A simplified method for human height estimation in video surveillance

Rahul Rachh (1401057), Kedar Acharya (1401081),
Janki Desai (1401103), Shloak Agarwal (1401105)

***Abstract*—This paper presents a simple method for estimating human height in video surveillance. Most cameras for video surveillance are installed in high positions at a slightly tilted angle, it is possible to retain human height using the tilt angle. Certain parameters such as reference height, detected pixel height and display resolution are taken into account. These parameters are calculated using a visual perception based formula proposed in the paper. The experimental results show that the proposed method can detect human height with the absolute error of 7 cm from ground.**

***Index Terms*—Soft bio-metrics, Human height estimation, Video surveillance, SVM, HOG Descriptor.**

I. Introduction

A very important problem in archaeology and forensic science is the problem of height estimation. Existing approaches that deal with this problem involve simple regression formulas based on statistical methods. Our goal is to investigate the potential of more complex methods, such as Support Vector Machines (SVMs) using HOG Descriptor. We believe that the good performance of HOG Descriptor on other problems can make them ideal for solving the problem of height estimation. Most of the other papers use regression which needs to train the data and takes a lot of time. Our proposed method is simple and takes only one parameter: Tilt angle.

Advances in the image resolution and quality of digital cameras in the last few years have increased the image analysis capability of modern video surveillance systems. The main advantage of this method is it provides the simplest solution for height estimation in video surveillance in comparison with other methods. The challenging task of detection is done using HOG descriptor while height is estimated using a simple formula which will be discussed later on.

II. Literature Review

HOG (Histograms of Oriented Gradients)is a type of feature descriptor to generalize the object in such a way that the same object (in this case a person) produces as close as possible to the same feature descriptor when viewed under different conditions. The entire person is represented by a single feature vector, as opposed to many feature vectors representing smaller parts of the person.

The HOG person detector uses a sliding detection window which is moved around the image. At each position of the detector window, a HOG descriptor is computed for the detection window. This descriptor is then shown to the trained SVM, which classifies it as either person or not a person.[3]

To recognize persons at different scales, the image is sub-sampled to multiple sizes. Each of these sub-sampled images is searched. Thus, image is detected.After that a function draws a rectangle around the human body. The height is then estimated using a formula discussed below.[2]

III. Discussion

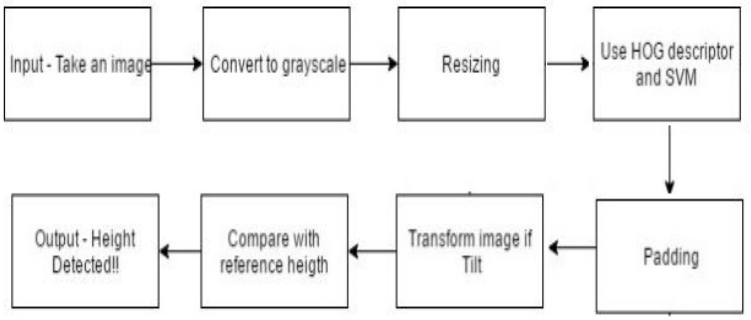

Fig. 1: Flow Chart

As shown in the Fig 1. input image or the video is converted into gray scale and then re-sized into 400 X 400 to increase computation speed .After that, HOG descriptor is used with SVM to classify and detect the human body. Image is then padded to increase the accuracy of detection of human body from head to toe. Geometric and trignometric formula is developed for tilt. It is transformed in such a way so that the tilt image is transformed to zero angle plane. Then the predicted height is calculated which is then compared with reference image to provide output that is detected height. Here height estimation was carried out for 2-4 people using video and static images. Only 1 camera was used with tilt known.

In Fig. 3, Fig. 4 and Fig. 5 different cases of camera with respect to a person are depicted.

IV. Proposed Method

This method is based on concept of visual perception. Parameters we used are explained in Fig. 2. and 3. and they are discussed below

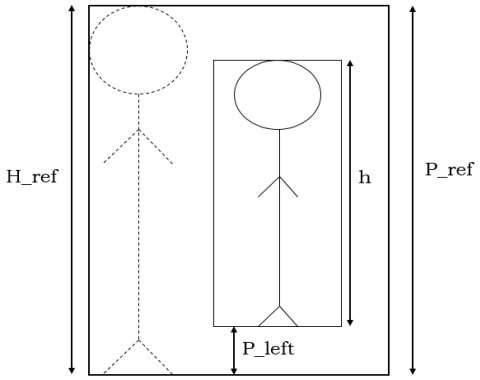

Fig. 2: Parameters

Fig. 3: Zero Tilt Angle

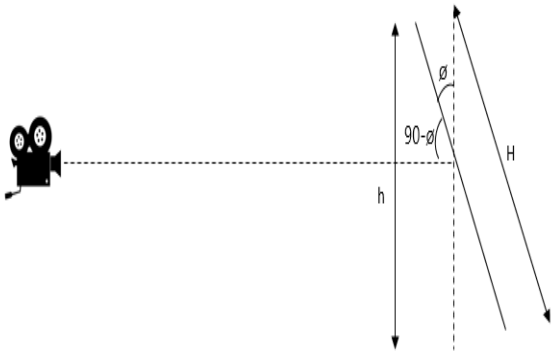

Fig. 4: Person is tilted making 'Phi' angle with camera

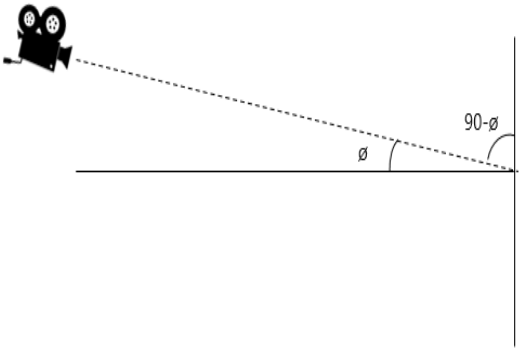

Fig. 5: Camera is tilted making 'Phi' angle

- 'P_ref' is Pixel reference and equal to total pixel present in height of screen.
- 'H_ref' is reference height of the person in 'cm' and pixel height equivalent to 'P_ref'.
- 'h' is pixel height detected by HOG.
- 'P_left' is difference between 'P_ref' and bottom most coordinate of detection box.
- 'x' is height more than horizontal height.
- 'Height' is predicted height.

According to visual perception if referred person moves to and fro with respect to frame of camera there will be change in its pixel height. If person moves backward there will be decrease in pixel height of person and there will be increase in P_left value then,

$$P\_left = P\_ref - \frac{y + h - pad\_h}{\cos\theta} + x$$

$$h = \frac{h}{\cos\theta}$$

where (y+h-pad_h) is the coordinate of detection box with respect to y-axis/scale. According to visual perception, predicted height is directly proportional to pixel height (h) and inversely proportional to (P_ref - $(2 \times P\_left)$)

$$Height = \frac{h - (2 \times pad\_h) \times H\_ref}{P\_ref - (2 \times P\_left)}$$

V. Current Challenges

- Surface of Ground
  - In case of ground surface not being on the same level, the floor not being flat or a shining floor there will be substantial error while estimating height.
- Walking habit of human.
  - All humans have different walking habits, some walk with bowed head, some with forward leaning pose, such cases provide error in detection which will cause substantial error in result.

VI. Results

In the images below we have included cases with false positive as well as false negatives detected using video surveillance. The accuracy for our project is 7 cm to estimate height.

- Here only one person detected from 2 and a object is falsely detected on ground.

Fig. 6

- Here due to sloppy surface there is 10-12 cm error in height detection.

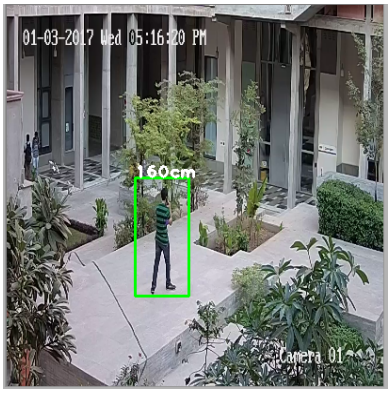

Fig. 7

- Here only one person is detected out of 2.

Fig. 8

- Here the person is not detected due to camouflage effect

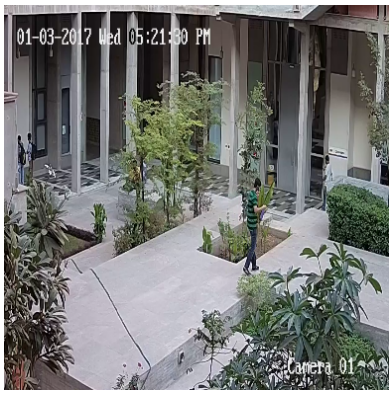

Fig. 9

- Here the person is perfectly detected and measured.

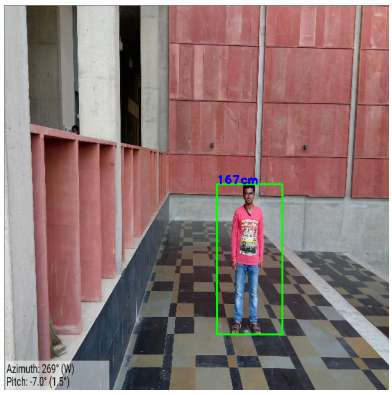

Fig. 10

- Here both persons are perfectly detected.

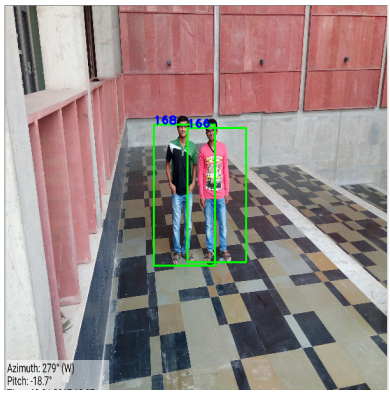

Fig. 11

- Perfect detection in video surveillance.

Fig. 12

- Here the person is perfectly detected and measured even with dark colour.

VII. Future Work

- Accuracy
  The target is to minimise the absolute error and achieve accuracy.
- Low Resolution
  To design it in such a way that it can detect a person having low resolution and distorted image.
- Night Vision
  Detect people at night and also those people wearing clothes which blends in the background.

VIII. Conclusion

This paper proposes a simple method for estimating the human height in video surveillance.The proposed method requires neither any special calibration object nor a special pattern on the ground, such as parallel or perpendicular lines. Only one parameters are retained in the camera model,making the estimation of parameters more efficient. In addition, the proposed method does not rely on computing vanishing points, which is difficult to estimate in practice. The experimental results show that the proposed method can predict the human height from the pixel values at the top of the head and near foot points in the video. The experimental results show that the mean absolute error is only about 5 cm from ground truth data in a walking human-based evaluation.

References

[1] Shengzhe Li, Van Huan Nguyen, Mingjie Ma, Cheng-Bin Jin, Trung Dung Do and Hakil Kim, *A simplified nonlinear regression method for human height estimation in video surveillance* Li et al. EURASIP Journal on Image and Video Processing, 2015.

[2] Pedestrian Detection OpenCV. *PyImageSearch*. N.p., 07 Sept. 2016. Web. 28 Feb. 2017.

[3] *HOG Person Detector Tutorial* HOG Person Detector Tutorial Chris McCormick. N.p., n.d. Web. 28 Feb. 2017.

[4] *Camera calibration With OpenCV*. Camera calibration With OpenCV OpenCV 2.4.13.2 documentation. N.p., n.d. Web. 08 May 2017.

[5] *Finding extreme points in contours with OpenCV.* PyImageSearch. N.p., 17 Mar. 2016. Web. 08 May 2017.