

“A Vision Based Static Hand Gesture Alphabet Recognition”

Ankita Chavda¹, Sejal Thakkar²

GCET, VallabhVidyanagar, Anand, Gujarat, India¹

GCET, VallabhVidyanagar, Anand, Gujarat, India²

Abstract- The work presented in this paper has goal to develop a system for automatic translation of static gestures of alphabets in American Sign Language. The required images for the selected alphabet are obtained using a digital camera. Color images will be segmented and converted into binary images. Morphological filtering is applied to binary image. After that the feature extraction is applied by Modified Moore Neighbor Contour Tracing Algorithm. Then area, centroid, and shape of the object are used as features. Finally based on these features and scan line feature, rule based approach is used to recognize the alphabet. This system only requires the live image of hand for the recognition. This system is able to recognize 26 selected ASL alphabets.

Keywords: Hand Gesture Recognition, Human Computer Interaction, American Sign Language

I. INTRODUCTION

Gesture recognition is an area of active current research in computer vision. Body language is an important way of communication among humans. People perform various gestures in their daily lives. It is in our nature to use gestures in order to improve the communication between us. The User Interface of personal computer has evolved from a text based command line to graphical interface with keyboard and mouse input. However, they are inconvenient and unnatural [1]. The use of hand gestures provides an attractive alternative to these cumbersome interface devices for Human Computer Interaction (HCI) [3].

The sign language is the fundamental communication method between the people who suffer from hearing defects [4]. For an ordinary person to communicate with deaf people, a translator is usually needed for the sign language into natural language or vice versa. Sign language can be considered as a collection of gestures, movement, posters and facial expressions corresponding to letters and words in natural language [4]. A large variety of techniques have been used for modeling the hand.

A hand gesture is an expressive movement of body part which has a particular message to be communicated between a sender and receiver [8]. A gesture is scientifically categorized into 2 parts: 1) Static gesture and 2) Dynamic gesture. For example if umpire wants to declare six then he will give the static sign with one finger

open and if it is four then he will give the continuous hand movement that is example of dynamic gesture.

The direct use of the hand as an input device is an attractive method for providing natural Human-Computer Interaction. Three approaches are commonly used to interpret the gestures for Human-Computer Interaction.

1) *Vision based Approach:* In Vision based method, the system requires only camera to capture the image required for the interaction between human and computer. This approach is very simple but so many challenges will be raised because of complex background, lighting variations, and other skin color object with the hand object [3].

2) *Data Glove based Approach:* Data glove approach uses sensor device for capturing hand position and motion. These approaches can easily provide exact co-ordinate of palm and finger. These devices are quite expensive and inefficient in virtual reality [3].



Fig. 1 (a) Data Glove based, (b) Vision based, (c) Colored marker based
(From web gallery)

3) *Colored marker Approach:* Colored markers are gloves that wear by human hand with some colors to direct the tracking process of hand to locate the palm and finger, which extract the geometric features to form the hand shape [3]. This technology is simple to use and low cost as compare to data glove but limits the naturalness level for interaction between human and computer.

A. Application Domains

There is a large variety of applications which involves the hand gesture which can be used to achieve natural human computer interaction for virtual environments. This technology can be used for deaf people. We present an overview of some of the application domains for gesture interaction.

Virtual Reality: Virtual reality applications use gestures to enable realistic manipulations of virtual objects using

3D display or 2D display simulation with 3D interaction [1], [2], [9].

*Robotics & Tele-presence:*Tele presence and robotics applications are typically situated within space exploration and military based research projects [1]. The robots have to co-operate with humans in uncertain and complex environment. The gestures used to interact and control the robots by cameras located on the robot [2]. So, these gestures can control the hand and arm movement of robots.

*Desktop & Tablet PC Application:*In desktop computing applications, gesture can provide an alternative interaction to the mouse and keyboard [1]. The task of manipulating graphics and editing documents can be done by many gestures using pen based gestures. Recently Android Tablets and windows based portable computers are using eyesight technology [2].

Vehicle Interfaces: These secondary control in vehicle is broadly based on the premise that taking the eye off the road to operate conventional secondary controls, but it can be reduced by using hand gestures [1].

Health care: Instead of touch screen or computer keyboard, hand gesture recognition system enables doctors to manipulate images during medical procedures [1]. A Human-Computer Interface is very important because by this surgeon controls medical information, avoiding patient contamination, operating room and other surgeons. Hand gesture recognition system offers a possible alternative.

*Games:*Hand gesture recognition is used to track player's hand and body position to control movement and orientation of interactive games objects such as car [1]. Play station 2 has introduced EyeToy, a camera that tracks hand movements for interactive games [2].

*Sign Language:*Sign language is an important case of communicative gestures [1], [2]. ASL is the fourth most used language in the United States only behind English, Spanish and Italian [4]. Sign language for the deaf is an example that has received significant attention in gesture literature [9].

II. VISION BASED GESTURE RECOGNITION

Computer vision based techniques have potential to provide more natural and non-contact solution, there are non-invasive and are based on the way human beings perceive information about their surroundings. It is difficult to design a vision based interface for generic usage, but it is feasible to design such an interface for a controlled environment [1]. For static hand gesture, it is possible to select some geometric or non-geometric features. Since it is not easy to specify features explicitly, image is taken as input and features are selected implicitly and automatically by recognizer. The approaches to vision

based hand gesture recognition can be divided into 3 categories:

1) *Model based Approaches:*Model based approach generates model speculation and evaluate them based on the available visual observations. This can be performed by formulating an optimization problem whose objective function is to measure divergence between the expected visual cues and actual ones [1]. This approach attempts to infer the pose of palm and the joint angles. Such an approach would be idle for realistic interactions in virtual environments. Generally, the approach is to search kinematic parameters that can bring the 2D projection of a 3D model of the hand into correspondence with an edge based image of the hand [2]. A common problem with this model is feature extract. The edges that can be extracted are usually from the exterior boundaries. This approach requires homogeneous and high contrast background relative to hand.

2) *Appearance based Approaches:*Appearance based model are derived directly from the information contained in the images and have traditionally been used for gesture recognition [1]. No explicit model is needed for hand, which means that no internal degrees to be specifically modeled. Unlike model based approach, differentiating between gestures is not straight forward. Therefore, the gesture recognition involves some sort of statistical classifier based on features that represent the hand [2]. Some of application use multi scale features like color detection. Appearance based approach is also known as View based approach as the gestures are modeled as a sequence of views [4].

3) *Low Level Features based Approaches:*In many gesture applications, mapping is required between input video and gesture [2]. Therefore the full reconstruction of the hand is not essential for gesture recognition. So, many approaches have utilized the extraction of low level image measurement that are fairly robust to noise and can be extracted quickly [2]. The low level features can be Centroid, Shape of object, height and boundaries.

A. Classification of Hand gestures

Hand gesture can be classified using the following two approaches:

1) *Rule based Approaches:*Rule based approaches consist of a set of manually encoded rules between feature inputs. A set of features are extracted from given input gestures and compared to encoded rules and the rule that matches the input is given as an output gesture [2]. The low level features predicates of the hand are defined for each of the actions under consideration. When the predicate of gesture is satisfied over a fix number of consecutive frames the gesture is returned [2]. A problem of rule based approach is that they rely on the ability of

human to encode the rules. But, if the rules are encoded well then this algorithm is fast as compare to any other and accurate also.

2) *Machine Learning based Approaches*: A machine learning approach is popular to treat a gesture as the output of stochastic process. The Hidden Markov Model and Artificial Neural Network have received the most attention in literature for classifying gesture [2].

B. American Sign Language

American Sign Language (ASL) is a complete language that employs signs made with the hands, facial expression and postures of the body. ASL is the visual language, means it is not expressed through sound but rather through combining hand shape with movement of hands and arms. ASL has its own grammar that is different from other sign languages like English [4]. ASL consists of approximately 6000 gestures of common words or proper nouns. Finger spelling uses one hand and 26 gestures to communicate the 26 alphabets [4]. The 26 alphabets of ASL are shown in Fig.2.

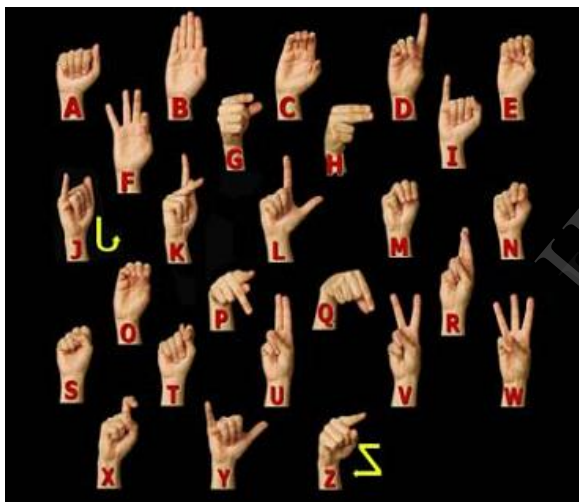


Fig. 2 American Sign Language Finger Spelling Alphabet [4]

A number of recognition techniques are available for vision based hand gesture recognition. Communicative gestures are intended to express an idea or a concept [7].

III. SYSTEM DESIGN AND IMPLEMENTATION

A system is designed to virtually recognize all static sign of American Sign Language (ASL). The users are not required to wear any type of sensor glove or color marker glove. But, since the different signers vary their hand shape size, operation habit and so on, which brings more difficulties in recognition [4]. That is why we require independent sign language recognizer to improve the system robustness. The combination of feature extraction method with image processing and rule based approach have led to successful implementation of ASL recognition

system with MATLAB [4]. The system is having 2 phases: Feature extraction phase and classification phase as explained in Figure.3. The feature extraction phase includes various image processing techniques which involves algorithm to detect and isolate desired portion of an image. The classification phase includes that based on what; we are classifying the signs into alphabet.

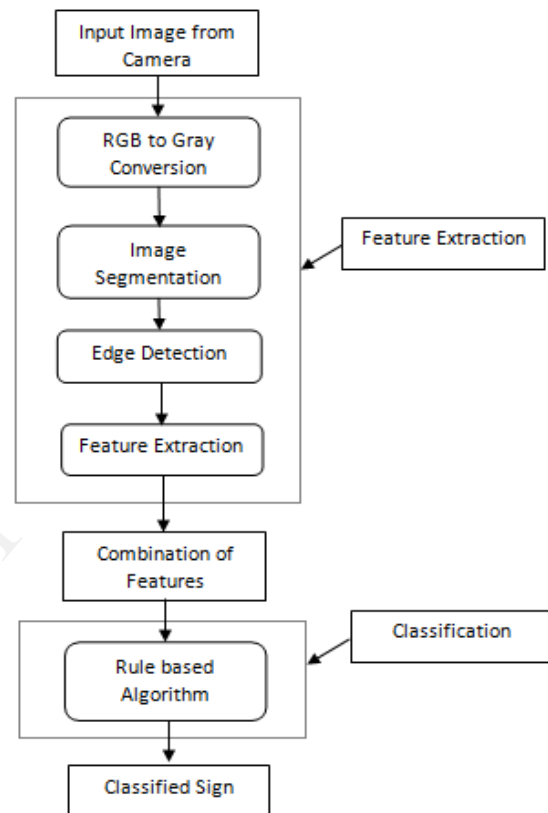


Fig. 3 System Overview

A. Feature Extraction Phase

In feature extraction phase, the first step is to take an image from digital camera or web cam. Once the frame picture is captured, it requires RGB to Gray conversion as our taken image is in RGB form. After the Conversion, we are having Gray scale image and now it is converted into binary by using OTSU's segmentation algorithm. These binary image pixels are having only 2 numeric values that are 0 as black and 1 as white. The binary images are often produced by thresholding a gray scale image from background. The result of this is sharp and clear details of an image. The goal of edge detection is to extract the boundary of desired object for shape details. Edge detector defines the location of our desired features in image. Canny edge detector is very effective algorithm but it is providing extra details more than needed [4]. For this reason we have selected Modified Moore Neighbor

Contour Tracing Algorithm for getting required detail only from the image as feature extraction method [10].

B. Classification Phase

The classification phase includes different features of feature extraction method and here we will apply rule based approach for recognition. In our application we have consider 3 features; Centroid, Roundness and Scan line features for recognition. We have collected all the values of different features and then based on that we will make the rules for recognition. This classification method is totally based on rules. Learning based approach is much time consuming method, so we are using rule based approach for fast recognition.

C. Implementation Procedure

The Proposed algorithm of static hand gesture recognition for our system which is implemented using MATLAB is explained below:

*Step-1:-*Capture the image from the camera using `videoinput()` function.

*Step-2:-*Convert into RGB to Gray image using `rgb2gray()` function.

*Step-3:-*Apply OTSU Segmentation algorithm using `graythresh()` function.

*Step-4:-*Convert into the Binary image using `im2bw()` function.

*Step-5:-*Apply the Morphological filtering using `strel()`, `imclose()`, `imfill()` functions and many more.

*Step-6:-*Apply the feature extraction by Modified Moore Neighbor Contour Tracing algorithm with `bwboundaries()` function.

*Step-7:-*Finding the Centroid using `regionprops()` function as per the equation

$$X_c = \frac{\sum_{i=1}^N X_i}{\text{area}}, \quad Y_c = \frac{\sum_{i=1}^N Y_i}{\text{area}}$$

Where, X_i and Y_i represents the X-coordinate and Y-coordinate of each boundary pixel of an image respectively [5].

*Step-8:-*Finding the roundness of desired object by formula

$$\text{Metric} = 4 * \pi * \text{area} / \text{perimeter}^2;$$

If the metric value is near to 1 then shape is round and if not then object is not round.

*Step-9:-*Using the scan line feature, we are dividing the frame into 4 regions and based on the appearance of these signs in that particular region, the alphabet will be recognized.

IV. EXPERIMENTAL RESULTS AND ANALYSIS

The performance of the recognition system is evaluated by testing its ability to classify the signs based on rule based recognition.

A. Experimental Results

We have tested our hand gesture recognition system for different signs. Here, i have wear white gloves for reducing skin color detection algorithm process, which is time consuming. The result is as shown in figure.

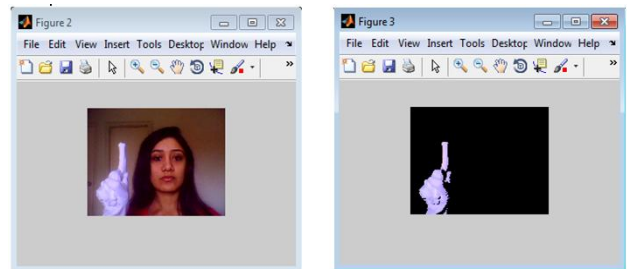


Fig. 4 Segmentation of gray scale image of gesture "D"

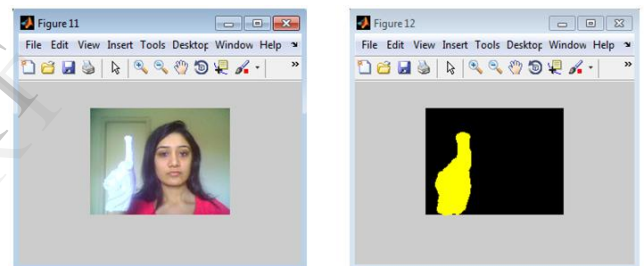


Fig. 5 Morphological filtered gesture "D"

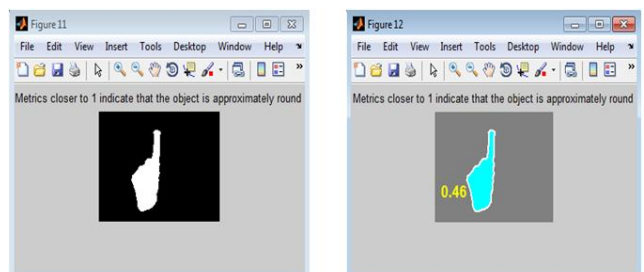


Fig. 6 Roundness of gesture "D"

Here, Fig 4 is showing the Segmentation of gray scale image but we have not applied any filter yet. So, image is somehow blurring. In Fig 5, we have applied some morphological filtering. After that we have found Centroid feature and based on that centroid we have found roundness of an object that is shown in Fig 6. But by only centroid and roundness we are not getting exact accuracy in rule based approach. So, we have added one more feature that is Scan line feature.

In Scan line feature, we have divided our GUI frame into four regions. Lower two regions containing palm and

upper two regions containing fingers. So, we can apply the rules; based on appearance of fingers and palm in particular region and based on that we can recognize alphabet that is shown in Fig 7.

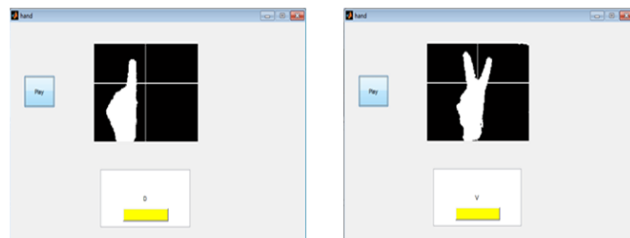


Fig. 7(a) Recognized gesture “D”, (b) Recognized gesture “V”

So, this hand gesture recognition system is accurate as well as fast algorithm for recognition of alphabet.

V. RESULT EVALUATION AND DISCUSSION

The performance of the recognition system is evaluated by testing with different signs. We have combined all the features and made rules by which we can recognize the gesture. As we are taking live image from camera, we have to take care of light and complex background.

TABLE I

THE PERFORMANCE OF RECOGNITION OF ASL

Letters	Recognized Letters	Unrecognized Letters	Recognition Rate (%)
A	7	1	87.5
B	7	1	87.5
C	6	2	75
D	8	0	100
E	7	1	87.5
F	6	2	75
G	5	3	62.5
H	7	1	87.5
I	8	0	100
J	5	3	62.5
K	6	2	75
L	7	1	87.5
M	7	1	87.5
N	5	3	62.5
O	6	2	75
P	6	2	75
Q	4	4	50
R	7	1	87.5
S	5	3	62.5
T	7	1	87.5
U	8	0	100
V	8	0	100
W	8	0	100
X	7	1	87.5
Y	6	2	75
Z	6	2	75

In the present work, the signs of ASL are captured by a digital camera. We have taken 8 samples for each gesture. We have calculated recognized and unrecognized letters for each alphabet and will calculate the rate of efficiency based on it. The recognition rate (%) = (No of Recognized Letters/ Total No of Samples) * 100 [4], [6]. This system is giving the gesture recognition very fast and accurate.

VI. CONCLUSION

In the present paper, a simple model of static hand gesture of ASL recognition system using rule based approach with MATLAB tool is discussed. The evaluation results show that the proposed method allows fast and reliable recognition with accuracy. Our proposed algorithm is Scale invariant means the actual size of hand and its distance from camera do not affect interpretation [5]. The overall performance of this system is near about 81% as describe in Table I. Using more features for feature extraction, we can improve the efficiency of our system. In learning based approach, the learning time is too high for just 26 alphabets. So, as compared to learning base algorithm, our proposed algorithm is fast and also accurate. Future work will include more features for feature extraction will be added to improve the system efficiency.

REFERENCES

- [1] G. Simion, V. Gui, and M. Oteşteanu, “Vision Based Hand Gesture Recognition: A Review”, INTERNATIONAL JOURNAL OF CIRCUITS, SYSTEMS AND SIGNAL PROCESSING, 2009
- [2] G. R. S. Murthy & R. S. Jadon, “A REVIEW OF VISION BASED HAND GESTURES RECOGNITION”, International Journal of Information Technology and Knowledge Management July-December 2009, Volume 2, No. 2, pp. 405-410
- [3] Noor Adnan Ibrahim & Rafiqul Zaman Khan, “Survey on Various Gesture Recognition Technologies and Techniques”, *International Journal of Computer Applications* Volume 50, July 2012
- [4] Vaishali S. Kulkarni, Dr. S.D.Lokhande, “Appearance Based Recognition of American Sign language Using Gesture Segmentation” (IJCSSE) International Journal on Computer Science and Engineering Vol. 02, No. 03, 2010, 560-565
- [5] Asanterabi Malima, Erol Özgür, and Mujdat Cetin, “A FAST ALGORITHM FOR VISION-BASED HAND GESTURE RECOGNITION FOR ROBOT CONTROL”, Faculty of Engineering and Natural Sciences, Sabanci University, Tuzla, İstanbul, Turkey.
- [6] Md. Atiqur Rahman, Ahsan-Ul-Ambia and Md. Aktaruzzaman, “Recognition Static Hand Gestures of Alphabet in ASL”, COPYRIGHT © 2011 IJCIT, ISSN 2078-5828 (PRINT), ISSN 2218-5224 (ONLINE), VOLUME 02, ISSUE 01, MANUSCRIPT CODE: 110749
- [7] Sushmita Mitra, Senior Member, IEEE, and Tinku Acharya, Senior Member, IEEE “Gesture Recognition: A Survey” IEEE TRANSACTIONS ON SYSTEMS, MAN, AND CYBERNETICS—PART C: APPLICATIONS AND REVIEWS, VOL. 37, NO. 3, MAY 2007

- [8] Noor Adnan Ibrahim & Rafiqul Zaman Khan, "COMPARATIVE STUDY OF HAND GESTURE RECOGNITION SYSTEM", International Journal of Computer Applications Volume 50, July 2012
- [9] Noor Adnan Ibrahim & Rafiqul Zaman Khan, "HAND GESTURE RECOGNITION: A LITERATURE REVIEW", International Journal of Artificial Intelligence & Applications (IJAA), Vol.3, No.4, July 2012
- [10] Ratika Pradhan, Shikhar Kumar, Ruchika Agarwal, Mohan P. Pradhan & M. K. Ghose, "Contour Line Tracing Algorithm for Digital Topographic Maps", Department of CSE, SMIT, Rangpo, Sikkim, INDIA.

First Author: Ms. Ankita Chavda is pursuing her M.E. from G. H. Patel College of Engineering and Technology.

Second Author: Ms. Sejal Thakkar is working as Assistant professor in department of information technology at G. H. Patel College of Engineering and Technology.

