Infoclass: Public.

Driver Drowsiness and Distraction Detection by Sensor Fusion (D4SF).

FFI Project Report

Project Partners:

Title of the report: Driver Drowsiness and Distraction Detection by Sensor Fusion (D4SF).

Authors: Arne Nåbo, Saab Automobile AB, Johan Karlsson and Benny Nilsson, Autoliv Development AB and Anna Anund, VTI.

Project reference number: 2009-00047.

Publication date: 2010-03-15. Issue 1.

Contact person: Arne Nåbo.

Project responsible: Autoliv Development AB.

Table of contents

SUMMARY

Loss of control is a fact if the driver diverts from the driving task for just a few seconds. A continuous attention to driving is a must in order to drive in a safe way. But it is also known that humans are easily being distracted or drowsy. These traffic safety problems have received lot of attention lately. Several recent studies indicates that about 10-20 % of all accidents are likely sleepiness-related, and distraction has been a contributing factor of up to 78-80% of all accidents and incidents.

The project aim was to contribute to traffic safety by promoting an alert and attentive driver by technical means in the vehicle.

A vehicle was equipped with a system for detection of visual inattention and sleepiness. The main sensors in this system were two cameras; one looking at the driver's head- and eye behavior for physiological signs of distraction and sleepiness and one looking forward at the road to see the effect of these on driving behavior. The objective was to get a more robust detection of visual inattention and sleepiness than if only one sensor was used. This was the interpretation of 'sensor fusion'. The vehicle also had a warning system to bring the driver back to attentive driving and to prompt a sleepy driver to take relevant countermeasures.

Tests of the system took place on real roads. The reason for this was to have as much similarity as possible to real driving conditions.

A total of 43 drivers participated in field studies of sleepy drivers. Each driver drove during the day, evening and night, each time for approximately 90 minutes. Subjective sleepiness was assessed every 5 minutes using the Karolinska Sleepiness Scale (KSS). Physiological sleepiness was continuously measured using EEG, EMG and EOG recordings.

Best drowsiness detection performance was by using a combination of long eye blinks, head pitch and time-of-day. Adding either the steering- or lane tracker data did not give any clear performance benefit.

The distraction detection performance was very dependent on a good driver monitoring system. Although data treatment did help overcome some of its shortcomings, an improved capability to track eye movements would mean a lot to the overall performance. However, when the driver monitoring system worked well (as it did on several drivers) the detection worked fine and with very few false detections.

SAMMANFATTNING

För en säker körning måste föraren ha en kontinuerlig uppmärksamhet på vägen. Ett par sekunders distraktion kan annars medföra att föraren förlorar kontrollen. Samtidigt är också känt att människan har svårt att hålla koncentrationen uppe under en längre tid och här är distraktion och sömnighet kända fenomen. Dessa trafiksäkerhetsproblem har den senaste tiden fått mycket uppmärksamhet. Ett flertal nyligen genomförda studier pekar på att ungefär 10-20% av alla olyckor sannolikt är sömnrelaterade, och att distraktion har varit en bidragande faktor i upp till 78-80% av alla incidenter och olyckor.

Projektets syfte var att bidra till trafiksäkerhet genom att hålla föraren uppmärksam och alert med hjälp av ett tekniskt system.

En bil utrustades med ett system för detektion av visuell ouppmärksamhet och sömnighet hos föraren. De två viktigaste sensorerna i detta system var två kameror; en riktad mot förarens huvud och ansikte för att se fysiologiska tecken på distraktion och sömn, och en riktad mot vägen framför bilen för att se effekter av distraktion och sömn på körbeteendet. Målet med detta var att få en mer robust detektion av visuell ouppmärksamhet och sömnighet än om bara en sensor används. Detta var förklaringen av begreppet ”sensor fusion”. Bilen var även utrustad med ett varningssystem för att få föraren tillbaks i uppmärksam körning och en sömnig förare att vidta nödvändiga åtgärder.

Systemet testades i verklig trafik för att få så stor överensstämmelse med verkliga körförhållanden som möjligt.

43 förare deltog i fältstudier inriktade mot sömnighet. Varje förare körde tre pass, ett på dagen, ett på kvällen och ett på natten. Varje pass tog c:a 90 minuter. Subjektiv sömnighet utvärderades var 5:e minut med hjälp av Karolinska sömnighetsskalan (KSS) och fysiologisk sömnighet analyserades från inspelningar av EEG, EMG och EOG.

Det som gav den bästa detektionen av sömnighet var att analysera en kombination av långa blinkningar tillsammans med huvudnickning och tid på dygnet. Att även använda data från rattaktivitet eller position på vägen gav inget klart bidrag.

Vad gäller detektion av distraktion så var den mycket beroende av hög prestanda på sensorn (kamerasystemet). Även om viss signalbehandling lindrade en del av de tillkortakommanden som fanns hos sensorn så skulle en ökad prestanda vad gäller ögonrörelsemätning betyda mycket för detektionen av distraktion. För de förare där sensorn fungerade bra (som det gjorde för flertalet förare) så uppvisade distraktionsdetektionen bra prestanda med mycket få falsklarm.

1. Introduction.

Loss of control is a fact if the driver diverts from the driving task for just a few seconds. A continuous attention to driving is a must in order to drive in a safe way. But it is also known that humans are easily being distracted or drowsy. These traffic safety problems have received lot of attention lately.

Epidemiological studies based on self-reporting or in-depth crash investigations indicate that about 10-20% of all crashes are likely sleepiness-related (Horne and Reyner, 1995, Maycock, 1997, Stutts et al., 1999, Stutts et al., 2003). It has also been demonstrated in post-crash interviews that night driving, prior night sleep of less than five hours, and sleepiness level before the crash are major predictors of the risk of being involved in a road crash (Connor et al., 2002). In addition, distraction has been a contributing factor of up to 78-80% of all crashes and incidents (Klauer et al., 2006).

This research area has previously been addressed in the IVSS program (Intelligent Vehicle Safety Systems) in the projects "*DROWSI – Drowsiness intervention*" and "*Driver attention – dealing with drowsiness and distraction*". This project has taken the learning from these in order to further explore the phenomena of drowsiness and distraction, and to find appropriate countermeasures.

2. Project objective.

This research projects aim was to contribute to traffic safety by supporting a driver in order to be awake and attentive by technical means in the vehicle.

The main objectives were to a) build a system capable of detecting the driver's status regarding drowsiness and distraction, b) install this system in a real vehicle and c) execute field test in order to get valid data in real driving conditions.

The hypothesis was that using data from multiple sensors would lead to a more accurate and robust diagnosis of the drivers status than using just a single sensor. Finally, the project evaluated different approaches on how this data should be used together (here called "sensor fusion") in order to reach the best result.

3. Development of a Drowsiness and Distraction Detection System.

Sensor set up.

A vehicle (Saab 9-3 Sedan) was equipped with sensors able to deliver data relevant for detection of drowsiness and distraction. The main sensors were;

a) A driver monitoring system (Smart Eye AntiSleep) consisting of a camera tracking the drivers head- and eye movements.
b) Forward looking cameras mounted inside the windscreen (MobilEye lane tracker and Autoliv SVS) tracking road markings.
c) Steering wheel grip sensor (From I.V.S.) measuring the drivers’ grip force on the steering wheel rim. (Reported in a separate project).
d) Precision steering wheel angle sensor, an optical encoder with a resulting resolution of about 70 increments per degree.

The system (hereafter called D4SF system) was also making use of data already available through the vehicles own information system (CANdata), such as vehicle speed, steering wheel angle, yaw rate, etc.

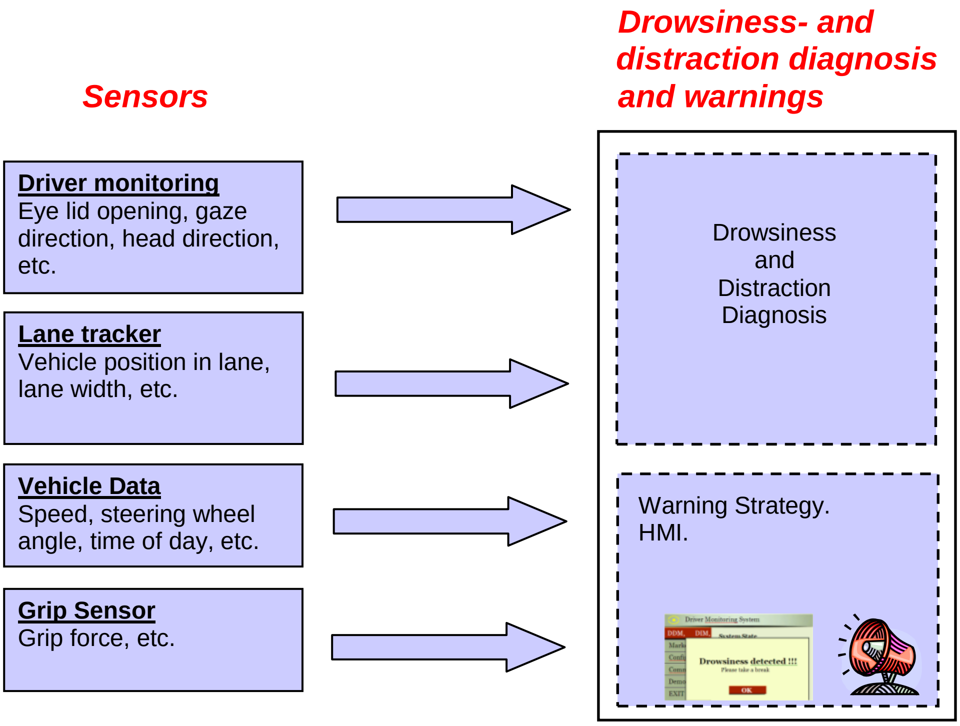

Figure 1. Description of the D4SF system functional concept. Data from the different sensors is fused together in order to reach diagnosis of drowsiness and distraction. The warning system consists of a warning strategy and a HMI.

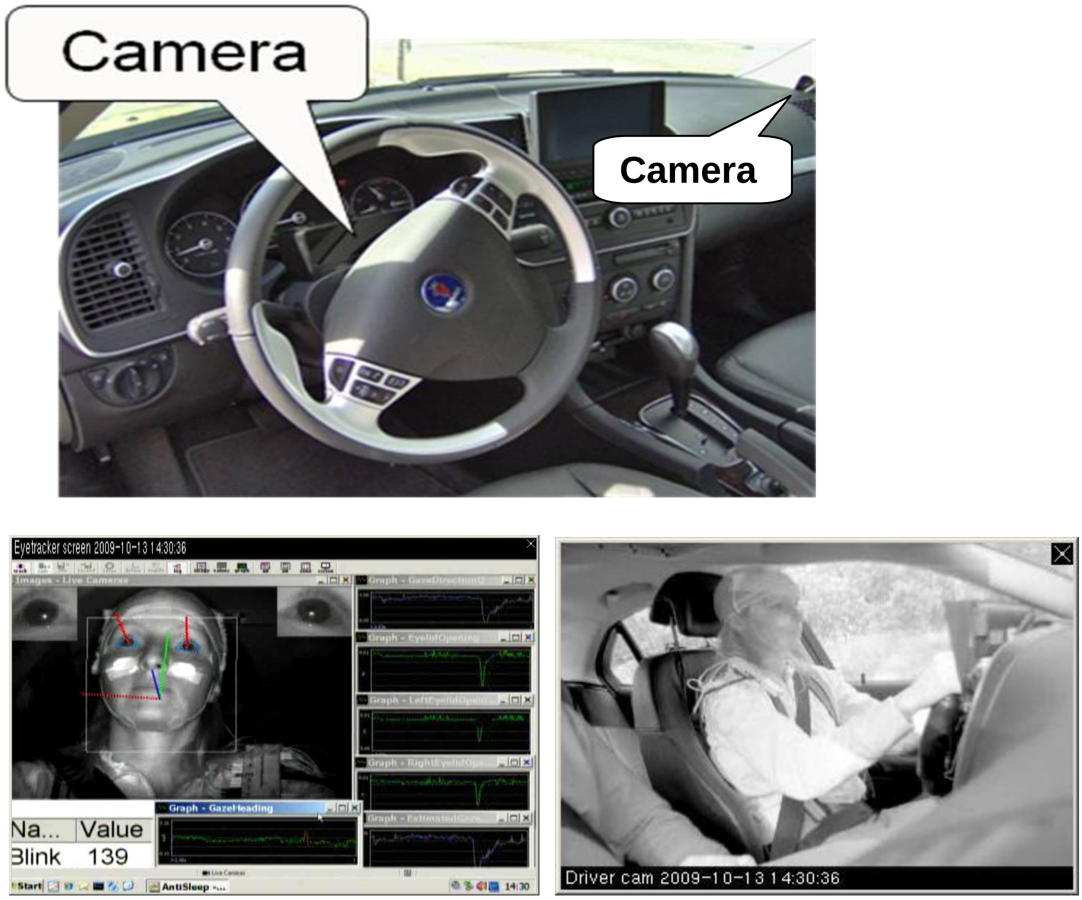

Figure 2. Location of- and output from the Smart Eye AntiSleep, as well as the cabin camera.

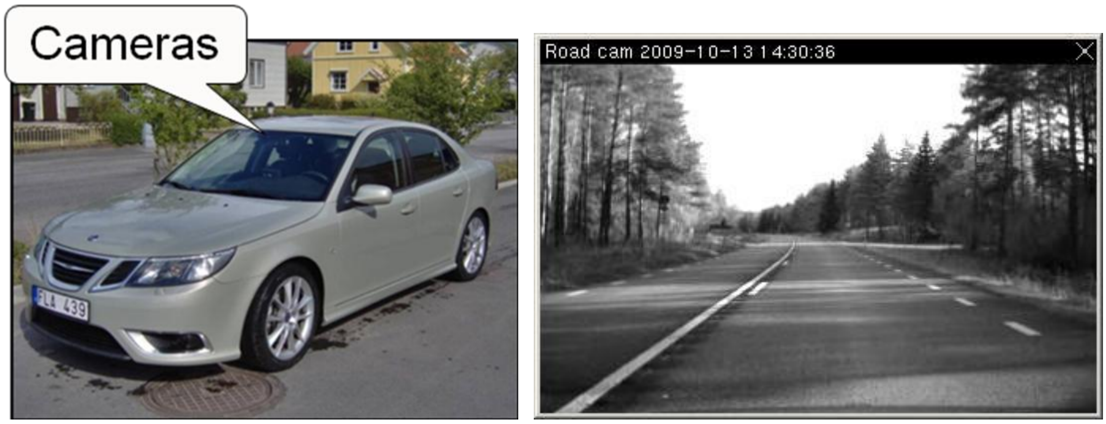

Figure 3. Location of- and output from one of the forward looking cameras.

Drowsiness indicators.

Next step was to develop indicators of drowsiness. These indicators could consider physiological signs or driving behavioral signs of a driver being drowsy. The indicators developed and used in this work for drowsiness were;

a) Long blink indicator: Average of blink duration.
b) Lane tracker indicator: Vehicle position variability.
c) Steering variability: Steering wheel position variability.
d) Head pitch indicator: Variability in head pitch.
e) Three-process model: The Three-process model, here taking only the time-of-day as input to the algorithms.

The three process model is a mathematical model for predicting sleepiness based on previous sleep patterns and a standardized circadian rhythm. (Åkerstedt 2004). In this implementation, information about previous sleep was omitted. Variability in driving performance indicators were calculated using Sandberg´s generic variability indicator (Sandberg, 2008).

Sensor fusion.

As a final step in the drowsiness detection procedure, the "sensor fusion" part, a Support Vector Machine (SVM) was trained to classify combinations of indicator values as "drowsy" or "non-drowsy" using real data and KSS drowsiness values from the field tests. SVM is a supervised learning method using pre-annotated data to create a model for classification. With the help of SVM's it is possible to fuse information from several sensors into the classifier to achieve better performance.

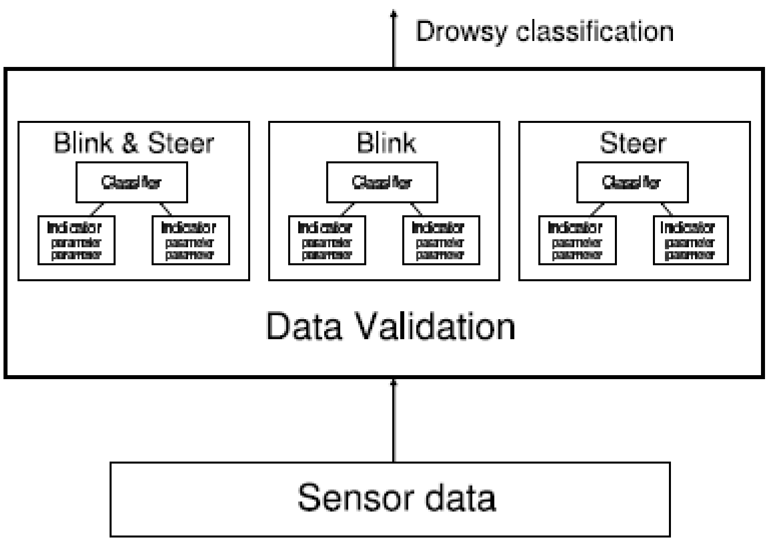

Figure 4. Drowsiness classification using Support Vector Machines as a method.

Distraction indicator.

The indicator for distraction was based on “eyes-off-road”. When the driver had been visually inattentive (not looking at the road ahead) for a period of time, he/she was classified as distracted. The indicator used data from both eye gaze- and head behavior, and their time history in order to make the final diagnosis.
The distraction indicator was mainly developed in the Saab Driving Simulator where the work could be done efficiently in a controllable and safe environment.

Driver warnings.

The vehicle was equipped with a number of actuators in order to explore what modalities were suitable to use as driver warnings. A visual warning was presented on top of the instrument panel and text messages and/or symbols were presented on a screen mounted in the center of the vehicle. Audible warnings could be presented as tones or speech. Tactile warnings could be presented as vibrations through the seat or the seat belt.

Figure 5. a) Visual warning on top of the instrument panel consisting of flashing red LED:s. b) Messages presented on a screen located in the center of the vehicle.

The warning actuators were controlled by a warning management system. The purpose of this system was to inhibit warnings that could be experienced as inappropriate or redundant.

4. Field Studies.

Field testing was conducted for various purposes. These were the main test activities:

- Initial testing: A first experiment with the aim to collect raw data on driver’s head- and eye behavior and driving behavior in different stages of sleepiness. The data was used to work with the performance from each sensor, but also as a starting point for designing the sensor fusion system.

- Data collection for algorithm development (called experiment I and experiment II). The experiments included measurement of driver physiological such as EEG, EOG and ECG, used for ground truth when looking into sensitivity and specificity. Two such tests have been done; one with 21 participants and another with 22 participants.

- System design testing: During the development of the sensor fusion system there were "engineering tests", where the functionality of the system was tested as the development progress. Special focus was on the distraction detection.

- Validation: There were extensive experiments with the D4SF system in full operation, including warnings. Drivers were employees at Autoliv and Saab, which used the vehicle for business trips or commuting.

The aim with the experiments was to collect data from alert and sleepy drivers, using all available sensors and for all algorithms at the same time. The data from experiment I (spring 2009) was used for further development of the sleepiness fusion algorithms before the start of experiment II (autumn 2009). The data from experiment II was used for validation of the system including fusion of sensors, and even further development of the drowsiness detection.

The data collection was based on a dose – response design in order to make sure to obtain data from each driver, during different levels of sleepiness and different time of the day. This was true for both experiment I and experiment II. The only major difference between the two experiments was that the driving time was extended with about 15 minutes per driver during experiment II.

Figure 6. The experimental vehicle with sign (Mätning= data collection in progress)

Procedure.

In the front seat was a test leader responsible for the dual command but also for the capture of the KSS values (Karolinska Sleepiness Score, see fig. x) manually (only used as a backup). The participants were prompted to report KSS every 5 minutes when the word "Sömnig?" (Sleepy?) appeared on a touch screen positioned to the right of the driver. Beneath this, there was an abbreviated version of the KSS scale, to aid the drivers' assessments. When the prompting screen was not visible, the screen was entirely empty in experiment I, while in experiment II, there was a small digital clock in the lower right corner, replicating the function of the original clock in the car, which was partially covered by the display.

Figure 7. The screen prompting the driver to report KSS (Karolinska Sleepiness Score). Daytime and nighttime appearance.

There was also a test leader in the back seat responsible for the data acquisition system and for recording of the KSS. Two displays were present in the back seat: a touch screen used to control the data acquisition and enter KSS and ORS data, or annotate specific events like overtaking. The second screen was displaying the Smart Eye system and made it possible for the test leader to keep track of changes in eye lid behavior (see figure 2).

<u>Experiment design.</u>

The experiment design was a repeated measure design, with an underlying idea of dose – response regarding circadian and sleep/wake components. In total 15 experiment days (24 hours) were performed, 7 days for experiment I and 8 days for experiment II. Two participants who were scheduled to participate in experiment II had to decline for personal reasons. Each 24 hours 3 participants were involved.

For experiment I in total 21 participants was recruited. For experiment II in total 22 participants were recruited and fulfilled the experiment. For both experiments there was an equal distribution for sex. In experiment I the average age was 44.8 years old (sd 7.8) and in experiment II it was 45 years old (sd 8.2).

For both experiments the participants were recruited with help of a random sample from the National register of vehicle owners. One inclusion criteria was that they should be in the age range of 30-60 years old and had a driving experience of more than 5000 km last year. Participants received a compensation of 3000 SEK.

All participants had a pre-visit at the laboratory. They filled out a background questionnaire and signed an agreement of confidentiality. The driver also filled out an informed consent form.

This experiment was based on the governmental approval (N2007/5326/TR), and ethical approval done earlier (EPN:142-07). An extra ethical approval was sent to Regional etikprövningsnämnd (2009-03-22) and accepted (EPN 142-07 T34-09).

At the experimental day the participants drove 3 times during a 24 hour period. They stayed at VTI the time between driving sessions. They were instructed to maintain a sleep/wake pattern for the three days before the test, in order to minimize confounding. The participants used the Karolinska sleepiness score (KSS) during the test. In order to be well trained the KSS was instructed during a pre-test visit at VTI. The drivers were also encouraged to practice at home. At the arrival for the test day the subjects were trained once more and it was underlined that they should report KSS once each 5th minute and that score should represent an average of how they have been feeling the past 5 minutes.

Data collection during field experiments.

Data were collected in several ways:

- Pre-driving questionnaire
- Post-driving questionnaires
- KSS-ratings
- ORS-ratings (observer rating scale, only Experiment I)
- Physiological data (EEG, EOG, EMG and ECG)
- Smart Eye AntiSleep
- Vehicle data
- Video recordings

Figure 8. EEG, EOG and EMG electrode positions

The physiological data was send to Stressforskningsinstitutet for artefact treatment, spectral analysis and scoring of Karolinska sleepiness scale (KDS). The KDS is a method that classifies polysomnographic data according to the presence of alpha or theta activity and slow eye movements. The classification is performed in twenty-second epochs, and yields measurements of 0-100%, in steps of 10%, where every 10% represents a two-second period with signs of drowsiness in EEG or EOG data.

Driving behaviour data recorded during the drive was both from internal sensors (CANdata) and external sensors (GPS and Outwards looking camera measurement systems).

Mini-Field Operational Test.

Testing was also performed as a "mini FOT" where three Saab employees used the vehicle "as their own" for 4 days. The first two days the warnings were inactive (baseline phase) and the last two the warnings were active (treatment phase). After the test there was a structured interview about the experiences with the system, acceptance and preferences.

5. Results.

Drowsiness detection.

As can be seen in table 1, drowsiness detection performance with a combination of steering wheel and eyelid indicators reached close to 80 %, counting the mean of sensitivity and specificity. Note that test data are completely unseen data, not a subset of training data.

Table 1. Performance scores from optimization runs for drowsiness detection. AS = indicator based on blink behavior, TPM = Three process model, Steer = indicator based on steering wheel variability.

	Training data			Test data		
Model	Fitness	Sensitivity	Specificity	Fitness	Sensitivity	Specificity
AS + TPM	0.79	0.64	0.94	0.81	0.67	0.96
AS + Steer + TPM	0.74	0.55	0.93	0.8	0.65	0.95

It should be noted that this represents the classification results on a forced binary classification, i.e., there is a result expected from the algorithm for each epoch of data which fulfills 'data goodness' criteria such as vehicle speed and sensor data availability. Also note that in this particular case, fitness actually drops when adding the steering wheel variability indicator.

A number of drivers have used the vehicle for business trips etc. during the various stages in the development process, and related their experience of warnings to the development team. There seems to be a trend towards better performance in terms of fewer false alerts and more timely detections of actual drowsiness, but there is no quantitative analysis available to support this observation.

Distraction detection.

It is very difficult to do performance testing of distraction in a field test because of lack of "ground truth" data, in particular reference data on 'false negative' are missing. Instead, most of the testing has to be done systematically in a controlled environment. In this project the Saab Driving simulator was used for this purpose. It had the exact same set-up of the driver monitoring system as in the test vehicle. Test with 49 participants were done in order to study availability, accuracy and precision of the AntiSleep sensor (Ahlström et.al. In press). This test showed that the system was capable of tracking the head- and eye in the area of interest for detection of sleepiness and distraction of a driver.

Results from the mini FOT showed that there were big differences in performance depending on participant characteristics. For some participants the AntiSleep showed poor performance but for others good. Also, the test revealed some other areas were improvement would be favorable, for example when and how to initialize the system. For the drivers where the AntiSleep worked well the distraction detection showed good performance. Only a few false detections could be observed (1-3 per hour driven).

Warnings during mini-FOT.

The warnings in the test vehicle were set to be rather strong on purpose in order to test the upper limit of driver acceptance. This meant that each warning had multiple modalities. Also, drowsiness warnings were repeated as long as the drive was regarded as drowsy. Some of the drivers reported that they experienced the warnings to be too strong. Especially the first time they had a warning was experienced as startling. Also, when having warnings repeated frequently, as for example once a minute a drowsiness warning, was experienced as very annoying.

6. Fulfillment of FFI Traffic safety objectives.

Driver Impairment, cognition and acceptance. (*Nedsatt körförmåga hos förare & Förarens kognition och tolerans*): The project has addressed driver impairment regarding drowsiness and distraction. Special focus has been on adapting systems to real road conditions and to collect data during real driving. The results reveal some difficulties when trying to apply laboratory results to real conditions.

Test/validation methods for safety systems (e.g. field studies). (*Metoder för test/validering av säkerhetssystem (t.ex. fältstudier)*): A vehicle was equipped with a state-of-the-art measuring system tailored to collect all data of relevance in an effective way. It facilitates supervised data collection as well as unsupervised. It has made it possible for the project to use the vehicle for a variety of purposes (functional tests, controlled tests, field operational tests and demonstrations).

Intelligent Safety Systems (e.g. integration of different safety systems). *(Intelligenta säkerhetssystem (t.ex. integration av olika säkerhetssystem)):* The developed system takes use of several systems also used for other safety functions, such as the seat belt and the lane tracker (used for lane departure warning). Additional to this it also use already available data from the vehicles own information system (engine, brakes, speed, etc.).

Zero vision/reduce number of fatalities and serious injuries in traffic. *(Nollvisionen/Reducera antalet dödade och allvarligt skadade i trafiken)*: The project aimed to contribute to traffic safety by promoting an alert and attentive driver by technical means in the vehicle.

Swedish and Swedish automotive industry competitiveness (strenghten brands and cost effective solutions). *(Konkurrenskraft för Sverige och den svenska fordonsindustrin (Förstärkning av varumärken och kostnadseffektiva lösningar)*: The project team lined up a complete chain of competence for an innovation system – from science to product. (VTI – science and research, Smart Eye - technology forefront, Autoliv – automotive supplier and Saab Automobile – OEM).

7. Conclusions and recommendations.

Drowsiness detection.

Best performance with current implementation of indicators is the combination of long blink, head pitch and three-process model. Adding either the steering- or lane tracker indicator doesn't give any clear performance benefit. This doesn't necessarily mean that these indicators are unusable, only that the current implementations of them are not good enough. Furthermore, driving performance based indicators may also detect driving impairment from other causes than drowsiness, which can reduce the performance score.

One example of a problem with the generic variability indicator is the current implementation of the steering wheel indicator. It optimizes to be sensitive to very small corrections with the steering wheel. If the vehicle dynamics or environment changes, the indicator stops working as intended.

Finally – the current implementation of the fitness function using sensitivity and specificity might not be the best solution. For example, looking at data from both experiments, the Lane-Steer-TPM combination had almost the same fitness value as Long blink-Head pitch-Steering-TPM. But when looking at the number of found drowsy periods, the Long blink-Head pitch-TPM combination had almost 30% better performance. Thus, there's clearly some room for improvements of the fitness function.

Distraction detection.

The distraction detection performance was very dependent on a good performance of the Driver monitoring system. Although signal- and data treatment helped to overcome some shortcomings, an improved capability to track eye movements would mean a lot to the overall performance. However, when the driver monitoring system worked well (as it did on several drivers) the detection worked fine and with very few false detections.

Driver warnings.

Some of the drivers reported that they experienced the warnings to be too strong. Especially the first time they had a warning was experienced as startling. This emphasizes the need for careful tuning of warnings; to be mild in order to gain user acceptance but not too weak so that it can be missed, and strong enough to be effective but not too strong in order to gain user acceptance. There is also a need to look deeper into the semantics of the drowsiness warning and the timing of it in relation to drowsiness status. How should an early warning be expressed? How will the driver react on a strong warning when very sleepy?

Field tests.

The field tests were carried out successfully with only minor data loss due to technical problems. Also, most of the planned tests could be executed and only a few participants were excluded (could not participate due to personal reasons) and only a few test runs were excluded due to poor weather conditions. All test procedures and staffing worked well and can be recommended for future tests. What is needed is further testing in other road environments. So far tests have been done on one road type only (two-lane rural road) and additional data collection must be done, primarily on motorways.

8. Project deliverables.

1. Test vehicle equipped with a D4SF system including warnings and a data logging system for field tests.
2. A D4SF system including warnings installed in Saab Driving simulator.
3. SW module for distraction detection based on AntiSleep system.
4. SW module for drowsiness detection based on AntiSleep system.
5. SW module for drowsiness diagnose with input from multiple indicators.
6. SW module for driver warning management.
7. Field test database from experiment I and experiment II.
8. Report: Experiment I and II, set-up and execution.
9. Report: Drowsiness detection specification.
10. Report: Drowsiness performance.
11. Report: Distraction specification.
12. Report: Test vehicle specification.
13. Report: Owner’s manual for D4SF.

.

9. References.

IVSS - DROWSI - Drowsiness Intervention. *www.ivss.se.*
IVSS- Driver attention - dealing with drowsiness and distraction. *www.ivss.se.*

Ahlström, C., et.al. Report on head- and eye gaze measurement; performance of a one- camera system and a three-camera system. *Project report ViP-VisualEyes. (in press) www.vipsimulation.se.*

Anund, A. Sleepiness at the Wheel. *Ph.D. Thesis, department of Public Health Sciences NASP – National Prevention of Suicide and Mental Health. Karolinska Institutet. Stockholm*. 2009.

Åkerstedt, T., Folkard, S. & Portin, C, Predictions from the three-process model of alertness*, Aviation Space, and Environmental Medicine, 75 (2004), pp. A75-A83*

Connor, J., Norton, R., Ameratunga, S., Robinson, E., Civil, I., Dunn, R., Bailey, J. and Jackson, R. Driver sleepiness and risk of serious injury to car occupants: population based case control study. *Br. Med. J.* , 2002, 324 1125.

Horne, J. A. and Reyner, L. A. Driver sleepiness. *J Sleep Res*, 1995, 4 (suppl 2): 23-29.

Klauer, S. G., Dingus, T. A., Neale, V. L., Sudweeks, J. and Ramsey, D. The impact of driver inattention on near-crash/crash risk: An analysis using the 100-car naturalistic driving study data In, Washington DC: NHTSA, 2006.

Maycock, G. Sleepiness and driving: The experience of U.K. car drivers. *Accid Anal Prev*, 1997, 29: 453-62.

Sandberg, D. Analysis and Optimization of Systems for Detecting Sleepiness in Drivers. *Licentiate thesis, Chalmers University of Technology*, 2008.

Stutts, J., Wilkins, J. W., Osberg, S. V. and Vaughn, B. V. Driver risk factors for sleep-related crashes. *Accid Anal Prev*, 2003, 35: 321-31.

Stutts, J. C., Wilkins, J. W. and Vaughn, B. V. Why do people have drowsy driving crashes? Input from drivers who just did. In. AAA Foundation for Traffic Safety, Washington, D.C., 1999: 85.