Griesche, S., Dziennus, M.: (2013): Images in mind – Design metaphor and method to classify driver distraction in critical situations. Contribution to 7. VDI-Tagung "Der Fahrer im 21.Jahrhundert", 5.-6. Nov. 2013 Braunschweig, Germany, In: VDI-Berichte 2205, pp. 85 – 99, VDI Verlag, ISBN 978-3-18-092205-8

Images in mind – Design metaphor and method to classify driver distraction in critical situations

Stefan Griesche, Marc Dziennus, DLR, Braunschweig

Abstract

The paper presents a driver model which classifies visual distraction based on the detection of atypical driving behavior. The model forwards the information to a driver adaptive collision mitigation system (CMS) and activates the acoustic warning earlier in case of distraction. Therefore the model requires the knowledge of the normal driving behavior. For that reason we introduce a design metaphor. We use the human memory and its ability to build up mental representations. Based on the idea to interpret multivariate time series as gray level images we adapted the concept of mental images to learn a situation based normal driving behavior. The model transfers the property of the long term memory to store, to interfere and to forget prototypes of mental images. We compare the stored prototypical image with the current image to obtain a distraction index. If the index reaches a certain threshold value the acoustic warning is presented.

1 Introduction

Latest studies of the Federal Statistical Office [1] confirm the positive trend of a decreasing number of traffic fatalities per year in Germany. To continue this positive development many OEM's offer already a various number of advanced driver assistance systems (ADAS) for their middle class to prevent accidents. Examples of such ADAS are blind spot warners, lane-keeping assists or emergency brake assists. Even though the costs for ADAS could be reduced over the last years the market share of ADAS is still low [2].

The EU-project interactive with its sub project "cost-efficient emergency intervention for collision mitigation" (EMIC) [3] addresses the issue of acceptance and costs of ADAS. The goal is to make ADAS available for more costumers and hence to increase the market share in the lower price segment. Therefore the sub project concentrates on two uses cases: a collision mitigation system (CMS) and an emergency steer assist (ESA). In this paper we will focus on CMS. More precisely we want to improve the acceptance of the CMS. Therefore the partners of the subproject developed a driver model. The driver model optimizes the warning and activation time of the CMS to avoid and to reduce false alarms as shown in fig. 1.

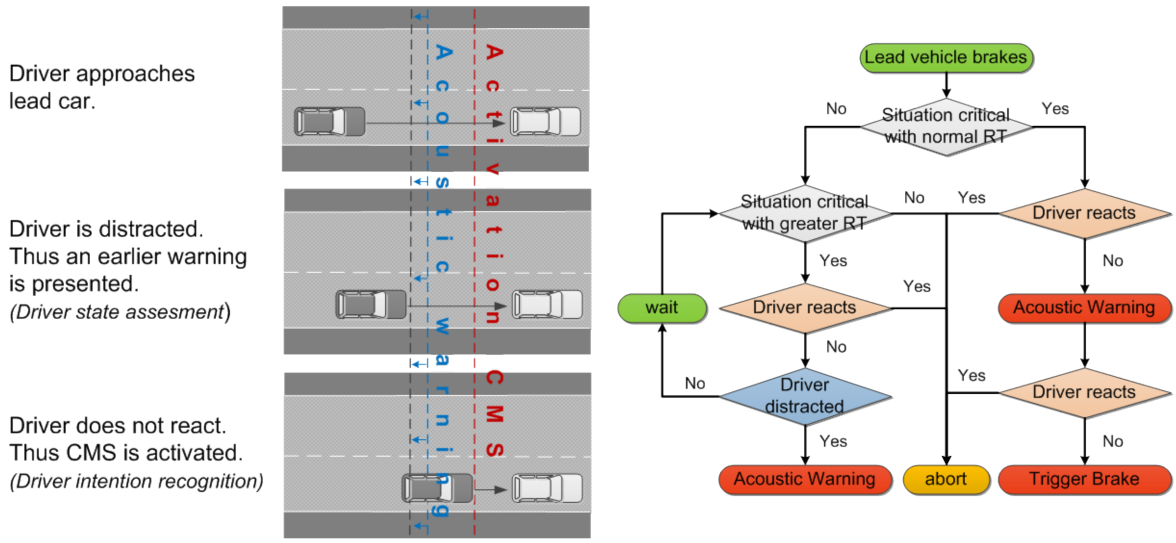

Figure 1: Process flow of the CMS

The driver model consists of two parts: a driver intention recognition [4] and a driver state assessment. We will explain in the detail the driver state assessment part. The aim of the driver state assessment is to detect visual distracted drivers. Our model takes the cost efficient approach of the subproject into account and does not require additional sensors like eye tracking cameras. Instead we refer to works of [5] or [6] and estimate distraction based on the detection of atypical driving behavior. The detection of atypical driving behavior requires the knowledge of the natural driving behavior. Thus the main task of the driver model is to learn the natural driving behavior. We will explain the learning in detail after the next chapter. Before that we will introduce a design metaphor. This design metaphor will help us to describe and to understand the concept and structure of our driver model. For the evaluation of the model we will concentrate on critical situations in car following as one typical use case of CMS.

2 Human memorizing as design metaphor for modeling natural driving behavior

2.1 Motivation

In the literature many approaches can be found to model the natural driving behavior ([7-11]). Thereby the underlying structure/method of the model highly depends on the examined problem. Common approaches are rule based methods ([10-13]) machine learning algorithms ([6], [14-15]) or combination of both ([7], [16]). Our driver model has to face the problem to learn the natural driving behavior in car following situations. This means the model has to be able to deal with temporal information/sequences and a couple different input parameters (time headway, time to collision, velocity/acceleration ego car, velocity/acceleration lead car, steering angle, brake pedal position, etc.)

Clearly drivers differ in their driving style especially in car following situations. Therefore rule based methods, like fuzzy logic which require a priori knowledge are less applicable to the problem. In contrast adaptive fuzzy logic (like in [4] or [10]) considers the individual differences and overcomes the drawback. Moreover if an adaptive fuzzy logic model uses a state machine and graphs it is able to represent temporal information as explained in [4]. But the high number of possible states and input parameters of our use case makes the method complex and hard to analyze.

Neural networks share these two issues. If we include the temporal information to a neural network it gets large and not traceable anymore. Even though neural networks are motivated by the metaphor of the human brain their learning behavior is a black box as remarked in [17]. The gap of methods which are able to handle temporal information/sequences, to learn individual behavior and are easily understandable motivated us to develop a new approach using a design metaphor. The design metaphor is motivated by the human memory and the concept of mental representation. The advantages of the design metaphor are first to help to understand the mathematical part of learning and pattern recognition of the driver model more easily and second to overcome the handicap of complexity and traceability.

2.2 The design metaphor

For the description of the design metaphor we orientate on works in neuroscience (see [18-21]). Our aim is not to develop a driver model which exactly uses the processes of the human memory. Rather we want to gain benefit by general concepts of memorizing temporal information and developing mental representations. An ability of the human memory is to cluster single experiences made with items or imaginations context based into categories. We call a concept the mental representation of a category [22]. Additional we define a prototype as the average concept of all experiences made within a category [22]: Thus the prototype is a representative of a category.

The idea of a mental representation linked with a prototype is the first idea we transfer to our model. We will use mental representations of prototypes to learn the natural driving behavior context based. In our case the context/category is a description of the current car following situation. We will explain the derivation of a category in more detail in the next chapter. The clue to link mental representations and driving behavior is possible by the idea to transform multivariate time series in a map as done for instance by [23]. Figure 2 illustrates the transformation.

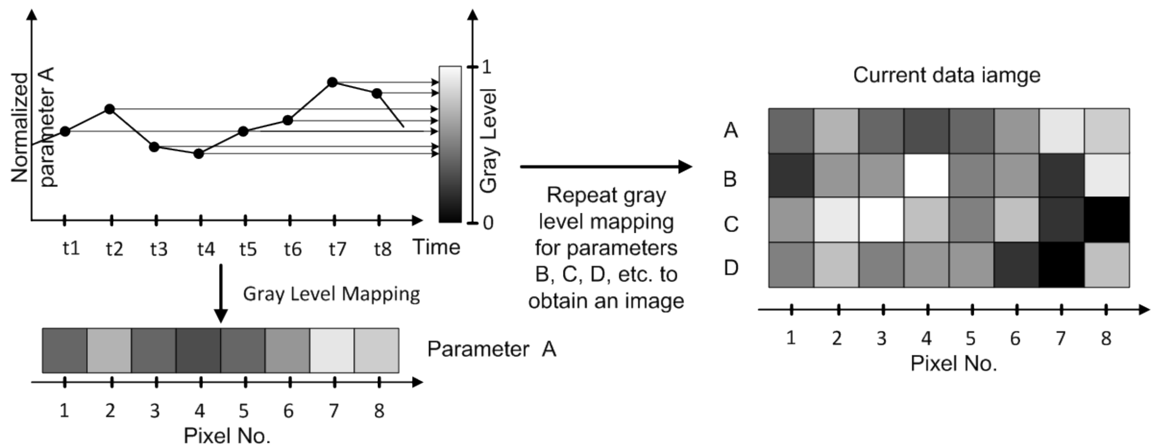

Figure 2: Transformation of discrete multivariate time series into a gray image

To stick with the design metaphor we refer to the term image rather than map. Thus speaking in terms of the design metaphor the natural driving behavior is defined as the median image and is a prototype for the current car following situation. Additional to the idea of mental representations we also want to adopt the aspect of how we store knowledge in our memory (i.e. how we build up a representation) and how we remember the mental representations. The advantage for our model is a simple understandable process to organize/store temporal information/sequences of incoming data and learn from the data at the same time. Consequently we apply the human memory structure to our model.

After Baddeley [18] the memory consists of three parts: sensory memory, working memory and long term memory (LTM). We will refer to a comment of Baddeley [18] and will rather use the term short term memory (STM) because we will focus on the storage capacity. For simplicity we assume the transfer of knowledge between STM and LTM is based on the repetition of knowledge [21]. Additional we consider the some aspects of working memory, more precisely the visual spatial sketchpad by Baddeley.

We take over the concept of rehearsal (i.e. repetition to avoid forgetting), the limited capacity (i.e. the possibility to forget knowledge) and the ability to store images in the memory. The last point justifies the idea of images as mental representations. Regarding the LTM we transfer the act of remembering [20]. This includes interference of representations, fuzzy representations and also the forgetting of representations. The properties of the STM and LTM allow us to save computational memory, to learn relevant knowledge and be online capable. In the next step we will explain the implementation of the design metaphor to our driver model.

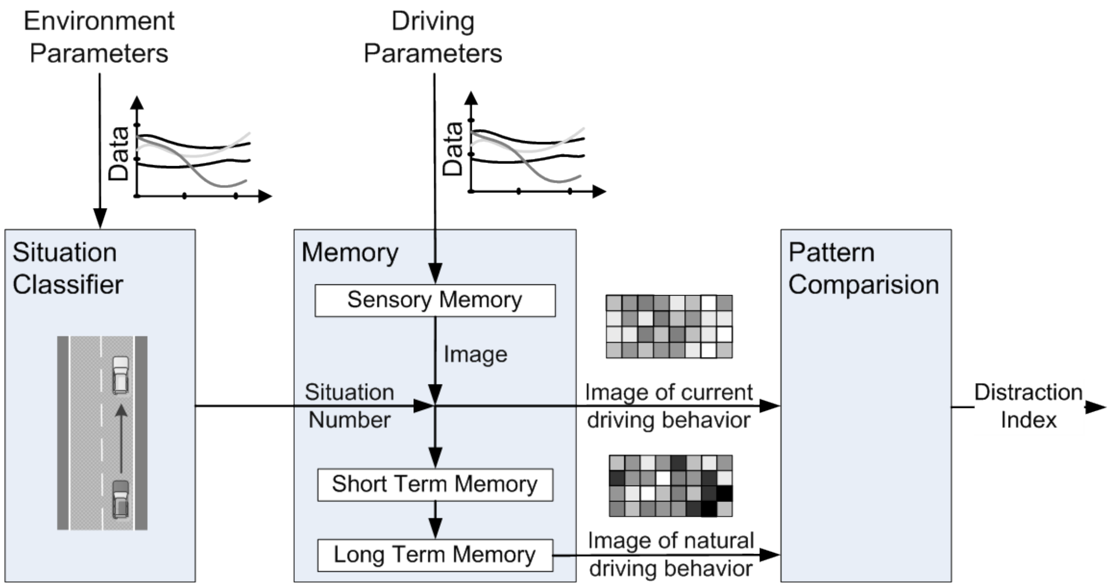

Figure 3: Schematic structure of the distraction classifier as part of the driver model

3 The distraction classifier

The distraction classifier is the part of the driver model which classifies the distraction. The output of the classifier is a distraction index which is a linear value between 0 and 1. The linearity reflects the varying duration of visual distraction from short to long. The classifier consists of three parts which are motivated by the design metaphor. Fig. 3 illustrates the interaction between the three parts. We explain the implementation of each part in the following. The first part is the situation classifier.

3.1 Situation Classifier

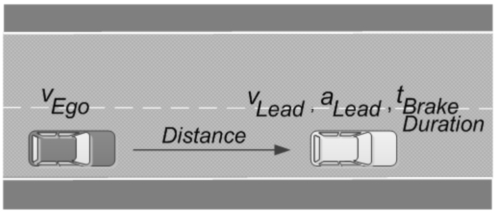

Figure 4: Input parameters for the situation classifier: Current velocity of the ego car (v_ego), current velocity/ acceleration of the lead car (v_lead/a_lead), duration of the braking maneuver from the lead car (t_BrakeDuration) and the distance between ego car and lead car

The situation classifier clusters and describes the situation/context around the vehicle. The classifier analyzes the environment parameters and transforms the current value into a state number. Table 1 explains the transformation in more detail. Each parameter has a predefined number of states. All states together of the involved parameters define a so called situation number (SN) as table 2 shows. If we recall the design metaphor the situation number represents a category of our mental representation.

Table 1: Transformation of input parameter distance in its state number

Distance	State
> 50m	0
< 50m & >= 25m	1
< 25m & >= 15m	2
< 15m & >= 10m	3
< 10m & >= 5m	4
< 5m	5

Table 2: Composition of the situation number based on example states of the inputs

Parameter	State
v_Ego	3
v_Lead	2
a_Lead	1
t_BrakeDuration	2
Distance	1
Situation Number:	*32121*

3.2 Memory

The memory part is the heart of the algorithm. In the first step the sensory memory transforms the incoming data of the driving behavior (like steering angle, brake pedal position or lateral deviation) into the current data image. Corresponding to the metaphor we only consider the last seconds of the time series. The model than connects the image with current situation number and forwards it to the STM. In other words according to the design metaphor we store an experience made (current driving behavior) to its corresponding category (current situation number). This allows us in the following to build up a prototype.

The challenge of the memory is to handle the huge number of possible situations. For example if the situation classifier uses five parameters and each parameter has five states a maximal number of 5^5 = 3125 situation numbers are possible. Hence it is impossible due to the limited training time to collect data for all situations and built up proper prototypes. That's why it is utmost important to have a strategy how to save data and learn from data. At this point we refer to the introduced design metaphor. We use the idea of a STM and a LTM to store only representative data.

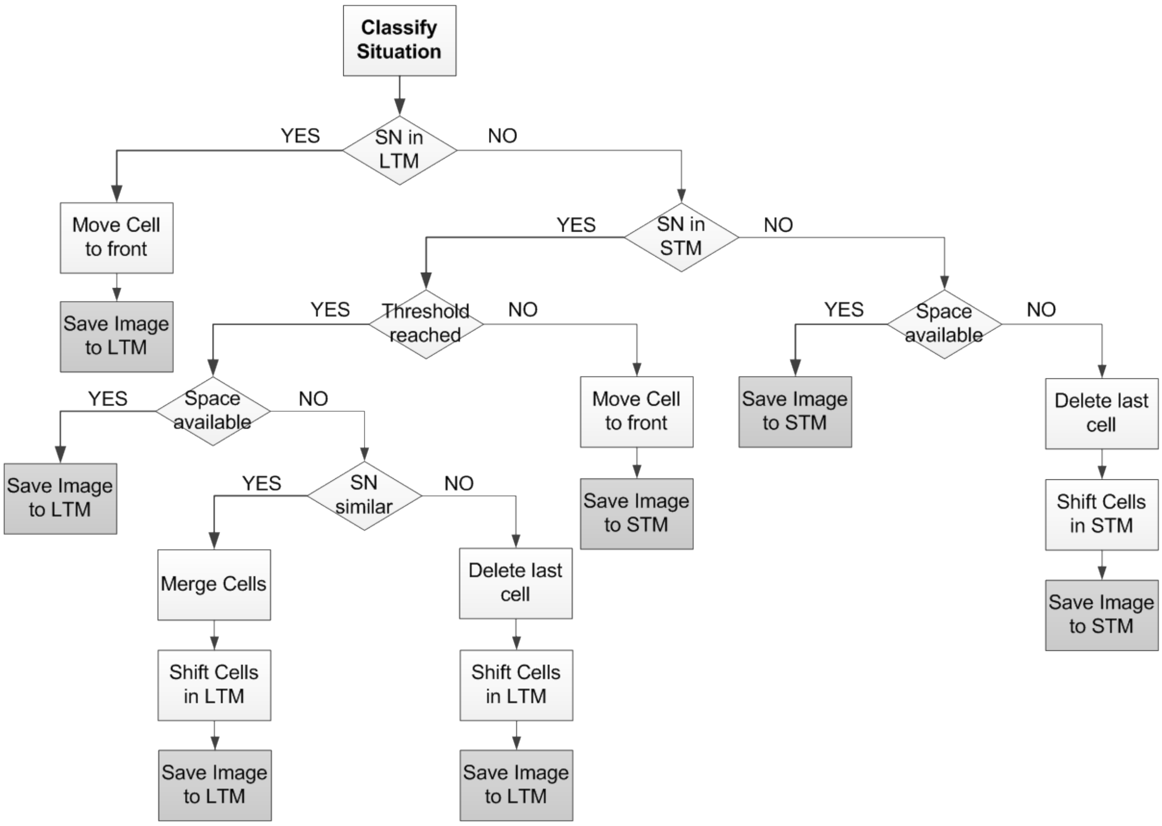

Figure 5: Process flow of the memory as decision tree

We call the place to store data a (memory) cell. A memory cell associates the situation number with its images. We model the described properties (storing & remembering) of the STM and LTM from the previous chapter by deleting, moving and comparing cells (see fig. 5).Therefore we have additional to define activated cells and similar cells. We call an activated cell a cell whose situation number is equal to the current situation number calculated by the situation classifier. We call a similar cell a cell whose situation number only differs at one digit by +/- 1 compared to the current situation number calculated by the situation classifier. With these definitions we can explain the structure of the STM/LTM. The structure of the STM is motivated by the concept of rehearsal. Consequently we limit the number of cells. Moreover we order cells in a stack to add a temporal component as shown in fig. 6.

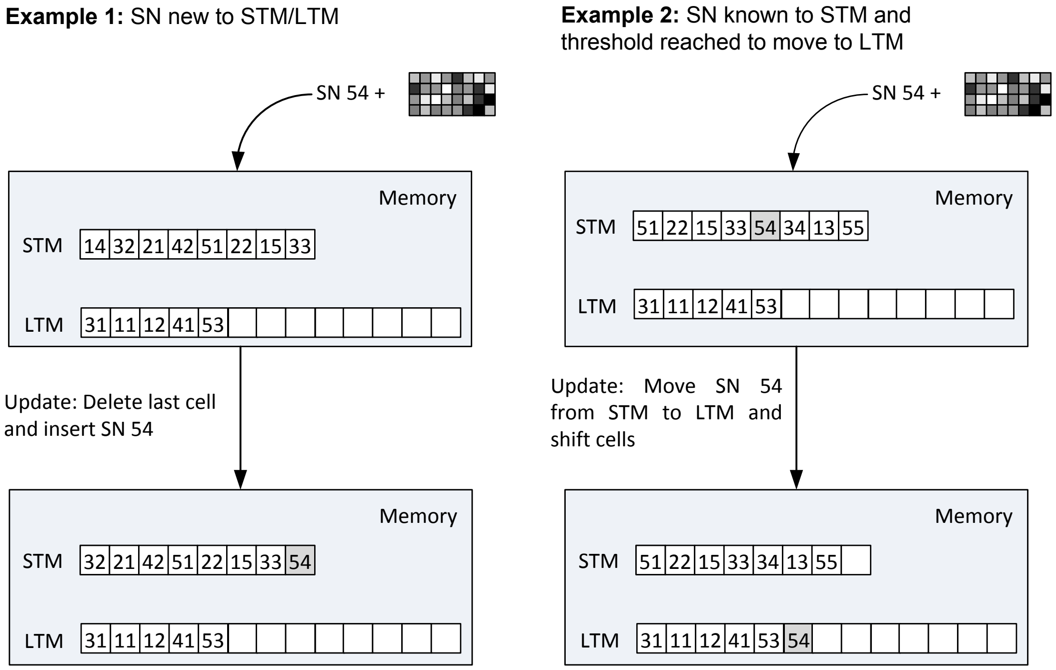

Figure 6: Illustration of the process flow on two arbitrary examples

We move an activated cell to the front of the stack to emphasize a recent occurred situation. In case all cells are occupied we delete the last cell in the stack to forget irrelevant knowledge and to have new space available. All together the STM reduces the amount of data and filters relevant data for the representation of the natural driving behavior. The model transfers knowledge from STM to LTM when a fixed number of activations for a cell occurred, for example ten. The stored knowledge in LTM than allows us to build up a representation about the natural driving behavior for a given situation. In contrast to the human LTM [20] we limit the capacity of the LTM within the model as well. The aim of the limitation is firstly to fasten up the algorithm and secondly to add interference and fuzziness to the model behavior. More precisely if we transfer knowledge from STM to LTM and all cells of the LTM are occupied we have to generate new space to store the transferred knowledge. In contrast to the STM we do not immediately delete the last cell. Rather we search for similar cells within the LTM to share knowledge. In case we find a similar cell we merge both cells, i.e. we merge the stored images of both cells to extend the examples of representations for the current situation number.

The merging of both cells describes the fuzziness and interference of our approach. In terms of our design metaphor this means older images/concepts are fuzzy or vanish and hence influence our representation about the natural driving behavior. Combining all ideas we are finally able to model a situation/context based natural driving behavior. Therefore we consider all images stored in a cell and calculate for each pixel the median value as illustrated in fig. 7. The result is a median image which represents the natural driving behavior.

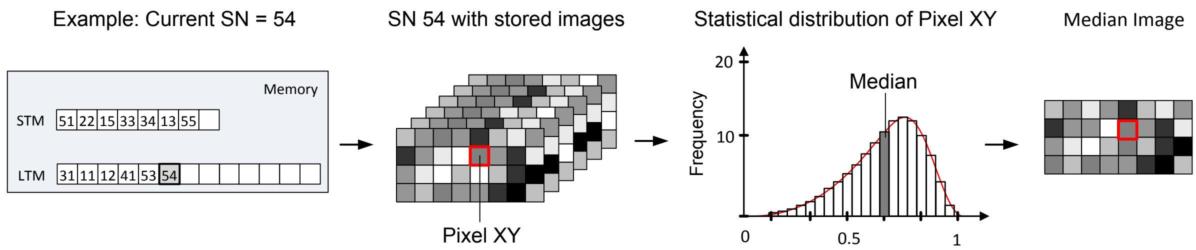

Figure 7: Derivation of the median image: The model searches the current SN in the memory, looks up the associated images and calculates based on the statistical distribution the median value pixel-wise

The derivation of the median image as representative of the natural driving behavior allows us to define and to detect atypical driving behavior and hence to detect and to classify distraction. This work is done by the last part of the distraction classifier.

3.3 Pattern Comparison

The pattern comparison part uses the advantage of the design metaphor and does pattern recognition by image processing algorithms. The basic idea is to take the median image as prototype for the current situation and to compare it with the current driving behavior image. Since we deal with different input parameters it is crucial to have a comparable measure between those. We decided to use the knowledge of the statistical distribution from the memory and to calculate the percentile value for each pixel of the current image. Than we can easily calculate the difference and get an offset image which can be further analyzed. Fig. 8 demonstrates the derivation of the offset image. Furthermore fig. 9 shows three typical approaches of pattern recognition in image processing for the analysis of the offset image.

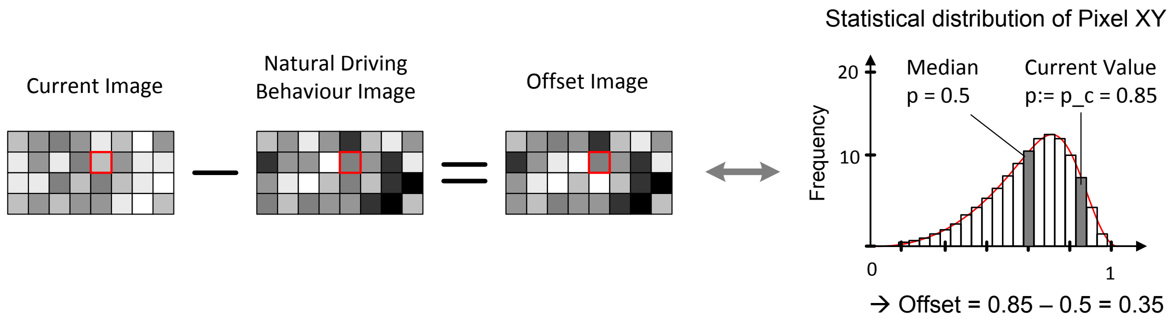

Figure 8: Derivation of the offset image: For each pixel the model computes the percentile value p_c and subtracts it from the median

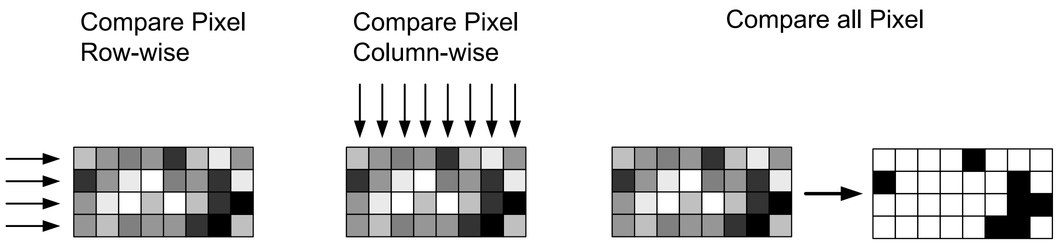

Figure 9: Different variants of pattern recognition in image processing

We decided to apply the first variant and compare the pixel row-wise using an integral approach. Moreover we suggest to add an image filter. The filter allows to weight every single pixel differently with a weight w_{ij} to emphasize the different importance of parameters on the one side and on the other side the different temporal importance. Consequently the final formula to compute the distraction index (DI) is given by:

$$DI = \sum_{i=0}^{m} \int_{t_0}^{t_n} \left(w_i(t) * \left(p_{c_i} - 0.5\right)^2 \right) dt \;\; with \;\; w_i(t_j) = w_{ij};\; p_{c_i}(t_j) = \; p_{c_{ij}} \;\; for\ all\ j = 1 \ldots n$$

Where m equals the number of parameters used and n equals the number of time stamps considered. The distraction index itself is a linear value between 0 and 1. The linearity reflects the varying duration of visual distraction from short to long. Finally we consider the case that a situation number is unknown to the memory after a defined training time. For that reason we assume the situation itself is atypical and set the distraction index by default to 1.

3.4 Uncertainty

The concept of observability in control theory motivated us to consider the uncertainty of the estimated distraction index. Our aim is to reduce false estimations by an additional parameter. The term uncertainty describes the quantity of the driving parameters involved in the estimation. If a driving parameter as a high uncertainty we force the algorithm to ignore the parameter for its estimation of the distraction index. In contrast if the uncertainty is low we include the parameter in the estimation. As explained in the section before we estimate distraction by analyzing atypical behavior of the driver. This requires a clear description of the natural driver behavior to detect any changes within the behavior. As explained our approach defines the natural driver behavior as the median image stored in the memory.

This allows us to build a linkage to the design metaphor. If the memory stores too many different images the memory makes too many different experiences for the category. Hence the memory is unable to build a clear mental image of a prototype. From a more technically point of view we know for any median image and for any of its pixel its statistical distribution. This allows us to calculate the density of the distribution around the median value. We define the density in this case as the distance between the maximal gray level and the minimum gray level which is equal to the right and left border of the distribution. If we connect the maximum borders and the minimum borders respectively of each pixel in each row we define a parameter tube. The tube is shown in fig. 10

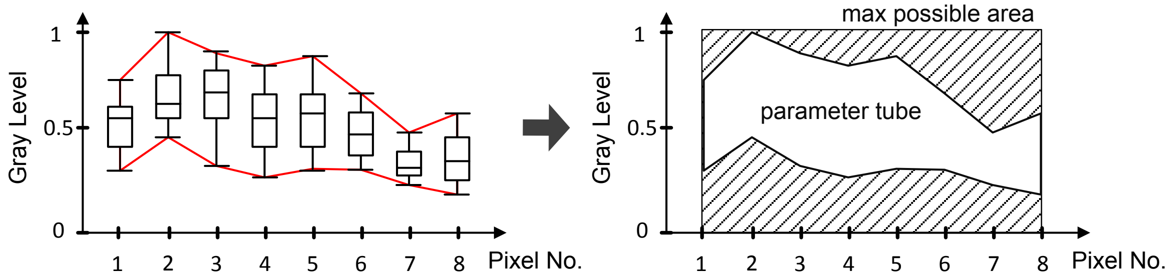

Figure 10 Definition of the parameter tube and derivation of the an uncertainty index

In the next step we compute the area covered by the tube. The ratio of this area and the maximum possible covered area is defined as the uncertainty index $u_i$of the current considered parameter for a given situation number. We add the uncertainty index to our formula and get an updated formula to calculate the distraction index.

$$DI = \sum_{i=0}^{m} u_i \int_{t_0}^{t_n} \left(w_i(t) * \left(p_{c_i} - 0.5\right)^2\right) dt \quad with \;\; w_i(t_j) = w_{ij};\; p_{c_i}(t_j) = \; p_{c_{ij}} \;\; for\, all\, j = 1 \ldots n$$

4 Conclusion & Outlook

In this paper we presented a driver model for the optimization of a CMS. Our proposed driver model adapted to the driver state and activated the acoustic warning earlier in case of visual distraction. We discussed the gap of methods which are able to:

1. learn from multivariate time series
2. recognize patterns
3. understand/analyze easily

For that reason we introduced a design metaphor. Motivated by the idea of mental representations and human memorizing we developed a new method to classify driver distraction online. The basic idea was to transform time series in an image and to connect it to a situation number. This allowed us to build categories and to develop prototypes. We compared the prototypes (median images) with the current example (current image) to detect patterns. More precisely we detected an atypical driving behavior. Through the design metaphor we were able to use algorithms from image processing for the comparison. We finished the description of the distraction classifier by introducing an uncertainty index. The index is a simple but powerful measure to estimate the quality of parameters online.

The evaluation of our approach is currently on going. Therefore we conducted a simulator study in the dynamic driving simulator at the DLR with 15 subjects (12 male, 3 female, average age: 28.9, age range: 19-44). The scenario was a three lane highway with intervals of visual distraction monitored by an eye tracking system. We will present results in the corresponding poster of the conference. After this evaluation we plan to investigate the behavior of the driver model on real traffic data and cognitive distraction as a reasonable next step. Furthermore the ability of the model to observe the normal driving behavior will used in the context of partly automated vehicles to improve system behavior and acceptance.

Acknowledgement

“This work was also supported by the European Commission under interactIVe, a large scale integrated project part of the FP7-ICT for Safety and Energy Efficiency in Mobility. The authors would like to thank all partners within interactIVe for their cooperation and valuable contribution”.

References

[1] https://www.destatis.de/DE/PresseService/Presse/Pressemitteilungen/2013/02/PD13_068_46241.html (accessed 26.08.2013)

[2] http://www.bosch-presse.de/presseforum/download/de/7966ks-d_Anlage_Befragung_Fahrerassistenz.pdf (accessed 26.08.2013)

[3] http://www.interactive-ip.eu/project/subprojects/EMIC (accessed 26.08.2013)

[4] Bauer, C.: A driver specific maneuver prediction model based on fuzzy logic, FU Berlin Diss. 2012

[5] M.A.S. Kamal, T. Kawabe, J. Murata and M. Mukai *Driver-Adaptive Assist System for Avoiding Abnormality in Driving*, 16th IEEE International Conference on Control Applications Part of IEEE Multi-conference on Systems and Control, 2007, Singapore, China

[6] Tango, F., Minin L., Montanari, R., & Botta, M.:*Non-intrusive Detection of Driver Distraction using Machine Learning Algorithms*. XIX European Conference on Artificial Intelligence (ECAI). 2010, Lisbon, Portugal

[7] Tango F., Minin L., Tesauri F., Montanari R.: Field tests and machine learning approaches for refining algorithms and correlations of driver's model parameters. Applied Ergonomics 41 (2010) pp.211–224

[8 Schreiner F. (1999). Automatische Führung des Kraftfahrzeugs mit fahreradaptiven Eigenschaften, Dissertation, Universität der Bundeswehr München, Forschungsberichte VDI Reihe 12, Nr. 404

[9] Feraric J. P. (1996). Echtzeitfähige Modellierung des individuellen Fahrerverhaltens zur Realisierung adaptiver Unterstützungsfunktionen in einem Monitor- und Warnsystem, Dissertation, Universität der Bundeswehr

[10] Martirosjan A., Griesche S.: Driver intention modelling for partly automated vehicles- The benefit and necessity of a driver and situation adaptive approach. Reports of the DLR-Institute of Transportation Vol. 19 (2012) pp. 49-55

[11] Poitschke T., Laquai F., Rigoll G., Jacko J.A (Ed.): Agent-Based Driver Abnormality Estimation. Human-Computer Interaction, Part III (2009) pp. 189–198.

[12] Piechulla W., Mayser C., Gehrke H., König W.: Reducing drivers mental workload by means of an adaptive man–machine interface. Transportation Research Part F 6 (2003) 233–248

[13] Löper, C, Kelsch, J. & Flemisch F.: Kooperative manöverbasierte Automation und Arbitrierung als Bausteine für hochautomatisiertes Fahren. Automatisierungs-,

Assistenzsysteme und eingebettete Systeme für Transportmittel, GZVB, Braunschweig (2008) pp. 215-237

[14] Ersal T., Fuller H. J. A., Tsimhoni O., Member, IEEE, Stein J. L., Fathy H. K.: Model-Based Analysis and Classification of Driver Distraction Under Secondary Tasks. IEEE Transactions on Intelligent Transportation Systems, Vol. 11, No. 3 (2010) pp. 692-701

[15] Lethaus F., Baumann M., Köster F., Lemmer K.: Predicting driver intent by means of pattern recognition based on gaze behaviour. Human Centred Automation (pp. 35-47). Maastricht, Netherlands: Shaker Publishing 2012

[16] Onken R., Feraric J.P.: ADAPTATION TO THE DRIVER AS PART OF ADRIVER MONITORING AND WARNING SYSTEM. Accid. Anal. and Prev. Vol. 29, No. 4.(1997) pp. 507-513

[17] Paola J. D. & Schowengerdt R. A.: A review and analysis of backpropagation neural networks for classification of remotely-sensed multi-spectral imagery. Int. J. Remote Sensing Vol. 16, 1995, pp 3033-3058

[18] Baddeley A.: Working Memory: Theories, Models, and Controversies. Annu. Rev. Psychol. 63 (2012) pp. 1-29

[19] Atkinson R. C., Shiffrin R. M.: Human memory: a proposed system and its control processes. ThePsychology of Learning and Motivation: Advances in Research and Theory (pp. 89–195) New York: Academic 1968

[20] Anderson M. C., Bjork R.A., Bjork E.K.: Remembering Can Cause Forgetting: Retrieval Dynamics in Long-Term Memory. Journal of Experimental Psychology: Learning, Memory, and Cognition Vol.20 No.5 (1994) pp.1063-1087

[21] Craik, F.I.M., Lockhart. R.S.: L.evels of processing: A framework for memory research. Journal of Verbal Learning and Verbal Behavior. II. (1972) pp 671-684

[22] Hintzman, D. L. (1988). "Schema Abstraction" in a Multiple-Trace Memory Model. Psychological Review, 93, (1986) pp. 411-428

[23] Peng R.D.: A Method for Visualizing Multivariate Time Series Data. Journal of Statistical Software Volume 25 (2008)

[24] Stolzmann W., Proette A., Sokoll A.: Reaktionszeitdifferenzen bei aufmerksamen und abgelenkten Fahrern. 4. Tagung Sicherheit durch Fahrerassistenz, 2010, Munich, Germany

[25] Dozza M.: What factors influence drivers' response time for evasive maneuvers in real traffic? Accident Analysis & Prevention (2012) in press