International Journal of Pure and Applied Mathematics
Volume 118 No. 18 2018, 3277-3285
ISSN: 1311-8080 (printed version); ISSN: 1314-3395 (on-line version)
url: http://www.ijpam.eu
Special Issue
ijpam.eu

Modified HOG based on-road vehicle detection method

Yehia Zakaria[1], Hossam E. Abd El Munim[2], M Ghoneima[1], Sherif Hammad[2]
[1]Department of Mechatronics, Faculty of Engineering, Ain Shams University
Email: {yehia.zakaria,maged_ghoneima}@eng.asu.edu.eg
[2]Department of Computer and Systems, Faculty of Engineering, Ain Shams University
Email: {hossameldin.hassan,sherif.hammad}@eng.asu.edu.eg

Abstract

A new vehicle detection method based on modified histogram of oriented gradient (HOG) is introduced. Three variants of the modified HOG based feature are used to train linear and nonlinear support vector machine (SVM) classifiers. A comparison between classifiers based on proposed HOG variants and the conventional one is done. The classifiers are trained and tested on samples extracted from KITTI dataset. The comparison between the results show that the proposed method increased the discriminative power and improved the detection rates. Finally, the best performing classifiers in the comparison are used in detection of vehicles in image frames provided by KITTI dataset.

Keywords: Vehicle detection, Histogram of Oriented Gradient, HOG, Compass Gradient, Support Vector Machines, SVM

1.Introduction

In the last few years, making driving safer has been a target for automotive industry. While it was reported that most of the road crashes are due to human factors [1], manufacturers are keen to provide advanced driver assistance systems (ADAS) (i.e. collision alert, lane departure assistant, driver alertness check, ...) to decrease the probability or human error and a more long-term target is to replace the human driver with automated one that will make the roads secure and save more lives. A main module in any ADAS or autopilot is the perception module that provides the vehicle driver (human or computer) with information about the surrounding environment, depending on sensors like cameras, radars and lidars. A whole picture of the vehicle surrounding is provided for the vehicle driver to take proper actions according to the changes in environment. One core function in the perception module is the vehicle detection, as surrounding vehicles are one of the dynamic obstacles which represent a prospective collision threat that needs to be taken care of. Cameras as a passive sensor are widely used in the perception module of vehicles. A lot of research efforts were done to develop methods for vehicle detection using either stereo or monocular cameras. The method introduced in this paper depends on data from monocular camera. Vehicles are visually different objects that have a lot of variations in colour, size and shape. In on-road vehicle detection unlike in surveillance cameras, the camera scene is always changing which makes the vehicles background to be visually complex, besides the challenges of detection based on visual information such as illumination variation and occlusion. All the mentioned reasons make vehicle detection from camera information a non-trivial task such that a lot research effort was paid to approach this problem. Vehicle detection using monocular camera can be classified into two main approaches, the first is motion-based at which sequence of images are used in detection process and the second is appearance-based at which only one image can be used to detect vehicles. One of the motion-based techniques used in vehicle detection is optical flow which is used in [2] for detection of overtaking vehicle to use this information in lane changing decision. Also in [3] optical, flow is used along with hidden Markov model (HMM) in vehicle detection in different lighting conditions. Other motion based techniques such as dynamic scene modelling was used in [4] along with hypothesis testing and robust information fusion to detect overtaking vehicles. Dynamic background modelling based on sparse optical flow was used in [5] to detect overtaking vehicles. Appearance-based vehicle detection is a two-stage process which includes feature extraction and classification [6]. A lot of features were used in vehicle detection. Symmetry feature of vehicles was used in their detection [7]. In [8], the shadow underneath the vehicle was used as a feature for vehicle detection in traffic scene. In [9], a fusion of both symmetry and shadow features was used in vehicle detection. Edge-based constraint filter was used in [10] to segment vehicle from background as a step in vehicle detection and tracking. Combination of features (symmetry, vertical edges, taillight and shadow) was used in [11] to detect vehicles in different weather and lighting conditions. Haar like features which were used in the well know Viola-Jones face detector [12], were used in vehicle detection in [13]–[15]. Also it was used to detect independent vehicle parts in [16] full review of the methods used in vehicle detection can be found in [17] and [16].

The following parts of this paper is organized as follow: Related work is described in Sec. 2. The proposed HOG variant along with experiments done and results of the comparison between the proposed HOG with the original one are described in Sec. 3. Discussions of the results and conclusions are in Sec. 4.

2. Related Works

One of the features that is widely used in object detection is the histogram of oriented gradient (HOG). The HOG feature is firstly introduced by Dalal and Triggs [18] with application on pedestrian detection but it had been used in various object detection applications including vehicle detection [19], [20], and [21]. Since then efforts had been made to improve the discriminative power of HOG introducing different changes on Dalal and Triggs HOG version. In [22], Zhang et al. introduced the local structured HOG (LSHOG) fused with local structured Local Binary Pattern (LSLBP) and applied it for object detection. In [23], Cheon et al. made use of the symmetry in HOG feature vector for symmetric objects such as vehicles and introduced a new HOG variant called symmetric HOG. In [24], Kim et al. concatenated position and intensity data for the original HOG to form a position and intensity HOG (πHOG). Recently Kassani el at. [25], introduced the soft HOG (sHOG) that depends on random selection of cells position with symmetric features and applied it on traffic sign detection. However, all the mentioned efforts maintained the same gradient calculation method as in conventional HOG. The goal of this paper is to introduce a new HOG variant using compass gradient mask in the calculation of HOG that is proved to increase the discriminative power of the original HOG on vehicle detection. The classification part is done using the support vector machines (SVM) [26] which is widely used in literature in combination with the HOG feature [23], [27].

3. Proposed work

3.1 Overview on HOG calculation

The original HOG computation is done in five steps as described in [18]. First the image goes through colour normalization and gamma correction. Then the image is divided as a grid of cells. The cells are grouped into larger overlapping blocks at which the cells can belong to more than one block. Figure 1 shows an example for dividing an image into cells of size 16x16 at which each cell has 256 pixels and blocks of size 2x2 which means that each block contains 2 cells in each direction. The blocks in the figure have overlapping ratio of 50% at which half of the block cells are shared with the neighbour block. Cell size and block size are parameters to be determined by the user according to the image size and amount of details needed to be captured.

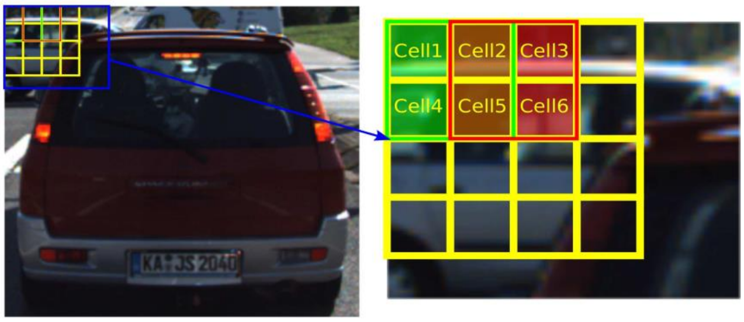

Figure 1: The image is divided into cells (outlined in yellow), the figure shows cells of cell size is 16x16, cells are grouped in overlapping blocks (outlined in green and red), the block size shown in the figure is 2x2 at which the block has 2 cells in each direction the overlap ratio between blocks is equals to 50 % where 2 cells (cell2 and cell5) contribute in the result of both highlighted blocks.

The second step which is the main concern in this paper is computing the image gradient. It was mentioned in [18] that various gradient masks were compared but a combination of different gradient masks, which is proposed by this paper is not mentioned. In [18], the approach was used with application on pedestrians which are visually very different from vehicles. After calculating the gradient, a histogram of the orientation of the gradient is done. Number of bins in the histogram is also a user parameter. The Third step is calculating weighted vote using tri-linear interpolation method at which each pixel not only contributes in the results of the cell where it's located but also to the nearby cells and nearby bins in the histogram. The weight of contribution is determined according to the spatial location of the cell as well as the gradient value. As shown in Fig. 2, the center pixels denoted by $(x_0, y_0), (x_0, y_1), (x_1, y_0), (x_1, y_1)$ have contribution in one another results according to their location, and also they contribute in the results of nearby orientation bins according to their gradient orientation values. The fourth step is the normalization of results over overlapping spatial blocks and the last step is to collect the HOG's of the detection window to form the feature vector representing this window.

3.2 Compass HOG

In the conventional HOG, the gradient of the image calculated using the intensity change in two directions only (vertical and horizontal). This will result in losing data from the image and make the feature less discriminative because it ignores the intensity diagonal changes. On the other side, the proposed compass HOG uses gradient calculated from the change in all compass eight directions, which make it more descriptive for the image and hence more discriminative feature. Compass gradient is used to create three different HOG variants, the first one is calculated by concatenating the HOG features calculated using the four-compass gradient.

$$compHOG = [HOG1\ HOG2\ HOG3\ HOG4]$$

At which $HOG1$, $HOG2$, $HOG3$, and $HOG4$ are the HOG features calculated using the gradient in four directions described in Fig. 3. The feature vector in this HOG variant is four times the length on the conventional HOG. Other HOG variant is formed by averaging the gradients in four compass directions. The third HOG variant is formed by taking the direction of the gradient that has the maximum magnitude. The latter two HOG variants have the same vector length of the conventional HOG.

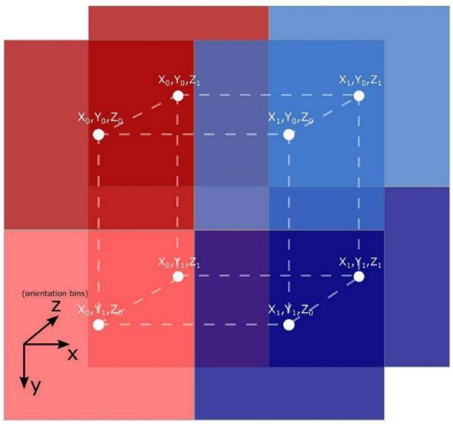

Figure 2: White points represent the cells center contributes in the tri-linear interpolation, Z is the dimension of orientation bins, the pixels contributes to the results of each cell according to its spatial location and gradient value.

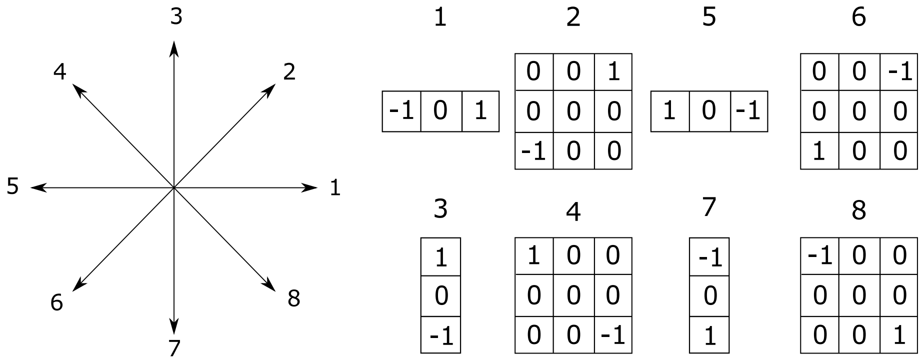

Figure 3: Filters used in calculating gradient in four directions

3.3 Experiments

3.3.1 KITTI Vision Benchmark [28]

In experimental work the KITTI dataset was used in training and testing. Introduced by Geiger et al. [29] in 2012 KITTI dataset aimed to bridge the gap between lab produced datasets and real life problem. KITTI dataset contains realistic data collected by driving around a city, so it has most of the challenging problems addressed by autonomous driving vision system. KITTI data acquisition setup is a car equipped with two sets of high resolution stereo cameras (Color set and grayscale set) and a GPS/IMU system that is used for accurate vehicle localization.

The dataset contains a visual odometry sequence of length 39.2 Km also it has 7481 stereo pairs for training with and 7518 stereo pairs for testing. The training images come with around 200 thousand labeled objects of different classes such as cars, vans, pedestrians and cyclists. The dataset could be used in a lot of application such as stereo matching, optical flow estimation, 3D visual odometry and 2D/3D object detection and pose estimation. That makes KITTI dataset one of the main datasets for testing state-of-the-art work related to vision systems for autonomous vehicles. The vehicles in the dataset are classified according to their detection difficulty into three levels, the vehicle detection difficulty is determined based on the size of the vehicle in the frame, the occlusion percentage and the truncation from the frame. The proposed approach is tested on the fully visible sedan vehicles.

3.3.2 Samples preparation

To test the proposed method, the training frames provided by the KITTI dataset is divided equally into two groups, one for training and the other for testing. Using the ground truth of the dataset, fully visible sedan vehicles extracted from the training and testing groups to be used as positive training and verification samples respectively. Negative samples are extracted randomly from the frame parts that have zero overlap with all vehicles in the frame. Fig. 4 shows positive samples.

3.3.3 Feature extraction

The extracted samples are then preprocessed before feature extraction. The sample is resized so that its height becomes one of the fixed heights which are 32, 64 and 128 pixels. The height is chosen based on the original size of the sample, the width is then determined so that the sample maintain the average aspect ratio. The next step is calculating the different HOG feature variations for each sample, the cell size is changed for each sample size so that all the samples have the same feature vector length.

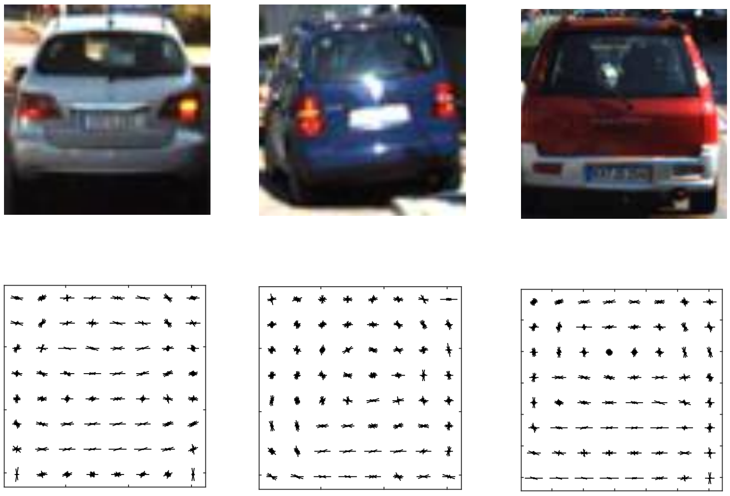

Figure 4: Positive samples for vehicles (above) and their HOG representation (below)

3.3.4 SVM classifier training

Linear SVM and nonlinear SVM classifiers are trained for car detection. Gaussian kernel was used as a kernel for the nonlinear SVM. The numbers of samples used are 2382 and 10226 for positive and negative samples respectively.

3.3.5 Testing using patches

Linear and nonlinear SVM classifiers based on different HOG variants are trained for detecting vehicle rear view and tested using 1657 positive test sample and 10113 negative samples. ROC curve is used to compare the performance of classifiers; the curve was plotted by choosing different thresholds for the classifiers confidence level. As shown in the ROC curve (Figure 5) both classifiers show similar relative behavior between the different HOG variants however all the nonlinear classifiers outperform their corresponding linear classifiers. All classifiers based on compass HOG performance exceeds that of the convenient HOG, the compass HOG classifier shows the best performance followed by the compass HOG with maximum gradient selection. The best performance achieved by nonlinear SVM based on compass HOG classifier shows a true positive rate of 89.56 % and true negative rate of 92.55 % with AUC equals to 0.9814.

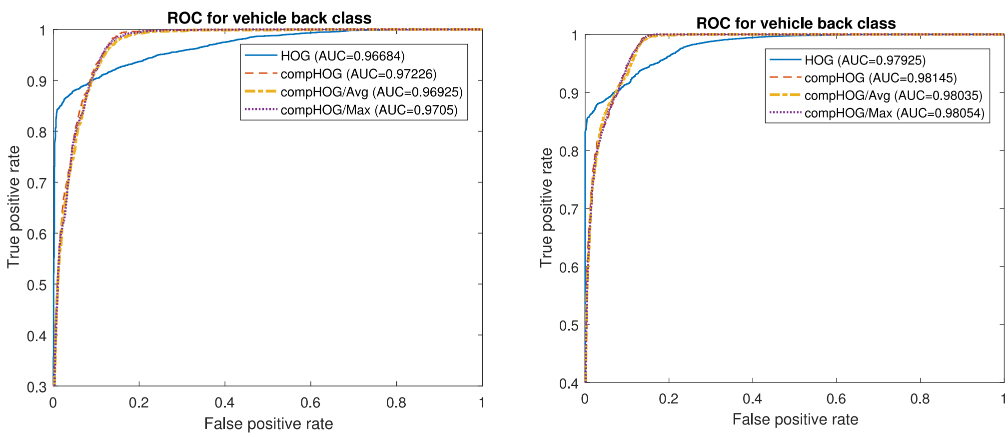

Figure 5: ROC curve for compared classifiers in the vehicle back class Linear SVM (Left), Nonlinear SVM (Right)

3.3.6 Vehicle detection

In this section, the classifiers which give the best performance in each class are used together to apply vehicle detection on frames provided by KITTI dataset. The exhaustive search technique is used to provide vehicle candidate for the classifiers, a sliding window with different sizes scan the image frame, the sliding windows height used in detection range from 40 to 180 with average step of 20 pixel, the window width is calculated according to aspect ratio associated with each classifier. The next step is to calculate compass HOG feature vector and provide it as an input for SVM classifiers which check if the window has a vehicle or not. The dimensions and locations of the detected vehicles from the three classifiers are then recorded. A non-maximum suppression process is done to reduce overlapping detected windows to only one window, the suppression process is done according to the confidence level of the classifier at each window, the window with the maximum confidence level is left while other windows that overlap with it with an overlap percentage higher than certain threshold are suppressed.

Samples for vehicle detection are shown in figure 6, the results show that the detection algorithm is effective in detecting vehicles with near and medium distance range which is the range covered by the sliding windows used in detection, also the detector shows robustness in vehicle detection of different shapes, colors and at different lighting conditions.

4.Conclusion and Future Direction

We have presented a new HOG based vehicle detection approach. We modified the feature by adding the derivatives from the diagonal direction. A SVM classifier was involved to detect vehicles in different video frames. We demonstrated an exhaustive results and validation section. The experiments showed the success of the proposed compass HOG approach against the conventional one. In the future, we will consider GPU implementation of the proposed technique as well as involving an object tracking approach to achieve real time performance.

Figure 6: Results for vehicle detection tested on KITTI image frames, true positive detection (Green), false positive detection (Red)

References

[1] ITF, "Road Safety Annual Report 2016."

[2] J. Diaz Alonso, E. Ros Vidal, A. Rotter, and M. Muhlenberg, "Lane-Change Decision Aid System Based on Motion-Driven Vehicle Tracking," *IEEE Trans. Veh. Technol.*, vol. 57, no. 5, pp. 2736–2746, Sep. 2008.

[3] A. Jazayeri, H. Cai, J. Y. Zheng, and M. Tuceryan, "Vehicle Detection and Tracking in Car Video Based on Motion Model," *IEEE Trans. Intell. Transp. Syst.*, vol. 12, no. 2, pp. 583–595, Jun. 2011.

[4] Y. Zhu, D. Comaniciu, M. Pellkofer, and T. Koehler, "Reliable Detection of Overtaking Vehicles Using Robust Information Fusion," *IEEE Trans. Intell. Transp. Syst.*, vol. 7, no. 4, pp. 401–414, Dec. 2006.

[5] Junxian Wang, G. Bebis, and R. Miller, "Overtaking Vehicle Detection Using Dynamic and Quasi-Static Background Modeling," 2005, vol. 3, pp. 64–64.

[6] S. Sivaraman and M. M. Trivedi, "Looking at Vehicles on the Road: A Survey of Vision-Based Vehicle Detection, Tracking, and Behavior Analysis," *IEEE Trans. Intell. Transp. Syst.*, vol. 14, no. 4, pp. 1773–1795, Dec. 2013.

[7] B. Dai, Y. Fu, and T. Wu, "A Vehicle Detection Method via Symmetry in Multi-Scale Windows," 2007, pp. 1827–1831.

[8] C. Tzomakas and W. von Seelen, "Vehicle Detection in Traffic Scenes Using Shadows," IR-INI, INSTITUT FUR NUEROINFORMATIK, RUHR-UNIVERSITAT, 1998.

[9] C. Hoffmann, "Fusing multiple 2D visual features for vehicle detection," 2006, pp. 406–411.

[10] N. Srinivasa, "Vision-based vehicle detection and tracking method for forward collision warning in automobiles," 2003, vol. 2, pp. 626–631.

[11] Y.-M. Chan, S.-S. Huang, L.-C. Fu, and P.-Y. Hsiao, "Vehicle Detection under Various Lighting Conditions by Incorporating Particle Filter," 2007, pp. 534–539.

[12] P. Viola and M. Jones, "Rapid object detection using a boosted cascade of simple features," 2001, vol. 1, p. I-511-I-518.

[13] Wen-Chung Chang and Chih-Wei Cho, "Online Boosting for Vehicle Detection," *IEEE Trans. Syst. Man Cybern. Part B Cybern.*, vol. 40, no. 3, pp. 892–902, Jun. 2010.

[14] D.-Y. Chen, G.-R. Chen, and Y.-W. Wang, "Real-time dynamic vehicle detection on resource-limited mobile platform," *IET Comput. Vis.*, vol. 7, no. 2, pp. 81–89, Apr. 2013.

[15] M. Gabb, O. Lohlein, R. Wagner, A. Westenberger, M. Fritzsche, and K. Dietmayer, "High-performance on-road vehicle detection in monocular images," 2013, pp. 336–341.

[16] S. Sivaraman and M. M. Trivedi, "Vehicle Detection by Independent Parts for Urban Driver Assistance," *IEEE Trans. Intell. Transp. Syst.*, vol. 14, no. 4, pp. 1597–1608, Dec. 2013.

[17] Zehang Sun, G. Bebis, and R. Miller, "On-road vehicle detection: a review," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 28, no. 5, pp. 694–711, May 2006.

[18] N. Dalal and B. Triggs, "Histograms of oriented gradients for human detection," in *2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)*, 2005, vol. 1, pp. 886–893 vol. 1.

[19] L. Mao, M. Xie, Y. Huang, and Y. Zhang, "Preceding vehicle detection using Histograms of Oriented Gradients," in *2010 International Conference on Communications, Circuits and Systems (ICCCAS)*, 2010, pp. 354–358.

[20] J. Arróspide, L. Salgado, and M. Camplani, "Image-based on-road vehicle detection using cost-effective Histograms of Oriented Gradients," *J. Vis. Commun. Image Represent.*, vol. 24, no. 7, pp. 1182–1190, Oct. 2013.

[21] G. Yan, M. Yu, Y. Yu, and L. Fan, "Real-time vehicle detection using histograms of oriented gradients and AdaBoost classification," *Opt. - Int. J. Light Electron Opt.*, vol. 127, no. 19, pp. 7941–7951, Oct. 2016.

[22] J. Zhang, K. Huang, Y. Yu, and T. Tan, "Boosted local structured HOG-LBP for object localization," in *CVPR 2011*, 2011, pp. 1393–1400.

[23] M. Cheon, W. Lee, C. Yoon, and M. Park, "Vision-Based Vehicle Detection System With Consideration of the Detecting Location," *IEEE Trans. Intell. Transp. Syst.*, vol. 13, no. 3, pp. 1243–1252, Sep. 2012.

[24] J. Kim, J. Baek, and E. Kim, “A Novel On-Road Vehicle Detection Method Using πHOG,” *IEEE Trans. Intell. Transp. Syst.*, vol. 16, no. 6, pp. 3414–3429, Dec. 2015.
[25] P. H. Kassani, J. Hyun, and E. Kim, “Application of soft Histogram of Oriented Gradient on traffic sign detection,” in *2016 13th International Conference on Ubiquitous Robots and Ambient Intelligence (URAI)*, 2016, pp. 388–392.
[26] C. Cortes and V. Vapnik, “Support-Vector Networks,” *Mach. Learn.*, vol. 20, no. 3, pp. 273–297, Sep. 1995.
[27] H. T. Niknejad, A. Takeuchi, S. Mita, and D. McAllester, “On-Road Multivehicle Tracking Using Deformable Object Model and Particle Filter With Improved Likelihood Estimation,” *IEEE Trans. Intell. Transp. Syst.*, vol. 13, no. 2, pp. 748–758, Jun. 2012.
[28] J. Fritsch, T. Kuehnl, and A. Geiger, “A New Performance Measure and Evaluation Benchmark for Road Detection Algorithms,” in *International Conference on Intelligent Transportation Systems (ITSC)*, 2013.
[29] A. Geiger, P. Lenz, and R. Urtasun, “Are we ready for autonomous driving? The KITTI vision benchmark suite,” in *2012 IEEE Conference on Computer Vision and Pattern Recognition*, 2012, pp. 3354–3361.

Authors Biography

Yehia Zakaria: He received his B.Sc. in Mechatronics from faculty of Engineering Ain shams University in 2012. He is currently a teaching assistant and master student in the faculty. His field of research in computer vision and autonomous systems.

Dr. Hossam E. Abd El Munim: He received the B.Sc. and M.S. degrees in electrical engineering from Ain Shams University, Egypt, in 1995 and 2000, respectively. He received his Ph.D. degree in Electrical and Computer Engineering from the University of Louisville, Louisville, Kentucky. He joined the Computer Vision and Image Processing Laboratory (CVIP Lab) at the University of Louisville on June 2002, where he has been involved in the applications of image processing and computer vision for medical image analysis. He is now an Associate Professor at the Computer and Systems Engineering Department, Faculty of Engineering, Ain Shams University in Cairo, Egypt. His current research interests include image modeling, image segmentation, 2D and 3D registration, visualization, and surgical simulation, about which he has authored or co-authored more than 30 technical articles including ICIP, MICCAI, ICCV, CVPR, BMVC, and TIP, TPAMI. He received an excellence in visualization award from Silicon Graphics in 2004, a US National Science Foundation travel grant to attend the International Conference on Computer Vision (ICCV) on 2005, and a first place in the University of Louisville Engineering Competition (E-Expo) in 2006. Also, he was awarded the Outstanding ECE Graduate Student Award in the University of Louisville, 2007.

Dr. Maged Ghoneima is an Assistant Professor leading the Mechatronics Engineering Department at Ain Shams University. He graduated with honors from the Faculty of Engineering – Ain Shams University in 1997. He obtained his PhD from Northwestern University in Chicago in 2001. After that he worked in the Electronics design sector in leading companies like Intel Corporation and Nvidia Corporation. He has more than 10 issued and pending patents, as well as more than 40 conference and journal publications.

Professor Dr. Sherif Ali M. Hammad graduated from the Faculty of Engineering – Ain Shams University in 1983. He obtained his master and PhD degrees from Ain sham University in 1987 and 1992 respectively. He is a professor in department of computer and systems. He was the dean of faculty of Engineering – Ain shams University from 2012 to 2014 and He is the former minister of scientific research in Egypt.