REAL TIME TRACKING OF MOVING PEDESTRIAN IN SURVEILLANCE VIDEO

Mr D. Manikkannan[1], A.Aruna[2]
Assistant Professor[1], PG Scholar[2]
Department of Information Technology[1], Adhiparasakthi Engineering College[2], Melmaruvathur

ABSTRACT-Pedestrian detection is an active area of research in video Surveillance. It remains a quite challenging problem in many applications where many factors cause a mismatch between source dataset used to train the pedestrian detector and samples in the target scene. In this paper, we propose a novel scene-specific detector with adaptation model for merging plentiful source domain samples with scared target domain samples to create a scene-specific pedestrian detector that performs as well as rich target domain simples are present. Our approach combines the prioritized score based transferability. It derived from the prediction consistency with the source classifications, to selectively choose the positive confidence in source domains to the target domain. Through clustering based multimode models are used for SVM classifier with KNN based feature selections. Experimental results show that our approach can improve the pedestrian detection rate, and also with the insufficient labeled data in target scene. Finally, the selected samples are processed with PCA transform to get the resulting features for classification. Performance is based on different parameters by using the support vector machines (SVMs) serving as the classifier. The proposed method, the accuracy of the SVM classifier can be improved significantly.

Keywords: pedestrian detection, SVM Classifier, KNN, PCA Transform

I. Introduction

Pedestrian detection is a key problem in computer vision, which has promising applications in various domains such as surveillance video, robotics etc. Pedestrian detection technique into vehicle-mounted monitoring system helps driver to timely stop the vehicle avoiding the traffic collision, by which it can identify pedestrians in the front of vehicle and also can be applied into apartment, supermarket, museum and scenic spots to count the number of pedestrians and figure out the pedestrian flow statistics for business optimization and with the development of society and technique. The complex application scenarios are demanded to be handled. For example, the speed requirement of embedded device of vehicle should be fast, and the detection accuracy in crowded area of supermarket should be high.

Various methods have been proposed to the pedestrian detection for last decades, such as Histograms of Oriented Gradients (HOG) Feature, Integral Channel Feature (ICF), SVM Classifier (SCC), Local Binary Patterns (LBP) . For the integral channel feature (ICF) can be traced back to the computer vision research. It has been used to efficiently compute the histograms of oriented gradients. The histogram can be computed by quantizing an image into multiple channels. For histograms of oriented gradients divides the original image into dependence blocks where each block is further divided into smaller cells. In that each cell, the histograms of oriented gradients are computed respectively. From the cell to block, all HOG are connected to a total histograms of oriented gradient for the original image. The SVM classifier trains a set of classifiers from simple and complex. The detection windows beyond to target are excused, and the complex and hard detection windows are delivered to SVM classifier. This approach saves the computing consume and sorting time. Only the positive sample are needs to compute its feature which improves the computational efficiency. Finally, for the local binary patterns (LBP) is used to describe the texture of image feature. The selected pixel are considered as the threshold value which is assigned to 1 or 0. The probability value of 8 pixels is 256 and it counts the value of histograms of oriented gradients of each image area when using LBP method. Based on existing researches, this paper proposes to design a fast pedestrian detection for video surveillance system.

ISSN: 2278 – 7798

II. RELATED WORK

Various methods have been proposed to pedestrians detecting. They have a good performance in the upright holistic pedestrian detection. Papa Georgiou *et. al.*, [5] proposed a sliding window-based target detector, where the SVM classifier [9] is used to the object identification of multi-scale Harr feature. Viola and Jones *et. al.*, [6] improved the real-time face recognition system at the running speed level. Their approach mainly computes the picture integral to reduce the redundancy. It only calculates the minor features for simple classifier which can rapidly remove most of negative samples for accelerating the target task checking speed. This method is regarded as a basis research for successor. Most researches are contributed to the new features extraction and its utilization. Dalal and Triggs *et. al.*, [2] given a feature using histogram of oriented gradient (HOG), and applied it into pedestrian detection. The method has a good effectiveness since it has reduced the miss-detection ratio at least one magnitude, comparing to the Harr-based detector. Zhu [3] combined the HOG feature and Viola-Jones method via calculating the gradient integral and adopting the cascaded classifier. This approach achieved pedestrian detection from the 320*240 px image which helps reducing the detection time consumption to 0.1 second. In latter, Wang *et. al.*, [7] added LBP features of image texture descriptions to HOG features to increase detection rate to 97.9% when the miss-detection ratio of negative windows is 0.0001 using INRIA dataset.

Other features also are researched and applied into specific pedestrian detection systems, such as color features and running features. Dollar et. *al.*, [8] proposed the concept about integral channel features which extracted gradient and color channels from the transformed image. After that, the accumulative integral value of special channels area is selected as image feature, which has a better effectiveness during its applications. In the improvement aspect of learning and classification algorithms, Majid *et. al.*, [9] proposed an IKSVM classifier in which IKSVM classifier had a better effectiveness than the liner SVM classifier in general. Turnel *et. al.*, [10] proposed a Riemann flow-based classifier for pedestrian detection. To handle various kinds of problems in Viola-Jones-based cascaded classifier, soft-cascaded classifier is proposed in papers [11, 12]. The strong classifier consists of some week classifiers trained by Boosting algorithm. The classifier is sorted with a threshold value which is compared to the total response value computed by all current classifiers. If it is less than the threshold value, the classifier immediately outputs negative value.

To improve the detection speed of pedestrian detection, these are mainly concerning on reducing the feature computing time or classifier computing time. For the former, integral feature is used, for example, Zhu *et. al.*, adopted HOG feature, Dollar *et. al.*, adopted integral channel feature. For the latter, the approximate calculation of adjacent scale features is used. Due to the different pedestrian has different size in scenarios, the image should be zoomed for several times. Thus, the zoomed image needed to be recomputed again. It made the computing task increased rapidly. They also found that the channel value is changeable with exponential when the image is zoomed. The approximately calculating the feature can reduce the computing time. Thus, Benenson *et. al.*, introduced this ideal to channel computing that multi-image scales were replaced by multi-classifiers. It made feature computing only execute at only once [31].

Adopting GPU helps to reduce the classifier's computing time. Due to the detection computing is usually independent, sliding window detection of multi-image scale intensions can be computed in parallel. Therefore, GPU accelerates the speed of recognition. Beneson *et. al.*, employed GPU to increase the pedestrian detection speed among 20FPS and 100FPS.

The current pedestrian detection is often to distinguish the upright holistic pedestrian detection. But when the pedestrian flow is intensive the algorithm performance will be suddenly drop. To handle this problem, many researchers proposed some solutions. For example, Wang *et. al.*, proposed a HOG features-based algorithm for overlap detection method in SVM response. Zeng *et. al.*, applied PCA-based mutil-scales HOGLBP features to body's head and shoulder detection. Felzenszwald *et. al.*, proposed a body model-based LatSVM algorithm to pedestrian detection, which wined a good score in PASCAL VOC target recognition content.

III.METHODOLOGY

3.1 The System Requirements of Pedestrian Detection

We propose a novel spectral -based domain adaptation method to select data from multiple source domains and learn the target model for pedestrian detection. In many cases source domains refer to the labeled samples from generic scenes, while the target domain refers to the data from target scene. Here it is always assumed that the source (trained) and target

ISSN: 2278 – 7798

video samples are the same, while the source and target are different in spectral domain. Through spectral information changes between two successive frames can be monitored effectively.

In following paragraphs, we will briefly introduce the pedestrian detection requirement of video surveillance system.

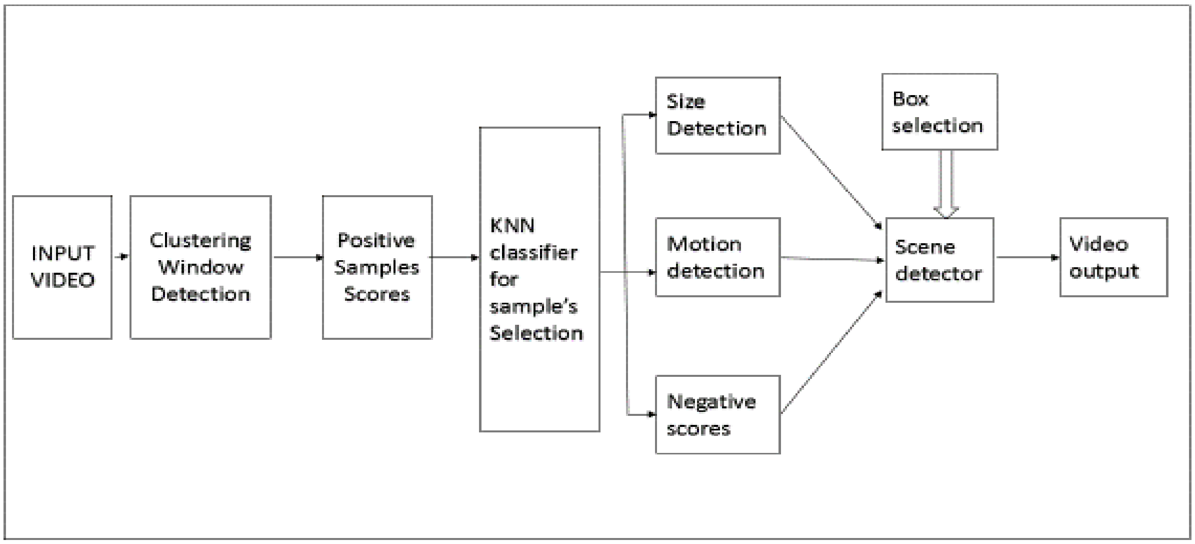


Fig.1 Block diagram of pedestrian detection system

3.2. Video Monitoring

The camera used for video monitoring plays a vital role in pedestrian detection. It is considered as the image source. However, camera only records the video stream of monitoring area. Thus, the video should be readable and replayed, and the camera parameters should be settable. To this purpose, multiple video cameras need to be connectable and they can be encode/decoded for replay. For the camera parameters setting, it should consist of two type parameters. The first parameter acts on camera which includes video format, resolution ratio and code rate. The other parameter acts on pedestrian detection, which includes the parameter about detection area.

3.3. Datasets.

We tested our detector on INRIA data set. INRIA containing 1805, 64×128 images of humans cropped from a varied set of personal photos. Fig. 5 shows some samples. The people are usually standing, but appear in any orientation and against a wide variety of background image including crowds. Many are bystanders taken from the image backgrounds, so there is no particular bias on their pose. The database is available from *http://lear.inrialpes.fr/data* for research purposes.

ISSN: 2278 – 7798

Fig.2 Sample images from INRIA dataset

3.4. Feature extraction

HOG features, proposed by Dalal and Triggs, and are adopted for our application. To extract HOG features, we firstly calculate the gradient orientations of the pixels in the cells. Then in each cell, we calculate a 9-dimensional histogram of gradient orientations is the features. Each block is represented by a 36-dimensional feature vector, which is normalized by dividing each feature bin with the vector module. Each sample is represented by 105 blocks (420 cells), corresponding to a 3780-dimensional HOG feature vector.

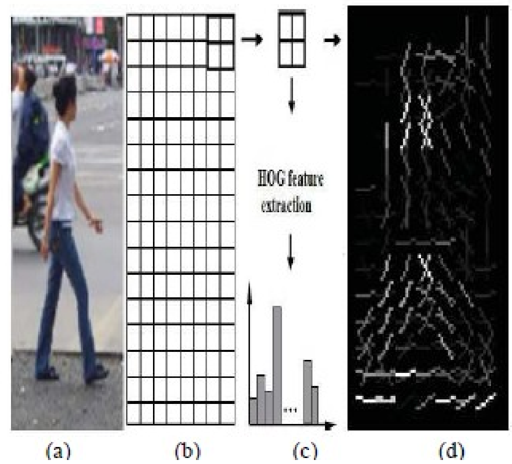


Fig.3. HOG feature extraction. (a) Human example. (b) HOG cells. (c) HOG feature extraction in block. (d) Visualization of HOG features multiplying with SVM norm vector.

3.5. SVM model with detection.

Given a set of training samples, from these samples train an Intersection Kernel SVM model with HOG features. The block diagram is shown in fig.4.Which summarizes the linear SVM training model. In this paper applied histogram equalization and median filtering of radius equal to 3 pixels on the test image as a pre-processing. The test image is repeatedly reduced in size by a factor of 1.1 resulting in an image pyramid. Sliding windows extracted from each layer of the pyramid. In each window, the hog features are extracted and tested with the IK-SVM to decide whether it is a human or not. Human detection block diagram is represented in fig.4

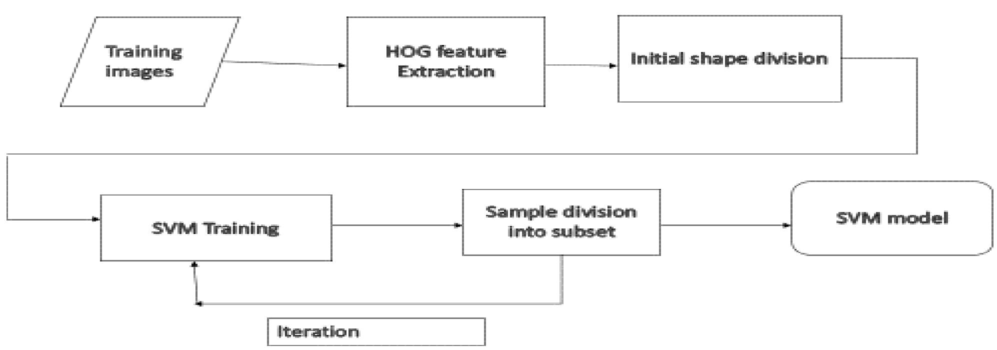


Fig.4 Human detection block diagram

ISSN: 2278 – 7798

3.6 Principal Component Analysis

Principal Component Analysis (PCA) is a standard technique for dimensionality reduction and has been applied to a broad class of computer vision problems, including feature selection, object recognition and face recognition. While PCA suffers from a number of shortcomings such as its implicit assumption of Gaussian distributions and its restriction to orthogonal linear combinations, it remains popular due to its simplicity. The idea of applying PCA to image patches is not novel.

3.7 k-NN Algorithm

The K nearest neighbor (kNN) classifier is an extension of the simple nearest neighbor (NN) classifier system. The nearest neighbor classifier works based on a simple nonparametric decision. Each query image *I*q is examined based on the distance of its features from the features of other images in the training database. The nearest neighbor is the image which has the minimum distance from the query image in the feature space. The distance between two features can be measured based on one of the distance functions such as, city block distance d1, and Euclidean distance d2 or cosine distance dcos

$$(d_1(x,y)) = \sum_{i=1}^{N} | X_i - Y_i |$$

$$(d_2(x,y)) = \sqrt{\sum_{i=1}^{N} | X_i - Y_i |}$$

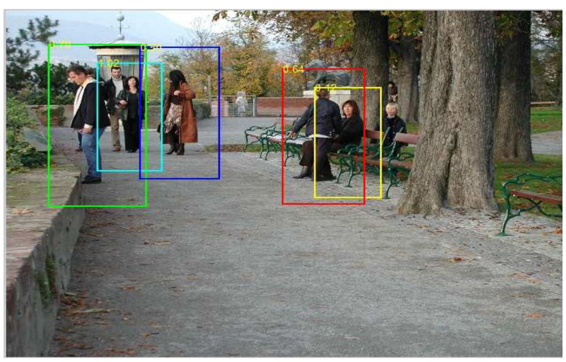
$$(d_1(x,y)) = 1 - \frac{\vec{x}.\vec{y}}{|x|.|y|}$$

K nearest neighbor algorithm uses K closest samples to the query image. Each of these samples belongs to a known class Ci. The query image Iq is categorized to the class CM which has the majority of occurrences among the K samples. The performance of the kNN classifiers highly related to value of the k, the number of the samples and their topological distribution over the feature space. Many approaches are introduced to improve the performance of the kNN systems using wavelet techniques, Cluster-Based Trees and Tolerant rough sets and so on. In this paper we show ensemble PCA based techniques can be used to improve the performance of the system.

IV.RESULTS AND ANALYSIS

The proposed human detection algorithm was tested on the INRIA dataset to illustrate its performance on a benchmark dataset. There are very few data samples available at very small scales of the input image (corresponding to very large pedestrians in the input images) to model the enclosing hyper sphere of the normalcy class. So the data samples from the smallest eight scales of each input image are grouped together before modelling the normalcy class, and then the anomalies (pedestrians) are obtained for these eight scales together. A subset of the INRIA dataset consisting of 230 images with sizes 480 × 640 and 640 × 480 are used to test the proposed algorithm. Fig.6 shows the final bounding boxes in the images representing the human detections. As shown in this figure, the proposed algorithm is capable of detecting human in urban and rural scenes. However, the number of false alarms appears to be higher in urban scenes, as exemplified in figures 4a, d, and f. This observation is due to the fact that some of the detection windows in the urban scenes have local spatial structures that are quite different from the majority of the image. So, these windows are deemed to be anomalies along with humans. At present, the detection rate of the proposed algorithm is around 74% at 1 false alarm per 4 images. The false alarm rate will drop sharply in rural scenes with less clutter.

ISSN: 2278 – 7798



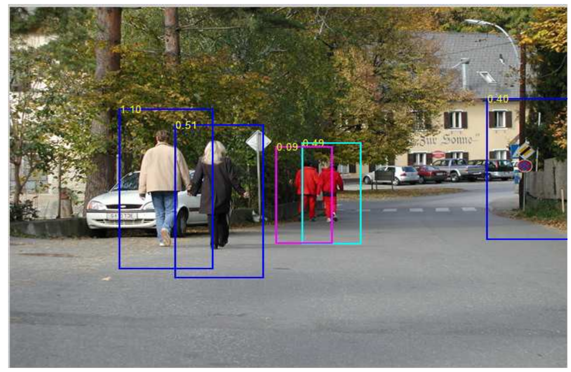


Fig.5 Pedestrian detected output obtained for set of input images from INRIA dataset.

V. CONCLUSION

In order to implement the fast pedestrian detection, this paper gives an overall design and its details. To meet real requirements, the fast pedestrian detection system consists of four part functional functions. It first introduces the overall design of system architecture and analyzes the relation between each module. For the detail function of each module, the target function is to be implemented under comparing previous studies in order to show the advantages and disadvantages. Based on this strategy, the detail plan of implementation for the fast pedestrian detection is discussed. For the future work, we will give a system prototype and execute some experiments to demonstrate the feasibility of our proposed.

References

[1] F. Porikli. Integral histogram: A fast way to extract histograms in Cartesian spaces. Proceedings of the International Conference on Computer Vision and Pattern Recognition. (2005) June 20-25, Vol.1:829-836; San Diego, CA, USA.

[2] N. Dalal and B. Triggs. Histograms of oriented gradients for human detection. Proceedings of the International Conference on Computer Vision and Pattern Recognition. (2005) June 20-25, Vol.1:886-893; San Diego, CA, USA.

[3] Q. Zhu, S. Avidan, M. Yeh, and K. Cheng. Fast human detection using a cascade of histograms of oriented gradients. Proceedings of the International Conference on Computer Vision and Pattern Recognition. (2006) June 20-25, Vol.2:1491-1498; New York, NY, USA.

[4] Ojala, T., Pietikäinen, M. and Harwood, D., A Comparative Study of Texture Measures with

ISSN: 2278 – 7798

Classification Based on Feature Distributions. Pattern Recognition.19 (3):51-59 (1996).
[5] C. Papageorgiou and T. Poggio. A trainable system for object detection. Intl. Journal of Computer Vision. 38(1):15-33(2000).
[6] P. A. Viola and M. J. Jones, “Robust real-time face det. Intl. Journal of Computer Vision. 57(2):137-154(2004).
[7] X. Wang, T. X. Han, and S. Yan.An HOG-LBP human detector with partial occlusion handling. Proceedings of the 2th International Conference on Computer Vision. (2009) September 27 - October 4, 32039, Kyoto, Japan.
[8] P. Doll´ ar, Z. Tu, P. Perona, and S. Belongie. Integral channel features. Proceedings of the British Machine Conference (2009) 7 September, 1-11 London, UK.
[9] S. Maji, A. Berg, and J. Malik. Classification using intersection kernel SVMs is efficient. Proceedings of the Conf. Computer Vision and Pattern Recognition (2008). 24-26 June, 1-11, Anchorage, Alaska, USA.
[10] O. Tuzel, F. Porikli, and P. Meer.Pedestrian detection via classification on riemannian manifolds. IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 30(10), pp.: 1713-1727(2008).
[11] Bourdev, L., Brandt, J.: Robust object detection via soft cascade. Proceedings of the International Conference on Computer Vision and Pattern Recognition. (2005) June 20-26, 236-243; San Diego, California, USA.
[12] Zhang, C., Viola, P.: Multiple-instance pruning for learning efficient cascade detectors. U.S. Patent US20090018980 A1, July 13 (2007).
[13] Dollar, P., Belongie, S., Perona, P.: The fastest pedestrian detector in the west. Proceedings of the British Machine Conference (2010) 31 August- 3 September, 1-11 Welsh, UK.
[14] P. Dollár, R. Appel and W. Kienzle. Crosstalk Cascades for Frame-Rate Pedestrian Detection. Proceedings of the European Conference on Computer Vision (2012) 7-13 October, 645-659, Florence, Italy

ISSN: 2278 – 7798