ISSN (Online) : 2394-6849

International Journal of Engineering Research in Electronics and Communication Engineering (IJERECE)
Vol 4, Issue 4, April 2017

Illumination, Attire and Posture Independent Pedestrian Detection

[1] Roshan Baba Kawade (Author)
[1] Department of Electronics and Telecommunication
Government College of Engineering, Amravati,Maharashtra, India,

***Abstract*-- Pedestrian detection is the process of detecting a person from image or a live processing video. This information is sent to user in the forms of sound or video alarm. It helps the user in taking necessary decisions. Pedestrian detection process comes under security and safety aspect.**

This paper introduces a concept of machine learning and its consequences. It combines two algorithms, one is motion based detection and another is feature based detection. Both the algorithms complement with each other. Proposed method increased overall accuracy of detection. Motion based detection is a fast detection method which increased speed of detection. Motion based detection faces illumination problem in real time scenarios. This issue is discussed and solved in this paper. Histogram of oriented gradient (HOG) features are used in feature based detection with support vector machine. Support vector machine is a machine learning technique which is capable to classify given features. Depending upon the shapes, appearances and postures of pedestrians, HOG features will be changed. This make detection process more complex

***Keywords*--Histogram of oriented gradient (HOG), Support vector machine (SVM), Advance driver assistance system (ADAS), Region of Interest (ROI), Automated Teller Machine (ATM), Normalised cross-correlation (NCC)**

I. INTRODUCTION

Vision based human detection is a popular technology widely used in various fields. It includes pedestrian detection in ADAS and human activity detection in security. It is a challenging task as pedestrians are having different appearances and postures. Different environmental conditions also make detection more difficult. It generates false alarm for user. This is undesirable in security and safety constrains. Significant growth in population has increased the number of vehicles on road. Now, the traffic congestion is a global concern and we happen to see many accidents all around, most of which are while crossing the road. Places like banks, ATM machines and government offices are being targeted for robberies. These conditions should be controlled. Pedestrian detection plays a major role in solving these problems. It alerts user when pedestrians are identified. So, probable chances of accidents and robberies will be reduced greatly.

In pedestrian detection speed and accuracy are the major areas of concentration. So, researchers are more concentrating in pedestrian detected at higher speed and accuracy. Following sections consist of literature survey on pedestrian detection, fundamentals of HOG and SVM, Proposed method, results, conclusion and discussions.

II. LITERATURE SURVEY

Basic step in object detection is to find out motion in successive frames. If a prominent motion of person is observed, then it confirms that a pedestrian is present there. This algorithm has been developed at 'University of Liege' to analyse the shape of these silhouettes (Pattern and shape) in order to detect human. This method requires stable background frame in order to detect complete motion. In real time, it generates more false detections. It also affects with illumination variation throughout a day. It is highly applicable in video surveillance field where background is stable.

Pedestrian detection tried with Haar based feature using Adaboost classifier. [3] Haar features basically useful in finding certain shapes of body like eyes, head and shoulder. Adaboost is the short form of 'Adaptive Boosting'. This technique has proposed by Viola-Jones in 2001. As it concentrates on certain body parts, it failed in detection of complete pedestrian body.

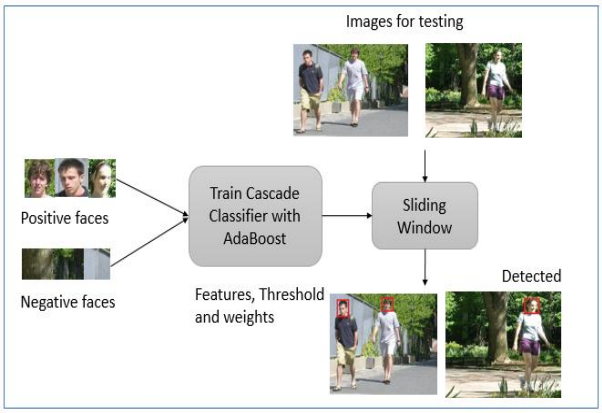


Figure 1. Viola-Jones Face detection

ISSN (Online) : 2394-6849

International Journal of Engineering Research in Electronics and Communication Engineering (IJERECE)

Vol 4, Issue 4, April 2017

The feature based detection is another approach for pedestrian detection. [1] In which features are used to accurately classify pedestrians. HOG and SVM based pedestrian detection has suggested by 'Navneet Dalal' and 'Bill Triggs' in 2005. They have implemented pedestrian detection with greater accuracy but failed in speed of detection. After that, various improvements have suggested by the researchers. These improvements are based on increase speed of extracting the features, eliminate unnecessary feature vectors and combining desired features with other features.

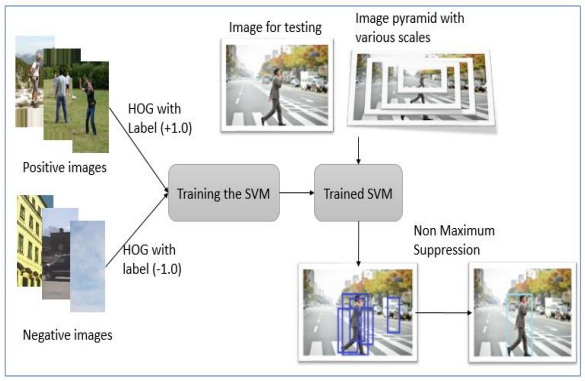


Figure 2. HOG and SVM based Pedestrian detection.

The researchers concluded that there is always a trade-off between speed and accuracy in case of pedestrian detection. If speed is considered, then accuracy will degrade and if accuracy has taken into consideration, then speed of detection will hamper.

Sr. No	Detection Techniques	Accuracy	Speed
1	Motion based Detection.	Low	High
2	Viola- Jones (HAAR and adaboost).	Low	High
3	HOG and SVM (Dalas and Trigg)	High	Low

Table I. Comparison table for Pedestrian detection

Basic fundamentals for pedestrian detection are described below. It includes HOG and SVM basic descriptions.

III. HISTOGRAM OF ORIENTED GRADIENT

HOG is a feature descriptor used in image processing for object detection. It extracts feature of every pixels from image. These features contain oriented gradient information of every pixel. For pedestrian detection, standard (64*128) sized window is considered. Number of steps involved in HOG computation are specified below,

1) Input source images or video frames are considered. Standard window is slide over that source image. Different numbers of feature sets will be generated. These features are used to classify in SVM.

2) Gamma correction

Gamma correction is needed to normalize overall brightness of the image. With the gamma correction, image perceives more information. It helps in clearly finding feature vectors from the image.

3) Every (64*128) sized image is divided into numbers of non-overlapping cells. Standard (8*8) sized cells are considered. (8*16) number of cells are available in that window.

4) Block

Four neighbouring cells are combined to form one block. Every block is of size (32*32). There are (7 *15) number of blocks are available for the given window.

5) Gradient and Orientation

There are various masks are available to find the gradient and orientation from the image. But simple differentiation masks along x- direction and y- direction are used i.e.

$$G_x = [-1 \quad 0 \quad 1] \qquad G_y = \begin{matrix} -1 \\ 0 \\ 1 \end{matrix}$$

These masks are convolved over an image to get absolute magnitude of gradient and orientation of that gradient. Corresponding Gradient and orientation is given by,

$$G_x = [-1 \quad 0 \quad 1] \qquad G_y = \begin{matrix} -1 \\ 0 \\ 1 \end{matrix}$$

For every cell, magnitude and direction is computed and then the histogram has been found out.

6) Computation of the histogram

For every pixel's orientation, corresponding orientation bins is found out and orientation magnitude (G) is voted to that bins. It always considers 9 bins of orientations. Every bin is separated out from each other by 20 deg.

ISSN (Online) : 2394-6849

International Journal of Engineering Research in Electronics and Communication Engineering (IJERECE)

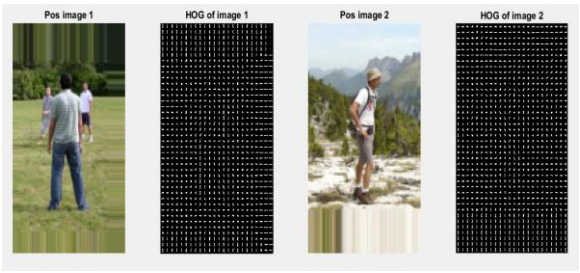
Vol 4, Issue 4, April 2017

7) Histogram normalization

Every bin of histogram is scaled by (*1.5). This is called as histogram normalization. It is an important step to efficiently finding HOG features.

8) Block normalization

This is a final step for finding HOG features. Histogram of four cells are concatenated into a vector of 36 components (4*9). After that, it divides this vectors by its magnitude to normalize respective vector of the block.



(a) (b) (c) (d)

Figure 3. HOG feature visualisation of positive image. (a)(c) contains pedestrian images and (b)(d) Corresponds HOG features.

IV. SVM

SVM is the machine learning technique. It is a discriminative classifier formally defined by a separating hyperplane. Hyperplane is a boundary line for classification of positives and negatives. Here positive consists the features of pedestrian images and negatives contain features of non-pedestrian images. Hyperplane gives largest minimum distance to training examples. Twice of this distance is called 'Margin' in SVM theory.

Hyperplane has function,

$$f(x) = \beta_0 + \beta^T x.$$

Magnitude of function | f(x)| is always 1.

$$|f(x)| = |\beta_0 + \beta^T x| = 1$$

Where, x- training examples, that are closed to hyperplane and called 'support vectors'

Distance from the point x and hyperplane (β,β_0) is given as,

$$Distance = \frac{|\beta_0 + \beta^T x|}{||\beta||} = \frac{1}{||\beta||}$$

$$Margin = \frac{2}{||\beta||}$$

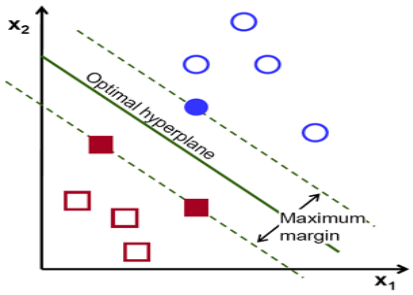


Figure 4. SVM classification for positives (Blue circles) and negatives (Red squares)

Depending on classification techniques used in SVM, SVM is classified as linear kernel SVM, non-linear kernel SVM, Polynomial kernel SVM and RBF kernel SVM. These kernel SVM is named on the basis of hyper plane line used for classification.

V. PROPOSED METHOD

For the pedestrian detection application, mostly HOG and SVM is implemented. But it requires higher processing time for processing of an individual frames. So, it is not suitable for real time applications. Researchers are still struggling to obtain pedestrian detection in higher accuracy and speed.

Figure. 5 represents block diagram of proposed method with various steps in boxes. Two routes are presented, one is illumination free motion based detection and other method is HOG and SVM based detection. Both processes are complimenting with each other. In some cases, motion detection provides good results and in other cases, HOG and SVM technique gives satisfactory result. Both the processes give individual outputs and based on combined observation, pedestrian detection is confirmed. Combining both the algorithms increase overall accuracy in detection.

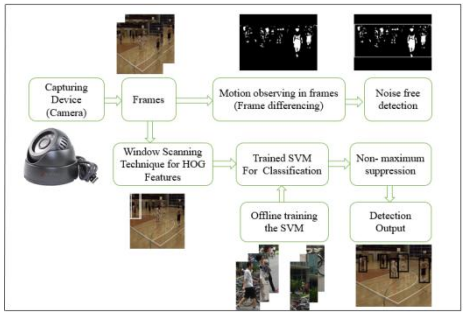


Figure 5. Block diagram of proposed method.

ISSN (Online) : 2394-6849

International Journal of Engineering Research in Electronics and Communication Engineering (IJERECE)
Vol 4, Issue 4, April 2017

A. *Camera*

Real time live streaming videos are captured using camera. Camera selection is an important step in pedestrian detection. For this given application, Z-Vision dome Shaped IR night vision camera with 3.6mm lens is used.

B. *Motion Capturing*

This is one algorithm to detect pedestrian. It finds motion from successive frames. These motions are considered as pedestrian motion and based on that classification is done. Frame differencing method highly suffers with illumination variation. For eliminating of illumination noise, Normalised cross correlation between successive frames are found out. This is compared with threshold value. Depending on that, pixels are classified whether it is belonged under illumination and shadow effect or actual motion. Output gives illumination free pedestrian detection.

C. *Pedestrian Dataset*

For feature based detection, pedestrian datasets collection is required for pre-training the SVM. These datasets contain collection of positive images and negative images. Online datasets are available to download. For example, INRIA pedestrian dataset, Caltech pedestrian dataset, MIT dataset and University of Pennsylvania pedestrian dataset.

D. *Train SVM for definite Dataset*

Training of the SVM is needed before actual implementation. Once the training is completed, SVM is ready to apply on real time online classification. The various features of image are given to SVM for training. Training should be balanced between positives and negatives so that, SVM will not be biased towards any one. Training plays an important role in classification. It pre-classifies datasets by labelling positives as (+1.0) and negatives as (-1.0). This helps the SVM for classification.

E. *Hard negative mining*

When a trained SVM is apply over unknown video data's, then it provides multiple overlapping boxes and false boxes as well. As the SVM has not trained for that particular data, it provides false detection. These false detections should be skipped while detection process. So, again SVM has to train for given positives and negatives with respective to current data. This is called Hard negative mining of SVM.

F. *Non-maximum suppression*

This technique suppresses the false overlapping boxes from detected output. If boxes are 80% and above overlapped with each other, then such boxes are eliminated from the output. This is called non-maximum suppression of blocks. This technique is always applied after detection.

G. *Audio-visual alarm*

Once pedestrian is identified, then rectangular block is drawn over it or audio alarm is generated. This provides alert to user.

VI. RESULTS

Result obtained by feature based detection,

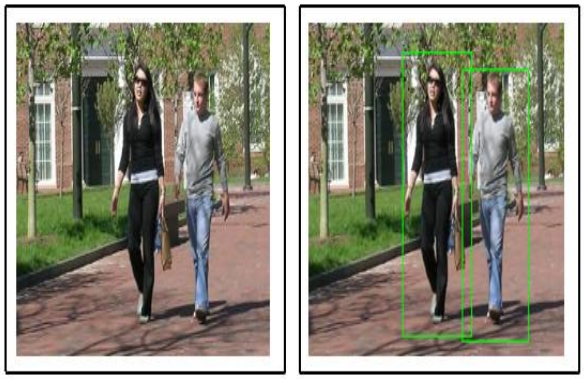


Figure 6. a) Source image b) Detected image

Figure 7. a) Source image b) Detected image.

Figure 8. a) Source image b) Detected image

ISSN (Online) : 2394-6849

International Journal of Engineering Research in Electronics and Communication Engineering (IJERECE)
Vol 4, Issue 4, April 2017

These source images are provided by University of Pennsylvania, Philadelphia. Corresponding detected outputs using HOG and SVM are specified in side image of Figure. 6,7,8. There are some false detections

in Figure 8 output detected image and some blocks are not cover entire pedestrians.
Live streaming results,

Motion based detection:

Figure.9 Motion based detection.

Feature based detection:

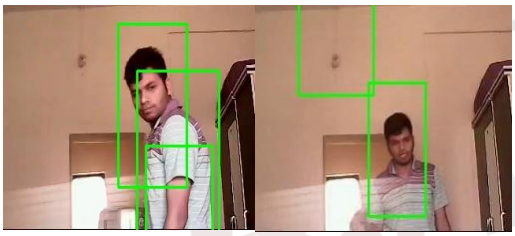


Figure 10. Feature based detection

VII. DISCUSSION

Pedestrian detection faces various problems such as, illumination changes, variation in posture of pedestrians, variation in clothing and variation in shapes of pedestrians.

Illumination Changes:
During entire day, multiple illumination changes happens. This directly effects on image quality. Sometimes image quality becomes worse that information cannot be retrieved. This problem always occurs while working on an image. Some pre-processing techniques like blurring of image, contrast stretching and histogram equalization also gives an improved result. In this method, overall image quality is increased. This results in clearly finding pedestrian from the image.

Normalized cross correlation(NCC) is useful to detect illumination changes. It has been commonly used to evaluate the degree of similarity between two images to be compared. The main advantage of the normalized cross correlation over the cross correlation is that, it is less sensitive to linear changes in the amplitude of illumination in the two compared images. NCC is calculated to estimate how similar the color properties of each motion pixel in current and previous frame and finally reclassified them as either illumination changed pixel or motion pixel. Initially, normalized cross- correlation between previous frame and current frame has been found out. It gives intensity variation information within the current frame and previous frame. This NCC value is compared with defined level of NCC to classify whether defined foreground pixels are from actual motion or noise generated motion.

NCC is given as,

$$NCC = \frac{PC(i,j)}{EP(i,j) \times ET_{(i,j)}}$$

Where,

$$PC(i,j) = \sum_{m=-N}^{N} \sum_{n=-N}^{N} p(i+m, j+n) \times T_{i,j}(m,n)$$

$P(i,j)$is a pixel intensity in previous image

$T_{i,j}(m,n)$is a neighborhood of pixel in current image.

$$EP(i,j) = \sqrt{\sum_{m=-N}^{N} \sum_{n=-N}^{N} (p(i+m, j+n))^2}$$

$$ET(i,j) = \sqrt{\sum_{m=-N}^{N} \sum_{n=-N}^{N} (T_{i,j}(m,n))^2}$$

Variation in Postures, Clothing and Shapes:

Pedestrians possesses different postures, attires and shapes which creates a problem in feature based detection. These problems can be eliminated by training the SVM. A large set of training data includes various postures, shapes and attires of pedestrians are used to train SVM. This helps in efficient classification of pedestrians. Following steps are generally used to overcome above problem,

ISSN (Online) : 2394-6849

International Journal of Engineering Research in Electronics and Communication Engineering (IJERECE)
Vol 4, Issue 4, April 2017

Increase numbers of positives and negatives with different postures and appearances of pedestrian images.

Increase in scaling of the image helps to detect multi-scaled pedestrian from an image. This process is called pyramid of images.

Image pre-processing like gamma correction, blur and filtering should be performed before sending it to SVM. This will help in overcoming the environmental conditions.

Training images should not be biased. There should be a balanced proportion between number of positives and negatives.

VIII. CONCLUSION

Pedestrian detection with combining motion based detection and feature based detection provides quite good result. Motion based detection helps in improving speed of detection and feature based detection gives higher accuracy in detection. Overall this algorithm gives good accuracy with improved speed. Illumination effects are eliminated from the motion based detection technique using NCC thresholding. Variation in postures and attire problem in feature based detection can be solved using efficient SVM training.

REFERENCES

[1] Navneet Dalal and Bill Triggs, "Histogram of oriented gradient for Human detection", IEEE Computer Society Conference on Computer Vision and Pattern Recognition, CVPR 2005, pp 1-7.

[2] Chi Qin Lai and Soo Siang Teoh, "A Review on Pedestrian Detection Techniques Based on Histogram of Oriented Gradient Feature", IEEE Student Conference on Research and Development 2014, pp1-6.

[3] P. Viola, M. J. Jones, and D. Snow. "Detecting pedestrians using patterns of motion and appearance" The 9th ICCV, Nice, France, volume 1,2003, pp 734741.

[4]Liming Wang, Jianbo Shi, Gang Song, I-fan Shen, (2007). Penn-Fudan Database for Pedestrian Detection and Segmentation [Online]. Available: https://www.cis.upenn.edu/~jshi/ped_html/

[5]Syaimaa Solehah Mohd Radzi, ShahrulNizam Yaakob, Zulaikha Kadim, Hon Hock Woon, "Extraction of Moving Objects Using Frame Differencing, Ghost and Shadow Removal". Fifth International Conference on Intelligent Systems, Modelling and Simulation 2014, pp 229-234