International Journal of Scientific Research & Engineering Trends
Volume 8, Issue 3, May-Jun-2022, ISSN (Online): 2395-566X

Alphaneumeric Hand Gesture recognition System

Prof. Megha Beedkar, Mr.Barve Saurabh, Mr.Jadhav Rajan, Mr.Raibole Vinay, Ms.Sonawane Pooja
Department of Electronics and Telecommunications,
Genba Sopanrao Moze College of Engineering,
Savitribai Phule Pune University,
Balewadi, Pune

Abstract- The project introduces an application using computer vision for Hand gesture recognition. A camera records alive video stream, from which a snapshot is taken with the help of interface. The system is trained for each type of count hand gestures (one, two, three, four, and five) OR Alphabets (A, B and C) at least once. After that a test gesture is given to it and the system tries to recognize it.

Keywords- Human interface devices, Automation, CNN, Controller.

I. INTRODUCTION

With the rapid growth of artificial intelligence technology, many intelligent applications have been developed such as smart TV and intelligent robots. The most natural way for humans to communicate with these intelligent systems is dynamic gestures. In recent years, air writing has become one of the most popular dynamic gestures. It is defined as writing alphanumeric with hand or finger movements in a three-dimensional (3D) free space.

Air writing is particularly useful for user interfaces that do not allow the user to type on the keyboard or write on the touch pad/touch screen or for text input for intelligent system control. Air-writing recognition is closely related to motion gestures or sign language recognition. Motion gesture recognition methods can be roughly divided into two categories: device based and device-free. . However, the requirement for handheld or worn devices by contrast, in the device-free method, users do not need to hold or wear any devices; hence, this method is more convenient than the device-based method.

Device-free methods can be further divided into vision-based and radio-based methods. The former utilizes 2D or 3D cameras to capture gesture input images. The latter uses radio sensors such as radar or Wi-Fi to obtain gesture signals. Air writing can be realized in three manners isolated, connected, and overlapped air writing. In isolated writing, the letters are written in an imaginary box with fixed height and width in the field of view of an image, one at a time. In connected writing, multiple letters are written from left to right, which is similar to writing on a paper.

In the last manner, one can write multiple letters stacked contiguously one over another in the same imaginary box. We study the isolated writing style in this paper. Isolated writing is the most essential and popular method. Motion characters are isolated alphanumeric letters written in a uni stroke.

The steps involved in air-writing recognition generally include hand/finger tracking, feature extraction and classification. The fundamental problems in isolated writing include (a) tracking of hand and/or fingers, (b) segmentation of writing acts (or push-to-write), (c) restrictions on the users' writing due to the limitation of an imaginary box, and (d) infraclass variability of the writing patterns of a letter.

For vision-based methods, the first problem has been addressed, but different solutions must be used for 2D and 3D image sensors. 2D camera-based systems often utilize colour markers on fingers to increase tracking performance since finger tracking without markers is challenging. 3D camera based systems address the hand/finger tracking problem well simply using the depth information provided by 3D image sensors such as Kinect, Leap Motion Controller (LMC), or Intel Real Sense camera. Air writing lacks a reference position on the writing plane and thus lacks the beginning and end points of a stroke. Therefore, it needs to automatically detect the start and end coordinates of the characters written in the air. This is referred to as segmentation of writing acts, or the so-called push-to-write problem. One of the possible solutions is to use a specific posture to signal the endpoint of a writing act, e.g., a fist posture.

However, this will increase the number of gestures that users must remember. When depth information is available, the segmentation of writing acts can be done by merely using a depth threshold. In summary, 3D camera-based systems address the first two problems more conveniently than 2D camera-based systems. However, 3D systems are more complex and expensive.

The imaginary box limits the range of writing. It reduces the variations of letter input such as position, scaling or rotation of the written image. This alleviates the burden of the subsequent processing. Nevertheless, from the users' perspective, this method causes inconvenience and restrictions of users on writing. In this paper, we design a simple yet effective air-writing recognition approach based on deep convolutional neural networks (CNN's) using a single low-cost 2D web camera. Our approach solves the first three problems in a convenient manner. Furthermore, it can work in a real-time smart-TV-like environment. The major contributions of this article are: (a) A robust air-writing trajectory acquisition algorithm based on a web camera. The algorithm combines skin and moving features to detect the moving skin region and then applies the Cam-shift algorithm to track the moving hand. It performs hand tracking only, thus avoiding the complicated procedures for finger tracking. In addition, the proposed algorithm solves the push-to-write problem without using a delimiter.

Furthermore, it does not utilize an imaginary box; hence, users can write freely in the air without any restrictions. (b) A novel data prepossessing scheme. The scheme normalizes the x and y coordinate sequences of the writing trajectory and then combines them into 1D and 2D arrays. The two types of data arrays are employed to train 1D-CNN and 2D-CNN. These simple data arrays make the designed CNN's The algorithm combines skin and moving features to detect the moving skin region and then applies the Cam-shift algorithm to track the moving hand.

It performs hand tracking only, thus avoiding the complicated procedures for finger tracking. In addition, the proposed algorithm solves the push-to-write problem without using a delimiter. Furthermore, it does not utilize an imaginary box; hence, users can write freely in the air without any restrictions. (b) A novel data prepossessing scheme. The scheme normalizes the x and y coordinate sequences of the writing trajectory and then combines them into 1D and 2D arrays. The two types of data arrays are employed to train 1D-CNN and 2D-CNN.

These simple data arrays make the designed CNN's simpler and more effective than the use of complex written images. (c) A CNN-based air-writing recognition system using a low cost web camera. It achieves real-time recognition with a high accuracy of more than 99% and very low network complexity. It outperforms the popular approaches using written images as input.

Recent developments in computer software and related hardware technology have provided a value added service to the users. In everyday life, physical gestures are a powerful means of communication. They can economically convey a rich set of facts and feelings. For example, waving one's hand from side to side can mean anything from a "happy goodbye" to"caution". Use of the full potential of physical gesture is also something that most human computer dialogues lack. The task of hand gesture recognition is one the important and elemental problem in computer vision. With recent advances in information technology and media, automated human interactions systems are build which involve hand processing task like hand detection, hand recognition and hand tracking. This prompted my interest so I planned to make a software system that could recognize human gestures through computer vision, which is a sub field of artificial intelligence.

II. SYSTEM ARCHITECTURE

Any hand gesture recognition system requires some basic stages. These stages are-

- Gesture module and image acquisition,
- Hand segmentation,
- Tracking,
- Feature extraction and
- Classification

1. Gesture Module and Image Acquisition:

The video frames of the user performing the gestures can be captured with the help of a low-cost webcam or webcam built-in in a personal laptop computer. In our work, the videos have been captured at a resolution of 720x480 at 40 framesper second (fps).

2. Hand Segmentation:

Proper hand segmentation from other body parts and background is very crucial for the overall efficacy and the effectiveness of any vision based hand gesture recognition system. In this work, a skin colour based hand segmentation

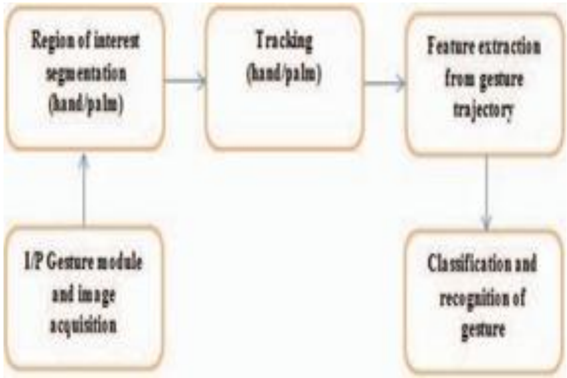

Fig 1. Block diagram for hand gesture recognition system.

3. Hand Tracking:

Tracking in computer vision refers to the technique which constantly monitors the consecutive positions/locations of the region of interest (ROI) (hand in this case). In order to find the gesture trajectory, the centroid / center of gravity (CoG) of the segmented hand is found out first. The centroid can be found out by moment calculation. A moment is a gross characteristic of a contour computed by integrating or summing over all of the pixels of the contour [12]. In this work, the only contour left out after segmentation is hand contour as hand contour is of interest here for hand gesticulation

4. Feature Extraction:
Feature extraction is very important in terms of giving input to a classifier. Transforming the input data set into a set of features is called feature extraction. Feature extraction is a kind of dimensionality reduction technique [13].

III. EXPERIMENTAL RESULTS

1. Tracking Results of Gestures:
The hand tracking with drawn gesture alphabet (left) along with its centroids (right) are shown in Fig. 4. The right side image shows the different positions of the hand centroid at different time-instants during the gesticulation of the alphabet.

Fig 2. Gesture alphabet ‘A’ (left) and its tracking result (right).

2. Isolated and Continuous Gestures:
We have considered eight isolated gestures viz. ‘A’, ‘B’, ‘C’, ‘D’, ‘E’, ‘F’, ‘O’ and ‘Q’ and six continuous gestures viz. ‘BA’, ‘AD’, ‘CA’, ‘CB’, ‘CF’ and ‘DB’. Fig. 5 shows the isolated gesture ‘B’ and a continuous gesture ‘BA’ considered in the work.

3. Tabulated Results of Isolated and Continuous Gestures:
Let S and T denote the number of trained samples and tested samples, respectively and C_S, R_S and W_S denote the that of correctly spotted samples, rejected samples and wrongly spotted samples, respectively. RR is recognition rate.Overall recognition rate (ORR) in %.

Table 1. Isolated Gesture Confusion Matrix.

Gesture	S	T	C_S	R_S	W_S	RR(%)=$\frac{C_S}{T}\times 100\%$
A	30	30	27	1	2	90.0%
B	30	30	26	2	2	86.7%
C	30	30	29	1	0	96.7%
D	30	30	25	2	3	83.3%
E	30	30	29	1	0	96.7%
F	30	30	29	0	1	96.7%
O	30	30	28	1	1	93.3%
Q	30	30	29	1	0	96.7%

Table 2. Continuous Gesture Confusion Matrix (MLP).

Gesture	S	T	C_S	R_S	W_S	RR(%)=$\frac{C_S}{T}\times 100\%$
BA	30	30	26	2	2	86.7%
AD	30	30	25	2	3	83.3%
CA	30	30	28	2	0	93.3%
CB	30	30	27	1	2	90.0%
CF	30	30	29	1	0	96.7%
DB	30	30	26	3	1	86.7%
AA	30	30	26	2	2	86.7%

Table 3. Continuous Gesture Confusion Ma-Trix (FTDNN).

Gesture	S	T	C_S	R_S	W_S	RR(%)=$\frac{C_S}{T}\times 100\%$
BA	30	30	25	1	4	83.3%
AD	30	30	25	3	2	83.3%
CA	30	30	27	1	2	90.0%
CB	30	30	28	2	0	93.3%
CF	30	30	29	0	1	96.7%
DB	30	30	24	4	2	80.0%
AA	30	30	25	1	4	83.3%

IV. CONCLUSION

A test paper checker was developed in this study that will recognize a handwritten.Text using Intelligent Character Recognition (ICR) for Alphanumeric Characters, Which is based on OpenCV for image processing and SVM for classification. It has a 93.0769% accuracy of correct recognition of characters.

For future work, deep learning algorithms such as convolutional neural networks will be implemented.

REFERENCE

[1] Sheng-Yu Peng,KanoksakWattanachote ,Hwei-Jen Lin,Kuan-Ching Li, ‘Real- Time Hand Gesture Recognition System for Daily Information Retrieval from IJSER © 2015 http://www.ijser.org International Journal of Scientific & Engineering Research, Volume 6, Issue 3, March-2015 191 ISSN 2229-5518 Internet’; Fourth International Conference on Ubi-Media computing 978-0-7695-4493-9/11 2011 IEEE DOI 10.1109/UIA.2011.45.
[2] Siddharth S. Rautaray, Anupam Agrawal, ‘Real Time Multiple Hand Gesture Recognition System for Human Computer Interaction’ I.J. Intelligent Systems and Applications, 2012, 5, 56-64 Published Online May 2012 in MECS.
[3] Jagdish Lal Raheja, Radhey Shyam, Umesh Kumar and P Bhanu Prasad, ‘Real- Time Robotic Hand Control using Hand Gestures’; in Proceedings of the Inter- national Conference on Machine Learning and Computing, pp. 12- 16, 2010.

[4] Yuvraj V. Parkale ‘Gesture based operating system control ’2012 Second International Conference on Advanced Computing & Communication Technologies. 978-0-7695-4640-7/12 © 2012 IEEEDOI 10.1109/ ACCT.2012.58.
[5] Nandkishor P. More, Sheetal S. Thokal, Aniket S. More, Bhagyavati V. Patil https://deepai.org/machine-learning-glossary-and-terms/pytorch