A. Sanofer Nisha *et al*, International Journal of Computer Science and Mobile Applications,
Vol.2 Issue. 1, January- 2014, pg. 142-155 **ISSN: 2321-8363**

PART-BASED PEDESTRIAN DETECTION AND TRACKING USING HOG-SVM CLASSIFICATION

Miss. A. Sanofer Nisha[1]

ME II Year

Department of Computer Science and Engineering

Government College of Engineering

Tirunelveli

Mrs. K. Thulasimani[2]

AP/CSE

Department of Computer Science and Engineering

Government College of Engineering

Tirunelveli

Abstract

Pedestrian detection is an essential and challenging task in any intelligent video surveillance system. Despite the challenges, pedestrian detection remains an active area of research in computer vision and a number of approaches have grown steadily in recent years. The major challenge is the development of reliable on-board pedestrian detection systems. This paper proposes a pedestrian detection system that uses a part-based approach. The system is based on a two-stage classifier. Candidates are detected using a Haar cascade classifier and then part-verification is done using the cascade object detector. The Histogram of oriented gradients is found for the whole input and it is combined with the classification done using Support Vector Machine (SVM). The location of the human is detected using a background subtraction algorithm. The detected human are traced by determining the centroid point of the blob and keeping track of its location. The tracking provides the identity of the human location and the location that the human has traversed. The robustness of this system relies on the combination of a HOG and SVM classification.

Keywords: Classifiers, Detectors, Pedestrian Detection, Part-Based Approach, Tracking.

---

1. INTRODUCTION

Pedestrian detection is a key problem in computer vision, with several applications that have the potential to positively impact quality of life. People are among the most important components of a machine's environment, and endowing machines with the ability to interact with people is one of the most interesting and potentially useful challenges for modern engineering. Detecting and tracking people is thus an active area of research[10]-[18]. Just in the US, nearly 5,000 of the 35,000 annual traffic crash fatalities involve pedestrians [6], hence the considerable interest in building automated vision systems for detecting pedestrians [7]. Applications include robotics, entertainment, surveillance, care for the elderly and disabled and content-based indexing. In this paper, we focus on a particular type of system, pedestrian protection systems (PPSs). The objective of a PPS is to detect the presence of both stationary and moving people in a specific area of interest in order to prevent collisions.

The main challenges of a PPS[19] involve detection of pedestrians. These challenges are summarized by the following points:

- Various styles of clothing in appearance.
- The presence of occluding accessories.
- Frequent occlusion between pedestrians.
- Performance in terms of system reaction time and robustness.

Tracking is a method to build a track from a temporal sequence of data. Tracking applies to concurrent inflow of data. Generally tracking is the observing of persons or objects on the move and supplying a timely ordered sequence of respective location data to a model e.g. Capable to serve for depicting the motion on a display capability. Tracking can be employed in a lot of fields such as Computer graphics, Automotive engineering and Satellite tracker.

Though a large amount of objects can be tracked and there exist lots of methodologies to track, human tracking has remained an active area of research for a large amount of time. Because humans are the most integral part of the environment.

A. Sanofer Nisha *et al*, International Journal of Computer Science and Mobile Applications,
Vol.2 Issue. 1, January- 2014, pg. 142-155 ISSN: 2321-8363

The diagram is a simplification that covers the structure of most of the systems, so particular module organizations presented in some papers, for example, interchanging tracking and verification stages, have not been included.

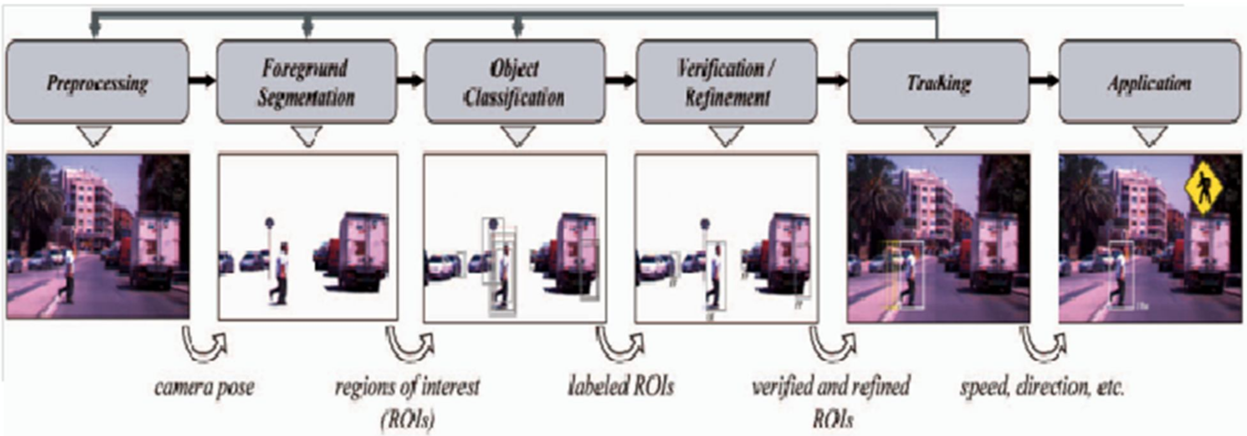

Fig1. Architecture for an On-board Pedestrian detection system

This architecture is then used as a common framework in which we review the different proposals in the literature, making it easier to understand the requirements, responsibilities, and advantages of the techniques in each module.

In this paper, a new approach is proposed for detecting humans. The humans are detected using a haar cascade detector. The HOG and SVM combination makes the system more accurate. The SVM classifiers have been learnt for entire pedestrians [9] and also for rigidly connected assemblies of sub-images [8]. Next, the human is tracked using a background subtraction algorithm that finds the centroid point of the object and keeps track of the path.

2. EXISTING METHODS

A important stream of research within computer vision which has gained a lot of importance in the last few years is the understanding of human activity from a video. Understanding human activity has applications in various fields, the most important of which is surveillance.

2.1 Human detection

Before the complexity of human activity can be understood, we first need automatic methods for finding humans in an image or a video. Once the human is detected, depending on the application, the system can do further processing to go into the details of understanding the human activity.

The relevant literature can be divided into two techniques:

- The Method that requires background subtraction or segmentation and
- The techniques which can detect humans directly from the input without such pre-processing.

2.1.1 Using Background Subtraction

Background subtraction techniques usually find the foreground object from the video and then classify it into categories like human, animal, vehicle etc., based on shape, color, or motion or other features.

The real-time Pfinder system [1] for detecting and tracking humans was invented. The background model used a gaussian distribution in the YUV space at each pixel, and the background model is continually updated. The person is modeled using multiple blobs with spatial and colors components and the corresponding Gaussian distributions. Since the blob was dynamically changing, its spatial parameters were constantly estimated with a Kalman filter.

A system[2] which first detects moving objects by computing optic flow only in regions selected by frame differencing. The optic flow velocity is then used to compute a time to collision with respect to a fixed reference point in the image. The proposed method does not have real-time performance.

A method where a moving object is classified as human based on the spatial uniqueness [3] of the image motion, temporal uniqueness of the human motion, and the temporal motion continuity. First, the moving object is detected by background subtraction, and then the three factors are evaluated.

A system that works based on a shape-based approach [4] for classification of objects is used following background subtraction based on frame differencing. The goal is to detect the humans for threat assessment. The target intruder is classified as human or animal or vehicle based on the shape of its boundary contour.

A method to detect and track a human body in a video. First, background subtraction is performed to detect the foreground object, which involves temporal differencing [5] of the consecutive frames. Other features of the technique include shadow removal.

2.1.2 By Direct Detection

Direct techniques operate on (features extracted from) image or video patches and classify them as human or non-human. We can also classify techniques based on the features which are used to classify a given input as human or not. These features include shape (in the form of contours or other descriptors), color (skin color detection), motion, or combinations of these.

A technique that focused on detecting periodic motions and is applicable to the detection of characteristic periodic biological motion patterns [25] such as walking. The video from a moving camera is first stabilized and frame differencing and thresholding is performed to detect independently moving regions.

A paper that uses the fact that the relative positions[26] (geometric distances) of various body parts are common to all humans, although the pixel values may vary because of the clothes or the illumination.

A paper that deals with the direct detection of humans from static images[27] as well as video using a classifier trained on human shape and motion features was invented. The training dataset consists of images and videos of human and non-human examples. The paper restricts itself to the case of pedestrians (where humans are always in upright walking poses).

A paper that focuses on human motion patterns [28] for robust detection since they are relatively independent of appearance and environmental factors was in practice. The method is not suitable for detecting partially occluded humans.

2.2 Human Tracking

Tracking is a method to build a track from a temporal sequence of data. Tracking applies to concurrent inflow of data.

A system where the multiframe validation[29] and tracking algorithm relies on Kalman filter theory to provide spatial estimates of detected pedestrians and Bayesian probability to provide an estimate of pedestrian detection certainty over time.

The system is to improve robustness in multiobject tracking by coupling object detection and tracking in a non-Markovian hypothesis selection framework[30]. Our approach does not rely on a Markov assumption, but can integrate information

over long time frames to revise its decision and recover from mistakes in the light of new evidence.

3. PROPOSED SYSTEM

Pedestrian detection is an essential and challenging task in any intelligent video surveillance system. The major challenge is the development of reliable on-board pedestrian detection systems. This paper proposes a pedestrian detection system that uses a part-based approach. The system is based on a two-stage classifier. The Haar cascade classifier detects humans based on Haar-like features and the cascade object detector verifies the humans by doing part verification. It does upper body and lower body verification and combines the results to verify the identity of a human. Part-based detection systems perform extremely well with occlusion. The number of high false positives per frame(FPPF) is reduced. In this project, blob tracking is used to calculate the centroid point of human using which the human can be tracked.

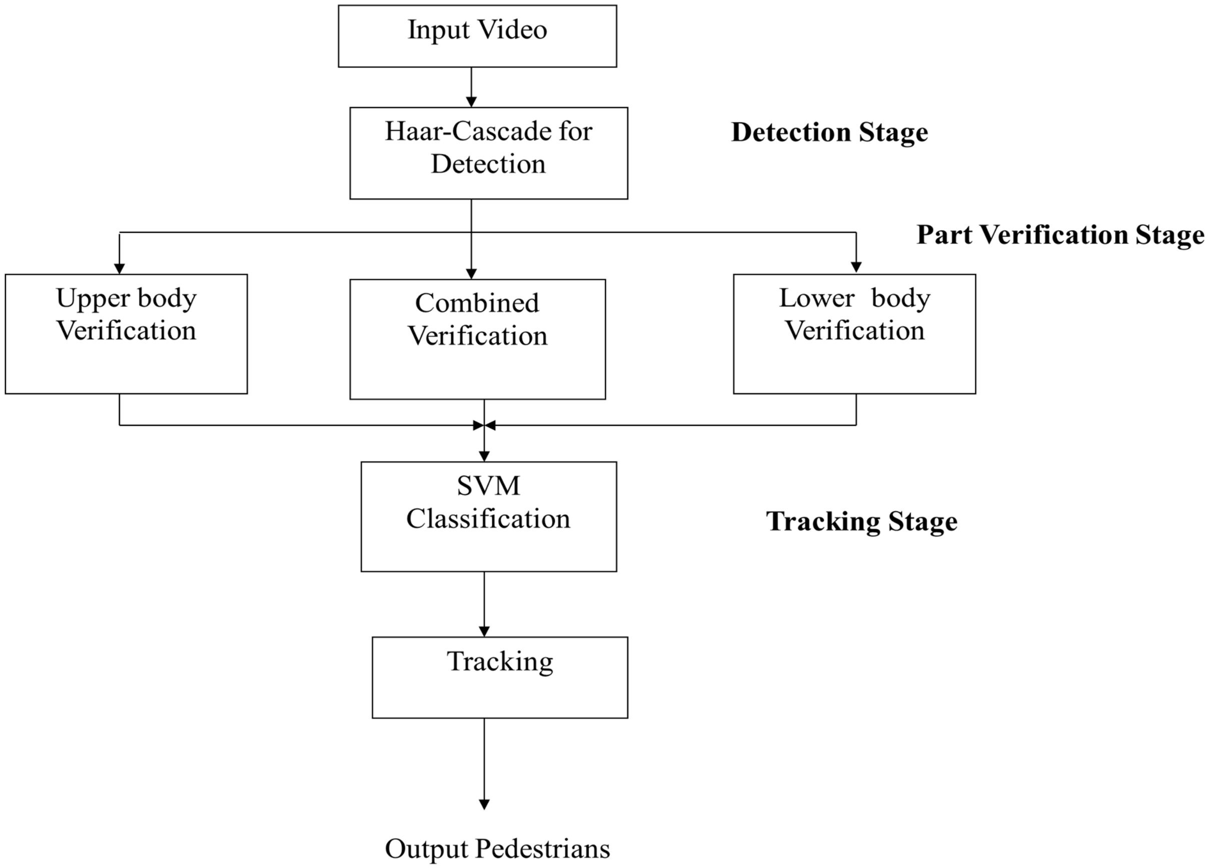

Fig 2. System flow diagram for human detection and tracking

The System consists of the following stages:

Detection Stage

The detection stage is implemented using the haar-cascade detector. The input video is processed frame-by-frame and the objects are detected. Unlike HOG features, Haar-like features do not benefit from having much background included. Another interesting element in the training phase is the choice of data sets used to train the cascade classifier. When the bounding boxes of candidate pedestrians have been obtained, they are passed to the verification stage.

The Haar-Cascade detector is trained using closely cropped humans. Hence the ROI is obtained by the haar-cascade detector which is passed to the verification stage for verifying the body parts. The detected humans are passed to the verification stage for part verification.

Part-Verification Stage

The function of this stage is to verify the identity of the human by processing the whole human and its body parts to confirm the human identity. In this project, the upper body and lower body are used to verify the identity of the human.

The cascade object detector detects the parts of the body. There are many optional properties such as ScaleFactor and MinSize used for defining the range of the detector. The CascadeObjectDetector can be used to detect various parts such as nose, mouth and so on. In this project, the upper and lower bodies of the human si detected.

Fig 4. Detection and upper and lower bodies of the human

The cascade object detector detects the upper and lower bodies of the human which increases the accuracy of the detection.

HOG-SVM Classification

The HOG descriptors are found for the whole input and it is used for classifying using SVM. Most of the papers use the HOG-SVM solution presented by Dalal and Triggs [20]. It is much more robust and generally detects pedestrians in harder situations, while keeping a low number of false positives. Its problem lies in processing speed. We combine the speed of the Haar detector with the robustness of a part-based HOG-SVM detector. The base for the method used in this paper was first presented by Geismann and Schneider [21], but is also covered by others in various versions [22], [23].

	1	2	3	4	5
1	0.2437	0.1092	0.6101	0.0990	0.4876
2	0.2257	0.1126	0.6341	0.0972	0.4611
3	0.2255	0.1136	0.6336	0.0966	0.4616
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]

Fig 5. HoG descriptors for the images

Our algorithm uses classic dense HOG descriptors (as opposed to the sparse descriptors used in [21]). They are calculated using integral images in an effort to speed up the process, as described in [24]. Since HOG works best if some amount of background is introduced to the detection window, the ROIs are resized appropriately from the tight boxes that are returned by the detection stage. Then the content of the ROIs is scaled so it matches the size the SVMs were trained with. At this point the HOG is calculated and passed on to the SVMs. As in the detection stage, each SVM is trained with the INRIA training set.

Tracking Stage

The tracking stage is used to detect or trace the path of the object. In this system, the movement of the human is tracked using a background subtraction algorithm. The centroid point of the blob is calculated and the tracking mechanism is performed using the centroid point. It determines the motion and displays the path traversed by the human.

Fig 6. Tracking of a single human

The tracking can be done for single and multiple humans. The tracking process is done by doing background subtraction for the input video. The centroid point of the human is found out and the path is calculated from time to time.

The tracking process can be improves further by tracking multiple humans which is done by taking a lot of considerations.

Fig 7. Tracking of multiple humans

This is done by doing background subtraction for the input video. The humans found by the background subtraction algorithm are known as blobs. The centroid points of the blobs are calculated and the locations of the blobs are tracked based on time. Hence the sequential motion of the humans can be determined by drawing bounding box around the object.

The analysis of the false positives per frame versus true positives per frame is calculated and the performance is evaluated.

REFERENCES

1. C. R. Wren, A. Azarbayejani, T. Darrell, and A. P. Pentland. Pfinder: real-time tracking of the human body. IEEE Transactions on Pattern Analysis and Machine Intelligence, 19(7):780–785, 1997.

2. H. Elzein, S. Lakshmanan, and P. Watta. A motion and shapebased pedestrian detection algorithm. IEEE Intelligent Vehicles Symposium, pages 500–504, 2003.
3. T. Haga, K. Sumi, and Y. Yagi. Human detection in outdoor scene using spatio-temporal motion analysis. International Conference on Pattern Recognition, 4:331–334, 2004.
4. D. J. Lee, P. Zhan, A. Thomas, and R. Schoenberger. Shape-based human intrusion detection. SPIE International Symposium on Defense and Security, Visual Information Processing XIII, 5438: 81–91, 2004.
5. Jianpeng Zhou and Jack Hoang. Real time robust human detection and tracking system. IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 3:149 – 149, 2005.
6. U. Shankar, "Pedestrian roadway fatalities," Department of Transportation, Tech. Rep., 2003.
7. D. Geronimo, A. M. Lopez, A. D. Sappa, and T. Graf, "Survey on pedestrian detection for advanced driver assistance systems," IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 32, no. 7, pp. 1239–1258, 2010.
8. Mohan, C. Papageorgiou, and T. Poggio. Example-based object detection in images by components. ieee *Transactions on PAMI*, 23(4):349–361, 2001.
9. C. Papageorgiou and T. Poggio. A trainable system for object detection. *International Journal of Computer Vision*, 38(1):15–33, 2000.
10. P. Dollar, C. Wojek, B. Schiele, and P. Perona, "Pedestrian detection: An evaluation of the state of the art," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 34, no. 4, pp. 743–761, Apr. 2012.
11. D. Geronimo, A. Lopez, A. Sappa, and T. Graf, "Survey of pedestrian detection for advanced driver assistance systems," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 32, no. 7, pp. 1239–1258, Jul. 2010.
12. A. Møgelmose, A. Prioletti, M. M. Trivedi, A. Broggi, and T. B. Moeslund, "A two stage part-based pedestrian detection system using monocular vision," in *Proc. 15th IEEE Int. Conf. Intell. Transp. Syst.*, Sep. 2012, pp. 73–77.

13. A. Prioletti, P. Grisleri, M. Trivedi, and A. Broggi, "Design and implementation of a high performance pedestrian detection," presented at the IEEE Intell. Veh. Symp., Gold Coast, Australia, 2013, Paper WePO2T1.25.

14. N. Dalal and B. Triggs, "Histograms of oriented gradients for human detection," in *Proc. IEEE Comput. Soc. Conf. CVPR*, 2005, vol. 1, pp. 886–893.

15. S. Krotosky and M. Trivedi, "On color-, infrared-, and multimodal-stereo approaches to pedestrian detection," *IEEE Trans. Intell. Transp. Syst.*, vol. 8, no. 4, pp. 619–629, Dec. 2007.

16. A. Broggi, P. Cerri, S. Ghidoni, P. Grisleri, and H. Jung, "A new approach to urban pedestrian detection for automatic braking," *IEEE Trans. Intell. Transp. Syst.*, vol. 10, no. 4, pp. 594–605, Dec. 2009.

17. P. Geismann and G. Schneider, "A two-staged approach to vision-based pedestrian recognition using Haar and HOG features," in *Proc. IEEE Intell. Veh. Symp.*, 2008, pp. 554–559.

18. T. Gandhi and M. Trivedi, "Pedestrian protection systems: Issues, survey, and challenges," *IEEE Trans. Intell. Transp. Syst.*, vol. 8, no. 3, pp. 413–430, Sep. 2007.

19. D. Geronimo, A. Lopez, A. Sappa, and T. Graf, "Survey of pedestrian detection for advanced driver assistance systems," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 32, no. 7, pp. 1239–1258, Jul. 2010.

20. N. Dalal and B. Triggs, "Histograms of oriented gradients for human detection," in Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on, vol. 1. Ieee, 2005, pp. 886–893.

21. P. Geismann and G. Schneider, "A two-staged approach to vision based pedestrian recognition using Haar and HOG features," in Intelligent Vehicles Symposium, 2008 IEEE. IEEE, 2008, pp. 554–559.

22. X. Yuan, X. Shan, and L. Su, "A Combined Pedestrian Detection Method Based on Haar-Like Features and HOG Features," in Intelligent Systems and Applications (ISA), 2011 3rd International Workshop on. IEEE, 2011, pp. 1–4.

23. W. Yongzhi, X. Jianping, L. Xiling, and Z. Jun, "Pedestrian Detection Using Coarse-to-Fine Method with Haar-Like and Shapelet Features," in Multimedia Technology (ICMT), 2010 International Conference on. IEEE, 2010, pp. 1–4.

24. F. Porikli, "Integral histogram: A fast way to extract histograms in cartesian spaces," in Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on, vol. 1. IEEE, 2005, pp. 829–836.

25. R. Cutler and L. S. Davis. Robust real-time periodic motion detection, analysis, and applications. IEEE Transactions on Pattern Analysis and Machine Intelligence, 22(8):781–796, 2000.

26. Akira Utsumi and Nobuji Tetsutani. Human detection using geometrical pixel value structures. Fifth IEEE International Conference on Automatic Face and Gesture Recognition, page 39, 2002.

27. P. Viola, M. J. Jones, and D. Snow. Detecting pedestrians using patterns of motion and appearance. IEEE International Conference on Computer Vision, 2:734–741, 2003.

28. H. Sidenbladh. Detecting human motion with support vector machines. Proceedings of the 17th International Conference on Pattern Recognition, 2:188–191, 2004.

29. I. Alonso, D. Llorca, M. Sotelo, L. Bergasa, P. de Toro, J. Nuevo, M. Ocaña, and M. Garrido, "Combination of feature extraction methods for SVM pedestrian detection," *IEEE Trans. Intell. Transp. Syst.*, vol. 8, no. 2, pp. 292–307, Jun. 2007.

30. B. Leibe, K. Schindler, and L. Van Gool, "Coupled detection and trajectory estimation for multi-object tracking," in *Proc. 11th IEEE ICCV*, Oct. 2007, pp. 1–7.

