Generalized Omnipresence Detection

Haretha Winmalar D, Vani A K, Sudharsan R, Hari Krishna R
School of Electrical and Electronics Engineering,
SASTRA University.

Abstract—Identification and Tracking of a person in a video are useful in applications such as video surveillance. Two levels of tracking are carried out. They are Classification and monitoring of individuals. The human body's color histogram is used as the basis for monitoring individuals. Our project can detect a human face in a video and store the detected facial features of the Local Binary Pattern Histogram (LBPH). In a video, once a person is detected, it automatically track that individual and assigns a label to that individual. We use the stored LBPH features to track him in any other videos. In this paper, we proposed and compared the efficiency of two algorithms. One constantly updates the background to make it suitable for illumination changes and other uses depth information with RGB. This is the first step in many complex algorithms in computer vision, such as identification of human activity and behavior recognition. The main challenges in human/object detection and tracking are changing illumination and background. Our work is based on image processing and also it learns the activities and stores them using machine learning with the help of OpenCV, an open source computer vision library.

Index Terms—LBPH, OpenCV, RGB

I. INTRODUCTION

Monitoring the activities and collecting information for managing or security purpose is called surveillance. This includes observing from a distance by closed-circuit television (CCTV). This uses a set of algorithms to track an individual. Surveillance of the public using CCTV is common in many areas around the world. CCTV has many applications in various industries. It is an excellent solution for video surveillance that not only improves security and enables tracking, but is also very useful in general observation. CCTV Surveillance Cameras are designed for high-quality images and can function just as well in indoor and outdoor configurations. Through merely looking into recorded videos, it is easy to look for people passing within the range of security cameras. With the increasing threats of terrorism and thefts, the advanced video surveillance has been used. Although the security camera reduces human effort, it takes some time to search for a person through the entire video. Conjunction of machine learning and image processing is used to teach a computer how to identify an individual and continuously track that individual in the video. After scanning through different videos, our program gives output like- person marked as subject1 is viewed in camera1 video, camera2 views subject1 in the video. By recognizing him / her in the video taken by multiple cameras, we have tracked an individual. Overall, our project is about tracking people in videos.

In the first step, video is the input. Our project's objective is to identify and recognize individuals and monitor them in a given series of videos. It can not, in one instance identify a number of people at a time and automatically find a specific person. Users are engaged in heavy workloads in a traditional video surveillance network. Since the user needs to identify the particular person who moves from

one video camera to another, there is a limited number of offenders to find. Identifying a person from a number of monitoring positions also increases the user's workload. In such situations, the user is required to carefully identify the person.

Literature Survey

Better Histograms of Oriented Gradients characteristics are used to represent image edge details in [1]. The Author used background subtraction detection with Histograms of Oriented Gradients to monitor in real-time, achieving the necessary accuracy and meeting real-time demand. The Author has also used the HOG feature which reflects the images' edge features. It reduces the impact of illumination. The improved approach significantly reduces energy consumption and increases the speed of detection. Navneet Dalal suggested that using locally defined gradient orientation characteristics [2] in a complex overlapping grid helps in detecting an individual. And also, it reduces false-positive levels than the Haar wavelet-based detector. After the face detection phase, comes the phase of face recognition. But for this phase there is a prerequisite, to recognize a person firstly, we have to train our machine for certain images of that person. Here training is done using Haar cascades and Local Binary Pattern Histograms (LBPH)[4]. A state estimation strategy is planned to grow Dlib connection tracker abilities alongside the utilization of the human location in [3] To resolve the issue of neglecting to follow an object constantly if that object vanishes because of impediments and all of a sudden return. The methodology recommended demonstrated preferable outcomes over utilizing Dlib correlation tracker legitimately with human recognition and was assessed in real-life situations. The methodology can be additionally improved for different cases and different assortment of human trackers.

II. PROJECT METHODOLOGY

In a smart surveillance system, human detection helps to differentiate between moving objects in a video sequence. In this section, before presenting the output to the user, we will address the various stages of the device. This gives an outline of how the framework really functions and how it generates the output.

A. Overview of the method

Real time Human detection in surveillance videos and tracking consists of the following stages: the human detection stage, the face detection stage and the face recognition stage. The detection process usually takes place in two steps: the identification of objects and classification of objects. When the object is detected, it could be distinguished as a human utilizing the shape based or texture-based features.

Points, boxes and blobs are the information extracted for shape based approach. Eishita et al [5] Presented a mis-leadingly basic strategy for following articles after aggregate or halfway impediment utilizing shape, colour and texture, regardless of whether the items have similar hues and textures. It is then normally viewed as an issue of matching templates. Various applications, such as smart rooms and visual surveillance, are likewise accessible for detecting human from static images. Detecting human from static images in a scene which is crowded is defined in [6]. The technique models an individual human as an assembly of body parts using the body features, which are nothing but silhouette-oriented features.

Regardless, the verbalization of the human and the distinctions in the discernment's experienced lead to a wide scope of conceivable body characteristics, thus it's hard to recognize an individual moving from other moving objects by utilizing a shape-based methodology. This issue is solved by applying template matching. Edge based high-dimensional histograms are utilized by surface-based methodologies, for example, the histograms of oriented gradient (HOG). And human regions are detected using support vector machine (SVM).

B. Approach

A sequence of videos is the input in the program. In the first stage, it detects human in the video. It first puts up a rectangle around them when a human is detected and tracks the detected individual in the video. Our program then tries to detect the individual face in the given video in the second phase. Face detection is autonomous of human detection in a video. Once the program recognizes the face, the next step is to identify the individual correctly. For a number of individuals, we had already trained our machine by extracting faces from some video clips. Now our system is trying to identify individuals in the video with the aid of a trained machine. Whether it is effective or not, it goes to the fourth stage, i.e. human tracking, after face recognition.
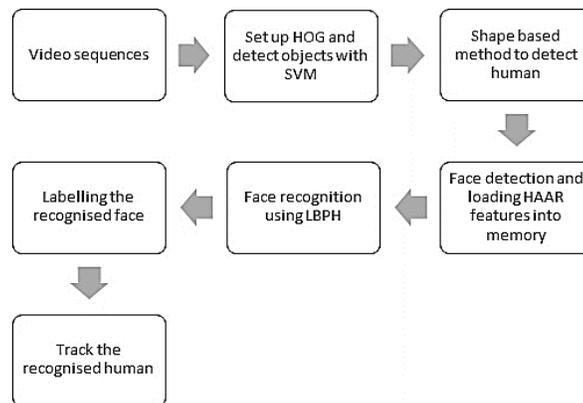

Fig. 1. Flow chart

After the input is taken, the next step is the setting up a Histogram of Oriented Gradients (HOG) which defines the features used for detecting individual with the help of a Support Vector Machine (SVM). HOG [7] implemented a texture-based algorithm that uses edge-based features and then to detect the human-body regions, we use SVM.

As it starts detecting people in the video, the program starts drawing a rectangle around the human. The second step is the face detection. For face detection, the program uploads the HAAR features into the memory, which stores all the features of faces and detects the human faces using these facial features. The third step is the face recognition phase. But there is a prerequisite for this step, first of all to recognize an

individual, we must train our machine for certain images of that person. Here training is done using Haar cascades and Local Binary Pattern Histograms (LBPH)[4].

First of all, Haar features are used to detect faces in video frames, now each face detected is treated as a set of data to train our machine and LBPH features corresponding to each face are stored in a register. It is a simple solution for the face recognition problem, which can be recognizing both the front face and side face. LBPH is an efficient text descriptor that labels each pixel's neighborhood pixels by threshold and produces the result as a binary number. Then the LBPH was combined with the histogram. To represent the face images, we can use a simple data vector. Therefore, by recognizing that this person has passed through this camera, our program detects human, detects faces, recognizes faces and monitors human. So, we got overall output like subject1 went through camera1 and he/she also went through camera2 and so on.

III. Algorithms

In this section, the algorithms used in our implementation are briefly explained.

A. Histogram Oriented Gradient (HOG)

The fundamental thought is that neighborhood object appearance and shape can regularly be portrayed or maybe well by the circulation of local intensity gradients or edge directions, even without exact information on the relating gradient or edge positions. This is achieved by segmenting the image window into tiny temporal regions (" cells") that accumulate over each cell's pixels a 1-D histogram of edge orientations or gradient directions. The representation is generated by the combined histogram inputs. The image is segmented into a small part which is called as cells. Each cell can have multiple pixels and a histogram of gradient is created for each pixel. It is very much useful to standardize the local responses before using them for improved invariance to be shadowing illumination, etc. This can be achieved by collecting a measure of local energy histograms across much larger spatial regions (" blocks") and using the effects to normalize all block cells. The uniform descriptor squares are alluded to as Histogram of Oriented Gradient (HOG) descriptors. It is especially appropriate for distinguishing human in recordings.

B. Support Vector Machine (SVM)

Support Vector Machines is generally regarded as a method of classification but can also be used in identification and regression. The program handles face in different conditions like lighting and shadows. SVM moves into higher dimensional space so that individuals can be categorized. Linear classification is a bit trickier task compared to nonlinear classification, but the resulting accuracy of linear classification is not a good large data set. Linear classification can be explained by equation 1.

i.e. $y_j f(x_j) > 0$ for a correct classification.

Given training data (x_j , y_j)f or $j = 1…N$; with $x_j \in Rd$ and
$y_j \in (1,-1)$ learn a classifier $f(x)$ such that
≥ 0

$f(x_j) =$
$y_j = +1$
< 0 $y_j = -1$

C. Haar Cascade

The study of pixels in the picture in squares by function is based on the Haar Wavelet technique. This uses machine learning methods to obtain a high level of precision from what is called" training data. This uses principles of" integral image" to measure the detected" features. "Initially, to train the classifier, it needs many positive images (with faces) and negative images (without a face). Then we must extract characteristics from it. Features are just numerical image details that can be used to distinguish between one image and another; For example, a histogram (distribution of intensity value) is one of the features that can be used to define several image features without even looking at the image, such as darker or brighter image, range of image intensity, contrast, etc. To detect faces in a picture, we will use haar features.

Let's look at a scenario where the task is to detect a face. Most part of the picture is a non-face region. The classifiers concept is implemented here. Rather than applying all the features simultaneously to a window, split the characteristics into different stages and append them one after the other. So, if a region fails during the first phase itself, it is not necessary to apply the remaining classifiers. We just apply the second stage to the regions that pass the first phase, and so on.

D. Linear Binary Pattern Histogram (LBPH) face recognizer

LBP can be used in conjunction with histograms to represent face images using a simple data vector.

- Divide the concerned window into cells
- It can also be interpreted as a 3x3 matrix. And has a pixel intensity (0 ~ 255).
- The matrix central value is taken and used as threshold
- If the middle pixel value is greater than the neighbor pixel value, write "1." If not, write "0." (Eg: 250 >70 = 1, 70 < 50 = 0 always consider the pixel value).
- Now, only binary values (ignoring the central value) will be included in the matrix. If middle pixel value is higher than the neighbor's value, write "1." If not, write "0." (Eg:250 > 70 = 1, 70 < 50 = 0 always find the pixel value). Binary value from the line-by-line matrix to the current binary value from each position (e.g. 10001101).
- We will consider the histogram computation for each number frequency over the cell and conjunction of which pixels are smaller and larger than the centre). Then the histogram is normalized.
- All cells must concatenate (normalized) histograms. It gives the window's function vector.

OpenCV library provides the LBPH. We used OpenCV to present a Haar Cascade Classifier that is used to detect faces. The classification of haar cascades uses the LBPH algorithm.

IV. **Results**

We check our method with various video sequences in different environments to show the efficiency of the proposed algorithm. In this section, when using our program against the test videos, we provided details of the results obtained by us. Using the display screenshots, we explained the output of each test case video. We have already trained the machine to look up for the person and they are labeled. We provided the output for test videos that represent different scenarios for each of the test videos. Given below are the screenshots of the stages output such as human detection, face detection and face recognition.

Fig. 2. Output of video1

The screenshots of the program output of the second video given by the program are shown in Figure 3. Initially, all of the humans are detected in the video. Then the person who is already trained in our samples enters the frame and he is labeled as 1. The program recognizes the person 1 from the training samples and successfully detects the face of him/her. From the video frame, the program successfully detects the human
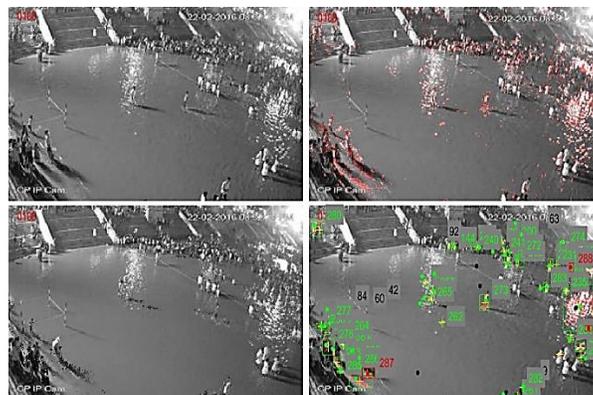

Fig. 3. Output of video2: Human detection

as the entire body is visible and draws a rectangle around him/her as seen in figure 4. Thus, the program detects and labels the face of person 1 until the time it is present in the frame.

Fig. 4. Output of video2: single person detection

 Since the first frame we can see here, both individuals have been identified correctly by the system. As can be seen in the figures given above, it has successfully detected the faces and has drawn a rectangle around both person 1 and person 2. When they step in the visibility of the camera, the rectangles are immediately drawn and they're always present and they move along with them until they're present in the frame. Till the last frame, it continues to recognize them. It is then moved towards the next camera and detects there also.

V. Conclusion

Because of its wide range of applications, one of the major topics in vision research is the precise recognition of human beings in a surveillance video. We presented a reliable human detection and tracking system that can function in different environments in real-time. The program succeeds in identifying a person's presence in a video and is also able to detect these people's faces. The program uses a pre-trained model to recognize people in a video and allows the user to monitor any person in a video series if the program succeeds in recognizing the name of the person on the terminal.

At the conclusion of this article, a discussion will be held to point out the future work necessary in surveillance videos to enhance the human detection process. We plan to develop detectors of body parts in the future, continue work on monitoring and conduct analysis of behavior. And also, in future, it has been planned to include a feature where the machine learns itself from the previously recorded videos and identifies the same person in any other future videos. This will help the user much easier in human detection and to reduce the number of false positives.

References

[1] Human detection and object tracking based on Histograms of Oriented Gradients Zhang S., Wang X. (2013) Proceedings - International Con-ference on Natural Computation, , art. no. 6818189 , pp. 1349-1353.

[2] Navneet Dalal and Bill Triggs. Histogram of oriented gradients for human detection, 2005.

[3] G. Gamage, I. Sudasingha, I. Perera and D. Meedeniya, "Reinstating Dlib Correlation Human Trackers Under Occlusions in Human Detection based Tracking," 2018 18th International Conference on Advances in ICT for Emerging Regions (ICTer), Colombo, Sri Lanka, 2018, pp. 92-98.

[4] Ahonen, T., Hadid, A., and Pietikainen, M. Face Recognition with Local Binary Patterns. Computer Vision - ECCV 2004 (2004), 469481.

[5] FZ Eishita, A Rahman, SA Azad, A Rahman, Occlusion handling in ob-ject detection. Multidisciplinary computational intelligence techniques: applications in business, engineering, and medicine. IGI Global. (2013). doi: 10.4018/978-1- 4666-1830-5.ch005

[6] B Wu, R Nevatia, Detection of multiple, partially occluded humans in a single image by Bayesian combination of edgelet part detectors, in IEEE International Conference on Computer Vision (ICCV 2005) (IEEE, Piscataway, 2005), pp. 90–97

[7] N Dalal, B Triggs, Histograms of oriented gradients for human detection, in IEEE Conference on Computer Vision and Pattern Recognition (CVPR 2005) (IEEE, Piscataway, 2005), pp. 886–893