International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 08 Issue: 04 | Apr 2021 www.irjet.net p-ISSN: 2395-0072

Real – Time Human Detection and Tracking in Motion Environment

Ajinkya Bam[1], Prakash Choudhary[2], Jatin Bhoir[3]

[1,-3]Dept. of Computer Engineering, M.G.M's College of Engineering and Technology, Kamothe, Navi Mumbai, Maharashtra, India

---

Abstract - *Detecting humans in films and videos is a challenging problem owing to the motion of the subjects, the camera and the background and to variations in pose, appearance, clothing, illumination and background clutter. We develop a detector for standing and moving people in videos, testing several different motions coding schemes and showing empirically that orientated histograms give the best overall performance. Use of human modelling to recognize and monitor human activity in the scene such as human walking, running etc is tracked. In addition to videos, detection from a static image can also be carried out by providing image as an input instead of a live feed from a CCTV footage respectively. Human detection in videos (i.e., series of images) plays an important role in various real-life applications (e.g., visual surveillance and automated driver assistance). The task of human detection in a series of images is challenging due to various reasons. One of these reasons is the variation of human size in the video frame. This results from changing the altitude of the platform that the camera is attached to during the task. Accuracy and short training time are the two important factors that should be taken into consideration to get a robust human, nonhuman classification system. The current Covid-19 Pandemic has altogether increased the need of such sustainable system that is Real time human detection to avoid any mishap and limit the spread of the virus, with the help of detected persons required actions can be taken by the concerned authorities respectively.*

Key Words: *Human detection, video streams, HOG, SVM.*

1. INTRODUCTION

Finding people in images has attracted much attention in recent years for practical applications such as visual surveillance. The detection of a human being is important for abnormal event detection, human gait characterization, people counting, person identification and tracking, pedestrian detection, gender classification.

1.1 MOTIVATION:

Object detection and tracking is of utmost importance for different kinds of applications such as safety, surveillance, man-machine interaction, driving assistance system, traffic monitoring. Due to the ongoing worldwide Covid-19 Pandemic, it is the upmost need of an hour to implement sustainable systems to provide smooth functioning.

1.2 Problem Statement

Automatic human detection and tracking is an important feature of video surveillance systems. It can improve a system's performance in fields such as security, safety, human activity monitoring etc. Human detection systems can have different goals such as detecting the presence of humans, recognition of abnormal behavior (falls, climbing, running, etc.), identification of specific individuals, etc. The concept of human detection and tracking is to give a computer system the ability of finding humans precisely in images or videos. Numerous algorithms and techniques have been developed for improving the performance of human detection. Recently Deep learning has been highly explored for computer vision applications. Human brain can automatically and instantly detect and recognize humans. But when it comes to computer, it is very difficult to do all the challenging tasks on the level of human brain. The main focus in this automated detection is that human activities are recognized correctly and based on human body part motion, the human activity analysis can happen. If abnormal behavior is detected, an alarm can be triggered. To identify individuals, face detection and recognition can also be applied.

1.3 Scope

Detecting human beings accurately in a visual surveillance system is crucial for diverse application areas including abnormal event detection, human gait characterization, congestion analysis, person identification, gender classification and fall detection for elderly people. The biggest use of such system is during the ongoing Worldwide Covid-19 Pandemic where intelligent tracking of mass gathering is of utmost importance to avoid the community spread of the disease. As the system comes with Real-time human counting, based on the statistics the task of the governing body gets reduce significantly to identify crowdy places or streets with the help of the proposed system respectively

2. ALGORITHMS AND TECHNIQUES

2.1 Current system

The most successful and popular vector-form feature: histograms of oriented gradients (HOG) It is shown that the HOG features are based on the contrast of silhouette contours against the background. Despite all the difficulties on human detection, a lot of work has been done recent years. First, we may use different features such as edge, Haar features and gradient orientation features; second, we may use different classifiers such as Nearest Neighbor, Neutral Network, SVM and Adaboost the second step of human detection is designing classifier. Large generalization ability and less classifying complexity are two important criteria for selecting classifiers. Linear support vector machine (SVM) and AdaBoost are two widely-used classifiers satisfying the criteria. So the traditional approach of AdaBoost for face detection and has demonstrated both high recognition accuracy and fast run-time performance. However, in most cases the classification accuracy is lower than that of the first proposed algorithm based on HOG+ SVM.

2.2 Challenges/Issues with Current System

The most difficulty in building a robust system for human detection and tracking is the amount of variation in the videos and several other factors contribute to this as given below

• Device creates dependencies on viewpoint such that even a small change in the object's position or orientation with respect to the camera centre may change its appearance considerably. An object detector system must handle the issues of viewpoint and scale changes.
• Most natural object classes have large within-class variations. For example, for humans both appearance and pose change considerably between images and differences in clothing create further changes
• Background clutter is common and varies from image to image. Examples are images taken in natural settings, outdoor scenes in cities and indoor environments. The system must be capable of distinguishing object class from complex background regions.
• Object color and general illumination varies considerably, for example direct sunlight and shadows during the day to artificial or dim lighting at night.
• Partial occlusions create further difficulties because only part of the object is visible for processing

2.3 PROPOSED SYSTEM:

To overcome the problems in the existing system, we shall develop an Automated detection and tracking system over manual system. Initially system consists of reading an image and detecting the existing people in this image, using HOG. Secondly the program must read a video streams recorded in the hard drive or collected by the webcam then tracking and detection is done. Initially the program carried out consists in reading an image and detecting the existing people in this image, using HOG, HOG optimized descriptor or Haar features. Secondly the program must read a video streams recorded in the hard drive or collected by the webcam or by an external camera; then make the same treatment for each video frame. We must test if the passed arguments are a static image or a recorded video stream. If there are no arguments the function directly will seek the existence of a webcam or a camera installed on the computer.

HOG for person detection in video: In the case of a video stream, we treat the flow components as independent images, we take their local gradients separately and we compute HOGs as in static image

In the case of a video stream, we treat the flow components as independent images, we take their local gradients separately and we compute HOGs as in static image

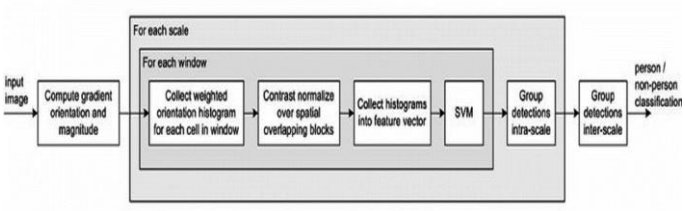

Fig 1: HOG for person detection in video

2.4 Algorithm

The currently available Human Recognizer Algorithms in OpenCV are:
• SIFT: Scale Invariant Feature Transform
• SURF: Speeded-Up Robust Feature
• HOG: Histogram of Oriented Gradient
For our purpose, we would be using the last algorithm (HOG)

HOG (Histogram of oriented gradients):
HOG is a feature descriptor used in computer vision and image processing for the purpose of object detection. This is one of the most popular techniques for object detection, to our fortune, OpenCV has already been implemented in an efficient way to combine the HOG Descriptor Algorithm with Support Vector Machine or SVM. The main idea behind the histogram of oriented gradient is that the local appearance and shape of object in an image can be described by the intensity distribution of gradients or direction of the contours. The implementation of these descriptors can be obtained by dividing the image into small connected regions, called cells. Then, for each cell we

compute a histogram of gradient directions or edge orientations for all pixels of the cell. The combination of these histograms is the descriptor. The HOG descriptor has some key advantages. Since it operates on localized cells, the method maintains the invariance to geometric and photometric transformations. We implemented the HOG descriptor on python and using OpenCV libraries. The HOG descriptor focuses on the structure or the shape of an object. Now you might ask, how is this different from the edge features we extract for images? In the case of edge features, we only identify if the pixel is an edge or not. HOG is able to provide the edge direction as well. This is done by extracting the gradient and orientation (or you can say magnitude and direction) of the edges. Additionally, these orientations are calculated in 'localized' portions. This means that the complete image is broken down into smaller regions and for each region, the gradients and orientation are calculated. We will discuss this in much more detail in the upcoming sections Finally, the HOG would generate a Histogram for each of these regions separately. The histograms are created using the gradients and orientations of the pixel values, hence the name 'Histogram of Oriented Gradients'.

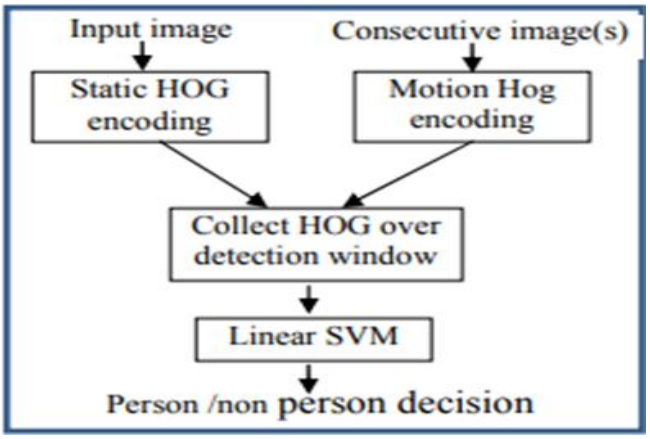

Fig 2: Flow of Motion Hog

The HOG descriptor technique counts occurrences of gradient orientation in localized portions of an image - detection window, or region of interest (ROI). Implementation of the HOG descriptor algorithm is as follows: 1. Divide the image into small connected regions called cells, and for each cell compute a histogram of gradient directions or edge orientations for the pixels within the cell. 2. Discretize each cell into angular bins according to the gradient orientation. 3. Each cell's pixel contributes weighted gradient to its corresponding angular bin. 4. Groups of adjacent cells are considered as spatial regions called blocks. The grouping of cells into a block is the basis for grouping and normalization of histograms. 5. Normalized group of histograms represents the block histogram. The set of these block histograms represents the descriptor.

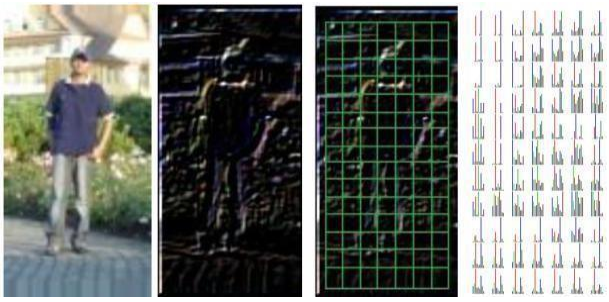

Fig 3: Algorithm Implementation

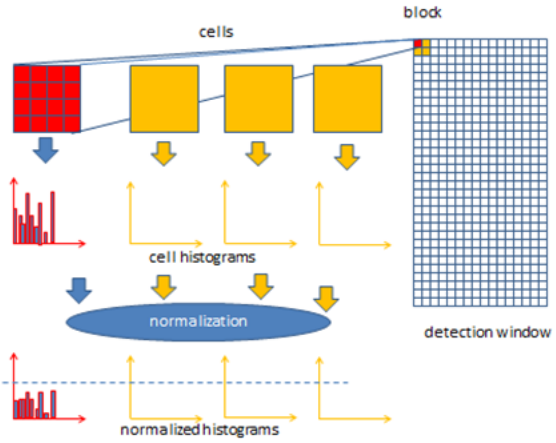

Fig 4: Histograms of Oriented Gradients principle.

Computation of the HOG descriptor requires the following basic configuration parameters:
- Masks to compute derivatives and gradients
- Geometry of splitting an image into cells and grouping cells into a block
- Block overlapping
- Normalization parameters the recommended values for the HOG parameters are:
- 1D cantered derivative mask [-1, 0, +1]
- Detection window size is 64x128
- Cell size is 8x8
- Block size is 16x16 (2x2 cells)

3. CONCLUSIONS

3.1 Limitations

DEFORMAL PART MODEL:

Deformable Part is a discriminatively trained, multi-scale model for image training that aim at making possible the effective use of more latent information such as hierarchical (grammar) models and models involving latent three-dimensional pose.

3.2 Applications

In security and surveillance, human presence detection are often used to detect intruders, whether it's in your personal home, an office or retail setting, or even an industrial environment. With computers, human presence detection is also used for security purposes. In this pandemic situation where number of peoples are limited to certain count by government to avoid spreading of covid counting the detected human can help to regulate the laws. The scenes obtained from a surveillance video are usually with low resolution. Most of the scenes captured by a static camera are with minimal change of background. Objects in the outdoor surveillance are often detected in far field. Most existing digital video surveillance systems rely on human observers for detecting specific activities in a real-time video scene. However, there are limitations in the human capability to monitor simultaneous events in surveillance displays

REFERENCES

[1] Paresh M Tank, Hitesh A Patel. Survey on Human Detection Techniquesin Real Time Video. IJCSI International Journal of Computer Science Issues Vol. 7, Issue 5, May 2018.

[2] N. Dalal and B. Triggs, "Human detection using oriented histograms of flow and appearance," in European Conference on Computer Vision (ECCV), 2006, pp. 428–441 .

[3] Mohamed E Hussein, W. Abd-Almageed, Yang Ran, Larry S. Davis. Real-Time Human Detection, Tracking, and Verification in Uncontrolled Camera Motion Environments, 10.1109/ICVS.2006.52 Computer Vision Systems, 2006 ICVS '06. IEEE International Conference

[4] Manoranjan Paul, Shah M E Haque, and Subrata Chakraborty 2013 Human detection in surveillance videos and its applications, EURASIP Journal on Advances in Signal Processing 2013(1):25 .DOI: 10.1186/1687-6180-2013-176