Optimized Scale-Invariant Hog Descriptors for Tobacco Plant Detection

[1]THIRTHE GOWDA M. T.[1]*, [2]CHANDRIKA J.
[1]Department of Computer Science and Engg, Government Engineering College, Hassan, Karnataka, INDIA-573202
[2]Department of Computer Science and Engg, Malnad College of Engineering, Hassan, Karnataka, INDIA-573202

Abstract: The histogram of gradient (HOG) descriptor is being employed in this research work to demonstrate the technique of scale variant to identify the plant in surveillance videos. In few scenarios, the discrepancies in the histogram of gradient descriptors along with scale as well as variation in illumination are considered as one of the major hindrances. This research work introduces a unique SIO-HOG descriptor that is approximated to be scale-invariant. With the help of the footage that is captured from the tobacco plant identification process, the system can integrate adoptive bin selections as well as sample resizing. Further, this research work explores the impact of a PCA transform that is based on the process of feature selection on the performance of overall recognition and thereby considering finite scale range, adoptive orientation binning in non-overlapping descriptors, as well as finite scale range are all essential for a high detection rate. The feature vector of HOG over a complete search window is computationally intensive. However, suitable frameworks for classification can be developed by maintaining a precise range of attributes with finite Euclidean distance. Experimental results prove that the proposed approach for detecting tobacco from other weeds has resulted in an improved detection rate. And finally, the robustness of the complete plant detection system was evaluated on a video sequence with different non- linearity's that is quite common in a real-world environment and its performance metrics are evaluated.

Keywords: scale-invariance, HOG descriptor, tobacco detection, PCA transform.

Received: February 20, 2021. Revised: July 5, 2021. Accepted: July 11, 2021. Published: July 23, 2021.

1. Introduction

In any crop management, with the increase in the area under cultivation, the diversity and type of weeds will also increase. These weeds are not only affecting the major crop growth but also result in deterioration of the overall crop yield. The exclusion of weeds can be carried out manually or by using herbicides. However, following any of these traditions in the whole of the cultivated land will lead to soil poisoning and unnecessary costs. To avoid these problems highly précised agriculture technology needs to be applied. Among the various methodologies investigated vision-based systems are widely preferred in many real-time crop monitoring applications. In general, detecting plants in vision input is a computationally intensive task since a vast amount of pixels need to be processed and there are numerous challenges due to various issues such as occlusions, shading, and illuminations, etc. Furthermore, owing to its variable appearance, illumination variations, as well as cluttered background, the vision inputs necessarily entail a greater number of features to be approximated. SIFT descriptors [2] are approximated on uniformly spaced cells to resolve all of these problems as well as implement robust feature sets such as edge orientation histograms [1]. Most of the detection process focuses on images with improved spatial conditions. But for the real-time plant detection process, it is essential to consider real environment conditions and non linearity's over camera video sequences. And the machine vision systems for most realm videos are providing very limited real-time detection properties. Moreover faces several practical difficulties

to simultaneously identify one particular crop from different types of plants.
This research work is divided into the following sections. The summary of various object detection techniques are presented in Section 2, a comprehensive description of the HOG technique is demonstrated in Section 3, a detailed analysis, as well as experimental validation of the proposed tobacco plant detection process, is presented in Sections 4 and 5 respectively, finally, the conclusion of the research work is presented in Section 6.

2. Previous Work

The recent advancements in image processing techniques as well as the growing availability of wide sets of features to encode local data of an object of interest gives way for several improvements for the object detection process. Previous work in plant detection systems has follows two methodologies one is to locate plant regions using trajectory measures or detect based on gradient vectors and their direction. A hybrid artificial neural network-Ant Colony that is based on vision is designed to categorize the potato plants from various types of weed [3]. To approximate the feature vector for every object part, the orientation histograms along with a gradient threshold are integrated and are employed in [4]. Histogram of gradient features has proven its discrimination for the various object recognition and detection. HOG feature attributes describe the spatial appearance of the object using vector coordinates and also exploit its spatial correlation relationship with other regions for object detection. Even though HOG has been used in various techniques of object recognition, it is unable to identify objects in various postures. The above drawback can be overcome by making use of multi-class SVMs as well as discriminatively trained models [5]. Mirror invariant descriptors [6] are designed as per the magnitude under flip as well as rotation to maintain uniqueness and thereby retaining the necessary properties of scale-invariant. Some of the local feature sets like Continuous wavelets [7], as well as SIFT [8,], are used as key point-based detection techniques in addition to the scale-invariant that is based on the context of shape.
3. HOG descriptor

HOG are high-level features and are approximated from decomposed patch image gradients. To analyze the identification process of tobacco plants along with finite discrimination, a Histogram of gradient feature vector which is developed from every frame is been utilized. The framework for optimizing the identification of the plant is been developed using feature vectors. Though it has been used in most human detection models the normalization is always required to discriminate the tobacco plant from weeds with appropriate gradient orientations. The HOG descriptor is primarily analyzed using the following two architectures and they are rectangular HOG with square spatial cells as well as a circular histogram of gradient partitioning cells in a log-polar manner.
The proposed system is designed to extract the optimized feature subsets that can avoid a large number of redundant features of the original HOG and also removed some insignificant features for improved detection efficiency. This allows human detection robust to non-linearity's and also reduces the calculation time observably without a decrease in the accuracy.
The major contributions of the SIO-HOG system are as follows:

- This method can accommodate a wide range of illumination changes with a consistent detection rate since scale-invariant HOG descriptors from different coordinates are used for object recognition.
- The feature descriptors are invariant over any image transformations and scaling.
- Variance-driven feature subsets give a better detection performance.

3. Spatial / Orientation Binning

Particularly, the HOG descriptors work efficiently on a detection window with 8x8 pixel cells, and each of the pixel cells is composed of decomposed overlapping spatial blocks [12]. Here input frames are decomposed into various rectangular blocks as well as the orientation bins are formulated as evenly spaced vectors with the angle of deviations from 0°–180° with a step size of 20° to 40°. Further, to minimize the spatial redundancy as well as to normalize the final feature sets spatial windows are divided into non-overlapping spatial blocks. During the computation of the gradient vector from each block, the total of 64 gradient vectors is generated by representing each pixel in the given window and formulate the 9-bin histogram. There exist 20 degrees for every bin because the Histogram varies from 0 to 180 degrees.

4. Tobacco Plant Detection

The general framework of the tobacco plant detection system is shown in Figure 1. Initially, the concept of macroblock conversion is taken into account, and further preceded by the extraction process of HOG and is divided into two phases: intensity varies depending on the selection of bin as well variations in the appearance depending on re sampling. changes. The process of design is structured in the following manner: Initially, the pre-computing stage is initiated, and this phase produces bins as well as sample sizes such as rectangular and angular bins. The HOG extracted features are given as input to the subsequent phase, and the selection of features is achieved over the extracted feature sets and thereby maintaining feature vectors with high variance. It ensures the presence of the most significant feature and discards the redundant sets. At last, the output of the classification system is given as input to the bounding boxes on object detected windows of the specified size.

```
Algorithm: SIO HOG tobacco plant detection
TH_limit  -  Threshold level
HOG       -  Histogram of gradient computation
B0        -  Macroblock size
SIO – HOG (Input Frame, tobacco plant detection)
/Input –Frame /Output –object detection spatial region
for i=1: N number of frames
 for P=1:Bo:cell
   for Q=1:Bo:cell                    -- B0 Orientation Bin
     Vector(i,:) = HOG(i,Bo)
   end
end
    Mout = Vector [Matching Vector (i) - Vector (i-1)]
 if (Mout <  TH_limit && it_count < it_limit)
    Bo = down_sample(Bo);
    it_count = it_count +1;
 else
    Bo  = Bo;
    it_count = it_count;
 end
end
```

4.1 Pre-computaion stage

For reliable object detection over poor illumination as well as rapid scale changes, we make use of adoptive bin selector as well as sampling size in the HOG generation stage as follows:

• *Finite feature subset:* The features of HOG are very simple and can be efficiently approximated from the image patches.

• *Bin size:* The magnitudes of HOG feature vectors are proportional to the pixel values in the input image. As a result, variation in the intensity level will affect the HOG description. As the number of bins increases, the feature vector becomes more sensitive.

4.2 Feature selection stage

To improve the plant detection rate with the least possible false detection the HOG feature vectors used for detecting the various plants are normalized by analyzing its variance level. Principal component analyses (PCA) are utilized to choose the dominant as well as potentially effective features, and further, a hierarchical correlation matrix, as well as variance computation, are employed to achieve the task. This often aids in the detection of very small objects, and also helps in detecting a large variety of postures as well as appearances.

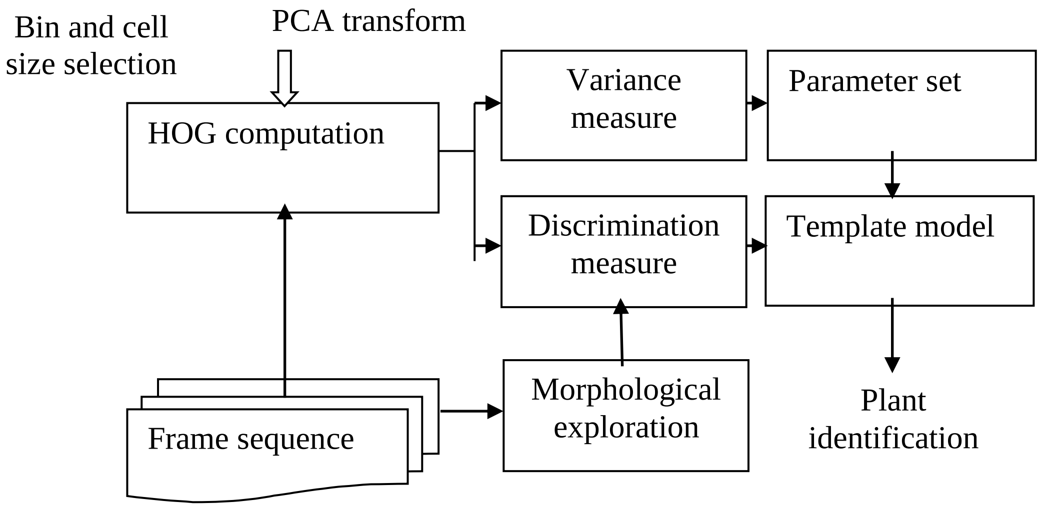

Figure 1. The general architecture of the plant detection system

4.3 Plant detection

Input frame I is first partitioned into multiple non-overlapping 8 × 8 patches Pi, i = 1,…,N . For each patch, a 9-bin HOG descriptor with the quantization angles j × 20°, j = 0...8 is used to represent its local spatial information. However, the HOG descriptor lacks the discrimination properties when the plant leaf spatial patterns are correlated with weeds, and thus shape feature attributes are incorporated with its morphological characteristics. Figure 4 shows that the large cell size of an object contains peak bin angles and its corresponding HOG approximates the gradient direction.

The local similarity measurement for $\{X_i,X_j\}$ on the HOG (r, θ) space P:
$P(r,\theta)=P(r,\theta)+Sim(X_i,X_j)$

This CHOG distribution at each coordinate point (x, y, q) is formulated as a histogram with M = 9 bins,

FCHOG(x, y, q) = [v1(x, y, q), v2(x,y,q) . . . , vM(x, y, q)] .

Where v1(x, y, t) $=\sqrt{lv/|lv|}+e$ denotes the normalized HOG descriptor, e- denotes the small constant for the normalization factor. While v2(x, y, t) . . . vM(x, y, t) respectively corresponds to leftward, rightward, upward, downward orientation models.

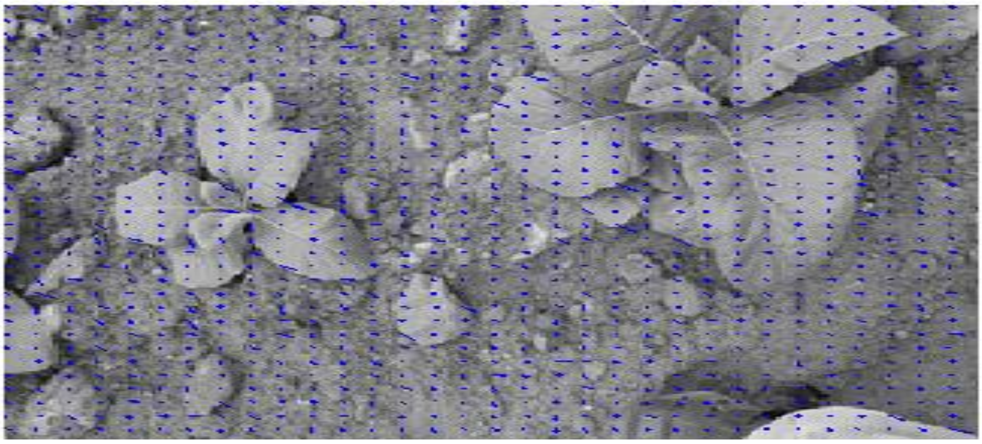

Figure 2 HOG of the target model

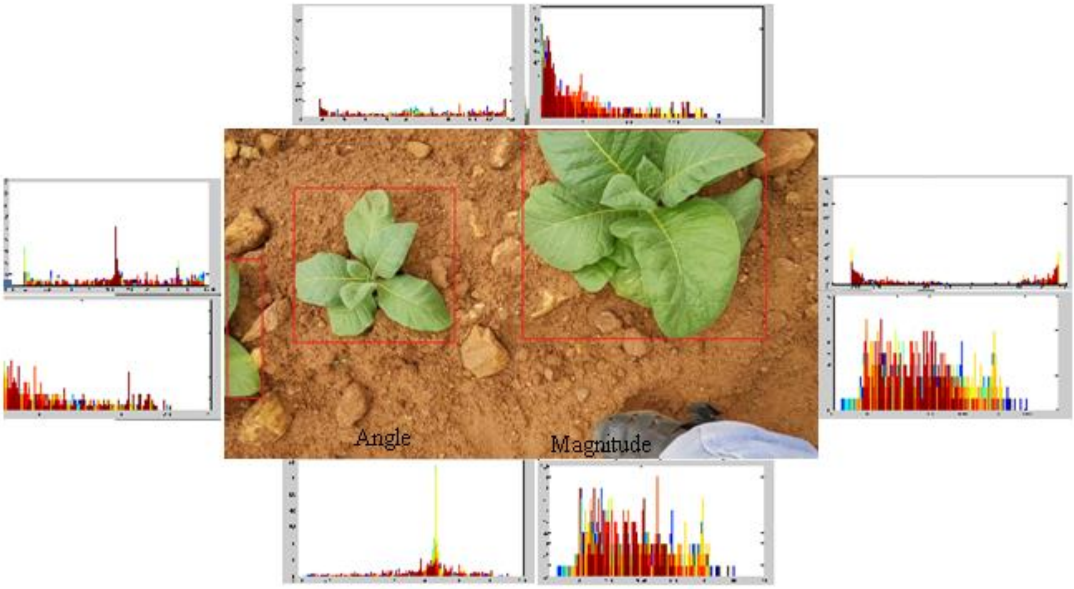

Figure 3 Angle and magnitude plot of CHOG model

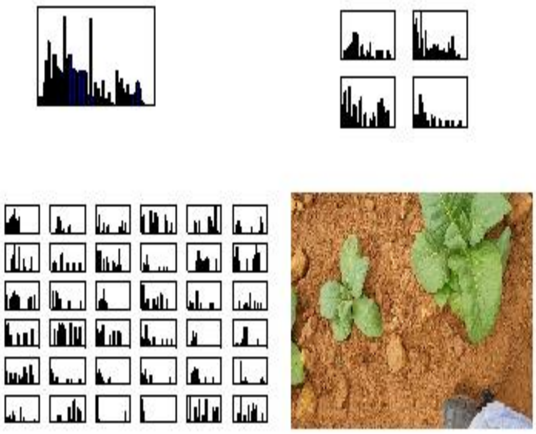

Figure 4 Histogram bins of variable cell size of an object

The extraction of HOG feature vectors is carried out over a 64x128 pixel search window with a cell size of 8x8 and thereby enabling the capturing of detailed local features. The dataset of 10 positive samples was collected, as well as the trade-off between window sizes and detection rate was examined. During plant detection, some of the plant samples are appeared as small; including weeds, and the HOG descriptor of these limited windows offers only minor details. Thus, it is also found that plants with the least samples are not influencing the overall detection rate performance.

4.4 Feature extraction

According to the histogram gradient, the plant identification templates are set up specifically from non-overlapping image patches in the search window. The gradient magnitude is defined in the specified direction and for every size of the cell, the histogram values are described. As previously mentioned, histograms with 9 bins that are ranging from 0 ◦ to 180◦ appear to be more preferable for the process of object detection. Since the histogram bins are successfully approximated and patches of the image are rectangular, the HOG feature vector can be constructed as follows: approximation of integrated samples is carried out with the help of non-overlapping blocks to extract the gradients and thereby saves time during the process of object detection.

5. Experimental Results

The performance analysis of object detection using a HOG descriptor that is based on gradient along with an adoptive bin selection technique is carried out with the help of the KNN classifier. To choose the class with the highest probability there exists no prior information about the objects. When compared to the conventional HOG model, the proposed plant detection rate of the scale-invariant Histogram of gradient shows enhanced performance, and any substantial variations in the intensity level that is approximated before the computation of HOG gradients significantly affect the true positive samples and it also demonstrated that the fine-scale consists of the majority of the available spatial orientation data and thereby helps in minimizing the illumination sensitivity as well as scale variations. The approximation of HOG is carried out at the finest image scale. Further, the gradients are rectified for adoptive orientation binning, as well as it is made invariant to the blurred spatial object region.

The implemented algorithm's detection performance is evaluated using the following metrics: sensitivity, specificity, as well as accuracy:

Sensitivity = True Positive / (True Positive + False Negative)

Specificity = True Negative/ (False Positive + True Negative)

Accuracy = TP + TN / TP+ TN + FP + FN

Table I Attributes considered for the Experimental setup

Attributes	Features
Total Number of videos Tested	30
Video Category	Set1: High-density plants Set2: Poor illuminated condition Set4: Partially occluded videos
Video Type	AVI format
Frame size	320x240

TABLE II
Performance metrics of PCA feature reduction.

DATA set	Block decomposition	Eigenvectors selected
EHOG model-PCA (Carl Vondricket a.2013)	Overlapping model	28
SIO-HOG model - PCA	Non-overlapping model	22

TABLE III
Average detection rate performance of proposed HOG class in the tobacco dataset.

DATA set (SIO-HOG model)	False detection rate	Average detection rate performance (accuracy)		
		Sensitivity	Specificity	Accuracy
High-	19%	89%	87%	94%

density plants				
Poor illuminated condition	10%	83%	81%	91%
Partially occluded plants	14%	81%	79%	89%

6. Conclusion

The unique SIO-HOG descriptors that are introduced in this research work are obtained by transforming the invariant histogram of gradient descriptors to scale as well as illumination variations. To demonstrate the configuration effectiveness of the HOG descriptor, various feature variations, as well as illumination variations, were studied. The performance metrics of the suggested model of HOG are evaluated for the detection of the tobacco plant and the experimental results show better performance when compared with the data sets that are obtained from the real fields. Moreover, dimensionality reduction proves the effectiveness of feature selection in overall detection accuracy. Since it is analyzed plant detection with real-time world data it is well proved that the presented tobacco plant detection system outperforms other HOG descriptors with constant bin selection. This work can be extended to carry out a plant growth analyzes system to prove the consistency of the proposed HOG model over a period of time.

Reference

[1]. W. T. Freeman and M. Roth. Orientation histograms for hand gesture recognition. Intl. Workshop on Automatic Face and Gesture-Recognition, IEEE Computer Society, Zurich, Switzerland, pages 296–301, June 1995.

[2]. D. G. Lowe. Distinctive image features from scale-invariant keypoints. IJCV, 60(2):91–110, 2004

[3]. P. Viola, M. J. Jones, and D. Snow. Detecting pedestrians using patterns of motion and appearance. The 9th ICCV, Nice, France, volume 1, pages 734–741, 2003.

[4]. K. Mikolajczyk, C. Schmid, and A. Zisserman. Human detection is based on a probabilistic assembly of robust part detectors. The 8th ECCV, Prague, Czech Republic, volume I, pages 69– 81, 2004.

[5]. Sabzi, S., Gilandeh, Y. A., & Javadikia, H. (2018). Developing a machine vision system to detect weeds from potato plant. Journal of Agricultural Sciences, 24(1), 105-118.

[6]. P. F. Felzenszwalb, R. B. Girshick, D. McAllester, and D. Ramanan, "Object detection with discriminatively trained part-based models," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 32, no. 9, pp. 1627–1645, 2010.

[7]. Xiaojie Guo and Xiaochun Cao, “MIFT: A framework for feature descriptors to be mirror reflection invariant,” Image and Vision Computing, vol. 30, no. 8, pp. 546–556, 2012.

[8]. Reddy, Vijaya Kumar, Kiran Kumar Siramoju, and Pradip Sircar. "Object Detection by 2-D Continuous Wavelet Transform." *Computational Science and Computational Intelligence (CSCI), 2014 International Conference on*. Vol. 1. IEEE, 2014.

[9]. Zhao, Wan-Lei, and Chong-Wah Ngo. "Flip-invariant SIFT for copy and object detection." *IEEE Transactions on Image Processing* 22.3 (2013): 980-991.

[10]. K. Mikolajczyk and C. Schmid. Scale and affine invariant interest point detectors. IJCV, 60(1):63–86, 2004.

[11]. Kanezaki, Asako, Yusuke Mukuta, and Tatsuya Harada. "Mirror reflection invariant HOG descriptors for object detection." *Image Processing (ICIP), 2014 IEEE*

International Conference on. IEEE, 2014.

[12]. Pedersoli, Marco, and Tinne Tuytelaars. "A scalable 3d hog model for fast object detection and viewpoint estimation." *3D Vision (3DV), 2014 2nd International Conference on*. Vol. 1. IEEE, 2014.

[13]. Chayeb, A., et al. "HOG based multi-object detection for urban navigation." *Intelligent Transportation Systems (ITSC), 2014 IEEE 17th International Conference on*. IEEE, 2014.

[14]. Chen, Lu, Panfeng Huang, and Jia Cai. "Object detection for non-cooperative targets using HOG-based proposals." *Robotics and Biomimetics (ROBIO), 2015 IEEE International Conference on*. IEEE, 2015.

[15]. Singh, Shekhar, and S. C. Gupta. "Human object detection by HoG, HoB, HoC and BO features." *Parallel, Distributed and Grid Computing (PDGC), 2016 Fourth International Conference on*. IEEE, 2016.