Analysis of the Vehicle Maneuver and Driver Emotion: Methodology and Results Discussion

Varvara Shushkova[1], Alexey Kashevnik[2,3], Yulia Rzhonsnitskaya[1], Alexey Blazhenov[1]
[1]ITMO University, St. Petersburg, Russia
[2]SPC RAS, St. Petersburg, Russia
[3]Petrozavodsk State University (PetrSU), Petrozavodsk, Russia
shushkova.vv@gmail.com, alexey.kashevnik@iias.spb.su, july.rzhony@gmail.com, a_blazh@mail.ru

Abstract—**Nowadays, systems for detecting driver behaviour are being actively developed in order to reduce road accidents. Emotions are one of the factors that influence driver's decisions. This paper explores the correlation between the type of performed maneuver and the emotion experienced by the driver. A methodology to match the maneuver being performed with the emotion is proposed. Experiments show that when aggressive braking is performed, a person's fear level increases in a few seconds. Moreover, it was found that emotions do not appear only at the moment of the maneuver. Experiments have shown that the greatest expression of emotions appears in 4 seconds before the start of the maneuver and ends in 3 seconds after the maneuver is performed.**

I. Introduction

Today a lot of research is aimed at developing a system to detect the current state of the driver. This is due to the constant increase in traffic and the number of accidents on the roads [1]. Developing such systems will help to increase safety on the roads because driver behavior remains one of the most important factors in traffic [2]. This can be achieved by alerting the driver to react to a potentially dangerous situation at a time. For instance, the system can analyse a driver's facial expression and determine if a driver is experiencing stress, fatigue or other negative emotions. If the system detects that a driver is stressed or tired, it can suggest that they take a break, pull over and rest. This can help prevent accidents related to fatigue or decreased concentration. Emotions are also one of the factors that can influence a driver's behavior on the road.

It is now established that emotions are powerful factors in decision making to a large extent [3]. Various domains show important patterns in the mechanisms by which emotions influence decisions and choices. Often, these patterns can be predicted and attempted to influence them. Thus, the authors of [4] showed that positive emotions significantly increase the influence of distractions compared to neutral and negative states. As a consequence, the driver's experience of positive emotions leads to a wider distribution of attention and distraction from the traffic situation.

This paper proposes to investigate the relationship between the driver's behavior on the road and the emotions they feel. This will allow to notice in time situations when the driver is distracted or experiences bright emotions that affect his/her decisions. Moreover, the moment of emotional occurrence as a response to the intention to make a maneuver can characterise the degree of a person's fatigue.

The goal of this paper is to investigate the correlation between a driver's emotional response and the type of road maneuver they perform. The DriverMVT [5] dataset is used to solve the problem. This dataset represents more than 1500 minutes of recordings of 9 drivers in traffic. There is also numeric data: accelerometer, gyroscope, and magnetometer.

The rest part of the paper is organized as follows: Section 2 represents a review of existing video detection approaches, driver behavior detection methods, as well as datasets that are used for the described tasks. Section 3 contains a description of the presented methodology for comparing the maneuver being performed and the driver's emotion, Section 4 describes the obtained results and shows the correlation between the maneuver and the driver's emotion. Section 5 summarizes the results obtained and describes future plans and limitations.

II. Related work

This section reviews existing approaches for face emotion recognition. After that, the existing solutions for driver behavior detection are reviewed and the datasets that are used for this task are described.

A. Video emotion detection

Face emotion recognition is becoming an increasingly relevant topic in the modern world. Detection of human emotions allows to determine their psychological state and make a forecast of their behaviour. Human emotion analysis is applicable in many fields, but today there is no single system that can universally identify human states [6].

Facial expressions are vital identifiers of human feelings because they correspond to emotions [7]. In most cases (about 55% of cases) [8], facial expressions are a non-verbal way of expressing emotion and can be used to predict human behaviour.

The authors of [7] [9] have reviewed articles of the last years, where research on emotion recognition was conducted. They distinguish the main steps of the pipeline for solving the problem of emotion recognition: preprocessing (picture clarity and scaling, contrast adjustment, extra enhancement process), extraction of facial (various texture discriminators, external methods, geometry-based methods), and emotion classifier.

Both classical machine learning methods and deep learning tools are used as emotion classifiers. The authors of [10] use a support vector machine (SVM) to solve the classification problem. The convolutional neural networks (CNN) are used in [11] [12] [13]. The authors offer their modifications of the architecture that improve the quality of the models.

Today a popular trend in emotion recognition is real-time emotion detection. The authors of [11] propose a model of the CNN and LSTM that works effectively under uneven illumination and head rotation (up to 25°), different backgrounds, and different skin tones. Authors [12] propose a new model for emotion detection using Reinforcement Learning to determine the context of conversational clips and incorporate information about the subject area of the conversation.

This paper proposes to solve the emotion detection problem, which is essentially also real-time emotion detection.

B. Driver behaviour detection

Driver behavior is a set of different habits, behaviors, and actions of a driver while driving, which can be divided into five types: normal or safe driving, aggressive driving, distracted driving, drowsy driving, and drunk driving [14]. The topic of aggressive driving behavior detection and prediction of drivers' future actions is particularly relevant. In [15] concluded that there is a correlation between driving conditions or driver behavior and driver state. The authors of [2] proposed a system to classify the driver's driving behavior into 5 types from the [14] using acceleration, gravity, throttle, and speed data. As a result of using such a system, it is possible to notify the driver of aggressive driving behavior at the proper period.

Another issue that needs attention is driver distraction. To solve this problem, the authors use the analysis of the driver's facial expressions and physiological characteristics, such as heart rate, blood saturation, and pressure. In [16] authors propose to detect driver's mental fatigue and drowsiness using the XSENS motion capture system. In this study, a novel modified BiLSTM was developed, trained, and tested. 3D head angular acceleration data with current time information were used as input data. The problem of distracted driver detection is solved in [17] by using a hybrid CNN framework for a dataset with images of drivers in traffic.

There are studies in which the authors evaluated driver driving behavior based on physiological measures. In [15], authors explore the correlation between maneuvers and vital signs. Vital sign indicators are heart rate, oxygen saturation, and blood pressure. The authors confirm the hypothesis that there is an effect of performing a road maneuver on heart rate and blood pressure as a result of the arousal and tension produced by such events.

C. Datasets

Datasets for driver behavior prediction tasks are a set of video sequences describing the traffic in progress. Some studies also use cameras with sensors, such as the Kinetic v2 [5] [18] [19]. With the help of such devices, it is possible to track the work of muscles, and the degree of their tension, which gives an additional field for analysis. Usually, such datasets contain various labels characterizing the surrounding environment: type of maneuver being performed, speed of traffic, information about other traffic participants, and data from different types of sensors.

The 2018 Honda Research Institute Driving Dataset (HDD) [19] was introduced. It is a dataset of over 104 hours of traffic on the streets of San Francisco. The authors propose a novel 4-layer approach for marking up the dataset annotation motivated by human factors: Goal Orientation, Stimulus, Reason, and Attention.

Another dataset containing traffic records is METEOR: A Massive Dense & Heterogeneous Behaviour Dataset for Autonomous Driving [20]. METEOR was introduced in 2021. It contains videos of traffic in India. The video shows the traffic situation as seen by the driver while driving. The dataset is about 100 GB with 1000 minutes of one-minute videos. The videos contain labels with various maneuvers on the road, such as overtaking, rule-breaking, yielding, etc.

Also, some studies present datasets for specific tasks. For instance, the Driver Anomaly Detection (DAD) dataset [21] is proposed to solve the problem of detecting anomalous driver actions such as falling asleep and consequently turning the wheel rapidly. It represents 763 minutes of video recordings of 31 drivers carrying out road traffic on a road simulator.

The DriverMVT [5] dataset also consists of a driver's video sequence in traffic. DriverMVT dataset contains video recordings of 9 drivers engaging in traffic. The total duration of the videos is more than 36 hours. The dataset also contains information about heart rate, driver's head position, accelerometer, and gyroscope data. This dataset can be used to determine the driver's condition while driving and to identify correlations between the driver's driving style and his/her condition.

Another dataset that also contains smartphone sensor measurements is the Driver Behaviour Dataset [22]. This dataset records data from the accelerometer, linear acceleration, gyroscope, and magnetometer using the phone application. For these measurements, there is a mapping with information about what maneuver the driver is currently performing. Among the maneuvers, it is possible to see aggressive left/right turns, overbuilding, acceleration, decelerating, or no maneuver.

In order to prevent accidents, it is necessary to predict what decisions a driver will make during traffic. Such decisions can be predicted based on the detection of a driver's emotions, since emotions are one of the factors that influence the decisions made. This solution can be integrated into existing road hazard alerting systems.

III. Methodology

The main goal of the study is to investigate the correlation between road maneuver and the driver's emotions. The general scheme of data preparation for problem solving is presented in Fig.1.

Firstly, it is necessary to train a maneuver classifier to identify the maneuver in progress using the accelerometer

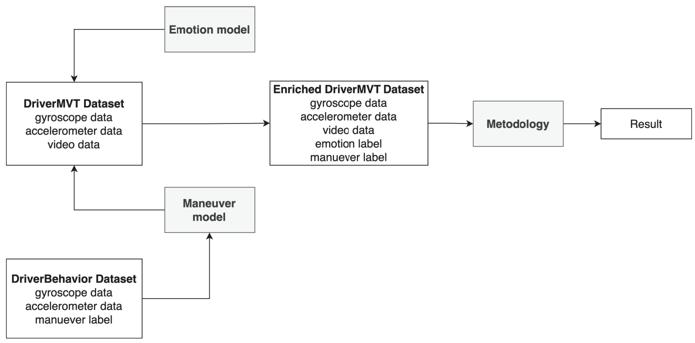

Fig. 1. The scheme of data preparation for problem solving

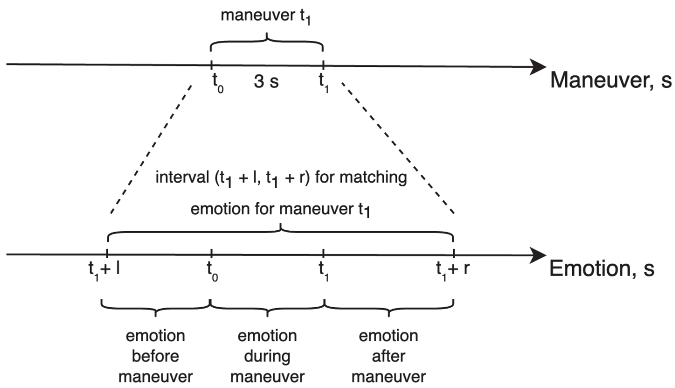

Fig. 2. The model of matching vehicle maneuvers with the driver's emotions at that moment

and gyroscope data. For this purpose, the DriverBehavior-Dataset [22] is used, which contains information about the maneuver in progress at a point in time and data from the accelerometer and gyroscope. This dataset has 3 x,y,z values for the accelerometer and 3 values for the gyroscope. The available data was aggregated to a granularity of 3 seconds (it was assumed that the maneuver duration is 3 seconds and, using a window of length 3 seconds for shifts, a sequence of maneuvers was obtained). The classifier was trained on the labels of the maneuver that corresponds to the last second of these three seconds.

The trained classifier is applied to the DriverMVT: In-Cabin Dataset for Driver Monitoring including Video and Vehicle Telemetry Information dataset [5]. There are 9 participants in the video data: 2 females and 7 males. About 29% of the recordings in this dataset correspond to a moment of no driving. There are recordings where the driver uses the phone while driving, and videos where the driver has a high level of fatigue are also presented. This dataset contains video data and a set of text files that have metrics synchronized with the video for each second. This file contains accelerometer and gyroscope data, but no maneuver label. The classifier trained in the previous step is applied to the available data to identify the maneuver at a point in time.

Then a face emotion recognition open-source model [23] is applied to the video sequence to classify 6 standard emotions (fear, neutral, aggression, happiness, disgust, sadness, surprise). It is a convolution neural network trained on the FER2013 dataset [24]. As an output of the model, numeric values between 0 and 1 are received, which show the expression of the emotions at the current moment. Then it is necessary to compare the obtained emotions with the maneuver being performed at a moment in time.

Next, logic is implemented to match the emotion being experienced to the maneuver being performed. The Fig. 2 is a model of matching vehicle maneuvers with the driver's emotions at that moment.

Let t_0 be the beginning of the maneuver, t_1 be the end. Then (t_0, t_1) is a maneuver with maneuver type per second t_1, $Maneuver$ is the time axis of moments and $Emotion$ is the time axis of emotions. It is necessary to complete the mapping between the moment $maneuver\ t_1$ ending at time t_1 and some $emotions$ around the maneuver $maneuver\ t_1$. It is necessary to compare which emotion is related to the maneuver. For this purpose, it is required to analyze the values of emotions at the different time intervals in connection with the maneuver and to verify the hypothesis that the emotion appears not only during the maneuver but also before or after it.

To achieve this goal, we introduce the left l and right r boundary definitions. These are the values of the time intervals (measured in seconds) from $t_1 + l$ and to $t_1 + r$ respectively, and t_1 may or may not be included in this interval (in case l and r are both negative, we look at the emotions in the interval before the maneuver starts).

Based on the comparison of the obtained values of emotional expressions and the maneuvers performed at the moment, a dataset is assembled. In this dataset, each maneuver at time t_1 is matched with emotion at time $t_1 + i$, where $i = l...r$. In this way, we can explore the driver's emotions at different points in time concerning the maneuver. This allows to evaluate to study the correlation between emotion and the maneuver, not only at the moment of maneuvering but also at moments when the driver is only going to maneuver or when the maneuver is already completed.

In the beginning, the experiment was performed on each driver individually. This allows to ignore the physiological characteristics of each driver (reaction speed, intensity of emotional expression) and find common patterns of influence of the type of maneuver and emotional expression.

So, several experiments are performed:

1) To explore the relationship between emotion and the type of maneuver being performed, the average value of emotional expression over the time interval (t_1+l, t_1+r) is analyzed Fig. 3. This experiment is repeated several times for different values of parameters l and r. In this case, l is taken from the range $(-10, -6)$ and r from $(0, 3)$. Based on the results of this experiment, it can be concluded that each driver's driving style (e.g., aggressive or non-aggressive) affects the driver's emotional state.

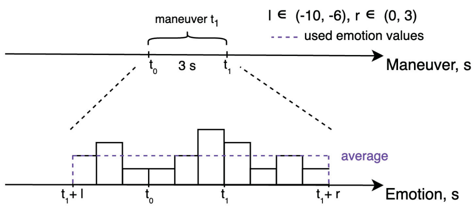

Fig. 3. The scheme of the average emotional expression score for each maneuver type

2) To study changes in emotion over time, in the surrounding area of the maneuver. It is necessary to consider the average value of emotional expression for each type of maneuver at i second relative to the end of the maneuver (t_1). This logic is shown in Fig. 4. For example, if there are two maneuvers of the same type at different time intervals, the average of the second-by-second values of emotional expression for these maneuvers is taken. In general terms, the emotion at the time $t_0 + i, i = l...r$ is evaluated for a specific range of l and r. This allows observing when a person shows emotion during maneuver performing check the hypothesis that some emotions appear for some time interval before the maneuver is performed.

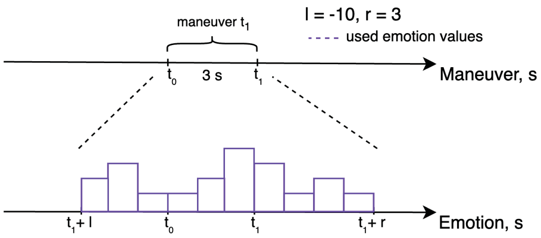

Fig. 4. The scheme of emotional expression estimation at each second around the performed maneuver

To research the relationship between emotions and the type of maneuver performed by all drivers, it is necessary to standardize the available measures of emotional expression for each driver. This is because all people have different states of emotional rest and different measures and rates of emotion in similar situations.

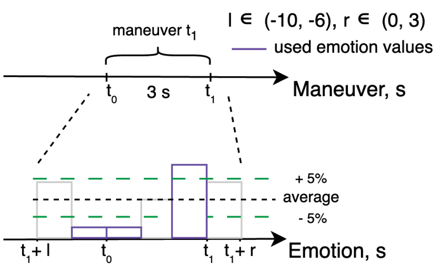

Fig. 5. The scheme of the algorithm for estimating the emotional expression without taking into account the individual characteristics of the driver

For this purpose, the relative magnitude of emotional expression is calculated. Fig. 5 shows the scheme of the algorithm for estimating emotional expressions without taking into account the individual characteristics of the driver. First, the average value of the expression of each emotion for each driver is calculated. Then the rows are filtered and the values that are outliers concerning 10% of the window relative to the calculated average are left. Based on these outliers, the average for each maneuver is calculated. The obtained values are averaged. Then the relative change of emotion on a particular maneuver is taken. This approach allows us to look at the trend of the driver's emotion change, minimizing the driver's characteristic features related to emotional expression.

Python 3.10 was used for implementation. Since a pre-trained model is used for emotion classification, significant computational resources are not required. All transformations can be performed using a CPU with 16 Gb RAM.

IV. Results

Three types of experiment were conducted. The first two were conducted for each driver separately:

- evaluation of the average emotional expression for each type of maneuver
- studying the seconds of emotion appearance depending on the type of maneuver

The third experiment was conducted for all drivers at the same time taking into account individual physiological features.

- studying the level of emotional expression depending on the performed maneuver for all drivers

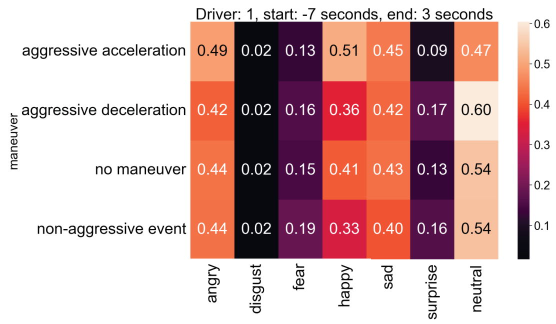

Fig. 6. Driver 1. The average expression of emotion by maneuver

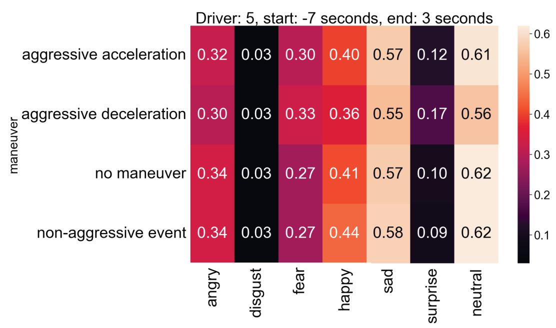

Fig. 7. Driver 5. The average expression of emotion by maneuver

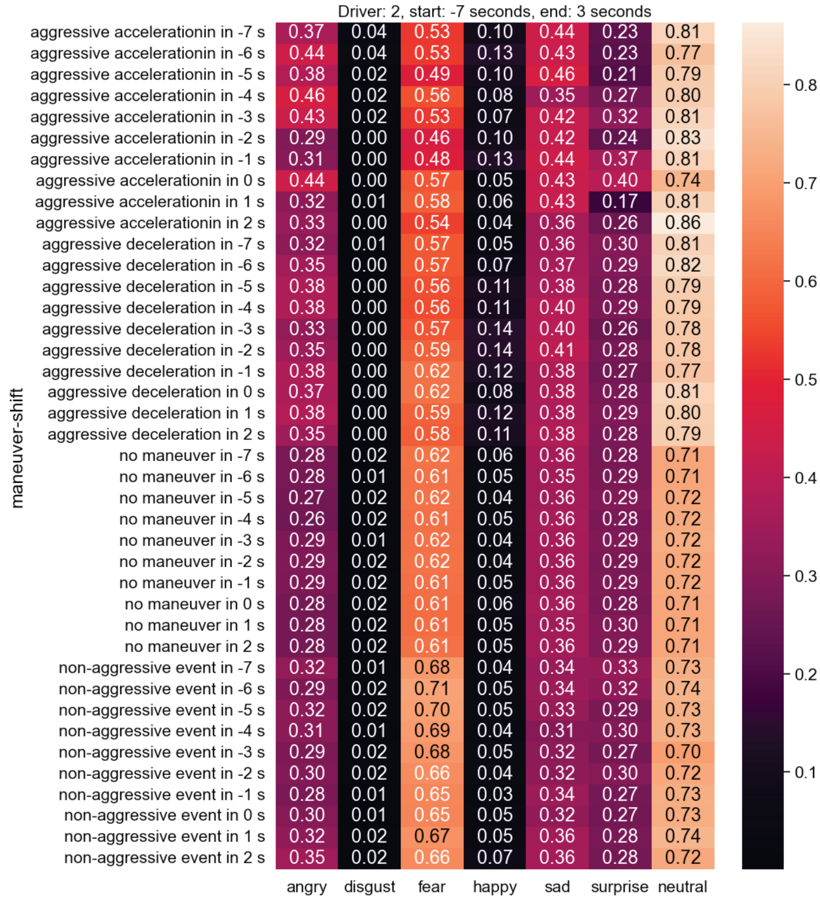

Fig. 8. Driver 2. The average of emotional expression per-second

First, the experiments were performed on each driver separately. The average value of emotional expression at a given interval (l, r) was analyzed. As parameters, l from the range $(-10, -6)$ and r from $(-1, 3)$ were enumerated. For example, for a pair $(-10, 1)$ given that the maneuver lasts 3 seconds, we look at 7 seconds before the maneuver starts and 1 second after the maneuver is performed.

As a result of this experiment, it was found that aggressive deceleration causes the greatest change in emotion. Fig. 6 and Fig. 7 show that fear and surprise increase during the aggressive deceleration maneuver. Also, both drivers have decreased anger during aggressive acceleration. Disgust does not change in two drivers and, in general, is not particularly noticeable during driving. However, the severity of disgust differs by a factor of two between drivers. This again illustrates the individual differences between drivers and the importance of taking these characteristics into account in experiments. Sadness also does not change depending on the maneuver being performed.

After the first experiment, it was logical to investigate the time of emotional occurrence for different types of performed maneuvers. Next, the second experiment was conducted. In this experiment, the mean value of emotion in each second relative to the moment of the end of the maneuver was investigated. The results for 2 drivers are presented in Fig. 8 and Fig. 9. From the presented results, it is clear that specific emotions appear and pass away at specific moments relative to the end of the maneuver t_1.

For instance, aggression increases during aggressive maneuvers, but not for the whole considered period. Moreover, at some moments, the driver notices that it is necessary to react, which leads to increased surprise, which may pass before the maneuver or remain until the time of the maneuver. It is also important to underline that when the maneuver is not performed, the driver's emotions are stable and no changes are observed.

At non-aggressive maneuvers both drivers do not change the level of expression of neutral emotional state for the whole interval under consideration. In general, this indicates the concentration of attention during driving, when the driver is focused and monitors what is on the road. At the same time, during aggressive decelerations, Driver 2 reacts to the maneuver more steadily than Driver 5. Thus, at the end of the aggressive maneuver, Driver 5's neutral state expression level decreases.

According to the results of the first experiment, the optimal range for analysis was obtained: $(-7, 3)$. However, the values of emotional expression for different drivers cannot be

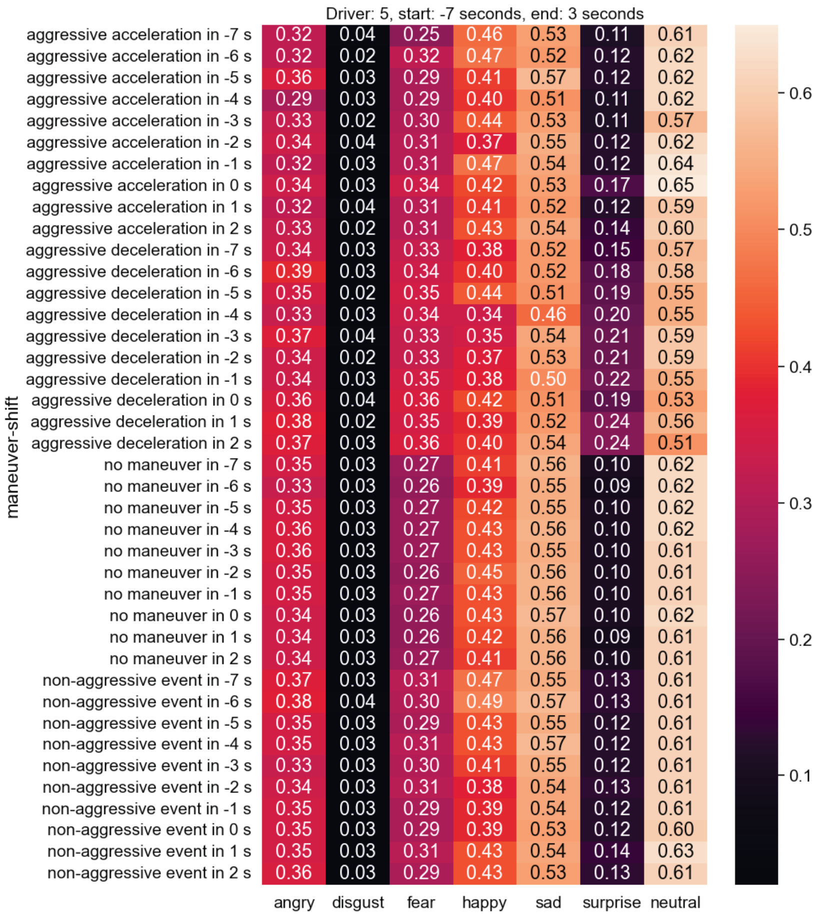

Fig. 9. Driver 5. The average of emotional expression per-second

compared due to physiological and psychological peculiarities. Therefore, to normalize the results obtained and to be able to compare them, relative values of emotional expression were considered.

Next, a third experiment was conducted. This experiment allows to estimate the expression of driver's emotions by minimizing the individual characteristics of the driver. It is impossible to get rid of such factors completely, but it is possible to reduce their influence. This allows to compare the values of emotional expressions between drivers.

To calculate the average relative value by emotion, the following steps are performed for each driver:

- calculation of the average value of emotional expressions for the whole period
- filtering of emotional expression values within the range of $\pm 5\%$ of the average
- calculating the average of the remaining values within the maneuver
- calculating the relative value as the ratio of the average of the remaining values within the maneuver and the average value of emotional expression for the whole period

Fig. 11 illustrates the average relative value of the driver. From the results obtained, the following conclusions can be drawn for most of the considered drivers:

- It can be noticed that during aggressive acceleration drivers can be divided into 2 groups. The first group is characterized by an increase in surprise and a decrease in anger. This can be interpreted as the first driver's reaction to a traffic situation that requires a response. The other group of drivers reacts in the opposite way. Surprise decreases and anger increases. Such opposite reactions can be explained by the opposites of anger and surprise. Some people react to unexpected situations with surprise and others with anger.
- At the same time, surprise does not increase before aggressive acceleration, which is logical. This is explained by the fact that a person usually plans acceleration independently, without any apparent influence of external factors.
- For the majority of drivers, neutrality is increased for aggressive maneuvers. This indicates that the driver is focused on the traffic situation.
- It is difficult to identify any trends for happiness. Reactions of drivers in this situation are very individual. However, for a non-aggressive event and driving without maneuvering, the values are at the level of average values

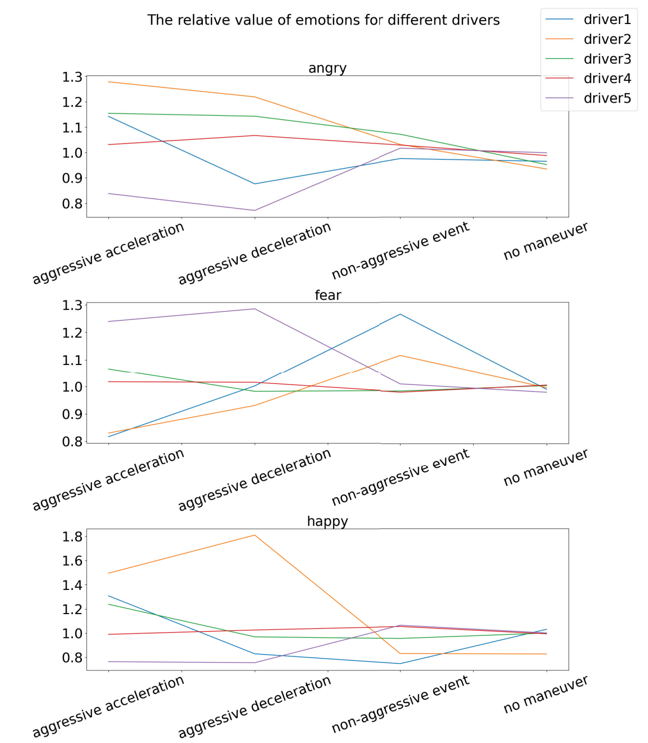

Fig. 10. The average relative value of emotional expression by the driver

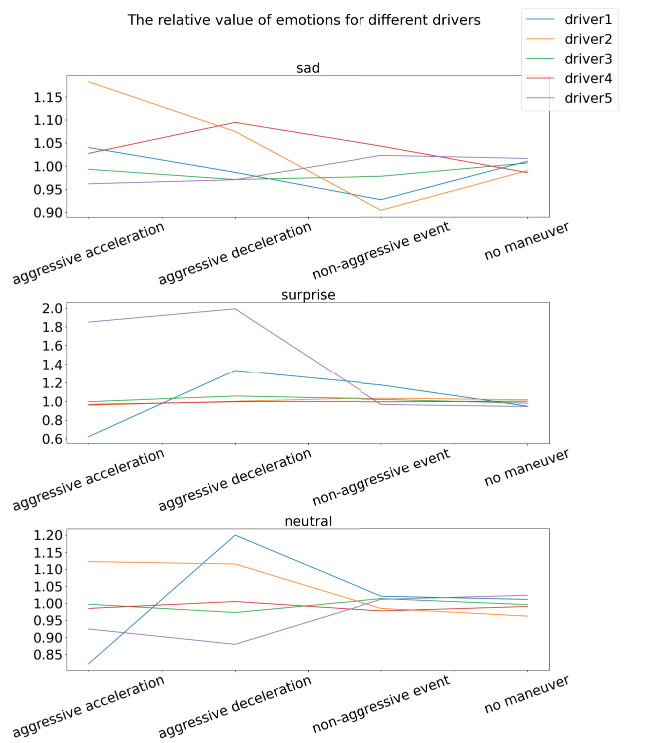

Fig. 11. The average relative value of emotional expression by the driver

for the whole time period.

- For fear, a general increasing trend is tracked for minor maneuvers and for aggressive deceleration.

From the obtained results, we can conclude that for the majority of drivers, the tendency to change emotions during different maneuvers is observed. However, there are moments when individual peculiarities of drivers appear. For a more general analysis, it is necessary to use a larger sample of drivers.

V. Conclusion

During analysis of the vehicle maneuvers and driver emotions, we conducted three types of experiments. The first two are designed to study emotions for each driver separately. The third experiment allows to compare the emotional reactions of the drivers minimizing their psychological and physiological peculiarities. To summarize, it is possible to say that for fear and surprise clear patterns can be identified. For happiness patterns are not observed in this sample of drivers, but there is similar behavior on specific maneuvers.

Experiments show that when aggressive braking is performed, a person's fear level increases in a few seconds. Moreover, it was found that emotions do not appear only at the moment of the maneuver. Experiments have shown that the greatest expression of emotions appears in 4 seconds before the start of the maneuver and ends in 3 seconds after the maneuver is performed. The correlation between fear and surprise has also been established. In the execution of aggressive inhibition. The increase in one emotion leads to a decrease in the other.

The results obtained in this research can be integrated into existing systems for alerting drivers in case of dangerous traffic situations. In this study, it is found that there is a correlation between the road maneuver being performed, and the emotions experienced by the driver. Based on the analysis of emotions while driving, the system can predict the dangerous maneuvers that the driver is going to perform.

In the future it is possible to apply the proposed methodology to a larger number of drivers and maneuvers. This will expand the possibilities for building a system that can predict driver behaviour taking into account the driver's emotions. As a consequence, it will reduce the probability of road accidents.

Acknowledgment

This work was supported by the Russian State Research FFZF-2022-0005.

References

[1] W. H. Organization *et al.*, "Road safety," 2020.

[2] M. Shahverdy, M. Fathy, R. Berangi, and M. Sabokrou, "Driver behavior detection and classification using deep convolutional neural networks," *Expert Systems with Applications*, vol. 149, p. 113240, 2020. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S095741742030066X

[3] J. S. Lerner, Y. Li, P. Valdesolo, and K. S. Kassam, "Emotion and decision making," *Annual review of psychology*, vol. 66, pp. 799–823, 2015.

[4] G. Rowe, J. B. Hirsh, and A. K. Anderson, "Positive affect increases the breadth of attentional selection," *Proceedings of the National Academy of Sciences*, vol. 104, no. 1, pp. 383–388, 2007.

[5] W. Othman, A. Kashevnik, A. Ali, and N. Shilov, “Drivermvt: In-cabin dataset for driver monitoring including video and vehicle telemetry information,” *Data*, vol. 7, no. 5, 2022. [Online]. Available: https://www.mdpi.com/2306-5729/7/5/62

[6] S. Mohanta and K. Veer, “Trends and challenges of image analysis in facial emotion recognition: a review,” *Network Modeling Analysis in Health Informatics and Bioinformatics*, vol. 11, 09 2022.

[7] W. Mellouk and W. Handouzi, “Facial emotion recognition using deep learning: review and insights,” *Procedia Computer Science*, vol. 175, pp. 689–694, 2020, the 17th International Conference on Mobile Systems and Pervasive Computing (MobiSPC),The 15th International Conference on Future Networks and Communications (FNC),The 10th International Conference on Sustainable Energy Information Technology. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1877050920318019

[8] C. Marechal, D. Mikolajewski, K. Tyburek, P. Prokopowicz, L. Bougueroua, C. Ancourt, and K. Wegrzyn-Wolska, “Survey on ai-based multimodal methods for emotion detection.” *High-performance modelling and simulation for big data applications*, vol. 11400, pp. 307–324, 2019.

[9] B. C. Ko, “A brief review of facial emotion recognition based on visual information,” *Sensors*, vol. 18, no. 2, 2018. [Online]. Available: https://www.mdpi.com/1424-8220/18/2/401

[10] M. Bartlett, G. Littlewort, M. Frank, C. Lainscsek, I. Fasel, and J. Movellan, “Recognizing facial expression: machine learning and application to spontaneous behavior,” vol. 2, pp. 568–573 vol. 2, 2005.

[11] A. Hassouneh, A. Mutawa, and M. Murugappan, “Development of a real-time emotion recognition system using facial expressions and eeg based on machine learning and deep neural network methods,” *Informatics in Medicine Unlocked*, vol. 20, p. 100372, 2020.

[12] K. Zhang, Y. Li, J. Wang, E. Cambria, and X. Li, “Real-time video emotion recognition based on reinforcement learning and domain knowledge,” *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 32, no. 3, pp. 1034–1047, 2021.

[13] N. Mehendale, “Facial emotion recognition using convolutional neural networks (ferc),” *SN Applied Sciences*, vol. 2, no. 3, p. 446, 2020.

[14] G. A. M. Meiring and H. C. Myburgh, “A review of intelligent driving style analysis systems and related artificial intelligence algorithms,” *Sensors*, vol. 15, no. 12, pp. 30 653–30 682, 2015. [Online]. Available: https://www.mdpi.com/1424-8220/15/12/29822

[15] W. Othman, B. Hamoud, A. Kashevnik, N. Shilov, and A. Ali, “A machine learning-based correlation analysis between driver behaviour and vital signs: Approach and case study,” *Sensors*, vol. 23, no. 17, 2023. [Online]. Available: https://www.mdpi.com/1424-8220/23/17/7387

[16] S. Ansari, F. Naghdy, H. Du, and Y. Pahnwar, “Driver mental fatigue detection based on head posture using new modified relu-bilstm deep neural network,” *IEEE Transactions on Intelligent Transportation Systems*, vol. PP, pp. 1–13, 08 2021.

[17] C. Huang, X. Wang, J. Cao, S. Wang, and Y. Zhang, “Hcf: A hybrid cnn framework for behavior detection of distracted drivers,” *IEEE access*, vol. 8, pp. 109 335–109 349, 2020.

[18] Y. Li, C. Lan, J. Xing, W. Zeng, C. Yuan, and J. Liu, “Online human action detection using joint classification-regression recurrent neural networks,” in *Computer Vision–ECCV 2016: 14th European Conference, Amsterdam, The Netherlands, October 11–14, 2016, Proceedings, Part VII 14*. Springer, 2016, pp. 203–220.

[19] V. Ramanishka, Y.-T. Chen, T. Misu, and K. Saenko, “Toward driving scene understanding: A dataset for learning driver behavior and causal reasoning,” in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2018.

[20] R. Chandra, M. Mahajan, R. Kala, R. Palugulla, C. Naidu, A. Jain, and D. Manocha, “Meteor: A massive dense & heterogeneous behavior dataset for autonomous driving,” *arXiv preprint arXiv:2109.07648*, 2021.

[21] O. Kopuklu, J. Zheng, H. Xu, and G. Rigoll, “Driver anomaly detection: A dataset and contrastive learning approach,” pp. 91–100, January 2021.

[22] https://github.com/jair-jr/driverBehaviorDataset.

[23] https://github.com/JustinShenk/fer.

[24] I. J. Goodfellow, D. Erhan, P. L. Carrier, A. Courville, M. Mirza, B. Hamner, W. Cukierski, Y. Tang, D. Thaler, D.-H. Lee, Y. Zhou, C. Ramaiah, F. Feng, R. Li, X. Wang, D. Athanasakis, J. Shawe-Taylor, M. Milakov, J. Park, R. Ionescu, M. Popescu, C. Grozea, J. Bergstra, J. Xie, L. Romaszko, B. Xu, Z. Chuang, and Y. Bengio, “Challenges in representation learning: A report on three machine learning contests,” in *Neural Information Processing*, M. Lee, A. Hirose, Z.-G. Hou, and R. M. Kil, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2013, pp. 117–124.