Car Detection Methodology in Outdoor Environment Based on Histogram of Oriented Gradient (HOG) and Support Vector Machine (SVM)

S. Guzmán[1], A. Gómez[1], G. Diez[1], D.S. Fernández[1,2]

{sebastian.guzman, alexander.gomezv, german.diezv, david.fernandez}@udea.edu.co

[1]Grupo de investigación GEPAR, Facultad de Ingeniería, Universidad de Antioquia UdeA, Calle 70 No. 52-21, Medellín, Colombia.

[2]Departamento de Ingeniería Electrónica y de Telecomunicaciones, Facultad de Ingeniería, Universidad de Antioquia UdeA, Calle 70 No. 52-21, Medellín, Colombia.

Keywords: Computer vision; surveillance; finite state machine (FSM); histogram of oriented gradient (HOG); support vector machine (SVM)

Abstract

Car detection is considered one of the best solution that used to reduce the congestion problem in big cities. There are many attempts to solve the problem based on Millimeter-wave radar, and magnetic loop sensor. However, car delectation based on video camera is more robust because it is more accurate and gives information can be easily understood by human. In this work, we propose methodology for car detection in outdoor environment. Our method integrates HOG features and SVM to determines whether there is a car or not in the captured frame. Extensive experiments were performed by changing SVM parameters to gain 99% of successful classification. Moreover, we develop a new database with multiple environment noise, hard conditions and proved methodology in another dataset[13] that reached similar results.

1 Introduction

Urban traffic has been increasing through years resulting in an important problem of vehicle congestion. Big cities does not have space to build more lanes and strategies like forbid vehicle use are less feasible each day. Current solutions use information and control of existing city infrastructure to design strategies for reduce congestions problem based on detect and track cars that cross street. Approximations like Millimeter-wave radar used in adaptive cruise control (ACC)[1] that are based on frequency shift in the received frequency-modulated waveform signal are used to quantify distance to the detected object. Another strategy is magnetic loop sensor (MLS)[2] sense metal of the cars that cruise above it, however those sensors just measure some variable such as distance of an object, induction and time change signals like voltage of some object that does not necessarily be a car just is in front of the sensor. Video cameras have been used for surveillance and traffic monitoring because they provide information that human easily can understand. Vision-based systems for traffic analysis are flexible and provide solutions to sense vehicle speed, quantity and traffic [3].

Urban vision-based systems deal with problems like low camera resolution, real-time analysis and intrusion (pedestrians, motorcycles and shadows) [4]. Many approximations to detect and track cars have been done, in [5] a feature matching between frames using expected displacement and background subtraction got 17.2% of error. [6] Reached 15% of error using geometric properties found in traffic scenes. A Markov chain Monte Carlo method is proposed in [7], they treated each frame as a model-based segmentation problem and used Viterbi algorithm for track, this method got 88% of efficiency. Recent solutions [8] use Bayesian inference and background subtraction reach a 5% of error at day without rain or dust.[9] used a SVM approach with Three-dimensional extended histograms of oriented gradients got 88% of efficiency using i-LIDS Parked Car data set[10]. In [11] stereo cue generates potential target locations, after extended histogram of oriented gradients are extracted and feed a SVM that reach a 3% of error in the UIUC dataset[12]. A Haar -like feature detector with a validation based on histogram of oriented gradients and SVM is proposed in [13] to detect back of cars in a highway, they reach a 91% of true detection rate.

In this paper a methodology for car detection based on histogram of oriented gradient (HOG) previously used for pedestrian, car detection and support vector machine (SVM) is evaluated. A new database for car recognition in multiple light, weather and intrusion conditions was developed. SVM with linear and Gaussian Kernel with multiple configurations were proved to evaluate features performance in the new dataset and standard one[13].

The rest of the paper is organized as follows: in section 2 Segmentation method, feature extraction and classification are explained, section 3 contains details about the experiment, in section 5 are the results and finally conclusions are presented.

2 Methodology

In this section the methodology is explained, Fig.1 summarize the process. Due to outdoor environment presents a lot of natural noise that can affect detection process a Gaussian smooth filter with mask and standard deviation of 3 is applied each time a new frame is captured.

Figure 1: Methodology Diagram

2.1 Object detection

In this section the initial trigger for object detection program is explained. A single street and one directional traffic flow is assumed since our interest is car detection in a single image.

A basic background subtraction method is implemented based on frame differencing [4], consist in a pixel by pixel absolute difference between two consecutive frames then this result is thresholding using an heuristic value of 80 and used as foreground mask. When threshold value get smaller more false positives pass to classification step on the contrary if this value increases any car will be detected. Frame differencing approach is very fast to compute but sensitive to noise itself, a region of interest(ROI) solution is make based on [14] , the two rectangular ROI's are putting one front another in the street traffic direction and the triggered is based on percentage of white pixels in background subtraction method implemented.

A finite state machine (FSM) was developed to filter intruders such as pedestrians or bicycles found on size criteria (a single person or bicycle is smaller than a car) and direction of the movement (pedestrians does not follow street direction always). Noise produced by shadows, abrupt illumination changes or periodic movements either follow a predictable movement in the street direction hence the FSM follow this reasoning, a detailed diagram is shown in Figure 2. Initial state change when ROI1 is activated and state 2 when ROI2 is trigger while ROI1 is activated this happens when an object follows a continuous movement in street direction . Once the system is in state 3 an image that include both ROI's are obtained Figure 2. This method does not allow to detect static cars, just those who pass through interest zones.

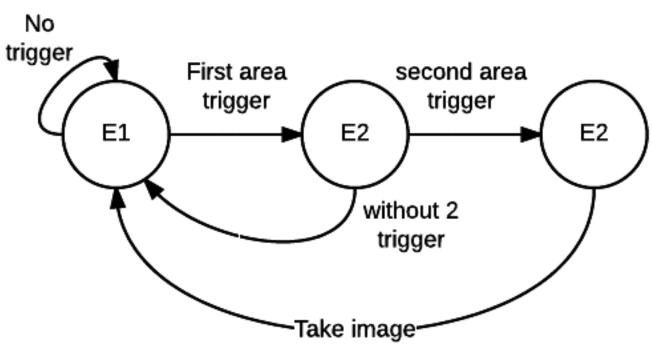

Figure 2: Methodology Diagram

2.2 Classification

2.2.1 Characterization

The image is converted to grayscale and histogram of oriented gradients (HOG) was obtained. HOG are a feature descriptor used for object recognition [15], this vector is extracted imposing a grid on image and calculating normalized histogram gradient orientation in each cell of the grid. To calculate gradient each cell is convolved with a 1-D kernel [-1,0,1] and applied along nine directions called bins then each pixel receive a weighted vote based on it intensity and gradient orientation, this values are accumulated into regions called cells, finally blocks of cells are normalized using the Equation (1) to eliminate local variation in illumination and foreground-background. Where Vn is normalized vector, V is an un-normalized vector result of joint un-normalized cells and ∈ is a coefficient for preventing division by zero problems.

$$Vn = \sqrt{\frac{V}{||V|| + \epsilon}} \quad (1)$$

2.2.2 Features extraction

High dimension of features vector represents a problem in the classification step, because of the curse of dimensionality [13]. To avoid this problem a dimension reduction is necessary. Feature extraction consists of mapping the input vector of observations x with n characteristics into a new feature description z with least characteristics, intended to be informative, non redundant, facilitating the subsequent learning. The system used Principal component analysis (PCA) to apply feature extraction. PCA is a statistical procedure that uses an orthogonal transformation to convert the set of observations x of possibly correlated variables into a set of values of linearly uncorrelated variables called principal components. The principal components are orthogonal because they are the eigenvectors of the covariance matrix, which is symmetric [13]. In the system the PCA is applied to HOG features discarding the five percent of the variance, reducing the dataset dimension features.

2.2.3 Classifier

Support vector machine is a kind of learning model which look for the best decision boundary. This boundary is located between the two classes to classify. In support vector machines the decision boundary is chosen to be the one for which the margin is maximized. This margin is the shortest distance from the decision boundary and any train point [16].

After the model is train and the Equation (2) is used to determinate if a new sample belongs to a class one or class two.

$$y(x) = \sum_{n=1}^{N} a_n t_n k(x_n, x_m) + b \quad (2)$$

Where the conditions in the Equation (3) Equation (4) and Equation (5), are satisfied:

$$a_n \geq 0 \quad (3)$$

$$t_n y(x_n) - 1 \geq 0 \quad (4)$$

$$a_n\{t_n y(x_n) - 1\} = 0 \quad (5)$$

The points with an=0 are not into the sum and can be discarded. Other points are known like **Supports Vectors**. These vectors are samples nearest to decision boundary. These samples are used to make the classification, considering these vectors enough to build the separating hyper plane.

3 Experiments and result

3.1 database description

The system was test using a proprietary car database. This database was built using three cameras with 640x480 resolution, these cameras pointing towards 3 streets Figure 3. A total of 11940 images samples were taken. There were 5970 positive samples and 5970 negative samples. Positive samples were images with car photographs, and the negative samples 5970 were photographs with sunny-street, shady-street, sidewalk, pedestrians and motorcyclists. In each image a region of interest with a 390 x 140 dimension was extracted manually to ensure that only the car was on the image and then the image was saved in a RGB format.

(a) (b)

Figure 3: Sample dataset images (a) negative images (b) positive images

3.2 Experimental setup

Cross-Validation with ten folds has been used. PCA transformation matrix was extracted from the train folds and applied to each validation sample. SVM models were trained varying the type of kernel between linear and Gaussian, box constraint from 0.01 to 100, and gamma parameter between 0.01 and 100 for the rbf kernel.

3.3 Results

Table 1 shows efficiency percent, confidence interval of each combination of parameters and can be made the following remarks: For the RBF kernel, the efficiency improves as the parameter gamma decreases and increases the Box constraint(C) parameter; The best result was obtained using a rbf kernel with gamma = 0.01 and C = 100 and 10 and for the linear kernel, efficiency has always been close to 98% in this database.

Since SVM fails when Gamma factor increase the classes are not disperse that means that HOG features separate classes even with outdoor noise. A linear kernel in support vector machine can solves classification problem, this issue lets evaluate SVM faster than a gaussian kernel. Table 2 shows best performance than reported in [11] using same dataset and HOG instead of extended HOG features. Although in [11] is discussed that HOG is not discriminative enough experimental results show best performance in this task. Hence similar results was obtained in 2 different datasets robustness of features to solve car detection problem is proved.

RBF					
	BOX CONSTRAINT				
Gamma	**0,01**	**0,1**	**1**	**10**	**100**
0,01	95,168%± 2,705	98,317%± 0,983	98,970%± 0,679	99,079%± 0,641	99,087% ±0,653
0,1	88,961%± 6,925	97,571%± 1,440	98,953%± 0,543	98,953%± 0,543	98,953% ±0,543
1	50,251%± 0,306	50,260%± 0,317	59,489%± 5,069	59,941%± 5,271	59,941% ±5,271
10	50,034%± 0,043	50,251%± 0,493	51,533%± 1,542	51,533%± 1,542	51,533% ±1,542
100	51,147%± 3,628	51,365%± 3,581	51,432%± 1,518	51,441%± 1,536	51,441% ±1,536
LINEAR					
-	98,652%± 0,749	98,710%± 0,655	98,333%± 0,839	98,233%± 0,921	98,233% ±0,921

Table 1: Experiment results

	BOX CONSTRAINT				
Gamma	**0,01**	**0,1**	**1**	**10**	**100**
0,01	89,131%± 3,423	96,873%± 2,877	99,347%± 1,046	99,447%± 0,687	99,447%± 0,867
0,1	96.511%± 1,987	96,637%± 2,081	98,701%± 1,411	98,829%± 1,463	98,829%± 1,463
1	50,503%± 1,026	50,503%± 1,026	52,579%± 3,555	52,579%± 3,555	53,579%± 3,555
10	99,09%± 0,8,21	99,099%± 0,821	50,378%± 0,821	50,503%± 1,026	50,503%± 1,026
100	50,254%± 0.513	50,254%± 0.513	50,254%± 0.513	50,254%± 0,513	50,254%± 0,513
LINEAR					
-	88,486%± 3,524	92,996%± 3,537	98,706%± 1,002	99,605%± 0,839	99,605%± 0.839

Table 2: results in UIUC dataset[13]

4 Conclusions

In this work a methodology for detect and count cars was evaluated, classification results show HOG features as good descriptor to car identification problem reaching 0.9909 of efficiency with 0.0065 standard deviation in aggressive outdoor conditions. A FSM was useful and efficient basic filtering and trigger method to outdoors problems like shadows, abrupt illumination changes or periodic movements. A new dataset for car detection and benchmark was development with 5970 of positive and 5970 of false images.

In future work the features will be proved with other datasets, classifiers and circumstances to test features performance. Motorcycle and pedestrian classes will be included in the classifier.Acknowledgements

We thanks to the engineer Juan Diego Velez, chief of Recursos de Apoyo e Informática DRAI, also to GEPAR research group for their support.

References

[1]. X.Mao, D. Inoue, S. Kato, and M. Kagami, “Amplitude-modulated laser radar for range and speed measurement in car applications,” IEEE Trans. Intell. Transp. Syst., vol. 13, no. 1, pp. 408–413, Mar. 2012.

[2] TAI, Jen-Chao, et al. Real-time image tracking for automatic traffic monitoring and enforcement applications. Image and Vision Computing, 2004, vol. 22, no 6, p. 485-501.

[3] R. Cucchiara, M. Piccardi, P. Mello, Image analysis and rule-based reasoning for a traffic monitoring system, IEEE Transactions on Intelligent Transportation Systems 1 (2) (2000) 119–130.

[4] BUCH, Norbert; VELASTIN, Sergio A.; ORWELL, James. A review of computer vision techniques for the analysis of urban traffic. Intelligent Transportation Systems, IEEE Transactions on, 2011, vol. 12, no 3, p. 920-939.

[5] S. Messelodi, C.M.Modena, and M. Zanin, “A computer vision system for the detection and classification of vehicles at urban road intersections,” Pattern Anal. Appl., vol. 8, no. 1/2, pp. 17–31, Sep. 2005.

[6] MASOUD, Osama; PAPANIKOLOPOULOS, Nikolaos P. Using geometric primitives to calibrate traffic scenes. Transportation Research Part C: Emerging Technologies, 2007, vol. 15, no 6, p. 361-379.[7] SONG, Xuefeng; NEVATIA, Ram. Detection and tracking of moving vehicles in crowded scenes. En Motion and Video Computing, 2007. WMVC'07. IEEE Workshop on. IEEE, 2007. p. 4-4.

[8] Nieto et al.: Vehicle tracking and classification in challenging scenarios via slice sampling. EURASIP Journal on Advances in Signal Processing 2011 2011:95.

[9] N. Buch, J. Orwell, and S. A. Velastin, “Three-dimensional extended histograms of oriented gradients (3-DHOG) for classification of road users in urban scenes,” in Proc. BMVC, London, U.K., Sep. 2009.

[10] Imagery Library for Intelligent Detection Systems, i-LIDS,Home Office Scientific Development Branch. [Online]. Available: http://www.ilids.co.uk/

[11] H. Feng, Y. Shan, R. Cekander, H. S. Sawhney, and R. Kumar. "A two-stage approach to people and vehicle detection with hog-based svm." In Performance Metrics for Intelligent Systems 2006 Workshop, pp. 133-140. 2006.

[12] A. Shivani, A. Awan, and D. Roth. "Learning to detect objects in images via a sparse, part-based representation." Pattern Analysis and Machine Intelligence, IEEE Transactions on 26, no. 11, pp. 1475-1490, (2004).

[13] W. Huan and H. Zhang. "A Hybrid Method of Vehicle Detection based on Computer Vision for Intelligent Transportation System."International Journal of Multimedia & Ubiquitous Engineering 9, no. 6 (2014).

[14] MATTHEWS, N. D., et al. Vehicle detection and recognition in grayscale imagery. Control Engineering Practice, 1996, vol. 4, no 4, p. 473-479.

[15] Dalal, N., & Triggs, B. (2005, June). Histograms of oriented gradients for human detection. In Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on (Vol. 1, pp. 886-893). IEEE.

[16] Bishop, C. M. (2006). Pattern recognition and machine learning. springer.