TECHNISCHE UNIVERSITÄT
CHEMNITZ

Real Time Recognition of Non-Driving Related Tasks in the Context of Highly Automated Driving

S. Enhuber[1], T. Pech[1], B. Wandtner[2], G. Schmidt[2], G. Wanielik[1]

[1] Professorship of Communications Engineering
Technische Universität Chemnitz, Germany

[2] Opel Automobile GmbH

stephan-michael.enhuber@etit.tu-chemnitz.de

Gefördert durch:

Bundesministerium
für Wirtschaft
und Energie

aufgrund eines Beschlusses
des Deutschen Bundestages

Introduction

Definition of vehicle automation levels

SAE International J3016

SAE level	Name	Narrative Definition	Execution of Steering and Acceleration/ Deceleration	*Monitoring* of Driving Environment	Fallback Performance of *Dynamic Driving Task*	System Capability *(Driving Modes)*
***Human driver* monitors the driving environment**						
0	**No Automation**	the full-time performance by the *human driver* of all aspects of the *dynamic driving task*, even when enhanced by warning or intervention systems	Human driver	Human driver	Human driver	n/a
1	**Driver Assistance**	the *driving mode*-specific execution by a driver assistance system of either steering or acceleration/deceleration using information about the driving environment and with the expectation that the *human driver* perform all remaining aspects of the *dynamic driving task*	Human driver and system	Human driver	Human driver	Some driving modes
2	**Partial Automation**	the *driving mode*-specific execution by one or more driver assistance systems of both steering and acceleration/ deceleration using information about the driving environment and with the expectation that the *human driver* perform all remaining aspects of the *dynamic driving task*	**System**	Human driver	Human driver	Some driving modes
***Automated driving system* ("system") monitors the driving environment**						
3	**Conditional Automation**	the *driving mode*-specific performance by an *automated driving system* of all aspects of the dynamic driving task with the expectation that the *human driver* will respond appropriately to a *request to intervene*	System	**System**	Human driver	Some driving modes
4	**High Automation**	the *driving mode*-specific performance by an automated driving system of all aspects of the *dynamic driving task*, even if a *human driver* does not respond appropriately to a *request to intervene*	System	System	**System**	Some driving modes
5	**Full Automation**	the full-time performance by an *automated driving system* of all aspects of the *dynamic driving task* under all roadway and environmental conditions that can be managed by a *human driver*	System	System	System	**All driving modes**

Source:
http://www.sae.org/misc/pdfs/automated_driving.pdf

Driver Availability in Highly Automated Driving (SAE Level 3)

- The driver state influences the transition process and driver's performance when there is a request to intervene.
- The transition process is highly affected by engagement in non-driving related tasks.
- Non-driving related tasks can affect different aspects of the driver state:
  - **Sensory state**
  - **Motoric state**
  - Cognitive state
  - Arousal level
  - Motivational conditions

Marberger et. al 2017: „Understanding and Applying the Concept of "Driver Availability" in Automated Driving". In: Advances in Human Aspects of Transportation: Proceedings of the AHFE 2017 Conference on Human Factors in Transportation, July 17-21, 2017, Los Angeles, California, USA. Band 597.

Analysis of Non-Driving Related Tasks

- vehicle cockpit is a relative defined environment
- non-driving related tasks consist of sub-tasks
- sub-tasks forming sequences

Analysis of Non-Driving Related Tasks

- vehicle cockpit is a relative defined environment
- non-driving related tasks consist of sub-tasks
- sub-tasks forming sequences
- for example.: non-driving related task „phoning“

sub-task 1		sub-task 2		sub-task 3		sub-task 4		
glance to mobile phone	→	grab phone	→	accept call	→	put phone to ear	→	…

Analysis of Non-Driving Related Tasks

- vehicle cockpit is a relative defined environment
- non-driving related tasks consist of sub-tasks
- sub-tasks forming sequences
- for example.: non-driving related task „phoning“

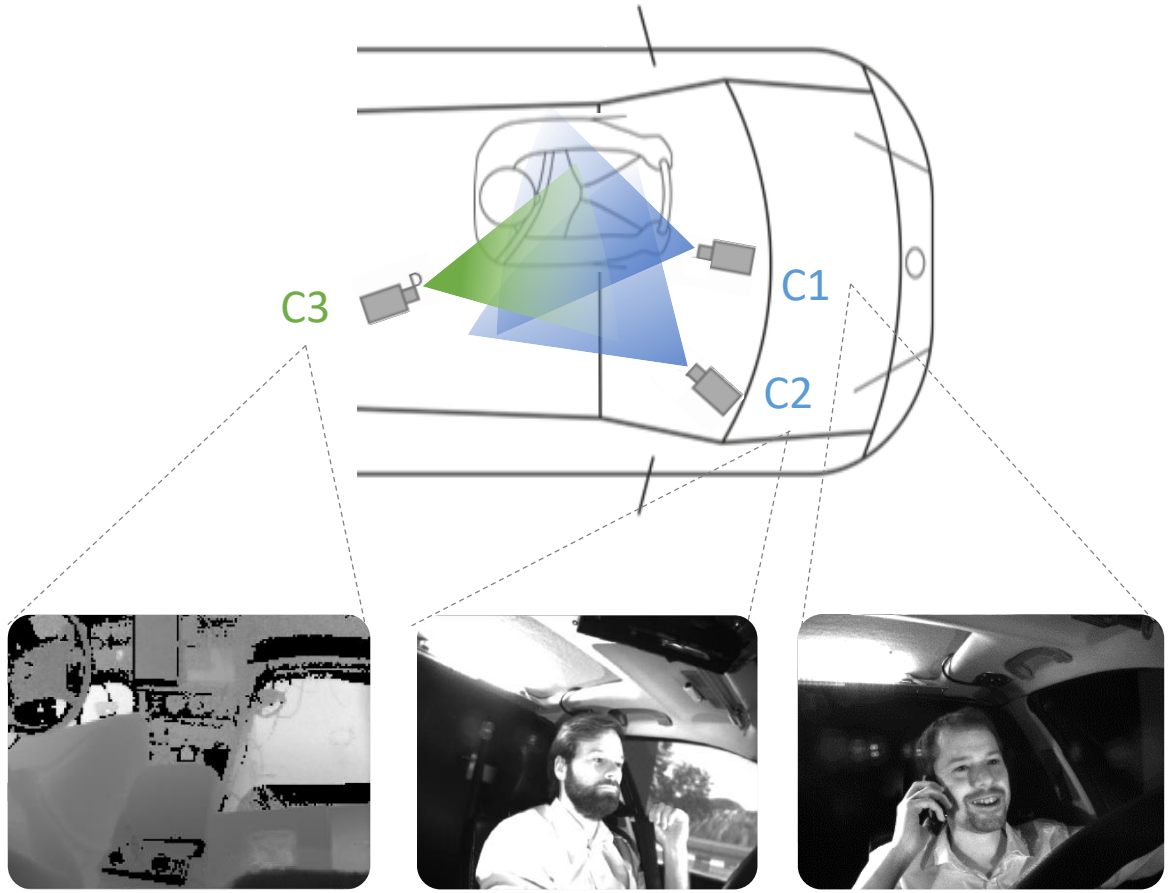
sub-task 1		sub-task 2		sub-task 3		sub-task 4	
glance to mobile phone	→	grab phone	→	accept call	→	put phone to ear	→ …

- sub-task sequence is defined by different glance and hand movements
- features and feature sequences are intra- and interindividual variable

Data Acquisition

- driving simulator environment
- sensor setup:
  - 2x 2D-cameras **C1, C2**
  - 3D-camera **C3**

- measurement data of 43 test persons
- set of several non-driving related tasks

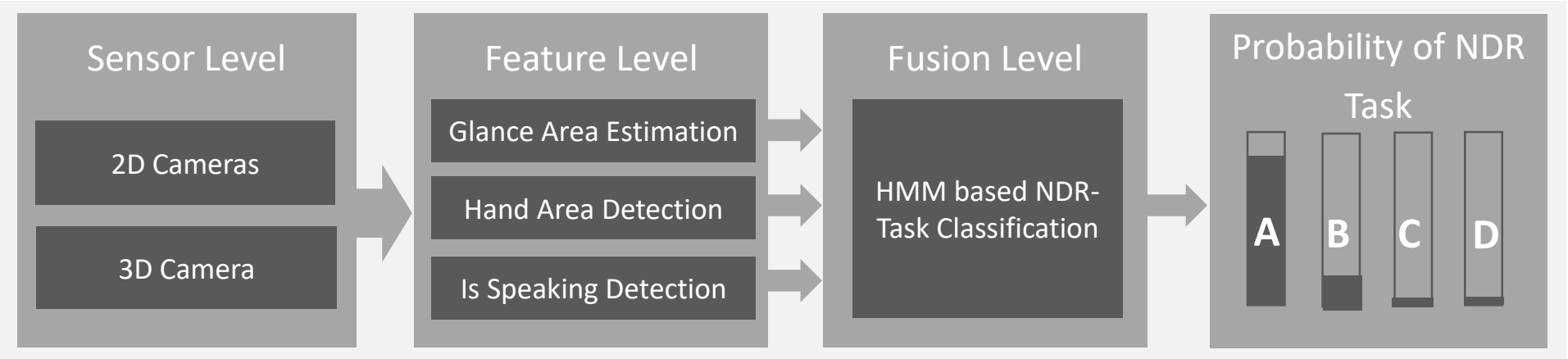


Used Set of Non-Driving Related Tasks

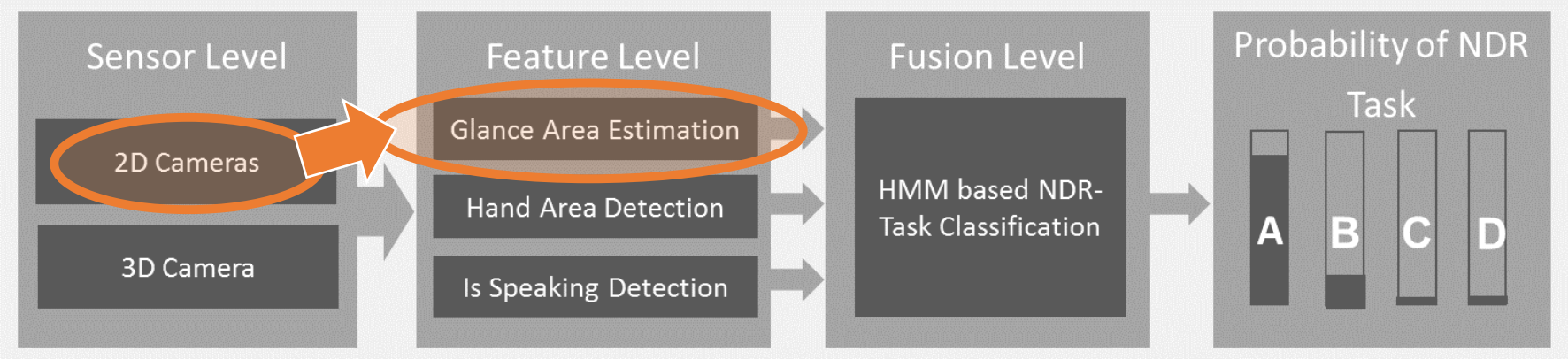
Number	Non-driving related task	Description
1	Repeating spoken text	Auditory presented sentences, repeating verbally
2	Reading out text	Written sentences presented on tablet computer (attached in center console), reading out loud
3	Texting on tablet (handheld)	Transcribing text on tablet computer, attached in center console
4	Texting on tablet (mounted)	Transcribing text on tablet computer, performed handheld
5	Reaching for object: passenger seat	Searching for specific Lego bricks and placing these in a box on the passenger seat
8	Cell-phone talk (handheld)	Receiving a call from the experimenter
9	Baseline	Observing Traffic

Description of Algorithm

System Overview

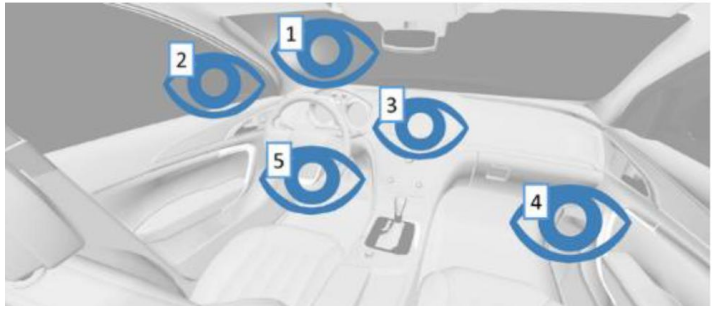


Description of Algorithm

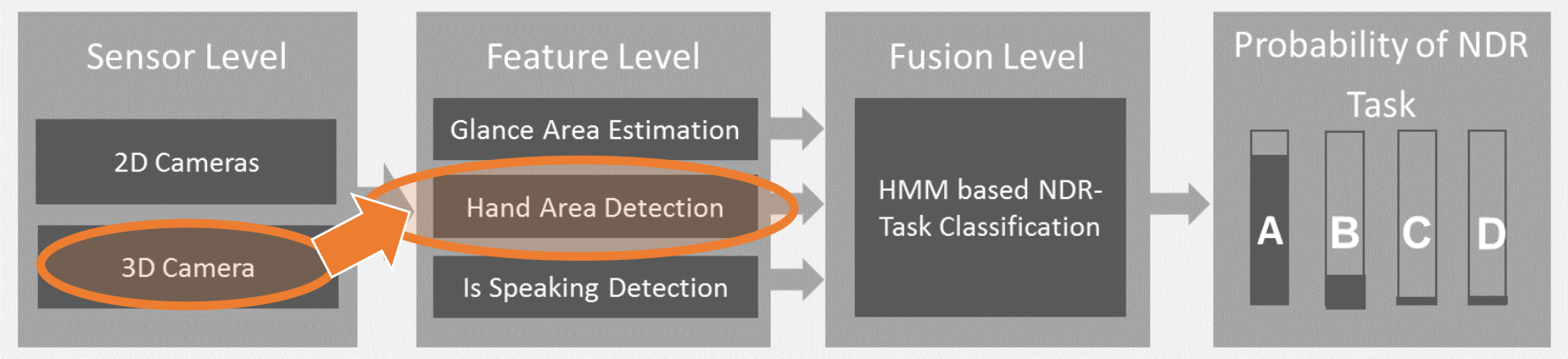


- Glance area estimated by head tracking measurements using naïve Bayes Classifier.
- Assignment of measured head orientations to a specific in-vehicle area.
- 5 classes C_i with $i = 1, \dots, 5$ – respectively in-vehicle areas
- determination of $p(\boldsymbol{x}|C_i)$ and $P(C_i)$ with Maximum-Likelihood-Estimation

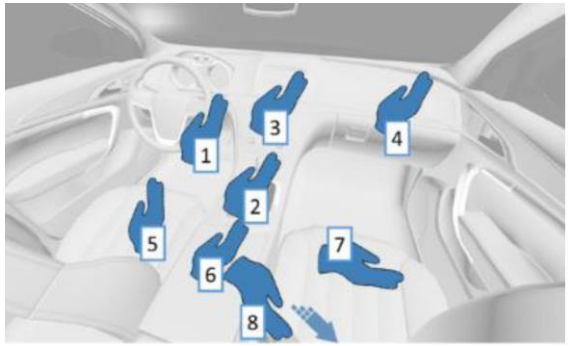
$$P(C_i|\boldsymbol{x}) = \frac{p(\boldsymbol{x}|C_i) \cdot P(C_i)}{\sum_{j=1}^{N} p(\boldsymbol{x}|C_j) \cdot P(C_j)}$$

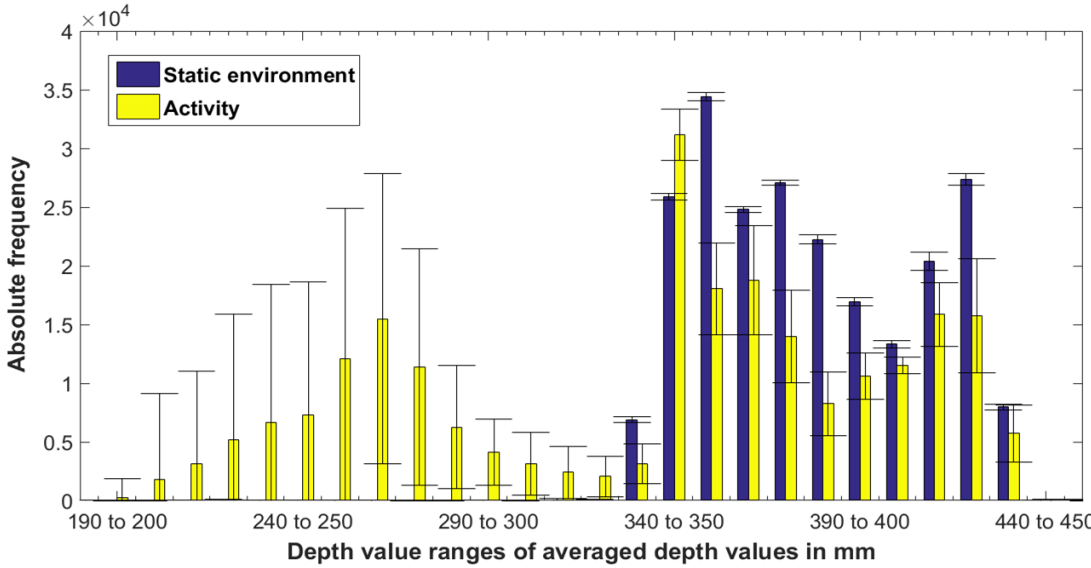


Description of Algorithm

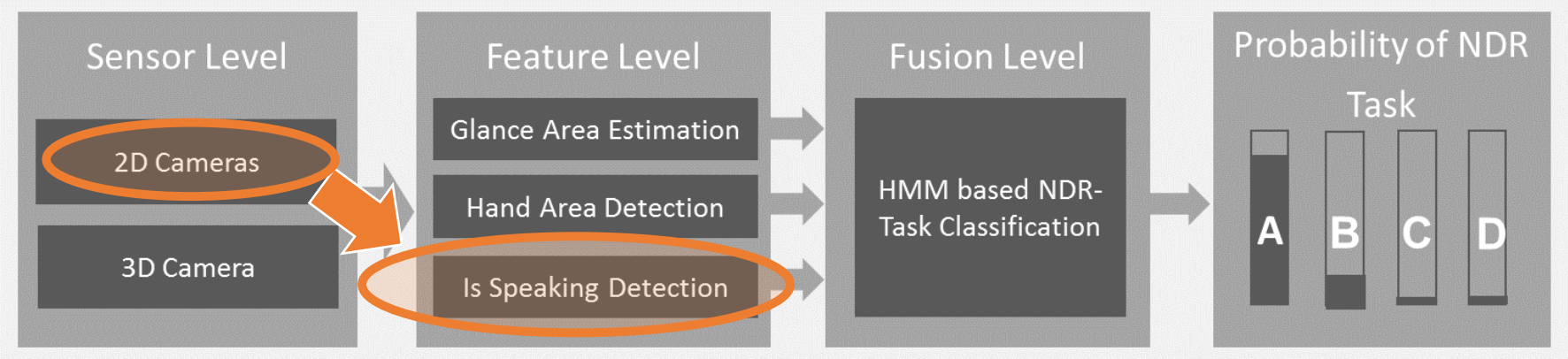


- detecting characteristic hand areas of non-driving related task
- using 3D depth camera
- grid based depth value segmentation of occupied areas



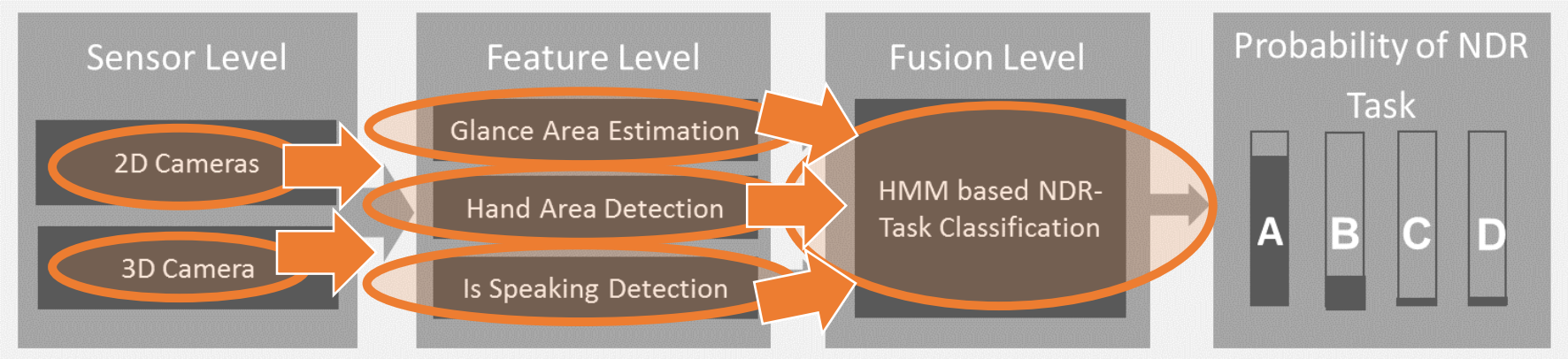


Description of Algorithm



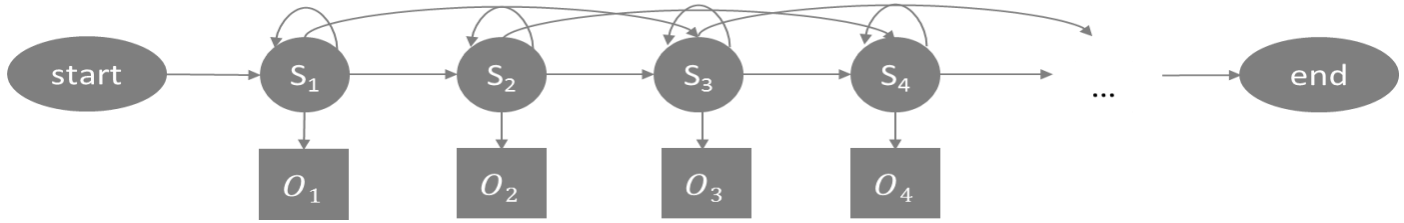
- detect if the driver is speaking using facial mouth landmarks from head tracking
- distance between upper lip and lower lip landmarks
- using variance analysis of 2 seconds sequence to determine if the person is speaking or not

Description of Algorithm

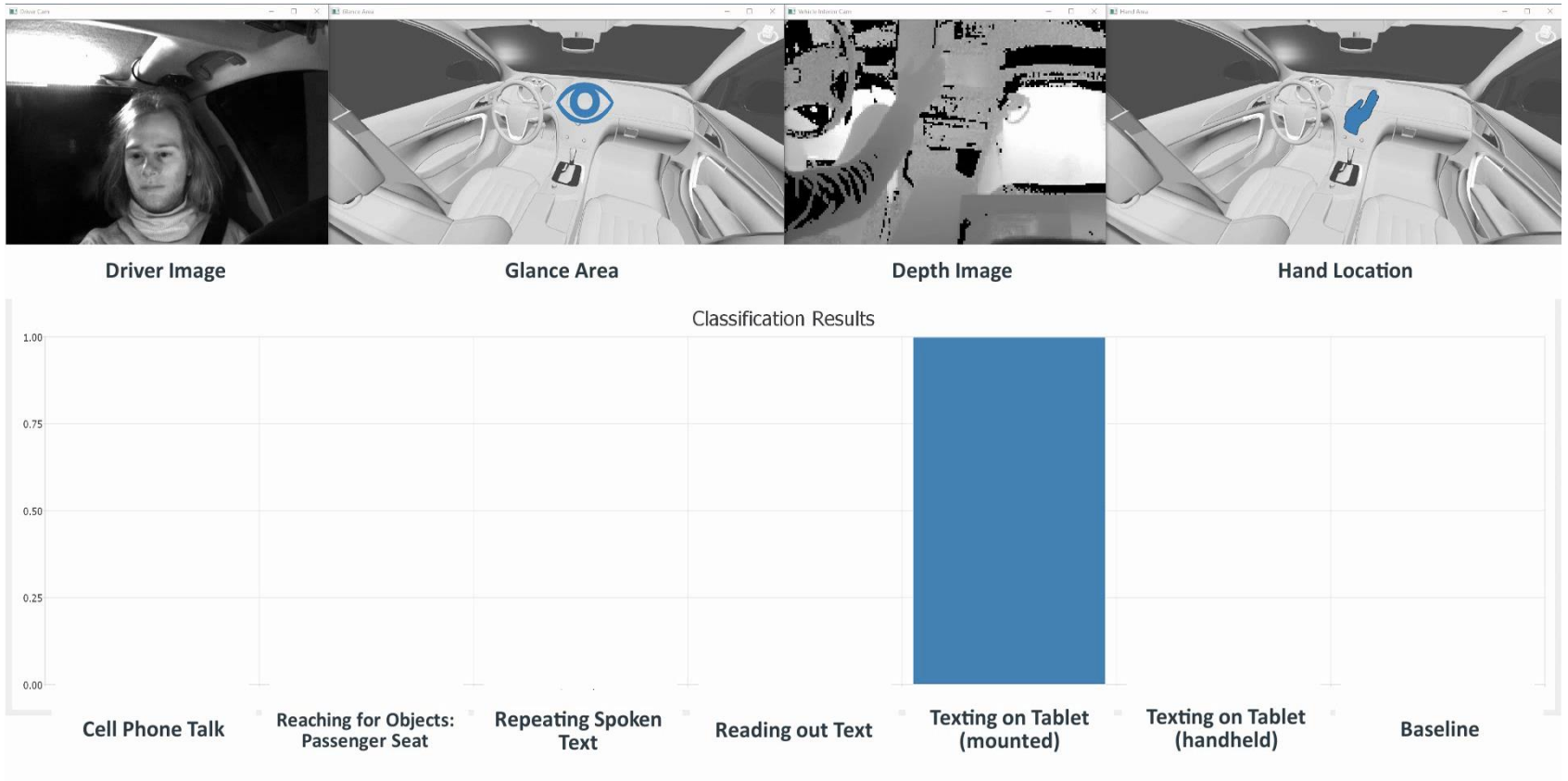


- non-driving related tasks classification using trained Hidden Markov Models (HMMs)
- take temporal dependencies of features into account
- input sequence created by sliding window
- HMM with max production probability fits best to observed input sequence

exemplary left-right structure of HMM



Results – Demonstration Video

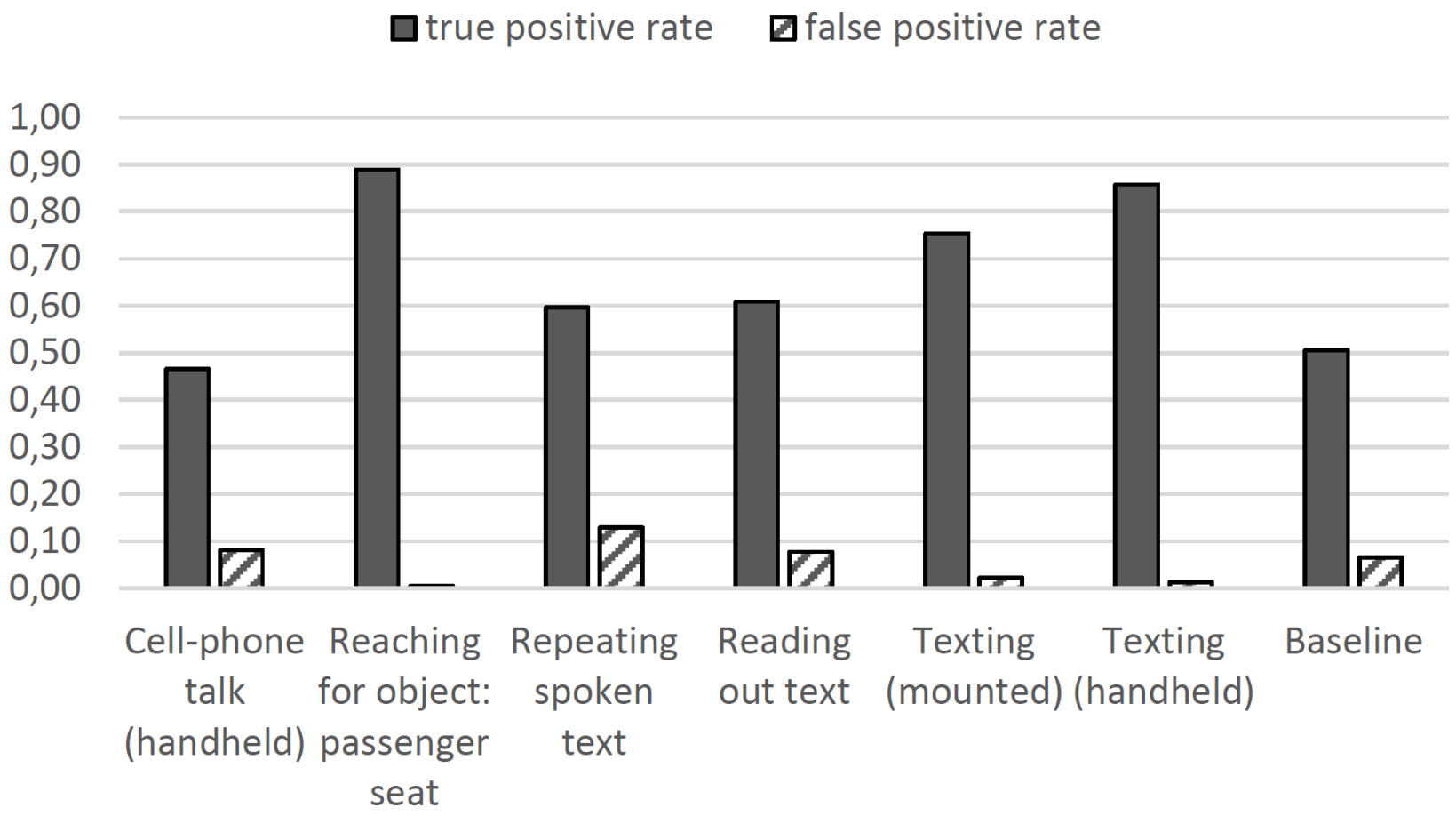


Results – Statistics

- leave-one-out cross-validation of 39 test persons

Issues:

- hand on head detection from 3D camera FOV
- speaking detection from facial landmarks



Results Update – Statistics

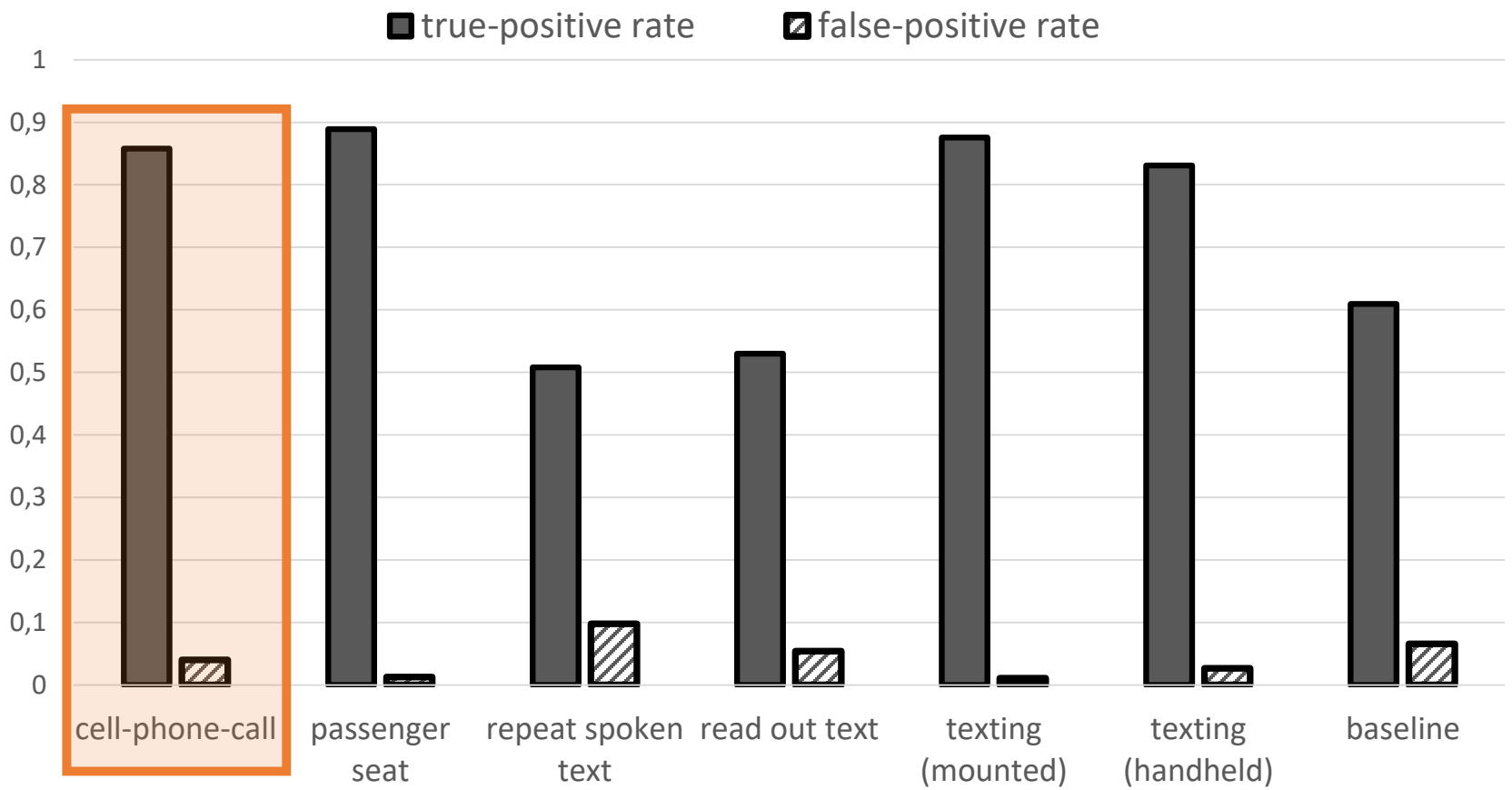
- leave-one-out cross-validation of 39 test persons

- Update:

add Hand area detection using Support Vector Machine from 2D input image

→ improved hand on head detection

→ better results for cell phone talk



Outlook

- transfer and test algorithm in real vehicle – demonstration video

Conclusion

- HMMs are feasible to detect NDR-task
- hand and glance positions are stable features to distinguish between the most critical NDR-tasks
- the necessary measurement feature to recognise NDR-tasks are related to the human information processing channels

Thank you for your attention!

stephan-michael.enhuber@etit.tu-chemnitz.de

Professorship of Communication Engineering
Technische Universität Chemnitz, Germany

This work results from the joint project Ko-HAF - Cooperative Highly Automated Driving and has been funded by the Federal Ministry for Economic Affairs and Energy based on a resolution of the German Bundestag.