Available online at www.sciencedirect.com

ScienceDirect

Procedia Technology 25 (2016) 333 – 340

Procedia Technology

Global Colloquium in Recent Advancement and Effectual Researches in Engineering, Science and Technology (RAEREST 2016)

A Novel Approach to Detect Pedestrian from Still Images Using Random Subspace Method

Priya V V [*], Rekha P [b], Reshmi K C [c], Manoj kumar S [d], Indhulekha K [e]

[*&c]*M.TECH Student,Royal College of Engineering & Technology,Chiramanangad P.O, Akkikavu, Thrissur, Kerala,680604, vvpriya07@gmail.com*

[b] *Assisstant professor,Royal College of Engineering & Technology,Chiramanangad P.O, Akkikavu, Thrissur, Kerala,680604, rekha087@gmail.com, manoj99iyer@gmail.com , indusreesan@gmail.com*

Abstract

Pedestrian detection from still images is a terribly troublesome task. Human detection is the crucial part within the systems of humanistic image reclamation, visual scrutiny, pedestrian detection, and posture recognition, home automation, robot sensing. Detecting humans is a stimulating task due to major difficulties scrolling back from the wide variability of the target, like the form, wear or pose; and thereafter the external factors, like situation, illumination, and partial occlusions. This paper detects the humans using Random Subspace Method (RSM). The detection process is only in the still images no motion information is used. By using random subspace method detects the pedestrians. To implement these using mainly three types of datasets PobleSec, INRIA and Daimler Multicue dataset, additionally used linear SVM for classification.

Peer-review under responsibility of the organizing committee of RAEREST 2016

*Keywords:*Occlusion; Ensemble Classifier; Support Vector Machines; GLCM; DCT; DWT; Random Subspace Method;

1. Introduction

Pedestrian detection system is a vital and important task in computer vision. Computer vision is a field that includes methods for procuring, processing, scrutinizing, and understanding images. Pedestrian recognition is one among the foremost difficult issues within the field of computer vision. Pedestrian detection is an essential and vital task in any intelligent video surveillance system. Pedestrian detection is an on-going research area in computer vision in recent years [2]. Numerous approaches have been proposed. Depending on the movement and distance from the camera, there are mainly two methods that are used to detect pedestrians on image motion analysis and texture analysis respectively. Occlusion is the process that there is something want to see, but can't, due to some

Peer-review under responsibility of the organizing committee of RAEREST 2016
doi:10.1016/j.protcy.2016.08.115

property of sensor setup or some event. For example, in human tracking system the occlusion occurs if a human that tracking is unseen (occluded) by another object. Suppose two persons walking past one another, or an automotive that drives underneath a bridge. During this case once an object disappears and reappears once more [3].

Pedestrians ought to be recognized in extremely dynamic scenes since each pedestrian and camera area unit in motion that obscures pursuit and movement analysis. Motion analysis is employed in computer vision, image process, high-speed photography and machine vision. Texture analysis denotes to the characterization of regions in an image by their texture content. Holistic detection means object is considered as a whole and process it [7]. Part-based detection is part of the object is takes and process it separately and finally combined each parts. Patch-based detecting process, extracted local features are used to match against the codebook entries, and every match casts one vote for the pedestrian hypotheses detection process as patch wise. Motion-based detection selects the moved objects [20].

Detecting pedestrians in an image has tested to be a difficult task for several researchers as a result of the wide variability in diagnoses. Posture, clothing, size, background, and weather all are impactful on the presence of an image. The presence of pedestrians exhibits terribly high inconsistency since they wear totally different garments, carry totally different objects, and have a substantial vary of sizes particularly in terms of height[4][15]. Pedestrians should be recognised in outdoor urban circumstances, i.e., they have to be detected within the context of a encumbered background, urban areas are more multifaceted than highways under a wide range of illumination, and weather circumstances that vary the quality of the sensed information e.g., shadows and distinction within the colour spectrum[16]. Additionally pedestrians may be part occluded by common urban components, like parked vehicles or street furnishings. Furthermore, pedestrians seem at totally different view angles [8].

Most of the previous efforts in pedestrian classification assume full visibility of pedestrians in the scene. Component-based approaches which represent a pedestrian as an ensemble of parts can only alleviate this problem to some extent without prior knowledge. The key to successful detection of partially occluded pedestrians is additional information about which body parts are occluded. Background subtraction techniques usually find the foreground object from the image and then classify it into categories like human, animal, vehicle etc., based on shape, colour, or motion or other features. The direct detection methods the relative positions or geometric distances of various body parts are common to all humans, although the pixel values may vary because of the clothes or the illumination. The technique uses a structure known as the distance map. These techniques used part based method.

The proposed method mainly focused on holistic method, i.e., object is considered as a whole. The major benefits of the proposed approach is generic, it can be applicable to any class of objects. The second benefit is that the random subspace classifiers are trained in the original space, no further feature extraction is required. Next benefit is that the detection is done on monocular intensity images, unlike other methods for which stereo and motion information are mandatory and during training, we only require a subset of images with and without partial occlusion; other detection methods require delineation of the occluded area. Traffic monitoring is a key issue to be addressed in day-to-day's life. The major applications are in driver assistant systems, video surveillance system. The intelligent vehicle reduces the number of accidents between pedestrians and vehicles.

The primary objective of the proposed system is to detect the pedestrians from still images using Random Subspace Method. The Random Subspace Method (RSM) is used for handling occlusion. RSM is a good learning method. This method has various advantages over the other methods. It does not require manual labelling of body parts, and also it does not require additional data like stereo and motion. The main advantage of this method is that it can be extended to other object classes too. In this method the window is described by a block based feature vector which includes the features of all the blocks. The resulting feature vector is evaluated by the holistic classifier. If the result from the holistic classifier is not clear, then the occlusion inference process is applied. In the occlusion inference process, for each block a discrete label is obtained which determines whether it is part of the pedestrian or background. Then segmentation is applied to remove the spurious response and to obtain a spatially coherent region. After segmentation only the blocks which are having the same property are grouped together. If they are not having the same features their features are distributed everywhere depending upon the upper bound of the features.

To detect the human from still images first of all convert the image into blocks. In this method the window is described by a block based feature vector which includes the features of all the blocks. The resulting feature vector is evaluated by the holistic classifier. Then extract features using texture analysis by using the Random subspace method to select the features [9]. The linear SVM was used as the base classifier. The random subspace method is used to select each block randomly. Then apply probabilistic algorithm to select maximum feature from the selected portions. The probabilistic algorithm gives more accurate detection of humans [5-6].

The following sections, In Section 2 introduce proposed method. Section 3 presents implementation steps. Section 4 relates the result and discussions. Conclusion are summarised in section 5. Section 6 contains the papers, books, referred during the preparation of this paper.

2. Proposed Method

2.1. Proposed outline

The overview of the proposed system is shown in Fig.1. This system presents a general outline for detecting objects. In this system detect pedestrian using Random Subspace Method (RSM). The main advantage of RSM is that it is trained in the original space and no further feature extraction is required. In the random subspace method, classifiers are constructed in random subspaces of the data feature space. RSM is associate ensemble classifier that consists of many classifiers each functioning in a subspace of the original feature space, and outputs the category supported the outputs of those individual classifiers. Random subspace method has been used for decision trees random decision forests, linear classifiers, support vector machines, nearest neighbours and different styles of classifiers.

The RSM may benefit from using random subspaces for both constructing and aggregating the classifiers. When the number of training objects is relatively tiny compared with the data dimensionality, by constructing classifiers in random subspaces one may solve the tiny sample size problem. The subspace dimensionality is smaller than within the original feature space, whereas the number of training objects remains a similar. Therefore, the relative training sample size will increase. When data have many redundant features, one may acquire higher classifiers in random subspaces than within the original feature space .

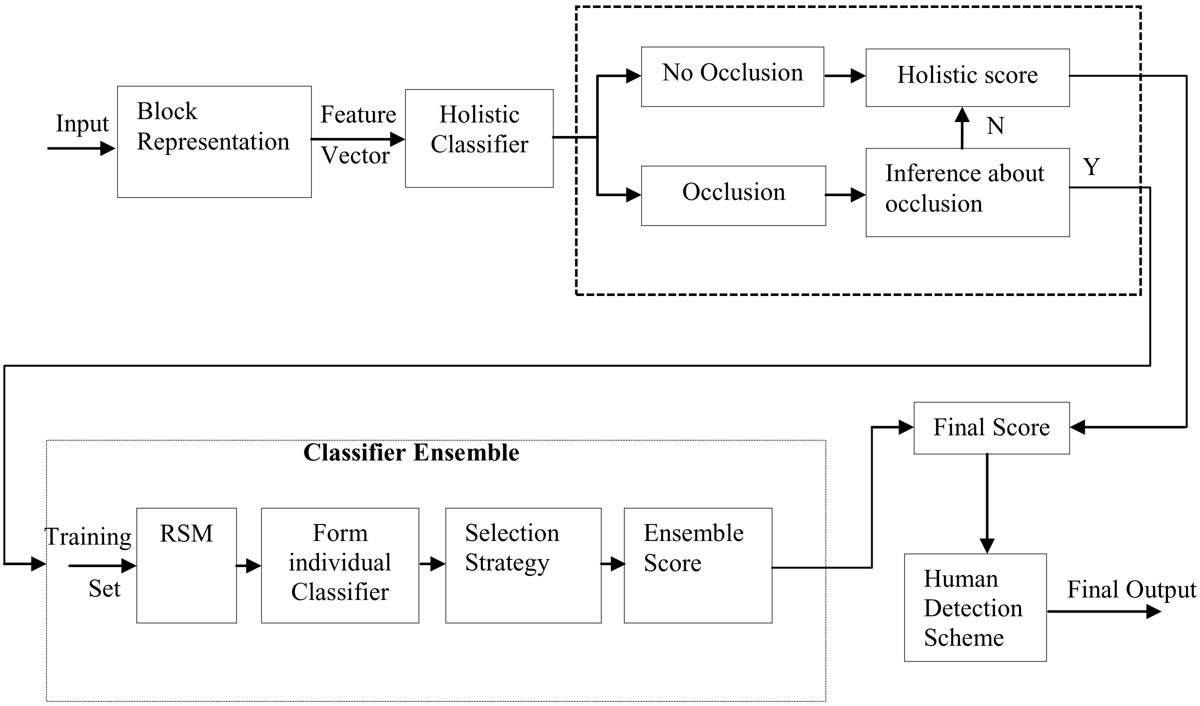

Fig. 1. Overview of proposed system

2.2. Occlusion handling using ensemble classifier

In general-purpose detection of occlusion regions is difficult and important because one-to-one correspondence of imaged scene points is needed for many tracking and video segmentation. If there is a partially occluded human figure in the frame sequence, we make use of a method similar to HOG-LBP [7][9][12]. The response from a classifier is perceived as ambiguous if it is near 0. When the output is ambiguous, an occlusion inference process is applied. This is based on the responses obtained from the features computed in each block [13].

The algorithm for the occlusion inference and the posterior reasoning is described in Occlusion Detection Algorithm. Javier Marin et al. [1] propose an ensemble of local classifiers using a Random Subspace Classifier which is primarily based on blocks. It involves finding the random subset of individual blocks until different subsets are found. After finding this random space this subspace values are each trained by individual classifiers [14] [19]. This process of classifying using individual classifiers is very costly and it is not efficient [16]. By using Algorithm1 detect the occlusion and Algorithm 2 refers to posterior reasoning and Occlusion Inference. In Algorithm3, refers to the feature extraction process with the help of RSM. Suppose E(x) is the ensemble and g_i is the local ensemble and w_i is the weight [1][17]. By using eqn (1) we can calculate the ensemble,

$$E(X) = \sum_{i \in s} w_i g_i(X) \tag{1}$$

Algorithm1: Occlusion Detection Algorithm
Input: Set of Blocks B_i
Output: Detect Occlusion
Procedure:

1. *Begin*
2. *for each i 1,….,n*
3. *Calculate $h_i(bl_i)$;*
4. *s_i = sign(h_i);*
5. *end S = {s_1,…..,s_n};*
6. *If S close to 1 No occlusion*
7. *else occlusion interference process*
8. *end*

Normally, partial occlusions can vary considerably in terms of shape and size. In the Occlusion detection process first of all we calculate the holistic score, and then find the number of signed position in the image. If S is close to '1' then found no occlusion. Otherwise found occlusion inference process.

Algorithm2: posterior reasoning and Occlusion interference
Input: B_1, . . . ,B_m
Output: Found partial occlusion
Procedure:

1. *Begin*
2. *for each occlusion interference object*
3. *N = Select N-possible classifier (with best score);*
4. *End*
5. *if score (N> occlusion threshold score) Object is Occluded.*
6. *else No Occlusion .*

Rather than using this classifiers method which uses a fuzzy based approach which is based on choosing the threshold for a particular dataset. Algorithm3 represents Random subspace classifiers pseudo-code [1].

Algorithm3: Random Subspace Classifiers.
Input: Training dataset DS = {(uj, vj)|1 ≤j≤ n}, Th
Output: f1, . . . ,fT
Procedure:

1. *Begin*
2. *I := {1, . . . ,m};*
3. *J := {Ø};*
4. *k := 1;*

5. *while* $k \leq Th$ *do*
6. *Randomly select a subset* $Jk \subset I$ *with* $Jk \neq \emptyset$;
7. *Given Jk generate the according* $(p1, \ldots ,pm)$;
   $(p'1, \ldots , p'm):=seg(p1, \ldots , pm)$;
8. *Obtain* $J'k$ *from* $(p'1, \ldots ,p'm)$;
9. *if* $| \Sigma p'I \neq m \wedge J'k \in J$ *then*
   Train fk in $Dk = \{(P'k(uj), vj)|1 \neq j \neq n\}$;
10. $J := J \cup \{J'k\}$;
11. $k := k + 1$;
12. *end*

3. Implementation details

3.1. Block representation of image

In the pre-processing stage, images could be segmented into adjacent regions based on texture properties of each region. Image segmentation is the process of splitting an image into multiple parts. This is typically used to identify objects or other significant information in digital images [18]. The goal of segmentation is to make straight forward and/or change the representation of an image into something that is more momentous and easier to analyze. Image segmentation is usually used to locate objects and boundaries in images. After segmentation process we get a set of segments that collectively cover the entire image, or a set of contours extracted from the image. Each of the pixels in a region is alike with respect to some features or computed property, such as color, intensity, or texture. Neighboring regions are significantly different with respect to the same characteristics. In this project set the 20% of image size.

3.2. Texture analysis and feature extraction

Major goals of texture analysis in computer vision area unit to know, model and process texture. To extract the features from image use transforms and textures. In transform mainly used wavelet transform and in texture analysis DCT and GLCM methods are used. In addition to this used area and orientation. Feature extraction involves reducing the number of resources needed to explain an oversized set of knowledge by using Grey Level Co-occurrence Matrix (GLCM). Examining the texture and consider the spatial relationship of pixels and form a matrix [10]. The Discrete Wavelet Transform (DWT) provides necessary information for scrutiny and amalgamation of original signal with relevant reduction in the computation time.

In DCT, the extraction of the most illuminating features with highly prejudiced ability to improve the classification accuracy and reduce complexity. In area based analysis to select the area find the region position values and compare it with the values. If the region position value is greater than 10000 pixels and if it is less than 100 pixel select it as edge points and neglect it. Other portions are processed. There are two types of orientation, horizontal and vertical. In this paper select the vertical orientation. If the orientation value is greater than 50 then select it as a human [20].

3.3. Support vector machine(SVM)

SVM may be a classification method that has evidenced to be terribly economical in such case of high dimensional information. It's been developed to sight a person's targeted in an exceedingly 128 × 64 single image. The SVM are for the most part used for the pattern recognition, the regression and also the density estimation. SVM may be a binary classifier supported supervised learning which provides higher performance than different classifiers. After occlusion detection each object has to be compared with the pre-defined data set to identify the presence of the human. Using SVM create a structure of datasets. In this paper we create a structure of features and comparing it with existing features. Using the SVM create a structure of features. The main application of using SVM classifier is that, we can increase the detection rate of human.

4. Results and discussions

The following section describes the result and performance analysis.

4.1. Results

The input images contain set of partial occluded and non-occluded persons. Fig. 2 shows the input and output images without using RSM. Fig. 3 shows the input images and output images using RSM respectively. The detection rate of human without using RSM is 60%. The detection rate of human after implementing RSM is 75.6%. By using RSM we can improve the detection rate.

Fig. 2 Input and output Images without using RSM

Fig. 3Input and Output Images using RSM

5. Performance evaluation

In this section, present several performance evaluation metrics that have been used for human detection and tracking. The basis for comparing the strengths and weaknesses of different human detection and tracking methods is to evaluate their results on a set of experiments with known ground-truth data using the same performance metrics. Within the performance analysis get 75.6% result. To evaluate the result consider five type of images and calculate the True Positive value, False Positive value, False negative value and PPV. Table 1 shows the performance evaluation. Fig 4 (a) shows the graph for the performance evaluation and Fig (b) shows the graph for PPV value.

Table1 Performance Evaluation with RSM

Type of Image	True Positive %	False Positive %	False Negative %	PPV
Type1	100	0	0	1
Type2	75	0	25	1
Type3	70	10	30	0.9
Type4	33	0	77	1
Type5	100	50	0	0.65

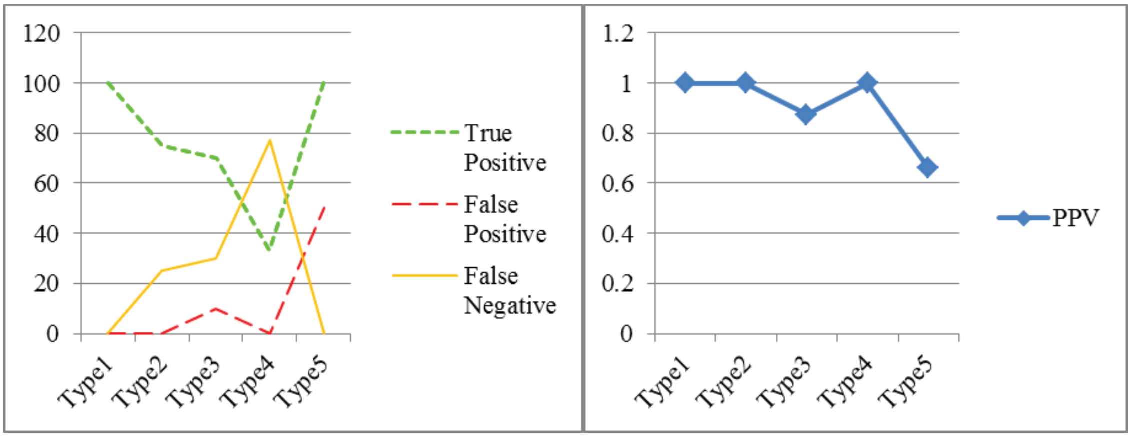

Fig. 4 (a) graph of performance evaluation;(b) graph of ppv

The human detection metrics that we used include:

- True Positive (TP)
- False Positive (FP)
- False Negative (FN)
- Positive Predictive Value(PPV)

- True Positive (TP):A true positive test result is one that detects the condition when the condition is present. True positive rate is the percentage of pedestrians detect from original image. In this paper we get the mean TP value is 75.6%.
- False Positive (FP): A false positive test result is one that detects the condition when the condition is absent. False positive rate is the percentage of detection of pedestrians by mistake. In this paper we get the mean FP value is 12%.
- False Negative (FN): The rate of False Negative is that the pedestrians are present in the image but it is not detected. In this paper we get the mean FN value is 26.4%.

- Positive Predictive Value (PPV): PPV is the value that predicted for detection of pedestrians. It can be calculated as True Positive value divided by the sum of True Positive value and False Positive value. PPV value is varies from 0 to 1.

6. Conclusion

Intelligent vehicles represent a key technology for minimizing variety of accidents between pedestrians and vehicles. Problems occurring throughout planning such a system should be overcome i.e. real time detection of fixing targets in out of doors situations that is uncontrolled. Pedestrian detection isn't straightforward task, consequently a far more than investigator operating for this method. In this paper, we gave a general approach for human detection from still images with the presence of partial occlusion. The tactic was supported a changed random mathematical space classifier ensemble. The tactics are often simply extended to alternative objects, and permits incorporating alternative block-based descriptors. The linear SVM was used because the base classifier. Evaluate this approach on three giant datasets, INRIA, Poblesec and Daimler Multicue dataset.

References

[1] Javier Marín, David V azquez, Antonio M. Lopez, Jaume Amores, and Ludmila I. Kuncheva M. Enzweiler and D. Gavrila, "Occlusion Handling via Random Subspace Classifiers for Human Detection," IEEE Trans. vol. 44, no. 03, pp. 342–354, March 2014.

[2] D. Ger´onimo, A. M. L´opez, A. D. 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.

[3] P. Doll´ar, 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.

[5] P. Viola and M. Jones, "Robust real-time face detection," Int. J. Comput. Vision, vol. 57, no. 2, pp. 137–154, Jul. 2004.

[6] N. Dalal and B. Triggs, "Histograms of oriented gradients for human detection," in Proc. CVPR, San Diego, CA, USA, 2005, pp. 886–893.

[7] X. Wang, T. Han, and S. Yan, "An HOG–LBP human detector with partial occlusion handling," in Proc. ICCV, Kyoto, Japan, 2009, pp. 32–39.

[8] S. Walk, N. Majer, K. Schindler, and B. Schiele, "New features and insights for pedestrian detection," in Proc. CVPR, San Francisco, CA, USA, 2010, pp. 1030–1037.

[9] Y. Pang, H. Yan, Y. Yuan, and K. Wang, "Robust CoHOG feature extraction in human-centered image/video management system," IEEE Syst. Man Cybern. B, Cybern., vol. 42, no. 2, pp. 458–468, Apr. 2012.

[10] T. Ojala, M. Pietik¨ainen, and D. Harwood, "A comparative study of texture measures with classification based on featured distributions," Pattern Recognit., vol. 29, no. 1, pp. 51–59, Jan. 1996.

[11] T. Ahonen, A. Hadid, and M. Pietik¨ainen, "Face description with local binary patterns: Application to face recognition," IEEE Trans. Pattern Anal. Mach. Intell., vol. 28, no. 12, pp. 2037–2041, Dec. 2006.

[12] Y. Yu, J. Zhang, Y. Huang, S. Zheng, W. Ren, K. Huang, and T. Tan, "Object detection by context and boosted HOG-LBP," in Proc. PASCAL Visual Object Challenge Workshop, Proc. Eur. Conf. Comput. Vision, Crete, Greece, 2011.

[13] D. Comaniciu and P. Meer, "Mean shift: A robust approach toward feature space analysis," IEEE Trans. Pattern Anal. Mach. Intell., vol. 24, no. 5, pp. 603–619, May. 2002.

[14] L. Pishchulin, A. Jain, C. Wojek, M. Andriluka, T. Thormaehlen, and B. Schiele, "Learning people detection models from few training samples," in Proc. CVPR, 2011, pp. 1473–1480.

[15] D. V´azquez, A. M. L´opez, D. Ponsa, and J. Marín, "Virtual worlds and active learning for human detection," in Proc. ICMI, Alicante, Spain, 2011, pp. 393–400.

[16] B. Kuneva, A. Torralba, and W. T. Freeman, "Evaluation of image features using a photorealistic virtual world," in Proc. ICCV, Barcelona, Spain, 2011, pp. 2282–2289.

[17] L. I. Kuncheva, Combining Pattern Classifiers: Methods and Algorithms. New York, NY, USA: Wiley-Interscience, 2004.

[18] T. Gao, B. Packer, and D. Koller, "A segmentation-aware object detection model with occlusion handling," in Proc. CVPR, Colorado Springs, CO, USA, 2011, pp. 1361–1368.

[19] T. K. Ho, "The random subspace method for constructing decision forests," IEEE Trans. Pattern Anal. Mach. Intell., vol. 20, no. 8, pp. 832–844, Aug. 1998.

[20] J. Marín, D. V´azquez, D. Ger´onimo, and A. M. L´opez, "Learning appearance in virtual scenarios for pedestrian detection," in Proc. CVPR, San Francisco, CA, USA, 2010, pp. 137–144.