CrossMark

Driver aggressiveness detection via multisensory data fusion

Omurcan Kumtepe[1]* , Gozde Bozdagi Akar[1] and Enes Yuncu[2]

Abstract

Detection of driver aggressiveness is a significant method in terms of safe driving. Every year, a vast number of traffic accidents occur due to aggressive driving behaviour. These traffic accidents cause fatalities, severe disorders and huge economical cost. Therefore, detection of driver aggressiveness could help in reducing the number of traffic accidents by warning related authorities to take necessary precautions. In this work, a novel method is introduced in order to detect driver aggressiveness on vehicle. The proposed method is based on the fusion of visual and sensor features to characterize related driving session and to decide whether the session involves aggressive driving behaviour. Visual information is used to detect road lines and vehicle images, whereas sensor information provides data such as vehicle speed and engine speed. Both information is used to obtain feature vectors which represent a driving session. These feature vectors are obtained by modelling time series data by Gaussian distributions. An SVM classifier is utilized to classify the feature vectors in order for aggressiveness decision. The proposed system is tested by real traffic data, and it achieved an aggressive driving detection rate of 93.1 %.

Keywords: Driver aggressiveness, Driving behaviour, Road safety, Road line detection, Lane detection, Vehicle detection

1 Introduction

Traffic accidents has become an important problem in the last few decades due to increasing number of vehicles on the roads. Every year, 1.24 million fatalities occur due to traffic accidents globally [1]. Some of these traffic accidents are caused by physical reasons such as road and vehicle conditions. However, mostly, human factor is effective in the occurrence of traffic accidents. Among the human factors, aggressive driving behaviour constitutes a huge portion of traffic accident reasons. According to a report of the American Automobile Association Foundation for Traffic Safety, published in 2009, 56 % of traffic accidents occur due to aggressive driving behaviour [2]. Moreover, traffic accidents brings about billions of dollars of economical cost for people, governments and companies [1]. For these reasons, reduction of the number of traffic accidents is an important issue. Considering human factors, detection of aggressive driving behaviour could

help in reducing the number of traffic accidents by giving necessary warnings to drivers and related authorities.

Aggressive driving behaviour is defined as an action "when individuals commit a combination of moving traffic offences so as to endanger other persons or property" by the National Highway Traffic Safety Administration (NHTSA) [3]. Aggressive driving behaviour is a psychological concept that does not have a quantitative measure. However, there exist some certain behaviours associated with aggressive driving such as excess and dangerous speed, following the vehicle in front too closely, in other words tailgating, erratic or unsafe lane changes, improperly signalling lane changes and failure to obey traffic control devices (stop signs, yield signs, traffic signals, etc.) [3]. Also, in [4], it is stated that lane changing and acceleration are the characteristics of driving behaviours that define driving style. Therefore, detecting these behaviours and constituting features from these information can yield quantitative information about the driving style of the driver.

Although these behaviours are indication of driver aggressiveness, detection of these behaviours in real time is a challenging task. Existing methods in the literature

*Correspondence: omurcan@metu.edu.tr
[1] Department of Electrical and Electronics Engineering, Middle East Technical University, Dumlupınar Blv. No:1, 06800 Ankara, Turkey
Full list of author information is available at the end of the article

Kumtepe *et al. EURASIP Journal on Image and Video Processing*   (2016) 2016:5

Page 2 of 16

mostly based on driving simulator data which do not work for real-time aggressive driving behaviour detection and do not fully reflect the real-world conditions [5]. There also exist sensor platform-based methods in literature; however, these methods do not consider vehicle following distance and lane following pattern which are very significant for indicating driver aggressiveness. The proposed system enables detection of driver aggressiveness in real time by considering a wide range of aggressiveness-associated driving behaviours.

In this paper, we proposed an automated aggressive driving behaviour detection system that works in real time. The system performs robust operation with simple and low complexity algorithm in order to be able to work efficiently in real time. Multisensory information is used by this system in order to extract features that characterize the related driving session. The system collects data about lane following, vehicle following, speed and engine speed patterns which are important for aggressive driving detection since aggressive driving behaviour is associated with sudden lane changes, tailgating and abrupt acceleration/deceleration. Features that are extracted utilizing these data are used to train an SVM classifier. The classifier is trained with annotated data so that aggressiveness decision can be modelled regarding the subjective point of view, that is, aggressive driving behaviour, which is a subjective and psychological phenomenon, can be modelled quantitatively. The system uses different types of features and feature extraction methods that works in real time; therefore, the system can create a decision at the end of each session. Session length is a design parameter which will be discussed in test results.

The organization of this paper is as follows: The next part describes the related work about aggressive driving behaviour detection. It is followed by the proposed method description and its advantages and novelty. Then, the test results are presented and concluding remarks are given.

2   Related work

Aggressive driving behaviour detection has been examined via different approaches in recent years [5]. The simplest method for detecting aggressive driving behaviour is to conduct surveys about the driving experience or psychological mood. In literature, there exist some methods that are based on observing the behaviours of subjects in the simulator environment. In [4], subjects are requested to drive via a simulator with different scenarios which contain events such as traffic light existence, intersection crossing and frustrating environment. Then, the findings are illustrated by probabilistic models. Similarly Danaf et al. [6] use a simulator environment to collect data about driving behaviour and expresses anger (or aggressiveness) as a dynamic variable. Hamdar et al. [7] define and develop

a quantitative aggressiveness propensity index in order to model driving behaviour by testing its proposal with a driving simulator. The main drawback of these works is that they are using a synthetic environment to measure the driving behaviour. Therefore, they do not fully reflect the real-world conditions and reactions of driver in real traffic environment.

In order to acquire real world data, Gonzalez et al. [5] propose a sensor platform-based system to detect driver aggressiveness. Their method monitors external driving signals such as lateral and longitudinal accelerations and speed and models aggressiveness as a linear filter operating on these signals [5]. Johnson and Trivedi [8] use sensor data which is obtained by a smart phone in order to characterize the driving style. Kang [9] examines driver drowsiness and distraction by collecting visual information such as eye gaze and yawning and physiological data such as ECG signals.

Satzoda and Trivedi [10] use multisensory information in order to analyse the drive and certain driving events such as lane changes, mean speed, etc. However, no interpretation is given about the aggressiveness of driver. Jian-Qiang and Yi-Ying [11] present a dangerous driving behaviour detection scheme using a CCD camera to acquire visual information about driving behaviour and identifies dangerous driving style. Nevertheless, the system uses only visual information and tries to identify the driving with a few features. The work presented in [12] exploits a sensor and a camera platform to detect independent driving events such as lane departure, acceleration, zig-zag driving, etc. Then, it uses a fuzzy technique to indicate whether the driving is dangerous. Although the system shows good results for identifying different driving events, the presented work focuses on dangerous driving rather than aggressiveness and does not propose any technique to verify aggressiveness with subjective observations.

Besides the systems that are specialized on detecting driver aggressiveness, there also exist advanced driving assistance systems (ADAS) in literature. ADAS are very popular in recent years and used in order to provide assistance to the driver about the current driving conditions such as lane departure or forward collision possibility [13]. They are used for collecting data about the driving and for warning the driver by giving feedback about the driving behaviour. However, ADAS do not interpret the driving data to reach an aggressiveness conclusion.

3   Proposed method

As indicated in [3] and [4], aggressive driving is associated with certain behaviour such as sudden lane changes, tailgating behaviour, speed and acceleration basically. Therefore, aggressive driving behaviour can be identified by observing these events. In order to obtain quantitative

Kumtepe *et al. EURASIP Journal on Image and Video Processing* (2016) 2016:5

Page 3 of 16

measures of these events, lane following, vehicle following, speed and engine speed patterns of a driving session is collected and processed automatically. As a result of the process, four different feature types are obtained to represent the related driving session. Lane deviation and forward car distance are extracted as visual information and vehicle and engine speed as sensor information. Since the operation of the system is in real time, robust and algorithmically simple methods are used for extraction of the related information. These information are collected and feature vectors are retrieved. Obtained feature vectors are given to a pre-trained classifier to detect aggressive driving behaviour. The overall system flow can be seen in Fig. 1.

3.1 Road line detection

In order to find the position of the host vehicle, which is the equipped and examined vehicle, inside the road lane, road line detection is required. Drivers who change lanes suddenly and continuously and do not follow the lane properly may involve in aggressive driving attitude. Therefore, detecting the position of the host vehicle inside the lane by detecting the road lines is an important information. For road line detection problem, non-uniformity of road lines is the major challenge [14]. In order to accomplish road line detection task with a robust operation to non-uniformities in road lines, we used a method based on temporal filtering and inverse perspective mapping which is a robust, simple and low-cost method and proper for the real-time operation of the overall system. However, in order to decrease the computation load and satisfy the real time operation condition, we modelled road lines with straight lines

instead of curves which provides sufficient results for our application.

In recent years, many different techniques and studies are conducted on road line detection, mainly caused by the current interest in advanced driving assistance systems and autonomous driving systems. Road line detection algorithms in the literature mostly consist of two stages, preprocessing and detection stages. In preprocessing stage, different image processing techniques are used in order to provide enhanced data for detection task. Preprocessing methods in the literature can be exemplified as follows. Somasundaram et al. [15] use transformation from RGB colour space to HSV colour space for reducing redundancy. Morphological filtering is used in [16]. In [17], canny edge detector is used to indicate and emphasis road lines, and in [18], Gaussian smoothing is used to eliminate noise. Transformation to binary image is used in [19]. Inverse perspective mapping and road segmentation methods are used as preprocessing in [20, 21]. A study conducted by Jung et al. [22] proposed constructing spatiotemporal images which exploits the temporal dependency of the video frames.

The widely used method for line detection after preprocessing stage is Hough transform. Hough transform is a generic line detection algorithm and used to find road lines [23, 24]. Borkar et al. [21] use a gaussian template matching method after Hough transform in order to increase the detection efficiency. Wang et al. [25] proposed using B-snakes to detect and track road lines. Ridge detection is performed in [20] with convolution with a Gaussian kernel. Another study proposed in [26] combines the self-clustering algorithm (SCA), fuzzy C-mean and fuzzy rules to process the spatial information and
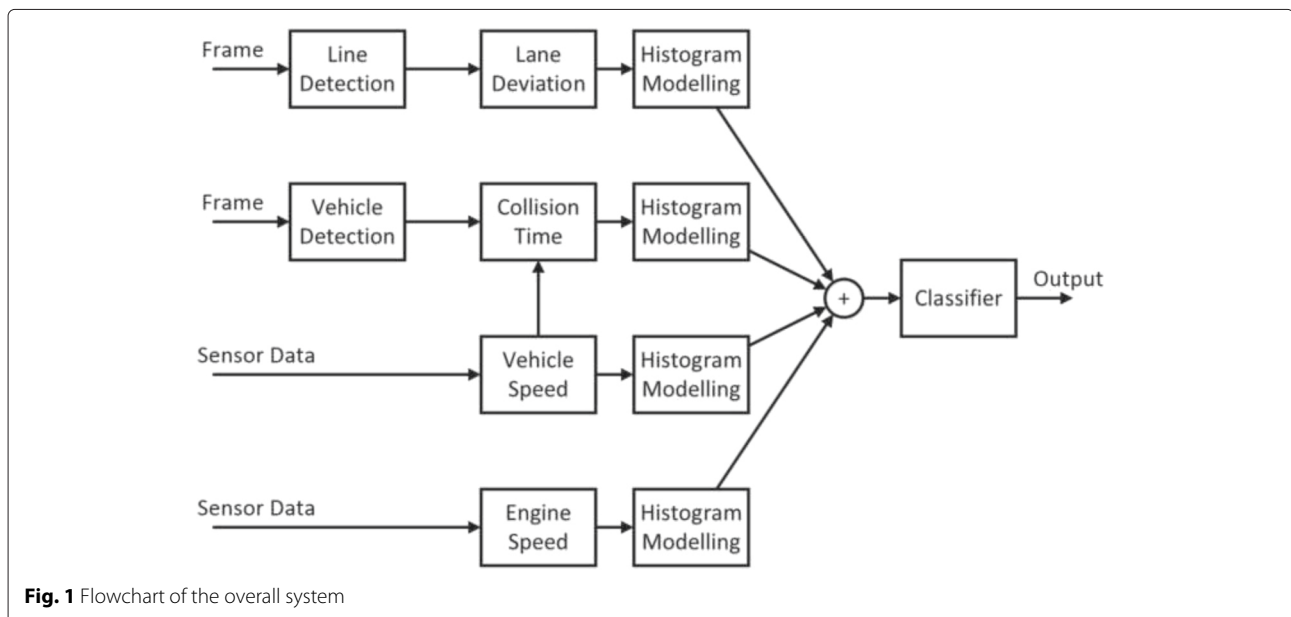

Fig. 1 Flowchart of the overall system

Canny algorithms to get good edge detection. Wu et al. [19] exploits angular relations between lane boundaries. Mu and Ma [27] uses piecewise fitting and object segmentation method to indicate road line positions. In [21], template matching in inverse perspective mapping applied images is used with a tracking scheme.

Although these presented methods perform promising results, they are not fully useful for our application. In our case, the main concern is to use a robust, simple and low complexity algorithm to detect the position of the host vehicle between two road lines correctly and fast enough to work in real time. Although these methods performs well regarding the detection rate, they are not providing low complexity, simplicity and robustness together, that is, a well-performing method may require high computation power and complex implementation which is a big disadvantage for real-time applications. Therefore, the main objective is to implement an algorithmically simple method which performs a robust operation. In order to satisfy this condition, we use temporal filtering and inverse perspective mapping which is a simple method as well as having low complexity and high robustness as explained in [14].

One of the most important problems regarding the robustness of the system is non-uniformity of road conditions [21]. In order to overcome the problems that are caused by shadows, different light conditions and discontinuities on the road line, a method based on temporal filtering is used [14, 21] with inverse perspective mapping which gives robust, fast and simple results.

First, the captured image is temporally filtered in order to eliminate dashed lines and discontinuities according to (1)

$$I_k^{'}(x,y) = \max\{I_k(x,y), \dots, I_{k-K}(x,y)\} \tag{1}$$

where I_k represents the current frame, I_{k-K} represents the Kth previous frame and (x,y) are pixel coordinates. K is chosen according to the frame rate and dashed line length so that all road lines can be seen as a continuous line as in Fig. 2.

Then, the gradient image of $I_k^{'}(x,y)$ is calculated and the high-gradient pixels are cleared from $I_k^{'}(x,y)$ to obtain $I_k^{''}(x,y)$. This operation gives the low gradient pixels which represent the road plane. Then, the mean and variance values of $I_k^{''}(x,y)$ is calculated so that the mean intensity value of road part can be known. Once these values are obtained, the pixels that are representing road plane are cleared from the image $I_k^{'}(x,y)$. This operation helps to eliminate noise and indicate road lines better. A simple derivative filter $F = [-1\ 0\ 1]$ is used to indicate the lines. After this operation, binary image is obtained using an adaptive threshold according to Otsu's method [28].

Inverse perspective mapping is an efficient method for road line detection. Camera placed at the front of a vehicle gives the road lines as straight lines intersecting at the horizon level. However, inverse perspective mapping enables the road lines to be seen as parallel lines. Moreover, since monocular vision system is used, inverse perspective mapping will be exploited to measure the distance between vehicles. In order to achieve inverse perspective mapping, four points are chosen in the filtered image and they are mapped to four other points in the birds-eye perspective assuming the surfaces are planar as in Fig. 3. This mapping procedure results in a 3×4 H matrix that contains the transformation parameters. This matrix is calculated before the operation and loaded to the system. Then, during the operation, inverse perspective mapping is done by transforming each ith point using H matrix as in (3).

$$p_k^i = HP_k^i \tag{2}$$

Since the aforementioned procedures work well enough to indicate the line positions, a simple procedure is done to locate road lines. Horizontal projection of the image is taken in a limited region so that the line locations appear as peaks in the horizontal projection vector. Nieto et al. [14] solve the line localization problem with a parametric curve fitting. However, since we exploit simple methods for the sake of real time application, we modelled the road lines with simple lines. This procedure is based on the assumption that curved roads are seen as straight up to a certain distance. And the region whose horizontal projection taken is chosen to minimize the noise in peak detection as shown in Fig. 4.

One last step that is used to increase the stability and accuracy of line detection is tracking the detected lines with kalman filter [29]. This tracking scheme includes denoising with Kalman filter as well as keeping the visibility counts of lines and recovering missing detections for a specific frame. This scheme significantly improves the efficiency of the overall process.

The two closest detected lines from the camera center, which is defined beforehand as a pixel value according to horizontal positioning of the camera, are chosen as own lane boundaries. The horizontal position of the camera center from the lane boundary is determined as parameter between -50 and 50 for each frame I_k.

The presented method is tested with real set-up data which includes different road conditions such as shadows, occlusion and road curve. As can be seen in the sample figures (Fig. 5), line detection method show robustness to these environment conditions.

In order to test the accuracy and reliability, the presented method is tested with Borkar's dataset [21]. In Borkar's dataset, there exist video sequences containing driving sessions at urban road, metropolitan highway and

Kumtepe *et al. EURASIP Journal on Image and Video Processing* (2016) 2016:5

Page 5 of 16
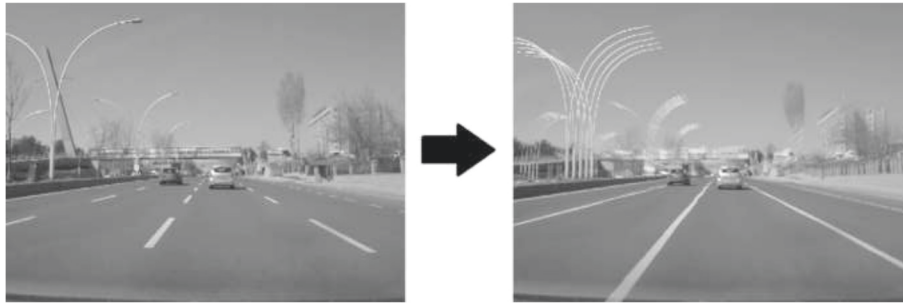
**Fig. 2** Raw image and temporal filtered image

isolated highway whose ground truth road line positions are provided. Since the main aim of road line detection module is to extract the position information of the host vehicle inside the lane, these ground truth values are used to determine the ground truth values of lane position. For this task, pixel value of camera center is estimated by visual inspection and position information is calculated accordingly. As can be seen in Fig. 6 for different video sequences, lane position information is determined accurately. In order to quantify the accuracy of the position information over a video sequence, mean absolute error mean absolute error (MAE) values are calculated. As indicated in [30], MAE can be used for measuring estimation accuracy of driving signals such as speed, orientation, etc. Hence, for each presented sequence in Fig. 6, a MAE value is calculated as in Eq. 3 and shown in Table 1. Regarding the mean absolute error values of the sequences, it can be said that for different conditions, the presented method performs lane position detection with a limited error rate. To illustrate MAE value for isolated highway, data is found as 1.51 which means that average error of lane position detection is 1.51 in −50:50 scale for a frame which is a very small error rate.

$$\text{MAE} = \sum_{i=1}^{N} \frac{|\text{LaneDeviation}_{GT}(i) - \text{LaneDeviation}_{\text{measured}}(i)|}{N}$$

(3)

We compared our lane deviation detection results with other methods in the literature which is tested for Borkar's dataset. As presented in [22], Jung et al. stated that lane detection rate for their method and Borkar's method are as in Table 2. We tested our method with Borkar's dataset with video sequences containing different conditions and presented the results in Table 2. Our method provides similar results with existing methods, performing better for urban dataset which is more critical in terms of aggressiveness detection. Moreover, the line deviation values over frames will be represented as distributions which is explained in the "Feature extraction and classification" section. This process will further compensate the deteriorating effect of errors regarding aggressiveness detection.

3.2 Vehicle detection

Vehicle detection process is required in order to find the distance between host car and other cars that can be seen from the camera. This distance will be used to build up a feature which characterize tailgating or unsafe following distance behaviour. For vehicle detection task, we used a simple and robust approach for the sake of real-time operation and we employed histogram of oriented gradients (HOG) features with a cascade classifier. We also improved the algorithmic efficiency and accuracy of vehicle detections by exploiting lane detection results since we
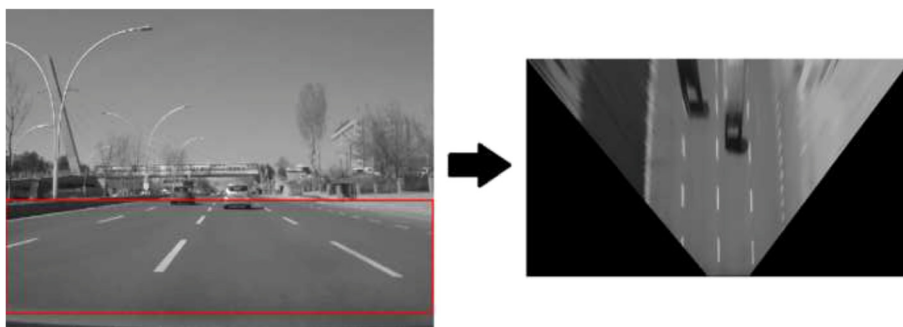

Fig. 3 Four corner point selection and perspective transformation. The corners of the *red box* represent the four chosen points, and these points are mapped to the corners of the *trapezoidal region* in the figure at the *right*

Kumtepe *et al. EURASIP Journal on Image and Video Processing*   (2016) 2016:5

Page 6 of 16
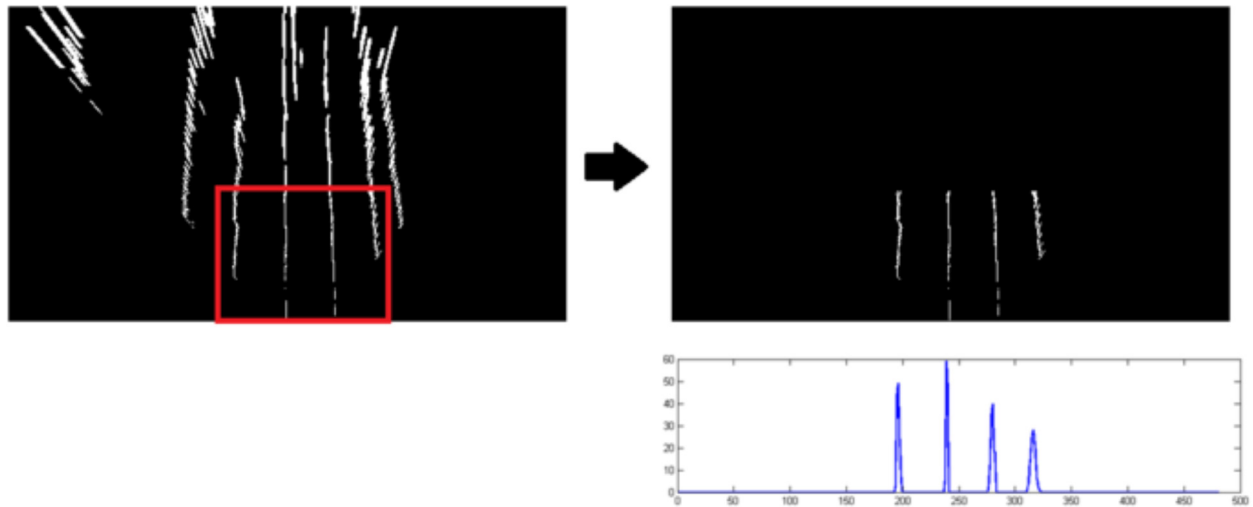
Fig. 4 Line position detection by horizontal projection. The figure at the *top left* represents the processed and transformed image. The *red box* represents the limited interest region. The figure at the *top right* is the masked version according to interest region. The graph at the *bottom right* is the horizontal projection of the image

Fig. 5 Correctly detected lines in different frames

Kumtepe *et al. EURASIP Journal on Image and Video Processing* (2016) 2016:5

Page 7 of 16
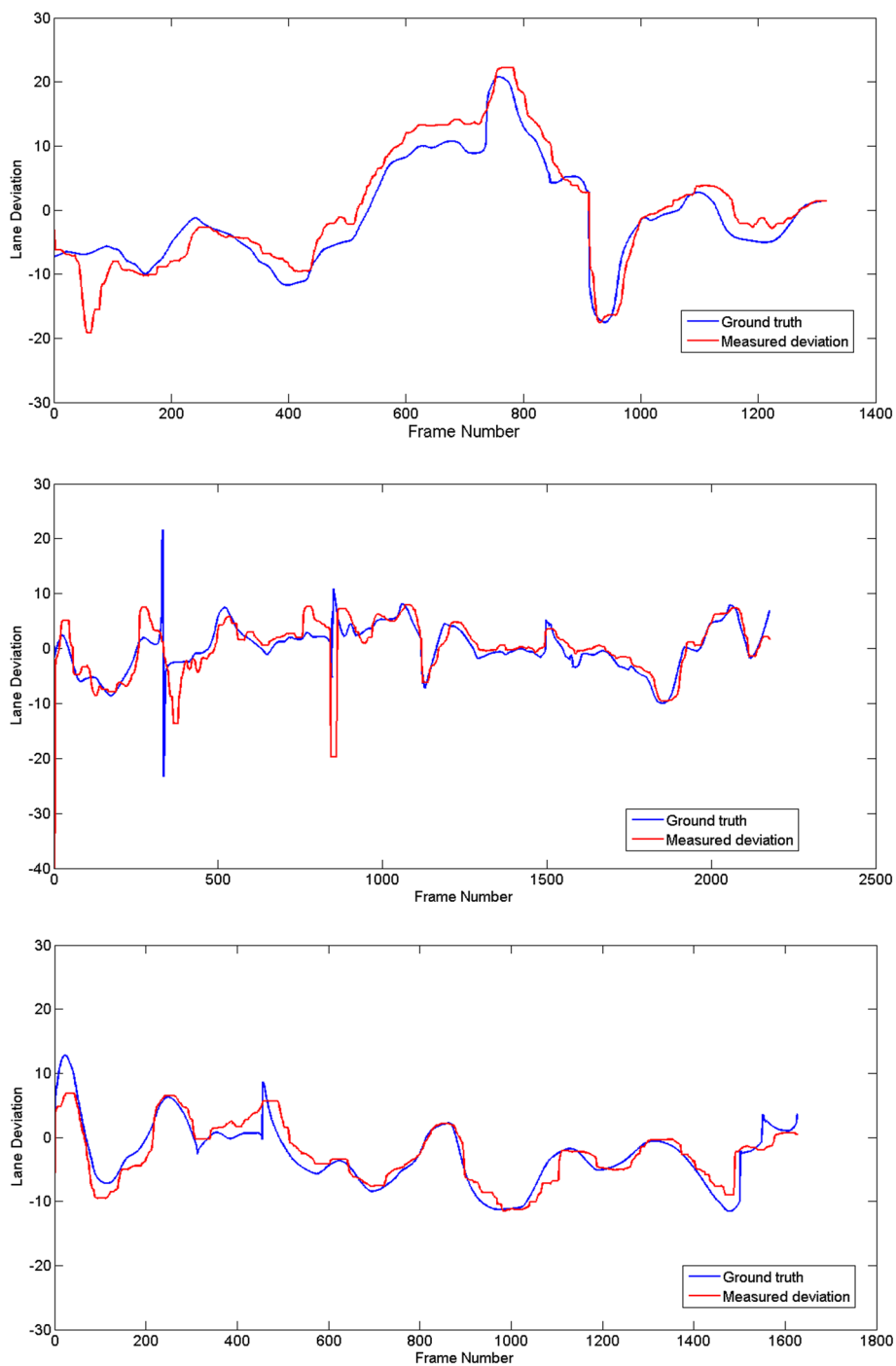
Fig. 6 Comparison of lane position detection with ground truth values in Borkar's dataset. The figure at the *top* belongs to video sequence of urban area in low traffic condition, the figure in the *middle* belongs to video sequence of metro highway in dense traffic condition and the figure at the *bottom* belongs to isolated highway in moderate traffic condition

are interested in only the vehicles which are in the same lane with host vehicle. This condition enabled us to run vehicle detection process in a specific region of interest.

There exist different approaches in previous studies about on-road vehicle detection. In most of the previous studies, vehicle detection is associated with forward

Kumtepe *et al. EURASIP Journal on Image and Video Processing* (2016) 2016:5

Page 8 of 16

Table 1 Mean absolute error values for different road and traffic conditions

Road condition	Mean absolute error
Urban area, low traffic	2.47
Metro highway, dense traffic	1.96
Isolated highway, moderate traffic	1.51

collision warning systems (FCWS) which is a part of driving assistance systems (DAS). In these systems, vehicle detection and distance estimation can be performed by radars or simple sensors as explained in [31, 32]. Another alternative of radar sensors are lidars. Lidars are also used for this task collaboratively with radars [33]. However, the state-of-the-art forward collision systems are based on camera-based platforms and image processing techniques. In literature, among on-road vehicle detection methods, Kim et al. [34] do vehicle detection by scanning the image so as to find a shadow region by the help of some morphological operations. The work presented in [35] depends on the active training of images represented by Haar-like features. In [10], HOG features are extracted from the frames, then a support vector machine (SVM) classification is utilized to find the vehicles. Considering forward vehicle distance estimation, both [34] and [10] use inverse perspective mapping to find the distance of the target vehicles. Other than these monocular camera-based methods, there exist studies that depend on stereo vision. The method presented in [36] detects objects in both images by motion segmentation and determine the vehicle distance by creating a depth map. Kowsari et al. [37] use Haar-like feature extraction, a feature classification with the power of stereo vision. Similarly, Seo et al. [38] use an omnidirectional camera and stereo vision techniques for vehicle detection and distance estimation. As can be seen in these studies, stereo vision methods give good results for estimating vehicle distance while increasing the hardware and computation complexity.

For our application, we employed HOG feature extraction since it is known to be a robust approach for object detection. And a cascade classifier detection technique is utilized in order to detect vehicle because cascade classifier is a robust and fast method which is proper for real time applications. In order to determine the distances of

Table 2 Correct detection rate of different methods of Borkar's dataset

Category	Borkar (%)	Jung (%)	Proposed (%)
Isolated highway	98.24	98.31	93.92
Metro highway	98.12	98.33	95.04
Urban	87.12	90.52	93.31

detected vehicles, inverse perspective mapping is used. Since it involves a training process its performance can be improved by using required number and variety of samples during the training phase.

In order to achieve object detection, first, vehicle images from real traffic data are collected to train a classifier. Image patches that contain a rear images of vehicles are cropped from collected images and tagged as positive. During this phase, different types of vehicles are chosen as samples in order to increase the accuracy. On each of these samples, HOG features are calculated and fed to the classifier.

A cascade classifier is trained according to the process that is described in [39]. During the implementation, each new-coming frame is scanned with a sliding window in different scales; HOG features are calculated over these windows and fed to the classifier to be tested. According to classifier result, detected objects are located by a bounding box. This process may create some false positive that are appearing for a few consecutive frames. Therefore, a Kalman tracking scheme [29] as described in previous section is used to track the detected objects. This process improves the detection rate and eliminates false positives. Some examples of vehicle detection can be seen in Fig. 7. In these figures, it can be seen that different types of vehicles are correctly detected.

In order to find the following distance, a vehicle is chosen as the target vehicle (if there exist a vehicle in the scene). The target vehicle is determined as the nearest vehicle in the own lane of host vehicle. The inverse perspective mapping information, that were found in the previous section, is used to transform the position of the target vehicle to the birds-eye view perspective which enables us to determine the distance between host vehicle target vehicle in pixel units. This difference in pixel units is converted to metric unit with a constant C which is predefined according to the perspective transformation values before the overall process.

The presented vehicle detection method is a well-known and simple scheme, and it gives satisfying results regarding our problem definition. So as to assess the performance of the method, we utilized LISA-Q Front FOV Dataset [35] which contains three different annotated video sequences. In [35], the presented method is tested with LISA dataset and the results are given according to several performance metrics. The details of these metrics can be found in [35].

Since the ultimate aim of the method is to find the distance between host vehicle and target vehicle, we reduced the region of interest in the front view image according to the results of lane detection. In other words, we aimed to detect the vehicles which are in the same lane with the vehicle. To accomplish this, we eliminated the other detections which are in different lanes but ours.

Fig. 7 Examples of detected vehicles

This approach improved the results significantly for each dataset. In Tables 3, 4 and 5, performance results of our method with region of interest selection and comparison with the given method in [35] can be seen for dense, urban and sunny datasets, respectively.

As can be seen in these tables, proposed method performs an average accuracy over 95 % for different conditions. Combining lane detection results with vehicle detection results significantly improved the performance by increasing true positive rate in dense dataset which includes dense traffic images. Furthermore, it decreased the false positive rate in all cases by outperforming the benchmark results in two datasets.

3.3 CAN bus data acquisition
Most of the new cars are equipped with a controller area network (CAN) bus which enables the communication between different microchips and sensors inside the

vehicles. It became mandatory in the USA for the cars that are produced after year 1996. CAN bus has a standardized physical connector and a protocol so that the vehicle data can be obtained using the CAN bus port for analysis and diagnosing purposes. In our application, vehicle speed and engine speed are used as the sensor-based information since certain patterns of these information are associated with aggressive driving behaviour. As indicated in [3], abrupt acceleration and deceleration can be an indication of aggressive driving. Therefore, vehicle and engine speed values are exploited for characterizing driver aggressiveness. In order to collect these data, external sensors can be used as performed in [5]. Instead of using external sensors, CAN bus system of the host vehicle can provide this information [10] with a proper adapter as shown in Fig. 8.

In order to read vehicle and engine speed data from the CAN bus of the vehicle, a proper adapter is used and related data is obtained with timestamps during driving in

Table 3 Performance evaluation of different methods for dense dataset

Method	TPR (%)	FDR (%)	FP/frame (%)	TP/frame (%)
Sivaraman's method [35]	95.0	6.4	0.29	4.20
Our method without lane selection	78.4	43.0	2.44	3.23
Our method with lane selection	89.9	9.8	0.09	0.85

Table 4 Performance evaluation of different methods for urban dataset

Method	TPR (%)	FDR (%)	FP/frame (%)	TP/frame (%)
Sivaraman's method [35]	91.7	25.5	0.39	1.14
Our method without lane selection	99.0	36.5	0.57	0.99
Our method with lane selection	99.0	20.6	0.25	0.99

Kumtepe *et al. EURASIP Journal on Image and Video Processing* (2016) 2016:5

Page 10 of 16

Table 5 Performance evaluation of different methods for sunny dataset

Method	TPR (%)	FDR (%)	FP/frame (%)	TP/frame (%)
Sivaraman's method [35]	99.8	8.5	0.28	3.17
Our method without lane selection	98.7	25.9	1.04	2.97
Our method with lane selection	98.7	4.8	0.05	0.99

order to synchronize the CAN bus data with visual data. Vehicle and engine speed data are collected with a period of 1 s. Therefore, in order to use this data combined with a higher frequency visual data (i.e. 10 fps frame rate), it is up-sampled by a factor of 10.

3.4 Feature extraction and classification

The aforementioned stages are performed to collect information about the behaviour of the driver in the traffic. These collected information is utilized by a feature extraction and classification stage in order to determine whether the related driving session is aggressive or not. For the characterization of the driving session, four different features are chosen considering the aggressive driving indicating behaviours as explained in the "Proposed method" section. These features are as follows:

- Lane deviation
- Collision time
- Vehicle speed
- Engine speed

The line detection results and lane position determination are used to construct lane deviation feature which characterizes the abrupt lane changing and not following the lane properly. The information obtained from the CAN bus, vehicle and engine speed is directly used as the features since drivers who show aggressive driving behaviour tend to drive with high and varying speed, therefore changing engine speed abruptly. The last feature which characterizes the tailgating and unsafe following distance behaviours is the collision time. Collision time feature defines the duration to collision if the vehicle in front would stop suddenly. Therefore, this feature utilizes both speed and target vehicle distance information. Collision time is calculated with a unit of seconds according to (5) where d_k is the distance of the target vehicle in meters and v'_k is the vehicle speed in meters per second at that instant.

$$\text{Collision time}(k) = \frac{d_k}{v'_k} \qquad (4)$$

Considering all features that characterize the driving session, their variation pattern in a certain amount of time is more informative for us rather than the time series signal itself in terms of driver aggressiveness. For instance, the frequency that a driver changes lanes is a more important information than the lane position value at a specific time frame. Therefore, we represented time series signals as density functions and modelled them using Gaussian mixture model (GMM) which is a powerful technique for density representation [40]. Since we are handling the collected data by batch process, Gaussian modelling provides an effective representation of driving

Fig. 8 CAN bus serial port adapter

Kumtepe *et al. EURASIP Journal on Image and Video Processing* (2016) 2016:5

Page 11 of 16

data. The works presented in [5] and [40] use Gaussian modelling of driving signals for making inferences about driving profiles and present effective results in terms of accuracy.

For our application, each feature is transformed into density functions (i.e. histograms). These histograms are filtered with a median filter in order to eliminate noisy data. Then, they are normalized so that all histograms represent the frequency of the data in the same base. A sample representation of an aggressive and smooth data can be seen in Fig. 9.
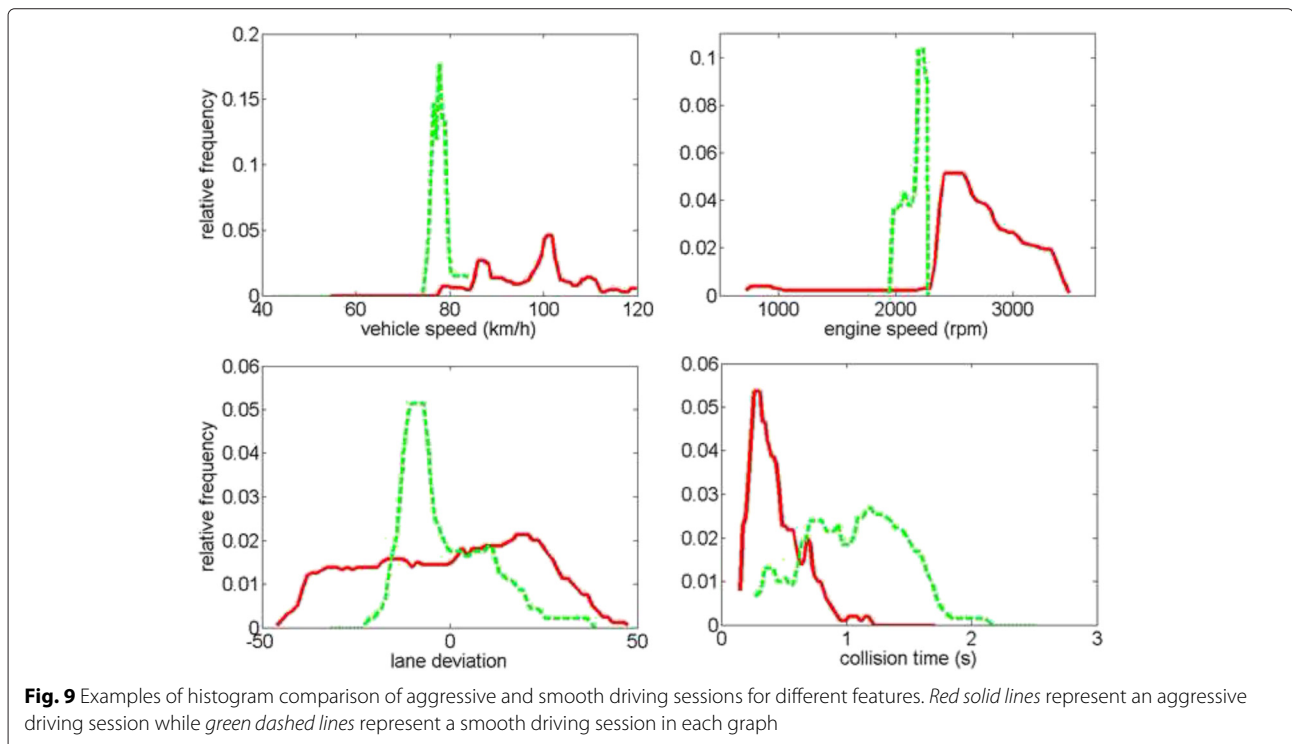
During the experiments, we observed that the density functions of driving signals have one dominant Gaussian component. Hence, we modelled histograms using one GMM component which is denoted by a mean μ and a standard deviation σ value which are enough for representing a Gaussian distribution. GMM components of density function are estimated using maximum likelihood estimation. Each driving feature provided one μ and one σ value. Then, these four mean and four standard deviation values are utilized to construct a feature vector consisting eight dimensions. An SVM classifier is employed [41] in order to classify the feature vectors to determine whether a driving session involves aggressive driving behaviour.
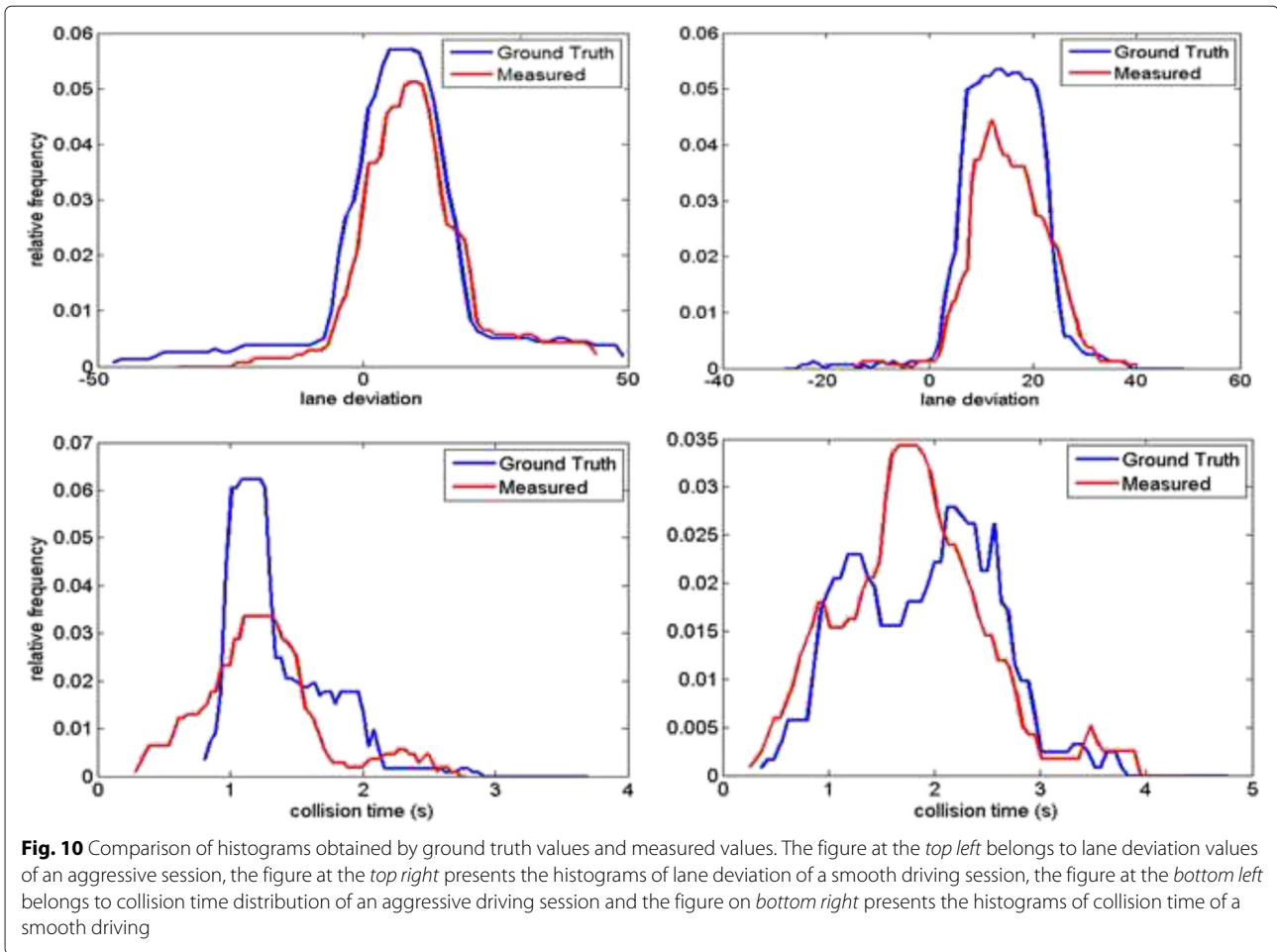
Although the presented feature extraction methods are proven to be reliable and comparable with the methods in the literature, the performance of lane deviation detection and collision time estimation modules will effect the result

of the aggressiveness classification. Nevertheless, the histogram representation of the features provides robustness to the process and reduces the deteriorating effect of missing detections in line detection and vehicle detection stages. In Fig. 10, histogram modelling of lane deviation and collision time values of an aggressive and a smooth driving session is presented. Mean and standard deviation values of these histograms are presented in Tables 6 and 7 with mean absolute error values between ground truth and measured time series signals. The data presented in Table 6 belong to the sample aggressive session whose histogram is given in Fig. 10, while the data presented in Table 7 belong to the smooth session. As can be seen in these tables, the effect of errors in the detection stage can be eliminated significantly utilizing the histogram modelling.

4 Experimental results

For test purposes, a mobile set up is constructed in order to collect visual and CAN bus data by vehicle. For visual data collection, a portable mini computer (Fig. 11) and a CCD camera (Fig. 12) is used. By this platform, video frames are captured at 10 fps with a resolution of 800×600 pixels. For CAN bus data collection, the adapter in (Fig. 8) is connected to CAN bus port of the vehicle and data is acquired through the serial port of the mini computer. The data collected from CAN bus is obtained at each second. Therefore, data is interpolated so that the sensor data exist for each frame. So as

Fig. 9 Examples of histogram comparison of aggressive and smooth driving sessions for different features. *Red solid lines* represent an aggressive driving session while *green dashed lines* represent a smooth driving session in each graph

Fig. 10 Comparison of histograms obtained by ground truth values and measured values. The figure at the *top left* belongs to lane deviation values of an aggressive session, the figure at the *top right* presents the histograms of lane deviation of a smooth driving session, the figure at the *bottom left* belongs to collision time distribution of an aggressive driving session and the figure on *bottom right* presents the histograms of collision time of a smooth driving

to synchronize the visual and CAN bus data, the data is timestamped.

Utilizing this set up, real traffic data is collected at different times of the day so that different traffic conditions are included in the dataset. The dataset also includes different road conditions with occlusions, shadows and different illumination. Whole dataset contains driving sessions of six different drivers. During driving, three different observers annotated the last 40 s as aggressive or smooth. The majority voting of the observers are recorded as the ground truth of the related driving session.

One important parameter that effects the performance of the proposed method is the duration of the driving session. In other words, how long multisensory

data is required in order to efficiently determine if that driving session is aggressive? In order to answer this question, the collected data is tested with driving sessions with lengths 40, 80, and 120 s. From the whole collected dataset, a total of 83 driving sessions including 41 aggressive and 42 smooth sessions having a duration of 40 s, 51 driving sessions including 22 aggressive 29 smooth sessions having a duration of 80 s and 22 driving sessions including 11 aggressive 11 smooth sessions having a duration of 120 s are tested according to proposed algorithm.

Due to the limited amount of data, k-fold cross validation technique is used for performance assessment. According to this technique, test samples are chosen randomly among the samples; the remaining samples are

Table 6 Comparison of ground truth and measured features of the sample aggressive driving session

	μ_{lane}	σ_{lane}	$\mu_{collision}$	$\sigma_{collision}$	MAE_{lane}	$MAE_{collision}$
Measured	14.88	7.57	2.01	0.81	3.23	0.58
Ground truth	14.07	8.01	1.93	0.89		

Table 7 Comparison of ground truth and measured features of the sample smooth driving session

	μ_{lane}	σ_{lane}	$\mu_{collision}$	$\sigma_{collision}$	MAE_{lane}	$MAE_{collision}$
Measured	9.64	10.73	1.47	0.52	6.78	0.46
Ground Truth	6.70	14.38	1.41	0.48		

Kumtepe *et al. EURASIP Journal on Image and Video Processing* (2016) 2016:5

Page 13 of 16

Fig. 11 Mini computer used in data collection

used for training the SVM classifier. This process is performed 10 times, and at each run, the classifier results are compared with the ground truth. For the 40-s-long samples, 20 of them; for the 80-s-long samples, 15 of them; and for the 120-s-long samples, 9 of them are chosen randomly as test samples. In Tables 8, 9 and 10, the related confusion matrices of the test results are given for 40-, 80- and 120-s-long samples, respectively.

According to the test results, it is observed that the proposed method achieved 91, 94 and 82.2 % detection rate for 40-, 80- and 120-s-long samples, respectively. As can be inferred from these results, 80-s-long driving sessions are more efficiently representing the driving characteristics while 40-s samples may not allocate enough data or 120-s samples may contain confusing data.
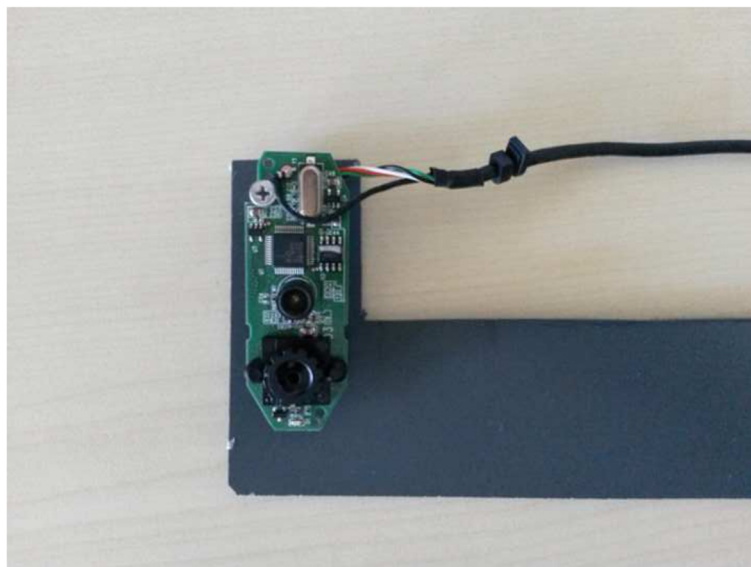
Fig. 12 Camera to capture visual information

Kumtepe *et al. EURASIP Journal on Image and Video Processing* (2016) 2016:5

Page 14 of 16

Table 8 Confusion matrix of aggressiveness classification for 40-s-long data

		Predicted class	
		Aggressive	Smooth
Real class	Aggressive	90	3
	Smooth	15	92

Proposed aggressiveness detection method also tested with real-world data from 100-car dataset [42]. This dataset is the output of a naturalistic driving study and collected via instrumented vehicles in a large scale. In the publicly available part of this dataset, some driving sessions which are approximately 30-s long are given with narratives. These narratives explain the events in the driving session. We investigated these narratives and selected the ones which can be interpreted as an aggressiveness involvement and which cannot. According to narratives, the ones which include aggressive and sharp actions are annotated as "aggressive" and the ones which includes stable actions as "smooth". We selected a total of 76 driving sessions according to narratives and tagged 40 of them as aggressive and 53 of them as smooth. In Table 11, some sample narratives of 100-car data and their interpretation is presented.

The vehicle speed, lane deviation and collision time data are directly present at 100-car dataset. However, instead of engine speed, gas pedal position data is used due to the direct correlation between them. Using these information, the aforementioned feature extraction procedure is applied to the data. In order to validate the reliability of the 100-car data, k-fold cross validation technique is utilized. In each run, 29 of the 93 driving session samples are chosen randomly to train an SVM classifier, and this procedure is repeated 10 times. The classifier achieved a correct detection at an average rate of 93.1 %. Confusion matrix of this process can be seen in Table 12.

5 Conclusions

In this paper, a driver aggressiveness detection method is presented. The proposed method utilizes multisensory information to conceive feature vectors, and using these, feature vectors classify the driving session as aggressive or smooth. The aggressiveness classifier is trained with data annotated by observers and performs classification using data collected in real-world conditions. The paper also studies the required driving session duration that can be efficiently decided if it involves aggressive driving behaviour. According to test results the proposed system performs good results in terms of detecting driver aggressiveness since it considers different driving behaviours in a real time operation. As a future work, the proposed system

Table 9 Confusion matrix of aggressiveness classification for 80-s-long data

		Predicted class	
		Aggressive	Smooth
Real class	Aggressive	67	5
	Smooth	4	74

Table 10 Confusion matrix of aggressiveness classification for 120-s-long data

		Predicted class	
		Aggressive	Smooth
Real class	Aggressive	33	7
	Smooth	9	41

Table 11 Sample driving sessions with their narratives and aggressiveness interpretation

Sample number	Narrative of the session	Aggressiveness
8354	Subject vehicle is driving relatively fast in the left lane as the traffic is merging into right lane. Lead vehicle is decelerating with right turn signal on, preparing to merge into right lane, and the subject vehicle must brake to avoid hitting lead vehicle in the rear. Subject vehicle is trying to get ahead of right lane traffic before merging.	Aggressive
8392	Subject vehicle is travelling in the rain and almost misses the intended exit. Subject vehicle enters the exit ramp at the last minute, nearly side swiping a vehicle already on the ramp beside it. Subject driver steered slightly left to avoid the crash and the other vehicle went ahead on the exit ramp.	Aggressive
8420	Subject vehicle is preparing to merge onto an exit ramp and a vehicle from the adjacent left lane realizes that they need to get onto the exit ramp also, and the lead vehicle suddenly crosses the subject vehicle's left lane line into the subject vehicle's lane. The subject brakes hard to avoid hitting the lead vehicle in the rear.	Aggressive
8374	Subject driver is talking/singing to herself and stops behind a line of cars at a light. A following vehicle approaches rapidly and almost hits the subject vehicle in the rear.	Smooth
8471	There are 2 left turn lanes with the subject driver in the far left lane. Vehicle 2 at the left turn lane to the right of the subject's vehicle starts to turn left and cuts the subject driver off.	Smooth
9059	Both the subject driver and lead vehicle are decelerating when the subject driver glances out his right side window. When the subject driver glances, the lead vehicle comes to a stop in front of him.	Smooth

Kumtepe *et al. EURASIP Journal on Image and Video Processing* (2016) 2016:5

Page 15 of 16

Table 12 Confusion matrix of aggressiveness classification for 100-car data

		Predicted class	
		Aggressive	Smooth
Real class	Aggressive	118	12
	Smooth	8	152

will be tested with more data to observe its performance with different classifiers. The system will be improved in order to provide a rate for driver aggressiveness in a granular approach. In other words, the measurement of aggressiveness level will be provided quantitatively.

Author details
[1]Department of Electrical and Electronics Engineering, Middle East Technical University, Dumlupınar Blv. No:1, 06800 Ankara, Turkey. [2]ISSD Electronics, METU Technopolis, Dumlupınar Blv. No:1, 06800 Ankara, Turkey.

References
1. World Health Organization. Violence and Iznjury Prevention and World Health Organization, *Global Status Report on Road Safety 2013: Supporting a Decade of Action*. (World Health Organization, Geneve, 2013)
2. Aggressive driving: Research update. Technical report, AAA Foundation for Traffic Safety, Washington DC (2009)
3. Aggressive Driving Enforcement: Strategies for Implementing Best Practices. Technical Report, US Dept of Transportation National Highway Traffic Safety Administration, Washington DC (March 2000)
4. T Toledo, HN Koutsopoulos, M Ben-Akiva, Integrated driving behavior modeling. Transp. Res. C Emerg. Technol. **15**(2), 96–112 (2007)
5. ABR Gonzalez, MR Wilby, JJV Diaz, CS Avila, Modeling and detecting aggressiveness from driving signals. Intell. Transp. Syst. **15**(4), 1419–1428 (2014)
6. M Danaf, M Abou-Zeid, I Kaysi, Modeling anger and aggressive driving behavior in a dynamic choice—latent variable model. Accid. Anal. Prev. **75**(0), 105–118 (2015)
7. SH Hamdar, HS Mahmassani, RB Chen, Aggressiveness propensity index for driving behavior at signalized intersections. Accid. Anal. Prev. **40**(1), 315–326 (2008)
8. DA Johnson, MM Trivedi, in *Intelligent Transportation Systems (ITSC), 2011 14th International IEEE Conference On*. Driving Style Recognition Using a Smartphone as a Sensor Platform, (Washington, DC, 2011), pp. 1609–1615
9. H-B Kang, in *Computer Vision Workshops (ICCVW), 2013 IEEE International Conference On*. Various Approaches for Driver and Driving Behavior Monitoring: A Review, (Sydney, NSW, 2013), pp. 616–623
10. RK Satzoda, MM Trivedi, Drive analysis using vehicle dynamics and vision-based lane semantics. IEEE Trans. Intell. Transp. Syst. **16**(1), 9–18 (2015)
11. G Jian-Qiang, W Yi-Ying, in *Intelligent Systems Design and Engineering Applications (ISDEA), 2014 Fifth International Conference On*. Research on Online Identification Algorithm of Dangerous Driving Behavior, (Hunan, 2014), pp. 821–824
12. B-F Wu, Y-H Chen, C-H Yeh, in *ITS Telecommunications (ITST), 2012 12th International Conference On*. Fuzzy Logic Based Driving Behavior Monitoring Using Hidden Markov Models, (Taipei, 2012), pp. 447–451
13. M Rezaei, M Sarshar, MM Sanaatiyan, in *Computer and Automation Engineering (ICCAE), 2010 The 2nd International Conference On*. Toward Next Generation of Driver Assistance Systems: A Multimodal Sensor-Based Platform, vol. 4, (Singapore, 2010), pp. 62–67
14. M Nieto, L Salgado, F Jaureguizar, J Arrospide, in *Image Processing, 2008. ICIP 2008. 15th IEEE International Conference On*. Robust Multiple Lane Road Modeling Based on Perspective Analysis, (San Diego, CA, 2008), pp. 2396–2399
15. G Somasundaram, Kavitha, K. I Ramachandran, in *Computer Science and Information Technologies CSIT, Computer Science Conference Proceedings (CSCP)*, ed. by DC Wyld. Lane Change Detection and Tracking for a Safe Lane Approach in Real Time Vision Based Navigation Systems, (Chennai, India, 2011). July 15 - 17, 2011
16. M-G Chen, C-L Ting, R-I Chang, Safe driving assistance by lane-change detecting and tracking for intelligent transportation system. Int. J. Inform. Process. Manag. **4**(7), 31–38 (2013)
17. M Oussalah, A Zaatri, H Van Brussel, Kalman filter approach for lane extraction and following. J. Intell. Robotics Syst. **34**(2), 195–218 (2002)
18. C Nuthong, T Charoenpong, in *Image and Signal Processing (CISP), 2010 3rd International Congress On*. Lane Detection Using Smoothing Spline, vol. 2, (Yantai, 2010), pp. 989–993
19. C-F Wu, C-J Lin, C-Y Lee, Applying a functional neurofuzzy network to real-time lane detection and front-vehicle distance measurement. IEEE Trans. Syst. Man Cybern. C. **42**(4), 577–589 (2012)
20. M Beyeler, F Mirus, A Verl, in *Robotics and Automation (ICRA), 2014 IEEE International Conference On*. Vision-Based Robust Road Lane Detection in Urban Environments, (Hong Kong, 2014), pp. 4920–4925
21. A Borkar, M Hayes, MT Smith, A novel lane detection system with efficient ground truth generation. IEEE Trans. Intell. Transp. Syst. **13**(1), 365–374 (2012)
22. S Jung, J Youn, S Sull, Efficient lane detection based on spatiotemporal images. IEEE Trans. Intell. Transp. Syst. **PP**(99), 1–7 (2015)
23. V Gaikwad, S Lokhande, Lane departure identification for advanced driver assistance. IEEE Trans. Intell. Transp. Syst. **16**(2), 910–918 (2015)
24. C Tu, BJ van Wyk, Y Hamam, K Djouani, S Du, Vehicle Position Monitoring Using Hough Transform. {IERI} Procedia. **4**(0), 316–322 (2013). 2013 International Conference on Electronic Engineering and Computer Science (EECS 2013)
25. Y Wang, EK Teoh, D Shen, Lane detection and tracking using b-snake. Image Vis. Comput. **22**(4), 269–280 (2004)
26. J-G Wang, C-J Lin, S-M Chen, Applying fuzzy method to vision-based lane detection and departure warning system. Expert Syst. Appl. **37**(1), 113–126 (2010)
27. C Mu, X Ma, Lane detection based on object segmentation and piecewise fitting. TELKOMNIKA Indones. J. Electr. Eng. TELKOMNIKA. **12**(5), 3491–3500 (2014)
28. N Otsu, A threshold selection method from gray-level histograms. IEEE Trans. Syst. Man Cybern. **9**(1), 62–66 (1979)
29. S Thrun, W Burgard, D Fox, *Probabilistic Robotics (Intelligent Robotics and Autonomous Agents)*. (The MIT Press, Cambridge, 2005)
30. R Danescu, F Oniga, S Nedevschi, Modeling and tracking the driving environment with a particle-based occupancy grid. IEEE Trans. Intell. Transp. Syst. **12**(4), 1331–1342 (2011)
31. Y-C Hsieh, F-L Lian, C-M Hsu, in *Intelligent Transportation Systems Conference, 2007. ITSC 2007. IEEE*. Optimal Multi-Sensor Selection for Driver Assistance Systems Under Dynamical Driving Environment, (Seattle, WA, 2007), pp. 696–701
32. M Satake, T Hasegawa, in *Vehicular Electronics and Safety, 2008. ICVES 2008. IEEE International Conference On*. Effects of Measurement Errors on Driving Assistance System Using On-Board Sensors, (Columbus, OH, 2008), pp. 303–308
33. Y Wei, H Meng, H Zhang, X Wang, in *Intelligent Transportation Systems Conference, 2007. ITSC 2007. IEEE*. Vehicle Frontal Collision Warning System Based on Improved Target Tracking and Threat Assessment, (2007), pp. 167–172
34. S Kim, S-y Oh, J Kang, Y Ryu, K Kim, S-C Park, K Park, in *Intelligent Robots and Systems, 2005. (IROS 2005). 2005 IEEE/RSJ International Conference On*. Front and Rear Vehicle Detection and Tracking in the Day and Night Times Using Vision and Sonar Sensor Fusion, (Alberta Canada, 2005), pp. 2173–2178
35. S Sivaraman, MM Trivedi, A general active-learning framework for on-road vehicle recognition and tracking. IEEE Trans. Intell. Transp. Syst. **11**(2), 267–276 (2010)

Kumtepe *et al. EURASIP Journal on Image and Video Processing*   (2016) 2016:5

Page 16 of 16

36. M Miyama, Y Matsuda, in *Signal and Image Processing Applications (ICSIPA), 2011 IEEE International Conference On*. Vehicle Detection and Tracking with Affine Motion Segmentation in Stereo Video, (Kuala Lumpur, 2011), pp. 271–276

37. T Kowsari, SS Beauchemin, J Cho, in *Intelligent Transportation Systems (ITSC), 2011 14th International IEEE Conference On*. Real-Time Vehicle Detection and Tracking Using Stereo Vision and Multi-View AdaBoost, (Washington, DC, 2011), pp. 1255–1260

38. D Seo, H Park, K Jo, K Eom, S Yang, T Kim, in *Industrial Electronics Society, IECON 2013—39th Annual Conference of the IEEE*. Omnidirectional Stereo Vision Based Vehicle Detection and Distance Measurement for Driver Assistance System, (Vienna, 2013), pp. 5507–5511

39. P Viola, M Jones, in *Computer Vision and Pattern Recognition, 2001. CVPR 2001. Proceedings of the 2001 IEEE Computer Society Conference On*. Rapid Object Detection Using a Boosted Cascade of Simple Features, vol. 1, (Kauai, HI, 2001)

40. A Wahab, C Quek, CK Tan, K Takeda, Driving profile modeling and recognition based on soft computing approach. IEEE Trans. Neural Netw. **20**(4), 563–582 (2009)

41. C Cortes, V Vapnik, Support-vector networks. Mach. Learn. **20**(3), 273–297 (1995)

42. VL Neale, TA Dingus, SG Klauer, J Sudweeks, M Goodman, An overview of the 100-car naturalistic study and findings. National Highway Traffic Safety Administration, Paper 05-0400 (2005)