Real-Time Pedestrian Detection System

Tinto Raj R V[1], Surya Priya .S[2]

P.G scholar, Department of CSE, Sarabhai Institute of Science & technology, Kerala, India

Assistant Professor, Department of CSE, Sarabhai Institute of Science & technology, Kerala, India

Abstract: ***This work is concerned with the challenging task of pedestrian detection in real world environments. That is, the aim is to successfully localize pedestrian in video surveillance system despite the presence of background clutter or partial occlusions. Even though pedestrian has many practical applications and has been an active area of research for many years, it has not been until recently that pedestrian recognition algorithms have become robust enough to deal with scene of realistic situations. This work proposes algorithms and algorithmic extensions, which further enhance detection accuracy compared to existing state- of-the –art approaches. Persons and pedestrian however are not rigid and their appearance changes greatly depending on the body articulation or pose. The variations of colors in clothing and textures add further difficulties. The proposed system has two phases learning phase and detection phase. In the learning phase create robust feature set that allows human form to be discriminated clearly even in cluttered background under different illumination. Object finding technique is based on Histogram of oriented gradients(HOG). Here shape based analysis is carried out. This thesis puts particular emphasis on pedestrian on detection in improved detection accuracy with decreased computational cost.***

Keywords: Visual features, object detection, Histogram of oriented gradients (HOG), pedestrian detection, real-time systems

1. Introduction

Pedestrian and person detection has received much attention in computer vision literature. Detecting people in images is a problem with a long history [14, 15, 16, 22], many pedestrian recognition approaches have been developed in the field of video surveillance or intelligent vehicles. More recently contributions also come from the object categorization community [2, 9, 10, 12]. These contributions focus on robust object models [6, 11, 13], and are typically less driven by application requirements such as real-time capabilities. The focus on robustness is also reflected in the fact that these models usually require higher resolution images [6]. The proposed work builds upon a successful pedestrian detection system in unconstrained environments. Accurate pedestrian detection would have immediate and far reaching impact to applications such as surveillance, traffic control, robotics, content based indexing, advanced human machine interfaces and automotive safety, among others [28].

This thesis focuses on the issue of pedestrian detection in images and videos. In practice, it involves the construction of human detectors, where the detectors search given images for human and localize them. A pedestrian detection system can be considered as a combination of two key factors: a feature extraction algorithm that transformers image regions to feature vectors and a detector that uses the computed features to make pedestrian/non-pedestrian decisions. This work targets general purpose human detectors that do not make strong contextual assumptions. More robust discriminant image descriptors simplify the classification task allowing human to be discriminated more easily with less training data and less computational costs.

The rest of the paper is organized as follows: In section II we discuss some of the related work in this field. In section III several major challenges for pedestrian detection are reviewed. In section IV discusses an overview of activity diagram of the proposed system in step by step procedure. In section V our pedestrian detection system is discussed giving brief description. In section VI we have explained experiment result followed by concluded in section VII.

2. Literature Review & Related Work

There is an extensive literature on object detection, but here I mention just few relevant papers on human detection [1, 3, 5, 7, 8, 19, 24]. Various approaches for pedestrian detection are discussed in [8]. However, different objects like bird, vehicle, etc may be present in the scene, so it is very important that we correctly distinguish human from other objects. Papageorgiou et al [22] proposed one of the first sliding window detectors, applying support vector machines (SVM) to an over-complete dictionary of multi-scale Haar wavelets. Viola and Jones [VJ] [4, 23] built upon these ideas, introducing efficient moving person detector, here more complex regions are trained using Adaboost, based on Haar like wavelets and space-time differences. Dollar et al. [17] proposed an extension of [VJ] where Haar-like feature are computed over multiple channels of visual data, and they provide a simple and uniform framework for integrating multiple feature types. In [20] this approaches was extended to fast multi-scale detection after it was demonstrated how features computed at single scale can be used to approximate features at nearby scales. In [6] they propose model for object detection have a multi-resolution structure. They treat the features in a scale dependent manner. Recently in 2013 authors [26] have presented a model for human detection in range images captured from a vertically oriented camera by analysis of 3D range data. In contrast, our detector uses a simpler architecture but appears to give significantly higher performance on pedestrian images.

3. Analysis of Problem

Pedestrian detection is a key problem in computer vision, with several applications that have the potential to positively impact quality of life. The main challenges involved in pedestrian detection are summarized by the following points:

- Variable appearance in terms of styles and cloth.
- The presence of occluding accessories.
- Frequent occlusion between pedestrians.
- Wide variety of articulated poses.
- Environmental situation like illumination change.
- Performance in terms of system reaction time and robustness.

In comparing different detection schemes, one notices that representation at the front end are progressively enriched, like more channels, finer scale sampling, enhanced normalized schemes. This has helped the dramatic improvements in detection accuracy. Unfortunately increase in detection accuracy has been paid for with increased computational cost.
Most recent detectors require multiple second to process a single image. Currently several human detection algorithms displays a good performance in controlled conditions. However, when these algorithms are applied to real world scenarios there is a sharp decrease in their performances. Hence the tasks of human detection in non-controlled environment remain unsolved.

4. Proposed Work

The pedestrian detection approach implemented in this proposed work is based on shape based analysis over the image at multiple scales, and running a human or non-human classifier in every position. It has two phase learning phase and detection phase.

Both phase involve feature extraction based on histogram of oriented gradients. The histogram of oriented gradient (HOG), uses normalized histogram of image gradients and orientation to model the local feature which comes from shape and appearance. So as to obtain a robust feature set that allows the human form to be discriminated, the local gradients are quantized. The quantization process is carried out according to orientation value corresponding to gradient angles. The values of histogram bins are derived from horizontal and vertical gradient magnitudes. Also the processing is effected in a spatial grid of cells with overlapping blocks. Within each overlapping block, a feature vector is extracted by gathering the normalized histogram from the adjacent and necessary spatial cell. The final feature vector comes from concentrated feature vectors of each block, followed by classification. This work puts particular emphasis on pedestrian detection in improved detection accuracy with decreased computational cost. The dynamic nature of our work is shown below using activity diagram.

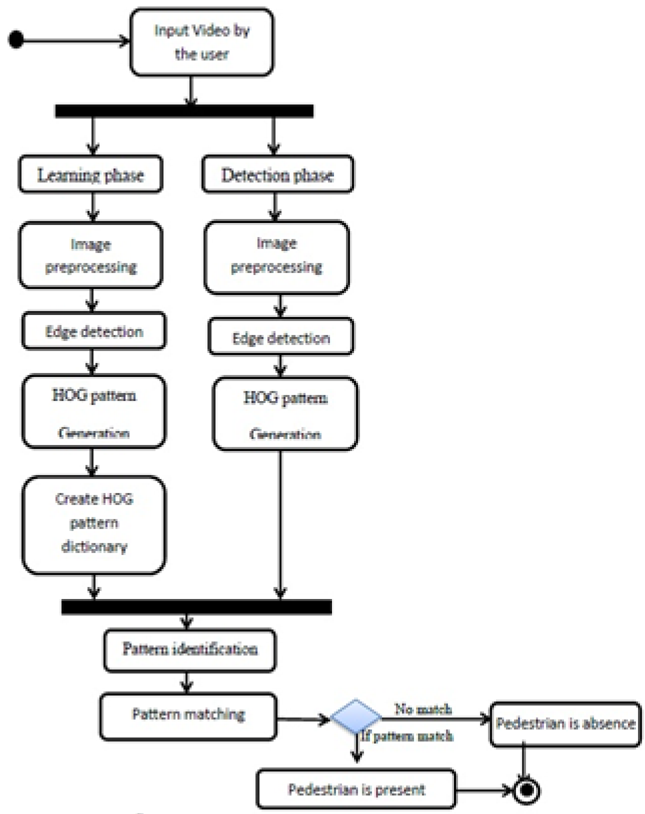

Figure 1: Activity diagram of pedestrian detection system

5. Pedestrian Detection

Firstly the video stream from a live webcam is captured. The noise in the video should be eliminated for better recognition results. Then study how the image frames on the basis of its resolution, RGB configuration, size, type etc. The output of this analysis will be image matrix that contains information about the basic histogram. Then perform pyramid analysis. Multi-resolution multi-orientation decompositions are one of the fundamental techniques of image analysis. Here develop a filter based representation to decompose images into information at multiple scales to extract texture/structure of human and to attenuate noise.

Edges are significant local changes of intensity in an image. The goal of edge detection is to produce the line drawing of a scene from an image of that scene and to extract important features from the edges of an image. Canny Edge Detector Algorithm [25], is used for edge detection. In this module pyramid features based on edges detector is extracted. The spatial pyramid method can be viewed is an updated version of a global appearance based method, or hybrid between local and global representation.

a) HOG pattern generation

Feature vector are generated by using Histogram of oriented Gradients (HOG). HOG are feature descriptors used for the purpose of object detection. This technique counts occurrences of gradient orientation in localized portion of an image. HOG is calculated as follows:

At the first, the algorithm converts an image to gray scale to remove color information. In the next step strength and edge degree are calculated. For a gradient computation, first gray scale image is filtered to obtain '*x*' and *'y'* derivatives of pixels. After calculating '*x*, *y'* derivatives (*Ix* and*Iy*), the magnitude and orientation of the gradient is also computed:

$|G| = \sqrt{Ix^2 + Iy^2}$ And $\Theta = arctan\frac{Ix}{Iy}$

Then image is split to local area called *cell area*, which is composed of 16×16 pixel square. Each pixel calculates a weighted vote for an edge orientation histogram channel based on the orientation of the gradient element centered on it, and the votes are accumulated into orientation bins over local spatial regions that we call *cells*. Cell can be either rectangular or radial. In the next step an edge histogram is built using edge degree and the strength calculated in the previous step. Gradient strength vary over the wide range of owing to local variations in illumination and foreground-background contrast, so effective local contrast normalization turns out to be essential for good performance. Different normalization schemes are evaluated and most of them are based on grouping cells into larger spatial blocks and contrast normalizing each block separately. Normalization introduces better invariance to illumination, shadowing, and edge contrast. It is performed by accumulating a measure of local histogram *energy* over local groups of cells that we call *blocks*. The result is used to normalize each cell in the block. Typically each individual cell is shared between several blocks, but its normalization is block dependent and thus different. The final descriptor is then the vector of all components of the normalized cell responses from all of the blocks in the detection window.

6. Experiment Result

The human detection is constructed via a method for classifying individual images region. It is divided into training and testing phases is used to make a binary classifier which gives human or non-human decisions for input image windows. The testing phase uses the classifier to perform a dense multi-scale scan reporting human decisions at each location of the testing images. The overview of training and testing phases is provided in figure (1).

There are several notable findings in this work. HOG is affected by gradient quality, the choice of number of bins, normalization method and so on. In order to improve the performance, strong edge information is needed. Also gradient should be calculated at the finest available scale in pyramid layer when HOG can give fine orientations on the other hand, strong local contrast normalization produces good result. Overlapping cells sizes are an important component in raising accuracy. The overlapping scheme makes each cell normalized several times with respect to different local supports. In processing, each cell appears four times with different normalizations. The cell size has an impact on the quantity of information. For pedestrian detection, the cell size which approximates human part-template size gives the better performance, such as detection rate raise from 83.45% (8×8 cell) to 98.26% (16×16 cell) in training data set.

7. Conclusion

This work has described a complete framework for the problem of detecting objects in images and videos. The proposed approach builds upon ideas in machine learning, computer vision and image processing to provide a general, easy to use and fast method for pedestrian detection. Our main contribution is the development of robust images feature sets for object detection tasks. And also proposed feature set based on well normalized grids of gradient orientation histograms. These features provide some invariance to shifts in object location and changes in shape and good resistance to changes in illumination and shadowing, background clutter and camera view point. For increasing detection rate, capturing fine detail with unsmoothed gradients and fine orientation voting, strong normalized and overlapping blocks are used. The descriptors do not involve any arbitrary thresholding of edges and they are relatively fast to compute.

Our analysis and experiments show that it is possible to inexpensively estimate features at a dense set of scales by extrapolating computations carried out expensively, that is improved detection accuracy has been accompanied by decreased computational costs. And the system has moderate memory consumption.

References

[1] PiotrDoll_ar, Ron Appel, Serge Belongie, and PietroPerona,"Fast Feature Pyramids For Object Detection"(IEEE Transaction on pattern analysis and machine intelligence, vol. 36, no. 8, august 2014)

[2] B. Alexe, T. Deselaers, and V. Ferrari, "What Is an Object?"IEEE Conf. Computer Vision and Pattern Recognition (CVPR), 2010.

[3] N. Dalal and B. Triggs, "Histograms of Oriented Gradients forHuman Detection," EEE Conf Computer Vision and PatternRecognition (CVPR), 2005

[4] P. Viola and M. Jones, "Rapid Object Detection Using a Boosted Cascade of Simple Features," IEEE Conf. Computer Vision and Pattern Recognition (CVPR), 2001.

[5] S. Walk, N. Majer, K. Schindler, and B. Schiele, "New Features and Insights for Pedestrian Detection," , IEEE Conf. Computer Vision and Pattern Recogn (CVPR), 2010.

[6] D. Park, D. Ramanan, and C. Fowlkes, "Multiresolution Models for Object Detection," 11th European Conf. Computer Vision ECCV), 2010

[7] Q. Zhu, S. Avidan, M. Yeh, and K. Cheng, "Fast Human Detection Using a Cascade of Histograms of Oriented Gradients," IEEE Conf. Computer Vision and Pattern Recognitio (CVPR), 2006

[8] P. Dollar, C. Wojek, B. Schiele, and P. Perona, "Pedestrian Detection: An Evaluation of the State of the Art,"IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 34, no. 4, pp. 743–761, 2012

[9] R. Fergus, P. Perona, and A. Zisserman, "Object Class Recognition by Unsupervised Scale-Invariant Learning," Proc. IEEE Conf. Computer Vision and Pattern Recognition (CVPR), 2003.

[10] S. Maji, A. Berg, and J. Malik, “Classification Using Intersection Kernel SVMs Is Efficient,” Proc. IEEE Conf. Computer Vision and Pattern Recognition (CVPR), 2008.
[11] L. Itti, C. Koch, and E. Niebur, “A Model of Saliency-Based Visual Attention for Rapid Scene Analysis,” IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 20, no. 11, pp. 1254-1259, Nov. 1998.
[12] Kart-Leong Lim, HamedKianiGaloogahi, “Shape Classification Using Local and Global Features” proceedings of Fourth Pacific-Rim Symposium on Image and Video Technology, 2010.
[13] M. Riesenhuber and T. Poggio, “Hierarchical Models of Object Recognition in Cortex,” Nature Neuroscience, vol. 2, pp. 1019-1025, 1999.
[14] T. Tsukiyama and Y. Shirai. Detection of the movements of persons from a sparse sequence of tv images. PR, 18(3-4):207–213, 1985.
[15] D. M. Gavrila and V. Philomin. Real-time object detection for “smart” vehicles. In ICCV, 1999.
[16] Y. Song, X. Feng, and P. Perona. Towards detection of human motion. In CVPR, 2000.
[17] P. Doll_ar, Z. Tu, P. Perona, and S. Belongie, “Integral Channel Features,” Proc. British Machine Vision Conf. (BMVC), 2009.
[18] Hong Liu, Tao Xu, Xiangdong Wang, and YueliangQian “Related HOG Features for Human Detection Using Cascaded Adaboost and SVM Classifiers” (19th International Conference, MMM 2013
[19] P. Doll_ar, R. Appel, and W. Kienzle, “Crosstalk Cascades for Frame-Rate Pedestrian Detection,” Proc. 12th European Conf. Computer Vision (ECCV), 2012.
[20] P. Doll_ar, S. Belongie, and P. Perona, “The Fastest Pedestrian Detector in the West,” Proc. British Machine Vision Conf. (BMVC), 2010.
[21] P.Dollar, R. Appel, and W. Kienzle, “Crosstalk Cascades for Frame–Rate Pedestrian Detection“, (12th European Conf.ComputerVision(ECCV), 2012).
[22] C. Papageorgiou and T. Poggio. A trainable system for object detection. IJCV, 38(1):15–33, 2000
[23] Hong Liu, Tao Xu, Xiangdong Wang, and YueliangQian, “Related HOG Feature for Human Detection Using Cascaded Adaboost and SVM Classifiers (19th International Conference, MMM,2013)
[24] P. Viola, M. Jones, and D. Snow, “Detecting Pedestrians Using Patterns of Motion and Appearance,” Int’l J. Computer Vision, vol. 63, no. 2, pp. 153-161, 2005.
[25] Rashmi, Mukesh Kumar and RoshiniSaxena, “Algorithm and Technique on Various Edge Detection: A Survey”, Int’l (SIPIJ) Vol.4, No.3, June 2013.
[26] RusiAntonovFilipov, Flavio Luis Cardeal Padua, Marco Aurelio Buonocarone, “Pylon grid: A fast method for human head detection in range images”,Journal of Neurocomputing(Elsevier),Vol.100, pp-74-85, 2013.
[27] Subra Mukherjee and Karen Das, “Omega Model for Human Detection and Counting for application in Smart Surveillance System”, Intl. journal of advanced computer science and applications, Vol.4, No.2,2013
[28] Manoranjan Paul, Shah M E Haque and SubrataChakaraborty, “ Human Detection in Surveillanace Videos and its Application- a review”, EURASIP journal on Advances in Signal Processing 2013:176
[29] R.ShanmugaPriya, S.Karthick. C.Thulasiyamal, “A Survey Of Various Techniques For Human Motion Detection From Video”, Intl. Journal of advanced Research, Volume 2 Issue 4,332-335 (2014)