GA Based PHOG-PCA Feature Weighting for On-Road Vehicle Detection

Nima Khairdoost, S. Amirhassan Monadjemi, Zohreh Davarzani, and Kamal Jamshidi

***Abstract*—Vehicle detection is an important issue in driver assistance systems and self-guided vehicles that includes two stages of hypothesis generation and verification. In the first stage, potential vehicles are hypothesized and in the second stage, all hypothesis are verified. The focus of this work is to classify vehicle candidate images into vehicle and non-vehicle classes. We extract Pyramid Histograms of Oriented Gradients (PHOG) features from a traffic image as candidates of feature vectors to detect vehicles. Principle Component Analysis (PCA) is applied to these PHOG feature vectors as a dimension reduction tool to obtain the PHOG- PCA vectors. Then we employ real coded chromosome Genetic Algorithm (GA) and linear Support Vector Machine (SVM) to classify the PHOG-PCA features as well as to improve their performance and generalization. Our tests show good classification accuracy of more than 96% correct classification on realistic on-road vehicle images.**

***Index Terms*—Feature weighting, GA, linear SVM, PCA, PHOG, vehicle detection.**

I. Introduction

Robust and reliable vehicle detection in images is the critical step for driver assistance systems and self-guided vehicles as well as traffic controllers. This is a very challenging task since it is not only affected by the size, shape, color, and pose of vehicles, but also by lighting conditions, weather, surroundings, and the surface of different roads. In [1], Principle Component Analysis (PCA) was used for feature extraction and linear Support Vector Machine (SVM) for classification of vehicle images. Goerick *et al.* [2] employed Local Orientation Code (LOC) to extract edge information of ROI and NNs to learn the characteristics of vehicles. Features extraction using Gabor filters was investigated in [3], Gabor filters provide a mechanism to extract line and edge information by tuning orientation and changing the scale. Sun *et al.* [4] performed the comparison of three methods of wavelet feature extraction which are all based on coefficients and grayscale space. In [5] a general object detection scheme was proposed using PCA and Genetic Algorithm (GA) for feature extraction and feature subset selection respectively.

Manuscript received July 19, 2012; revised August 21, 2012.

Nima Khairdoost was with Department of Computer Engineering, Faculty of Engineering, University of Isfahan, Isfahan, 81746, Iran (e-mail: n.khairdoost@gmail.com).

S. Amirhassan Monadjemi and Kamal Jamshidi are with Department of Computer Engineering, Faculty of Engineering, University of Isfahan, Isfahan, 81746, Iran (e-mail: monadjemi@eng.ui.ac.ir, jamshidi@eng.ui.ac.ir).

Zohreh Davarzani was with Department of Computer Engineering, Faculty of Engineering, Ferdowsi University of Mashhad, Mashhad, Iran (e-mail: z_davarzani@yahoo.com).

In this paper, we consider the problem of front and rear view vehicle detection from gray-scale images. The first stage of any vehicle detection system is to hypothesize the locations of potential vehicles in an image. Then, all hypothesis are verified. Both stages are important and challenging. Approaches to generate the hypothetical locations of vehicles in an images use motion information [6], symmetry [7], color [8], shadows [9], texture [10], vertical/horizontal edges [11], and stereo-based [12]. Our emphasis here is on improving the performance of the verification stage by employing a novel PHOG-PCA feature weighting method based on GA.

The rest of the paper is organized as follows: in Section II, the proposed method is described in detail. Section III shows the experimental results and the last section is the conclusions.

II. Proposed Method

In this paper, Pyramid Histograms of Oriented Gradients (PHOG) features are extracted from an image dataset as primitive features since they have shown good results in object recognition [13]. To improve the classification accuracy and reduce the dimensionality, we also apply PCA to the PHOG features to generate what we call the PHOG-PCA feature vector. Then we divide the samples into two parts of *Training Data* and *Test Data* as shown in Fig. 1.

It is well known that feature weighting is effective for pattern classification as shown in [14], [15]. It is expected that the classification accuracy can be further improved by weighting the proper first PHOG-PCA features since some local regions are less relevant in the vehicle detection than the others. For this purpose, we use a real coded chromosome GA feature weightener. The *Training Data* is divided into two parts of *data1* and *data2*. We employ Linear SVM for vehicle/ non-vehicle classification which is trained with the *data1* and then the *data2* is used for validation of the classifier. The classification accuracy is returned to the GA as one of the fitness factors. After the convergence of the GA, linear SVM is trained regarding the *Optimum Weights* and the *Training Data* then we test it with the *Test Data* and the classification accuracy of our proposed method is obtained. The overview of the vehicle detection system is shown in Fig. 1.

A. Pyramid Histograms of Oriented Gradients (PHOG)

PHOG descriptor is a spatial pyramid representation of HOG descriptor, and reached good performance in many studies, e.g. [13], [16], [17]. In this paper, PHOG features are extracted from vehicle and non-vehicle samples to represent by their local shape and spatial layout. As illustrated in Fig. 2,

DOI: 10.7763/IJIEE.2013.V3.276

the PHOG descriptor consists of a histogram of orientation gradients over each image sub-region at each resolution. For extracting PHOG features, edge contours are extracted using the Canny edge detector for entire image as shown in Fig. 2. Then each image is divided into cells at several pyramid level. The grid at resolution level l has 2^l cells along each dimension. The orientation gradients are computed using a 3×3 Sobel mask without Gaussian smoothing. Histogram of edge orientations within each cell is quantized into K bins. Each bin in the histogram represents the number of edges that have orientations within a certain angular range. Histograms of the same level are concatenated into one vector. The final PHOG descriptor for an image is a concatenation of all vectors at each pyramid resolution that introduces the spatial information of the image [13]. Consequently, level 0 is represented by a K-vector corresponding to the K bins of the histogram, level 1 by a 4K-vector, and the PHOG descriptor of the entire image is a vector with dimensionality $K\sum_{l\in L} 4^l$. The PHOG descriptor is normalized to sum to unity that ensures images with more edges are not weighted more strongly than others. Fig. 2 shows PHOG descriptor procedure and the PHOG features of example images. As can be seen, vehicle images have similar PHOG representations whereas non-vehicle images have different PHOG representations far enough from the vehicle ones.

B. Principal Component Analysis (PCA)

The total number of extracted PHOG features is rather high. Also, these features are probably irrelevant and redundant. PCA was applied in [18] for reducing the dimensionality of the feature vectors. PCA can be defined as the orthogonal projection of the input data onto a lower dimensional linear subspace, such that the variance of the projected samples is maximized. Dimension reduction and noise reduction are two advantages of employing PCA. In this paper, we utilize this idea to reduce the dimensionality of the feature vectors. The PCA algorithm can be summarized in the following:

Let $\{x_i \mid i=1,\ldots,N\}$ be a set of M-dimensional vectors. We compute the mean vector of input vectors that is defined as $\bar{x} = \frac{1}{N}\sum_{i=1}^{N} x_i$ and then we compute the covariance matrix Σ that is defined as follows:

$$\Sigma = \frac{1}{N}\sum_{n=1}^{N}(x_i - \bar{x})(x_i - \bar{x})^T \tag{1}$$

By solving the eigen equations of the covariance matrix Σ, the optimum projection matrix U is obtained

$$\Sigma U = \Lambda U, \qquad (UU^T = I) \tag{2}$$

and then the PCA scores for any PHOG features can be computed by using the following equation. We called these new features PHOG-PCA features.

$$y = U^T(x_i - \bar{x}) \tag{3}$$

In order to reduce the dimensionality, we just keep the first d principal axis that they keep the significant discriminant information.

C. Genetic PHOG-PCA Feature Weighting

GA is a probabilistic optimization algorithm and a branch of evolutionary algorithms. In the past, they have been used to solve different problems such as object detection [5], face recognition [19], speech recognition [20], image annotation [21], and attribute weighting artificial immune system (AWAIS) [14].

In this study, we utilized the GA for the PHOG-PCA features weighting. Since our goal in feature weighting is to reduce the classification error, we used the GA to find optimum weights for features that give the minimum classification error when used by the classifier. Thus, we formed a population consisting of the chromosomes representing the weights for features of two classes of vehicle and non-vehicle and used them in the GA process. The best chromosome is the one leading to the lowest test classification error. The procedure in finding optimum weights via the GA is as follows:

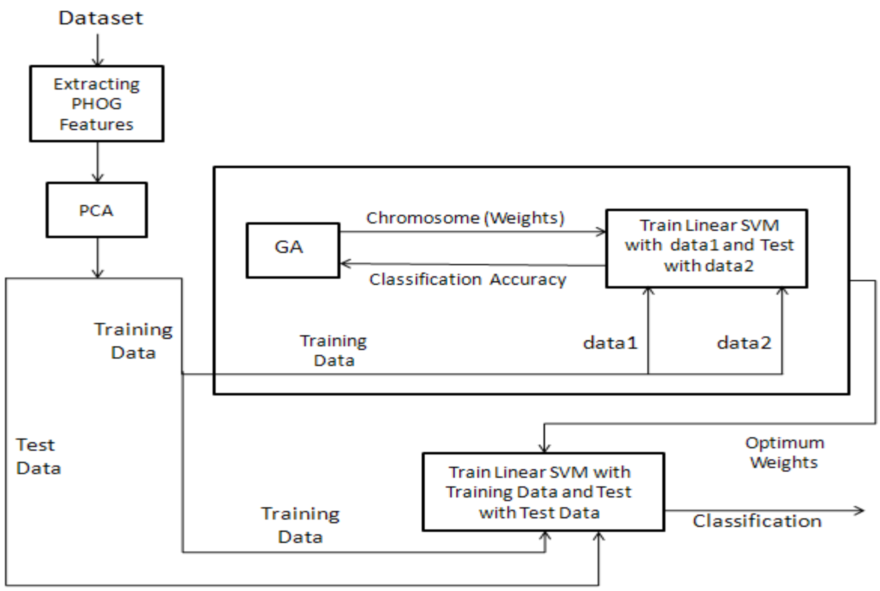

Fig. 1. The overview of the vehicle detection system

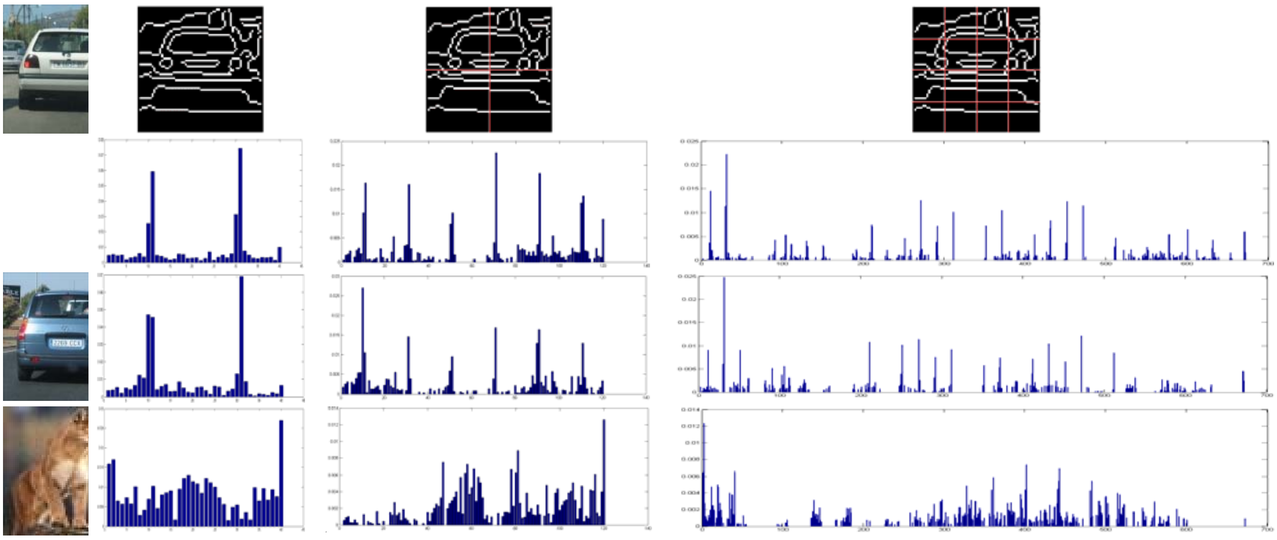

Fig. 2. Shape spatial pyramid representation. Top row: a vehicle image and grids for levels l = 0 to l = 2; Below: histogram representations corresponding to each level. The final PHOG vector is a weighted concatenation of vectors (histograms) for all levels. Remaining rows: another vehicle image and a non-vehicle image, together with their histogram representations.

1) Feature weighting encoding

Let the number of PHOG-PCA features be L, so each chromosome will be represented with L genes that each gene take values from the range of [0,10], which in our study is divided into 100 discrete levels.

2) Calculating the fitness of these chromosomes

We forced weights <1 in value to 0 during our trial. These embellishments resulted in GA-optimized classifiers with reduced feature sets. With some training data and regarding non-zero weights, linear SVM is trained using the chromosome whose fitness value is to be calculated. Then, some test data are presented to the trained classifier and classification accuracy is calculated in percentage form. The fitness function is as follows:

$$Fitness(c)= CA^4 (c) - \alpha(\frac{N(c)}{L}) \tag{4}$$

where c is the chromosome, $CA(c)$ is the classification accuracy using the linear SVM classifier. α represents the tradeoff between the two criteria (using α =0.01). $N(c)$ is the number of non-zero weights. And finally, L is the total number of features (which is fixed at 250 for all experiments). In our experiments, classification accuracy is often more than 75% so we used $CA^4(c)$ instead of $CA(c)$ because it can be more distinctive fitter chromosome than others.

3) Initial population

All the genes of the first chromosome are ‘10’, which means the weights of all the features are equal. The other chromosomes are generated randomly. In all of our experiments, we used 1000 generations and a population size of 800. In most cases, the GA converged in less than 1000 generations.

4) Crossover

We used uniform crossover, in this case each bit of the offspring is selected randomly from the corresponding bits of the parents. The crossover rate used in all of our experiments was 0.9.

5) Mutation

We choose uniform mutation that is, each bit has the same mutation probability. The mutation rate used in all of our experiments was 0.07.

6) Elitism

We used the elitism strategy to prevent fitness of the next generation be smaller than the largest fitness of the current generation, the best 30 chromosomes are preserved for the next generation automatically.

III. Experimental Results

A. Dataset

The vehicle dataset used contains 1646 non-vehicle images and 1648 front and rear view vehicle images. Some of these images are from the MIT vehicle dataset and the Caltech-101 dataset, while the rest images have been gathered with different types, poses, and colors (although all images were converted to grayscale). Some of the images contain the vehicle and other background objects. We converted all images to jpg format and normalized size of each image to 64x64 pixels (see Fig. 3).

B. Experiments

In our experiments, we used the linear SVM classifier and extracted PHOG features from all collected images with 3 levels of pyramids and 40 orientation bins in the range of [0, 360] in each level. Therefore the 3 level PHOG descriptor of an image is an 840-vector.

Also, we used a 7-fold cross-validation to estimate both the accuracy and generality of the linear SVM classifier. In this case, all of the examples are partitioned into 7 subsamples and the 7th subsample is retained as *Test data* while the remaining 6 subsamples are used as *Training Data*. Cross-validation is then repeated 7 times with all of the 7 subsamples used exactly once as the *Test data*. In first experiment, we used PHOG descriptors and linear SVM classifier. Table I shows the result.

In second experiment, we employed PCA to reduce the dimensionality of PHOG descriptors and produce the reduced PHOG-PCA features. We also used here the same folds in the first experiment for cross validation. We estimated linear SVM accuracy for different number of first PHOG-PCA features. Table II shows the result.

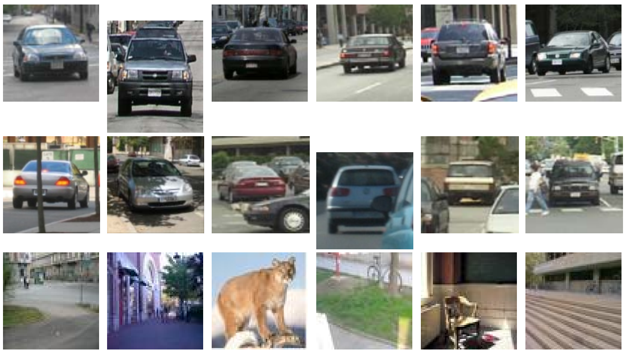

Fig. 3. Some vehicle and non-vehicle training sample images

TABLE I: CLASSIFICATION RESULTS WITH PHOG FEATURES

Number of Features	True Positive (%)	True Negative (%)	Classification Accuracy (%)
840	95.87	91.31	93.59

TABLE II: CLASSIFICATION RESULTS WITH EMPLOYING PCA TO REDUCE THE DIMENSIONALITY OF PHOG DESCRIPTORS

Number of Features	True Positive (%)	True Negative (%)	Classification Accuracy (%)
150	96.42	93.01	94.72
200	96.48	93.50	94.99
230	96.60	93.68	95.14
250	96.66	93.68	95.17
270	96.54	93.68	95.11
300	96.54	93.62	95.08
350	96.46	93.57	95.02
400	96.34	93.49	94.92

TABLE III: CLASSIFICATION RESULTS WITH EMPLOYING GA TO WEIGHT PHOG-PCA FEATURES (OUR PROPOSED METHOD)

Number of Features	True Positive (%)	True Negative (%)	Classification Accuracy (%)
241	97.32	95.97	96.65

As can be seen, the 250 first PHOG-PCA features shows the better performance so in the third experiment, we used the GA with the configuration that mentioned in section II.C to optimize weight of these features. We also used the same folds in the previous experiments for cross validation and in this case, we used 5 folds of Training Data (that we called Data1) for training of linear SVM and 1 fold of Training Data (that we called Data2) for validation of the learned classifier to guide GA during weight optimization. After the convergence of the GA, we trained linear SVM with the Training Data regarding the optimum weights and tested it using Test Data. Feature weighting when applied was also able to reduce the number of features from 250 to 241, and enhance accuracy of classification by 1.48%. Table III shows the result.

IV. CONCLUSION

We have proposed a three steps method for classifying vehicle candidate images into vehicle and non-vehicle classes. First of all, we extracted PHOG features from an image dataset as primitive features and then we applied the PCA to reduce the dimensionality of PHOG descriptors and

REFERENCES

[1] Q. B. Truong and B. R. Lee, "Vehicle detection algorithm using hypothesis generation and verification," in *Proc. ICIC (1)*, 2009, pp. 534-543.

[2] C. Goerick, N. Detlev, and M. Werner, “Artificial neural networks in real-time car detection and tracking applications,” *Pattern Recognition Letters*, vol. 17, pp. 335-343, 1996.

[3] T. R. Lim and A. T. Guntoro, “Car recognition using gabor filter feature extraction,” *Circuits and Systems*, *APCCAS’02*, vol. 2, pp. 451-455, 2002.

[4] Z. Sun, G. Bebis, and R. Miller, "Monocular precrash vehicle detection: features and classifiers," *IEEE Transactions on Image Processing*, vol. 15, no. 7, pp. 2019-2034, 2006.

[5] Z. Sun, G. Bebis, and R. Miller, "Object detection using feature subset selection," *Pattern Recognition*, vol. 37, pp. 2165-2176, 2004.

[6] C. Demonceaux, A. Potelle, and D. K. Akkouche, “Obstacle detection in a road scene based on motion analysis,” *IEEE Trans. on Vehicular Technology*, vol. 53, no. 6, pp. 1649– 1656, 2004.

[7] A. Bensrhair, M. Bertozzi, and A. Broggi, "A cooperative approach to vision-based vehicle detection," in *IEEE Intelligent Transportation Systems Conference Proceedings*, 2001.

[8] T. Xiong and C. Debrunner, “Stochastic car tracking with line- and color-based features,” *IEEE Transactions on Intelligent Transportation Systems*, vol. 5, no. 4, pp. 324–328, 2004.

[9] M. B. V. Leeuwen and F. C. A. Groen, "Vehicle detection with a mobile camera: Spotting midrange, distant, and passing cars," *IEEE Robotics and Automation Magazine*, vol. 12, no. 1, pp. 37-43, 2005.

[10] T. Bucher, C. Curio, and J. Edelbrunner, *et al.*, “Image processing and behavior planning for intelligent vehicles,” *IEEE Transactions on Industrial Electronics*, vol. 50, no. 1, pp. 62–75, 2003.

[11] Z. Sun, R. Miller, G. Bebis, and D. DiMeo, "A real-time precrash vehicle detection system," in *Proc. IEEE Int’l Workshop Application of Computer Vision*, Dec. 2002.

[12] U. Franke and A. Joos, “Real-time stereo vision for urban traffic scene understanding,” in *Proc. of IEEE Intelligent Vehicles Symposium (IVS ’00)*, 2000, pp. 273–278.

[13] A. Bosch, A. Zisserman, and X. Munoz, "Representing shape with a spatial pyramid kernel," In *Proc. of the International Conference on Image and Video Retrieval*, 2007.

[14] S. Ozşen and S. Guneş, "Attribute weighting via genetic algorithms for attribute weighted artificial immune system (AWAIS) and its application to heart disease and liver disorders problems," *Expert Systems with Applications*, vol. 36, pp 386-392, Jan. 2009.

[15] F. Hussein, R. Ward, and N. Kharma, “Genetic algorithms for feature selection and weighting, a review and study,” in *Proc. ICDAR*, 2001, pp. 1240-1244.

[16] B. Zhang, Y. Song, and S. U. Guan, "Historic chinese architectures image retrieval by SVM and pyramid histogram of oriented gradients features," *International Journal of Soft Computing*, vol. 5, no. 2, pp. 19-28, 2010.

[17] Y. Bai, L. Guo, L. Jin, and Q. Huang, "A novel feature extraction method using pyramid histogram of orientation gradients for smile recognition," in *16th IEEE International Conference on Image Processing (ICIP)*, 2009, pp. 3305 – 3308.

[18] T. Kobayashi, A. Hidaka, and T. Kurita, "Selection of histograms of oriented gradients features for pedestrian detection," in *Proc. ICONIP*, 2007, pp.598-607.

[19] C. Liu and H. Wechsler, "Evolutionary pursuit and its application to face recognition," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 22, no. 6, pp. 570-582, June 2000.

[20] B. T. Ongkowijaya and X. Zhu, "A new weighted feature approach based on GA for speech recognition," in *7th International Conference on Signal Processing (ICSP)*, 2004, pp. 663–666.

[21] T. Zhao, J. Lu, Y. Zhang, and Q. Xiao, "Image annotation based on feature weight selection," in *International Conference on Cyber worlds*, 2008, pp. 251–255.

Nima Khairdoost received the BS and MS degrees in Computer engineering from Ferdowsi University of Mashhad and University of Isfahan, Iran in 2008 and 2011, respectively. His research interests include Image processing, machine vision and pattern recognition as well as evolutionary algorithms.

S. Amirhassan Monadjemi was born 1968, in Isfahan, Iran. He received his BS degree at computer hardware engineering from Isfahan University of Tech., in 1991, his MS degree at computer engineering, machine intelligence and robotics from Shiraz University, Shiraz, in 1994 and his PhD degree in computer science, image processing and pattern recognition from University of Bristol, Bristol, UK. in 2004. He is now working as assistant professor at the Department of Computer Engineering, Faculty of Engineering, University of Isfahan, Isfahan, Iran. His research interests include image processing, computer vision and pattern recognition, computer aided learning and physical detection and elimination of viruses.

Zohreh Davarzani received the BS and MS degrees in Computer engineering from Ferdowsi University of Mashhad, Iran in 2008 and 2011, respectively. Her research interests include fuzzy logic, soft computing and Quantum computing

Kamal Jamshidi received the MS and PhD degrees in electrical engineering from Anna University of India in 1990 and I.I.T University of India in 2003, respectively. He currently is an assistant professor in the Engineering Department of University of Isfahan. His research interests include wireless sensor network and vehicular ad hoc networks as well as fuzzy systems and microprocessor based systems.