IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

Hand Tracking System

[1]Gaurav Saini, [2]Gaurav Mandhyan, [3]Er Akansha Singh

[1]Student, [2]Student, [3]Assistant Professor,

[12]Bachelor of Technology (Computer Science and Engineering), [2]Bachelor of Technology (Computer Science and Engineering)

[1,2]Shri Ramswaroop Memorial College of Engineering and Management, Faizabad Road, Lucknow, India

Abstract: The project introduces an application using computer vision for Hand gesture recognition. A camera records a live 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) at least once. After that a test gesture is given to it and the system tries to recognize it. Research was carried out on a number of algorithms that could best differentiate a hand gesture. It was found that the diagonal sum algorithm gave the highest accuracy rate. In the preprocessing phase, a self-developed algorithm removes the background of each training gesture. After that the image is converted into a binary image and the sums of all diagonal elements of the picture are taken. This sum helps us in differentiating and classifying differenthand gestures.

Index Terms **– Machine Learning, Augmented Reality, Neural Networks.**

I. INTORDUCTION

Vision Based scene perception (VST) is one of many fast-emerging areas in the intelligent transportation system. This field of research has been actively studied over the past decade. VSP involves three phases: detection, recognition and tracking of various objects of interest. Since recognition and tracking often rely on the results from detection, the ability to detect objects of interest effectively plays a crucial role in scene perception. In this thesis, we focus on three important classes of objects: traffic signs, cars, and cyclists. a typical on-road traffic scene with the detected objects of interest and illustrates some positive examples from the three mentioned classes.[1]

Most previous methods have designed specific detectors using different features for each of these three classes. The approach we claim here differs from these existing approaches in that we propose a single learning-based detection framework to detect all the three important classes of objects. The proposed framework consists of a dense feature extractor and detectors of these three classes. Once the dense features have been extracted, these features are shared with all detectors.[1] The advantage of using one common framework is that the detection speed is much faster, since all dense features need only to be evaluated once in the testing phase.

II. DIGITAL IMAGE PROCESSING

Image processing is reckoned as one of the most rapidly involving fields of the software industry with growing applications in all areas of work.[2] It holds the possibility of developing the ultimate machines in future, which would be able to perform the visual function of living beings. As such, it forms the basis of all kinds of visual automation.

Biometric systems are systems that recognize or verify human beings. Some of the most important biometric features are based physical features like hand, finger, face and eye. For instance, finger print recognition utilizes of ridges and furrows on skin surface of the palm and fingertips.[2] Hand gesture detection is related to the location of the presence of a hand in still image or in sequence of images i.e., moving images. Other biometric features are determined by human behavior like voice, signature and walk. The way humans generate sound for mouth, nasal cavities and lips is used for voice recognition. Signature recognition looks at the pattern, speed of the pen when writing one's signature.

III. HAND GESTURE DETECTION AND RECOGNITION

Hand detection is related to the location of the presence of a hand in a still image or sequence of images i.e. moving images. In case of moving sequences it can be followed by tracking of the hand in the scene but this is more relevant to the applications such as sign language. The underlying concept of hand detection is that human eyes can detect objects which machines cannot with that much accuracy as that of a human. From a machine point of view it is just like a man fumble around with his senses to find an object.[4]

- **Variations in image plane and pose**
  The hands in the image vary due to rotation, translation and scaling of the camera pose or the hand itself. The rotation can be both in and out of the plane.

- **Skin Color and Other Structure Components**
  The appearance of a hand is largely affected by skin color, size and also the presence or absence of additional features like hairs on the hand further adds to this variability.[4]

- **Lighting Condition and Background**
  As shown in Figure 1 light source properties affect the appearance of the hand. Also the background, which defines the profile of the hand, is important and cannot be ignored.

Figure 1: Light Source on hand

IV. RECOGNITION

Hand detection and recognition have been significant subjects in the field of computer vision and image processing during the past 30 years. There have been considerable achievements in these fields and numerous approaches have been proposed. However, the typical procedure of a fully automated hand gesture recognition system. Biometric technologies make use of various physical and behavioral characteristics of human such as fingerprints, expression, face, hand gestures and movement. These features are then processed using sophisticated machines for detection and recognition and hence used for security purposes. Unlike common security measures such as passwords, security cards that can easily be lost, copied or stolen; these biometric features are unique to individuals and there is little possibility that these pictures can be replaced or altered. In video conferencing system, there is a need to automatically control the camera in such a way that the current speaker always has the focus. One simple approach to this is to guide the camera based on sound or simple cues such as motion and skin color.[5]

Hand gestures are important to intelligent human and computer interaction to build fully automated systems that analyze information contained in images, fast and efficient hand gesture recognition algorithms are required.

Among the biometric sector hand gesture recognition are gaining more and more attention because of their demand regarding security for law enforcement agency as well as in private sectors such as surveillance systems.

V. SCOPE

I. METHODOLOGY

The scope of this project is to build a real time gesture classification system that can automatically detect gestures in natural lighting condition. In order to accomplish this objective, a real time gesture-based system is developed to identify gestures. This system will work as one of futuristic of Artificial Intelligence and computer vision with user interface. It creates method to recognize hand gesture based on different parameters.[6] The main priority of this system is to simple, easy and user friendly without making any special hardware. All computation will occur on single PC or workstation. Only special hardware will use to digitize the image (Digital Camera).

II. GENERIC OBJECT DETECTION

Object detection is a challenging but important application in the computer vision community. It has achieved successful outcomes in many practical applications such as face detection and pedestrian detection. Complete survey of object detection can be found in. This section briefly reviews several generic object detection methods. These frameworks achieve excellent detection results on rigid object classes. However, for object classes with a large intra-class variation, their detection performance falls down dramatically.[6] Recently, a new detection framework which uses aggregated channel features (ACF) and an AdaBoost classifier has been proposed in.[6] This framework uses exhaustive sliding-window search to detect objects at multi-scales. It has been adapted successfully for many practical applications.

III. WORK

We propose a single learning-based detection framework to detect all the three important classes of objects. The proposed framework consists of a dense feature extractor and detectors of these three classes. Once the dense features have been extracted, these features are shared with all detectors. The advantage of using one common framework is that the detection speed is much faster, since all dense features need only to be evaluated once in the testing phase. The proposed framework introduces spatially pooled features as a part of aggregated channel features to enhance the feature robustness to noises and image deformations. In order to further improve the generalization performance, we propose an object sub categorization method as a means of capturing the intra-class variation of objects.[6]

V. PROPOSED WORK

- **Hand Detection**

```
import numpy as np
import cv2
import argparse
import color_calculator as cc
import color_detection as cd
import video_detection as vd


def analyse_args():
    """Parses the args"""
    parser = argparse.ArgumentParser()
    parser.add_argument(
        '-i', '--input', default=None, help='Place ROI on the
left')
    parser.add_argument(
        '-l', '--left', action='store_true', help='Place ROI
on the left')
    parser.add_argument(
        '-s',
        '--shot',
        action='store_true',
        help='Not video, just a single shot')
    return parser.parse_args()


def main():
    """Main function of the app"""
    args = analyse_args()
    video_capture = cv2.VideoCapture(0)
    lower_color = np.array([0, 50, 120], dtype=np.uint8)
    upper_color = np.array([180, 150, 250], dtype=np.uint8)

    while True:
        _, frame = video_capture.read()
        frame = cv2.flip(frame, 1)

        cv2.putText(frame, 'Welcome', (0, 50),
cv2.FONT_HERSHEY_SIMPLEX, 2,
                    (255, 0, 0), 3, cv2.LINE_AA)

        cv2.imshow('VCOM Project', frame)
        key = cv2.waitKey(10)
        if key != -1:
```

```
                cv2.destroyAllWindows()
                video_capture.release()
                break

        if key == ord('v'):
            try:
                avg_color, max_sensibility =
    cc.captureCamera(args.left)
                vd.start(
                    avg_color,
                    max_sensibility,
                    video=not args.shot,
                    path=args.input,
                    left=args.left)
            except TypeError:
                print('Did not calculate the color bound.')
        elif key == ord('h'):
            cd.draw_contours(lower_color, upper_color)


    if __name__ == '__main__':
        main()
```

VI. RESULTS

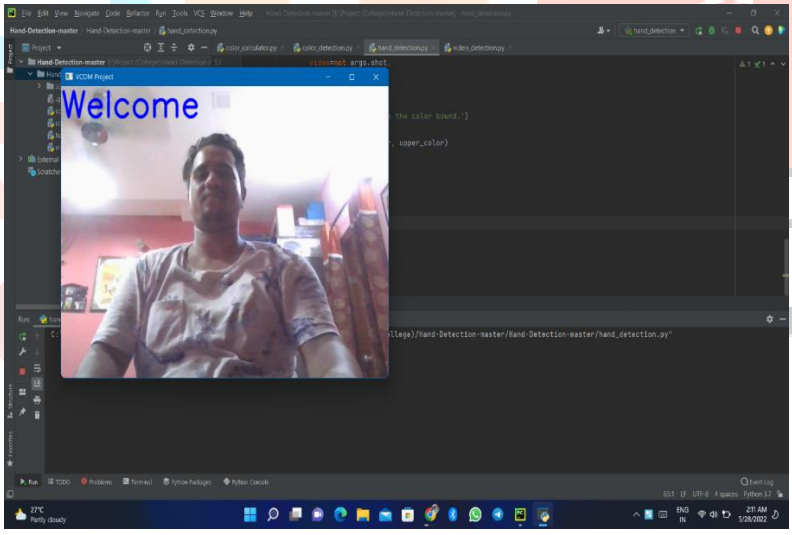

Figure 2: Project Screenshot 1

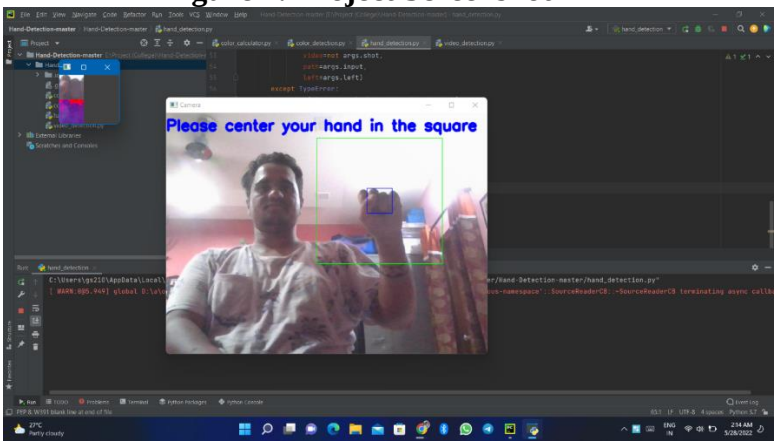

Figure 3: Project Screenshot 2

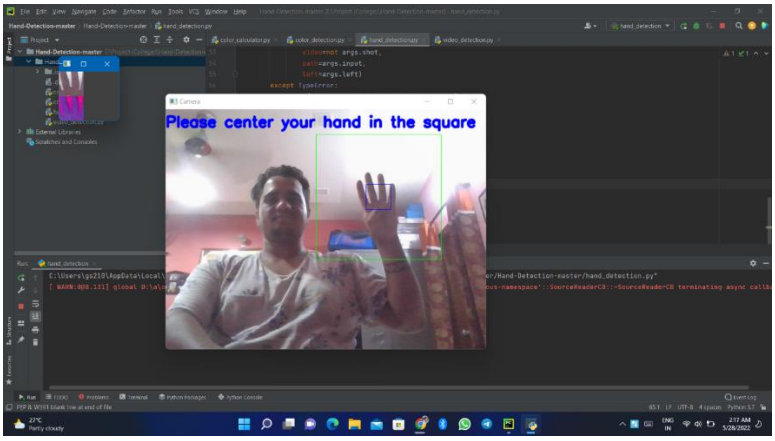

Figure 4: Project Screenshot 2

VII. CONCLUSION

Hand gesture recognition addresses a fault in interaction systems. Controlling things by hand is more natural, easier, more flexible and cheaper, and there is no need to fix problems caused by hardware devices, since none is required. From previous sections, it was clear to need to put much effort into developing reliable and robust algorithms with the help of using a camera sensor has a certain characteristic to encounter common issues and achieve a reliable result. Each technique mentioned above, however, has its advantages and disadvantages and may perform well in some challenges while being inferior in others.

VIII. ACKNOWLEDGEMENT

I would like to thank our guide Er. Akansha Singh for helping me throughout this research and for the continuous guidance during the development of this project.

REFERENCES

[1] S. Agarwal, A. Awan, and D. Roth, "Learning to detect objects in images via a sparse, part-based representation," IEEE Trans. Pattern Anal. Mach. Intell., vol. 26, no. 11, pp. 1475–1490, Nov. 2004.

[2] T. Ahonen, A. Hadid, and M. Pietikäinen, "Face recognition with local binary patterns," in Proc. Eur. Conf. Comput. Vis., 2004 pp. 469–481..

[3] J. Behley, V. Steinhage, and A. B. Cremers, "Laser-based segment classification using a mixture of bag-of-words," in Proc. IEEE Int. Conf. Intell. Robots Syst., 2013, pp. 4195–4200.

[4] A. Broggi, A. Cappalunga, S. Cattani, and P. Zani, "Lateral vehicles detection using monocular high resolution cameras on terramax," in Proc. IEEE Intell. Veh. Symp., 2008, pp. 1143–1148.

[5] https://lec.inf.ethz.ch/ifmp/2019/dl/additional/handouts/ ProgramTracing.html 4 November 12:00 PM

[6] Virtual & Augmented Reality For Dummies Book by Paul Mealy.

[7] Augmented Reality: Where We Will All Live Book by Jon Peddie

[8] Augmented Reality for Developers: Build Practical Augmented Reality Applications with Unity, ARCore, ARKit, and Vuforia Book by Jonathan Linowes and Krystian Babilinski.