Original Research Paper

Detection of a Moving Car Based on Invariant Moments

Abdul Monem Rahma and Nabaa Basil Abd

Department of Computer Science, University of Technology, Baghdad, Iraq

Abstract: This paper presents a proposed technique based on Invariant Moments to detect moving cars in front of Autonomous Cars in order to increase safety and reduce road accidents, thus saving lives, which are one of the most important matters in the context of Autonomous Cars. Object detection from a sequence of images is preparatory step but a crucial mission for computer vision applications in information extraction. In this work, a car database was constructed and used in background subtraction to define Region of Interest, then applying the invariant moments for RoI and dependence on the information of the database that was constructed to detect and recognize the moving car. The technique was tested and the experimental results showed that it could detect the moving objects successfully with detection rate over 87% and precision 97% and FoM 91%.

Keywords: Autonomous Car, Moving Object Detection, Background Subtraction, Invariant Moments

Introduction

The issue of raising road safety and preventing road traffic injuries has become worldwide. In addition, with the number of accidents increasing day by day, it has become important to take over the human errors. All of this could come to an end with Autonomous Cars. An Autonomous Car or a driverless car (sometimes called a self-driving car) is a robotic vehicle, which through feedback returned by various sensors operate independently. The purpose of these vehicles reduces risks, problems and costs that emerged from human intervention. It is designed to travel between destinations without a human operator. So, one of the major functions of Autonomous Car is the detection of moving object (cars) in front of it by using computer vision technologies. Even after research for several years, detection and tracking moving object is still an open research issue. Until today, it is still a great challenge to achieve an accurate, robust and high performance approach. How can one define the object to be detected and tracked is the difficulty level of this problem Shaikh *et al.* (2014), which help to make the decision for determining the path of Autonomous Cars and avoid accidents.

Detection Moving Object (Related work)

Detection of a moving object is a critical process in computer vision application, especially in video surveillance and Autonomous Car. Insignificant information in the scene is ignored in this process and focus attention to moving objects. To achieve this target through the past years, many methods and algorithms and techniques have been proposed either based on a predictive or probabilistic mechanism (Wang *et al.*, 2011; Warnell *et al.*, 2012; Mansour and Vetro, 2014; Thangarajah *et al.*, 2016). Optical flow, frame difference and background subtraction are the prime three methods of detecting moving objects in addition to many other methods using sophisticated techniques in combination with those basic ones.

Optical Flow

The change in optical flows properties by motion objects over time was exploited by Optical flow and it is suitable for the dynamic and static background. However, because of the poor anti-noise performance and its complex computation so, for real time processing, it requires a special hardware Yang *et al.* (2012).

Frame Difference Model (FDM)

In frame difference, or time difference, the moving regions are extracted by threshold of time difference in adjacent frames' pixels. The interest of this method is its rapid background update, good adaptive performance and the insensitive to the variation of light, but the disadvantage of this method that it cannot detect the moving objects which has conformable inter color and big size Yang *et al.* (2012).

Background Subtraction Model (BSM)

BSM is a method used to find movable object and to achieve this, the current frame image differentiating with background image, if the result is greater than the threshold. In this case, the pixels of the current image are moving pixels, otherwise the pixels are not moving. To define the moving objects, this model is completely relied on the background image. The correct constructing of background image determines slow movement of the objects and temporarily motionless objects. The resultant image of background subtraction contains noise if the background image does not contain maximum number of stationary pixels and it is not able to detect regions of objects properly (Hossain and Das, 2014).

Adaptive Background Subtraction Model (ABSM)

This method is similar to background subtraction but here, background image is adaptively updated over time. In this method, transforming the colored video to grey scale video is not necessary. Initially, creation background image then is modified according to the surrounding environment changes with the assist of learning rate (a). This method takes too much time to modify the background from time to time (Hossain and Das, 2014).

In addition, recent suggestions are using sophisticated techniques on the background subtraction in combination with those basic ones. To extract backgrounds from moving points, several methods are using advanced statistic models (Huang and Chen, 2013a; 2013b; Cheng et al., 2015; Guo et al., 2013). However, others proposed using neural networks or outlier detection models (Zhou et al., 2012; Huang and Do, 2014). According to their main characteristics, Bouwmans (2014) categorized these subtraction models further into 17 groups.

Gaussian Mixture Models (GMM)

In this method, the components of background image are accumulated as terms of Gaussian Distribution Functions. A pixel from the input image is decided whether it is from foreground or background by the Gaussian distribution functions. This statistical determinant is effective for minor changes in the background moreover and using mixtures of such functions makes the method multimodal Dong-Sun and Jinsan (2016).

Thangarajah et al. (2016) presented a method to update the threshold of GMM based BGS with regard to color distortion, illumination measures in pixel level and similarity. A threshold was set automatically for moving objects detection in video sequences.

Srivastav et al. (2017) proposed a technique to detect object which is able to reduce the holes problem that result from two frame differencing based on the differencing of three frame and background subtraction.

In this work, database is constructed, which contains features for number of objects (cars) used for objects detection. Also it is based on background subtraction to define Region of Interest (RoI), then applying the invariant moments for RoI, lastly using Sum of Square Error (SSE) between the calculated invariant moments and the invariant moments that is stored in the database as features in order to detect and recognize the car. The period between one image and another in the same sequence (same video) is controlled by the user of the system.

The Proposed Technique for Moving Object Detection

In Autonomous Car, the front view is acquired (as a video) using a camera mounted on the Autonomous Car then applying few processes in order to detect the moving object (car). Any video is a consecutive of sequent images or frames from which object (car) can be detected. One of important functions in Autonomous Car is detecting and tracking the moving object (car) in front of it. In this work, it is assumed, that the Autonomous Car is in the middle of the road also that just one moving car appears in the front view and it can move freely to the left, right, up and down according to traffic laws. Object detection in video sequence refers to the process of separating frame into background and moving objects according to features such as color, intensity, edge or motion, to locate and identify objects and estimates its velocity and location in frame.

This section describes all steps to detect and recognize the moving car, the method comprises four main tasks (construct car database, define region of interest and applying invariant moments to each regions of interest compute SSE).

Constructing Cars Database

Preparing a database containing images of specific roads has been done. These roads are recorded when they are empty. The images of empty roads are used for background subtraction in order to extract the area of interest as in (3.2). The database also contains features for several targets (cars) which represent objects characteristics such as invariant moments for (binary, grey, green) images for the same car, using these characteristics to detect and recognize the car.

Define Region of Interest

Define Region of Interest by using Background Subtraction between two images has been described. Two images used for subtraction process are defined as background image and foreground image. Background image is retrieved from database where it is stored and the foreground image from the sequence of images that obtained from the camera and the user controls the period of time between one image and another to the same sequence.

The colored image spatial information consists of color information stored in three various components (color channels) for each pixel, which considered as coordinates in some color space. Most of the present methods for background subtraction convert the spatial color information of background and foreground images to grey image of 0-255 intensity levels. The eventual grey image of background subtraction is called Differenced Image (DI) and it is given in Equation 1:

$$S(x,y) = |B(x,y) - F(x,y)| \tag{1}$$

Where, $S(x, y)$ refers to pixel intensity of ID at x^{th} row and y^{th} column of grey image. $B(x, y)$ and $F(x, y)$ refer to the pixels intensity of xth row and y^{th} column of background and foreground images respectively.

By setting the threshold Th, Binary image is obtained (converting the image to black and white) as in Equation 2, thereby extracting the moving object region from the image:

$$BI(x,y) = \begin{cases} 1 & if & S(x,y) > Th \\ 0 & otherwise \end{cases} \tag{2}$$

Where, $BI(x, y)$ refers to pixel intensity of xth row and y^{th} column of binary image.

Also, Converting the Binary image from Equation 2 to Green image $GR(x, y)$ of 0-255 intensity by using the spatial coordinate of non-zero value of the Binary image and the green component (channel) with the same spatial coordinate from foreground image as in Equation 3:

$$GR(x,y) \begin{cases} F(x,y)_{green} & if & BI(x,y) = 1 \\ 0 & & otherwise \end{cases} \tag{3}$$

Where, $GR(x, y)$ refers to pixel intensity of the xth row and y^{th} column of Green image and $F(x, y)_{green}$ denotes pixel intensity of the xth row and y^{th} column of the foreground image with just green component.

Applying the Invariant Moments

Pixel intensity is nothing but the pixel color value. Image moments can simply be described as some functions of the image pixel intensity. Moments are described with respect to their power as in raised to the power in mathematics.

The geometrical image features and description of object shape are represented by using a set of moments. Via moment functions, center of mass, area of an image and orientation information can be found. From geometric moments the properties of an image can be generated Ong *et al.* (2014).

If an image describing by 2D discrete intensity function $I(x, y)$ with non-zero values in the finite part of XOY plane then the geometrical moments of all orders $(p+q)$ are presented by Equation 4, Favorskaya *et al.* (2013):

$$m_{pq} = \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} x^p y^q I(x,y) \tag{4}$$

Where, M and N are the image dimensions and x, y are the region coordinates (pixel coordinates in digitize images).

Calculation of central moments of order $(p + q)$ is provided by Equation 5:

$$\mu_{pq} = \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} (x - \bar{x})^p (y - \bar{y})^q I(x,y) \tag{5}$$

Where, μ_{pq} denotes the central moment and $\bar{x}$, $\bar{y}$ is the gravity center coordinates (centroid) of object image, which are invariant to translations, which can be obtained using the Equation 6, Favorskaya *et al.* (2013):

$$\bar{x} = \frac{m_{10}}{m_{00}}, \bar{y} = \frac{m_{01}}{m_{00}} \tag{6}$$

Scale invariance is achieved by normalizations of moments according to Equation 7, where $\gamma = \frac{p+q}{2} + 2$:

$$\eta_{pq} = \frac{\mu_{pq}}{\mu_{pq}^{\gamma}} \tag{7}$$

It is known that seven non – linear functions φ_1 - φ_7 of normalized invariant Hu moments Favorskaya *et al.* (2013) are calculated by Equation 8:

$$
\begin{aligned}
\varphi_1 &= \eta_{20} - \eta_{02} \\
\varphi_2 &= (\eta_{20} - \eta_{02})^2 + 4\eta_{11}^2 \\
\varphi_3 &= (\eta_{30} - 3\eta_{12})^2 + (3\eta_{21} - \eta_{03})^2 \\
\varphi_4 &= (\eta_{30} + \eta_{12})^2 + (\eta_{21} + \eta_{03})^2 \\
\varphi_5 &= (\eta_{30} - 3\eta_{12})(\eta_{30} + \eta_{12})[(\eta_{30} + \eta_{12}) - 3(\eta_{21} + \eta_{03})^2] \\
&\quad + (3\eta_{21} - \eta_{03})(\eta_{21} + \eta_{03})[3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2] \\
\varphi_6 &= (\eta_{20} - \eta_{02})[(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2] \\
&\quad + 4\eta_{11}(\eta_{30} + \eta_{12}) - (\eta_{21} + \eta_{03}) \\
\varphi_7 &= (3\eta_{21} - \eta_{03})(\eta_{30} + \eta_{12})[(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2] \\
&\quad - (\eta_{30} - 3\eta_{12})(\eta_{21} + \eta_{03})[3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2]
\end{aligned}
\tag{8}
$$

Equation 8 is applied to the Grey image $S(x, y)$ and then to the Binary image $BI(x, y)$ and then to the Green image $GR(x, y)$ separately.

Computing the Sum of Square Error (SSE)

After applying the invariant moments for the object (car) in the image, we used the value of seven invariant moments and Sum of Square Error (SSE) to detect and recognize the car according to the information that stored in the cars database.

SSE is a measure of discrepancy between the calculated invariant moments in 3.3 and the invariant moments that are stored in the database for specific objects (cars). A small RSS indicates a tight fit of the object in the database to the object that extracted from the sequence of images. It is used as an optimality criterion in object selection. SSE is given in the Equation 9:

$$SSE = \sum_{i=1}^{n} (\varphi_i - \overline{\varphi_i})^2 \qquad (9)$$

where, $n = 1, \ldots, 7$ and SSE is the Sum of Square Error and φ_i is the seventh invariant moments calculated for the input image and $\overline{\varphi_i}$ is the seventh invariant moments which is stored in the database for specific car.

SSE is applied for Grey, Binary and Green images separately. Using the result of SSE on one of the three types (Grey, Binary and Green) of images to detect and recognize the car and we can use the results of SSE on the three types together to detect and recognize the car.

Case Studies and Experimental Results

The described technique above was implemented by using Visual Studio 2012 using C sharp programming language and used camera mounted on the car, which obtained the front view as a sequence of images. The system has been tested using a set of images up to (40) as samples, Fig. (1) And (2) show the steps of proposed technique for two samples taken from a two sequences of images. The evaluation of the proposed technique was tested using three performance matrices, namely Precision, Recall and Figure of Merit (FoM). These metrics are based on the following parameters, False Positive (FP), True Positive (TP) and False Negative (FN).

The Recall or Detection Rate: given by Equation 10 measures the percentage of predicted True Positive (TP)

as compared to the total number of actual positives in the ground truth:

$$Recall = TP / (TP + FN) \qquad (10)$$

The Precision: defined by Equation 11 measures the percentage of correct detection as compared to the total number of detection as positives:

$$Precision = TP / (TP + FP) \qquad (11)$$

Precision and Recall are very different perspective measures of the performance. Thus, a weighted harmonic mean measured jointly with recall and precision called Figure of Merit (FoM) provides a better performance evaluation defined by Equation 12:

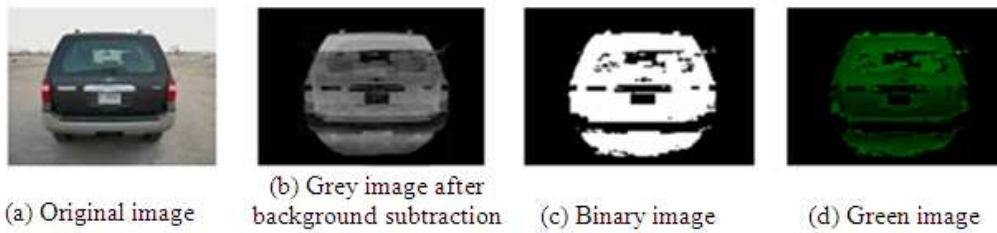$$FoM = (2 \times Recall \times Precision) / (Recall + Precision) \qquad (12)$$

Table (1) shows the performance of the proposed technique and other algorithms based on the above measures.

Figure 1 and 2 illustrate the steps of implementation of the proposed technique for two different cars. Image (a) is the original image derived from a series of images, which are the input images to the system. Image (b) is grey image resulted from background subtraction. Image (c) and (d) are the result from convert image (b) to binary and green respectively. Images (e) In Figure 1 and 2 Illustrate the applying of the invariant moments on images (b) (c) and (d) separately. Image (f) illustrates the result from computing SSE between the values from (e) in Fig. 1. and Fig. 2. and the invariant moments that was stored in the database to detect and recognize moving car as in (g)

Table (1) illustrates the performance comparison between the proposed technique and four other techniques based on Recall, Precision and Figure of Merit (FoM) matrices, where they obtained result show that the proposed technique has Recall, Precision and FoM matrices values higher than the other existing models, so that the proposed technique more robust than the other compared algorithms in terms of FoM.

Table 1: Performance comparison on various dataset for several algorithms

Algorithms	Recall	Precision	FoM
FDM	0.87	0.95	0.89
BSM	0.81	0.91	0.85
ABSM	0.75	0.93	0.84
GMM	0.76	0.76	0.76
Proposed	0.87	0.97	0.91

(a) Original image

(b) Grey image after background subtraction

(c) Binary image

(d) Green image

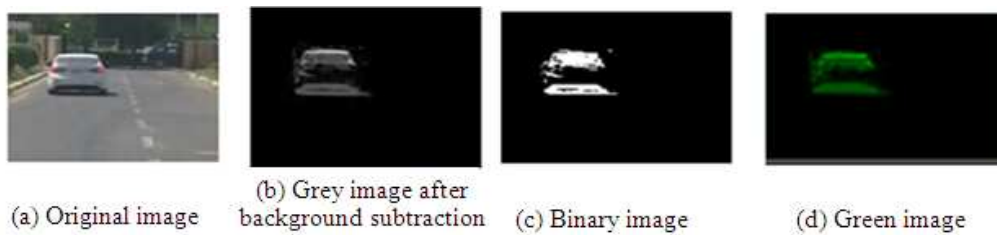| images	M1	M2	M3	M4	M5	M6	M7
Binary		0.000970897439	8.279829E-09	1.57690494E-09	1.61866354E-10	-6.73243141E-20	1.12014136E-14	4.64231984E-20
Grey		0.00232327217	5.126965E-08	3.36314869E-08	2.32411E-09	-1.514831E-17	4.50530157E-13	1.38826E-17
Green		0.00254244637	4.88498273E-08	4.758048E-08	3.513788E-09	-3.15711977E-17	6.263029E-13	3.26722677E-17

(e) Applying the invariant moment on (a) (b) (c) images

checking result

Binary Table	Thr Car Id is 10
Gray Table	Thr Car Id is 10
Green Table	Thr Car Id is 10

(f) Result of computing SSE

(g) Detection and recognition moving car

Fig. 1: Steps of the proposed technique for detection object (car1)

(a) Original image

(b) Grey image after background subtraction

(c) Binary image

(d) Green image

images	M1	M2	M3	M4	M5	M6	M7
Binary		0.010662701	1.30842282E-05	2.882645E-06	1.38536791E-06	-2.42548624E-12	4.99294828E-09	1.334753E-12
Grey		0.0214989558	5.32171325E-05	2.2426324E-05	1.472595E-05	-1.97186767E-10	1.07223087E-07	1.80922763E-10
Green		0.0214989558	5.32171325E-05	2.2426324E-05	1.472595E-05	-1.97186767E-10	1.07223087E-07	1.80922763E-10

(e) Applying the invariant moment on (a) (b) (c) images

checking result

Binary Table	Thr Car Id is 12
Gray Table	Thr Car Id is 12
Green Table	Thr Car Id is 12

(f) Result of computing SSE

(g) Detection and recognition moving car

Fig. 2: Steps of the proposed technique for detection object (car2)

314

Conclusion and Future Work

A new technique was introduced in this study by using computer vision techniques and using camera mounted on a car to detect and recognize one moving object in front of the car and gives information to the driver and alerts about the cars in front view. The proposed technique exploited the invariant moments to detect and track moving car. The results proved that proposed technique succeeded to detect and track object in the dataset used in the case study. Hence, the evaluation of the performance shows that the proposed technique more robust than the other compared algorithms in terms of FoM. By using this technique, it is possible to estimate the location of the car that in the front view and thus estimate its velocity based on addition information such as the velocity of the car that load the camera, thus avoid collision and traffic accidents. In addition, this technique can be updated by adding more objects to the database to detect and recognize any object and many types of vehicles.

Future work can be dedicated to detect more than one car and it will be implemented completely in the Autonomous Car.

Acknowledgement

The authors would like to thank Department of Computer Sciences of the University of Technology and for supporting this study.

Author's Contributions

Abdul Monem Rahma: Is contributed in the theoretical aspect of this work and writing this paper for publication purpose.

Nabaa Basil Abd: Is contributed in the theoretical aspect of this work in addition to the practical aspect and writing this paper for publication purpose.

Ethics

The authors confirm that this manuscript has not been published elsewhere and that no ethical issues are involved.

References

Bouwmans, T., 2014. Traditional and recent approaches in background modeling for foreground detection: An overview. Comput. Sci. Rev., 11-12: 31-66. DOI: 10.1016/j.cosrev.2014.04.001

Cheng, F.C., S.C. Huang and B.H. Chen, 2015. A hybrid background subtraction method with background and foreground candidates' detection. ACM Trans. Intelligent Syst. Technol., 7: 1-14. DOI: 10.1145/2746409

Favorskaya, M., D. Pyankov and A. Popov, 2013. Motion estimations based on invariant moments for frames interpolation in stereovision. Procedia Comput. Sci., 22: 1102-1111. DOI: 10.1016/j.procs.2013.09.196

Guo, J.M., C.H. Hsia, Y.F. Liu, M.H. Shih and C.H. Chang et al., 2013. Fast background subtraction based on a multilayer codebook model for moving object detection. IEEE Trans. Circuits Syst. Video Technol., 23: 1809-1821. DOI: 10.1109/TCSVT.2013.2269011

Hossain, W. and M.N Das, 2014. Moving object detection in dynamic backgrounds for surveillance system. Proceedings of IEEE International Conference on Advanced Communication Control and Computing Technologies, May 8-10, IEEE Xplore Press, India, pp: 1476-1479. DOI: 10.1109/ICACCCT.2014.7019348

Huang, S.C. and B.H. Chen, 2013a. Highly accurate moving object detection in variable-bit-rate video-based traffic monitoring systems. IEEE Trans. Neural Networks Learning Syst., 24: 1920-1931. DOI: 10.1109/TNNLS.2013.2270314

Huang, S.C. and B.H. Chen, 2013b. Automatic moving object extraction through a real-world variable-bandwidth network for traffic monitoring systems. IEEE Trans. Industrial Electronics, 61: 2099-2112. DOI: 10.1109/TIE.2013.2262764.

Huang, S.C. and B.H. Do, 2014. Radial basis function based neural network for motion detection in dynamic scenes. IEEE Trans. Cybernetics, 44: 114-125. DOI: 10.1109/TCYB.2013.2248057

Dong-Sun, K. and K. Jinsan, 2016. Moving object detection on a vehicle mounted back-up camera. Sensors. DOI: 10.3390/s16010023

Mansour, H. and A. Vetro, 2014. Video background subtraction using semi-supervised robust matrix completion. Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing, May 4-9, IEEE Xplore Press, Italy, pp: 6528-6532. DOI: 10.1109/ICASSP.2014.6854862

Ong, L.Y., S.H. Lau and V.C. Koo, 2014. A new approach of local feature descriptors using moment invariants. J. Comput. Sci., 10: 2538-2547. DOI: 10.3844/jcssp.2014.2538.2547

Shaikh, S.H., S. Khalid and C. Nabendu, 2014. Moving object detection using background subtraction. Springer Briefs Comput. Sci., 67: 32-32. DOI: 10.1007/978-3-319-07386-6_1

Srivastav, N., L.A. Shubh, K.G. Sandeep, R.S. Saurabh and C. Blessy et al., 2017. Hybrid object detection using improved three frame differencing and background subtraction. Proceedings of 7th International Conference on Cloud Computing, Data Science and Engineering - Confluence, Jun. 12-13, IEEE Xplore Press, India, pp: 613-617. DOI: 10.1109/CONFLUENCE.2017.7943225

Thangarajah, A.Q.M., J. Wu and J. Huo, 2016. A unified threshold updating strategy for multivariate Gaussian mixture based moving object detection. Proceedings of the International Conference on High Performance Computing and Simulation, Jul. 18-22, IEEE Xplore Press, Austria, pp: 570-574. DOI: 10.1109/HPCSim.2016.7568386

Wang, S.C., T.F. Su and S.H. Lai, 2011. Detecting moving objects from dynamic background with shadow removal. Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing, May 22-27, IEEE Xplore Press, Czech Republic, pp: 925-928. DOI: 10.1109/ICASSP.2011.5946556

Warnell, G., D. Reddy and R. Chellappa, 2012. Adaptive rate compressive sensing for background subtraction. Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing, Mar. 25-30, IEEE Xplore Press, Japan, pp: 1477-1480. DOI: 10.1109/ICASSP.2012.6288170

Yang, J.B., M. Shi, Q.M. YI, 2012. A new method for motion target detection by background subtraction and update. Phys. Proc., 33: 1768-1775. DOI: 10.1016/j.phpro.2012.05.283

Zhou, X., C. Yang, W. Yu, 2012. Moving object detection by detecting contiguous outliers in the low-rank representation. IEEE Trans. Pattern Anal. Mach. Intell., 35: 597-610. DOI: 10.1109/TPAMI.2012.132