A

PROJECT REPORT

ON

HAND GESTURE RECOGNITION FROM 3D IMAGES

BY:

ASHUTOSH S. VERULE

SHRIVARDHAN S. SURYAWANSHI

POONAM B. RAJPUT

UNDER THE GUIDANCE OF

PROF R.R.ITKARKAR

In Partial Fulfillment of

B.E. (E&TC)

University of Pune

[2013-2014]

JSPM's

RAJARSHI SHAHU COLLEGE OF ENGINEERING

Dept. of Electronics and Telecommunication Engineering

PUNE-33

CERTIFICATE

This is to certify that the project report entitled

"HAND GESTURE RECOGNITION FROM 3D IMAGES"

By

VERULE ASHUTOSH S.	B80373148
SURYAWANSHI SHRIVARDHAN S.	B80373136
RAJPUT POONAM B.	B80373063

is a record of bonafide work carried out by them in the department of Electronics & Telecommunication Engineering; under the guidance, in partial fulfillment for the award of degree in **Bachelor of Engineering in Electronics & Telecommunication Engineering** of University of Pune.

Prof. R. R. ITKARKAR
Project Guide

Prof. D. G. BHALKE
HOD (E&TC)

Dr. D. S. Bormane
Principal

JSPM's
Rajarshi Shahu College of Engineering, Pune

ACKNOWLEDGEMENT

We take this opportunity to thank the teachers and senior authorities whose constant encouragement made it possible for us to take up a challenge of doing this project. We express our deepest sense of gratitude towards our honorable Head of the Department, Prof. D. G. Bhalke for his constant encouragement for this work.

We are grateful to Prof. Mrs. R. R. Itkarkar for her technical support, valuable guidance, encouragement and consistent help without which it would have been difficult for us to complete this project work. She is a constant source of inspiration and information to us. We consider ourselves fortunate to work under the guidance of such an eminent personality.

We owe sincere thanks towards our Principal Dr. D. S. Bormane for his encouragement and guidance throughout the engineering course.

Last but not the least; we are thankful to our entire staff of Electronics and Telecommunication Department for their timely help and guidance at various stages of the progress of the project work.

Ashutosh Verule

Poonam Rajput

Shrivardhan Suryawanshi

CONTENTS

LIST OF FIGURES AND TABLES

PREFACE

Our project, "HAND GESTURE RECOGNITION FROM 3D IMAGES", consists of an unparalleled algorithm which recognizes the 3D gestures and according to the identified gesture, executes the respective windows application. Converting the hand gesture images to the computer manageable data is the most important part of this project. Features are extracted from hand gestures with the help of MATLAB. According to all the extracted features, a comparison of test image (input gesture image) is done with the train images in the data base and the desired output is produced.

A rapid emergence of 3D applications and virtual environments in computer systems has led to the rise of need for a new type of interaction device. Our proposed project helps this cause. Our project shows the possibility to perform non-trivial tasks from the recognized hand gestures using only a few well-known hand gestures, helping the handicapped negate the dependency on others to complete their mundane routine, selecting an alphanumeric value, replacing the human need of pressing buttons in elevators making them automatic, and other such fledgling applications.

The following chapters will give an idea about the project, its features, methodology, design, results, performance analysis, and the future scope.

Chapter 1: Introduction

This chapter describes the pertaining theory required for lucid understanding of the algorithm. It gives the introduction of sign languages along with the information about gesture recognition. Furthermore, the problem definition and objectives of "Hand Gesture Recognition From 3D Images" are explained in this section.

Chapter 2: Literature Survey

This chapter provides a brief review of the varied literature work studied for reference. A comparison has also been drawn between the different techniques used for gesture recognition on the basis of parameters like accuracy.

Chapter 3: Pre-requisites

The project demands to know a few concepts in advance before understanding the entire proposed system. Concepts like Eigen Values and Eigen Vectors are explained in this section. Along with these concepts, one must also know about the American Sign Language and its gestures before handling this project. This has also been briefed in this chapter.

Chapter 4: Methodology

The general block diagram along with its description is explained in this chapter. The different stages of operation in “Hand Gesture Recognition From 3D Images” are also explained in this chapter. The functionality of each block has also been explained briefly.

Chapter 5: Design

The algorithm designed for the proposed system of the project has been explained in detail in this chapter. It also consists of the flowchart describing the program flow of the system. Furthermore, techniques used for image segmentation, various functions used from MATLAB and so on, are also mentioned in this part of the report.

Chapter 6: Results And Analysis

The analysis of the system is done using images of each gesture with various backgrounds. The analysis considers the execution time required for processing each operation by the system, the accuracy of the system and its speed.

Chapter 7: Conclusions & Future Scope

No system can be designed as a perfect one in one go. Modifications of even the smallest magnitude, at even the smallest level can make a huge difference in the performance of the system. This chapter concludes the work with results and conclusions. It also throws a light towards the future work.

Chapter 8: Bibliography

This chapter lists down the various books referred during the course of the making of the project. It also lists the research papers referred for information regarding sign languages, gesture recognition, Euclidean distance, and so on.

ASHUTOSH VERULE

SHRIVARDHAN SURYAWANSHI

POONAM RAJPUT

ABSTRACT

A speedy rise in 3D applications and virtual environments in computer systems has led to the need for a new type of interaction device. “Hand Gesture” is a concept in human computer interaction, which has become popular in recent years, and can be used to develop such an interaction device. Hand gesture recognition is a relatively new field. In present times, much research is going on in the field of Artificial Intelligence in Natural language processing. Hand gestures, body postures are also the natural languages.

The use of hand gestures provides an attractive alternative to the cumbersome interface devices for human computer interaction (HCI). User’s generally use hand gestures for expression of their feelings and notifications of their thoughts. Gesture recognition, then, consists not only of the tracking of human movement, but also the interpretation of that movement as semantically meaningful commands.

Our proposed algorithm involves recognition of gestures from their 3D images by their acquisition, segmentation of the gesture information by removing background, extraction of the key features from the gesture using Eigen values and Eigen vectors, determination of the gesture enclosed within the image, and, executing the application as per the gesture identified.

Obviously a very efficient method is direct manipulation with bare hands, but, not every human is lucky enough to have hands. This project shows the possibility to perform non-trivial tasks from the recognized hand gestures using only a few well-known gestures, helping the handicapped negate the

dependency on others to go by their mundane routine, selecting any alphanumeric value without any key pressing, replacing the human need of pressing buttons in elevators making them automatic, and other such fledgling applications.

Our project also emphasizes on execution of windows based applications as per the recognized gestures, where applications such as MS word, MS excel etc. can be opened using defined gestures.

CHAPTER 1

INTRODUCTION

CHAPTER 1

INTRODUCTION

This chapter focuses on the fundamentals of sign language and gesture recognition systems.

1.1. Sign Language

A *gesture* is some specific motions of body parts that represent a meaningful data. *Sign Language* is a well-structured code gesture, every gesture having a unique meaning assigned to it. Many research works related to sign languages have been done as for example the American Sign Language, the British Sign Language, the Japanese Sign Language, and so on, which establish standards that are followed to consort a meaning to a particular gesture.

Interpreting is a complex process that requires a high degree of linguistic, cognitive and technical skills in English. Sign language interpreting, like spoken language interpreting, involves more than simply replacing a word of spoken English with a signed representation of that English word. Interpreters work in a variety of settings and situations. Interpreters provide services wherever a deaf or hard of hearing person needs to communicate with people who can hear but cannot sign fluently for themselves; such as in educational, medical field, theatre and legal settings; for conferences and conventions; or at corporations and institutions. Interpreters may also work as video relay interpreters, where deaf or hard of hearing individuals use an interpreter to communicate with anyone in the world over the telephone by the use of a webcam or video phone. Wherever communities of deaf people exist, sign languages develop. While they utilize space for grammar in a way that spoken languages do not, sign languages exhibit the same linguistic properties and use the same language faculty as do spoken languages.

Sign Language is the only means of communication for deaf people. Finding an experienced and qualified interpreter every time is a very difficult task and an unaffordable one too. Moreover, people who are not deaf, never try to learn the sign language for interacting with the deaf population. This becomes a cause of isolation of the deaf people.

But if a system can be programmed in such a way that it can translate sign language to text or audio format, the conversation gap between the normal people and the deaf community can be minimized. The project proposed is able to recognize the various alphabets and numbers of American Sign Language for Human Computer Interaction (HCI) giving more accurate results in least possible time.

Not only is gesture recognition limited to the use of deaf and handicapped population, but today's world also has many applications of such systems like robotics, gaming consoles, television control mechanisms and wheel chair automation. With such widespread applications, it is imperative for us to adapt to this changing technology and hence presenting a unique algorithm for gesture recognition from 3D images.

1.2. Gesture Recognition

A gesture can originate from any bodily motion or state but commonly originate from the face or hand. *Gesture recognition* is the concept with the goal of interpreting human gestures with the help of mathematical algorithms. Gesture recognition can be seen as a way for a computer to begin to understand human body language, thus building a richer bridge between machines and humans than primitive text interfaces or even GUIs (graphical user interfaces), which still limit the majority of input to keyboard and mouse.

Hence, gesture recognition is a process by which the system will know, what is going to be performed by the gesturer. Gesture recognition can be conducted with techniques from computer vision and image processing. Gesture recognition enables humans to interface with the machine and interact naturally without any mechanical devices. Using the concept of gesture recognition, it is possible to point a finger at the computer screen so that the cursor will move accordingly. This could potentially make conventional input devices such as mouse, keyboards and even touch-screens redundant. Hand gesture recognition system can be used for interfacing between computer and human using hand gesture.

Our project uses the concept of gesture recognition to apply it in the field of *image processing*. Various techniques and algorithms have been put forth and explained by different researchers all around the world, which will be studied in the literature survey. Gesture recognition using image processing involves the study of concepts like image segmentation, feature extraction and some mathematical calculations to produce the desired results.

1.3. Objectives

The *primary goal* of our project is to create a system which can identify specific human gestures from their 3D images.

Other major objectives of our proposed system are:

- *Image Segmentation* using *YCbCr* and *HSV* models
- *Feature extraction* using covariance matrix, Eigen values and Eigen vectors, covering all the features of the entire image
- Giving the respective text output, and, executing the application consorted to the identified gesture

CHAPTER 2

LITERATURE SURVEY

CHAPTER 2

LITERATURE SURVEY

2.1. Sign Languages: Types And Selection

There are different countries and their associations coming forward for the benefit of the deaf and dumb community who have designed and developed their respective sign languages. Today, hundreds of different sign languages exist for the vocally handicapped population. Some of the notable ones are as below:

- American Sign Language
- British Sign Language
- Chinese Sign Language
- Japanese Sign Language
- Indian Sign Language
- Sri Lankan Sign Language
- Argentin
- ian Sign Language
- Austrian Sign Language and so on….

American Sign Language was chosen amongst all of the above due to its following advantages:

- It is the most *predominant* sign language in the world
- All alphabets and numbers can be represented by using a *single hand*
- It is the *simplest* sign language to learn

- It is used all across the globe, from North America and Korea to the regions of South-East Asia

2.2. Necessity Of 3D Images

After examining the research work based on gesture recognition, it is noted that very less amount of work is done on 3D images. Two dimensional images are used to obtain the desired results in all the research works studied during the course of project. Hence, it is a good opportunity to explore a completely new dimension in the gesture images.

In 3D images, the third dimension can also be used, i.e., the depth of the image, where the inclination of hand or its tilting can be taken into account. So, with an added dimension, much more information can be stored within a single gesture. Many more gestures can be declared using the third dimension.

2.3. Work On 2D Gesture Recognition

Different approaches have been used by researchers for recognition of hand gestures which were implemented in varied fields.

- *Reza Hassanpour et al.* carried a detailed survey on the methods of analyzing, modeling and recognizing hand gestures in the context of the human computer interaction (HCI).[1]

- *Dharani Mazumdar et al.* used a new idea called "Finger-Pen", where a glove based system is developed by segmenting only one finger from the hand for proper tracking. Problems such as skin colour detection, complexity from large population in front of the camera, complex background removal and variable lighting condition are found to be efficiently handled by this system. [2]

- *Hazem Khaled et al.* used the background subtraction technique to extract the *Region Of Interest* (ROI) of the hand in which the fingertip is detected using logical heuristics equations that are applied on hand contour, convex hull and convexity defects points. This method improves the finger tip's detection by 52%.[3]

- *Joyeeta Singha et al.* proposed a system using Eigen value weighted Euclidean distance as a classification technique for recognition of various Sign Languages of India. The system comprises of four parts: Skin Filtering, Hand Cropping, Feature Extraction and Classification which gave a recognition rate of 97%.[4]

- *Cristina Manresa et al.* used a new algorithm to track and recognize hand gestures for interacting with a videogame. This algorithm consists of three main steps: hand segmentation, hand tracking and gesture recognition from hand features. The system's performance is evaluated by showing the usability of the algorithm in a videogame environment.[5]

- *Padmanabham Patki and Nagasrikanth Kallakuri* applied the approach of using various algorithms like Principal Component Analysis (PCA), Linear Discriminant Analysis (LDA), Support Vector Machines (SVM) considering linear, polynomial and sigmoid kernels and also Correlation Filter. An accuracy of up to 98.84% was attained using the above mentioned algorithms on a self-generated database. [6]

Many more papers were studied which also required external equipment like gloves or colored fingertip to identify the gesture. Thus, some of the above research work can be summarized and compared by using the following table:

AUTHORS	TITLE	RESEARCH WORK	DESIRED RESULTS
Padmanabham Patki & Nagasrikanth Kallakuri	British Sign Alphabet Recognition System	Using PCA, LDA, SVM And Correlation Filter For Gesture Recognition	Recognition Rate: 90%
Hazem Khaled, Sayed, El Sayed Mostafa & Hossam Ali	Hand Gesture Recognition Using Average Background And Logical Heuristic Equations	Background Subtraction Where The Fingertip Is Detected Using Logical Heuristics Equations	Recognition Rate: 84%
Dharani Mazumdar, Anjan Kumar Talukdar & Kandarpa Kumar Sarma	A Colour Fingertip-based Tracking Method For Continuous Hand Gesture Recognition	Use Of Finger pen and A Glove Based System For Segmentation	—
Joyeeta Singha & Karen Das	Indian Sign Language Recognition Using Eigen Value Weighted Euclidean Distance Based Classification Technique	Eigen Value Weighted Euclidean Distance As A Classification Technique For Recognition	Recognition Rate: 97%

CHAPTER 3

PRE-REQUISITES

CHAPTER 3

PRE-REQUISITES

For lucid understanding of the proposed system, knowledge about following subjects and concepts is required.

3.1. American Sign Language

Sign languages are an ancient form of languages that could be dated back to as early as the advent of the human civilization, when the first theories of sign languages appeared in history. Since then the sign language has evolved and been adopted as an integral part of our day to day communication process. In airports, a predefined set of gestures makes people on the ground able to communicate with the pilots and thereby give directions to the pilots of how to get off on the run-way. In the world of sports too, gestures are common. A single forefinger raised by the umpire in cricket signifies the batsman is out.

Furthermore, deaf people have over the years developed a sign language where all defined gestures have an assigned meaning. The language allows them to communicate with each other and the world they live in.

American Sign Language (ASL) is the predominant sign language of deaf communities in the United States and English-speaking parts of Canada. Besides North America, dialects of ASL and ASL-based creoles are used in many countries around the world, including much of West Africa and parts of Southeast Asia. Originated in the early 19th century, ASL use propagated widely via schools for the deaf community organizations.

Figure 1 depicts the gestures representing the numbers from 1 to 10 in American Sign Language, while figure 2 constitutes of all the alphabets in ASL.

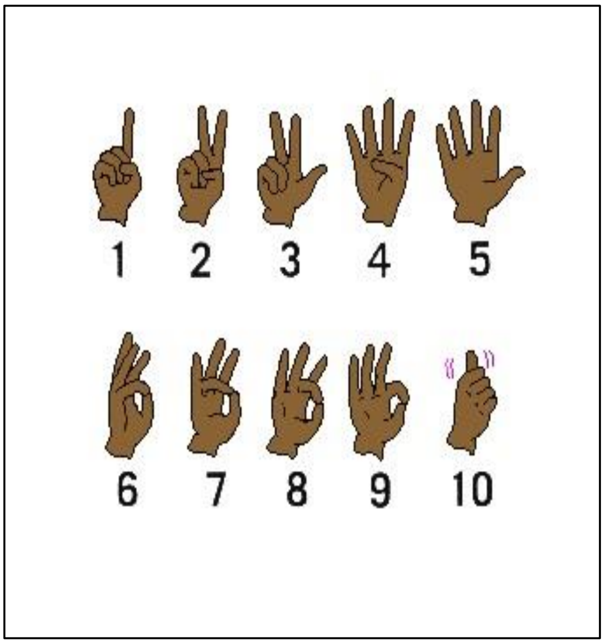


Figure 1: Numbers in American Sign Language

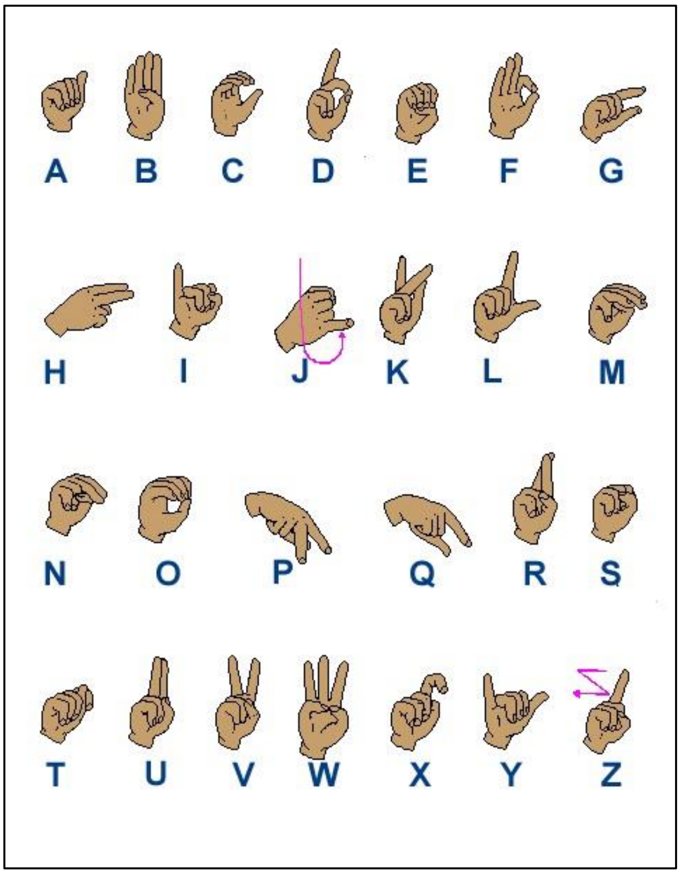


Figure 2: Alphabets in American Sign Language

3.2. Eigen Values And Eigen Vectors

- An *Eigen vector* of a square matrix A is a non-zero vector ν that, when the matrix is multiplied by ν, yields a constant multiple of ν, the multiplier being commonly denoted by λ. That is:

$$A\nu = \lambda\nu$$

  The number λ is called the *Eigen value* of A corresponding to *Eigen vector, ν.*

- Eigen vectors are set of basic functions which describes *variability of data*.
- Also, Eigen vectors are also a kind of coordinate system for which the covariance matrix becomes diagonal for which the new coordinate system is uncorrelated.
- Eigen values measures the *variance* of data of new coordinate system.
- Eigen values and Eigen vectors are a part of *linear transformations*.
- Eigen vectors are the directions along which the linear transformation acts by stretching, compressing or flipping and Eigen values gives the factor by which the compression or stretching occurs.
- The more the Eigen vectors the better the information obtained from the linear transformation.

CHAPTER 4

METHODOLOGY

CHAPTER 4

METHODOLOGY

The proposed system for "Hand Gesture Recognition From 3D Images" is explained along with the block diagram in this chapter.

4.1. Block Diagram

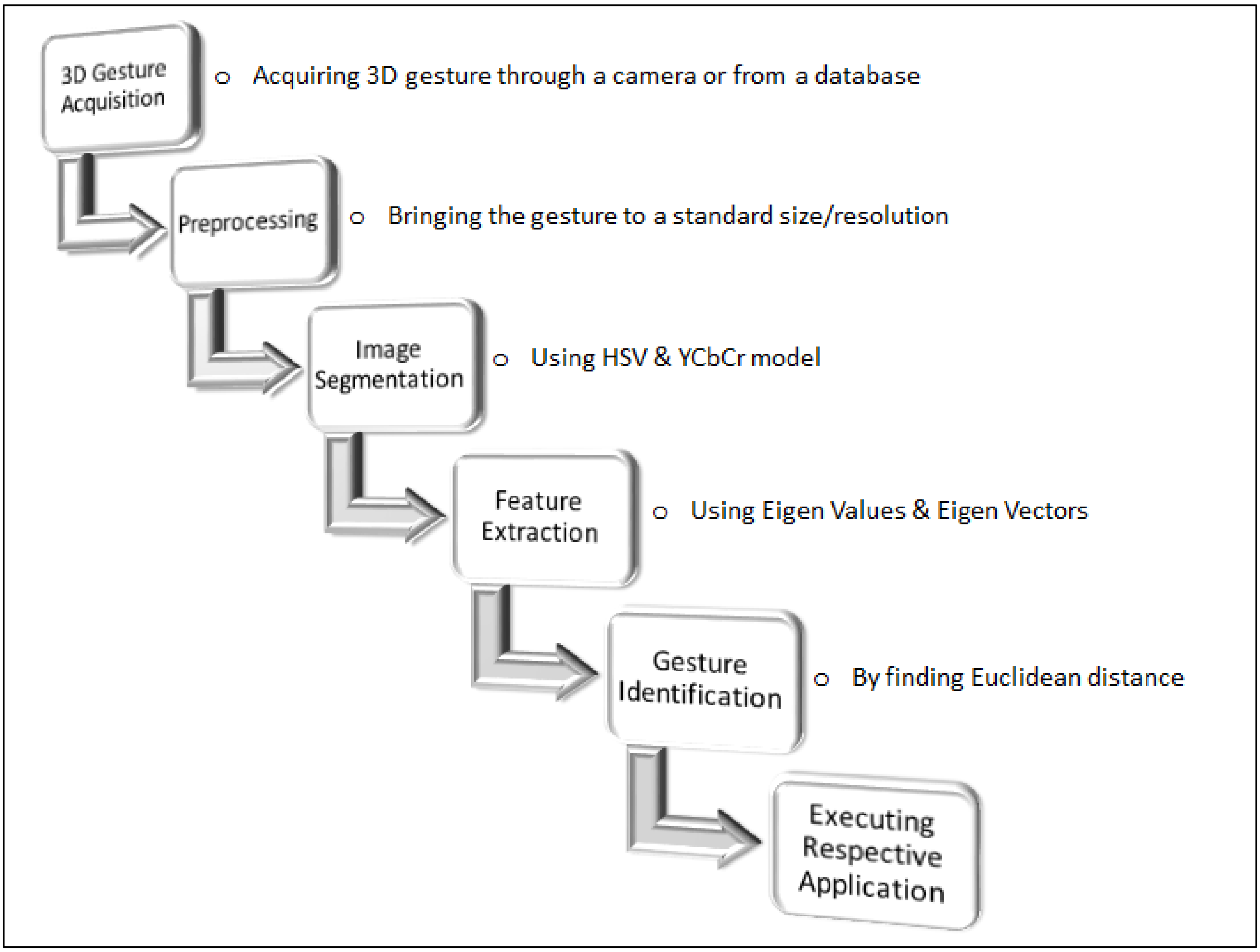


Figure 3: Block Diagram Of The Proposed System

4.2. Stages Of Operation

Our proposed system for "Hand Gesture Recognition From 3D Images" mainly consists of six stages. These six stages of operation are as below:

1] Acquiring 3D gesture image

2] Pre-processing

3] Image segmentation

4] Feature extraction

5] Projected test and train matrix comparison

6] Gesture identification and execution of respective application

4.2.1. Acquiring 3D gesture image

A three dimensional image of input gesture is at first taken as input through the digital camera. This image is stored in the test database. Also, prior to this, 3D train images of all the gestures are taken by the same camera and stored in the train database for further evaluation.

a. Camera specifications:

Name: Cybershot (DSC-HX-20V)

Type: Exmor R CMOS Sensor.

Size: 1 / 2.3 type (7.76mm).

Gross Pixels: Approx. 18.9 Mega Pixels.

Effective Pixels: Approx. 18.2 Mega Pixels.

Lens Type: Sony G Lens

Optical Zoom: 20x

Clear Image Zoom: 40x

b. *3D Still Image specifications:*

Extension: ‘.mpo’

18M (4,896 X 3,672) 4:3 mode

13M (4,896 X 2,752) 16:9 mode

4.2.2. Pre-processing

In pre-processing the resolution of the image is reduced i.e. changed as desired, because the resolution of any image for every camera differs. So to obtain this desired format the ‘imresize’ function of MATLAB was used. The image was scaled to 0.1, i.e., 10%.

MATLAB version specifications:

Name of the version: R2009B

Version number: 7.9.0.529

64-bit (win64)

4.2.3. Image Segmentation

This stage mainly does the work of hand detection and background removal from the image. The primary requirement before identifying and classifying the hand signs is to locate the hand in the frame, subtract the background and to be insensitive to lighting conditions. In order to obtain the location of the hand, skin colour recognition was used.

The algorithm used for the skin colour identification is to convert the image obtained, which is an image in RGB colour space, to YCbCr model and then set thresholds for the Cb and Cr colour dimensions. The thresholds used are:

$$77 < Cb < 131$$

$$121 < Cr < 160$$

Once the pixels that fall within the threshold limits are identified, all the other pixel values are zeroed. Thus a binary CbCr image is obtained from the thresholds.

Although the YCbCr model does detect skin colour effectively from the thresholds, there are certain limitations. The YCbCr model fails when the background comprises of flash/lighting effect, or shadow of the hand used, or when the colour of background is exactly similar to that of skin. Thus, another model called the HSV model was used along with the YCbCr model.

Figure 4: Test Image For Gesture ‘3’ Before YCbCr Segmentation

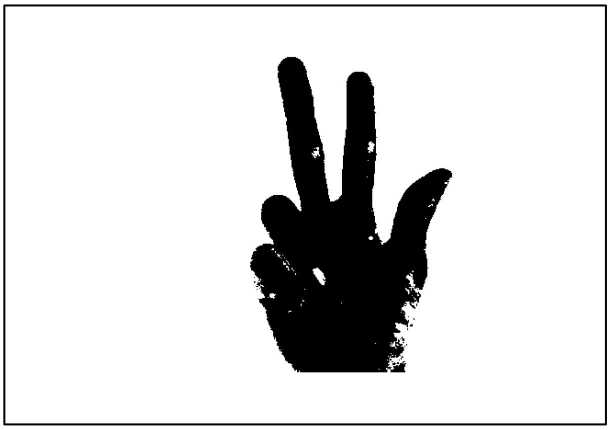


Figure 5: Test Image For Gesture '3'After YCbCr Segmentation

In HSV Image segmentation technique, the input RGB image is converted to the HSV image. The motive of performing this step is RGB image is very sensitive to change in illumination condition. The HSV colour space separates three components: Hue which means the set of pure colours within a colour space, Saturation describing the grade of purity of a colour image and Value giving relative lightness or darkness of a colour. Then, thresholds for hue and saturation are used. The brightness component is neglected since the effect of light in background like flash or shadow need not be included in the final binary image. The thresholds used for HSV binary image are:

$$0 < H < 0.25$$

$$0.05 < S < 0.9$$

After getting all the components that fall within the threshold limits, all the other pixel values are zeroed. Thus a binary HSV image is obtained from the thresholds.

Figure 6: Test Image For Gesture ‘7’ Before HSV Image Segmentation

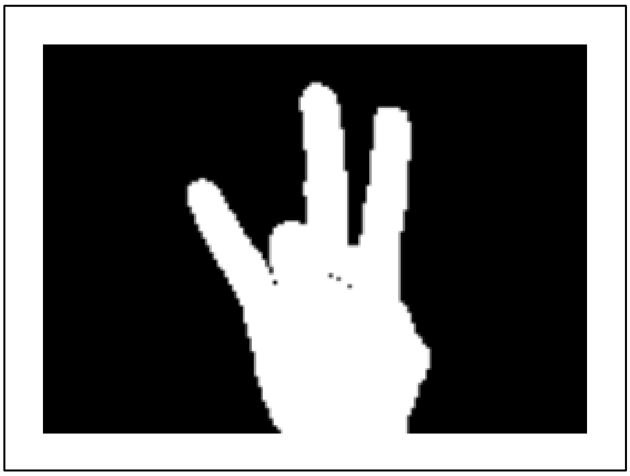


Figure 7: Test Image For Gesture ‘7’ After HSV Image Segmentation

A logical 'AND' operation is done between the binary HSV and the binary YCbCr image to get the most probable skin region (i.e. hand). Noise is minimized using morphological operations like erosion, dilation etc. Fig. 4 shows the processing from the test image to its equivalent binary image after segmentation.

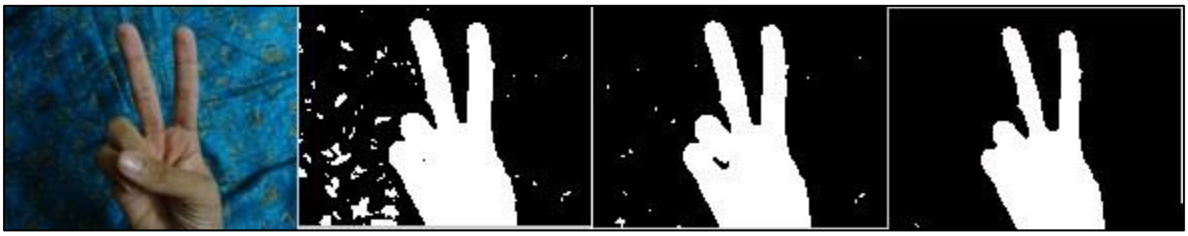


Figure 8: Image Segmentation
Test image followed by YCbCr binary image, HSV binary image and the final binary image (in order from left to right)

4.2.4. Feature Extraction

After successfully augmenting the hand portion from the image, all the features of the image must now be extracted, which are in the form of a matrix. This feature extraction is performed on both the train and test images in the following way:

- The train images are firstly converted to grayscale. Later reshape function is used to form the matrix of the grayscale images, in such a way that each column in the matrix represents the mean values of elements for each and every train image along each of its rows. The matrix will have exactly the same number of columns as the number of train images.

- A covariance matrix of the train images is formed by subtracting the mean value of each row in the above matrix from each element of every image.

$$CovMatTrainImage(i,j,k) = TrainImage(i,j,k) - Mean$$

- Using this covariance matrix, the Eigen values and Eigen vectors for all the train images are generated. This is done so as to have a detailed extraction of features from the images, without missing out on larger values within the image matrix. An Eigen matrix is then made out of these Eigen values and Eigen vectors. The formula used for this was,

  $$A\nu = \lambda\nu$$

  The number λ is called the Eigen value of A corresponding to Eigen vector, ν.

- Based on the Eigen matrix and the covariance matrix, a projected train image matrix is then generated which will have combined results of the covariance matrix and the Eigen matrix. This projected train image matrix will have values of all the train images in each of its columns respectively from which the image represented by that column can be displayed.

- Similar operations of gray scale conversion, reshape function, covariance and Eigen matrix generation are performed on the test image, so as to get the projected test image matrix representing the input gesture image.

4.2.5. Projected Test And Train Matrix Comparison

Based on the projected test and train image matrices, a Euclidean distance is found as the difference between the average values represented by the elements within them. The column in the projected train image matrix that deviates minimally from the projected test image matrix will have minimum Euclidean distance and vice versa.

Thus, the column with lowest Euclidean distance will represent the identified gesture image amongst all other train gesture images.

$$EucDist = ProjTestImg() - ProjTrainImg()$$

4.2.6. Gesture Identification And Execution Of Respective Application

Based on the identified gesture in the test image, the equivalent numerical and alphabetical output is given out on the screen, for example, an output of ‘1’ and ‘One’ for the test input as a gesture image representing 1 in ASL.

At the same time, a different Windows application is launched for every gesture identified. For example, if ‘1’ is the identified gesture, MS Word is opened in the system, or if ‘7’ is the identified gesture, MS PowerPoint is opened in the system, and so on.

In this way, using covariance matrix, Eigen values and Eigen vectors, and, Euclidean distance, any input gesture available in the train database can be identified by the above method.

CHAPTER 5

DESIGN

CHAPTER 5

DESIGN

For accurate and precise working of any system, it is necessary to follow the rules before beginning the coding for that system. Hence, it is necessary to first design the algorithm and flowchart for the system before starting its coding.

5.1. Algorithm

Our project presents an unparalleled algorithm as follows:

1. Start
2. *Acquiring 3D image* of the input gesture through an externally interfaced camera or a readymade database
3. *Pre-processing* the acquired image to bring it to a standard resolution or size
4. *Segmentation* of the image by *YCbCr* technique, negating the effects of background of the gesture
5. *Segmentation* of the image by *HSV* technique, negating the dynamic lighting conditions in the image
6. Combining the results of HSV and YCbCr image segmentation techniques and generating a single binary matrix from their results, by performing a logical AND operation between their results.
7. *Extracting features* to avoid an erroneous interpretation by using *covariance matrix, Eigen values and Eigen vectors*, covering all the features of the entire image
8. Generating a *projected test image matrix* for the input image
9. Generating a *projected train image matrix* for the train images from their difference with the test image

10. Calculating the *Euclidean distance* by finding the variation in projected test image matrix from the projected train images matrix.

11. Finding the image with the minimum Euclidean distance and declaring it as the identified gesture image

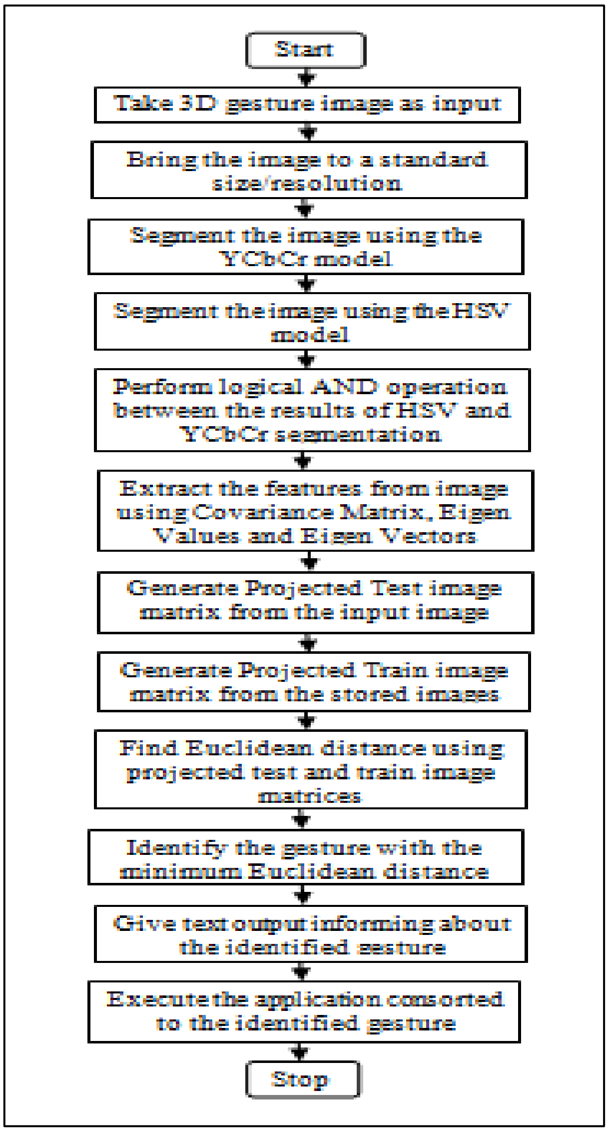
12. Giving the respective *text output* informing about the identified gesture

13. *Executing the application* consorted to the identified gesture

14. Stop

5.2. Flowchart

After understanding the algorithm, the program flow can be understood using the following flowchart.



CHAPTER 6

RESULTS AND ANALYSIS

CHAPTER 6

RESULTS AND ANALYSIS

With the help of adequate and apt image segmentation techniques of YCbCr and HSV models, the desired results can be obtained as gestures were correctly recognized by the system.

6.1. Results

Every input image was first segmented to remove the background noise and lighting effects, by using a combination of HSV and YCbCr segmentation models. After the segmentation and feature extraction, Euclidean distance was found with each train image with respect to the test image. This has been explained with an example below:

Example:

Consider that the input gesture is given as a 3D image of the number '*Nine'*, i.e. *'9'*. Now, after appropriate image segmentation and feature extraction, the calculation of Euclidean distance will be carried out. The results of Euclidean distance will be observed in the following manner:

euc_dist =

1.0e+006 *

6.1344 1.5051 3.0798 3.5246 4.9812 3.1846 3.2478 3.8272

0.7300 1.9502

Minimum Euclidean Distance is: 730042

The identified gesture enclosed in the image is: 9

As observed, the Euclidean distance with the ninth image of the train database is the least. Thus, the text output is given that the identified gesture is *'Nine'*.

In this way, the gestures can be identified based on the results of their Euclidean distance calculations.

6.2. Comparison Of Results With Various Backgrounds

Thus, by using covariance matrix, Eigen values and Eigen vectors, and, Euclidean distance, any input gesture available in the train database can be identified by the proposed method. Some of these results are shown in the Table 1.

Test Image	Euclidean Distance with Train image 1	Euclidean Distance with Train image 2	Euclidean Distance with Train image 3	Euclidean Distance with Train image 4	Euclidean Distance with Train image 5	Gesture Recognized	Application opened as per gesture
	2.3576 X e^{11}	**2.3571 X e^{11}**	2.3698 X e^{11}	2.3668 X e^{11}	2.3654 X e^{11}		MS Excel
	4.0995 X e^{11}	6.3094 X e^{11}	2.7352 X e^{11}	3.5371 X e^{11}	**1.0453 X e^{11}**		Mozilla Firefox
	**0**	5.8744 X e^{11}	4.7386 X e^{11}	5.7671 X e^{11}	2.3354 X e^{11}		MS Word

Table 1: Comparison of results with various backgrounds

As seen in the above table, the gesture is identified based upon its Euclidean distance with the train images. The image with minimum Euclidean distance (values shown in **bold**) is declared as the identified image.

To get a detailed idea about the performance and accuracy of the system, a comparison of results for each gesture was done. This comparison was done by taking the 3D image of the gesture with different backgrounds. The backgrounds not only differed in color, but also in designs and patterns.

It was observed that the *average recognition rate* for gestures was *92%*. This means that, if 100 3D images (with different backgrounds in colors, designs and patterns) of a single gesture were processed through our system, on an average 92 of these 100 gestures will be recognized successfully.

CHAPTER 7

CONCLUSIONS AND FUTURE SCOPE

CHAPTER 7

CONCLUSIONS AND FUTURE SCOPE

Thus, the entire project on “Hand Gesture Recognition From 3D Images” can be summarized as below.

7.1. Conclusions

A gesture can be identified after the removal of background by finding its covariance matrix. This covariance matrix can then be used to find the Eigen values and Eigen vectors from the matrix for each test and train image. After calculating the Eigen matrices of both test and train image, one can give the desired output declaring the identified gesture, by finding the minimum Euclidean distance between the test and train image, as shown in Table 1. The *average recognition rate* for the system was *92%*.

Thus, doing non-trivial tasks can become much easier using this technique of hand gesture recognition from 3D images.

7.2. Future Scope

Although our gesture recognition algorithm is advantageous, it has some limitations. These limitations, listed below, can be worked upon for future work.

- One of the limitations of our algorithm is, if the background happens to be of skin colour, then the YCbCr and HSV results fail to show complete background suppression. Some minor part of 'hand' is lost during segmentation. Techniques like Gaussian filtering can be used to do so [7].

- Furthermore, the execution time for our algorithm is around half a minute per gesture recognized since they are 3D gestures with large file size. Work can also be carried around this to make the algorithm much faster.

- The project can be implemented on 3D images with the help of Sharp C on microcontrollers, DSP processors and FPGA kits.

- Applications in the modern world:
  The project can be implemented in some real time applications like:-
  - Elevators
  - Wheel chair automation.
  - Automated vehicles.
  - Human computer interface.

CHAPTER 8

BIBLIOGRAPHY

CHAPTER 8

BIBLIOGRAPHY

A list of different books, research papers and websites referred during the course of the project is mentioned below.

8.1. Books Referred

A couple of *books* were referred for information about the different segmentation techniques and their comparison, as mentioned below:

1. A. K. Jain, "Fundamentals of Digital Image Processing", Pearson Education
2. Gonzalez and Woods, "Digital Image Processing", Pearson Education

8.2. Papers Referred

The different *research papers* referred for the project are as below:

1. Reza Hassanpour, Asadollah Shahbahrami, "*Human Computer Interaction Using Vision-Based Hand Gesture Recognition*", Journal of Advances in Computer Research, 2010.
2. Dharani Mazumdar, Anjan Kumar Talukdar & Kandarpa Kumar Sarma , "*A coloured fingertip-based tracking method for continuous hand gesture recognition*", International Journal of Electronics Signals and Systems (IJESS), ISSN: 2231- 5969, Vol-3, Iss-1, 2013.
3. Hazem Khaled, S. Sayed, El Sayed Mostafa ,Hossam Ali , "*Hand Gesture Recognition Using Average Background and Logical Heuristic*

Equations", International Journal Of Computers & Technology, ISSN 2277-3061, Vol. 11 No. 5, 2013

4. Joyeeta Singha, Karen Das, "*Indian Sign Language Recognition Using Eigen Value Weighted Euclidean Distance Based Classification Technique*", International Journal of Advanced Computer Science and Applications, Vol. 4, No. 2, 2013.
5. Cristina Manresa, Javier Varona, Ramon Mas and Francisco J. Perales, "*Hand Tracking and Gesture Recognition for Human-Computer Interaction*", Electronic Letters on Computer Vision and Image Analysis 5(3):96-104, 2005.
6. Padmanabham Patki, Nagasrikanth Kallakuri, "*British Sign Alphabet Recognition System*".
7. Aisha Meethian, B.M.Imran, "*Real time gesture recognition using hand tracking system based on GMM*".
8. Diedrick Marius, Sumita Pennathur, and Klint Rose, "*Face Detection Using Color Thresholding, and Eigenimage Template Matching*".
9. S. Chitra and G. Balakrishnan, "*Comparative Study for Two Color Spaces HSCbCr and YCbCr in Skin Color Detection*". Applied Mathematical Sciences, Vol. 6, 2012, no. 85, 4229 - 4238
10. Sofia Tsekeridou and Ioannis Pita, " *Facial Feature Extraction in Frontal Views Using Biometric Analogies"*
11. Amanpreet Kaur and B.V Kranthi, "*Comparison between YCbCr Color Space and CIELab Color Space for Skin Color Segmentation",* International Journal of Applied Information Systems (IJAIS) – ISSN : 2249-0868 Foundation of Computer Science FCS, New York, USA Volume 3– No.4, July 2012.

12. Jorge Alberto Marcial Basilio, Gualberto Aguilar Torres, Gabriel Sánchez Pérez, L. Karina Toscano Medina, Héctor M. Pérez Meana, *“Explicit Image Detection using YCbCr Space Color Model as Skin Detection”*, Applications of Mathematics and Computer Engineering, ISBN: 978-960-474-270-7
13. Yanjiang Wang, Baozong Yuan, *“A novel approach for human face detection from color images under complex background”*, Pattern Recognition 34 (2001) 1983}1992.
14. Douglas Chai and King N. Ngan, *“Face Segmentation Using Skin-Color Map in Videophone Applications”*, IEEE Transactions On Circuits And Systems For Video Technology, Vol. 9, No. 4, June 1999.
15. Rick Kjeldsen and John Kender, *“Finding Skin in Color Images”*, 1996 IEEE.
16. G. Kukharev and A. Novosielski, *“Visitor Identification Elaborating Real Time Face Recognition System”*, Proceedings Of 12th Winter School Of Computer Graphics, Plzen, February 2004

ACHIEVEMENTS

Following are the achievements of our project:

1. Paper on the project titled *"Application Control Using 3D Hand Gesture Recognition"* published at the *IEEE's International Conference On Convergence Of Technology (I2CT) 2014*. ISBN: 978-1-4799-3759-2

2. Paper on the project titled *"Hand Gesture Recognition From 3D Images"* published at the *IETE's National Level Undergraduate Conference On Emerging Trends In Electronics And Telecommunication Engineering (UGCON) 2014*.

3. Participated in *IETE's Project Competition 2014* held at *D. Y. Patil College Of Engineering, Pimpri, Pune*.