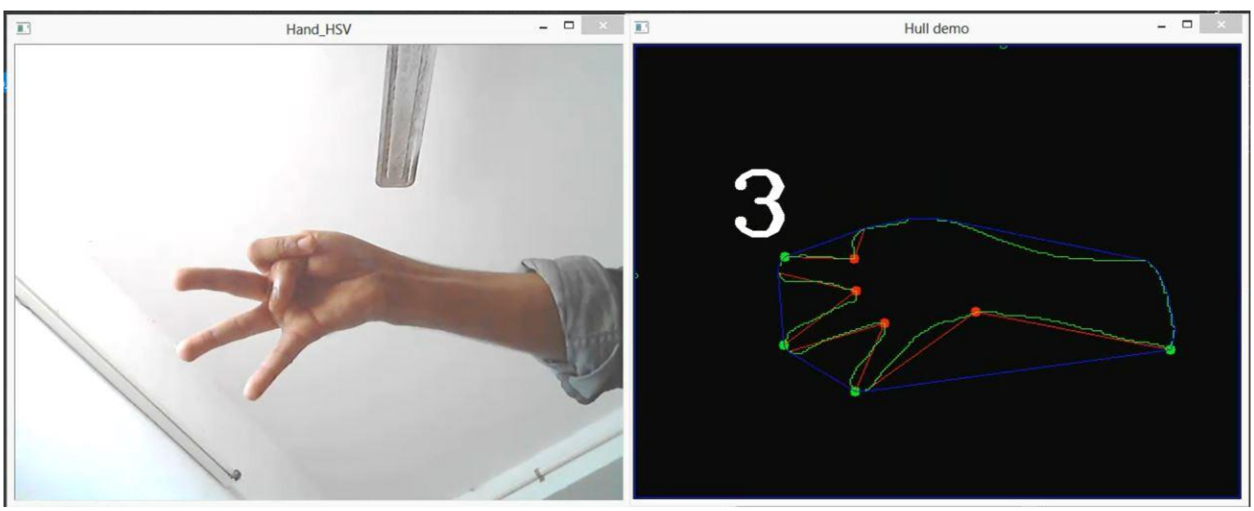

Figure 11 Result4

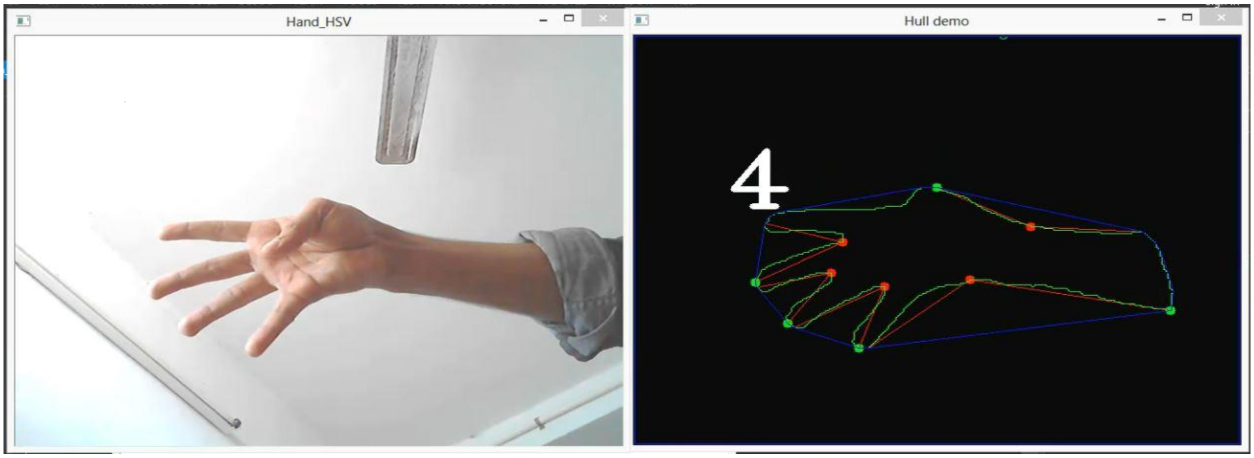

Figure 12 Result5

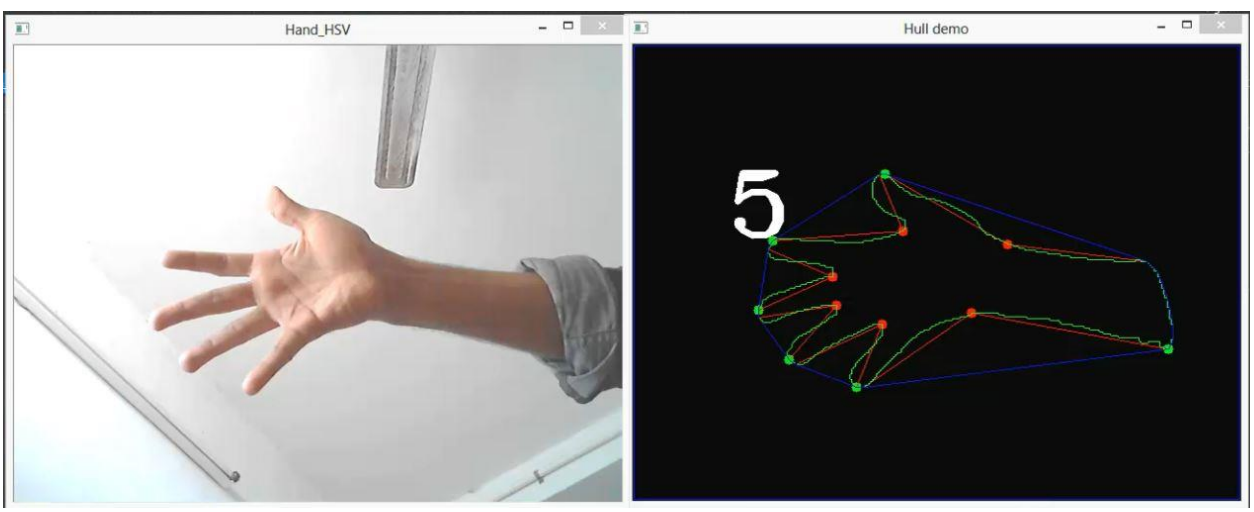

Figure 13 Result6

Chapter 7

Applications

1. Applications

Consumer Electronics: This is the field that will be mostly affected by this technology. Every consumable electronics that we use in our day-to-day life will be controlled by gestures. Is it overwhelming? Electronics like LCD display, washing machine, TV's to laptops and even to cameras, gestures will do the job for us.

Automated Homes: When it comes to our homes, we try to have it as comfortable as it can be. This technology just brings the right solution to all your problems. Now, we can easily control all the appliances in a room by just simple gestures making our life a lot easier after a hectic daily work.

Gaming: Gestures comes into play when we talk about computer games. It simply makes it user friendly and easy to control. Some common examples like Freeman traced a body position or hand motion to effectively control movements and the orientation of game objects which are interactive such as bikes.

Sign Language: The main application is sign language which will be like a communicative gestures. Sign languages being highly organizational, makes it suitable as beds for testing for implementing algorithms on optical vision. Concurrently, this is one of the best way to help the needy and disabled to interact with electronic systems. American Society of Sign language working for the deaf people is a remarkable example in the gesture technology literature.

Defense and Healthcare Sector: This is the most sensitive area that gesture technology targets. This will make these sectors efficient. For example, as a surgeon, being able to manipulate data in a touch-free way while treating someone will bring out enormous benefits.

Safe Driving: Gesture technology will bring a revolution to automobile sector where people will depend more on technology for safety. Isn't it fascinating that we could drive using gestures without actually driving manually? All sensor assistance, controls and parking can be done easily. This will bring new luxury, and high-end vehicles into the market. Now, coming year's gesture technology and touch less sensors will keep the market as well as consumers interested.

2. Analysis Parameters

The various analysis parameters are:

Dependability and Robust: In reality, visual signals or information may contains lot of noise, incompleteness, dynamic background, clutter, and occlusion. These systems should be independent from the user and should be robust against all the above mentioned factors.

Scalability: This technology should be generalized so that it could be scaled down and can be easily modified for various other applications. For example the system core for gesture technology should support all available desktop environments, robot navigation, Sign Language Recognition and also for virtual environment.

Computational Efficiency: Generally, gesture technology which works on vision based interaction requires real-time systems. The algorithms/techniques and hardware used should be fast, effective and cost efficient.

User's Acceptability: Error and malfunctions related to vision based technology should not be endured as it could led to fatal outcomes. When the system shows some error then it should not bring about much loss. Sometimes it is better to repeat some actions rather to allow the system to keep on making mistakes and taking wrong decisions.

Chapter 8

Conclusion and Future work

1. Conclusion

In current situation, numerous facilities and models are available for providing gesture inputs to any number of applications. But unfortunately there is a great demand in market to provide smart environment and input technologies corresponding to such advanced environment. Even though systems have been upgraded but bugs still prevails in input technologies. Currently there are many applications which provide gesture technology which can control systems from a distance without using any input devices like mouse or keyboard. Our above method showed some important steps of extracting raw gesture essential data from a noisy gesture data. These applications can be used to manipulate objects through hand gestures in some virtual environment is being suggested and implemented successfully in the present paper work providing a suitable effectual and user friendly HCI (Human Computer Interface).As these application provides the perfect flexibility to the consumers and more specifically to physically disabled users to express the gesture in accordance to their easy to use..

2. Future Work

This technology currently seems to be working fine with lot of applications, which can be considered as practicable and user friendly in comparison to the outdated HID's but it too has some demerits, which is less robust in tracking and recognition phase. Various attempts are taken to make these type of input systems robust for the users and having less constraints, and should work fine with hand gestures. But dependability and reliability can only be increased through more robust and efficient algorithms with will reduce noise and other dynamic motions. In this way we can get a system which will effectively convert hand gestures to commands for further processing. There is another aspect related to designing of gesture firmware which will be independent. This firmware should be useful in controlling gaming systems and also should be helpful in developing applications games and other applications dependent on the controlled through user defined gestures.

Bibliography

[1] F. W. C. L. Howe, "Comparison of Hand Segmentation Methodologies for Hand Gesture Recognition," *IEEE-978-4244-2328-6,* 2008.

[2] C. M. a. D. H. A. Elgammal, "Skin Detection – a Short Tutorial," Encyclopedia of Biometrics, Springer-Verlag Berlin Heidelberg, 2009.

[3] V. S. A. A. A. V. Vezhnevets, "A Survey on Pixel-Based Skin color Detection Techniques".

[4] C. W. W. Freeman, "Television control by hand gesture", IEEE international Workshop on Automatic Face and Gesture Recognition, Zurich," 1995.

[5] M. Z. Y. Guan, "Real-time 3D pointing gesture recognition for natural HCI," *Proceedings of the world congress on Intelligent Control and Automation, China,* pp. 2433-2436, 2008.

[6] Y. Y. L. D. A. Sepehri, "Employing the Hand as an Interface Device," *Journal of Multimedia,* vol. Vol. 1, 2006.

[7] Q. a. G. N. D. a. P. E. M. Chen, "Real-time vision-based hand gesture recognition using haar-like features," 2007.

[8] G. B. M. N. R. a. X. A. Erol, "Vision-based hand pose Estimation: A review," *Science Direct, Computer Vision and Image Understanding,* pp. 52-73, 2007.

[9] V. R. a. P. M. D. Comaniciu, "Real-time tracking of non-rigid objects using mean shift," *Computer Vision and Pattern Recognition,* vol. vol. 2, no. IEEE Conference on, 2:142-149, 2000.

[10] R. T. N. C. C. Burande, "Advanced Recognition Techniques for Human Computer Interaction," vol. vol. 2, pp. 480-483, 2010.

[11] Q. Chen, O. Ottawa Univ., N. D. Georganas and E. Petriu, "Real-time Vision-based Hand Gesture Recognition Using Haar-like Features," in *Instrumentation and Measurement Technology Conference Proceedings*, Warsaw, 2007.

[12] D. Molloy, "derekmolloy.ie," [Online]. Available: http://derekmolloy.ie/beaglebone/beaglebone-video-capture-and-image-processing-on-embedded-linux-using-opencv/.

[13] V. R. a. P. M. D. Comaniciu, "Real-time tracking of non-rigid objects using mean shift," *Computer Vision and Pattern Recognition,* vol. vol. 2, 2000.