Computer Vision Based Air Writing Recognition

AkashKumar M B[1], Ashwini R[2], Dhanaraj S[3], Shwetha B K[4], Dr.V Keerthika[5]
[1][2][3][4]Final year students,[5]Associate professor
Dept of CSE, Dayananda sagar Academy of Technology and Management,

Abstract

In this recent era, there is a continuous need for automated appliances. Air writing Recognition is a constant video based venture which permits the drawing and composing of English letters in order or number 0-9 through the air, before a web camera or versatile camera. To begin with, it tracks the fingertip in the video edges and afterward applies optical character acknowledgment innovation over the plottedpictures to perceive the composed letters inorder or numbers. This gives ahuman framework connection without necessities of any gadgets for contributing the character through the air. This venture is actualizing by using the Python and Open CV. Here, numbers of various formed characters are utilized. This venture is of programming based methodology and pertinently exceptionally basic, quick, and simple. It doesn't requires any sensors or any equipment's other than a web camera, and is can be utilized to actualize to any remaining dialects utilizing their particular informational indexes, but if any upsetting articles in the foundation it might decrease the outcomes to the clients.

Keywords : *Air-writing, EMNIST dataset, pen storkes, Hue-Saturation-Value*.

---

1. Introduction

Handwriting analysis is a classic, well-explored problem in the introductory machine learning that catches many of the important topics of conventional neural networks. When we discussed handwriting analysis during project work, we realized that handwriting analysis is a project that could be extended with other

machine learning concepts for an interesting combination applications.
This project is to use a combination of computer vision and handwriting recognition to create a system that acts as a virtual whiteboard. Our model recognizes gestures written in the air and converts it into text. Users would be able to write "airwords" facing a web camera either real time and have those gestures translated into letters or digits.

1.1 Motivation

We choose computer vision and handwriting analysis because we believe air writing recognition is a worthwhile topic to pursue. It allows for an new text input interface that does not require much more than the computer itself.

We were not the first to pioneer the idea. Other systems such as the HTC Vive [5] virtual reality (VR) system have products that follow a similar idea. The Vive has a virtual whiteboard experience that allows users to write in the air while immersed in VR, but the system comes at a high cost that also requires an addition expensive tracking system.

Objective

The objective of our project is to create a simple system that needs a computer and built-in webcam to recognize different letters and digits written in the air.

2. MODELS AND ALGORITHMS

2.2 Object Tracking

In our model, object tracking is realized with detection of the target object through image analysis of each video frame. OpenCV [2], an open source computer vision and machine learning library, provides both classic and state-of-the-art algorithms relevant to object tracking. Therefore, we decided to utilize assets from the OpenCV library to detect, track, and save the trajectory of the target object as its position shifts throughout the video

After setting the lower and upper boundaries of our target object, we preprocess each frame by resizing and reducing Gaussian noise. Then, we construct a binary mask around the object(s) and perform morphological transformations to clean up. Finally, we detect the contours and form an enclosing circle around the object, saving the center coordinates of the circle as our tracked points.

2.3 Setting Parameter Values

In order to detect the object, our system filters the image for items of a specific color. Our system searches for a color within a specified range to provide a margin to account for slight variations in object color. Our target object does not necessarily always exhibit the exact same color values, whether due to environmental lighting conditions or to the use of different objects that slightly differ in color values.

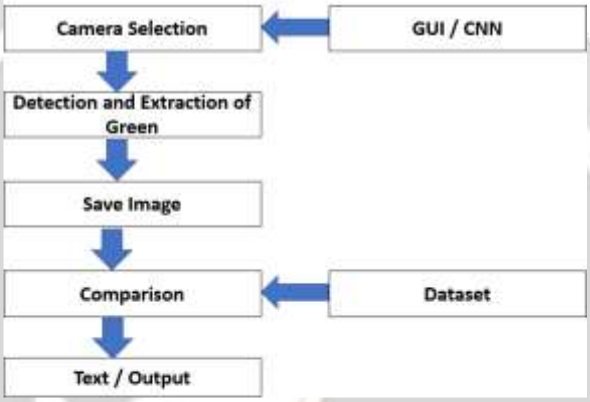

Figure 1. Block Diagram of proposed system Model

focuses are saved and overlaidonto a parallel picture that will be preprocessed and afterward passed as contribution to the penmanship arrangement model.

2.4 Frame Preprocessing

Since this system performs analysis on each frame in the video, we resize each frame while maintaining aspect ratio to increase efficiency and smoothness during run time. Next, we apply a Gaussian Blur to reduce the Gaussian noise, statistical noise with a probability density function equal to that of a Gaussian distribution, of the image. Finally, we convert the image from BGR to the Hue-Saturation-Value (HSV) color space to match our color boundaries that are set to HSV values. We chose to use the HSV color space, because the BGR color space values correlate with the amount of light hitting the object, whereas, image descriptions in terms of hue, saturation, and lightness are more relevant.

2.5 Mask Construction

To simplify detection and to match the training images in the EMNIST [4] dataset, which uses binary images, we create a binary mask over the objects found to be in the target color range. A series of morphological transformations, nonlinear operations related to the shape of the object in the binary image, are then performed on the binary mask to clean up the image and prepare it for contour detection.

We first perform an erosion operation that shrinks the boundaries of the object and removes small white noise. Next, we perform a dilation operation to expand the boundaries of the object. Although erosion and dilation might seem like a counter-intuitive sequence, as the former shrinks the object while the latter expands it, it actually is very reasonable. While erosion shrinks the object boundaries, it also removes noise and cleans up the image. Dilation then simply re-expands the boundaries without the noise.

2.6 Contour Detection

We find the contours using an OpenCV algorithm which calculates the hierarchy of contours in the image and compresses it. After obtaining a list of contours in the mask, we select the largest one and compute the minimum enclosing circle around that contour. If the enclosing circle has a radius larger than a predefined size and exists within the correct whiteboard margins, which will be further explained in the Depth Analysis section, then we update our list of tracked points. The tracked points are saved and overlaid onto a binary image that will be preprocessed and then passed as input into the handwriting classification model.

3. DEPTH ANALYSIS

3.1 Segmentation Problem

For our undertaking, we were goal- oriented in the sorts of characters we needed to distinguish. Maybe than stay with the exemplary 10 digits or 26 lowercase letters, we needed to make a framework that was prepared on 62 classes: 10 digits, 26 lowercase letters, and 26 capitalized letters. When chipping away at signal following, we understood that large numbers of the characters we needed to perceive required different "pen strokes" or portioned lines to compose. On the off chance that we were doing just lowercase letters and digits we might have had the option to stay with persistent info, yet numerous capitalized letters, (for example, letters A, E, and I) require portioned lines to be precisely positioned for exact separation for both computer vision and human discernment. To take care of this issue, we have concocted a few arrangements including flipping the item to a non- identified shading, covering the article, and utilizing foreordained keystrokes to demonstrate stroke beginnings and closures. Eventually, we chose the execute the technique we thought would be the most instinctive answer for the client the utilization the pen lift.

Figure 2. EMNIST Dataset

In the wake of noticing the normal composing propensities for clients, we understood that whether composing on paper or on vertical whiteboards, clients normally lifted their pens toward the finish of each stroke and supplanted it toward the start of the following stroke. We made an interpretation of that activityinto an adjustment of profundity from the camera. We needed clients to have theoption to lift their "pen" off an imperceptible whiteboard to flag the finish of a stroke, and a critical change inside and out made by the "lifting" activity would be the marker to the framework that the stroke was over Implementing profundity examinationended up being quite possibly the most intriguing difficulties of this venture. In the area of PC vision, profundity examination is being investigated on numerous fronts, the most well known being the stereoscopic methodology.

This technique, notwithstanding, requires two unique points of a similar field of vision to utilize optic calculation to locate profundity. At the point when understood, this methodology requires the base two distinct cameras to be mounted at a particular stature which isn't an arrangement the normal shopper would have. Since our center objective was to expand ease of use and availability, we imagined that it is to discover an answer for ascertaining profundity utilizing just one detecting gadget, for example, a computer web camera)

4. Result and Discussion

Our Handwriting classification network is prepared on the EMNIST (broadened MNIST) Dataset of written by hand number digits [4]. This dataset is a famous decision for penmanship order organizations, and we decided to utilize the"by class" design what isolates betweendigits, capitalized, and lowercase letters–to prepare on. The EMNIST dataset of more than 814,225 examples is now parted 90% preparing and 10% testing, yet we chose totake the preparation set of information to be both our preparation and our assessmentinformation. We split this dataset 8:2, preparing to assessment, utilizing the Scikit-Learn train_test_split strategy [6]. This subsequent division permitted us to save the first 10% test information for concealed testing.

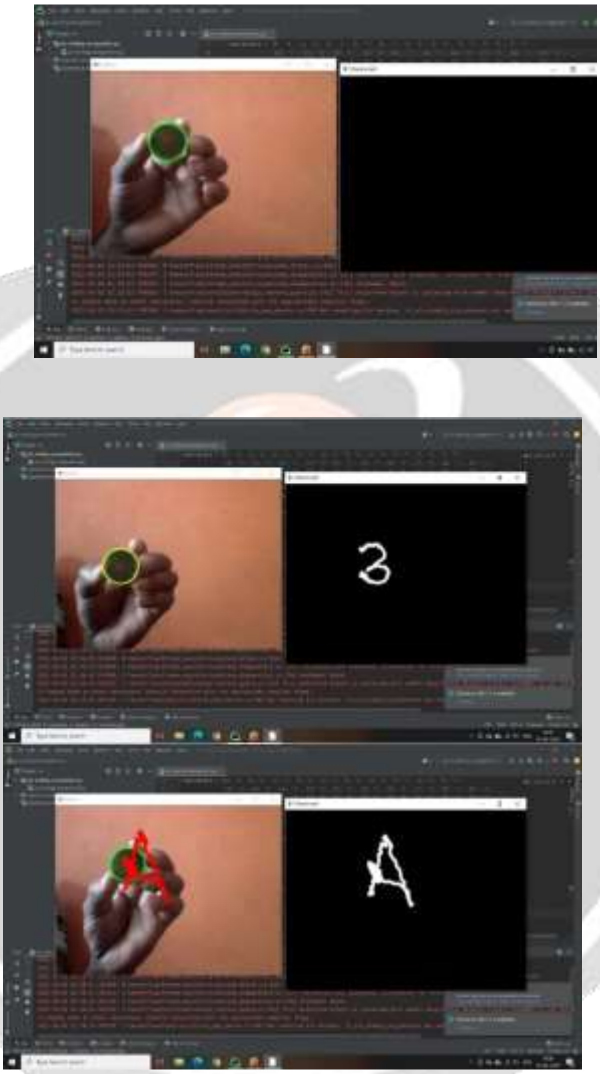

Figure 3 Input Generation

Conclusion

This proposed work presents a video-based calculation which permits composing ofEnglish letter sets and numeric numbers on air utilizing web camera. This paper has two primary undertakings: first it tracks the finger in the video casings and afterward apply CNN over plotted pictures to perceive the composed characters. nonetheless, this work gives a characteristic human framework connection so as to not need any console, mouse and so on for contributing the character. It simply requires a web camera and versatile camera for rearrange a finger. we have built up this undertaking utilizing OpenCV with Python language.

Future work

This venture at this stage were executed to perceive the English letter sets and numbers whatever draw on

the air utilizing Web camera or Mobile camera utilizing the language python and OpenCV libraries. In Future Geometric images concentrate to utilize CCTV camera image and video capturing.

Reference

1. Chinnam Datta Sai Nikhil, ChukkaUma Someshwara Rao, E. Brumancia, K. Indira, T. Anandhi,P. Ajitha," Finger Recognition and Gesture based Virtual Keyboard" Journal of International Conference on Communication and Electronics System, Vol. 4, PP. 1324-1328, July

   2020.
2. G. Bradski. 2000. The OpenCV Library. Dr. Dobb's Journal of Software Tools (2000).
3. Sugnik Roy Chowdhury, Sumit Pathak, M.D. Anto Praveena, "Gesture Recognition Based Virtual Mouse and Keyboard", Journal of International Conference on Trends in Electronics and Informatics, Vol. 5, PP.585-589, Aug 2020.
4. Gregory Cohen, Saeed Afshar, Jonathan Tapson, and André van Schaik. 2017. EMNIST: an extension of MNIST to handwritten letters. arXiv preprint arXiv:1702.05373
5. Dharmaraj Ojha, Rajesh George Rajan, "Histogram based Human Computer Interaction for GestureRecognition", Journal of International Conference on Electronics Communication and AerospaceTechnology, Vol.4, PP. 263- 266,2019.
6. Yung-Han Chen, Po-Chyi Su, Feng- Tsun Chien, "Air-Writing for Smart Glasses by Effective Fingertip Detection", Journal of Conference on Consumer Electronics, Vol.2, PP.381- 382,2019.
7. Jugannathan MJ, Surya M, Aqib Muhammed BT, RJ Poovaraghavan,"Finger recognition and gesture based augmented keyboad", International Journal of Advance Research Idea and Innovations in Technology, Vol. 4, PP.662-665, 2018.
8. Miroslav Hagara, Jozef Pucik, "Fingertip Detection for Virtual Keyboad Based on Camera", Czech Republic, Vol. 5, PP. 356- 360, 2013.
9. Mingyu Chen, Ghassan AlRegib, and Biing-Hwang Juang, "Air- Writing Recognition—Part II: Detection and Recognition of Writing Activity in Continuous Stream of Motion Data", This article has been accepted for inclusion in a future issue of this journal, Vol. 9, PP. 2168-2291, 2015.
10. W.S. Mada Sanjaya, Dyah Anggraeni, Kiki Zakaria, Atip Juwardi, Madinatul Munawwaroh, "The Design of Face Recognition and Tracking for Human-RobotInteraction", International Conferences on Information Technology, Vol. 6, PP. 315-320,2017.
11. HTC and Valve Corporation. 2016. HTC Vive. https://www.vive.com/eu/.