150 Jahre
culture of
excellence

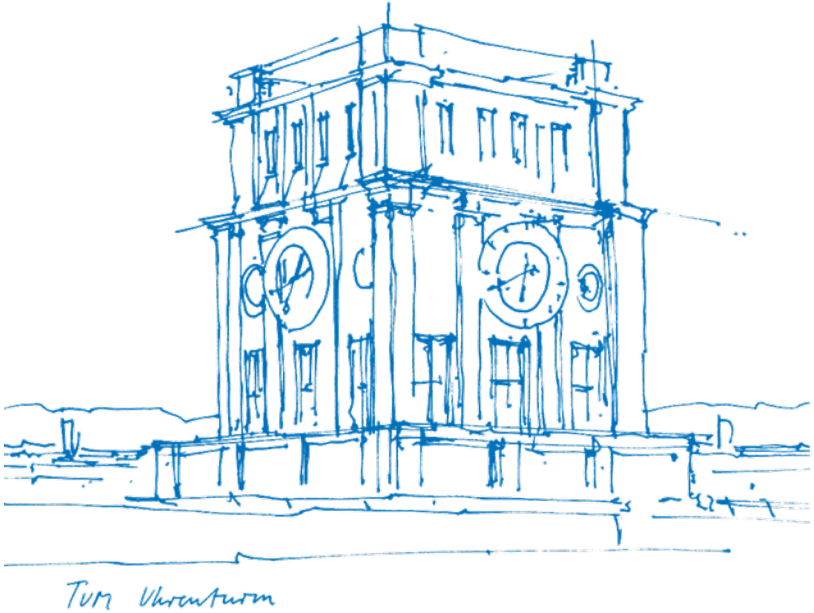
Integration of Driver Behavior into Emotion Recognition Systems: A Preliminary Study on Steering Wheel & Vehicle Acceleration

Sina Shafaei, Tahir Hacizade and Alois Knoll

Technical University of Munich

Department of Informatics

sina.shafaei@tum.de



† First International Workshop on Advanced Machine Vision for Real-life and Industrially Relevant Applications

Perth, Dec. 3nd 2018

Outline

Introduction

Objectives

Experiments

Results

Conclusion & Future Works

Introduction

Status Quo

- Current status of the emotion recognition systems in cars is mostly focused on facial-based approaches
- Modeling behavior of the driver in cabin has great impact on developing intelligent and autonomous driving
- According to *7-38-55* rule, 93% of human communication is performed through nonverbal means, which consists of facial expressions, body language and voice tone.
- Recent studies have shown a high level of correlation between driving behavior and emotional status

Introduction

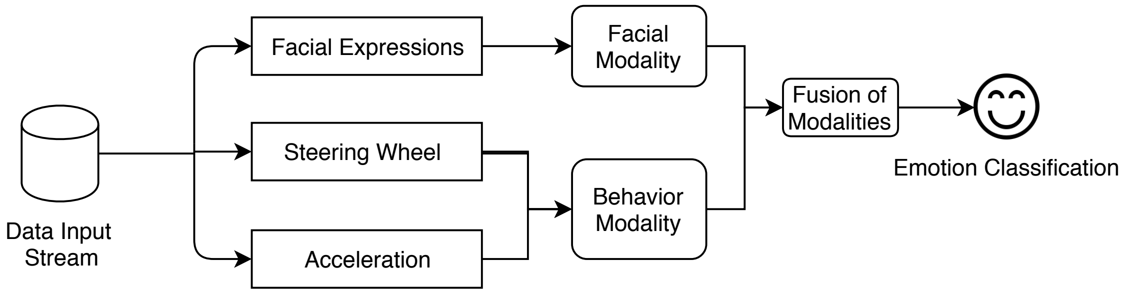
Main objectives of this work:

- How emotions affect the behavior of the driver?
- How to map the driving behavior, to the current emotional status?
- What are the benefits of multimodality in emotion recognition systems?

Objectives

Proposed system in this work:

- Facial approach is based on HOG (Histogram of Oriented Gradients) descriptors and SVM
- Acceleration/Deceleration and Steering Wheel usage are considered as the driving behavior-related modalities
- Decission-level fusion is used for combining the modalities (arousal-valence measure)



Objectives - Facial Approach

- Facial landmark detector was used to detect ROI,
- After detecting the face, HOG descriptors were used by applying a fixed size sliding window over an image pyramid build upon them,
- Model was build on a liner kernel SVM with decision function of One-Vs.-Rest,
- For training the model k-fold cross validation was used with k set to 10,
- CK+ and JAFFE databases were used for training,

```
featureVector ← init list
SVMClassifier ← load model
while newFrame is exist do
    frame ← FetchVideoStream()
    grayFrame ← GrayscaleImage(frame)
    if faceTracker(grayFrame).Score < threshold then
        face ← detectFace(grayFrame)
    else
        face ← faceTracker(grayFrame).Position
    end if
    ROIarray ← FetchROI(face)
    for each ROI in ROIarray do
        hog ← HOGDescriptor(ROI)
        featureVector ← featureVector + hog
    end for
    result ← SVMClassifier(featureVector)
end while
```

Experiments

Testbed:

► A real “SMART” car simulator, 15 participants, each driving 3 scenarios , 36 minutes on average for all of the scenarios,

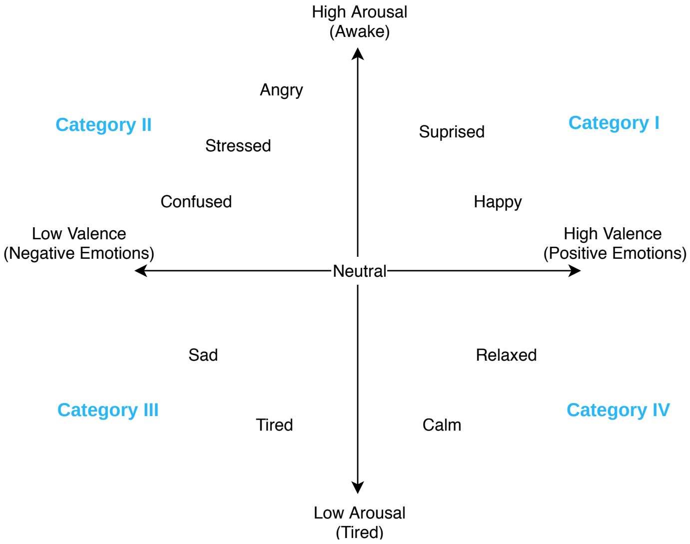
► Their attitude was different toward scenarios due to the prehistory condition narrated to them and predefined situations on the road,

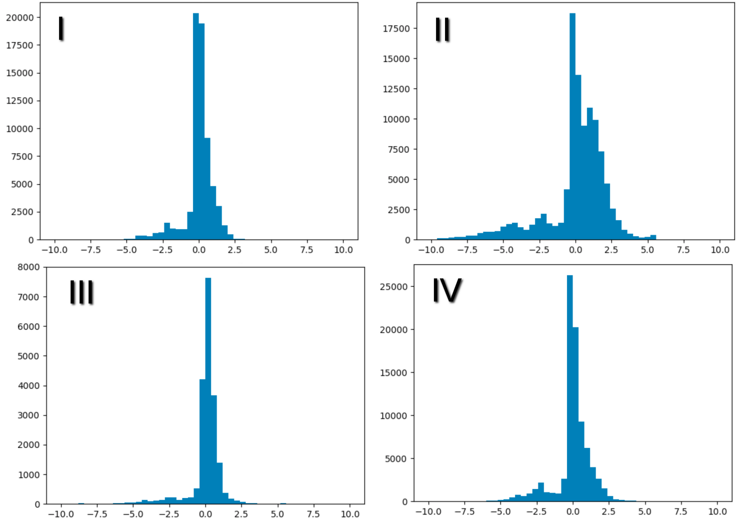
► Participants were asked right after each ride about their actual emotional status using a questionnaire,

► Facial expressions are recorded through a camera along with the respective signals of steering wheel and acceleration through the virtual test drive software of the simulator

Results

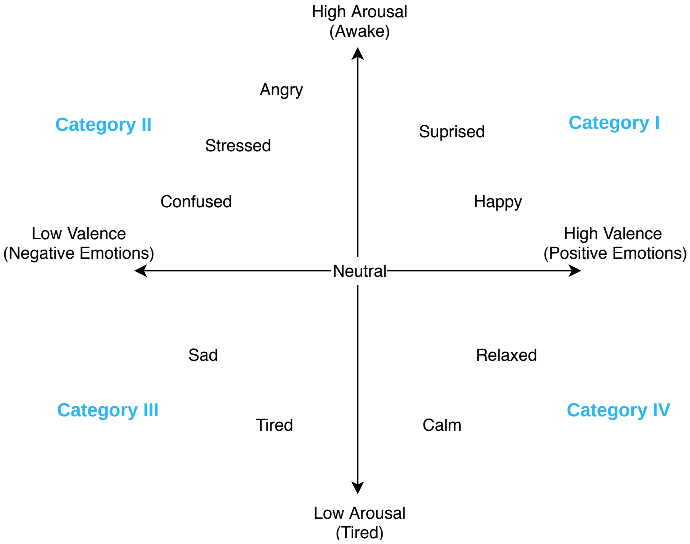
- Frequency distribution of vehicle acceleration in 4 groups of emotional status

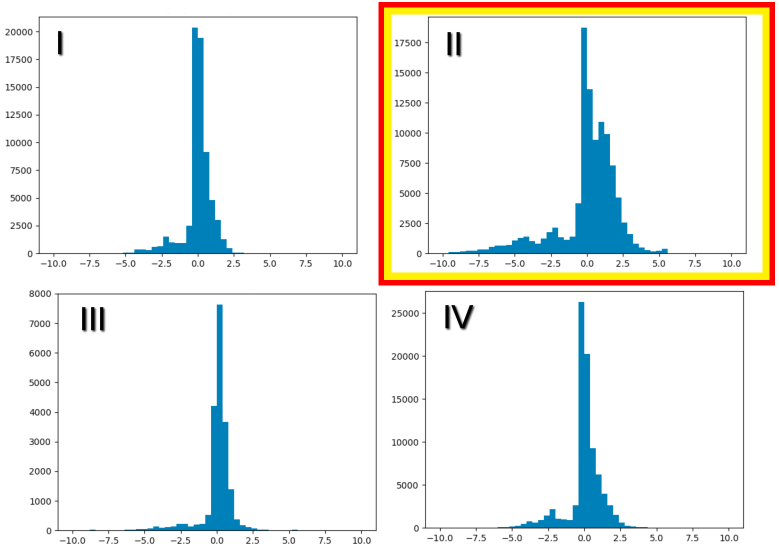




Results

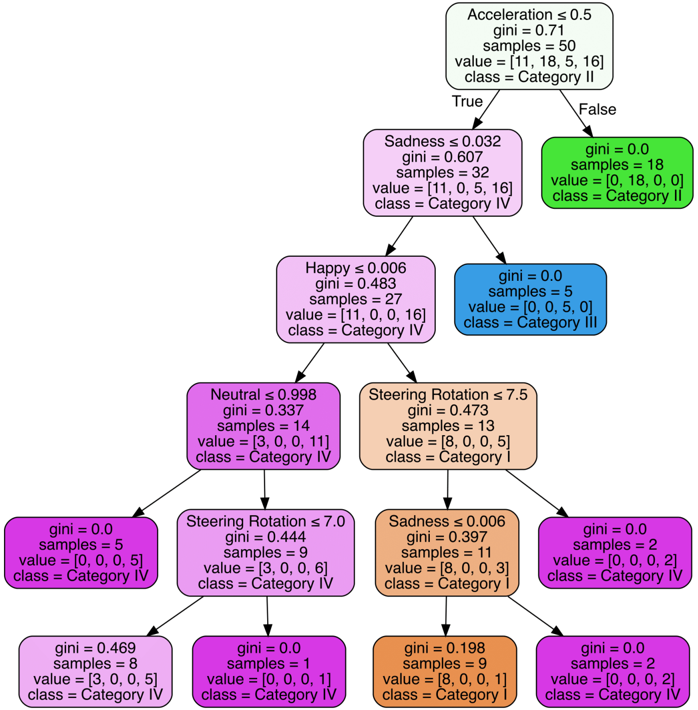
- Angry/stressed/confused drivers tend to accelerate/decelerate faster





Results

- ▶ 50 samples of collected data are considered,
- ▶ Decision tree of combining the 3 different module using 50 samples,
- ▶ A considerable impact of the vehicle acceleration from category II,
- ▶ All 18 samples of category II are grouped using only one condition from vehicle acceleration,
- ▶ Only samples from category I and IV are left un-grouped,



Results

Conditions formulated by steering wheel (SW) rotation and happy, neutral and sadness features from the facial expressions module, are considered together to form the feature vector,

- After analyzing a single decision tree, we use the same feature vectors from 50 samples to train a random forest classifier

Module	Vector index	Parameter Name	Value
VA	1	Acceleration	0 or 1
SW	2	Steering Rotation	0 to ∞
Facial Expression	3	Neutral	0 to 1
	4	Anger	0 to 1
	5	Disgust	0 to 1
	6	Fear	0 to 1
	7	Happy	0 to 1
	8	Sadness	0 to 1
	9	Surprise	0 to 1

Results

► The 77.27% of accuracy is obtained using multimodal emotion recognition system on data samples with 2 minutes of length,

► This condition is prone to errors and false predictions since in real-life situations the 2-minutes range could be easily falsified by situations like staying behind a red light,

► To mitigate this issue, we consider the decision taking step at the end of each ride by summarizing the emotion predictions performed for only sub-samples and choosing the most frequently felt emotion

Authors	Facial Method	Accuracy
J.F.Cohn and T.Kanade et al.	Active Appearance Models	83%
H. Alshamsi et al.	BRIEF Feature Extractor	89%
W.Swinkels et al.	Ensemble of Regression Trees	89.7%
Sébastien Ouellet	Convolutional Network	94.4%
R.A.Khan et al.	HOG-based	95%
M. F. Donia et al.	HOG-based	95%
Our Method	**HOG on ROI regions**	**93%**

Method	Accuracy	Precision	F1 Score	Recall
Facial-based Module	54.54%	54.75%	50.45%	49.86%
SW-based Module	37.5%	10.3%	13.6%	25%
VA-based Module	68.18%	35.51%	37.76%	41.37%
Fusion of All Three Modules	**77.27%**	**73.39%**	**73.59%**	**75.89%**

Results

Comparison of different unimodal and multimodal emotion recognition systems based on accuracy and different number of emotional classes

System	Type	Method	Classes	Accuracy
[1]	Unimodal	Electrodermal Activity (EDA)	3	70%
[2]	Unimodal	Facial Emotion Recognition.	6	70.2%
[3]	Unimodal	Speech Emotion Recognition	3	88.1%
[4]	Unimodal	Speech Emotion Recognition.	2	80%
[5]	Multimodal	EDA and Skin Temperature	4	92.42%
[6]	Multimodal	Speech & Facial Emotion Recognition	7	57%
[7]	Multimodal	Acoustic & Facial Emotion Recognition	3	90.7%
Our System	Multimodal	Facial and Vehicle Parameters	4	**94.4%**

Conclusion & Future Work

- Most of the studies on emotion recognition are based on unimodal approaches where only audio or visual is examined,
- Adding behavior related modalities increases the accuracy of the predictions and robustness of the systems,
- We have only investigated the impacts of integrating two modalities of acceleration and steering wheel usage but too many are left for the future,
- The proposed system was capable of classifying the emotions into 4 main categories with the final accuracy of 94.4%,
- We are going to extend our system by integration of other behavior-related modalities and will study the shared models among the drivers according to their emotional states

Thank You.

References

The State-Of-The-Art Unimodal/Multimodal Emotion Recognition Methods

- J. S. K. Ooi, S. A. Ahmad, Y. Z. Chong, S. H. M. Ali, G. Ai, and H. Wagatsuma, “Driver emotion recognition framework based on electrodermal activity measurements during simulated driving conditions,” in *Biomedical Engineering and Sciences (IECBES), 2016 IEEE EMBS Conference on*. IEEE, 2016, pp. 365–369.
- R. Theagarajan, B. Bhanu, A. Cruz, B. Le, and A. Tambo, “Novel representation for driver emotion recognition in motor vehicle videos,” in *Image Processing (ICIP), 2017 IEEE International Conference on*. IEEE, 2017, pp. 810–814.
- A. Tawari and M. Trivedi, “Speech based emotion classification framework for driver assistance system,” in *Intelligent Vehicles Symposium (IV), 2010 IEEE*. IEEE, 2010, pp. 174–178.
- C. M. Jones and I.-M. Jonsson, “Automatic recognition of affective cues in the speech of car drivers to allow appropriate responses,” in *Proceedings of the 17th Australia conference on Computer-Human Interaction: Citizens Online: Considerations for Today and the Future*. Computer-Human Interaction Special Interest Group (CHISIG) of Australia, 2005, pp. 1–10.
- M. Ali, F. Al Machot, A. H. Mosa, and K. Kyamakya, “Cnn based subject-independent driver emotion recognition system involving physiological signals for adas,” in *Advanced Microsystems for Automotive Applications 2016*. Springer, 2016, pp. 125–138.
- D. Datcu and L. Rothkrantz, “Multimodal recognition of emotions in car environments,” *DCI&I 2009*, 2009.
- S. Hoch, F. Althoff, G. McGlaun, and G. Rigoll, “Bimodal fusion of emotional data in an automotive environment,” in *Acoustics, Speech, and Signal Processing, 2005. Proceedings.(ICASSP’05). IEEE International Conference on*, vol. 2. IEEE, 2005, pp. ii–1085.