

## Emotion monitoring system for drivers

Z. Kowalczyk\* M. Czubenko\* T. Merta\*

\* *Gdańsk University of Technology, Narutowicza 11/12 80-233 Gdańsk*

kova@pg.gda.pl, micczube@pg.gda.pl, tommerta@gmail.com

**Abstract:** This article describes a new approach to the issue of building a driver monitoring system. Actual systems focus, for example, on tracking eyelid and eyebrow movements that result from fatigue. We propose a different approach based on monitoring the state of emotions. Such a system assumes that by using the emotion model based on our own concept, referred to as the reverse Plutchik's paraboloid of emotions, the recognition of emotions is carried out by means of a video camera and an external algorithm that recognizes real/internal emotions based on facial expressions. The final emotion is estimated by the Kalman filter, where the emotion is treated as measurement data. The aim of our future work is to determine the impact of the driver's emotional state on driving safety.

© 2019, IFAC (International Federation of Automatic Control) Hosting by Elsevier Ltd. All rights reserved.

*Keywords:* Affective computing, emotion recognition, car safety.

### 1. INTRODUCTION

The vast majority of automotive companies assume the introduction of autonomous vehicles in the near future. However, due to the last fatal accident (July 2018, Pheonix), the mood in society associated with autonomous cars is decreasing. This event proved it is still far to replace human drivers with machines. However most companies plans introducing autonomous cars in before 2025. Along with the development of autonomous cars, there is a high progress in systems that support drivers known as DMS (*Driver Monitoring System*) or ADAS (*Advanced Driver-Assistance System*). These include semi-autonomous systems such as lane assistant, active cruise control, automatic emergency braking, automatic parking, or driver concentration control. Both autonomous cars and driver support systems have one goal, which is to improve the safety of road users.

The police conducts special actions in periods of increased traffic on the road to provide security. Nevertheless during the last action *Znicz* (All Saints' Day in Poland) there were 408 traffic accidents with over 500 fatal injuries. Despite the development of technology, driving a car still may be dangerous. According to data from Virginia Tech, the risk of an accident is affected by the following motivations (Knapton, 2016):

- drug/alcohol – 35%
- mobile phone dialing – 20%
- reading/writing – 16%
- emotions: anger, sadness – 14%
- reaching for an object – 9%
- other – 6%.

It should be noted that the driver concentration control system which monitors human behavior, like falling asleep at the wheel, is currently produced by many car companies (e.g. Mercedes, MAN). Researchers are working on the development of new capabilities of such systems in the form of built-in modules as well as external products

(Akrouit and Mahdi, 2014). Despite car producers assure that their systems increase safety, there is probably no official statistic data which clearly indicates how much such a system affects the reduction of accidents.

#### 1.1 State of the art

Current solutions that monitor the driver mainly use cameras, eye-trackers, or contactless sensors, like a heartbeat monitor (Bergasa et al., 2006; Dong et al., 2011). There are also contact (wired) proposition based on data from electroencephalogram, electrocardiogram, electrooculogram, electromyogram, etc., however they require a probe or electrode to be connected to the driver's body, thus they are not useful in practice (Ursulescu et al., 2018; Wan et al., 2018).

Transport-related institutions, for example, the American Automobile Association Foundation for Traffic Safety, focus on lack of attention, which is why 5 categories of driver inattention status have been distinguished (Dong et al., 2011):

- attentive
- distracted (e.g. looking away from the roadway, being lost in thought, manually adjusting the radio volume)
- looked but did not see (cognitive distraction)
- sleepy
- unknown.

Clearly, the lack of attention is related to the term 'fatigue', which refers to 'a combination of symptoms such as impaired performance (disability) and a subjective feeling of drowsiness' or 'concerns the inability or disinclination to continue an activity, generally because the activity has been going on for too long' (Christopher Brill et al., 2005; Singh et al., 2011).

Many systems and algorithms have been created to check the driver status, such as PERCLOS, Ridy, Valeo, Smart Eye Pro, InSight, Driver Alert Control (Volvo), DD850

Driver Fatigue Monitor, etc. They usually suggest rest before continuing travel, or they perform an alarm when detecting sleeping or other dangerous condition (Dong et al., 2011; Ursulescu et al., 2018; Arun et al., 2011). These systems interpret the driver's external behavior, which is associated with fatigue. Our system of driver emotions recognition goes a step deeper, because it estimates the state of emotions that is internal.

The problem of recognizing emotions is relatively new. Of course appropriate detection algorithms can be found but they are mainly used for presentation of audiovisual content in entertainment, advertising or marketing (Mishra and Ratnaparkhi, 2018; Iyer et al., 2017; Kanluan et al., 2008). Various simulators of emotions which are related to automobiles (Cai et al., 2007) are also available. However, at present there are no systems that estimate the emotions of the car driver. Therefore, we propose a preliminary solution that aims to determine the emotional state of the driver. To achieve our goal we need to present basic elements related to emotions.

## 2. A LITTLE ABOUT EMOTIONS

There are many different models of emotions (Kowalczyk and Czubenko, 2016). The vast majority of them are used to describe an affective state of a man. Such models can be generally depicted in a certain space (2D or 3D). Dimensions (axes) of this space can be marked as:

- pleasure, degree of arousal and degree of relaxation (in the oldest model of emotions (Wundt, 1897))
- arousal and valence (Russell, 1980)
- activation and pleasure (a model called *circumplex* (Posner et al., 2005))
- serotonin, dopamine and noradrenaline (a neurotransmitter model (Lövheim, 2012))
- color and intensity (Plutchik, 2001).

The authors of the models presented specific points or areas in such spaces to locate specific, named emotional states. Their number varies from 3 to 28.

The issue of primal/basic emotions should also be presented here. They describe the number of possible emotional categories assuming no gradation. In such a case, the most popular model of primal emotions is the model created by Ekman et al. (2013) which distinguishes emotions such as: anger, disgust, fear, joy, sadness and surprise. These emotions are classified according to the facial expressions what is important in the context of this article. Basic emotions can be extended using the model of Plutchik (1980), which adds acceptance and anticipation to the emotions above (based on the reference to biological processes).

Some models associated with the concept of computational emotion systems are created to simulate the emotional state for robotic purposes. They focus mainly on modeling the processes of creating emotions based on external stimuli. The most famous models of this type include: FLAME (El-Nasr et al., 2000), EMA (Gratch and Marsella, 2004), Wasabi (Becker-Asano, 2008), xEmotion (Kowalczyk and Czubenko, 2017) but only FLAME and xEmotion supports grading. It is worth noting that the majority of computational emotion systems assume that emotions occur

in the form of a continuous process or individual labels. What's more, we can distinguish between different intermediate states of emotions (i.e. emotional involvement of the agent), despite the classical, sharp gradation of emotions (like annoyance, anger, rage).

## 3. THE MODEL OF EMOTIONS

The xEmotion is a part of the Intelligent System of Decision-making (Kowalczyk and Czubenko, 2011). The ISD is capable of constructing individual emotions for a virtual or robotic agent as a response to its observations and interactions. The xEmotion subsystem is based on several different theories of emotion: somatic (Zajonc et al., 1989; Damasio, 1994), appraisal (Lazarus, 1991), evolutionary Plutchik (2001). The xEmotion also includes a process personal emotions called equalia, but there is no need to model them for driver monitoring system. Thus we present only a simplified version of xEmotion subsystem. A complete description of xEmotion can be found to (Kowalczyk et al., 2019) while the details of the personal emotions can be found in (Kowalczyk and Czubenko, 2019).

The simplified version of the xEmotion subsystem has four components, shown in Figure 1. The subsystem not only has a division into appropriate color and intensity but also includes the duration of the emotional component (Oatley et al., 2012), so we can distinguish:

- emotional context of impressions (short-time pre-emotions, related to perceived and recognized features)
- emotional context of objects (short-time expressive sub-emotions, related to perceived and recognized objects)
- emotional state of the agent (classical emotion)
- mood (long-time nonlinear derivative of emotional evolution).

Pre-emotions are the most basic form of emotions. They are derived from the somatic theory of emotions which states that the response to the stimulus (e.g. tears) should appear first, emotions are created later corresponding to this response, what is a form of self-coupling. The pre-emotions are combined with the perception of some strictly predefined impressions. Sub-emotions are the second (main) emotional component in xEmotion subsystem. They are associated with objects or events recognized by the agent (appraisal theory of emotions). In the case of a human being, it is also possible to call sub-emotions by reminding, e.g. an unpleasant event. In this work the sub-emotional objects and events are determined by copying the emotional state of the user. The sub-emotions are suppressed over time (according to the Ebbinghaus's forgetting curve) or with a change of attention. Both of the above elements affect the proper emotional state of the agent, what is simply called the classical emotion. The classical emotion is suppressed over time due to the calming effect. The degree of unfulfillment of the agent's needs may influence the classical emotion at the same time (what is omitted in this work).

As the emotions suppress with time, the agent should be in a neutral state in most cases. In turn, emotional

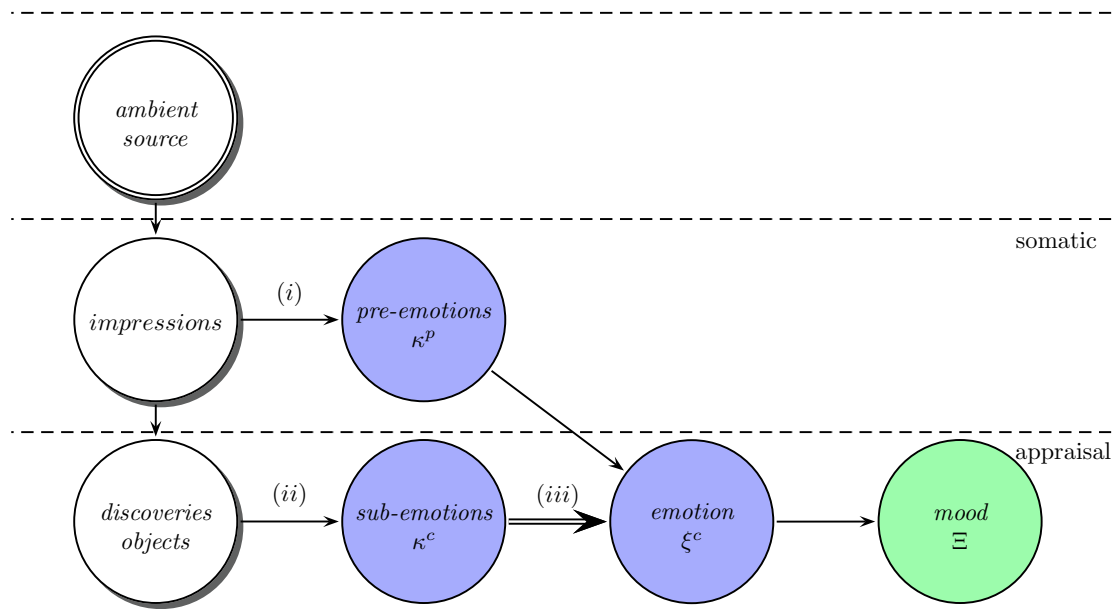


Fig. 1. Emotional components and basic relationships (Kowalczyk and Czubenko, 2017).

stimuli together with unfulfillment needs should lead to negative emotions and hence the impact of needs. The last emotional component is the mood, which is a kind of quasi-derivative of classical emotion. However, emotional stimuli together with unfulfillment needs should rather cause negative emotions, hence the impact of needs. The last emotional component is the mood, which is a kind of quasi-derivative of the classical emotion.

The emotional components, with the exception of mood, are based on Plutchik's inverted paraboloid with an added neutral state. This model is called the circle of emotions which is presented in Figure 2. The agent's condition can be described with one of 24 emotions in the circle of emotions which contains: 8 emotions with reduced intensity, 8 primal emotions and 8 emotions with increased intensity. All emotional states are based on the concept of fuzzy sets (Zadeh, 1965). Thus, every emotion can take a full value (distinct zone) in the kernel of the set, a dominant value (emotion zone) or dominated value (fuzzy zone). Such zones give us the possibility of grading emotions and inducing mixed emotions (in the place where emotions contact each other).

It should be noted that the emotions in the circle can be characterized by the intensity (radius) of color (angle). The mood is based on the TAWS function presented in (Kowalczyk and Czubenko, 2010) and takes one of three linguistic values (negative, neutral and positive).

#### 4. EMOTION RECOGNITION

Research on the recognition of facial expressions was first initiated by Ekman and Friesen (1978). He presented the coding system of mimic facial movements called FACS (*Facial Action Coding System*). Such a solution allows to unambiguously state the occurrence of certain emotions by joining individual action units. Thus, FACS became a basis for many different types of vision systems which recognize emotions – FER (*Facial Emotion Recognition*). Among them, we can distinguish systems based on:

- classic image processing algorithms
- classifiers of image features, e.g. SVM or K-means (Gupta, 2018; Ashwin et al., 2018)
- deep convolutional neural networks<sup>1</sup> (Gudi et al., 2015)
- transfer learning (Ng et al., 2015; Mollahosseini et al., 2016).

The presented methods have recently dominated the recognition of emotions due to the popularity of convolutional networks and transfer learning. Unfortunately, in both cases the images are processed automatically. This means that the image features which describe the emotions are extracted by convolutional filters, not based on detected AUs. So these systems try to recognize the artificial emotions marked on the pictures (i.e. played by actors). Therefore, such systems may not apply in real situation.

One of the emotion recognition systems based on AU is presented in (Kowalczyk and Chudziak, 2018). This system is based on color segmentation and morphological operations on the image. Localization of face and its characteristic points is achieved using a cascade of classifiers based on image features.

#### 5. THE SYSTEM

Our general proposal of emotion monitoring system assumes that the detection of emotion is based on a FER algorithm. Such an algorithm provide data related to the actual emotion. However, the actual measured value (actual value of emotion) does not reflect the actual emotional state of the driver. Therefore, our system should be provided with an appropriate state estimation method.

The general scheme of the system is presented in Figure 3. The system acquires video data from a camera which observes the driver. The camera should have a wide viewing angle to cover the face of the driver and also

<sup>1</sup> There are also systems that can recognize emotions on distorted images, e.g. (Jiang et al., 2017).

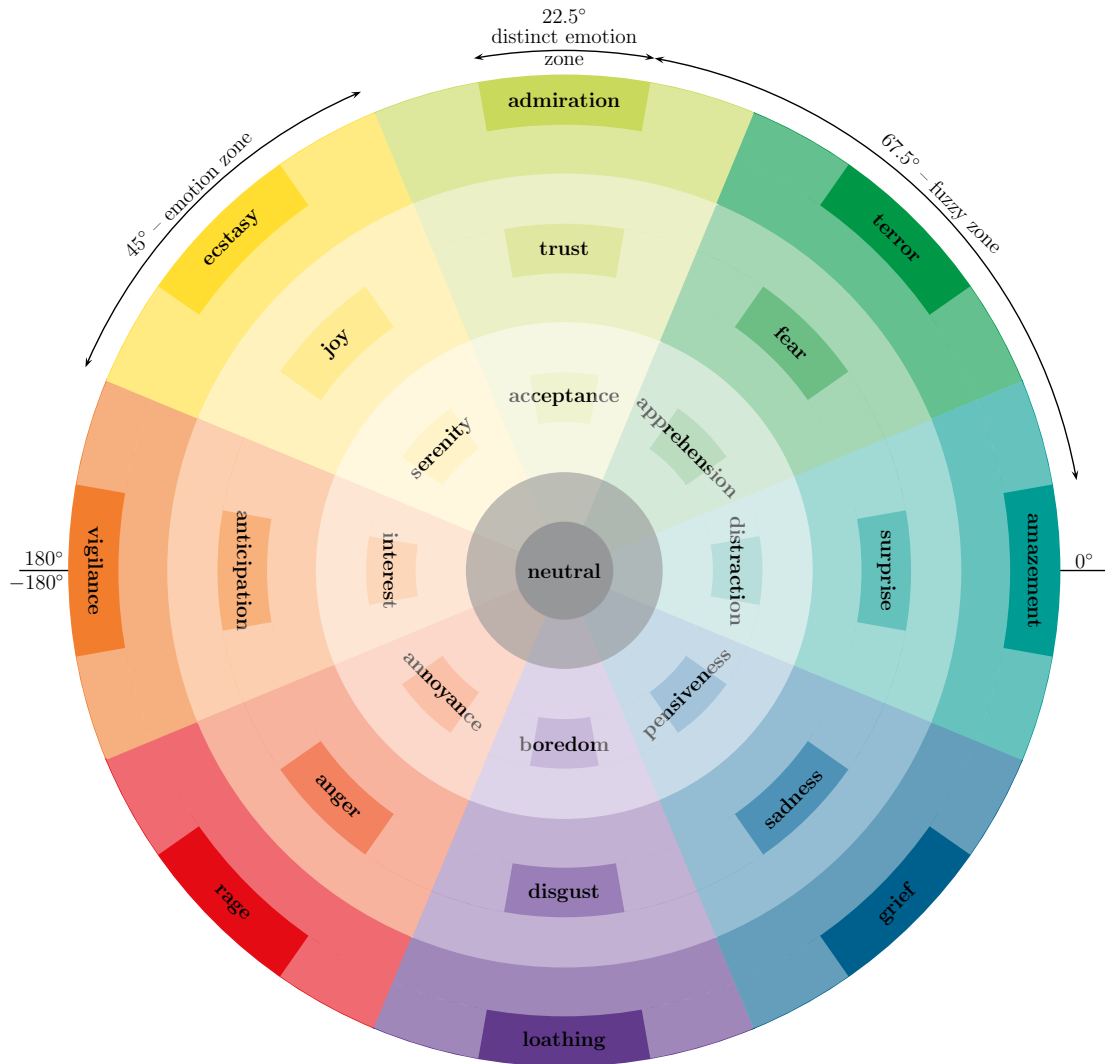


Fig. 2. The circle of emotions (Kowalczyk and Czubenko, 2017).

to provide a frontal perspective (*en face*). Video data is processed by the FER algorithm based on AU, which leads to a result which is a form of measuring the instantaneous value of the current set of emotions. As a result, we obtain the coordinates of the point on the circle of emotions corresponding to the instantaneous value of the emotion (the measured value). In this way, the measured value of the emotion variable is the basis for a proper emotion estimation. The whole algorithm repeats for consecutive frames from the camera. It is worth noting that the set of emotions visible in Figure 2 and operated by FER may not coincide. Therefore a conversion can be considered for a specific FER algorithm, which in the simplest case should omit unknown emotions.

The proposed system is flexible, in the sense that depending on the needs, different FER algorithms can be considered and the algorithm most suitable for use can be used. For example, Affectiva (FER) processes video data and determines the actual level of seven emotion indicators based on 20 metrics of facial expression. We can interpret 7 instantaneous values as sub-emotions, so their weighted geometric center is our proper measurement

value, which can be described in polar coordinates on the circle of emotions, as  $[r_m, \phi_m]^T$ .

To assess the state of emotions, we propose using a Kalman filter in which there are mechanisms to regulate the possible rate of change of emotions. When designing such a filter, one can rely on the uncertain probabilistic process described by the following expression:

$$\begin{aligned} \mathbf{x}(k+1) &= \mathbf{A}\mathbf{x}(k) + \mathbf{B}\mathbf{u}(k) + \mathbf{v}(k) \\ \mathbf{y}(k) &= \mathbf{C}\mathbf{x}(k) + \mathbf{w}(k) \end{aligned} \quad (1)$$

where  $k$  is the number of sample,  $\mathbf{v}(k)$  is the process noise,  $\mathbf{w}(k)$  is the measurement noise,  $\mathbf{A}, \mathbf{B}, \mathbf{C}$  are the matrices of the state model (matrix  $\mathbf{D}$  is not used),  $\mathbf{x}(k)$ ,  $\mathbf{u}(k)$  and  $\mathbf{y}(k)$  are state, control and measurement (observation) vector respectively. In this model, the noise is parameterized using the mean value and variance (process and measurement covariance matrices  $\mathbf{Q}$  and  $\mathbf{R}$ ):

$$\begin{aligned} \mathbf{v}(k) &\sim N(0, \mathbf{Q}) \\ \mathbf{w}(k) &\sim N(0, \mathbf{R}) \end{aligned} \quad (2)$$

Both types of noise are characterized by a zero average value and a given variance that are constant over time.

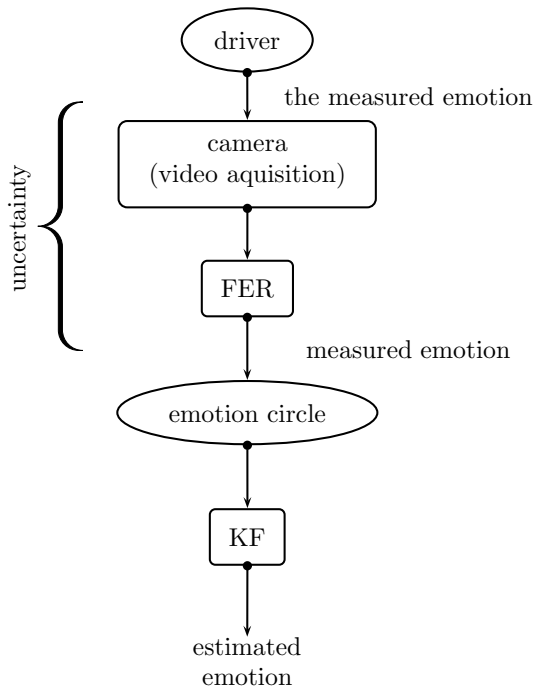


Fig. 3. A general scheme of the system.

In the proposed system, we focus on the estimation of the current emotion and the speed of emotion change. Thus, state variables are polar coordinates of the current value of emotions and corresponding derivatives. Assuming that the FER algorithm provides measurements of current emotions (as positions on the circle  $[r_m, \phi_m]^T$ ), one can observe all state variables. The state vector used and the transition and output matrices are as follows

$$\mathbf{x} = \begin{bmatrix} r \\ \phi \\ \Delta r \\ \Delta \phi \end{bmatrix}, \mathbf{A} = \begin{bmatrix} 1 & 0 & dT & 0 \\ 0 & 1 & 0 & dT \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \mathbf{C} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \quad (3)$$

where, due to the autonomous nature of this system, it has no input channel, i.e.  $\mathbf{B}$  is zero.

It is worth noting that the measurement noise is associated with elements such as blinking eyes, yawning, head movement. Therefore the matrix  $\mathbf{R}$  should be selected according to the habits of the driver.

## 6. CONCLUSIONS

The proposed system of emotion monitoring for drivers is based on the emotion model, FER algorithm based on AU, and Kalman filtration. This system give us the opportunity to observe the driver's emotions, in order to estimate the impact of emotions on driving safety. This solution is at an early stage of work, which is why simulation tests and experimental studies are necessary to evaluate the system. Of course, the effectiveness of measuring emotions depends primarily on the FER algorithm. Therefore, in future research work, these algorithms should be evaluated in the context of our system.

## REFERENCES

- Akrouf, B. and Mahdi, W. (2014). Spatio-temporal features for the automatic control of driver drowsiness state and lack of concentration. *Machine Vision and Applications*, 26(1), 1–13.
- Arun, S., Murugappan, M., and Sundaraj, K. (2011). Hypovigilance warning system: A review on driver alerting techniques. In *IEEE Control and System Graduate Research Colloquium*, 65–69. Shah Alam, Malaysia.
- Ashwin, D.V., Kumar, A., and Manikandan, J. (2018). Design of a Real-Time Human Emotion Recognition System. In N. Kumar and A. Thakre (eds.), *Ubiquitous communications and network computing*, volume 218 of *Lecture Notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering*, 177–188. Springer, Cham.
- Becker-Asano, C. (2008). *WASABI: Affect Simulation for Agents With Believable Interactivity*. Ph.D. thesis, Faculty of Technology, University of Bielefeld.
- Bergasa, L.M., Nuevo, J., Sotelo, M.A., Barea, R., and Lopez, M.E. (2006). Real-time system for monitoring driver vigilance. *IEEE Transactions on Intelligent Transportation Systems*, 7(1), 63–77.
- Cai, H., Lin, Y., and Mourant, R. (2007). Study on driver emotion in driver-vehicle-environment systems using multiple networked driving simulators. In *Driving Simulation Conference, North America 2007*, 1–9.
- Christopher Brill, J., Hancock, P., and D Gilson, R. (2005). Driver fatigue: Is something missing? 138–142. Rockport, Maine, USA.
- Damasio, A. (1994). *Descartes' Error: Emotion, Reason, and the Human Brain*. Gosset/Putnam, New York.
- Dong, Y., Hu, Z., Uchimura, K., and Murayama, N. (2011). Driver inattention monitoring system for intelligent vehicles: A review. *IEEE Transactions on Intelligent Transportation Systems*, 12(2), 596–614.
- Ekman, P., Friesen, W.V., and Ellsworth, P. (2013). What emotion categories or dimensions can observers judge from facial behavior? In A.P. Goldstein and L. Krasner (eds.), *Emotion in the human face: guidelines for research and an integration of findings*, 57–67. Elsevier Science.
- Ekman, P. and Friesen, W.V. (1978). *Facial Action Coding System: A technique for the measurement of facial action*.
- El-Nasr, M.S., Yen, J., and Ioerger, T.R. (2000). Flame - fuzzy logic adaptive model of emotions. *Autonomous Agents and Multi-Agent Systems*, 3(3), 219–257.
- Gratch, J. and Marsella, S. (2004). A domain-independent framework for modeling emotion. *Cognitive Systems Research*, 5(4), 269–306.
- Gudi, A., Tasli, H.E., den Uyl, T.M., and Maroulis, A. (2015). Deep learning based FACS Action Unit occurrence and intensity estimation. In *11th IEEE international conference and workshops on automatic face and gesture recognition (FG)*, 1–5. IEEE.
- Gupta, S. (2018). Facial emotion recognition in real-time and static images. In *2nd international conference on inventive systems and control (ICISC)*, 553–560. IEEE.
- Iyer, A.V., Pasad, V., Sankhe, S.R., and Prajapati, K. (2017). Emotion based mood enhancing music recommendation. In *2nd IEEE International Conference on Recent Trends in Electronics, Information Communica-*

- tion Technology (RTEICT), 1573–1577.
- Jiang, R., Ho, A.T., Cheheb, I., Al-Maadeed, N., Al-Maadeed, S., and Bouridane, A. (2017). Emotion recognition from scrambled facial images via many graph embedding. *Pattern Recognition*, 67, 245–251.
- Kanluan, I., Grimm, M., and Kroschel, K. (2008). Audio-visual emotion recognition using an emotion space concept. In *16th European Signal Processing Conference*, 1–5. Lausanne, Switzerland.
- Knapton, S. (2016). Which emotion raises the risk of a car crash by nearly 10 times?
- Kowalczyk, Z. and Czubenko, M. (2017). Emotions embodied in the SVC of an autonomous driver system. *IFAC-PapersOnLine*, 50(1), 3744–3749.
- Kowalczyk, Z. and Chudziak, P. (2018). Identification of Emotions Based on Human Facial Expressions Using a Color-Space Approach. In *Springer*, 291–303.
- Kowalczyk, Z. and Czubenko, M. (2010). Model of human psychology for controlling autonomous robots. In *15th international conference on methods and models in automation and robotics*, 31–36.
- Kowalczyk, Z. and Czubenko, M. (2011). Intelligent decision-making system for autonomous robots. *International Journal of Applied Mathematics and Computer Science*, 21(4), 621–635.
- Kowalczyk, Z. and Czubenko, M. (2016). Computational approaches to modeling artificial emotion—an overview of the proposed solutions. *Frontiers in Robotics and AI*, 3(21), 1–12.
- Kowalczyk, Z. and Czubenko, M. (2019). Qualia – A note on personal emotions representing as the temporal form of impressions. *Transactions on Affective Computing*, x(x), xx–xx. In review.
- Kowalczyk, Z., Czubenko, M., and Merta, T. (2019). Interpretation and modeling of emotions managing autonomous robots, based on the paradigm of scheduling variable control. *Engineering Applications of AI*, x(x). In review.
- Lazarus, R.S. (1991). *Emotion and Adaptation*. Oxford University Press, USA, New York.
- Lövheim, H. (2012). A new three-dimensional model for emotions and monoamine neurotransmitters. *Medical Hypotheses*, 78(2), 341–8.
- Mishra, P.P. and Ratnaparkhi, P.S. (2018). Hmm based emotion detection in games. In *3rd International Conference for Convergence in Technology (I2CT)*, 1–4. Pune, India.
- Mollahosseini, A., Chan, D., and Mahoor, M.H. (2016). Going deeper in facial expression recognition using deep neural networks. In *IEEE winter conference on applications of computer vision (WACV)*, 1–10. IEEE.
- Ng, H.W., Nguyen, V.D., Vonikakis, V., and Winkler, S. (2015). Deep Learning for Emotion Recognition on Small Datasets using Transfer Learning. In *Proceedings of the 2015 ACM on international conference on multimodal interaction - ICMI '15*, 443–449. ACM Press, New York, New York, USA.
- Oatley, K., Keltner, D., and Jenkins, J. (2012). *Understanding Emotions*. Blackwell Publishing, 2nd edition.
- Plutchik, R. (1980). A general psychoevolutionary theory of emotion. In R. Plutchik and H. Kellerman (eds.), *Emotion: theory, research, and experience*, volume 1, 3 – 33. Academic, New York.
- Plutchik, R. (2001). The nature of emotions. *American Scientist*, 89, 344.
- Posner, J., Russell, J.A., and Peterson, B.S. (2005). The circumplex model of affect: an integrative approach to affective neuroscience, cognitive development, and psychopathology. *Development and Psychopathology*, 17(3), 715–34.
- Russell, J.A. (1980). A circumplex model of affect. *Journal of Personality and Social Psychology*, 39(6), 1161–1178.
- Singh, H., Bhatia, J.S., and Kaur, J. (2011). Eye tracking based driver fatigue monitoring and warning system. In *India International Conference on Power Electronics 2010 (IICPE2010)*, 1–6. New Delhi, India.
- Ursulescu, O., Ilie, B., and Simion, G. (2018). Driver drowsiness detection based on eye analysis. In *13th International Symposium on Electronics and Telecommunications (ISETC)*, 1–4. Timisoara, Romania.
- Wan, W.H., Tsang, Y.T., Zhu, H., Koo, C.H., Liu, Y., and Lee, C.C.T. (2018). A real-time drivers' status monitoring scheme with safety analysis. In *IECON 2018 - 44th Annual Conference of the IEEE Industrial Electronics Society*, 5137–5140. Washington DC, USA.
- Wundt, W. (1897). Outlines of Psychology. In *Classics in the history of psychology*. York University 2010.
- Zadeh, L.A. (1965). Fuzzy sets. *Information and Control*, 8, 338–353.
- Zajonc, R.B., Murphy, S.T., and Inglehart, M. (1989). Feeling and facial efference: implications of the vascular theory of emotion. *Psychological Review*, 96(3), 395–416.