Hand Gesture Recognition

Hand Gesture Recognition To Automated Vocal Speech

Tamzid Ahmed[18−37503−1[C.S.E]]
email: tamjidahmed958@gmail.com

Bazlul Bari Mozharul Haq[18−37266−1[C.S.E]]
email: m220197@gmail.com

Joydeep Karmoker[16−33028−3[C.S.E]]
email: jkarmoker96@gmail.com

A thesis proposal presented for the degree of
Bachelor of Science

Computer Science and Engineering
American International University-Bangladesh
Bangladesh
20 Nov 2020

1 Abstract

This framework can be successful and can be utilized soon for correspondence between the individuals who do not comprehend the Sign language since it deals with the guideline of Gesture Recognition. The framework uses a characteristic portable camera for motion acknowledgment and securing; signal procured is prepared with the assistance of Algorithms like HSV model-(Skin Colour Detection), LargeBlob Detection, Flood Fill, and Contour Extraction The framework can remember one sign of the standard letters in order (A-Z) and numeric qualities (0-9). This framework's yield is exceptionally proficient, steady, and of high guess of signal handling and discourse examination.

2 Introduction

Human-Computer Interaction (HCI) is a discipline associated with designing, evaluating, and implementing computing systems that interactive for human use. A gesture is a non-verbal communication in which only visible body language actions communicate particular messages instead of speaking in spoken words. Gestures include body movements such as the hand, face, or other parts of the body. The primary purpose of developing such a system is to understand a person's hand or body movement and implement them to an intelligible voice speech. It also serves many applications, from virtual reality to sign language recognition.
Both non-vision and vision-based approaches are used to acquire hand gesture recognition. An example of a non-vision-based approach is the detection of finger bending with a pair of wired gloves. In general vision-based approaches are more natural as they require no hand fitting devices to determine it. Theoretically, the research classifies hand gestures into two types: static and dynamic gestures. Static hand gestures are defined as the hand's orientation and position in the space during an amount of time without any movement. If there is a movement during the time above duration, then it is called a dynamic gesture. Dynamic hand gestures are those gestures that involve the movement of body parts like waving of a hand, while static hand gestures include single formation without movement. The static expression includes the "Ok" signal. The process is done by jamming the thumb and index finger. Writing survey After exploration, the different advances engaged with signal acknowledgment include picture obtaining, pre-preparing, division, highlight focuses extraction, and order, trailed by the outcome. The paper conjures the utilization of different PC vision strategies and calculations, which are included in the assurance of hand motions.

3 Research Motivation

In the previous research, many techniques have been used to convert the hand gesture to speech. Nonetheless, they were restricted regarding their functionalities. Numerous methods required gloves with sensors, which not just made the application more complicated but also expensive. The goal of this project is to design a hand gesture recognition system. This project follows a vision-based approach for gesture recognition. So that we can use two different algorithms. One is the HSV model, which can detect the skin color of a person. Another one is LargeBlob detection that can also determine the body or hand movement of a person.
We are motivated for this research by searching many research papers and results that anyone can not do it so understandable. We want to do something new that can easily detect the hand gesture to an understandable speech. Some projects were heavily dependent on massive GPUs, making it difficult for a man to use the system. Additionally, there were some systems for discovery which required the object to be of particular skin color. Although there have been different procedures for changing over the hand signal to message yet not many spotlights on changing the motion over to both content and discourse with that too with restricted properties.

4 Objective

This research's primary goal is to present a framework of a tool that will recognize hand gestures and then automate the gestures meaning into words. Hand gestures are taken as input from the user and then converting the meaning of the hand gesture to spoken words as output. The overall objective is divided into the following sub-objectives:

- To provide an idea of how the selected algorithm processes the input images.
- How to utilize the processed image data to have a more accurate match with the database's data.
- How to generate more sensible sentences from hand gestures.

5 Research Methodology

5.1 Case Study Research Method

The diagram of hand signal acknowledgment is depicted in Figure 1. First, the hand is identified utilizing the foundation deduction strategy, and the consequence of hand recognition is changed into a paired picture. At that point, the fingers and palm are portioned to encourage finger acknowledgment. Besides, the fingers are distinguished and perceived. Last, hand signals are perceived utilizing a primary principle classifier.

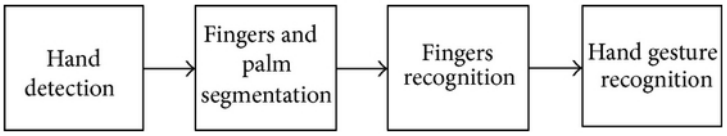

Figure 1: The overview of the proposed method for hand gesture recognition.

Hand Detection by HSV:
The first pictures utilized for hand motion acknowledgment in work are shown in Figure 2. The foundation of these pictures is indistinguishable. It is simple and powerful to identify the hand locale from the first picture utilizing the foundation deduction strategy. Nonetheless, now and again, there are other moving articles remembered for the aftereffect of foundation deduction. The skin tone can be utilized to separate the hand area from the other moving articles. The shade of the skin is estimated with the HSV model. The HSV (tint, immersion, and worth) estimation of the skin tone is 315, 94, and 37, individually. The recognized hand picture is resized to 200x200 to make the motion acknowledgment invariant to the picture scale.

Figure 2: The procedure of hand detection.

5.2 Experimental Results

Data Set:
In the tests, two informational indexes of hand signals are utilized to assess the proposed strategy's presence. Informational index 1 is a picture assortment of thirteen signals. For each signal, 100 pictures are caught. In this way, there is a sum of 1300 pictures for hand signal acknowledgment. All the signal pictures have a place with three females and four guys. The size of one signal picture is 640x480.

Another informational index is gathered from 10 subjects, and it contains ten signals for numbers 0 to 9. In this way, there is a sum of 10x10x10 cases. The informational index caught in jumbled foundations is an excellent test for hand motion acknowledgment. Plus, for each motion, the subject stances with varieties close by direction, scale, verbalization. We contrast our strategy and FEMD on the informational collection.
Performance Evaluation on Data Set 1:

1. Classification Accuracy:
   To quantify the proposed hand signal acknowledgment strategy exhibition, the characterization exactness is assessed in the investigations. In the preparation stage, the standards separating the thirteen signals are created. At that point, the standard classifier utilizes the principles to foresee the personality of the testing picture. In Figures 3, 4, 5, 6, and 7, the five motions' acknowledgment is illustrated. There are six subfigures in each figure: the pictures indicating the parallel hand picture, the palm print, the wrist line, the aligned hand picture, the palm veil, the distinguished fingers, and finger and signal acknowledgment separately. In the subfigure of finger and motion acknowledgment, the name of the motion is anticipated. The anticipated mark is appeared behind "Answer."

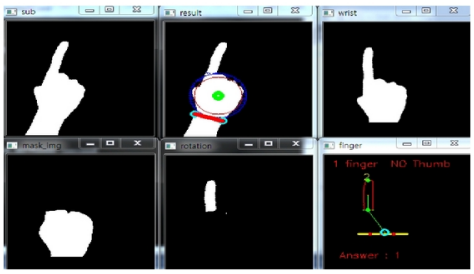

Figure 3: The recognition of the hand gesture 1.

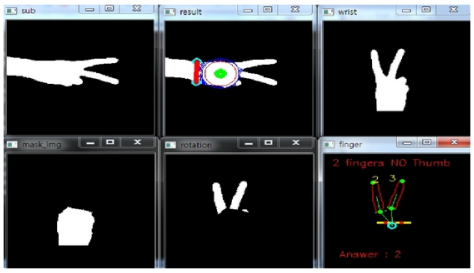

Figure 4: The recognition of the hand gesture 2.

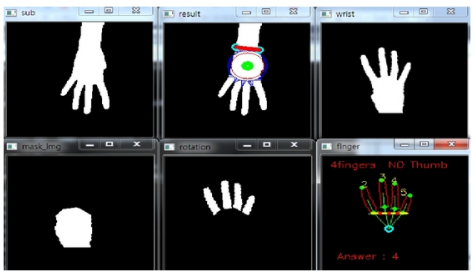

Figure 5: The recognition of the hand gesture 3.

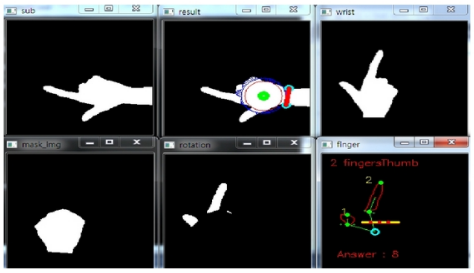

Figure 6: The recognition of the hand gesture 4.

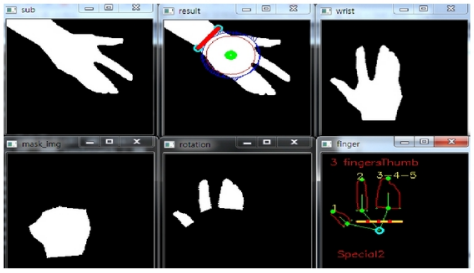

Figure 7: The recognition of the hand gesture 5.

The arrangement aftereffects the absolute of 1300 pictures is summed up with a disarray lattice in Table 1. In the disarray network, the main section and the last line are the marks of the signals. Different passages of the lattice record the quantities of the signal pictures anticipated as the comparing marks. For instance, for the primary line, the numbers 99 and 1 are in the sections relating to the marks 1 and 3, separately. It implies that there are 99 and 1 pictures anticipated as the marks 1 and 3 in the 100 testing pictures of the motion 1. Along these lines, for the testing pictures of signal 1, the order precision is 99%. As appeared in the disarray lattice, the proposed strategy performs well and acquires high arrangement precision. The all-out grouping exactness of the 1300 testing picture is 96.69%. In the disarray framework, the offers of S2 and S3 are misclassified as 5. The explanation is portrayed as follows: for specific offers of S2 and S3, the fingers do not remain shut. That is, there is an opening between two fingers. In this way, in these cases, the offers of S2 and S3 are misclassified as 5.

1	99		1									
2		94	6									
3		2	95	3								
4			4	95	1							
5				3	93	4						
6						100						
7	4						96					
8	2		5					92	1			
9									100			
0	1									99		
S1			1								99	
S2					2							98
S3					3							97
	1	2	3	4	5	6	7	8	9	0	S1	S2	S3

Table 1: The confusion matrix of hand gesture recognition on data set 1.

2. Time Cost:
The time cost for perceiving the motions is accounted for in Table 2. In the table, the unit of the time cost is second. An incentive in the subsequent line is the averaging runtime of 100 pictures of one signal. For the complete 1300 pictures, the average time cost to perceive hand motions is

0.024 seconds. The analyses are run on the PC Intel i7-2630 2.00 GHz CPU and 4 GB RAM. The proposed strategy is advantageous and can meet the prerequisite of ongoing applications.

1	2	3	4	5	6	7	8	9	0	S1	S2	S3
0.024	0.021	0.022	0.024	0.027	0.023	0.026	0.022	0.025	0.022	0.022	0.026	0.021

Table 2: The runtime of hand gesture recognition.

5.3 Relevance of case study method for this research

This technique can be used for understanding the language of dumb people—talk with hand gestures. So, in this method, the expression can be extracted and converted into natural automated language. The hand area is identified from the foundation by the foundation deduction strategy. At that point, the palm and fingers are divided. Based on the division, the fingers in the hand picture are found and perceived. A primary principle classifier refines the acknowledgment of hand signals. The exhibition of our technique is assessed on an informational index of 1300 hand pictures. The test results show that our methodology performs well and is suitable for continuous applications. Besides, the proposed technique beats the condition of-craftsmanship FEMD on a picture assortment of hand motions.
The presentation of the proposed technique profoundly relies upon the consequence of hand recognition. On the off chance that there are moving items with the shading like that of the skin, the articles exist in consequence of the hand identification and afterward corrupt the hand signal acknowledgment presentation. Be that as it may, the AI calculations can separate the hand from the foundation. ToF cameras give the profundity data that can improve the exhibition of hand recognition. Along these lines, in future works, AI strategies and ToF cameras might be utilized to address the perplexing foundation issue and improve the strength of hand recognition.

6 Significance of Research

Introducing a system that can convert sign language into words can be very beneficial to all those who have difficulty or cannot establish proper verbal communication. The primary use of sign language is for the deaf-mute community who cannot communicate with spoken language. Nowadays, unlike olden days mute people are becoming more outgoing and do not want to depend on anybody else for communication. So for such individuals, it is important that the general public around them can understand what they are trying to express through sign language. A system that can understand and then translate the sign language to verbally understandable language can break the barrier mute people have while communicating.
In Human-Computer Interaction (HCI) gesture recognization is a significant advantage. Hand gestures convey a major part of information transfer in our everyday life. Being able to create an understanding of hand gestures between human and machine can boost communication astonishingly. Feilds like Graphic design, Game development, 3D modeling, Automatic driving, etc, can substantially benefit from hand gesture recognition system.

7 References

Gestures - Social Sci LibreTexts
Hand Gesture Recognition - Sajjad Ansari
Hand Gesture Recognition - Anbumani, Vinod, Dinesh
Hand Gesture Recognition System - Jyothi Nayak