Human Object Detection by HoG, HoB, HoC and BO Features

Sumati Malhotra
Computer Science
Department, Student,
Panipat Institute of
Engineering & Technology,
Smalkha

Shekhar Singh
Computer Science
Department, Assistant
Professor, Panipat Institute of
Engineering & Technology

S. C. Gupta, PhD
Computer Science
Department, H.O.D, Panipat
Institute of Engineering &
Technology, Smalkha

ABSTRACT

Human object detection in image or video is always a challenge in computer vision which is hurdle in development of automatic cars and robots since machine till now is not able to categorize the object on its own. We have discussed the issues in human object detection algorithms in this paper and suggested a new feature extraction approach with SVM classifier. We have cascaded a new features set using four different features which provides color, edge, bar information along with minimization of false detection. These are HoG, HoC, HoB and BO respectively. With these features set we are able to get a good accuracy rate then previous work.

General Terms

Human Object Detection .

Keywords

HoG, HoB, BO,SVM, Human Detection

1. INTRODUCTION

Object detection in an image is a challenging task, with many applications that has attracted lot of attention in recent years. Consider the case of personal digital content analysis, where typical content is images taken during a vacation, at a party or at some family occasion. Statistics show that even digital camera owners who use their cameras only occasionally can take as many as 10,000 photos in just 2-3 years, at which point it becomes tedious to manually search and locate these photos. Intelligent digital content management software that automatically adds tags to images to facilitate search is thus an important research goal. Most of the images taken are of people, so person detection will form an integral part of such tools. For commercial film and video contents, person detection will form an integral part of applications for video on demand and automatic content management. In conjunction with face and activity recognition, this may facilitate search for relevant contents or searches for few relevant sub-sequences. Figure 1 shows some images containing people from a collection of personal digital images.

Our work on human object detection is based on modifications in feature extraction part. Previously a major group of researchers have worked on HoG features with SVM classification algorithm. Some of them improved the classification algorithm part or some of them improved features extraction part. Although both parts are equally important for correct detection yet. Features extraction is leading in them since once exact and relevant features are extracted, classification accuracy will be increased. So we have focused on this area. Previously a new set of features was extracted which combined the HoG, Hoc,HoB features

(paper is sent to you). In another paper available [8] , a new features set along with HoG is used to improve the accuracy and a significant improvement is visible in the paper. So in our work we will create a cascading of four features set for training and testing purpose to detect human object. These will be HoG+BO (block orientation)+HoC+HoB. A block diagram for these is shown in Figure 2 .

Figure 1. Some images from a collection INRIA static person detection data set

2. PROPOSED WORK

The human object detection in our work is based on the features extraction form the test image and classification through Support Vector machine (SVM). Our work is divided into two main modules: one is focused on features extraction and other one is using SVM for classification from the whole test image.

In this a cascading of features has been used. Three different features are extracted and saved in database for INRIA dataset. Features including Histogram of gradient (HoG),

histogram of color (HoC), histogram of bar (HoB) and for reducing false detection of human object block orientation (BO) are used. These all are cascaded and used collectively. The proposed architecture of work is shown in Figure . 2
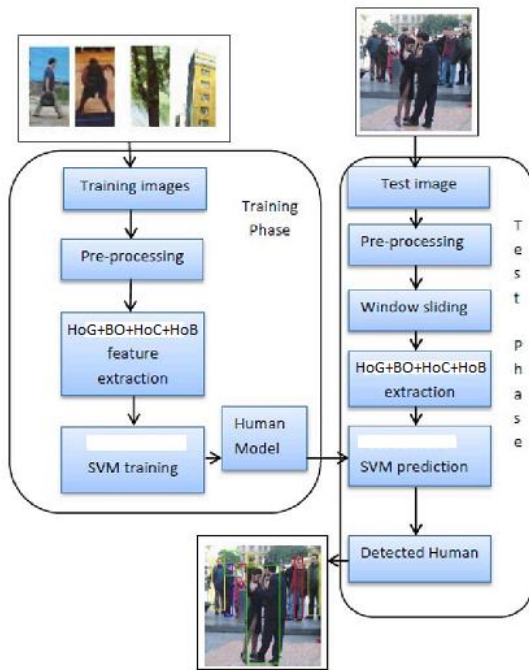

Figure 2 . Proposed methodology using four different features set

The test image is converted in to 648128 pixels and then features are extracted from it. To generate the database same size of images from INRIA dataset is taken. These four features set are discussed as:

2.1 Histogram of Gradient (HoG): This is also called first order gradient and related to edge information. The HOG features were originally introduced by Dalal & Triggs [7]. To obtain them, we need to compute the first order gradient at each pixel, aggregate the gradients to the corresponding cell, make a histogram on each cell, normalize the histogram along four directions, and finally concatenate all the normalized histograms to get the feature vector. However, we here use a modified HOG features suggested by Felzenszwal et al. [10], which mainly has two improvements from the original HOG: 1. The cell feature normalized along four directions are summed together, instead of concatenation, which reduces the dimensionality of feature vector to one-fourth; 2. A 4-dimensional texture feature vector is added for each cell .

2.2 HoC (histogram of color*):* it is also called zero order gradient. Though the three RGB channels are descriptors of red, green and blue, respectively, their tri-tuple is not a good representation for feature extraction, due to the mixture of pure color information and intensity information. To separate these two kinds of information, we convert RGB to Hue-Saturation-Intensity (HSI) color space. As the intensity information has already been used in HOG features (the computation of the first-order gradient), to avoid redundant information, we only retain the hue and saturation channels in HSI space, skipping the intensity channel. Fig. 3 is the schematic diagram of HSI color space. It can be seen that, without regard to intensity channel, the hue and saturation channels form a disk-shape space, where hue corresponds to angle and saturation corresponds to radius. If we map hue and

saturation to the orientation and magnitude of the first-order gradient in the HOG features, respectively, and follow the entire computation process of the HOG features, we can obtain the histograms of saturation over hue bins, which can describe the distribution of color in the image. These Histograms of Color (HoC) features are also cell-based, similar to the structure of the HOG features.
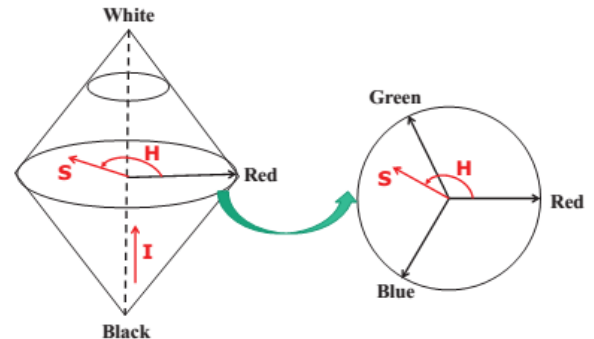

Figure 3. Schematic diagram of HSI color space [8]

2.3 Histogram of bar (HoB): these are second order gradient and related to bar information. The human object can be modeled like bar and blobs. So it may also be helpful in human detection. According to the definition of the kth-order gradients, the second-order gradients can be computed as follow:

$$r^* = \max_{\theta} \frac{\partial^2 I}{\partial^2 u}$$

$$\theta^* = arg \max_{\theta} \frac{\partial^2 I}{\partial^2 u}$$

where I is the intensity value of the input image, and u = (cos θ, sin θ) is the unit direction vector. By zeroing the derivative of the maximization item we can obtain $\theta^* = \frac{1}{2} \arctan(\frac{2.I_{xy}}{I_{xx}-I_{yy}}) r^* = I_{xx} cos^2\theta^* + 2I_{xy} cos\theta^* sin\theta^* + I_{yy} sin^2\theta^*$

After we get the second-order gradient at each pixel (x, y), we can follow the entire computation process of the HOG features, just with the first-order gradients replaced by second-order gradients, and then we can obtain the Histograms of Bar-shape (HoB) features, which can describe the distribution of bar-shapes in the image and also have similar structure with HOG features.

2.4 Block Orientation (BO): In HoG each image is divided into block size of 8*8 and for each such block 36 dimensional feature vector for first order gradient is obtained. Similarly in BO too all cells in the image are divided into up down and left right sub shells as shown in fig . 4
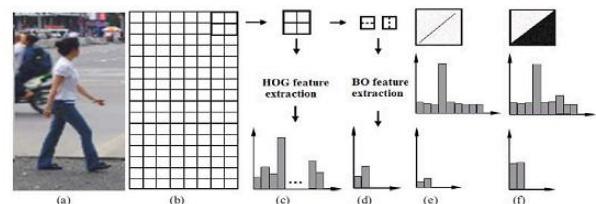

Figure 4. (a) Human example. (b) HOG and BO cells. (c)-(d) HOG and BO feature extraction (e) Stroke pattern with noise and its HOG and BO features. (f) Region pattern with noise and its HOG and BO features. [7]

The horizontal and vertical gradient are calculated by :

$$Bh = \max_{c \in \{R,G,B\}} \left| \sum_{x \in left subshell} I(x) - \sum_{x \in right subshell} I(x) \right|$$

$$Bv = \max_{c \in \{R,G,B\}} \left| \sum_{x \in up subshell} I(x) - \sum_{x down subshell} I(x) \right|$$

where $Ic(X)$ is one of the R,G and B color values at pixel X. The BO features are obtained by normalizing Bh and Bv.

These all four features are cascaded to form a complete features set for an image. We extract the features for all INRIA images which contains positive images with human object and negative images without human, and save those features in our database which will be used during SVM classification.

3. RESULTS

For classification purpose a sliding window approach is used in which the image region in window is used for testing purpose. The window size on image is chosen randomly which slides over whole image as shown in figure 5. It may happen that for chosen window size no human object is classified in any window because of size uncertainty of object, so keeping window size same, image size is reduced as shown in figure 5(b) and this process continues till complete human object or highest classification accuracy is not achieved.
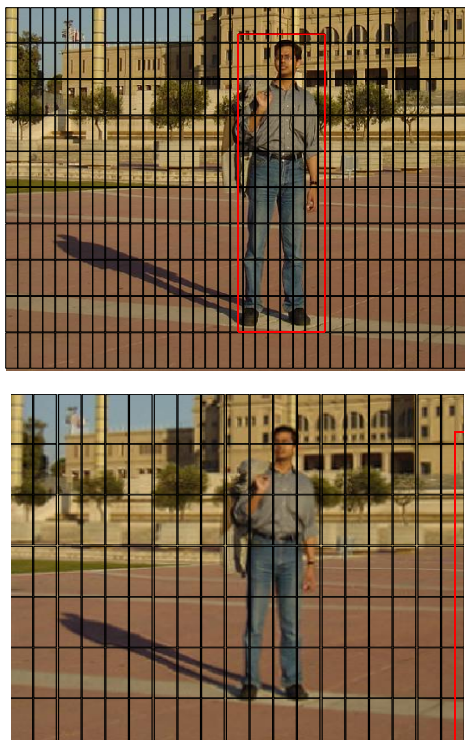
Figure 5 . (a) original image with many sliding windows of same size (b) reduced size image and sliding windows

We have tested our proposed features set on many INRIA images and compared the F-measure value with HoG+BO and HoG+HoC+HoB features set. A significant improvement

from previous work is achieved. Result of a test image is shown in figure 6 (a) and 6(b). The comparative table of F measure of four test images considered in this paper is shown in Table 1.

Table 1: Comparison of proposed scheme with previous hybrid features

F-measure->	Proposed Method	HoG+BO	HoG+HOC+HoB
Test Image1	2	1.2	0.3
Test Image2	2	1.8	0.8
Test Image3	2	1.2	0.2
Test Image 4	2	1.8	0.9

As is clear from the table we have proposed hybrid features is performing very well than previous scheme. HoG+BO is giving better results than third one since BO features are reducing the false recognition and so is our method.

The basis of F-measure and we have been able to achieve 40 % high f-measure than HoG+BO feature set and much more than HoG , HoB, HoC features .
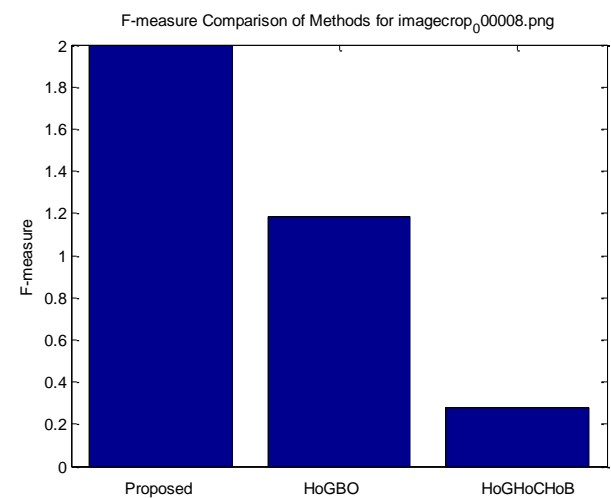
Figure 6(a) test image with detected human object in black rectangle (b) F-measure comparison with previous methods

Results have been successfully tested on multi object images too as shown in figure 7.
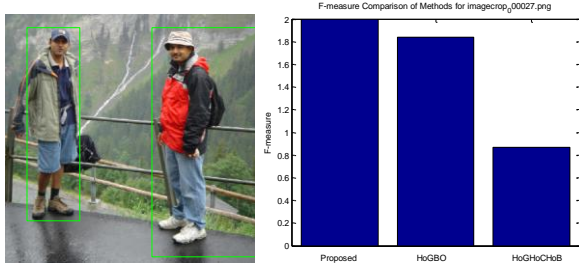
Figure 7. testing on multi object images

4. CONCLUSION

We have investigated the combination features/models for human detection and made two contributions. First ,we introduce the HOG features which consist of HOG features , color features (HoC) , bar-shape features (HoB) . Then these features are combined with one more with one more features set named Block Orientation (BO) to reduce the false

detection . A database of features set for INRIA images is generated and used in supervised SVM classification algorithm. results are compared on the basis of F-measure and we have been able to achieve 40% high f-measure than HoG+ BO feature set and much more than HoG, HoB, HoC features.

5. REFERENCES

[1] Hiyam Hatem, Zou Beiji and Raed Majeed," A Survey of Feature Base Methods for Human Face Detection", International Journal of Control and Automation Vol.8, No.5 (2015), pp.61 -78

[2] Bingquan Huo and Fengling Yin," Research on Novel Image Classification Algorithm based on Multi-Feature Extraction and Modified SVM Classifier", International Journal of Smart Home Vol. 9, No. 9 (2015), pp. 103-112.

[3] A. Satpathy, X. Jiang and H. L. Eng, "Human Detection by Quadratic Classification on Subspace of Extended Histogram of Gradients," in IEEE Transactions on Image Processing, vol. 23, no. 1, pp. 287-297, Jan. 2014.

[4] S. Varma and M. Sreeraj, "Object detection and classification in surveillance system," *Intelligent Computational Systems (RAICS), 2013 IEEE Recent Advances in*, Trivandrum, 2013, pp. 299-303.

[5] Jain Stoble B, Sreeraj M," Multi-posture human detection based on hybrid HOG-BO feature", Fifth International Conference on Advances in Computing and Communications, 2015

[6] Yunsheng Jiang and Jinwen Ma, "Combination features and models for human detection," 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Boston, MA, 2015, pp. 240-248.

[7] L. Spinello and R. Siegwart, "Human detection using multimodal and multidimensional features," Robotics and Automation, 2008. ICRA 2008. IEEE International Conference on, Pasadena, CA, 2008, pp. 3264-3269.

[8] N. Dalal and B. Triggs, "Histograms of oriented gradients for human detection," 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05), San Diego, CA, USA, 2005, pp. 886-893 vol. 1.
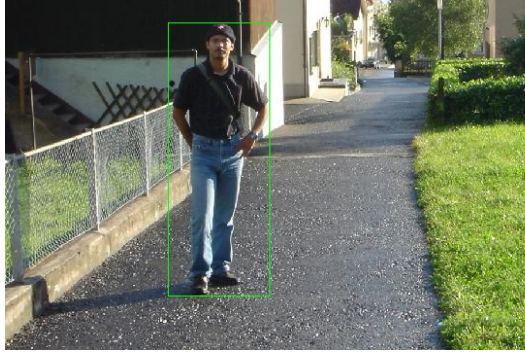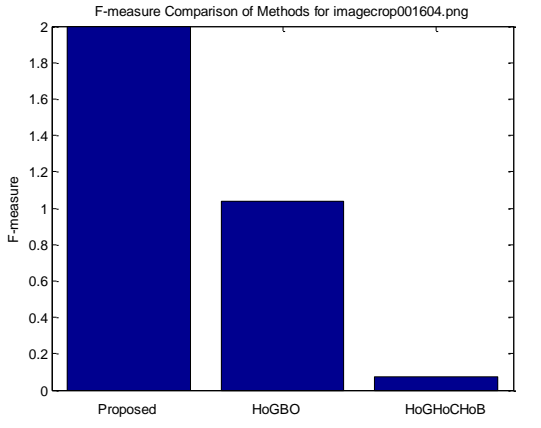
[9] M. Gupta, S. Kumar, N. Kejriwal, L. Behera and K. S. Venkatesh, "SURF-based human tracking algorithm for a human-following mobile robot," Image Processing Theory, Tools and Applications (IPTA), 2015 International Conference on, Orleans, 2015, pp. 111-116.
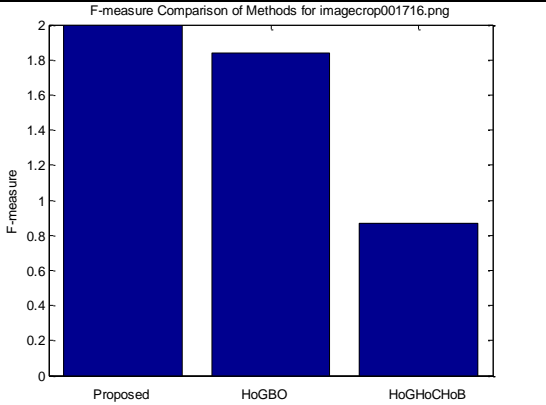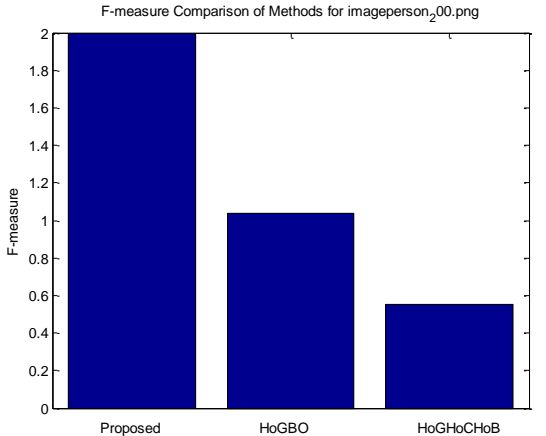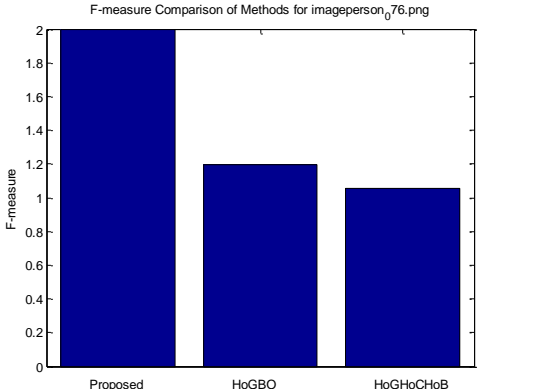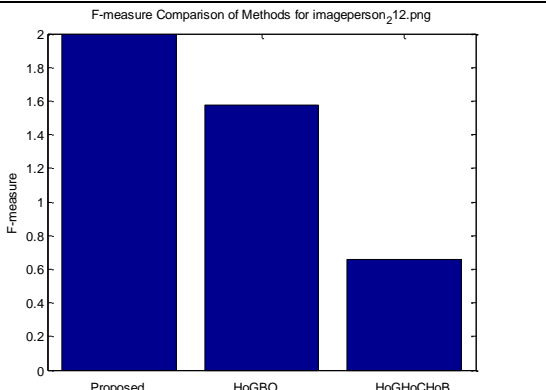
[10] Q. Ye, Z. Han, J. Jiao and J. Liu, "Human Detection in Images via Piecewise Linear Support Vector Machines," in IEEE Transactions on Image Processing, vol. 22, no. 2, pp. 778-789, Feb. 2013.

6. APPENDIX

A: Some more results of Test Images

Serial No	Detected Human Image	Comparison Bar plot
1		

2		
3		
4		
5		

6		
7		
8		
9		

10