IJMTST is UGC Approved Journal with Journal Id: 43137

International Journal for Modern Trends in Science and Technology, 8(S05):159-164, 2022

ISSN: 2455-3778 online
DOI: https://doi.org/10.46501/IJMTST08S0528
Available online at: https://www.ijmtst.com/vol8si05.html

Building a Air Canvas using Numpy and Opencv in Python

D. Vijendra Kumar, G.Vijaya Raj Siddarth, R. Venkata Satya Sravani, I.Vishnu Vardhan Reddy, Y. Lalitha Sri Naga Durga Vyshnavi

Department of Electronics and Communication Engineering, Godavari Institute of Engineering and Technology(A), JNTUK, Kakinada.

To Cite this Article
D. Vijendra Kumar, G.Vijaya Raj Siddarth, R. Venkata Satya Sravani, I.Vishnu Vardhan Reddy and Y. Lalitha Sri Naga Durga Vyshnavi. Building a Air Canvas using Numpy and Opencv in Python. International Journal for Modern Trends in Science and Technology 2022, 8(S05), pp. 159-164. https://doi.org/10.46501/IJMTST08S0528

Article Info
Received: 26 April 2022; Accepted: 24 May 2022; Published: 30 May 2022.

ABSTRACT

Writing in air has lately become one of the most interesting and difficult research areas in image processing and pattern recognition. A wide range of human-machine interactions benefit from this technology's inclusion. Numerous researches have investigated new methods and tactics for reducing processing time and improving recognition accuracy. In Computer Vision, object tracking is a crucial part of the task. Object tracking systems have become more popular as a result of faster computers, more affordable and higher quality video cameras, and the increasing need for automated video analysis. Typically, video analysis involves three major steps: identifying the object, tracking its movement from frame to frame, and lastly analyzing its behavior. Object tracking takes into account four issues: object representation, tracking feature selection, object identification, and object tracking. Object tracking algorithms are utilized in a variety of real-world applications, including video indexing, autonomous surveillance, and vehicle navigation. Motion-to-text converters for intelligent wearable electronics that can write in the air are the subject of this study. This initiative serves as a time capsule for transitory movements. It will use computer vision to track the path of the finger. Messages, emails, etc. maybe be sent using the text that has been generated. It will be a great help for deaf individuals to communicate. It's a good way to communicate without having to write.

Keywords:Air Writing, Character Recognition, Object Detection, Real-Time Gesture Control System, Smart Wearable, Computer Vision.

1.INTRODUCTION

For most people nowadays, digital art takes the place of traditional writing as their primary means of expression. We call it "digital art" when we use digital medium to convey our artistic expressions and messages. Digital manifestation is distinguished by its reliance on modern science and technology. There were different kinds of art forms before digital art was created. There are four types of art: visual art, audio art, audio-visual, and audio-visual imaginative. This includes everything from literature to painting to sculpture to architecture to music to dance and theatre. However, digital and traditional creative forms are inextricably intertwined. However, human life's fundamental necessities are the driving force behind societal progress. There is a similar phenomenon in art. The link between digital and traditional art is now so intertwined that we need to study both of them meticulously. The old-school method of writing is when you use a pen and paper or a blackboard. Developing a hand motion recognition system that can be used to digitally write has been the fundamental objective of digital art. It is possible to produce digital art by utilizing a keyboard, touch-screen surface, digital pen, stylus or even electronic hand gloves. Machine learning algorithms and python programming, on the other hand, are used to recognize hand gestures in order to facilitate human-machine communication organically. We're seeing a rise in the need for natural 'human-computer interaction (HCI) systems as technology progresses.

2. PROPOSED SYSTEM

This computer vision experiment uses an Air canvas, which allows you to draw on a screen by waving a finger equipped with a colourful tip or a basic coloured cap. These computer vision projects would not have been possible without OpenCV's help. There are no keypads, styluses, pens, or gloves needed for character input in the suggested technique.

2.1 FEATURES OF AIR CANVAS

- Can track any specific colored pointer.
- User can draw in four different colors and even change them without any hustle.
- Able to rub the board with a single location at the top of the screen.
- No need to touch the computer once the program is run.

2.2 SCOPE

The scope of computer vision and that of OpenCV is huge.Object (human and non-human) detection in both commercial as well asgovernmental space is huge and already happens is many ways.

Transportation - with autonomous driving in ADAS (Automated Driver Assist System) in traffic signs detection, pedestrian detection, safety features such driver fatigue detection etc.

Medical imaging - mammography, cardiovascular and microscopic image analysis (I'm not a medicine guy but I am hearing that a whole lot of computer imaging aided decision-making such as automated detection and counting of microorganisms will involve use of OpenCV)

Manufacturing - Ton of computer vision stuff there as well such as rotation invariant detection on a conveyer belt with detection of stoke of robotic gripping.

Public order and security - pedestrian/citizen detection and tracking, mob management, prediction of future events.

2.3 APPLICATIONS

Python may be used to quickly analyze photos and videos and extract meaningful information from them, thanks to the many methods provided in OpenCV. Other frequent uses include,

Image Processing:

There are several ways in which the OpenCV may be used to process and interpret images, such as altering their shape, colour, or extracting important information from the supplied picture and writing it into a new image.

Face Detection:

By employing Haar-Cascade Classifiers, either from locally recorded videos or photos or from live streaming through web camera.

Face Recognition:

In order to identify faces in the films, face identification was performed using OpenCV by generating bounding boxes (rectangles) and subsequently model training using ML methods.

Object Detection:

OpenCV and YOLO, an object identification method, may be used to identify moving or stationary objects in images and videos.

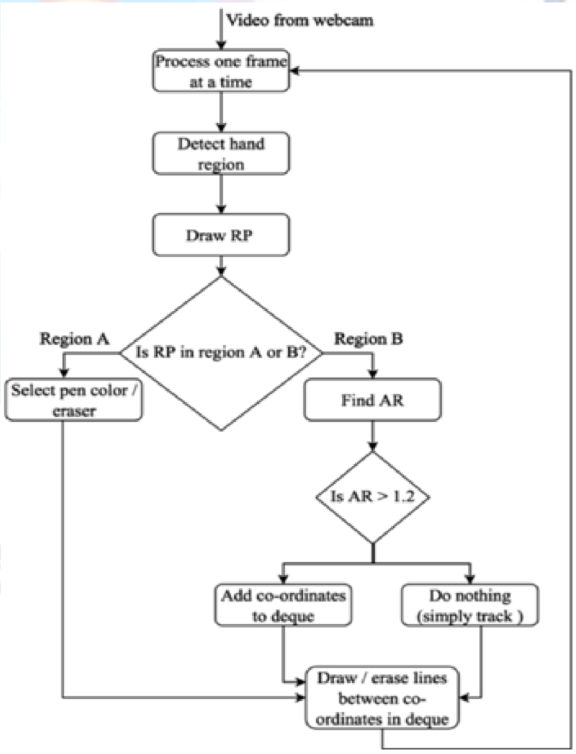

Figure1: Flow Chart of Proposed Method

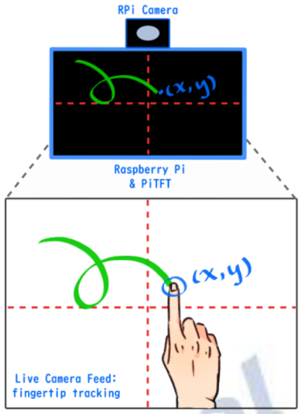

Figure2:Block Diagram of System

3. METHODOLOGY

This system needs a dataset for the Fingertip Detection Model. The Fingertip Model's primary purpose is used to record the motion, i.e., the air character.

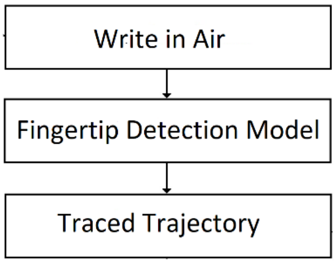

Figure3: Fingertip detection model

A. Fingertip Detection Model:

For air writing, all you need is a stylus or one of the many colored airpens on the market. But the system relies on the tip of a finger to operate it. We think that individuals should be able to write in the air without having to carry a pen around with them. Every frame has been analyzed using Deep Learning techniques to generate a list of coordinates for each fingertip.

B. Techniques of Fingertip Recognition Dataset Creation:

1. Video to Images: Hand motion films were recorded for two seconds and then transferred to a variety of diverse contexts. This footage was sliced up into 30 individual pictures, which can be seen in More than 2000 photos were taken. Data was manually annotated in this dataset. The most accurate model developed using this dataset has a precision of 99.5 percent. Dataset monotony resulted from using a same movie with identical surroundings to produce 30 different photos. It follows that for non-continuous environments like those in the sample, the model failed miserably..

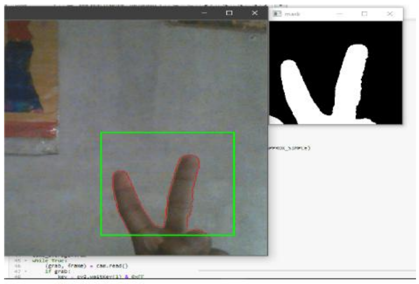

Figure4: Finger Tip Recognition Dataset Creation

2. Take Pictures in Distinct Backgrounds:

We built a new dataset to solve the problem of the prior method's lack of variety. This time, we were well aware of the fact that gestures were required in order to operate the system. To that end, we gathered the four hand gestures shown in Figure 4. The goal was to create a model that could recognize the tips of all four fingers with equal efficiency. The user would be able to manage the system by showing the number of fingers he has. One index finger is all that's required to start writing, and the user can now - quickly write by showing one index finger, convert this writing motion to e-text using two fingers, add space using three fingers, hit backspace using five fingers, and then show 1,2,3 fingers to select the first, second, and third predictions respectively. Show five fingers to exit the prediction mode. The photos in this collection totaled 1800. An auto-labeled dataset was created using a software that has previously been trained on this dataset Once the photos had been properly renamed, a new model was introduced. At least 94% of the time, it worked. Different backdrops proved to be suitable for the model, unlike the previous one.

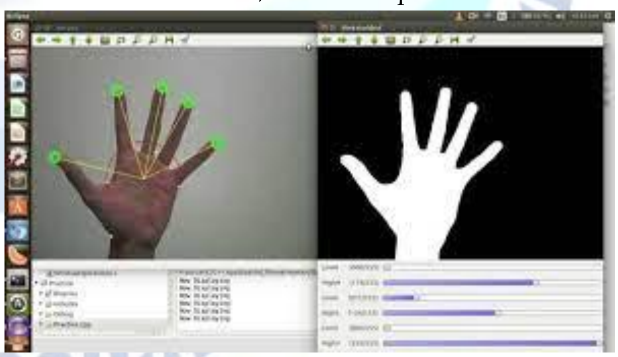

Figure5: Finger Tip Recognition Model

C. Fingertip Recognition Model Training: Divided into train and dev sets after completion of the dataset preparation and labeling process (85 percent -15 percent). Our dataset was trained using the Single Shot Detector (SSD) and the Faster RCNN pre-trained models. In comparison to SSD, the accuracy of the faster RCNN model was much higher. Please go to the Results Section for further details. Object detection in SSDs consists of two basic modules: one that suggests areas and another that classifies them. Detecting items in a single shot reduces processing time. It is often used for real-time object detections. To calculate region suggestions, Faster RCNN utilizes the output feature map from Fast RCNN. They are reviewed by a Region Proposal Network and forwarded to a Region of Interest pooling layer. The categorization and bounding box regression results are eventually shown in two completely linked layers [15]. We modified the final completely linked layer of Faster RCNN to detect the fingertip in the picture.

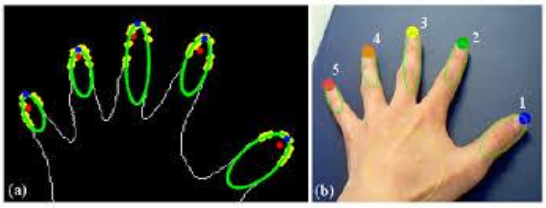

Figure6: Finger Tip Recognition Model

4. ALGORITHM

Step1: Start reading the frames and convert the captured frames to HSV colour space.(Easy for colour detection)

Step2: Prepare the canvas frame and put the respective ink buttons on it.

Step3: Adjust the track bar values for finding the mask of coloured marker.

Step4:Preprocess the mask with morphological operations.(Erotion and dilation)

Step5: Detect the contours, find the center coordinates of largest contour and keep storing them in the array for successive frames .(Arrays for drawing points on canvas)

Step6: Finally draw the points stored in array on the frames and canvas.

5. RESULTS

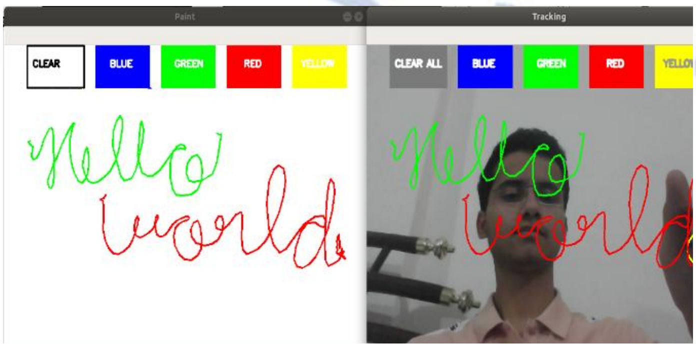

Figure-7(a)

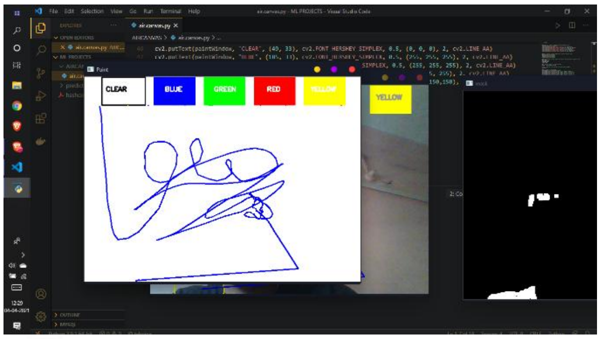

Figure-7(b)

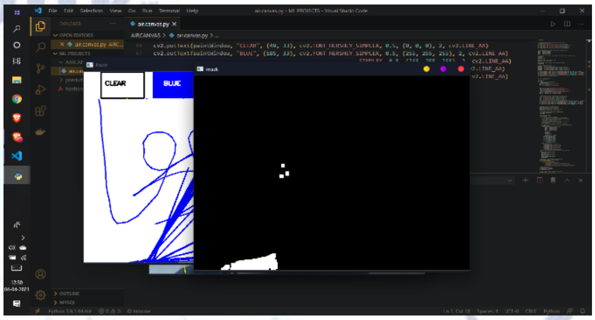

Figure-7(c)

Figure7(a)(b)(c): Final Output Images

6. CONCLUSION

An airwriting approach employing a laptop camera and a video-based pointing mechanism is shown here.. To begin with, the proposed technique tracks the colour of the fingertip in video frames and then applies OCR to the plotted pictures in order to identify the written letters. It also allows a natural human-system interface that does not need a keypad, pen, or glove for character input. It's all you need is a phone camera and a red hue to reorganize a finger tip. OpenCv and python were used to create an application for the studies. Using the suggested technique, an average accuracy of 92.083 percent may be achieved in the recognition of correct alphabets. The suggested solution resulted in a 50 ms per character delay in total writing time. Furthermore, the suggested approach may be used to all unconnected languages, but it has one important disadvantage that it is colour sensitive in such a manner that the presence of any red colour in the backdrop before commencing the analysis might lead to erroneous findings.

Future Scope

Computer Vision is the science of helping computers perceive and interpret digital pictures such as photos and movies. It's been a decades-long topic of intense investigation. Computer vision is getting better than the human visual cognitive system at spotting patterns from pictures. Computer vision-based technologies have surpassed human doctors' pattern recognition skills in the healthcare industry.

Let us examine the status of computer vision technology now and in the future.There are several aspects to consider when computer vision expands its effect on the human world. With further study and fine-tuning, computer vision will be able to do more in the future. The system will be simpler to train and can identify more from photos than it does presently.Computer vision will be used in conjunction with other technologies or subsets of AI to generate more attractive applications. Image captioning apps, for example, may use natural language generation (NLG) to understand things in the environment for visually impaired persons. Computer vision may help create artificial general intelligence (AGI) and artificial superintelligence (ASI) by processing information better than the human visual system. Computer vision is a growing sector linked to virtual and augmented reality (VR and AR). Recent market participants have shown a great interest in VR/AR fusion. This significant growth in attention is mirrored in the release of several cutting-edge technology items.

Conflict of interest statement

Authors declare that they do not have any conflict of interest.

References

[1] Y. Huang, X. Liu, X. Zhang, and L. Jin, "A Pointing Gesture Based Egocentric Interaction System: Dataset, Approach, and Application," 2016 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), Las Vegas, NV, pp. 370-377, 2016.

[2] P. Ramasamy, G. Prabhu, and R. Srinivasan, "An economical air writing system is converting finger movements to text using a web camera," 2016 International Conference on Recent Trends in Information Technology (ICRTIT), Chennai, pp. 1-6, 2016.

[3] Saira Beg, M. Fahad Khan and Faisal Baig, "Text Writing in Air," Journal of Information Display Volume 14, Issue 4, 2013

[4] Alper Yilmaz, Omar Javed, Mubarak Shah, "Object Tracking: A Survey", ACM Computer Survey. Vol. 38, Issue. 4, Article 13, Pp. 1-45, 2006

[5] Yuan-Hsiang Chang, Chen-Ming Chang, "Automatic Hand-Pose Trajectory Tracking System Using Video Sequences", INTECH, pp. 132- 152, Croatia, 2010

[6] Erik B. Sudderth, Michael I. Mandel, William T. Freeman, Alan S. Willsky, "Visual Hand Tracking Using Nonparametric Belief Propagation", MIT Laboratory For Information & Decision Systems Technical Report P2603, Presented at IEEE CVPR Workshop On Generative Model-Based Vision, Pp. 1-9, 2004

[7] T. Grossman, R. Balakrishnan, G. Kurtenbach, G. Fitzmaurice, A. Khan, and B. Buxton, "Creating Principal 3D Curves with Digital Tape Drawing," Proc. Conf. Human Factors Computing Systems (CHI' 02), pp. 121- 128, 2002.

[8] T. A. C. Bragatto, G. I. S. Ruas, M. V. Lamar, "Real-time Video-Based Finger Spelling Recognition System Using Low Computational Complexity Artificial Neural Networks", IEEE ITS, pp. 393-397, 2006

[9] Yusuke Araga, Makoto Shirabayashi, Keishi Kaida, Hiroomi Hikawa, "Real Time Gesture Recognition System Using Posture Classifier and Jordan Recurrent Neural Network", IEEE World Congress on Computational Intelligence, Brisbane, Australia, 2012

[10] Ruiduo Yang, Sudeep Sarkar, "Coupled grouping and tching for sign and gesture recognition", Computer Vision and Image Understanding, Elsevier, 2008

[11] R. Wang, S. Paris, and J. Popovic, "6D hands: markerless and-tracking for computer-aided design," in Proc. 24th Ann. ACM Symp. User Interface Softw. Technol., 2011, pp. 549–558.

[12] Maryam Khosravi Nahouji, "2D Finger Motion Tracking, Implementation For Android Based Smart phones", Master's Thesis, CHALMERS Applied Information Technology,2012,pp 1-48

[13] EshedOhn-Bar, Mohan ManubhaiTrivedi, "Hand Gesture Recognition In Real-Time For Automotive Interfaces," IEEE Transactions on Intelligent Transportation Systems, VOL. 15, NO. 6, December 2014, pp 2368-2377

[14] P. Ramasamy, G. Prabhu, and R. Srinivasan, "An economical air writing system is converting finger movements to text using a web camera," 2016 International Conference on Recent Trends in Information Technology (ICRTIT), Chennai, 2016, pp. 1- 6.

[15] Kenji Oka, Yoichi Sato, and Hideki Koike, "Real-Time Fingertip Tracking and Gesture Recognition," IEEE Computer Graphics and Applications, 2002, pp.64-71.

[16] H.M. Cooper, "Sign Language Recognition: Generalising to More Complex Corpora", Ph.D. Thesis, Centre for Vision, Speech and Signal Processing Faculty of Engineering and Physical Sciences, University of Surrey, UK, 2012

[17] Y. Huang, X. Liu, X. Zhang, and L. Jin, "A Pointing Gesture Based Egocentric Interaction System: Dataset, Approach, and Application," 2016 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), Las Vegas, NV, pp. 370-377, 2016

[18] Vladimir I. Pavlovic, Rajeev Sharma, and Thomas S. Huang, "Visual Interpretation of Hand Gestures for Human-Computer Interaction: A Review," IEEE Transactions on Pattern Analysis and Machine Intelligence, VOL. 19, NO. 7, JULY 1997, pp.677-695

[19] Guo-Zhen Wang, Yi-Pai Huang, Tian-Sheeran Chang, and Tsu-Han Chen, "Bare Finger 3D Air-Touch System Using an Embedded Optical Sensor Array for Mobile Displays", Journal Of Display Technology, VOL. 10, NO. 1, JANUARY 2014, pp.13-18

[20] Napa Sae-Bae, Kowsar Ahmed, Katherine Isbister, NasirMemon, "Biometric-rich gestures: a novel approach to authentication on multi-touch devices," Proc. SIGCHI Conference on Human Factors in Computing System,2005, pp.977-986

[21] W. Makela, "Working 3D Meshes and Particles with Finger Tips, towards an Immersive Artists' Interface," Proc. IEEE Virtual Reality Workshop, pp. 77-80, 2005.

[22] A.D. Gregory, S.A. Ehmann, and M.C. Lin, "inTouch: Interactive Multiresolution Modeling and 3D Painting with a Haptic Interface," Proc. IEEE Virtual Reality (VR' 02), pp. 45-52, 2000.

[23] W. C. Westerman, H. Lamiraux, and M. E. Dreisbach, “Swipe gestures for touch screen keyboards,” Nov. 15 2011, US Patent 8,059,101

[24] S. Vikram, L. Li, and S. Russell, "Handwriting and gestures in the air, recognizing on the fly," in Proceedings of the CHI, vol. 13, 2013, pp. 1179–1184.

[25] X. Liu, Y. Huang, X. Zhang, and L. Jin. "Fingertip in the eye: A cascaded CNN pipeline for the real-time fingertip detection in egocentric videos," CoRR, abs/1511.02282, 2015.