Significant HOG-Histogram of Oriented Gradient Feature Selection for Human Detection

3 authors, including:

Muhammed J. A. Patwary
Shenzhen University
18 PUBLICATIONS   **122** CITATIONS

Subrina Akter
International Islamic University Chittagong
14 PUBLICATIONS   **51** CITATIONS

Some of the authors of this publication are also working on these related projects:

Fuzziness-based semi-supervised learning View project

Significant HOG-Histogram of Oriented Gradient Feature Selection for Human Detection

Muhammed Jamshed Alam Patwary
Assistant Professor
Department of Computer Science and Engineering
International Islamic University
Chittagong Bangladesh

Shahnaj Parvin
Assistant Professor
Department of Computer Science and Engineering
International Islamic University
Chittagong Bangladesh

Subrina Akter
Lecturer
Department of Computer Science and Engineering
International Islamic University
Chittagong Bangladesh

ABSTRACT
Detecting human efficiently is an important field of research and has many applications such as intelligent vehicle, robotics and video surveillance. Histogram of Oriented Gradient (HOG) is one of the eminent algorithms for human shape detection. HOG features are extracted from all location of a dense grid on an image region and use linear Support Vector Machine (SVM) to classify the combined features. Although HOG gives an accurate description of the contour of human body, it requires a large computational time. We studied the fundamental idea and consider features that have high percentage to contain edge. In this proposed method we used difference of Gaussian to obtain the edge percentage of each feature. Then a threshold is used to remove features with low edge percentage. Selected features then classified using linear SVM. Experiments on INRIA dataset demonstrate that the proposed method not only reduce the dimension of the HOG features but also outperforms.

General Terms
Pattern Recognition, Image Processing.

Keywords
Significant, HOG, Feature Selection, Human Detection.

1. INTRODUCTION
Human detection is one of the most important fields of image processing. It plays a vital role in our practical life. It is used in the field of security surveillance, visual surveillance system and automated driving assistance etc. There are several ways for detecting human from an image and videos. Nowadays most of the researchers are trying to make it efficient and reduce its complexity. Sometimes it takes long time to detect human because of computational complexity. There are different methods available for detecting human; some of them are quiet efficient and HOG (Histogram of Oriented Gradient) is one of them. HOG method used many features and all the features are classified by SVM. SVM computes all the features and it takes long time for computation. Computational time is proportional to the number of features. We also used the HOG method but we reduce the features so that this method takes less computational time, consequently this method provides an efficient and reliable result.

In the detection process, system percepts the highly probable area in an image where all instances of human appearance is located, and it is done by searching a particular pattern in all locations of an image. The aim of the research is to detect human from images in a significant way.

Before computing the HOG features, proposed method find out the cell percentage using the whole data set. It helps to select only the significant cells whose HOG features need to extract. Consequently, dimension of the HOG features are reduced. Here Section II discusses about the previous work and section III explains the overall methodology of proposed system. Section IV is about experimental results and discussion and finally section V concludes the paper.

2. REVIEW OF PREVIOUS WORK
Different types of algorithms are proposed for human detection. An Equation that has some identical characteristics, used to represent the shape of upper contour of the human body is proposed in [1]. The equation is named as omega equation and used in images acquired from video sequence. In [2] the edge occurrence with magnitude and orientation is considered for detecting human in images. They extract histogram of oriented gradient features and employ linear support vector machine to classify human and nonhuman objects. Selection of Histograms of Oriented Gradients Features for Pedestrian Detection is another approach for detecting human [3]. In the mentioned paper, Difference of Gaussian and Principal Component Analysis (PCA) to reduce the dimensionality of histogram of oriented gradient features are used. Statistical Learning Theory is discussed about SVM in [4]. A pedestrian detector based on polynomial SVM using rectified Haar wavelets as input descriptors is explained in [5] and with a part (sub-window) based variant in [6]. For part based detection, Mikolajczyk et al [7] used combinations of orientation position histograms with binary-threshold gradient magnitudes. Ronfard et al [8] introduced an articulated body detector by incorporating SVM based limb classifiers over 1st and 2nd order Gaussian filters in a dynamic programming framework. This is similar to efficient matching of pictorial structures [9] and Probabilistic methods for finding people [10]. Viola et al [11] build an efficient moving person detector. This research aims to reduce the HOG features while detecting human from images.

3. PROPOSED METHODOLOGY
3.1 Overall Process
This section gives an overview of our proposed method that aims to reduce the HOG features while detecting human from images. In an image all the cell does not contain the edge information. Blocks of an image become important when the percentage of at least one cell of this block is high enough. Then we extract the features from the selected blocks. So before extracting HOG features, an iterative analysis executed all over the training data set and finds all the blocks which are important. Then these sets of blocks are used as a suggestion while extracting HOG features. This will eliminate all the unnecessary blocks. Blocks are consisting of four cells and edge information which are not distributed in all cells with the

same rate for this reasons all the blocks are not necessary to classify in the SVM. If at least one cell has percentage than a reliable threshold then the respective block will participate in the extraction process of the HOG features. Collected features then classified using linear support vector machine. After this learning process for a particular image data it will provide an output that if the given image is Human or non human. Figure 1 summarizes the overall process of our proposed method.
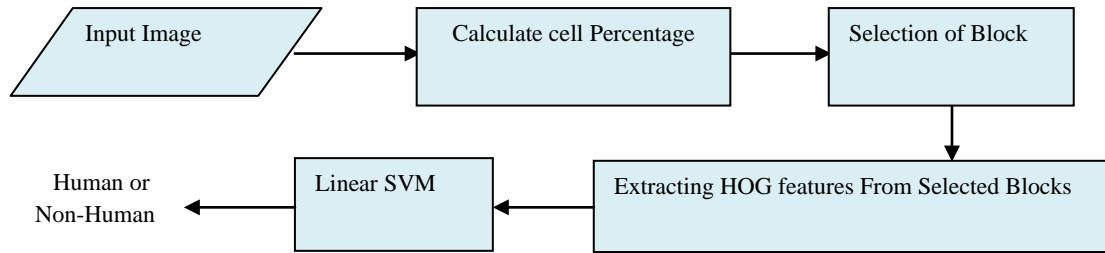

Fig 1: Overall View of the Process

3.2 Data Training Method

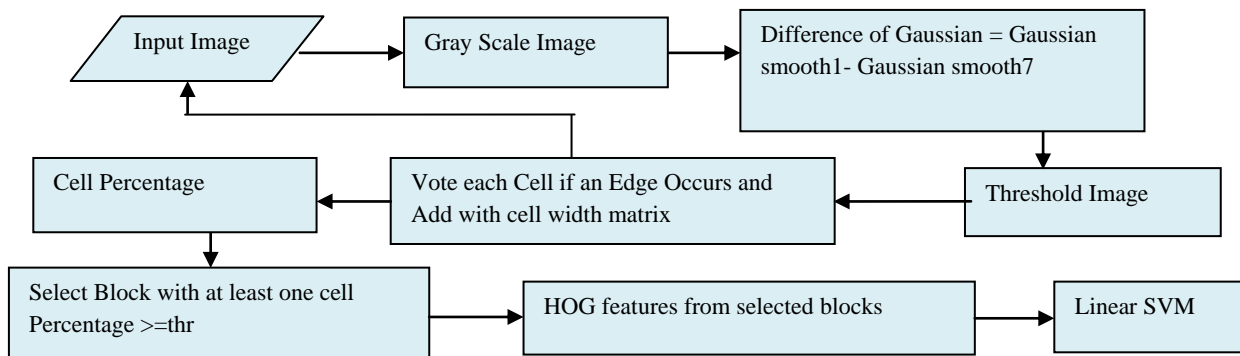The overall training process is described in figure: 2.

Fig 2: Training Process

3.3 Data Testing Method

Machine learning usually have to deal with big and uncertain data. We used pedestrian data set for train this method. To test this method we used 453 negative data and 1126 positive data.
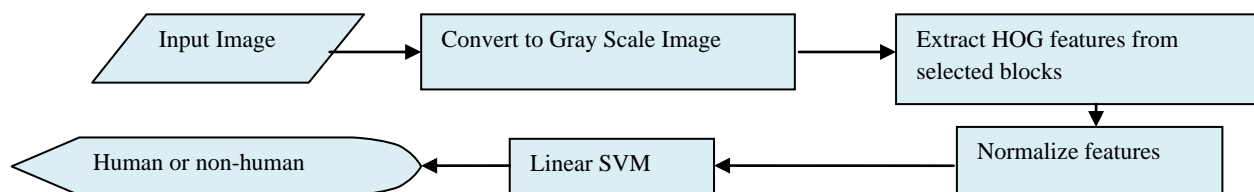

Fig 3: Testing Method

3.4 Basic Step of Preprocessing

a   b   c   d   e   f

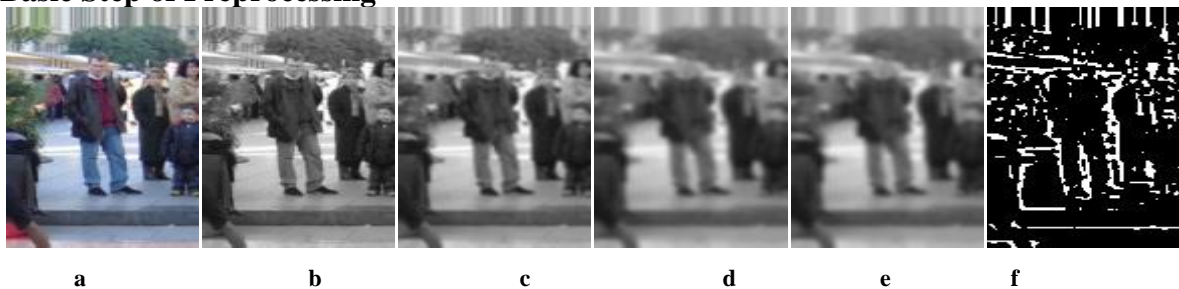**Fig 4: processing steps to find cell width (a) Main image (b)Gray scale image(c) Gaussian smooth 1 (d) Gaussian smooth 7 (e)Subtracted image (difference of Gaussian image) (f)Threshold Image**

3.5 Detection Process In Large Image

In this research the testing process is done by 64x128 image. To detect human in images that have dimension bigger than 64x128 need more engineering process. As our detection window is constant in size we have to resize our main image that means image which have width>65, height>128 we employ multi-scale version of the image. Here we create five levels of multi-scale image. Each image is 10% reduced in scale then the previous image. Each image is scaled keeping its aspect ratio preserve. Then slide the detection window with a dimension of 64x128 along with its width and height. To reduced the chances of sudden change sliding window move with 50% overlap of the previous window. Histogram of oriented gradient features from each window classify using SVM and denote the region where the chance is high getting human shape.

For multi-scale image processing we use the pyramid representation.Image is subject to repeated smoothing and sub sampling. Image pyramid representation shown in figure: 5.

3.6 Calculation Of Block Percentage

We evaluate the difference of Gaussian of the input image. The absolute output value of the difference of Gaussian filters at any pixel show high magnitude which has large variations. The difference of Gaussian method has sensitive at the extreme illuminated points. The pixel values will be accepted if the pixel values of the output DOG filter bigger then a threshold value. A cell width matrix with dimension stores the occurrence of the edge in a particular cell of the image. After iteration over all the training data set the cell width matrix represents each the total respective width of a cell. Using the cell width matrix calculate the percentage that represent the occurrences of edge in a particular cell. This percentage then used to select the blocks from which we have to extract the Histogram of Oriented Gradient Futures. This cell percentage ensures us the information that all the cells are not equally important to represent edge information of a human appearance. Figure: 6 show the cell percentage.

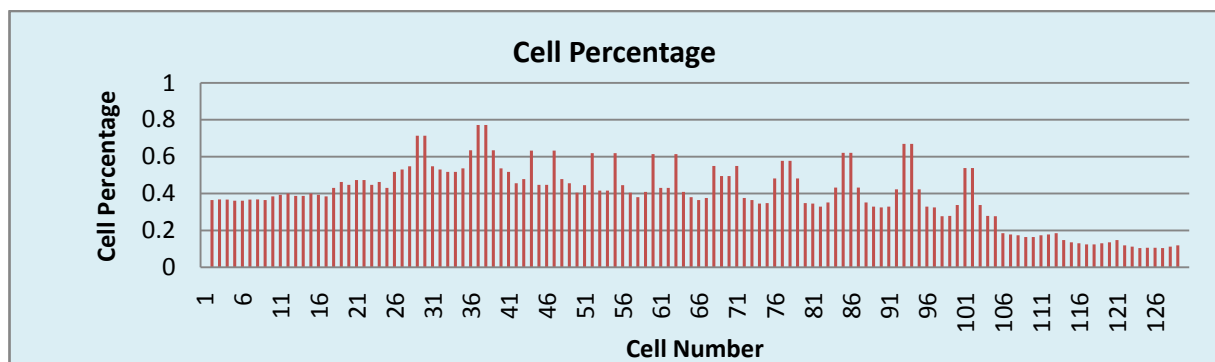**Fig 5: Image pyramid (a) main Image (b) reduced 10% (c) reduced 20% (d) 30% reduced (e) 40% reduced**

Fig 6: Cell percentage

3.7 Selection of Blocks

The cell percentage gives the system a guideline about which blocks need to be selected. If any of the cells in a block have the high chances to contain edge feature will be selected. So the percentage of containing edge in at least one cell of selected blocks will be high enough. So we can estimate that while detecting human appearance the selected block will have important edge information.

3.8 Rejected Block for Different Threshold

To find the total feature number we used this equation;

Feature = (bin*cell in a block*block along x axis*block along y axis)-(bin*cell in a block*eliminated blocks)

4. RESULTS AND DISCUSSION
4.1 Experimental Results

This research represents a technique that reduces the dimension of the HOG features. We execute some experiments that evaluate the performance of the procedure. The basic concept of this paper is to consider the basic probable area of an image and extract the important features. Extraction and training are mostly expensive steps in object

detection. Any procedure that reduces the feature dimensionality will be preferred. In this method dimensionality of the feature is reduced significantly but along with total success rate fall in a small rate than the existing system. But this can be acceptable as the visual impact is almost same as the main work. Selection of a block depends on cell percentage .Experimentally we found that blocks in the lower region are less important. We select threshold values that will define the lower acceptable percentage for a cell. If at least one cell has percentage greater than this threshold value will be accepted. Table [1] represents the actual number of features against different threshold value and figure: 7 shows the representation of the number of features with respect to threshold value. Table 2 shows the detail results.

Table 1: Results Sheet with Features and Threshold

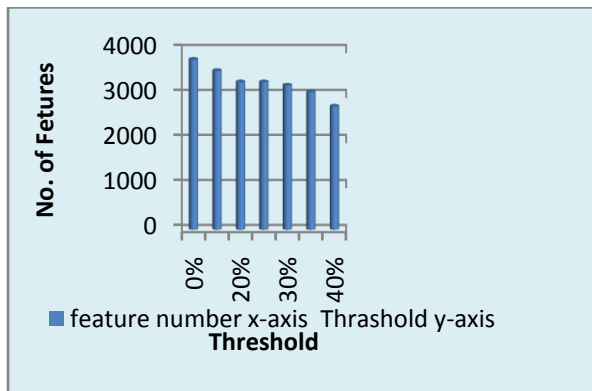| Threshold	Number of Feature
0%	3780
15%	3529
20%	3277
25%	3277
30%	3205
35%	3061
40%	2737

Fig 7 : Representation of Number of Features with Respect to Threshold Value

Table 2: Output result based on pedestrian data set (INRIA Data Set)

Threshold	Number of Feature	False Negative Rate	False Positive Rate	Success Rate
0%	3780	3.35%	1.45%	95.1868%
15%	3529	3.35%	1.58%	95.0601%
20%	3277	3.98%	1.83%	94.1735%
25%	3277	3.98%	1.83%	94.1735%
30%	3205	4.17%	1.70%	94.1102%
35%	3061	4.62%	2.15%	93.2235%
40%	2737	4.43%	2.08%	93.4768%

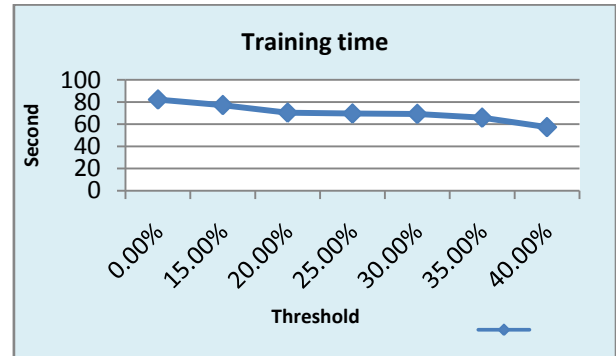## 4.2 Computational Time
A. Training Time:

Fig 8 : Computational Time While Training Process (Using 2000 positive and 12000 negative samples)
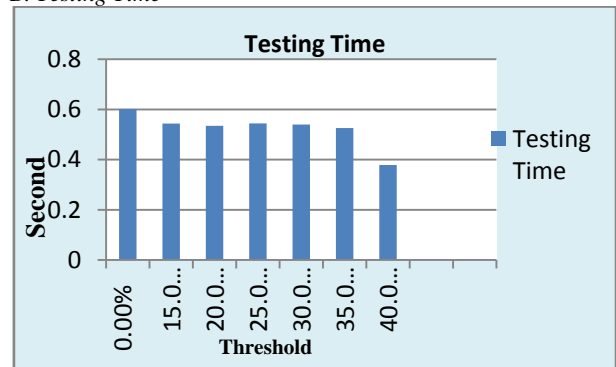
B. Testing Time

Fig 9 : Computational Time While Testing Process (Using 1132 positive and 400 negative samples)

4.3 Discussion
The experimental result of the proposed method is shown in table I that is needed much less HOG features than original HOG. HOG shows 95.18% success rate with using all features on the other hand our experiment shows success rate is 93.47%. The false negative rate in HOG is 3.35% when threshold is 0%. In our Experiment false negative rate is 4.43% shown in table [2]. On the other hand our computation time with fewer features is lesser than that of original HOG. So we can ignore the small variation of success rate as well as the false positive and false negative values. The proposed process in [2] reduces the dimension of HOG features. They collect HOG features from some specific key points and then utilize principal component analysis to reduce the dimensionality of the features. We suggest that after finding the key points using deference of Gaussian method and then selecting the most probable blocks, the principle component analysis could be utilized to reduce the dimensionality of the features. This process will help to increase the success rate.

5. CONCLUSION & FUTURE WORK
Detecting human is the most challenging task in the field of image processing. Human detection is the process of identifying human appearance in an image. Proposed method eliminates those blocks which have less significance in the image for extracting the histogram of oriented gradient features. We used the INRIA pedestrian data set for training. We conclude that HOG features can be reduced by considering useful features that contain edge information. And our proposed method has the ability to reduce the features.

Further optimization in this research can be done in many ways. We will try to speed up the detections rate and improve this method by reducing computational time so that detecting human from live video stream become easier. This technique used fewer feature dimensions then the existing system to compute but while testing the system all over success rate decrease in a tolerant range respect to the main system. This is a result due to the rate of false positive rate. We will try to reduce this false positive rate. In future we will try to increase the success rate.

6. REFERENCES

[1] Subra M., Karen D.:A novel Equation based classifier for detecting human in Images. In International Journal of Computer Applications (0975 – 8887) Volume 72– No.6, May 2013

[2] Dalal, N., Triggs, B.: Histograms of Oriented Gradients for Human Detection.In:IEEE Conference on Computer Vision and Pattern Recognition (CVPR) (2005)

[3] Takuya K., Akinori H., and Takio K.: selection of Histogram of oriented gradients. In M. Ishikawa et al. (Eds.): ICONIP 2007, Part II, LNCS 4985, pp. 598–607, 2008.
c Springer-Verlag Berlin Heidelberg 2008

[4] Vapnik, V.N.: Statistical Learning Theory. John Wiley and Sons, Chichester (1998)

[5] C. Papageorgiou and T. Poggio. A trainable system for object detection. IJCV, 38(1):15–33, 2000

[6] A. Mohan, C. Papageorgiou, and T. Poggio. Example-based object detection in images by components. PAMI, 23(4):349–361, April 2001

[7] Mikolajczyk, K., Schmid, C.: A performance evaluation of local descriptors. In: Proc. of Computer Vision and Pattern Recognition (2003)

[8] R. Ronfard, C. Schmid, and B. Triggs. Learning to parse pictures of people. The 7th ECCV, Copenhagen, Denmark, volume IV, pages 700–714, 2002.

[9] P. Felzenszwalb and D. Huttenlocher. Efficient matching of pictorial structures. CVPR, Hilton Head Island, South Carolina, USA, pages 66–75, 2000.

[10] S. Ioffe and D. A. Forsyth. Probabilistic methods for finding people. IJCV, 43(1):45–68, 2001.

[11] Viola, P., Jones, M.J., Snow, D.: Detecting pedestrians using patterns of motion and appearance. In: Proc of the 9th International Conf. of Computer Vision, Nice, vol. 1, pp. 734–741(2003).

.