

DRIVER ALCOHOL DETECTION SYSTEM FOR SAFETY (DADSS) – A STATUS UPDATE.

Abdullatif K. Zaouk

Michael Wills

QinetiQ North America

U.S.A.

Eric Traube

National Highway Traffic Safety Administration

U.S.A.

Robert Strassburger

Alliance of Automobile Manufacturers

U.S.A.

Paper Number 15-0276

ABSTRACT

The National Highway Traffic Safety Administration (NHTSA) and the Automotive Coalition for Traffic Safety (ACTS) began research in February 2008 to try to find potential in-vehicle approaches to the problem of alcohol-impaired driving. Members of ACTS comprise motor vehicle manufacturers representing approximately 99 percent of light vehicle sales in the U.S. This cooperative research partnership, known as the Driver Alcohol Detection System for Safety (DADSS) Program, is exploring the feasibility, the potential benefits of, and the public policy challenges associated with a more widespread use of non-invasive technology to prevent alcohol-impaired driving. The 2008 cooperative agreement between NHTSA and ACTS for Phases I and II outlined a program of research to assess the state of detection technologies that are capable of measuring blood alcohol concentration (BAC) or Breath Alcohol Concentration (BrAC) and to support the creation and testing of prototypes and subsequent hardware that could be installed in vehicles. This paper will outline the technological approaches and program status.

INTRODUCTION

Alcohol-impaired driving (defined as driving at or above the legal limit in all states of 0.08 g/dL or 0.08 percent) is one of the primary causes of motor vehicle fatalities on U.S. roads every year and in 2011 alone resulted in almost 10,000 deaths. There are a variety of countermeasures that have been effective in reducing this excessive toll, many of which center around strong laws and visible enforcement. Separate from these successful countermeasures, the National Highway Traffic Safety Administration (NHTSA) and the Automotive Coalition for Traffic Safety (ACTS) began research in February 2008 aimed at identifying potential in-vehicle approaches to the problem of alcohol-impaired driving. Members of ACTS comprise motor vehicle manufacturers representing approximately 99 percent of light vehicle sales in the U.S. This cooperative research partnership, known as the Driver Alcohol Detection System for Safety (DADSS) Program, is exploring the feasibility, the potential benefits of, and the public policy challenges associated with a more widespread use of non-invasive technology to prevent alcohol-impaired driving. The 2008 cooperative agreement between NHTSA and ACTS (the “Initial Cooperative Agreement”) for Phases I and II outlined a program of research to assess the state of detection technologies that are capable of measuring blood alcohol concentration (BAC) or Breath Alcohol Concentration (BrAC) and to support the creation and testing of prototypes and subsequent hardware that could be installed in vehicles.

Since the program’s inception it has been clearly understood that for in-vehicle alcohol detection technologies to be acceptable for use among drivers, many of whom do not drink and drive, they must be seamless with the driving task, they must be non-intrusive, that is, accurate, fast, reliable, durable, and require little or no maintenance. To that end, the DADSS program is developing non-intrusive technologies that could prevent the vehicle from being driven when the device registers that the driver’s blood alcohol concentration (BAC) exceeds the legal limit (currently 0.08 percent throughout the United States).

To achieve these challenging technology goals, very stringent performance specifications are required. These specifications have been formally documented in the DADSS Performance Specifications, which provide a template to guide the overall research effort. Another important challenge will be to ensure that the driving public will accept in-vehicle alcohol detection technology once it meets the stringent criteria for in-vehicle use. A parallel effort is underway to engage the driving public in discussions about the technologies being researched so that their feedback can be incorporated into the DADSS Performance Specifications as early as possible. The challenges to meet these

requirements are considerable, but the potential life-saving benefits are significant. An analysis of NHTSA's Fatality Analysis Reporting System (FARS) estimates that if driver BACs were no greater than 0.08 percent, 7,082 of the 10,228 alcohol-impaired road user fatalities occurring in 2010 would have been prevented.

The research effort that comprised the Initial Cooperative Agreement followed a phased process. The five-year Initial Cooperative Agreement began with a comprehensive review of emerging and existing state-of-the-art technologies for alcohol detection in order to identify promising technologies. Phase I, completed in early 2011, focused on the creation of proof-of-principle prototypes. The objective of Phase I was to determine whether there were any promising technologies on the horizon. Three prototypes were delivered and tested at the DADSS laboratory that yielded promising results for two of the three technologies.

The Phase II effort, begun in late 2011 and completed in late 2013, focused on the continued research of the technology to narrow gaps in performance against the DADSS Performance Specifications and meet the DADSS Performance Specifications within the needs of an in-vehicle environment.

Phase III and subsequent phases of research – the focus of the current Cooperative Agreement – will permit further refinement of the technology and test instruments as well as basic and applied research to understand human interaction with the sensors both physiologically and ergonomically – that is how these technologies might operate in a vehicle environment. At the culmination of this Agreement will be a device or devices that will allow a determination to be made regarding whether the DADSS technologies can ultimately be commercialized. If it is determined that one or more of these technologies can be commercialized, it is currently anticipated that the private sector will engage in additional product development and integration into motor vehicles.

The purpose of this paper is to outline the technological approaches taken in developing alcohol detection hardware. These approaches are founded on a clear understanding of the processes by which alcohol is absorbed into the blood stream, distributed within the human body, and eliminated from it. Not only must technologies under consideration quickly and accurately measure BAC, but the medium through which it is measured (e.g., breath, tissue, sweat, etc.) must provide a valid and reliable estimation of actual BAC levels. Alcohol absorption, distribution, and elimination measurement is a topic about which much has been written yet some large gaps in our understanding still remain. This paper will provide an overview of what is known regarding alcohol measurement via various methods and their implications for the decisions about which technologies deserve further study. The paper also will provide an overview of the current performance specifications developed to assess the in-vehicle advanced alcohol detection technologies and the rationale for them as well as an overall status of progress made to date.

ALCOHOL ABSORPTION, DISTRIBUTION, AND ELIMINATION IN HUMANS

The science of pharmacokinetics is concerned with the ways in which drugs and their metabolites are absorbed, distributed, and eliminated from the body (Jones, 2008). This is separate from pharmacodynamics which is the study of the physiological effects of drugs and their actions on the body (Buxton, 2006 see Jones paper). Ethyl alcohol or ethanol, more commonly referred to as alcohol, is only one of a family of organic compounds known as alcohols. Ethanol, referred to hereafter as alcohol, is highly soluble in the body's water, which makes up 50-60 percent of body weight. Even though alcohol is a central nervous system depressant, people perceive it as a stimulant and in the early stages it can produce feelings of euphoria (Jones, 2008). With the consumption of larger amounts of alcohol, performance and behavior can be impaired resulting in reduced coordination, loss of motor control, lack of good judgment, and at very high concentrations (greater than 0.4 g/dL) loss of consciousness and death. Figure 1 portrays schematically the pathways by which alcohol is absorbed into the blood stream, is distributed throughout the body, and eliminated from it.

After ingestion, alcohol enters the stomach where it is partially absorbed through the stomach wall (about 20 percent), and then to the small intestines where most of the absorption takes place (about 80 percent). Alcohol is then transported to the liver and on to the heart before it is distributed by the arteries throughout all body fluids and tissues. Alcohol easily passes the blood-brain barrier where it affects central nervous system functioning. The time required for reaching equilibrium depends on the blood flow to the various organs and tissues, but over time alcohol mixes completely with all the water in the body and reaches into all fluid compartments within the body.

The characteristics of alcohol's distribution and elimination can point to potential ways in which BAC can be measured. There are two mechanisms by which alcohol is eliminated from the body, metabolism and excretion. The liver is the primary organ responsible for the elimination of alcohol and it is where about 95% of ingested alcohol is metabolized. The remainder of the alcohol, about 2-5 percent, is excreted unchanged wherever water is removed from the body; through the skin in sweat, from the lungs in breath, from the eyes in tears and from the

kidneys in urine. As noted above, alcohol distributes completely into all the body's compartments so alcohol can be measured in vivo in bodily tissue.

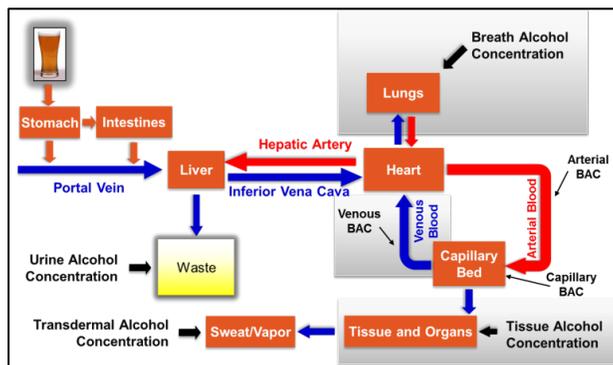


Figure 1. Alcohol absorption, distribution and elimination through the body

Methods used to measure blood alcohol concentration

For many years the only means to determine BAC was through blood and urine testing. As early as 1874 it was recognized that ingested alcohol can be measured in breath (Jones, 2008), and the smell of alcohol on breath is a well-known indication that someone has been drinking. Accurate measurement of alcohol in expired air has a physiological basis. Under normal lung function there is an efficient gas exchange between blood and gases, thus resulting in a close correlation between blood and gas concentrations of alcohol (Hök, 2006). Furthermore, a recent study (Lindberg et al., 2007) has established that the concentration of alcohol in breath is in very close agreement to that of alcohol in arterial blood (Figure 2), even though the gold standard for equating breath to blood alcohol is venous BAC. Of note is that arterial BAC is a better indication of brain alcohol and hence impairment than venous blood, so BrAC is particularly well suited as a measure of driver impairment.

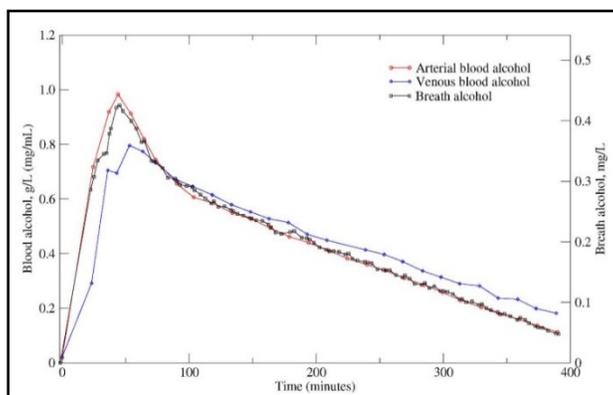


Figure 2. Pharmacokinetic profile in one subject showing concentrations of alcohol in arterial blood (ABAC), venous blood (VBAC) and breath after oral ingestion of 0.6 g of alcohol per kg body weight

Dr. Robert Borkenstein is recognized as the inventor of the first system that measured alcohol on a person's breath. In 1954, he invented the first breath testing device, which used chemical oxidation and photometry to measure alcohol concentration. Subsequently physiochemical methods were developed for the measurement of alcohol in breath such as gas chromatography, electrochemical oxidation, and infra-red analysis. Breath testing has flourished because it is non-invasive and, in contrast to urine and blood samples that have to be sent away for testing, provides on-the-spot results. As a result most countries have adopted breath testing both for roadside screening and evidential purposes to establish BAC.

In recent years a number of other approaches have been identified that could be used to measure alcohol in perspiration (either vapor phase or liquid phase) or from measurements of alcohol in a person's tissue. As noted below, these techniques have not yet been widely used to measure alcohol concentration.

TECHNOLOGICAL APPROACHES

One of the first tasks of the project team was to perform a comprehensive review of emerging and existing state-of-the-art technologies for alcohol detection (Ferguson et. al., 2010). Technology scans were undertaken through patent and literature reviews. Based on these reviews four categories of technologies were identified with potential for measuring driver BAC within the vehicle environment:

1. Electrochemical/Transdermal Systems

Electrochemical Systems are chemical-reaction-based devices such as transdermal and breath-based systems. Alcohol in the presence of a reactant chemical will produce colorimetric changes measured by spectral analysis or a semi-conductor sensor. In fuel cell systems (typically used in current technology – breath-alcohol ignition interlocks), exhaled air containing alcohol passes over platinum electrodes which oxidize the alcohol and produce an electrical current; the more alcohol in the air sample, the greater the electrical current. The electrical current level permits accurate calculation of breath alcohol concentration (BrAC) which can be converted to blood alcohol concentration (BAC) using a standardized conversion factor.

2. Tissue Spectrometry Systems

Tissue Spectrometry Systems allow estimation of BAC by measuring the alcohol concentration in tissue. This is achieved through detection of light absorption at a particular wavelength from a beam of Near-Infrared (NIR) reflected from within the subject's tissue. As classified herein, they are touch-based systems and require skin contact. Variations of tissue spectrometry systems include Michelson, Raman, Fabry-Perot, Laser Diode and Light Emitting Diode (LED) based devices.

3. Distant/Offset Spectrometry Systems

Distant Spectrometry Systems use an approach that is similar to Tissue Spectrometry, except that no skin contact is required. Infrared (IR) is transmitted toward the subject from a source that also has a sensor to receive and analyze the reflected and absorbed spectrum to assess alcohol concentration in the subject's exhaled breath.

4. Behavioral Systems

Behavioral Systems detect impaired driving through objective behavioral measures. These include ocular indices such as gaze and eye movement, driving performance measures, as well as other performance measures believed to be related to driving performance.

In addition to the technology scans, a Request for Information (RFI) was published as a means by which the DADSS program was first communicated to potential technology developers. The goal of the RFI was to establish the level of interest among technology developers in taking part in the research, the kinds of technologies available, and their states of development relevant to in-vehicle applications. Based on information gleaned during the RFI process, a subset of technology companies were selected to receive a Request for Proposal. Detailed evaluation of the proposals that were received resulted in awards to technology companies based on two of the technological approaches outlined above; tissue spectroscopy and distant spectroscopy. The electrochemical/transdermal and behavioral approaches were not being pursued due their limitations for the DADSS application.

Current breath-based measurement systems as well as transdermal systems that measure alcohol in vapor or liquid phase perspiration, utilize electrochemically-based fuel-cell technology which has several limitations. Fuel cells must be warmed up to breath temperature to meet accuracy requirements, which in cold climates can take several minutes to accomplish. Furthermore, fuel cell alcohol measurements experience drift over time and require recalibration within one year or less. Both of these aspects render fuel cell technologies unsuitable for every-day use by the general public. Transdermal fuel-cell based devices suffer from the same disadvantages; however, there is an additional concern due to the long lag time of peak alcohol concentration in sweat versus blood. Various studies have identified the lag times to be on the order of two hours or more and it is not clear how future technological approaches to measuring TAC at a point in time can address this fundamental physiological difference.

Interest in behavioral-based approaches to measuring alcohol impairment dates back to the 1970s when the government and industry collaborated on possible vehicle-based measures of impairment (Ferguson et. al., 2010). There are a large number of measurable behaviors that have been identified that are affected by alcohol, including eye movements, reaction times, and vehicle-based measures of impairment such as lane position variability/lateral position, changes in driving speed and speed variability, pedal and steering control, distance from the car in front, and delay in motor actions and responses such as braking reaction times. Researchers have examined the

relationship of BAC to changes in these behaviors; however, close correlations between these measures have not yet been established. Another issue is that behavioral task performance may change as a result of a variety of impairments, whether from fatigue, illness, alcohol, medications, illegal drugs, or other sources. Furthermore, in order to measure impairment there has to be some measure of “normal” abilities on the specific task that can act as a baseline measurement for comparison. It should be noted that other sources of impairment can result in unsafe driving, and research continues to identify those risks and determine potential countermeasures. However, the limitations outlined above would be hard to deal with in an unobtrusive device to measure alcohol.

PERFORMANCE SPECIFICATIONS

ACTS developed performance specifications to assess the in-vehicle advanced alcohol detection technologies that are being developed. The specifications are designed to focus the current and future development of relevant emerging and existing advanced alcohol detection technologies (Ferguson et. al., 2010). In addition to requirements for a high level of accuracy and very fast time to measurement, the influences of environment, issues related to user acceptance, long-term reliability, and system maintenance also will be assessed. The resulting list of specifications with definitions, measurement requirements, and acceptable performance levels are documented in the DADSS Subsystem Performance Specification Document. The accuracy and speed of measurement requirements adopted by the DADSS Program are much more stringent than currently available commercial alcohol measurement technologies are capable of achieving. As noted above, the devices would need to be seamless with the driving task and not inconvenience drivers. Translating that to appropriate performance specifications was approached by estimating the potential for inconvenience if reliability, accuracy, and time to measurement were set at various levels. Presented below are the processes used to derive them.

Reliability Developing an alcohol detection device as original equipment for the vehicle environment brings with it special challenges. Reliability is defined as the ability of a system or component to perform its required functions under stated conditions for a specified period of time. Levels of reliability that are too low would result in an unacceptable number of failures to operate the vehicle. It has been estimated that at the 3σ reliability (sigma - Greek letter σ - is used to represent the standard deviation of a statistical population) there could be the potential for 66,800 defects per million opportunities, where an opportunity is defined as a chance for nonconformance. The accepted level of reliability within the industry is 6σ . The term "six sigma process" comes from the notion that with six standard deviations between the process mean and the nearest specification limit, there will be practically no items that fail to meet specifications. In practice, 6σ is equivalent to 99.9997% efficiency. Processes that operate with "six sigma quality" over the short term are assumed to produce long-term defect levels below 3.4 defects per million opportunities.

Accuracy and Precision Accuracy is defined as the degree of closeness of a measured or calculated quantity to its actual (true) value (also referred to as the Systematic Error – SE). Precision is the degree of mutual agreement among a series of individual measurements or values (also referred to as the Standard Deviation – SD). To limit the number of misclassification errors, accuracy and precision must be very high, otherwise drivers may be incorrectly classified as being over the threshold (false positives), or below the legal limit (false negatives). To assure that drivers with BACs at or above the legal limit will not be able to drive, while at the same time allowing those below the limit to drive unhindered, SE and SD requirements at a BAC of 0.08 g/dL will need to achieve levels of 0.0003%. See Table 1 for the accuracy (SE) and precision (SD) requirements at other BACs.

Table 1. DADSS Performance Specifications (% BAC or % BrAC)

Ethanol concentration	SE	SD
0.020	0.0010	0.0010
0.040	0.0010	0.0010
0.060	0.0007	0.0007
0.080	0.0003	0.0003
0.120	0.0010	0.0010

Speed of measurement Another important performance requirement is that time to measurement be very short. Sober drivers should not be inconvenienced each and every time they drive their vehicle by having to wait for the

system to function. Current breath-based alcohol measurement devices can take 30 seconds or more to provide an estimate of BAC. However, it was determined that the DADSS device should take no longer to provide a measurement than the current industry standard time taken to activate the motive power of the vehicle. Thus, the subsystem should be capable of providing a reading of the current BAC and communicating the result within 325 msec. It should be capable of providing a second reading, if necessary, within 400 msec.

THE DADSS SUBSYSTEMS

Tissue Spectrometry: Takata-TruTouch Touch-based Subsystem

Also known as near-infrared (NIR) spectrometry, this is a noninvasive approach that utilizes the near infrared region of the electromagnetic spectrum (from about 0.7 μm to 2.5 μm) to measure substances of interest in bodily tissue (Ferguson et. al., 2010). NIR spectroscopy is the science that characterizes the transfer of electromagnetic energy to vibrational energy in molecular bonds, referred to as absorption, which occurs when NIR light interacts with matter. Most molecules absorb infrared electromagnetic energy in this manner. The specific structure of a molecule dictates the energy levels, and therefore the wavelengths, at which electromagnetic energy will be transferred. As a result, the absorbance spectrum of each molecular species is unique. Better-known applications include use in medical diagnosis of blood oxygen and blood sugar, but devices have been developed more recently that can measure alcohol in tissue (Ridder et al., 2005).

Although the entire NIR spectrum spans the wavelengths from 0.7 – 2.5 μm , TruTouch has determined that the 1.25-2.5 μm portion provides the highest sensitivity and selectivity for alcohol measurement. The 0.7-1.25 μm portion of the NIR is limited by the presence of skin pigments such as melanin that can create large differences among people, particularly of different ethnicities. In contrast, the longer wavelength portion of the NIR, from 1.25-2.5 μm , is virtually unaffected by skin pigment (Anderson et al., 1981). One other advantage of using this part of the spectrum is that the alcohol signal in the 1.25-2.5 μm region is hundreds of times stronger than the signal in the 0.7-1.25 μm part of the NIR.

For the 1st generation prototype, as shown in Figure 2, the measurement begins by illuminating the user's skin with NIR light which propagates into the tissue (the skin has to be in contact with the device). A portion of the light is diffusely reflected back to the skin's surface and collected by an optical touch pad. The light contains information on the unique chemical information and tissue structure of the user. This light is analyzed to determine the alcohol concentration and, when applicable, verify the identity of the user. Because of the complex nature of tissue composition, the challenge is to measure the concentration of alcohol (sensitivity) while ignoring all the other interfering analytes or signals (selectivity).

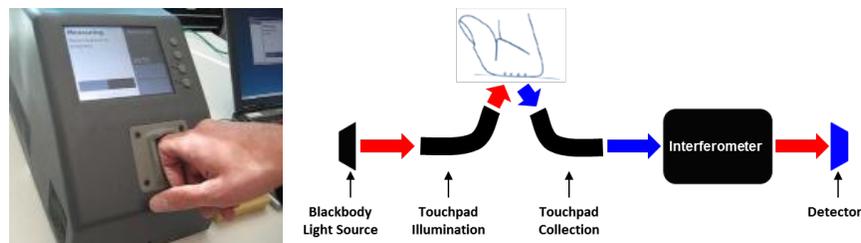


Figure 2. Touch-based subsystem 1st generation sensor and block diagram

Currently, the 2nd generation prototype is undergoing a fundamental change in system architecture; namely, a shift from a bulky spectrometer engine with moving parts to a fully solid-state sensor. This new approach, shown in Figure 3, requires extensive hardware and software research, the aims of which are to transform the touch-based sensor to improve suitability for long-term in-vehicle use and to improve the signal to noise ratio for better accuracy, precision, and shorter measurement times. The key enabling innovation is the ability to define an optimized subset of optical wavelengths which provide a high quality non-invasive alcohol measurement in humans. The 2nd generation uses modulated laser diodes to generate 40 unique wavelengths of light for alcohol measurement. The necessary laser diode target specifications were derived from an analysis of the human subject system data with accurate comparative reference data. The proposed design is also based on a Hadamard laser modulation scheme, a multiplexing technique, to improve signal to noise, along with re-design of the electronics, fiber-optical assembly, reference, touchpad and software controls to approach the necessary environmental and durability requirements for an automotive sensor device.

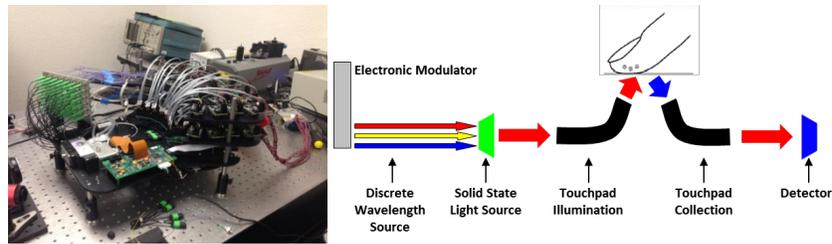


Figure 3. Touch-based subsystem 2nd generation sensor and block diagram

The focus of the current effort is to validate the new system architecture using 40 single laser packages with the goal of acquiring, verifying and integrating the full set of required multi-lasers packages into the 3rd generation benchtop system. Based on success of the 3rd generation sensor, the plan is to evolve and integrate into the DADSS research vehicle as show in Figure 4.

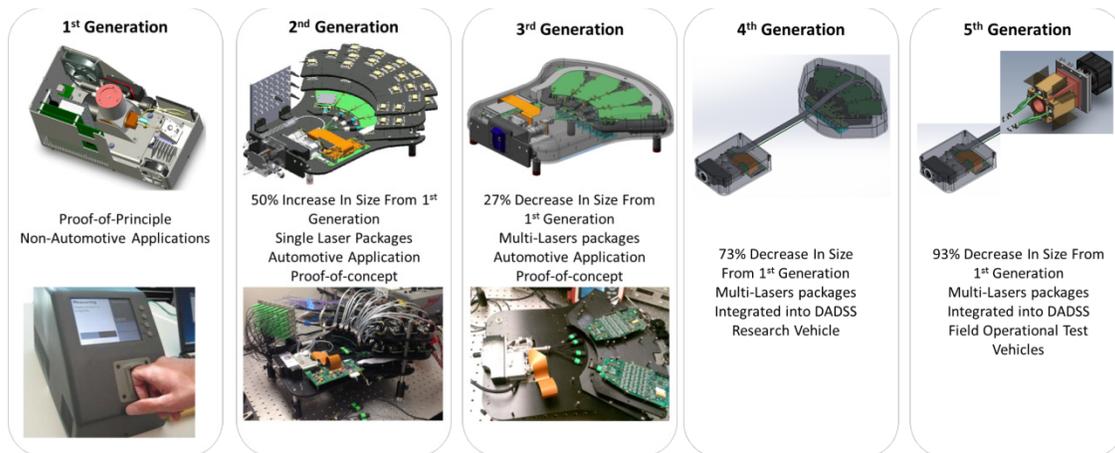


Figure 4. Evolution of solid state touch-based DADSS subsystem

Autoliv Breath-based Subsystem

Distant spectrometry systems use an approach similar to tissue spectrometry, in that they utilize the mid infrared (MIR) region of the electromagnetic spectrum (2.5-25 μm), except that no skin contact is required (Ferguson et. al., 2010). Infrared light is transmitted toward the subject from a source that receives and analyzes the reflected and absorbed spectrum to assess alcohol concentration in the subject's exhaled breath. There are a number of approaches under development that aim to remotely analyze alcohol in breath either within the vehicle cabin or around the driver's face without the driver having to provide a deep-lung breath sample.

As mentioned above, under normal lung function there is an efficient gas exchange between blood and gases, resulting in a close correlation between blood and breath alcohol concentrations (Hök, 2006) reflecting the very rapid equilibrium kinetics between pulmonary capillary blood and alveolar air (Opdam et al., 1986). In fact, as seen in Figure 2, BrAC measurements (converted to units of BAC) track arterial BACs throughout the blood alcohol time curve; only slightly below during the ascending curve, then virtually identical on the descending limb of the BAC time curve (Lindberg et al., 2007).

Current breath-based alcohol measurement techniques require direct access to undiluted deep-lung air, and therefore employ a mouthpiece. The challenge in measuring alcohol in breath from around the driver's face or within the vehicle cabin is that the breath is diluted with the cabin air. With funding from the Swedish Road Administration, Autoliv, Hök Instruments AB, and SenseAir AB have collaborated in the development of a non-contact method to measure alcohol in breath. The measuring principle of the sensor is to use measurements of expired carbon dioxide (CO_2) as an indication of the degree of dilution of the alcohol concentration in expired air. Normal concentration of CO_2 in ambient air is approximately 400 parts per million or 0.04%. Furthermore, CO_2 concentration in alveolar air is both known and predictable, and remarkably constant. Thus, by measuring CO_2 and alcohol at the same point, the degree of dilution can be compensated for using a mathematical algorithm. The ratio between the measured

concentrations of CO₂ and alcohol, together with the known value of CO₂ in alveolar air, can provide the alveolar air alcohol concentration.

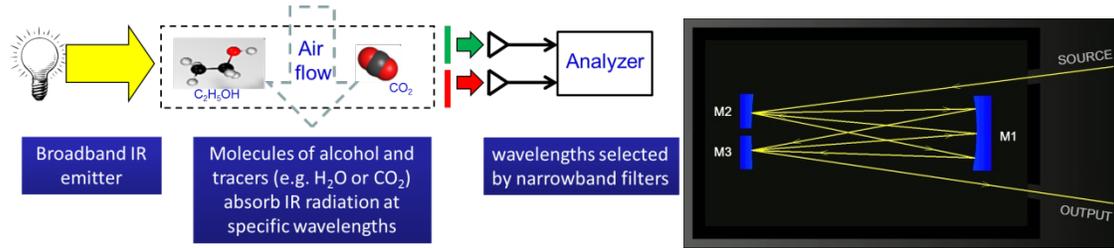


Figure 5. Breath-based sensor block diagram

The sensor technology under development by Autoliv and its partners uses infrared (IR) spectroscopy, which is superior to conventional fuel-cell devices in two ways. The IR-based sensors can be stable over the full product lifetime, eliminating the need for recurrent calibrations. Furthermore, the IR sensor is not as sensitive as the fuel-cell to major variations in ambient temperature. The 1st generation prototype uses a patented optical device in which multiple reflections of the IR beam within a closed space enables the calculation of alcohol concentration with high resolution. The expired breath from the driver is drawn into the optical module through the breathing cup. Once in the chamber, IR light is emitted from a light source and reflected by mirrors to increase the overall length of the IR optical path as shown in Figure 5, thus increasing the prototype’s resolution. Detectors in the module then measure the ethanol and CO₂ concentrations. For the purposes of human subject testing, the current device requires drivers to blow towards the sensor, which is positioned at a distance of 5 inches.

The 2nd generation sensor underwent incremental improvement that primarily involved a change in material composition of the sensor optical housing as well as significant improvements in mirror fabrication, coating, and integrated heaters designed to improve startup time, accuracy and precision. Significant progress was made in the 2nd generation with improvement to the startup time, dynamic accuracy and measurement performance at very low temperatures. The sensor underwent a series of Verification and Validation (V&V) tests as per the DADSS Performance Specification. The results from the V&V tests showed that there was no observed degradation or aging after these tests which simulated a vehicle life time of fifteen years.

The 2nd generation optical sensing element is too bulky and not suitable for vehicle integration. Further improvements are required to meet the DADSS specification. In addition, the improved sensor should be more robust when exposed to thermal gradients during the startup sequence. The focus of the current effort is to design, test, and validate a smaller, more robust optical sensor cavity that may be more easily packaged into a motor vehicle, with the objective remaining that the devices meet or exceed the DADSS Performance Specifications. Figure 6 shows the evolution of the Autoliv sensors.

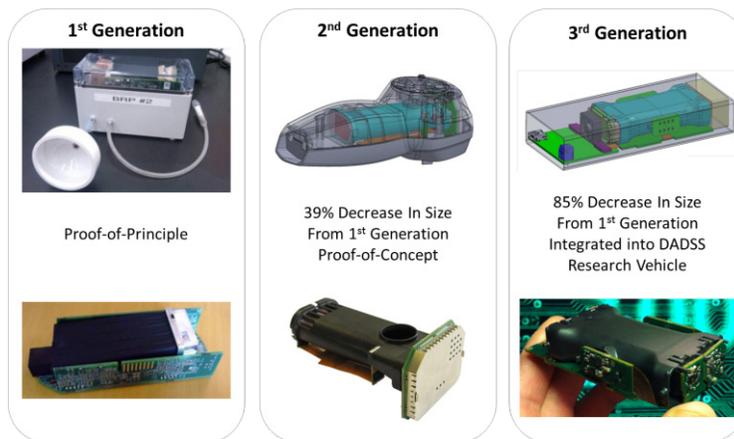


Figure 6. Evolution of Breath-based DADSS Sensor

TECHNOLOGY & MANUFACTURING READINESS LEVELS

To manage, measure, and assess the progress and maturity of both technologies during the research and development phase, Technology Readiness Levels (TRL) and Manufacturing Readiness Levels (MRL) will be used throughout the program. These Readiness levels provide common terms to define technology from concept to commercial production, and have a proven effectiveness from the aerospace and defense sectors. Independently, readiness levels can also assist with self-assessment, monitoring progress and planning goals and actions. The advantages of using readiness levels are:

- provide a common understanding of technology and manufacturing readiness status
- risk management
- make decisions concerning technology funding
- make decisions concerning transition of both technologies to the automotive industry

The readiness levels used for the program are based on the “Automotive Technology and Manufacturing Readiness Levels, A guide to recognized stages of development within the Automotive Industry” by Professor Richard Parry-Jones CBE, Co-Chairman of the Automotive Council. These levels were revised and updated by the DADSS Technical Working Group (TWG) to incorporate DADSS specific milestones to achieve demonstrated commercial feasibility as shown in Figure 5.

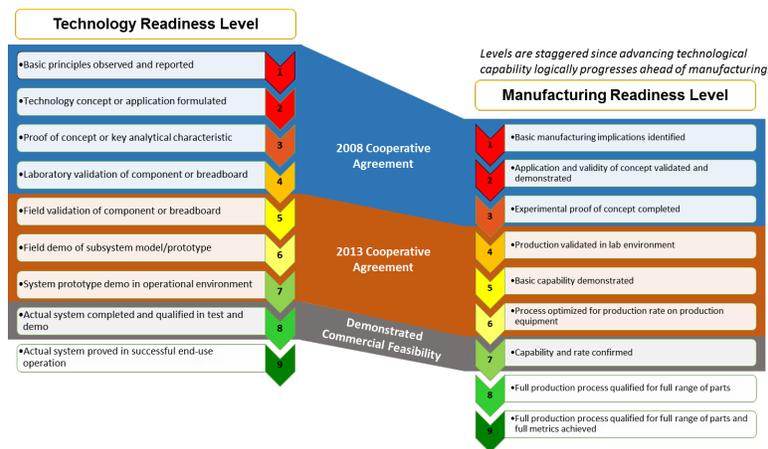


Figure 7. TRL/MRL Demonstrated Commercial Feasibility

Table 2 summarizes a preliminary evaluation of the “readiness” of the breath-based and touch-based technologies at the end of Phase II. As Figure 7 indicates, at the end of the 2008 cooperative agreement, the breath based technology achieved the expected TRL 4 level. The touch-based technology on the other hand was behind both TRL and MRL expectations. However, a number of technological challenges are ahead for the breath-based system relating to sampling in a vehicle cabin with the windows open and the air conditioning or heater on, which are not expected to be challenges that the touch-based system will need to surmount. Furthermore, the touch-based system readiness levels are anticipated to increase rapidly once development and integration of the full set of required multi-lasers packages is complete.

Table 2. Technology and Manufacturing Readiness Levels by Technology Type

Technology	TRL	MRL
Breath-based	4	4
Touch-based	3	2

STANDARD CALIBRATION DEVICE (SCD) DEVELOPMENT

Standard Calibration Devices (SCD) were developed to assess and document the accuracy and precision of the Phase I prototypes. Two different SCDs were developed for prototype testing; one breath-based and one touch-

based. There are two aspects that were addressed (Ferguson et. al., 2010). First, samples of simulated “breath” and “tissue” were developed to provide a calibrated (known) and consistent ethanol concentration in vapor and/or liquid to the prototype. These samples also had to provide reasonable facsimiles of human breath and tissue. As noted above, the DADSS Performance Specifications for accuracy (SE) and precision (SD) are significantly more stringent than current evidential calibration instruments, thus the sample sources of breath and tissue had to exceed the DADSS specifications by an order of magnitude. The second requirement necessitated the development of delivery methods so that the targeted samples could be effectively delivered to the prototypes.

Tissue Spectrometry SCD

An SCD sample that simulates human tissue must produce a consistent ethanol response from the sensor at all concentrations of BAC, mimic the average optical scattering properties of human tissue over the target NIR wavelength range, and maintain the test material at normal human skin temperature (34 °C). Figure 8 compares NIR reflectance of human versus simulated tissue and demonstrates the high level of concordance at the relevant wavenumbers. The system also must support varying concentrations of ethanol over the target BAC test range of 0.02% through 0.12 % BAC.

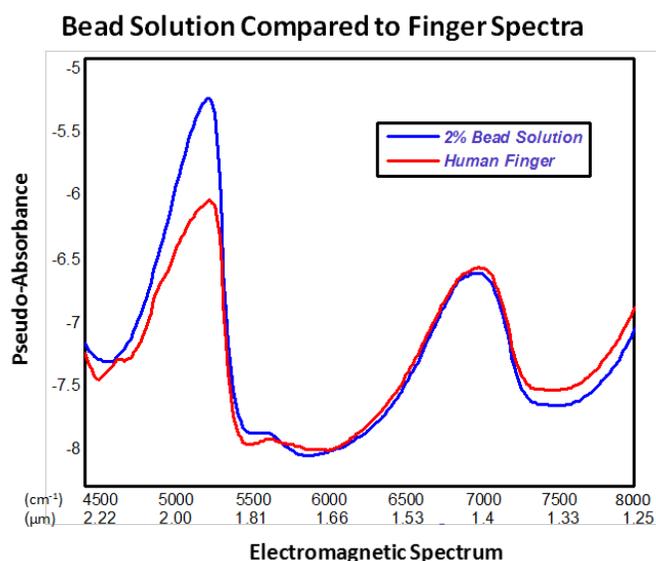


Figure 8. Comparison of NIR reflectance of simulated tissue solution with human tissue

Working with TruTouch Technologies, an SCD system was developed that comprised standardized aqueous test samples representative of human tissue and an electromechanical fluidic system for introducing the samples to the sensor. The standardized aqueous test samples are gravimetrically prepared solutions that use mono-dispersive polystyrene microspheres as an optical scattering agent. Quantities of ethanol in the solutions are certified by GC analysis to meet the required concentration levels after the beads are added. The simulated tissue solutions were stored in individual 15 mL vials. In addition to water and alcohol the “tissue” samples contain normal components of human blood such as urea, salt, and creatinine, as well as albumin that simulates blood density, microspheres that simulate the reflectance and scattering properties of collagen, and Triton that prevents the beads from clumping. An automated pipette delivery system made by Hamilton eliminates human error and operator variability, as well as improves accuracy and precision during gravimetric preparation of the solutions. The system consists of the Nimbus independent two channel work station, the pedestals and deckware custom containers, and a third party Mettler Toledo scale to accommodate 10, 20, 40, 100mL vials (Figure 9).

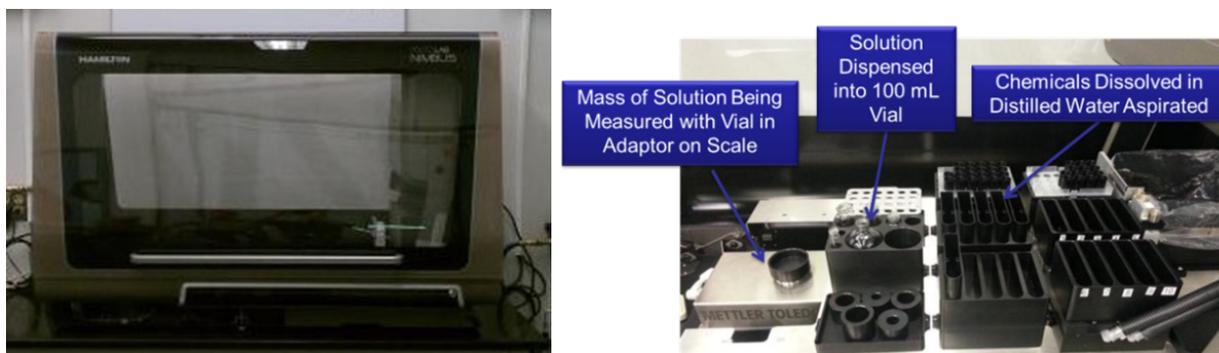


Figure 9. Hamilton Nimbus Automated pipette delivery system

The fluidic delivery system module was designed to easily attach to the TruTouch prototype sensor touch pad. The system module creates a liquid seal interface to support direct coupling between the optical sensor and the SCD test sample. The system also includes an agitation mechanism to prevent settling of the microspheres without introducing bubbles into the sample at the optical surface. The sealed system prevents evaporation loss, allows for sample removal, cleaning, and drying between sample measurements to prevent cross-contamination, and provides a reasonable degree of automation to avoid operator error. The prototype fluidics system is illustrated in Figure 9.



Figure 10. Liquid coupling interface and prototype delivery system

Distant Spectrometry SCD

The first step in the development of highly accurate breath samples was the production of standardized calibration dry gases (Ferguson et. al., 2010). Then the next step was to develop the DADSS dry gas mixture with the potential to exceed the DADSS Performance Specifications.

Two ethanol gas mixtures in 110 L pressurized bottles were developed in cooperation with ILMO Products Company:

1. Ethanol/Nitrogen (N₂)
2. Ethanol/N₂/5 % CO₂/16 % oxygen (O₂)

Each mixture was gravimetrically prepared at concentrations of 0.02, 0.04, 0.06, 0.08, 0.12 % BrAC. The mixtures were certified at ± 0.5 ppm (± 0.0002 % BrAC) by the vendor, exceeding the 0.0003 % BAC SE and SD when tested at 0.08 % BAC. In-house GC testing confirmed that the gas mixtures provided the levels of accuracy and precision for ethanol and other gases to the DADSS specifications over the complete range of gas concentrations. Additional testing verified acceptable shelf-life stability of the gas bottles.

Having validated that the dry gas mixtures complied with DADSS specifications, the next step was to humidify the gases to simulate human breath. Tests were conducted using a spirometer on a healthy male subject to measure the average flow rate and time of an exhaled breath. The ACTS team then developed a Wet Gas Breath Alcohol Simulator (WGBAS), shown in Figure 11, to add the necessary humidity.

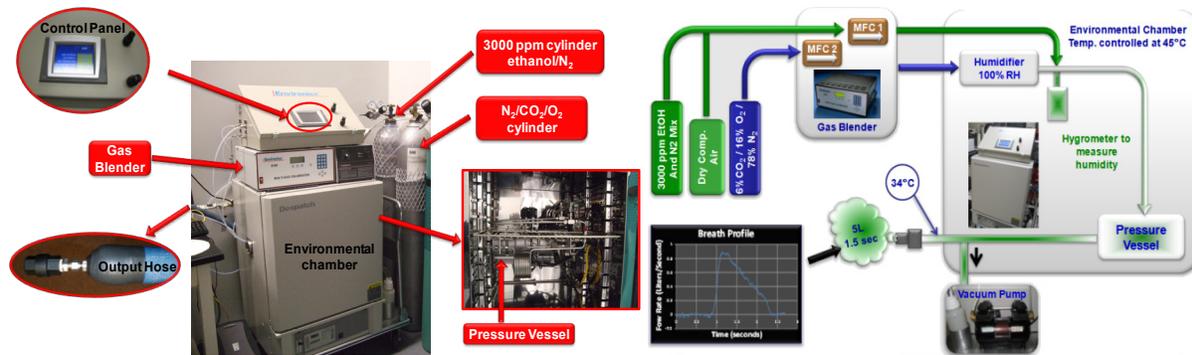


Figure 11. WGBAS Configuration and principal of operation

The WGBAS uses two dry gas sources: the first a mixture of N₂/CO₂/O₂, and the second a 3000 ppm cylinder of ethanol, balanced with nitrogen. Mass flow controllers (MFC) generate the range of humidified ethanol from 0.02 % BAC to 0.12 % BAC. The gas in the premixed cylinder of N₂/CO₂/O₂ enters the gas mixing module, flowing through humidifier metering valves located in an enclosure on top of the heated chamber. The proportional control of these valves allows the humidity to be adjusted. The ethanol/ N₂ mixture flows out of the second MFC and into the bypass line that flows around the humidifier. The humidified N₂/CO₂/O₂ mixture and the ethanol mixture meet before entering the hygrometer, which reports the dew point, humidity and gas temperature values, allowing for any necessary adjustments to obtain the required output of the humidifier. The humidified gas mixture then passes into the evacuated pressure vessel where it accumulates to a preset pressure, as monitored by an absolute pressure transmitter. When the preset pressure is met, the pressurized gas is expelled into the evacuated output tubing. As the gas leaves the system it is cooled to 34 °C, the dew point temperature of the mixture, at a rate of approximately 1 liter in 2-3 seconds, thus simulating a humidified gas flow of breath. Figure 11 shows the WGBAS principle of operation described above.

The SCD dry gas, when passed directly through the WGBAS, was capable not only of meeting but also exceeding the DADSS SEs and SDs. In the second set of verification tests, humidity was added to the mixed gases with an output dew point of 34 °C. The addition of humidity resulted in much larger SE values than the DADSS specifications and the SD values were influenced by differences in the ethanol concentration, with only the lowest ethanol concentration being able to meet and exceed the specifications.

The WGBAS was not used in the Phase I evaluation process due to its current early development status. The system will undergo additional enhancements in Phase II to improve accuracy and precision through the introduction of a closed-loop feedback system to control the amount of ethanol concentration mixed into the gas stream. Therefore, the system is planned to be used for prototypes evaluation in Phase II.

Verification Process

An SCD qualification process was developed to document that the breath and tissue sample performance meet the requisite performance specifications. Initially, components of the breath and tissue SCD were measured with a Gas Chromatograph (GC) using a Flame Ionization Detector (FID) to verify that the critical SEs and SDs were achieved. The primary function of a GC is to separate and detect chemicals in a gas flow passing through a thin column lined with specific coatings that interact with the components in the flow. The FID ignites the gases flowing out of the column with hydrogen gas. The detector then generates an electrical signal corresponding to the amount of ionized products from the combustion. The area under the curve of the electrical signal is integrated to correspond to the concentration of the gases at the column exit. Temperature, flow rate, and column selection influence the retention time of gas flow components.

The tubing and connections to the GC and the mass flow controller are heated to 34 °C, the exit temperature of human breath, to ensure the incoming gases are uniform and to assure consistent results. Several variables were found to affect and improve the GC ethanol measurement, including:

- Operating at low temperature vs. high temperature;
- Obtaining a homogeneous system, with uniform equilibrium temperature for each part of the system;
- Passivating (ability to treat a surface (typically metal) so that it is less reactive chemically) the sample line and regulator;

- Reducing the surface area of the sample line;
- Creating a constant backpressure on the actuator valve exhaust line;
- Producing a purging process for the regulator and sample loop.

The influence of the variables were quantified and examined before an optimal operating condition was obtained for the ethanol gas measurement process. Once the optimal operating conditions were identified, the dry and wet gases were then measured using the improved system shown in Figure 12.

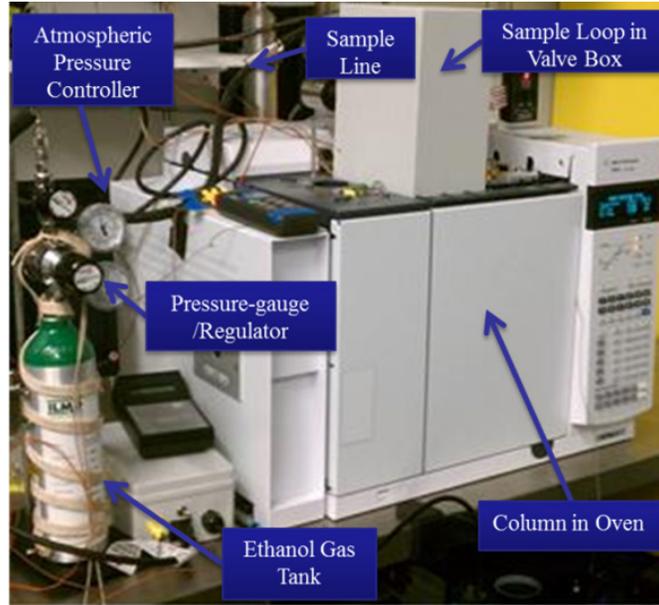


Figure 12. GC system used to measure ethanol gas

Figure 13 shows the dry gas ethanol gas measurement that is inaccurate and imprecise (left) compared with a measurement, using the developed measurement process, that shows accurate and precise measurements within the DADSS specification (right). Figure 14 shows the WGBAS measurements within the DADSS tolerance specification.

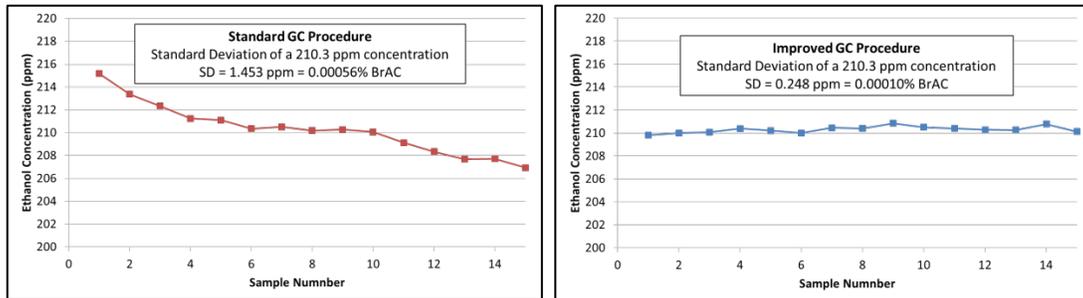


Figure 13. Standard and improved dry gas measurements comparison

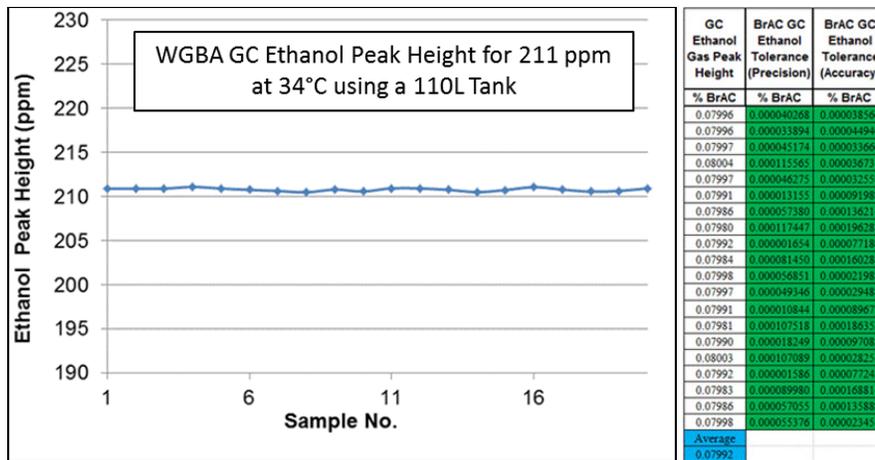


Figure 14. WGBAS accuracy and precision using the improved GC system

CONCLUSIONS

Significant progress has been made to identify DADSS technologies that have the potential to be used on a more widespread basis in passenger vehicles. Two specific approaches have been chosen for further investigation; tissue spectrometry, or touch-based, and distant/offset spectrometry, or breath-based sensors. Proof-of-principle prototype DADSS sensors have been developed, one designed to remotely measure alcohol concentration in drivers' breath from the ambient air in the vehicle cabin, and the other is designed to measure alcohol in the driver's finger tissue through placement of a finger on the sensor.

Progress also has been made to develop calibration devices for both breath-and touch-based bench testing that will be able to measure whether the DADSS devices can meet the stringent criteria for accuracy and precision. Unique standard calibration devices have been developed for both the breath- and touch-based systems that go well beyond current alcohol-testing specifications.

In summary, the DADSS Program so far has accomplished the goals set at the onset of the program. Prototype testing has indicated that there are potential technologies that ultimately could function non-invasively in a vehicle environment to measure a driver's BAC. Furthermore, the DADSS Program is on track to develop research vehicles to demonstrate the technologies by the end of 2015.

REFERENCES

- Anderson, R.R., Parrish, J.A. 1981. The Optics of Human Skin. The Journal of Investigative Dermatology, 77, 13-19.
- Ferguson, S.A., Traube, E, Zaouk, A, Strassburger, R. 2009. Driver Alcohol Detection System For Safety (DADSS) – A non-regulatory approach in the development and deployment of vehicle safety technology to reduce alcohol-impaired driving. Paper Number 09-0464. Proceedings of the 21st International Technical Conference on the Enhanced Safety of Vehicles.
- Ferguson, S.A., Zaouk, A., Strohl, C. 2010. Driver Alcohol Detection System for Safety (DADSS). Background and Rationale for Technology Approaches. Society for Automotive Engineers Technical Paper, Paper No. 2010-01-1580, Warrendale, PA: Society of Automotive Engineers.
- Ferguson, S.A.; Traube, E; Zaouk, A; Strassburger, R. 2011. Driver Alcohol Detection System For Safety (DADSS) – Phase I Prototype Testing And Finding. Paper Number 11-0230. Proceedings of the 22nd International Technical Conference on the Enhanced Safety of Vehicles.
- Hök, B., Pettersson, H., Andersson, G. 2006. Contactless measurement of breath alcohol. Proceedings of the Micro Structure Workshop. May, 2006, Västerås, Sweden.
- Jones, A.W. 2008 Biochemical and Physiological Research on the Disposition and Fate of Ethanol in the Body. Garriott's Medicological Aspects of Alcohol (Fifth Edition), 47.

Lindberg, L., Brauer, S, Wollmer, P., Goldberg, L., Jones, A.W., Olsson, S.G. 2007. Breath alcohol concentration determined with a new analyzer using free exhalation predicts almost precisely the arterial blood concentration. *Forensic Science International*. 168: 200-207.

Opdam, J.J., Smolders, J.F. 1986 Alveolar sampling and fast kinetics of tetrachloroethene in man. I Alveolar sampling. *British Journal of Industrial Medicine*, 43, 814-824.

Ridder, T.D., Hendee, S.P., Brown, C.D. 2005. Noninvasive alcohol testing using diffuse reflectance near-infrared spectroscopy. *Applied Spectroscopy*, 59, 181-189.

INTRODUCTION OF A SOLID STATE, NON-INVASIVE HUMAN TOUCH BASED ALCOHOL SENSOR

Leonard Cech
Mallik Nagolu
Dennis Rumps
TK Holdings, Inc.
U.S.

Ben Ver Steeg
Derek Treese
Bentley Laaksonen
Stephanie Tehseldar
Trent Ridder
TruTouch Technologies, Inc.
U.S.

Paper Number 15-0380

ABSTRACT

This paper presents an overview of the theory and implementation of a touch-based optical sensor (TruTouch sensor) for monitoring the alcohol concentration in the driver of a vehicle. This novel sensor is intended to improve driver safety by providing a non-intrusive means of notifying a driver when their blood alcohol concentration may be too high to operate a vehicle safely. The optical alcohol detection system has successfully completed several stages of development and validation. A commercially available, industrial version of the system (TruTouch 2500, or Mark 1) has undergone extensive clinical testing and field validation. Under the DADSS (Driver Alcohol Detection System for Safety) Program, a compact semiconductor version (Mark 2) of the optical system has been developed targeting use in consumer vehicles. Based on proven semiconductor laser technologies, the Mark 2 sensor system has demonstrated excellent spectral accuracy and precision and is currently undergoing laboratory validation testing. A demonstration vehicle version of the system has been designed and will be implemented following completion of the laboratory validation testing.

INTRODUCTION

The negative societal impact of alcohol (ethanol) impaired vehicle driving has been established through numerous clinical studies [1] and confirmed by accident statistics over many years [2]. Currently, the primary means of mitigating alcohol impaired driving is through education and legal enforcement. The percentage of alcohol in the blood circulatory system at any given time can be directly correlated with the neurological sensory, cognitive and reactive performance of the driver. Legal limits for alcohol concentration have been established for drivers of both private and commercial vehicles [3]. Today, the use of alcohol detection technology for driver safety is limited to law enforcement testing (post-accident, during traffic stops / checkpoints) or in the case of previously convicted DUI offenders through the installation of a breath based vehicle interlocks. These systems work well for the purpose of law enforcement but are unsuitable for routine use by consumers. The goal of the current sensor development is to produce a system that is seamlessly integrated into the vehicle's infrastructure, providing consumers with the knowledge of their alcohol concentration without imposing inconvenience to their daily driving experience. In order to accomplish this, a human machine interface (HMI) design has been proposed that incorporates the optical sensor into the vehicle start button. Advances in automotive buttons and touchscreens make it feasible to integrate skin based sensors and achieve appropriate visual, audio, and haptic feedback without compromising function [4]. During routine vehicle operation, the driver's alcohol concentration could be measured and communicated to the driver so that appropriate choices can be made (e.g. delay drive, alternate driver or alternate transportation method). The use of existing vehicle occupancy sensors combined with advanced signal processing supports a simple and practical anti-spoofing method, improving the safety and efficacy of the system.

SENSOR THEORY OF OPERATION

The TruTouch alcohol measurement technology has been validated using multiple approaches including *in vitro* (test tube) studies of multi-component samples, clinical studies involving alcohol dosing of humans, and real-world measurements by customers in a variety of challenging environments. The validation efforts are exemplified by multiple peer-reviewed journal articles and a strong intellectual property base.

Scientific Basis of the TruTouch Measurement

The TruTouch technology employs near-infrared (NIR) absorption spectroscopy to measure skin tissue. The NIR spectral region typically spans the portion of the electromagnetic spectrum between the visible, which is generally considered to end at 0.7 μm , and the infrared, which begins at 2.5 μm . However, for measuring alcohol *in vivo* (in human), some portions of the NIR are more advantageous than others. The features most commonly observed in the NIR are overtones and combinations of fundamental vibrations of hydrogen bonded to carbon, nitrogen, and oxygen [5,6,7,8,9].

The absorbance spectrum of alcohol shows features over the NIR region (see figure 1). The 1.25- μm 2.5 μm region contains the 1st overtone and combination bands of the carbon-hydrogen and oxygen-hydrogen bonds. The 0.7-1.25 μm region contains higher order overtones of these bonds. Examination of Figure 1 and its inset shows that the 0.7-1.25 μm region is 400 times weaker than the signal in the longer wavelength, 1.25- 2.5 μm region.

Furthermore, the utility of the visible region (0.3 to 0.7 μm) and the 0.7-1.25 μm part of the NIR are limited by the presence of skin pigmentation (melanin) that creates large differences between people, particularly of different ethnicities. In contrast, the longer wavelength region is virtually unaffected by pigmentation [10]. As a result of the larger signal and absence of pigmentation, the TruTouch technology is designed to measure the longer wavelength (1.25-2.5 μm) region.

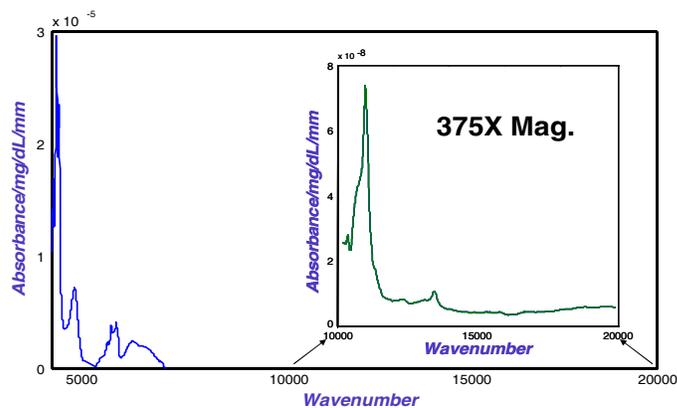
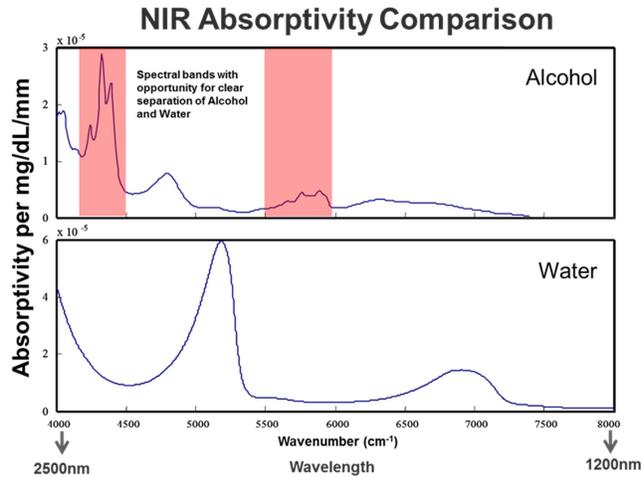
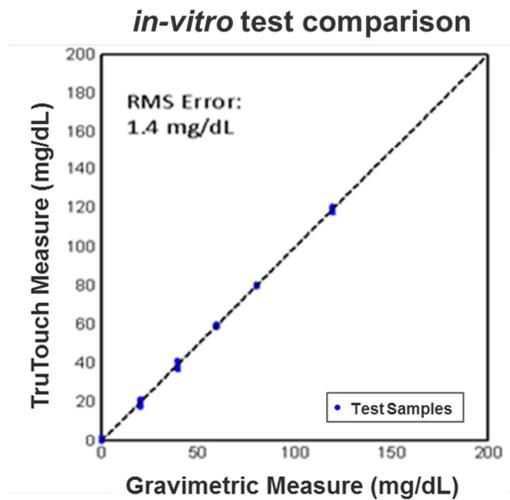


Figure 1. Absorptivity of alcohol in the NIR and visible.

In addition to the aforementioned advantages, the NIR spectral region (4000-8000 cm^{-1} or 1.25-2.5 μm) is of prime interest for non-invasive alcohol measurements because it offers specificity for a number of analytes, including alcohol and other organic molecules present in tissue, while supporting optical path lengths of several millimeters with acceptable absorbance characteristics [11,12,13,14,15]. Comparing NIR spectra (normalized to unit concentration) of alcohol and water collected using a TruTouch system, demonstrates the effect of molecular structure on NIR absorption spectra and indicates spectral regions of separation (see Figure 2a).



(a)



(b)

Figure 2. Comparison of Alcohol, Water in NIR (a); Ethanol Test Results: 98 in-vitro samples (b).

TruTouch systems (including Mark 1) are based on a Michelson interferometer Fourier Transform IR (FTIR) instrument that delivers NIR radiation to the skin and underlying tissue and collects the diffusely reflected signal using a fiber-based optical probe. The collected light contains spectral information which allows the determination of the subject’s alcohol concentration directly from the measurement. Specific details of the industrial version of the optical alcohol detection system can be found in several issued United States Patents and applications [16,17,18,19,20].

Laboratory and Clinical Validation

The objectives of any analytical measurement procedure are high sensitivity and high selectivity for the target analyte (e.g. alcohol concentration). Sensitivity refers to a method’s ability to respond to quantity changes in the target analyte, while selectivity is the extent to which a method erroneously responds to changes in interfering analytes (e.g. water, collagen, proteins, and other chemicals present in the body). Ensuring the

selectivity of an analyte measurement can be notoriously challenging in complex systems such as human tissue [21, 22]. Accordingly, careful design and controlled experiments are required to verify the validity of any measurement approach.

Historically, researchers have used *in-vitro* experiments to assess the sensitivity and selectivity of methods for quantifying analytes at physiological concentrations [23,24,25,26,27,28,29]. These experiments are useful diagnostics for the validity of a measurement approach because sample composition and the experimental conditions are controlled by the practitioner; allowing direct assessment of measurement sensitivity and selectivity. For laboratory validation of the alcohol sensor, an optically scattering tissue phantom was developed using 0.3 micron diameter polystyrene microspheres to mimic the optical properties of human skin.

To validate the Mark 1 sensor, a validation study comparing sensor measurements with tissue phantom samples containing gravimetric prepared ethanol, urea, creatinine, albumin, and saline was carried out. The study demonstrates the high degree of accuracy achievable with the touch based sensor (see figure 2b).

In Vivo Clinical Results

In order to demonstrate the accuracy of the sensor with human subjects, controlled clinical trials were conducted on the commercially available version of the system (TruTouch 2500, Mark 1). In these trials, venous blood samples were collected and sent to a certified forensic grade lab for gas chromatography analysis. Comparison data were collected on the TruTouch sensor, evidentiary grade breath sensor, and both compared against the venous blood samples. Alcohol excursions were induced in 108 subjects at Lovelace Scientific Resources (Albuquerque, NM) following overnight fasts. Written consent was obtained from each participant following explanation of the IRB-approved protocols (Quorum Review). Baseline blood and touch NIR alcohol measurements were taken upon arrival in order to verify zero initial alcohol concentration. The alcohol dose for all subjects was ingested orally with a target peak blood alcohol concentration of 120 mg/dL (0.12%). The mass of the alcohol dose was calculated for each subject using an estimate of total body water based upon gender and body mass [29]. The test results (see figure 3), indicate a strong correlation between the touch based sensor and venous blood measurements [30].

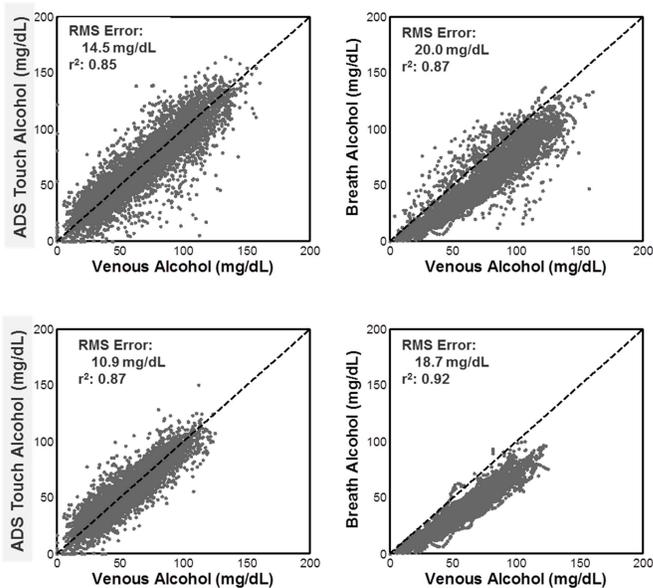
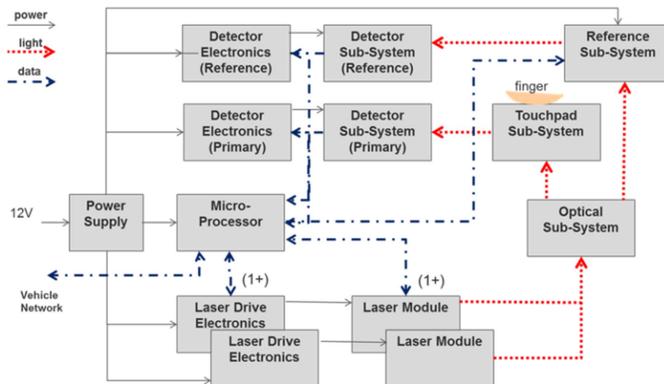


Figure 3. Human Subject (*in-vivo*) study results.

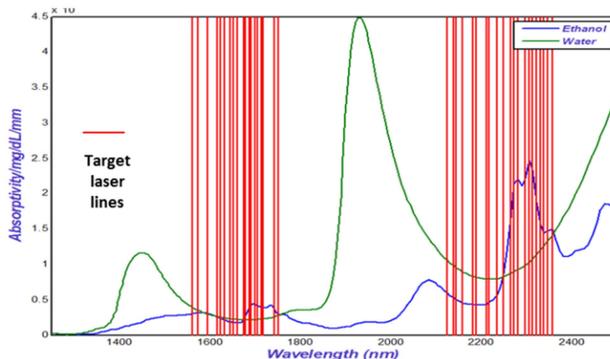
The top two plots compare the TruTouch sensor to venous blood measurements, and an evidential grade breath test to venous blood measurements, respectively. The bottom plots compare data, limited to the “elimination phase” of the alcohol excursion (e.g. after the initial rapid physiological alcohol uptake which is governed by gastric emptying and absorption of the alcohol in the small intestine during and after consumption).

SOLID STATE TOUCH SENSOR (MARK 2)

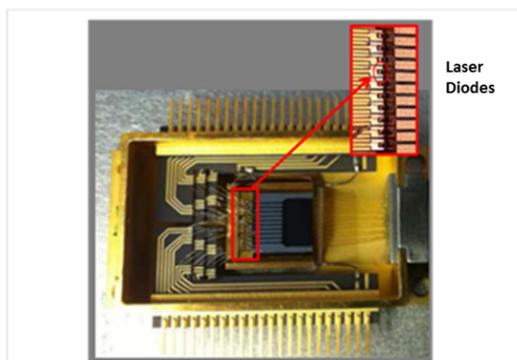
The Mark 2 sensor under development for potential application to vehicle operation uses standard automotive electronic design (see figure 4a) and discrete semiconductor lasers to encode the specific spectral information necessary for alcohol measurements.



(a)



(b)



(c)

Figure 4. Solid State Design (a), Laser line targets (b), Prototype Multi laser Module (c).

In contrast to the Mark 1 sensor, which measures a semi-continuous spectrum, the new design uses discrete, narrow-band spectral lines specifically chosen through analysis of several hundred thousand *in-vivo* alcohol tests. The laser wavelengths are targeted to spectral regions where ethanol and water absorbance levels are separable (see figure 4b), optimizing ethanol detection while avoiding the strong water absorbance features.

The use of multiple discrete lasers to interrogate spectral information allows for a highly integrated, compact optical module. For the Mark 2 design, a custom 12 laser prototype module was developed with individual controllable laser die mounted on ceramic substrates (see figure 4c). The inset detailed view shows the individual laser die mounted on the ceramic substrate. The design is compact, even at the prototype phase, and can be further integrated using packaging techniques developed for laser applications in other industries.

To show the spectral measurement accuracy of the Mark 2 system, several standard benchmark measurements have been performed and compared to both the Mark 1 system and to laboratory grade FTIR spectrometers (see figure 5). Measurements to date indicate good agreement with laboratory grade and Mark 1 system measurements. Additional testing is ongoing and planned including *in-vitro* and *in-vivo* studies similar to those previously described to verify performance that exceeds all previous FTIR systems and approaches performance targets for a vehicle based system.

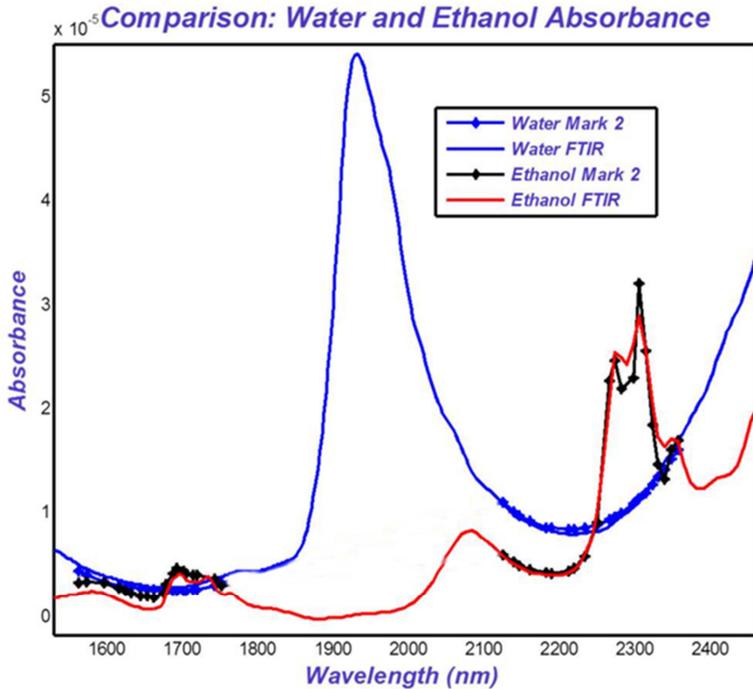


Figure 5. Mark 2 Measurement performance comparison compared to FTIR.

AUTOMOTIVE ADS SYSTEM CONSIDERATIONS

Although significant progress has been made towards establishing the feasibility of a non-invasive touch based alcohol measurement system, continued research and development is necessary to achieve a production automotive system that can meet aggressive targets for performance, measurement time, reliability and robustness. Several key considerations in the touch based design are explored further below.

Human Machine Interface A touch based sensor provides a natural opportunity to integrate the sensor touchpad into a standard starter ignition switch. Inclusion of proximity and/or touch sensors in the design supports the ability to enable alcohol measurements only when appropriate. Inclusion of a haptic actuator(s) and/or light(s) provides for natural HMI feedback (see figure 6). For example, light and haptic actuators can be used to provide driver feedback on proper finger placement, measurement initiation, measurement result and other desired HMI feedback.

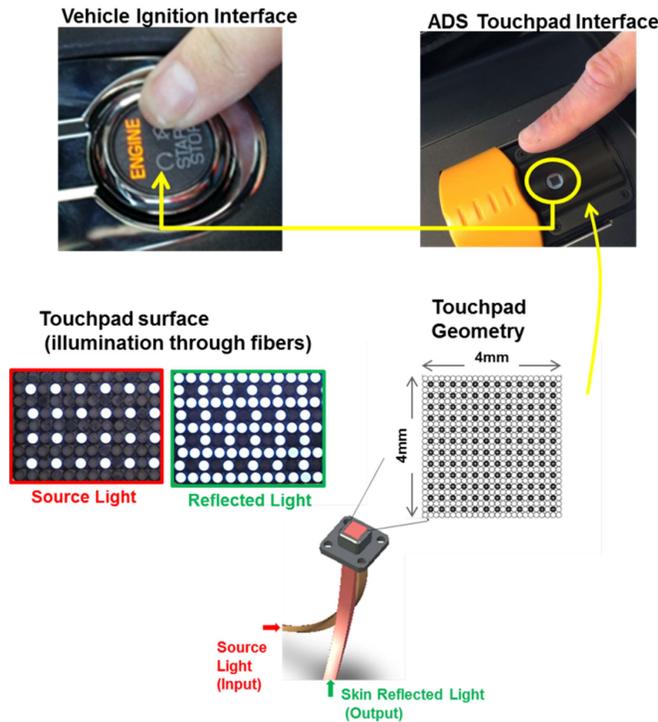


Figure 6. Prototype alcohol measurement ignition switch integration concept.

A vehicle ignition system provides one natural touch based measurement location; other viable locations throughout a vehicle exist where a driver/operator skin touch interface might be used for initial or periodic alcohol measurements.

Anti-Spoofing Concept

Based on the proposed ignition switch HMI integration, there is the potential for non-impaired occupants to try and test (start the vehicle) on behalf of another person, who intends to drive, but is impaired. To mitigate this situation, a simple anti-spoofing method can be achieved through the integration of an electric field Occupant Classification System (OCS) within the driver seat. Such systems are often used in production vehicles for the front passenger seat to satisfy a regulatory requirement to distinguish child seats from empty or full size human occupants. The information is often used to suppress airbag deployment in the case of child seats or empty seats. Electric field based OCS systems emit a controlled signal in proximity to the seat occupant. This harmless signal is influenced by the seated occupant and is transmitted through anything that the occupant touches. Because the ignition switch is touched by the driver, detection of this OCS signal can be used to distinguish the seated driver from others touching the ignition switch. The alcohol measurement control logic can be configured to require a new ignition touch (measurement) for all driver seat occupancy state changes (e.g. ingress/egress). This anti-spoofing concept could be further enhanced through signal integrity methods currently used in secure, safety automotive systems.

Environmental, Life and Ruggedness

The Mark 2 design is based on electro-optical components widely used in aerospace, defense, communications and commercial industries requiring high accuracy and precision in a rugged environment and over long operational life. Lasers and detectors of the type utilized have been verified for long term operation through military specification environmental and ruggedness testing [31, 32]. However, the Mark 2 application, introduces new technical and commercial challenges. For example, the measurement touch probe surface must be designed to operate with a wide range of finger surface chemicals and mechanical abuse such as scratching or impact. In addition, the system must be capable of reliable and accurate operation over full vehicle life, despite natural sub-system aging and drift. To mitigate these effects, the Mark 2 design incorporates an absolute chemical reference to allow for variance, bias and drift removal from measurements, over the course of the system life, in addition to providing a method to verify measurement quality and viability. Such techniques are widely used in other safety critical automotive sensors.

Commercial Challenges

The primary technical limiting factor in the development of the Mark 2 system is the fabrication and manufacturing base immaturity of the laser module and supporting optical interfaces. While advancements in the field of lasers is widespread [33,34], particularly due to new applications and markets evolving quickly, there are currently limited fully developed applications for semi-conductor lasers with the target wavelengths and optical powers targeted for the Mark 2 system.

To date, many of the target high laser wavelengths are only required in small volume specialized applications or in clinical research. On the contrary, many of the lower wavelength lasers required for the Mark 2 design are used in high volume, high reliability, and moderate to low cost packages for high bandwidth telecommunications. Lessons learned in the evolutionary technology and manufacturing growth of laser diodes can help accelerate the maturity and availability of higher wavelength diodes for use in this application. Research and development in non-invasive medical sensing is accelerating, driving new applications and markets; organically increasing manufacturing base and test investment necessary to lead to competition for low cost, high reliability automotive quality laser modules for non-invasive alcohol sensing.

CONCLUSIONS

Establishing the technical feasibility for a touch based sensor that could be used to accurately and precisely measure blood alcohol concentrations is a key initial step towards providing technical solutions to reduce alcohol impaired driving. The solid state Mark 2 system prototype provides a technically feasible architecture based on initial testing, with concepts to achieve naturalistic HMI and an anti-spoofing method. However, additional testing and design iteration are required towards a system that is capable of meeting automotive requirements.

ACKNOWLEDGMENTS

The authors would sincerely like to thank the DADSS program (and their sponsors) as well as Lovelace Clinical Research center for their continuing support on this project.

REFERENCES

- [1] Compton and Benning, DOT HS 812 117 Behavioral Safety Research February 2015 “Drug and Alcohol Crash Risk”
- [2] DOT HS 812 017 May 2014, “State Alcohol-Impaired-Driving Estimates”
- [3] Blood Alcohol Limits Worldwide, <http://www.icap.org/Table/BACLimitsWorldwide>
- [4] SAE# 2014-01-026, April 2014, Lisseman et al. “Touchscreen Concepts Revisited: Approaches and Possibilities”

- [5] Wheeler OH (1959). "Near Infrared Spectra of Organic Compounds". *Chem. Rev.* 59: pp.629-666.
- [6] Fletcher AN, Heller CA (1968). "The Alcohol Self-Association Dimer and the Absorption Band near 1.53 μm ". *Journal of Physical Chemistry* 72: pp. 1839-1841.
- [8] Iwamoto R, Shigetoshi A, Saito Y, Samura H (2001). "FT-NIR Spectroscopic Study of OH Groups in Ethylene-Vinyl Alcohol Copolymer". *Applied Spectroscopy* 55(7): pp.864-870.
- [9] Alam MK, Callis JB (1994). "Elucidation of Species in Alcohol-Water Mixes Using Near-IR Spectroscopy and Multivariate Statistics". *Analytical Chemistry* 66(14): pp.2293-2301.
- [10] RR Anderson, JA Parrish, "The Optics of Human Skin" *The Journal of Investigative Dermatology*, Vol. 77, No. 1., pp. 13-19, 1981
- [11] Wheeler OH (1959). "Near Infrared Spectra of Organic Compounds". *Chem. Rev.* 59: pp.629-666.
- [12] Fletcher AN, Heller CA (1968). "The Alcohol Self-Association Dimer and the Absorption Band near 1.53 μm ". *Journal of Physical Chemistry* 72: pp. 1839-1841.
- [13] Iwamoto R, Shigetoshi A, Saito Y, Samura H (2001). "FT-NIR Spectroscopic Study of OH Groups" in Ethylene-Vinyl Alcohol Copolymer. *Applied Spectroscopy* 55(7): pp.864-870.
- [14] Alam MK, Callis JB (1994). "Elucidation of Species in Alcohol-Water Mixes Using Near-IR Spectroscopy and Multivariate Statistics". *Analytical Chemistry* 66(14): pp.2293-2301.
- [15] Smith AW, Boord CE (1926). "Infrared Absorption in Ethers, Esters, and Related Substances". *J. Am. Chem. Soc.* 48: pp.1512-1520.
- [16] Abbink, RE et al. (2003). "System for noninvasive measurement of glucose in humans". U.S. Patent#6,574,490.
- [17] Messerschmidt RG, Abbink RE (2003). "Interferometer Spectrometer with Reduced Alignment Sensitivity." United States Patent #6,504,614.
- [18] T. Ridder, J. Maynard (2003). "Method and apparatus for optical spectroscopy incorporating a vertical cavity surface emitting laser as an interferometer reference," United States Patent #6,654,125.
- [19] C. Gardner, T. Ridder, W. Gruner (2004). "Non-invasive determination of direction and rate of change of an analyte," US Patent #7,016,713.
- [20] T. Ridder, J. Maynard, et al (2004). "Noninvasive determination of alcohol in tissue," US Patent App. #10/852415.
- [21] Brown C, and Ridder T, "A Definition of Multivariate Selectivity: Part 1: Theoretical and Practical Merits", *Applied Spectroscopy*, 59(6), pp. 787-803, 2005.
- [22] Ridder T, and Brown C, "A Definition of Multivariate Selectivity: Part 2: Experimental Applications", *Applied Spectroscopy*, 59(6), pp. 804-815, 2005.
- [22] McShane MJ, Cote G (1998). "Near-Infrared Spectroscopy for Determination of Glucose, Lactate, and Ammonia in Cell Culture Media." *Applied Spectroscopy* 52(8): pp.1073-1078.
- [23] Wabomba MJ, Small GW, Arnold MA (2003). "Evaluation of selectivity and robustness of near-infrared glucose measurements based on short-scan Fourier transform infrared interferograms." *Analytica Chimica Acta* 490: pp.325-340.
- [24] Arnold MA, Small GW, Xiang D, Qui J, Murhammer DW (2004). "Pure Component Selectivity Analysis of Multivariate Calibration Models from Near-Infrared Spectra." *Analytical Chemistry* 76(9): pp.2583-2590.
- [25] Rhiel M, Cohen MB, Murhammer DW, Arnold MA (2002). "Nondestructive Near-Infrared Spectroscopic Measurement of Multiple Analytes in Undiluted Samples of Serum-Based Cell Culture Media." *Biotechnology and Bioengineering* 77(1): pp.73-82.
- [26] Riley MR, Arnold MA, Murhammer DW (2000). "Effect of Sample Complexity on Quantification of Analytes in Aqueous Samples by Near-Infrared Spectroscopy." *Applied Spectroscopy* 52(2): pp.255-261.
- [27] Wehlburg CM, Haaland DM, Melgaard DK, Martin LE (2002). "New Hybrid Algorithm for Maintaining Multivariate Quantitative Calibrations of a Near-Infrared Spectrometer." *Applied Spectroscopy* 56(5): pp.605-614.
- [28] Melgaard DK, Haaland DM, Wehlburg CM (2002). "Concentration Residual Augmented Classical Least Squares (CRACLS): A Multivariate Calibration Method with Advantages over Partial Least Squares." *Applied Spectroscopy* 56(2): pp.615-624.
- [29] B.T. Davies and C.K. Bowen, "Peak Blood Alcohol Prediction: An Empirical Test of Computer Models," *Journal of Studies on Alcohol*, 61, 187-191 (2000).

- [30] Ridder T, Ver Steeg B, Laaksonen B, Radigan W, “Robust Calibration Transfer in Noninvasive Ethanol Measurements, Part II: Modification of Instrument Measurements by Incorporation of Expert Knowledge (MIMIK),” *Applied Spectroscopy*, 68(8), (2014)
- [31] ICSO 2010 Rhodes, Greece, International Conference on Space Optics 4 - 8 October 2010; S. Tornow et Al; “QUALIFICATION OF AN EVALUATED BUTTERFLY-PACKAGED DFB LASER DESIGNED FOR SPACE APPLICATIONS”
- [32] Proc. of SPIE Vol. 5465, Reliability of Optical Fiber Components, Devices, Systems, and Networks II, Y. Deshayes et Al; “Estimation of lifetime distributions on 1550 nm DFB laser diodes using Monte-Carlo statistic computations”
- [33] V. K. Kononenko; Stepanov Institute of Physics NASB 2010, “History and Developments of Semiconductor Lasers”
- [34] <http://www.laserfocusworld.com/articles/print/volume-51/issue-01/features/laser-marketplace-2015-lasers-surround-us-in-the-year-of-light.html>; “Laser Marketplace 2015: Lasers surround us in the Year of Light”

The RELATIONSHIP between BAC and BrAC of HEALTHY KOREAN MALE

Seung-Hwan, Yi

Sang-Ho, Lee

Department of Mechanical Eng., KNUT (Korea National University of Transportation)
Republic of Korea

Beom-Woo, Nam

Jeong-Seok, Seo

Bo-ram, Lee

Min-Choul, Ahn

Department of Psychiatry, Konkuk University Hospital
Republic of Korea

Byong-Do, Kang

Jeong-Min, In

KATRI (Korea Automotive Testing & Research Institute)
Republic of Korea

Paper Number 15-0410

ABSTRACT

BrACs (Breath alcohol concentrations) are often converted to the corresponding BACs (blood alcohol concentrations) by multiplying a partition ratio, Q . However, according to the previous researches, it has been revealed that it depended upon the nations. So, the partition ratio (or Q -factor) of healthy Korean adult males and its correlation to some variables including TBW (total body water), BMI (body mass index), BFM (body fat mass), and PBF (percentage of body fat) were revealed. The average of partition ratio did show particular difference around 100 when the subjects were divided with two sets: below and above the average of TBW. The partition ratio of Korean healthy males showed 1,913 (95 % confidence interval (CI) from 1,889 to 1,937) for whole time intervals. However, when Q was averaged after peak BACs, it gave 2,011 (95 % CI range from 1,982 to 2,040). Bland-Altman plots revealed the compatibility of measurement method of multi-gas analyzer, and the biases according to the partition ratios ($Q=2,100$ and $Q=1,913$) gave -0.0052 (95 % CI from -0.0059 to -0.0045) and -0.0004 (95 % CI from -0.0011 to $+0.0003$), respectively. From this study, the partition ratio of Korean healthy males has been reported for the first time with massive medical experiments.

INTRODUCTION

The vehicles are getting more and more important in modern life for traveling, commuting, and logistics etc.. As the numbers of automobiles increase, however, road traffic safety become a national-wide matter in order to diminish road traffic accidents and also fatalities. In Sweden, they declared Vision Zero slogan in order to eliminate any victim from the road traffic accidents [1]. Among the road traffic accidents, alcohol-related crashes and fatalities are the major issues around the world in terms of improving more safe road traffic situation. In order to alleviate alcohol-related accidents, most of nations use alcohol sensing apparatus for screening drunken drivers and it measures BrAC by using optical components or fuel-cell type devices [2, 3].

Currently, the breathalyzers for monitoring BrAC are widely used in the world in order to screen impaired drivers at the roadside. The regal limit for impairments are 0.08 % in U.S. and 0.05 % in most European countries. Although, Sweden adopted a 0.05 % BAC limit in 1950s, the regal limit of BAC has been recently lowered to 0.02 % in order to improve traffic safety further. By lowering regal limit, James and Robert reported that fatal crashes and severe personal injuries have been decreased and settled down more safe road traffic situation than ever [4]. In 1962, Korean government legitimated the road traffic law in order to enforce alcohol-impaired driving. After four decades, BrAC analysis for road traffic offences is regulated in 2006, setting a regal BAC limit of 0.05 % for driving. Also, there were several trials to decrease the regal limit of BAC to 0.03 % for road traffic safety and for decreasing alcohol-related fatalities.

Since the consumed alcohol is eliminated from the body also through the exhaled air that is coming from deep lungs [5, 6], the measurement of BrAC has been evaluated to analyze BAC for decades. The relationship between BAC and BrAC has been studied for long time, so BAC is currently converted from BrAC by multiplying by a BAC/BrAC ratio known as partition ratio or conversion factor, Q . The US NHTSA (National Highway Traffic Safety Administration) uses a conversion factor, Q , as 2,100 [7], this value was also adopted for converting BrAC to BAC in Korea. However, Jones and Andersson reported that most countries adopt a conversion factor of 2,000 to 2,300 [8]. Furthermore, the recent study showed that the conversion factor could be ranged from 2,225 to 2,650; Jones and Andersson reported the average conversion factor was 2,448 in their article [9], Pavlic et al. presented the time dependency of Q ranged from 2,225 to 2,650 [10], and Lindberg et al. showed that the BAC/BrAC ratio was 2,251 in case of Swedish subjects [1].

Since the conversion factor, Q , is relevant to BAC determined from breathalyzer and also it is relatively different from country to country, the aim of this study is to identify the BAC/BrAC ratio of Korean healthy males and its correlation to some variables: TBW, BMI, BFM, and PBF in this research.

MATERIALS AND METHODS

Subjects

One hundred and one individuals, whole healthy males, were enrolled in this study as paid volunteers. Ages ranged from 20 to 50 years and body weights ranged from 55 to 78 kg. Whole subjects were not heavy consumers of alcoholic beverages (less than two bottles of Soju (20 % (v/v) in their regular lifestyle). The study was approved by the Ethics committee of Konkuk University Hospital, Korea.

Experiment procedures

The volunteers were recruited by posting announcement through internet or public board. The volunteers were screened by psychiatric doctor with basic medical and psychiatric checkup containing physical examination, vital sign, CBC (complete blood count), LFT (liver function test), vital sign check, Alcohol Use Disorder Identification Test-Korea (AUDIT-K), CAGE (cut-down, annoyed, guilty, eye-opener) survey. Finally, 101 individuals were selected and participated in this research. Whole participants checked up their current physical status by measuring weight, height, TBW, BFM, PBF, and BMI, which were analyzed by InBody 720 (InBody Co., LTD.). Since it roughly took 2 minutes to analyze one breath sample, the volunteers were divided into ten groups. The volunteers belonged to each sub-group were randomly selected and each subject was asked to consume 0.35 mg/ml/kg or 0.7 mg/ml/kg Soju (30 % (v/v)) for 15 minutes. After consuming alcohol, they are allowed to rinse the mouth with drinking water in order to ensure the absence of mouth alcohol before the first testing, however, not allowed to drink water until 2 hours after consuming whole alcohol. The samples (one blood and two breath) were obtained at timed intervals of 15, 30, 45, 60, 90, 120, 180, 240 minutes after drinking alcohol in order to reveal the relationship between BAC and BrAC for healthy Korean males. The BAC/BrAC ratio was analyzed according to the four body index (TBW, BFM, PBF, and BMI). In order to avoid complexity in subsequent data interpretation, no food and mixing of different alcohol were allowed, furthermore, violent physical activity was not allowed during the experiment also.

Collection of blood and BAC analysis

After drinking of alcohol within 15 minutes, a blood sample of 2 mL was drawn from the proximal stopcock which is connected to the indwelling catheter at each timed interval as mentioned earlier and injected into a 3 mL Vacutainer tube (BD Franklin, Lake NJ, USA), containing EDTA (Ethylene-diamine tetra-acetic acid). The tubes were stored in a refrigerator at 4 to 6 °C and brought to Neodin medical Institute located in Seoul the day after the experiments were finished for each sub-group. Each delivered blood sample was analyzed by enzymatic methods (COBAS Integra 800, Roche USA) twice times in order to reveal BAC.

BrAC Measurements

One breath sample for each volunteer was collected with 3 liter non-odor bag (TK005-N-003, BMS Corp., Japan) at the same time the blood sample drawn from the indwelling catheter and analyzed with INNOVA-1312 multi-gas analyzer (LumaSense Technologies, Denmark). The analyzer used in this study consists of two main components:

optical and acoustical measurement units. By adopting photo-acoustic measuring principles, it could analyze gases from ppb to ppm level. So, this analyzer has been selected as a reference BrAC measurement apparatus in this study. However, in order to enhance the measurement reliability, the multi-gas analyzer was sent back to the manufacturer for adding a new filter (for ethanol measurement) and was calibrated in order to secure the accuracy of measurement. After measuring each breath sample three times with multi-gas analyzer, the average BAC has been converted ppm level to percentage level by multiplying conversion factors. The other breath samples for each person were measured by using four portable breathalyzers (AL9000, Sentech Corp., Korea) at timed intervals of 15, 30, 45, 60, 90, 120, 180, 240 minutes after drinking alcohol. Then four-measured values were averaged after finishing tests for comparison. Each portable breathalyzer was also calibrated before the experiments to alleviate the reliability issues raised in fuel-cell type breathalyzer.

Calculation of Q

The partition ratio, Q, was individually calculated for each subject from the ratio of the mean BAC value to the average BrAC value determined from the multi-gas analyzer in this study. However, the ratio of BAC to BrAC value, Q, assumed to be 2,100 in portable breathalyzer because Korean jurisdiction admitted this value currently. All statistical parameters such as average, standard deviations of average and 95 % confidence intervals for the calculated parameters were acquired with MS Excel 2013 and also Sigma Plot 12.5.

RESULTS and DISCUSSION

After arranging BAC results according to the elapsed times, Korean adult males show three characteristic alcohol metabolism patterns as shown in Figure 1: left-shifted, standard, and right-shifted patterns (five mixed patterns are excluded in this analysis). The numbers of subjects belonged to each category denote as n in Figure 1. Compared to Figure 1 b), left-shifted pattern reveals no peak of alcohol concentration in their blood, however, right-shifted (also standard) type presents a peak alcohol concentration after 90 mins later in this subject. As can be inferred from Figure 1, the alcohol metabolism of Korean adult males could be divided into three characteristic patterns.

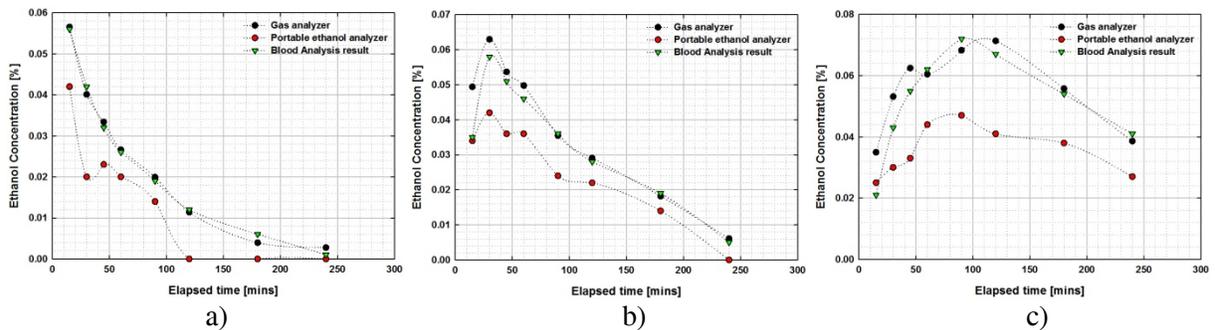


Figure 1. Three characteristic patterns of BAC: a) left-shifted (n=29), b) standard (n=52), c) right-shifted (n=15).

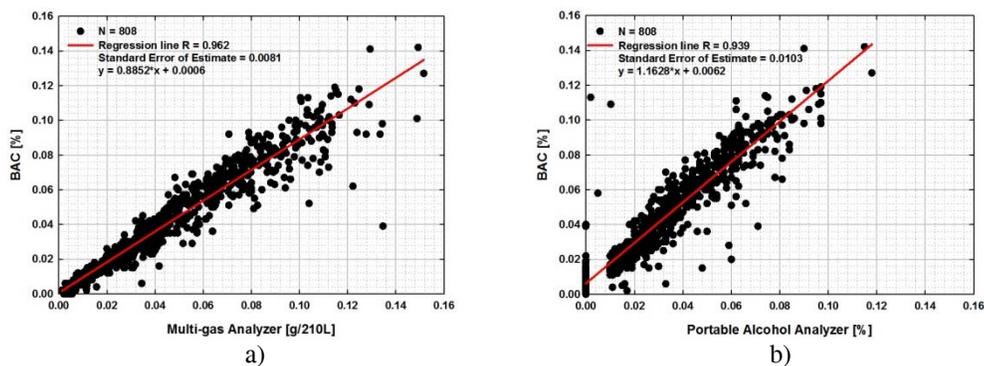


Figure 2. Relationship between BAC and BrAC with two different BrAC measurements.

Figure 2 shows the relationship between BACs and BrACs measured by two different apparatus: multi-gas analyzer and portable breathalyzer with an assumed partition ratio, $Q=2,100$ ($N=808$, numbers of whole measured values). In case of multi-gas analyzer, BACs are overestimated as depicted in Figure 1 a); regression line $y=0.8852x+0.0006$ with $R=0.962$. However, when breath alcohol concentrations are measured with portable breathalyzer, BACs were underestimated as can be seen in Figure 1 b); regression line $y=1.1628x+0.0062$, with $R=0.939$. Even though there was small bias that is roughly 10 % of the error (± 0.005 %) with the measurement of multi-gas analyzer, the data measured by breathalyzer showed little higher offset value in BAC measurement. Furthermore, even BACs had some meaningful values, the breathalyzer didn't monitor the actual BACs in some subjects. Since there were some differences between BAC and BrAC measurements in both cases when the partition ratio, Q , was used as 2100, the partition ratio was calculated according to the elapsed time. Low BACs which are less than 0.01 % are excluded in this analysis in order to increase the accuracy of partition ratio. Also, the average value of T_{max} , which means the average time that BAC reaches the highest value after consuming alcohol in this study, was 55 mins [11], so the partition ratios were calculated with this time reference and showed as in Figure 3. During the absorption period (which was less than 60 mins after consuming alcohol), average of Q was 1,779. However, after 60 mins (this time intervals belonged to the digestion of alcohol), the average partition ratio was 2,011. Furthermore, when whole data were calculated without the time limitation, the average partition ratio of healthy Korean males was calculated as 1,913.

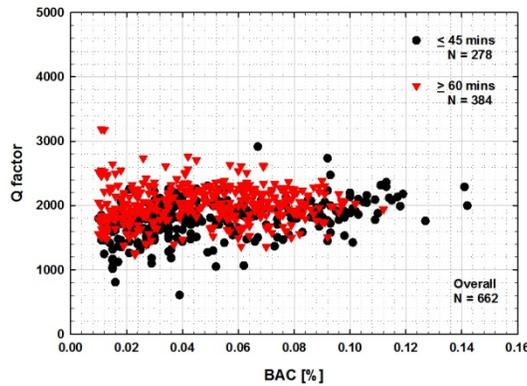


Figure 3. Partition ratio as a function of BAC

When total body waters were above and below the average value (42.4 Liters), the average of Q -factor was 1,903 (95% C.I. 1,870 to 1,938) and 1,999 (95% C.I. 1,966 to 2,033), respectively. In terms of BMI, the average of BMI was 25 for healthy Korean male. When BMI was above the average value, the partition ratio showed 1,935 (95% C.I. 1,887 to 1,983). When the partition ratios were categorized into body fat mass (BFM) and percentage body fat (PBF), the averages of partition ratio presented 1,950 (95 % C.I. 1,916 to 1,983) and 1,957 (95 % C.I. 1,924 to 1,990) when BFM and PBF are above the average values, respectively. The data related to partition ratio according to body index were listed in Table 1.

Table 1. Partition ratio according to body index (excluded BACs less than 30 mins)

Categories		BMI		TBW		BFM		PBF	
		Above Ave.	Below Ave.						
Average values		1,935	1,951	1,903	1,999	1,950	1,941	1,957	1,934
95 % C.I.	Lower Limit	1,887	1,924	1,870	1,966	1,916	1,905	1,924	1,898
	Upper Limit	1,983	1,979	1,938	2,033	1,983	1,976	1,990	1,979
Standard Deviation		315	284	308	267	285	304	284	305

The partition ratio ($Q=2,011$ after 60 mins later) would be an important factor in order to calculate the estimated BAC when the extrapolated BAC is needed to evaluate initial BAC value [12] by police. Also the average partition ratio, after drinking alcohol without time limit, would be essential to the manufacturer of BAIDs (breath alcohol ignition interlock devices) since the user of BAIDs would be requested to pass the rolling retest during driving [13, 14]. So, it would be valuable to know the relationship between BACs and BrACs as a parameter of partition ratio and showed their relationship in Figure 5. As can be seen in Figure 5, when

Q=2,100 was adopted to calculate the BACs, the actual BACs would be overestimated by the exhaled breath samples. The regression line showed $y = 0.8615x + 0.0023$ with $R = 0.944$ in this case (Q=2,100). It would be certain that the measurement results can be not favorable to Korean healthy males. However, when the partition ratio is less than 2,100, as denoted in Q=1,913 (depicted in red dots), BrACs product by Q showed more favorable matches that the previous results with regression line, $y = 0.9457x + 0.0023$, with $R = 0.944$.

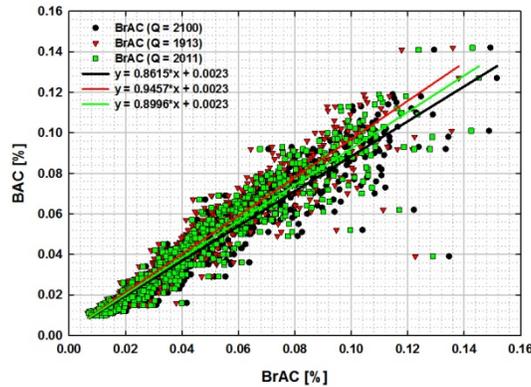


Figure 5. BAC vs. BrAC according to the partition ratio, Q.

In order to assure the capability of replacement of invasive BAC measurement, Bland-Altman plots according to the partition ratios (Q=2,100 and Q=1,913) have been presented in Figure 6. If there is no or small bias in Bland-Altman plot, it is well known that a new experimental apparatus or device could be replaceable in medical checkup etc. As described in Figure 6 a), when the partition ratio, Q, was used as 2,100, there was a bias with -0.0052 and showed its 95 % CI from -0.0059 to -0.0045. Also the limits of agreement ranged from -0.0242 to 0.0138. Even though it is not shown in this article, when Q=2,011 is adopted, the bias showed -0.0029 and its 95 % CI marked from -0.0036 to -0.0022. However, when Q=1,913 was multiplied to BrACs in order to calculate BACs, the bias showed -0.0004 and revealed its 95 % CI from -0.0011 to +0.0003. Furthermore, the limits of agreement of evaluation ranged from -0.0179 to +0.0171. Since the limit of agreement had offset toward negative value in case of Q=2,100, the measured values from multi-gas analyzer would clearly exaggerate BACs as mentioned in Figure 5. From the results mentioned above in Figure 5 and Figure 6, it is clear that the average value of partition ratio calculated with the whole time interval BAC/BrAC ratio will be more reasonable than the value of Q (=2,100) used in current breathalyzer for Korean adult males.

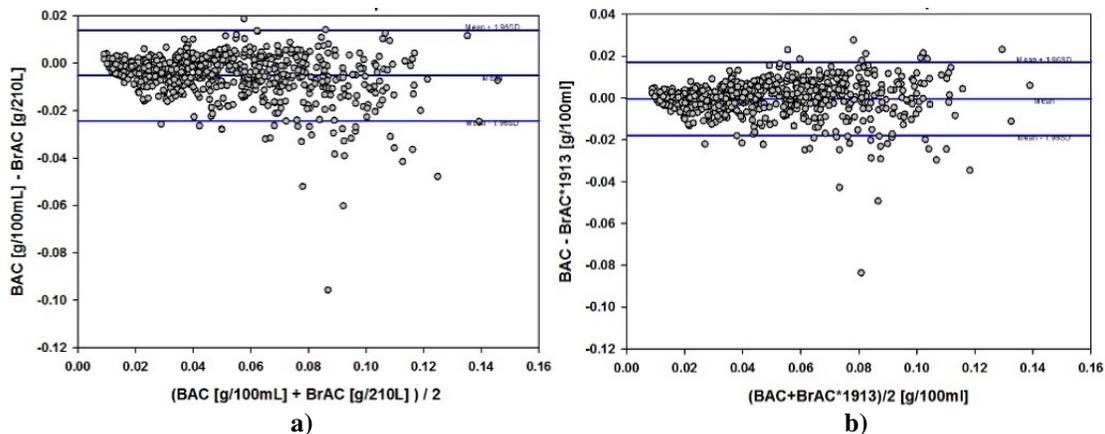


Figure 6. Bland-Altman plots according to the partition ratio, Q: a) Q=2,100, b) Q=1,913.

CONCLUSIONS

In order to reveal the relationship between BAC and BrAC of healthy Korean males, the huge medical experiments has been executed for the first time in Korean medical study. Korean healthy males showed three characteristic BAC patterns, however, more than 50 % subjects (52 out of 101 individuals) participated in this

study showed a standard pattern with BAC peak around 50 mins. Even though there was no big differences in partition ratio in terms of BMI, BFM, and PBF, however, there was a meaningful gap in partition ratio when TBW is divided into two categories: above and below the average value. Since the partition ratio differed from the conventional value ($Q=2,100$) in case of Korean healthy males, it would be better to consider a new value ($Q=1,913$) or execute more profound research activities in order to calculate BACs from BrACs for Korean healthy males. The more reasonable partition ratio included Korean females will be reported shortly in the near future.

ACKNOWLEDGEMENTS

This research is supported by a grant from Ministry of Land, Infrastructure and Transportation (Grant number: 13TLRP-C067560-05-000000).

REFERENCES

- [1] Swedish Transport Administration, 2012. "Analysis of road safety trends 2012: management by objective for road safety work, towards the 2020 interim targets", December 2013.
- [2] Lindberg, L., Brauer, S., Wollmer, P., Goldberg, L., Jones, A.W., Olsson, S.G., 2007. "Breath alcohol concentration determined with a new analyzer using free exhalation predicts almost precisely the arterial blood alcohol concentration.", *Forensic Science International*, 168, 200-207.
- [3] Edna, S., David, S., 2011. "An analysis of alcohol breath tests results with portable and desktop breath testers as surrogates of blood alcohol levels.", *Accident Analysis and Prevention*, 43, 2188-2194.
- [4] James, C.F., Robert, B.V., 2006. "The effectiveness of reducing illegal blood alcohol concentration (BAC) limits for driving: Evidence for lowering the limit to 0.05 BAC", *J. Safety Research.*, 37, 233-243.
- [5] Marion, P., Petra, G., Kathrin, L., and Walter, R., 2007. "Elimination rates of breath alcohol." *Forensic Science International*, 171, 16-21.
- [6] Hlastala, M.P., 1998. "The alcohol breath test-a review.", *American Physiological Society*, 401-408.
- [7] Federal Register, 1993. "Highway safety programs; Model specifications for devices to measure breath alcohol.", *Federal Register/Vol.58, No. 179, pp.48705-48710/Friday, September 17, 1993*
- [8] Jones, A.W., Andersson, L., 1996. "Variability of the blood/breath alcohol ratio in drinking drivers." *J. Forensic Science*, 41, 916-921.
- [9] Jones, A.W., Andersson, L., 2003. "Comparison of ethanol concentrations in venous blood and end-expired breath during a controlled drinking study.", *Forensic Science International*, 132, 18-25.
- [10] Pavlic, M., Grubwieser, P., Brandstatter, A., Libiseller, K., Rabl, W., 2006. "A study concerning the blood/breath alcohol conversion factor Q: concentration dependency and its applicability in daily routine." *Forensic Science International*, 158, 149-156.
- [11] Nam, B.W., Yi, S.H., Seo, J.S., Lee, B.R. 2014. "Alcohol pharmacokinetics in Korean healthy male", *J. European College of Neuropsychopharmacology*, 24, supplement 2, S661.
- [12] Altman, D.G., Bland, J.M. 1983. "Measurement in medicine: the analysis of method comparison studies", *Statistician*, 32, 307-317.
- [13] Federal Register, "Model specifications for breath alcohol ignition interlock devices (BAIIDs)", vol.57, no.67, Tuesday, April 7, 1992.
- [14] National Research Council Canada, "Technical standard for vehicular breath alcohol interlock devices in Canada", Project 54-A3948, August 9, 2011.

UNOBTRUSIVE BREATH ALCOHOL SENSING SYSTEM

Bertil Hök

Hök Instrument AB
Sweden

Håkan Pettersson

Autoliv Development AB
Sweden

Jonas Ljungblad

Hök Instrument AB
Sweden

Paper Number 15-0458

ABSTRACT

Although the vast majority of vehicle drivers are sober, drunk driving remains to be a major contributor to fatal accidents. Massive deployment of unobtrusive breath alcohol sensing systems could potentially save tens of thousands of lives worldwide every year by preventing drunk driving [1]. The work reported here is ultimately aiming at such a system. The technical performance of the present sensing system with respect to automotive requirements is summarized, and new results towards unobtrusive breath alcohol determination within vehicle compartments are presented.

Breath alcohol concentration (BrAC) can be determined unobtrusively if (i) the sensing system provides real-time signals with adequate accuracy corresponding to the local concentrations of both alcohol and a tracer gas, e.g. CO₂, (ii) the dilution of the breath is not excessive in relation to background concentrations, (iii) the sensor location can be seamlessly integrated into the interior of a vehicle cabin. All three of these aspects are addressed in the present paper.

More than a hundred prototypes based on infrared spectroscopy were fabricated and subjected to automotive qualification tests in the full temperature range -40 ... +85°C. In the majority of tests, adequate performance was noted. Measures are now being taken to fill remaining performance gaps. Test results with human subjects were positive and in accordance with expectations with respect to physiological variations. In-vehicle tests showed that for the best sensor position, passive breath samples allowed BrAC to be determined at a resolution of 2-4% of the US legal limit, providing proof-of-principle for unobtrusive testing. Nevertheless, vehicle integration remains to be the major technological challenge to the objective of deployment on a large scale of unobtrusive driver breath alcohol determination.

The feasibility of unobtrusive breath alcohol determination in vehicles, and adequate performance of a sensor system based on infrared spectroscopy have been experimentally demonstrated. The alcohol sensing system may advantageously be integrated into vehicles, and may also be combined with other technologies to monitor driver impairment.

INTRODUCTION

Although the vast majority of vehicle drivers are sober, drunk driving remains to be a major contributor to fatal accidents. Many informative and persuasive initiatives have been undertaken. Devices for the determination of breath alcohol concentration (BrAC) are commercially available for screening and evidential purposes, and alcohol interlocks are being increasingly used [2]. However, according to the driver alcohol detection system for safety (DADSS) initiative [3, 4], there is a need for radical improvement in order to make such devices acceptable on a

larger scale. The technology needs to be unobtrusive to the sober driver, and it should determine whether the driver's blood alcohol concentration (BAC) is above or below the legal limit with high accuracy. Deployment of such a technology on a large scale could potentially save tens of thousands of lives every year by preventing drunk driving.

Our research towards less obtrusive sensor systems for BrAC determination started in 2005. The envisioned system will unobtrusively and accurately detect alcohol in the driver's breath before the vehicle may be started, or while driving. In earlier publications, we have demonstrated methods and system solutions for contactless determination of BrAC [5, 6] in screening applications where sobriety is expected to be the norm. The physiological rationale of using a tracer gas, e.g. carbon dioxide (CO₂), for contactless determination was examined [7], and the usefulness of this technique in patients with reduced consciousness was demonstrated [8]. Recently, further progress towards unobtrusive and highly accurate BrAC determination in automotive applications has been demonstrated [9, 10, 11].

In this paper, an updated review of the methods and technology for unobtrusive and highly accurate breath alcohol determination is provided. New experimental results are presented on the technical performance of the sensing system with respect to automotive requirements. Results from human tests and in-vehicle unobtrusive testing are summarized and discussed in view of the overall objectives.

METHODS AND TECHNOLOGY

Basic system function

Breath alcohol concentration (BrAC) can be determined unobtrusively if (i) the sensing system provides real-time signals with adequate accuracy corresponding to the local concentrations of both alcohol and a tracer gas, e.g. CO₂, (ii) the dilution of the breath is not excessive in relation to background concentrations of both alcohol and the tracer gas, (iii) the sensor location can be seamlessly integrated into the interior of a vehicle cabin without undue influence from passengers or other sources of interference. These three aspects will be addressed throughout the paper.

The requirements of unobtrusiveness and high accuracy are seemingly contradictory. A key to resolving this contradiction is to introduce a two-step procedure, in which the first unobtrusive step is providing a preliminary result whether the driver's BrAC is above, say half the legal limit, or not. If below, the vehicle immediately becomes drivable ('green'). If BrAC is much higher than the legal limit, the drivability will be locked ('red'). A sober driver, and one with BrAC clearly above the legal limit, will thus perceive the system to be unobtrusive. If the unobtrusive BrAC reading is in the 'yellow zone' in between, the driver will be offered the possibility of providing an active breath test in order to determine BrAC with high accuracy.

Figure 1 schematically shows a typical time sequence starting by automatically switching on the sensor system when the car is unlocked. The sensor is staying in a standby mode until the door to the driver's seat is first opened and then closed. This is the point when the sensor is activated. The occurrence of a CO₂ peak is used as an indicator of a breath above the background level. If a corresponding peak of ethyl alcohol (EtOH) is detected at basically the same point in time, it is possible to estimate BrAC using the following equation

$$\text{BrAC} = \text{EtOH}_{\text{meas}} * \text{DF} = \text{EtOH}_{\text{meas}} * (\text{CO}_{2\text{et}} - \text{CO}_{2\text{background}}) / (\text{CO}_{2\text{meas}} - \text{CO}_{2\text{background}}) \quad (1).$$

The subscript 'meas' denotes the measured peak values, and 'CO_{2et}' the end tidal CO₂ concentration, which is believed to approach the alveolar concentration, typically 4.8±0.5 vol% [7, 12]. DF is the dilution factor, ranging from one in highly concentrated air close to the mouth of the subject, to large numbers at a large distance. The background CO₂ concentration is typically less than 0.1 vol%. The standard measurement unit for BrAC is mg/L, which relates to blood alcohol concentration (BAC, %) by the approximate relation 1 mg/L BrAC = 0.2%BAC [13]. The US legal limit of 0.08%BAC thus corresponds to a BrAC value of 0.4 mg/L.

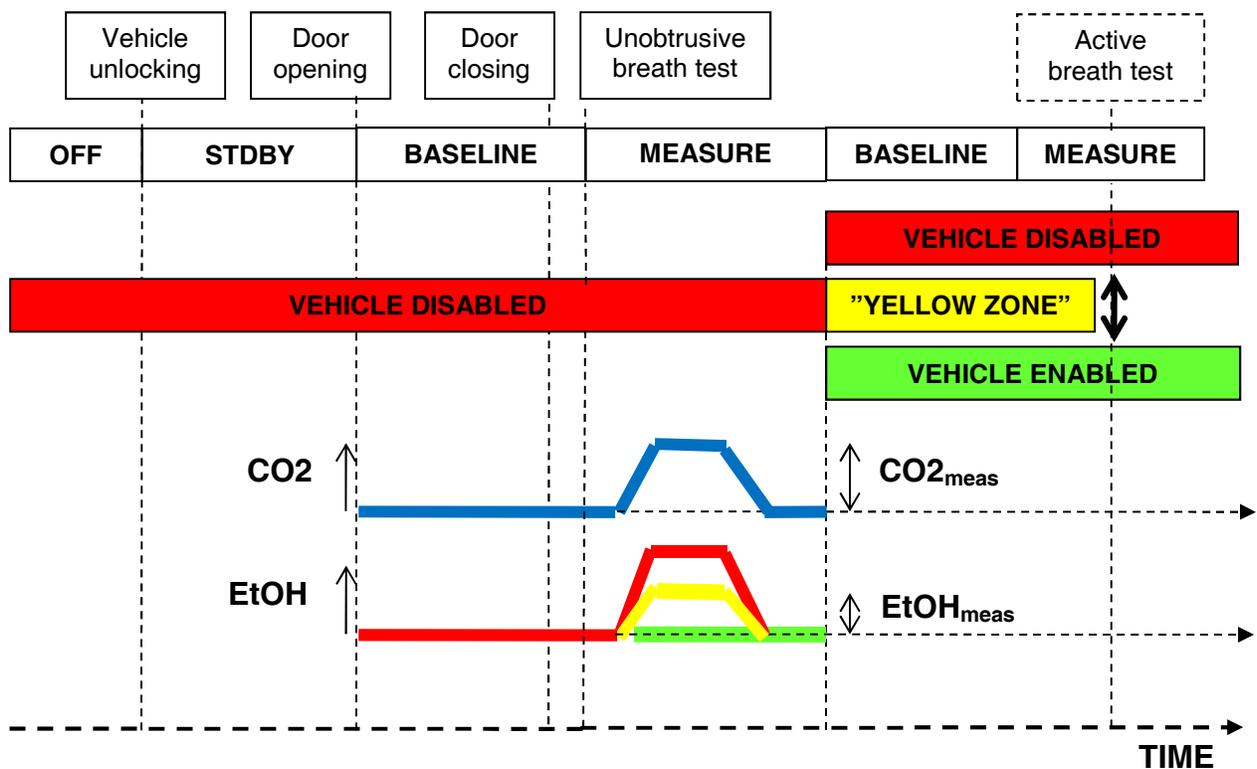


Figure 1. Schematic time sequence of unobtrusive breath test.

The active test to be performed when the unobtrusive test results in a “yellow zone” BrAC value, is expected to distinguish with a high accuracy whether or not the legal limit is exceeded. Then it is necessary to provide an undiluted breath sample ($DF=1$ in equation (1), independent of $CO_{2\text{meas}}$). It should be noted that continuous or intermittent monitoring using basically the same scheme is also possible during driving. Accumulating data over time adds to the accuracy of the system.

System implementation

The sensor system includes the following critical parts:

- Air inlet defining the sampling point at which air is continuously being withdrawn, and fed to:
- A measuring cell including optical and sensing elements for real-time infrared transmission measurement for the selective detection of CO_2 and EtOH, respectively
- Signal processor for digitizing the sensor signals into a standard, calibrated format corresponding to local gas concentrations
- Auxiliary sensor elements to distinguish between a true breath and possible interference
- Main processor performing algorithms for breath recognition and BrAC determination, including eq. (1).

Three prototype generations of the system have been implemented so far. The 3rd generation devices were miniaturized compared to generation 2, with approximate dimensions 120 x 40 x 20 mm, packaged for handheld use

vehicle integration. Figure 2 shows photographs of the unpackaged device, a handheld implementation, and a possible future integration of the device into the A-pillar of a vehicle.



Figure 2 Photographs of an unpackaged sensor (left), a handheld device (middle), and a device integrated into the A-pillar of a vehicle (right).

With a handheld device operated at 3-5 cm distance, the dilution factor DF will be in the range 1.5-2.5. For a less obtrusive breath at 15-20 cm distance DF is typically 5-10.

Experimental tests

Extensive tests have been performed on the device and system levels, with experimental settings involving both artificial and human objects. The device level included automotive qualification tests, and tests on human subjects. More details of these tests are provided in the Results section.

In-vehicle system tests were performed in order to provide an understanding of the critical boundary conditions relating to unobtrusive breath alcohol determination. The breath flow is expected to be influenced by other flow sources, including ventilation, air conditioning, passengers, and obstacles within a vehicle compartment. The in-vehicle tests included theoretical simulations using finite element methods, and experimental visualization of breathing pattern using a phantom and water mist as an optical contrast medium. A third method was to position sensor prototypes at various locations within a vehicle compartment, and to record and analyze the sensor signals upon entrance of a human subject into the driver's seat.

RESULTS

This section will provide a summary of results from automotive qualification tests, tests on human subjects, and in-vehicle tests.

Automotive qualification test results

When possible, the tests were performed according to industrial standards. However, in several cases more stringent specifications were adopted [14, 15] compared to requirements according to current industrial standards. This was especially the case when the requirements related to measurement accuracy and startup time were examined in view of unobtrusive and highly accurate BrAC determination. More than a hundred complete devices of generation 2 and 3 have been fabricated and tested.

The test results are summarized in table 1, including columns of the test types, relevant limit values, standards, and test result. In total 18 test types were included, all of which primarily relate to the device performance. The majority of tests were performed on generation 2 devices. Results from generation 3 are underway and will be added in due course.

Table1.
Results of automotive qualification tests.

Test	Limit	Standard	Test result
Unit-unit calibration error EtOH	±5%	[14, 15]	Pass
Unit-unit calibration error CO ₂	±5%	[14, 15]	Pass
Resolution	2µg/L	[14, 15]	Pass
Linearity	±2%	[14, 15]	Pass
Startup time at room temp	5 sec	[14, 15]	-
Startup time at -40°C	20 sec	[14, 15]	-
Power consumption	70W peak, 8W cont.	[14, 15]	Pass
Function test 0°C ... +85°C	±0.03 mg/L	[14, 15]	Pass
Function test -40°C	±0.03 mg/L	[14, 15]	-
Cross sensitivity	Acetone, ...	EN50436-1,2	Pass
Barometric pressure	0.8 ... 1.1 bar	EN50436-1,2	Pass
Manipulation, circumvention	-	EN50436-1,2	Pass
Vibration test	-	ISO16750	Pass
Mechanical shock	-	ISO16750	Pass
Accelerated aging	Corr. to 15 yrs of use	[14, 15]	Pass
Corrosive atmosphere	NO _x , SO _x ...	[14, 15]	Pass
EMC	200V/m immunity	[14, 15]	Pass
Application-like long term test	-	[14, 15]	Pass

Table 1 summarizes the fact that the results met or exceeded the requirements in the majority of tests. There is still a gap between actual and required performance at extremely low temperature, and some improvement is required for the startup time.

Human subjects study

The human subject tests were motivated by the fact that the proposed technique represents a new method in need of experimental evidence. The results summarized here have recently been published in more detail elsewhere [10].

Thirty adult volunteers with an age distribution from 19 to 70 years were enrolled for the test and provided their informed consent to participation. The study was approved by the Swedish Ethical Review Board in Uppsala (dnr 2013-089). Each subject was instructed to consume alcohol with a target intoxication level of 0.06 to 0.10 %BAC (BrAC 0.3 to 0.5 mg/L) within 15 minutes. The dosage was decided using body weight as the main parameter. During the elimination phase, the subjects performed breath tests every 20 minutes, using generation 2 devices both in a contacting mode or operation with a mouthpiece, and without a mouthpiece at a distance varying from 3 to approximately 15 cm. On each of these occasions a reference BrAC value was obtained with an evidential breath analyzer, Evidenzer (Nanopuls AB, Uppsala, Sweden). A total number of 1,465 breath tests were performed.

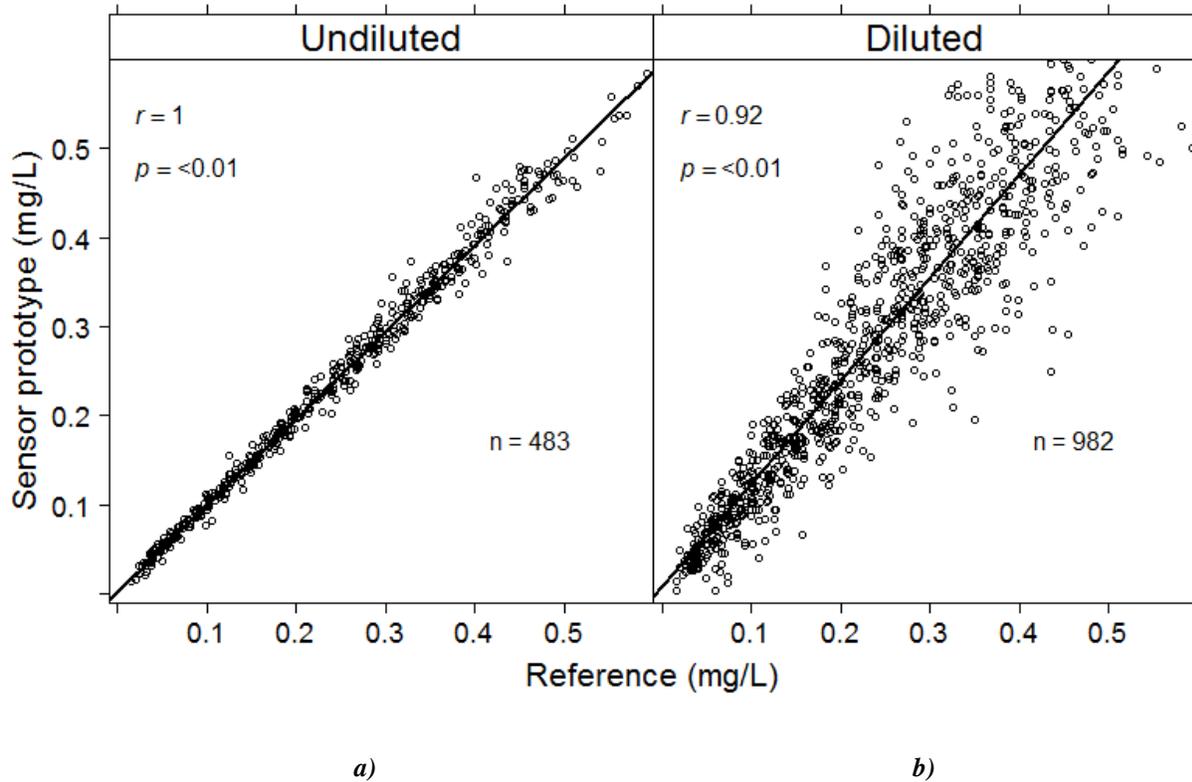


Figure 3 Results of human subjects test of thirty volunteers. The graph a) represents active tests with a mouthpiece and b) tests performed at 3-15 cm distance. Both graphs show measurement results with generation 2 devices (y-axis) compared to an evidential breath analyzer (x-axis).

The results of the human subject study are summarized in Figure 3. The correlation between active tests (undiluted, $DF=1$, graph 4 a)) is excellent with a correlation coefficient of 1.00, providing experimental support of the technical performance summarized in the previous subsection.

As shown in Figure 3 b), the human tests performed at 3-15 cm distance exhibited much larger variations than the undiluted breath tests. In this graph, eq. (1) was used for calculation of BrAC. A striking feature is that the distribution is shifted upwards from the identity line. This can be understood from a systematic deviation between the end tidal value and the alveolar CO_2 concentration [7, 11, 12]. The distribution has a funnel shape, increasing with concentration. This is a direct result of eq. (1) and the dominating variability of CO_{2et} . Not shown in the presented graphs is that the sensor distance did not influence the distribution, despite the large variation from 3 to 15 cm [11].

In-vehicle test results

A theoretical model was designed for simulation of relevant phenomena relating to in-vehicle air flow dynamics using finite element methodology (ANSYS). The model represented an idealization of a real-world occupant compartment geometry in order to enable the study of basic mechanisms and phenomena at moderate requirements of processing capacity and computing time. Figure 4 a) shows a simulated breath flow being deflected by a stronger guide flow passing to the left of the driver's head. The guide flow is attracting the breath flow, thus creating a possibility to control the flow direction.

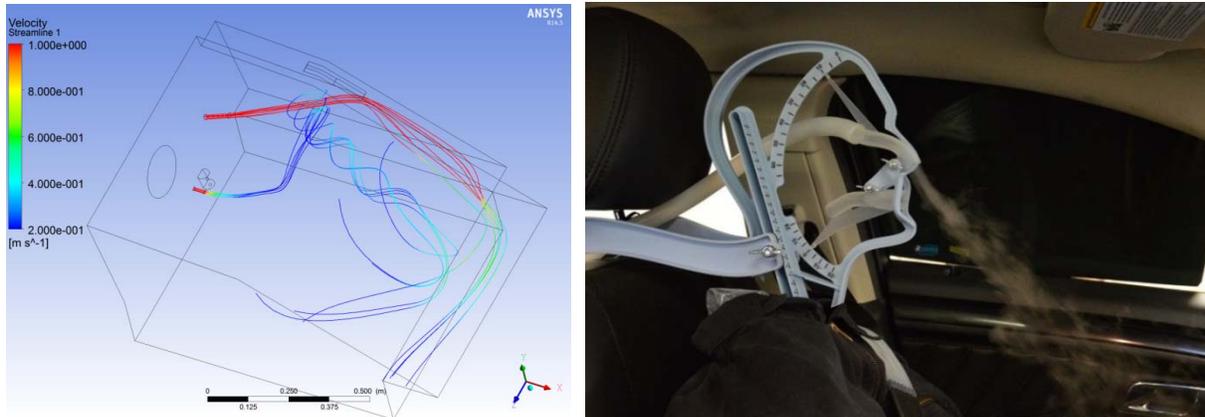


Figure 4. a) Visualization of in-vehicle breath flow using FEM simulation (left). b) Breath flow from driver phantom visualized by a mist of water droplets (right).

The experimental setup shown in Figure 4 b) was designed to enable air flow patterns to be visualized using a driver phantom, providing a realistic breathing pattern. A mist of water droplets injected into the flow was used as an optical contrast medium. Both mouth and nose breathing could be simulated using this setup.

The results of in-vehicle measurements using human subjects are summarized in Figure 5, showing graphs of measured EtOH concentration as a function of the dilution factor DF. A parabolic relationship between measured EtOH concentration and dilution is observed at given levels of BrAC, and is evident from eq. (1). Figure 5 a) also includes data of the range of dilution factors observed in a completely passive mode of operation. The in-vehicle tests were performed with volunteers instructed to control their exhalation either by nose or mouth. Data from the most favorable positions are included: Seat belt, sun shield, and steering column.

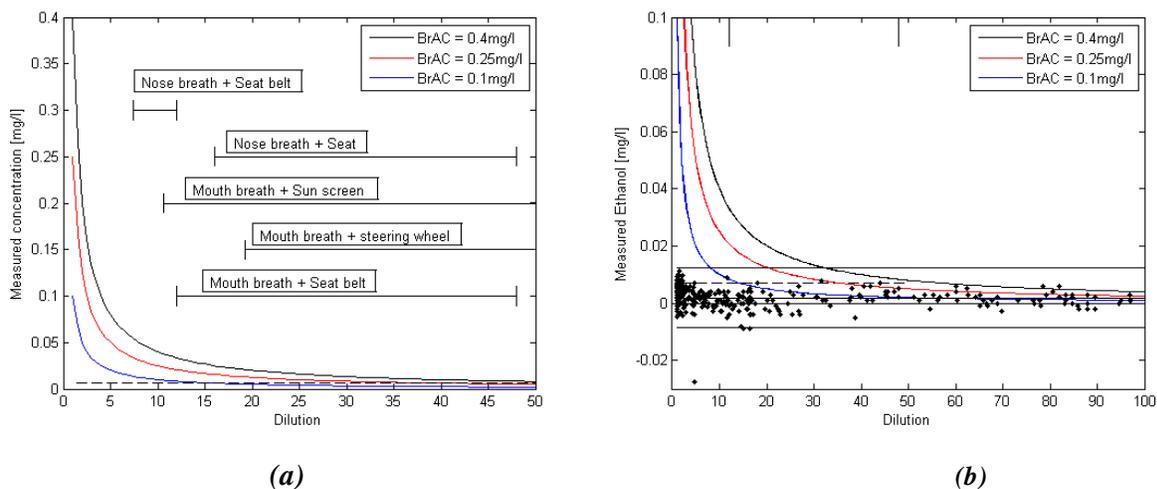


Figure 5. Measured EtOH concentration as a function of dilution for various settings.
a) Theoretical curves at different BrAC levels are superimposed with measured dilution data from various in-vehicle positions. The dashed line corresponds to the 3σ resolution.
b) Basically the same graph as a) but with different scaling, and measured data (dots, see text).

Figure 5 b) includes experimental data from the sensor positions seat belt, steering column, and side door. The 420 data points from 38 volunteers indicate that DF needs to be below 20-30 in order to obtain adequate resolution.

DISCUSSION

The automotive test results were positive for the majority of test cases conducted. Measures are now being taken to shut the remaining performance gaps concerning startup time and accuracy at extreme temperatures. This work is underway and is expected to provide overall fulfillment of the present specification [14, 15].

The system performance in human subjects is adequate in view of the suggested two-step procedure, in which the first unobtrusive step is a provisional BrAC determination (Figure 1). The observed variability of contactless measurements can be understood from the corresponding variability between individuals of alveolar CO₂ concentration [7].

The test results on unobtrusive in-vehicle determination indicated (Figure 4 a) that for the best sensor position in a vehicle setting, the seat belt position, typical dilution factors of 8-15 were observed by passive detection in several individuals, resulting in a BrAC resolution of 2-4% of the US legal limit. This observation is believed to constitute proof-of-principle for passive BrAC determination according to the scheme of Figure 1. However, the seat belt position is not considered suitable from an integration perspective. In other positions, the dilution is much larger.

Besides truly passive detection the option of a directed breath from a distance to a sensor integrated in the vehicle, as depicted in Figure 2 c), should also be considered. It may be argued that this option is also unobtrusive to the sober driver, since it only takes 1-2 seconds of the driver's attention, similar to pressing a button. This reflection notwithstanding, aspect (iii) related to vehicle integration remains as the most compelling technical challenge compared to (i) and (ii) (see subsection on basic system function) to the objective of unobtrusive breath alcohol determination.

Several initiatives are underway to integrate breath alcohol sensor systems with other technologies, including the already mentioned DADSS program [3, 4], and the Drive Me project directed towards evaluation of new technologies for autonomous vehicles [16].

CONCLUSIONS

The feasibility of unobtrusive breath alcohol determination in vehicles, and adequate performance of a sensor system based on infrared spectroscopy have been experimentally demonstrated. The alcohol sensing system may advantageously be integrated into vehicles, and may also be combined with other technologies including autonomous driving.

ACKNOWLEDGEMENTS

Financial contributions from the Automotive Coalition for Traffic Safety (ACTS), the National Highway Traffic Safety Administration, NHTSA, the Swedish Knowledge Foundation, and Vinnova, The Swedish Innovation Agency, are gratefully acknowledged. The authors also wish to express their thanks to all members of the project team at Autoliv, SenseAir, and Hök Instrument, for their excellent contributions.

REFERENCES

- [1] National Highway Traffic Safety Administration (NHTSA), *The Economic and Societal Impact of Motor Vehicle Crashes 2010*, Report DOT HS 812 013, May 2014.
- [2] Harding P, Zettl R (2008) Methods for breath analysis (Chapter 7). Garriott JC (Ed.) Garriott's Medicolegal Aspects of Alcohol (5th Ed.). Lawyers & Judges Publishing Company, Tucson, AZ. p. 229-253.
- [3] A Zaouk, Driver Alcohol Detection System for Safety, Transportation Research Board 90th Annual Meeting, Washington D. C., Jan. 25, 2011.

- [4] S Ferguson, A Zaouk, N Dalal, C Strohl, E Traube, R Strassburger, Driver Alcohol Detection System for Safety (DADSS) – Phase I Prototype Testing and Findings, Proc. 22nd Int. Conf. Enhanced Safety of Vehicles, February 2012, Paper No 11-0230.
- [5] B Hök, H Pettersson, G Andersson Contactless measurement of breath alcohol, Proc. Micro Structure Workshop, MSW 2006, May 2006, Västerås, Sweden.
- [6] B Hök, H Pettersson, Annika Kaisdotter Andersson, S Haasl, P Åkerlund, Breath Analyzer for Alcolocks and Screening Devices, IEEE Sensors Journal, vol. 10 (2010), p. 10-15.
- [7] A Kaisdotter Andersson, Improved Breath Alcohol Analysis with Use of Carbon Dioxide as the Tracer Gas, Ph. D. Thesis No. 83, Mälardalen University, Västerås-Eskilstuna, Sweden, September 2010.
- [8] A Kaisdotter Andersson, B Hök, D Rentsch, G Rücker, M Ekström, Improved Breath Alcohol Analysis in Patients with Depressed Consciousness, Med. Biol. Eng. Comput. 2011.
- [9] A Kaisdotter Andersson, B Hök, A Karlsson, H Pettersson, Unobtrusive Breath Alcohol Testing, International Conference on Alcohol, Drugs and Traffic Safety, ICADTS 2013, August 25-28, 2013, Brisbane, Australia, Paper No. 2Bi.
- [10] J Ljungblad, B Hök, M Ekström, Critical Performance of a New Breath Analyzer for Screening Applications, Intelligent Sensors, Sensor Networks and Information Processing (ISSNIP), 2014 IEEE Ninth International Conference on, Singapore, 21-24 April 2014, doi: 10.1109/ISSNIP.2014.6827626.
- [11] B Hök, J Ljungblad, A Kaisdotter Andersson, M Ekström, M Enlund, Unobtrusive and Highly Accurate Breath Alcohol Determination Enabled by Improved Methodology and Technology, J. Forensic Investigation, 2014;2(4) 8.
- [12] A B Lumb, Nunn's Applied Respiratory Physiology, 6th Ed., Elsevier, 2005, p. 157-158.
- [13] AW Jones (2010) The relationship between blood alcohol concentration (BAC) and breath alcohol concentration (BrAC): A review of the evidence. (Road Safety Web Publication No. 15). Department of Transport, London.
- [14] T Eklöv et al, Autoliv Design Goal Document for Alcohol Sensor (unpublished document, 2011).
- [15] T Eklöv et al, Autoliv Test Specification for Alcohol Sensor (unpublished document, 2012).
- [16] <http://www.slashgear.com/volvo-drive-me-project-brings-100-autonomous-cars-to-the-road-02307359/>.

MITIGATING DROWSINESS: LINKING DETECTION TO MITIGATION

Chris Schwarz

Timothy L. Brown

John Gaspar

Dawn Marshall

The University of Iowa

USA

John Lee

The University of Wisconsin

USA

Satoshi Kitazaki

The University of Iowa

USA

Julie Kang

National Highway Traffic Safety Administration

USA

Paper Number 15-0453

ABSTRACT

Drowsy driving contributes towards up to 24% of crashes and near crashes observed; 886 fatal crashes per year can be attributed to drowsy, fatigued or sleeping drivers. Drowsiness mitigation technology is composed of a detection algorithm and a mitigation component. This paper is primarily concerned with the latter, specifically for a driving simulation study about mitigating drowsy driving. The study is part of NHTSA's Driver Monitoring of Inattention and Impairment using Vehicle Equipment (DrIIVE) program. The detection algorithm incorporates time series probabilistic estimation using a Hidden Markov Model, so a drowsiness prediction at any time is dependent on a previous history of observations. Two mitigation methods are designed for testing in the simulation study. One is a three stage audio/visual alert that requires a driver response through a button press. The second is a binary haptic alert that uses a vibrating seat. Additionally, each mitigation will include three varying levels of sensitivity: a nominal model, an over-sensitive model, and an under-sensitive model. These variations will expose drivers to different numbers of false alarms while also potentially missing episodes of drowsiness. Various parameters in the detection algorithm were tested and the vote thresholds of two Random Forest models were selected for variation. It was observed how these parameters affected the output of the detection and mitigation system using previously collected drowsy driving data. Three specific levels were chosen as candidates for the experiment. It is hoped that the study will answer questions about how effective a mitigation system is at changing driving performance, whether drivers willfully ignore the mitigation, and how many alerts are too many.

INTRODUCTION

The National Highway Traffic Safety Administration (NHTSA) estimates that 83,000 crashes per year and 886 fatal crashes per year can be attributed to drowsy, fatigued, or sleeping drivers (NHTSA, 2011). The 100-car naturalistic driving study found that drowsy driving contributed to 22% to 24% of crashes and near-crashes observed (Klauer, Dingus, Neale, Sudweeks, & Ramsey, 2006). Other studies suggest that despite known dangers many drivers continue to drive drowsy and fall asleep behind the wheel (MacLean, Davies, & Thiele, 2003; McCartt, Rohrbaugh, Hammer, & Fuller, 2000). Technology may be able to address some of these risks.

Drowsiness mitigation technology consists of two subsystems, a drowsiness detection system and a driver feedback system. The drowsiness detection system or algorithm collects data from the driver or vehicle, processes this data with a detection algorithm, and makes predictions about the alertness of the driver. The feedback system activates when the detection system predicts that the driver is drowsy and alerts the driver in order to prevent a drowsiness related crash. With some exceptions, research on drowsiness mitigation technology has largely focused on the detection algorithm. This piece of the system is critical because it strongly influences drivers' trust and reliance on the mitigation technology and constrains the design space of the feedback system (Balkin, Horrey, Graeber, Czeisler, & Dinges, 2011).

Research on drowsiness detection algorithms can be differentiated by the input data, prediction algorithm, and ground truth definition of drowsiness. Input data typically consists of camera-based eye measures (Dinges & Grace, 1998; Grace et al., 1996; Ji, Zhu, & Lan, 2004), electric potential measures from the brain (Lal, Craig, Boord, Kirkup, & Nguyen, 2003; Lin et al., 2005; Wali, Murugappan, & Ahmmad, 2013), or driver input to the vehicle such as steering wheel angle (Krajewski & Sommer, 2009; McDonald, Lee, Schwarz, & Brown, 2013a; Sayed & Eskandarian, 2001). Prediction algorithms vary from simple thresholds (Dinges & Grace, 1998), to more complex graphical models (Ji et al., 2004). The ground truth definitions also vary between studies and range from general levels of drowsiness associated with lack of sleep (Sayed & Eskandarian, 2001; J. H. Yang, Tijerina, Pilutti, Coughlin, & Feron, 2009), to more episodic measures of drowsiness such as drowsiness-related lane departures (McDonald et al., 2013a). Recent research primarily focuses on innovations in the prediction algorithm dimension. One prominent development in this dimension is a transition from static prediction algorithms to time-based prediction algorithms (Ji, Lan, & Looney, 2006; G. Yang, Lin, & Bhattacharya, 2010; J. H. Yang et al., 2009). These time-based prediction algorithms allow predictions to account for well-understood temporal effects of drowsiness: for example, a drowsy driver is likely to stay drowsy and an alert driver is likely to stay alert. Additionally, they can be built around previously non-temporal (or static) algorithms to improve predictions (Ji et al., 2006, 2004). The success of these algorithms and their strong basis in the theory of drowsy driving suggests that it could be helpful to enhance other non-temporal models by incorporating them into temporal frameworks.

Mitigation systems are the critical link between the detection system and influencing driver behavior. While the detection system aims to accurately assess driver state, the aim of the mitigation system is to present driver state information to the driver in a way that is likely to persuade the driver to make choices that improve safety. This process involves the translation of the raw detection system outputs for use by the mitigation system. These systems can theoretically take many forms, from a simple audible chime or visual icon to more complex displays that relay different levels of performance or instruction to the driver. Although the same algorithm might be used across systems, the type of the interface will dictate the required adaptation of the raw data.

The topic of this paper is the design of a mitigation system to provide feedback to the driver about the system's perception of their state of drowsiness. The mitigation system should help the driver become more aware of their drowsiness. In the short term, it may help them to improve their driving performance; however, the ultimate desired effect would be to cause them to pause their trip and take a rest.

There are several drowsiness alert systems on the market currently (see Figure 1). Many are binary alerts that display a coffee cup icon and play a chime when the alert is triggered. Some systems attempt to provide a more continuous, or at least multi-level discrete, scale of drowsiness to the driver. Some systems require the driver to press a button to acknowledge the alert.

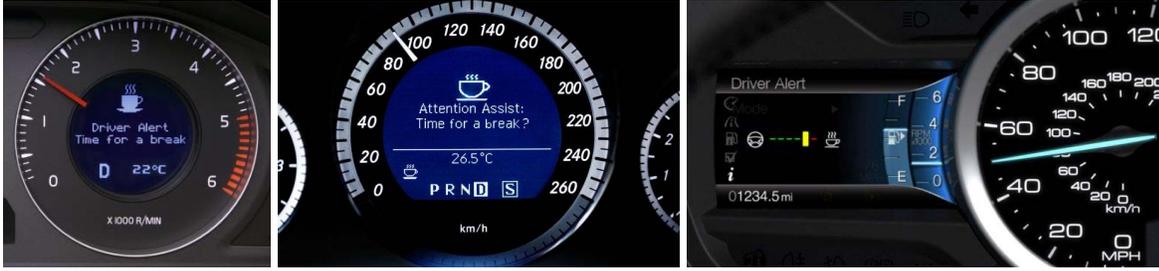


Figure 1. Example Mitigation interfaces from Volvo, Mercedes-Benz and Ford

PRIOR WORK

The NHTSA DrIIVE program focuses on the detection and mitigation of driver impairment from drowsiness and distraction. Several models were generated in phase 1 of the DrIIVE program, including a Bayesian Network, a time-to-lane-crossing (TLC) model, and a Random Forest model based on steering wheel angle (McDonald, Lee, Schwarz, & Brown, 2013b). A Random Forest model that incorporates temporal steering information into a static algorithm was trained on drowsy lane departure data (2013a), (Brown, Lee, Schwarz, Fiorentino, & McDonald, 2014).

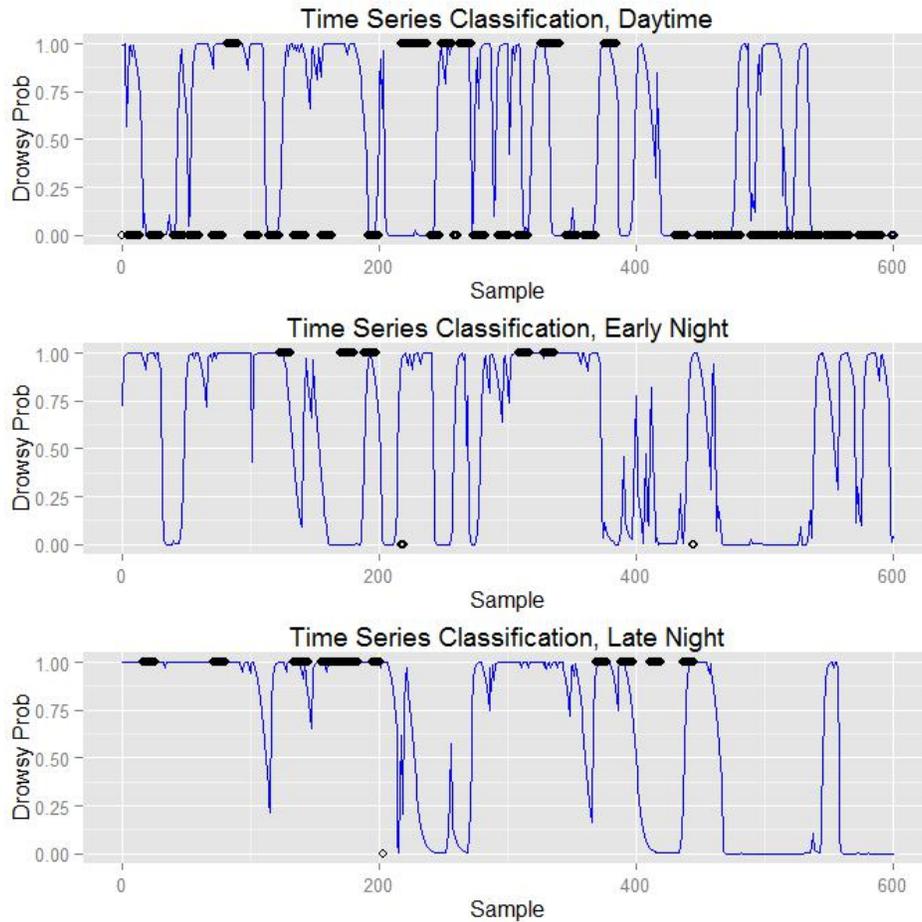


Figure 2. Example Time Series: ground truth markers and drowsiness HMM output

This initial algorithm was then extended by placing the static steering algorithm into a temporal prediction framework and exploring the effect of this approach on the timeliness of the detection algorithm (Schwarz,

McDonald, Lee, & Brown, *submitted*). The enhancements produced a set of Random Forest (RF) models that were fed into a Hidden Markov Model (HMM) capable of capturing the heuristic that an awake driver is more likely to remain awake in the near future, while a drowsy driver is likely to remain drowsy.

The Random Forest models were trained in the open source statistical software R (R Development Core Team, 2009) using the caret package (Kuhn, 2008). Normally, a classification is inferred using an RF model by running all the decision trees and using the majority vote as the output. However, if one keeps track of the vote count for each instance the model is run, then the vote count can be used as the continuous predictor in a Receiver Operator Characteristics (ROC) analysis. Then, an optimal threshold on the vote count may be computed from the ROC curve using Youden's Index (Powers, 2007). An optimal set of RF models was produced using vote thresholds of 162 votes for the steering RF model and 151 votes for the pedals RF model, where all RF models had 500 decision trees.

A Hidden Markov Model (HMM) was designed to include the effect of historical observations and accept inputs from the RF models, and was trained using the HMM library in R (Himmelman, 2010). A regular time interval of six seconds is selected as the model frequency. Two pieces of evidence are provided, one from the steering RF classification, and the other from the pedal RF classification. The output of the HMM, shown in Figure 2 is compared to a threshold to classify each time sample as a drowsy or awake. The threshold value was selected from an ROC curve to be 0.74. The RF models along with the HMM complete the drowsiness detection algorithm.

METHOD

Two mitigation systems were designed for use as between subject conditions in a new drowsiness mitigation study. The first is a three stage audio/visual alert with driver interaction through a button. The second is a binary haptic alert that vibrates the driver's seat. Three levels of the drowsiness detection system are included in the experimental design as a between-subjects dependent variable. The three levels will include a nominal design, a design that is more sensitive, and one that is less sensitive. These levels will expose drivers to different numbers of false alerts, while perhaps also failing to detect the drowsiness in some cases.

MITIGATION DESCRIPTIONS

The audio-visual alert is a three stage warning. The threshold value used to trigger each stage is the same for each stage. If drowsiness is detected while in the nominal state of no mitigation, then a stage 1 warning is issued. This warning is a white coffee cup icon with an OK button for driver acknowledgement (Figure 3a) and an audio chime that plays when the icon appears. Once the driver presses the button, the icon is removed. The mitigation will remain in stage 1 for a minimum period of time; and during that time the detection algorithm may remain in a classification of drowsy state or return to an awake state. If the detection algorithm classification returns to awake, then the mitigation will abate after a fixed delay. However, if another drowsy episode is detected before the mitigation abates, or the drowsy state persists for 60 seconds, then the mitigation escalates to stage 2. On entry into stage 2, a stage 2 warning is issued using the visual icon in Figure 3b along with an audio beep. This icon is removed once the driver acknowledges the warning with a button press. Exactly the same logic is applied during stage 2 until the mitigation either abates back to stage 1 or escalates to stage 3. A stage 3 alert consists of the icon in Figure 3c, and a repeated audio beep. There cannot be any more escalations from stage 3, but the warning may be reissued if the drowsy state persists or soon repeats. Only incremental escalations and abatements are allowed. This mitigation has the chance to capture the driver's attention by varying the stimulus on repeated warnings; but it also has the potential to be a nuisance to a driver who is already self-aware or not drowsy.

The haptic alert is a binary alert system that provides a counterpoint to the three stage alert. It also differs in modality by providing a haptic alert through seat vibration, thus making it a more subtle, and potentially less annoying, cue. The same logic for stage escalation is applied in the binary alert to either trigger the initial alert or repeat it after 60 seconds. Once the drowsy detection expires, the mitigation naturally abates back to the nominal driving state.

MITIGATION SENSITIVITY MANIPULATIONS

A significant question addressed in this paper is: how can we vary the sensitivity of the algorithm / mitigation system? Random Forest models are especially opaque and little intuition about why a given parameter set works is available to the designer. Hidden Markov models are slightly easier to intuit, but are nonetheless complicated. There are several choices that were considered, most of which were either discarded or found to not have a significant effect upon the final outcome of the classification and mitigation performance. The Random Forest models are considered first.

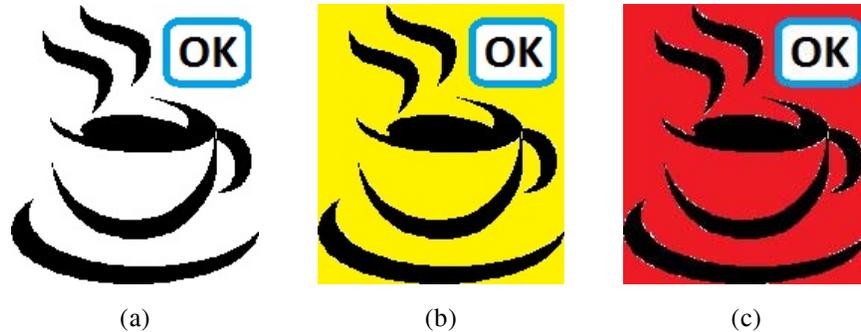


Figure 3. Visual Mitigation Icons: (a) stage 1 interactive, (b) stage 2 interactive, (c) stage 3 interactive

Two Random Forest models are used in the drowsiness detection algorithm, one for steering and one for pedals. Each Random Forest is composed of 500 decision trees; and each decision tree may have on the order of 100 nodes. Therefore, it is not feasible to try to tune those parameters individually. One could think about a strategy of retraining new RF models with the intent of changing the sensitivity; but then there may be other performance differences between them that are confounded with sensitivity. The other parameter one can think about tuning is the voting threshold for output classification. Normally, RF models are majority rule, meaning that more than 250 trees in a 500 tree RF would have to agree to set a class output. This vote threshold number may be allowed to vary and we used it as a threshold variable in an ROC analysis in our prior modeling work. The lower the value of the vote threshold, the more trees potentially need to be run to gather up the required number of votes, thus potentially increasing the computational demand of the model by some small amount. For example, with a threshold of 100, one may have to evaluate as many of 400 of the trees to guarantee that there are not 100 votes for drowsiness.

Hidden Markov Models have fewer parameters than RF models and they are more intuitive than trying to tune a decision tree. The state transition probabilities set the probability of an HMM changing state from awake to drowsy or vice-versa at any time step. The probabilities in each direction can be set independently. The emission probabilities set the chances that any of the observed variables of the HMM, or combinations thereof, indicate the value of the state. The state transition model is the base of the HMM with the prior probabilities, while the emission model conditions the state transitions with the presence of evidence. We estimated values for the emissions probabilities by counting the presence of RF model classification and their likelihood of correlating with a drowsy driving ground truth state. The final parameter that could be varied is the threshold we apply to the posterior probability, the output of the HMM, to set a final classification for the detection algorithm.

We chose not to attempt to tune the emission probabilities, for essentially the same reason we did not tune each decision tree. It would change the characteristics of the model and defeat the purpose of the machine learning training routines that optimize model parameters. We experimented with varying the state transition probabilities, the effect of which is similar to that of changing a low pass filter cutoff frequency that is filtering the HMM output (Figure 2). A lower transition probability will produce a more filtered signal that has a longer rise (or fall) time. Ultimately, the effect of varying these probabilities, while measurable, did not effect a significant change in the algorithm output.

Similarly, we explored the final HMM threshold value. This value was obtained previously as the optimal operating point on an ROC curve obtained after a model optimization process conducted on the RF and HMM models. This parameter is the easiest to understand, effectively dividing Figure 2 into an upper and a lower region that corresponds to drowsy and awake predictions, respectively. Unfortunately, the variation of this last threshold has the least effect out of all the parameter tuning that was tried. This is likely because most state transitions changed the posterior probability all the way from zero to one and vice versa. The number of cases where a transition changed direction partway was fewer than one might have expected. In that situation, we can only shift the edges of the state transition by a few seconds by varying the output threshold.

The parameters that had the greatest effect on the drowsiness detection algorithm were the vote thresholds of the two RF models. An RF model with a higher vote threshold simply requires more of its constituent decision trees to agree on the output classification. Setting this threshold above the majority value may be problematic because it may then be that neither class gathers enough votes to meet the threshold. Ten levels of parameters for the RF vote thresholds were set. Values for the Steer RF model are: {162, 170, 180, 190, 200, 210, 220, 230, 240, 250}. Values for the Brake RF are: {151, 160, 170, 180, 190, 200, 210, 220, 230, 240}. The parameters are always varied together, not independently. Level one values correspond to the optimal threshold obtained in prior work to optimize the ROC curve indicators of model performance. The subsequent levels step up the values of each threshold until the steering RF value reaches majority rule. Notice that the relationship between the two values is essentially maintained through the levels such that the steer RF threshold is always greater than the brake RF.

This particular range of parameters fits nicely with the goals of our model variation exercise. The optimal values gave the best performance when compared to the awake and drowsy ground truth data points; however, the majority of time history samples are not associated with any ground truth because there was no lane departure. Therefore, the algorithm performance at these points is difficult to judge. We did observe however, that many of these in-between points are classified as drowsy and thus contribute to the overall number of mitigation warnings. We would therefore consider this parameter set as being on the sensitive side. To make the models less sensitive, we wish to make it harder for the RF models to issue drowsy classifications, which means requiring more models to agree on drowsiness. Therefore we increase the values of the vote thresholds up to the majority rule value, but no further.

The different levels of RF models were run on the DrIIVE Phase I drowsiness data with all other parts of the detection algorithm held constant. Some simple metrics were calculated on the detection / mitigation system in order to compare across levels. A mitigation was considered to be in a ‘correct’ stage at each ground truth data point if it was in stage 0 (no mitigation) at an awake point or in any stage of mitigation at a drowsy point.

The system was designed to operate at speeds greater than 40 mph, so the percentage of time that the vehicle was traveling faster than this limit was calculated as a reference for other measures. A variable, *timeAtSpeed*, was calculated as the amount of time in the drive that the car was traveling above this limit. A variable, *timeInMitigation*, was calculated as the amount of time that the mitigation system was in any mitigation stage. Then a normalized measure was calculated as

$$Time\ in\ Mitigation\ (\%) = \frac{timeInMitigation}{timeAtSpeed} \times 100$$

Confining ourselves to only those samples with ground truth data, we counted which of those points were in the ‘correct’ stage of mitigation, as described above. This may be expressed in an indicator variable, I_c , of zeros and ones of length N , where N is the number of ground truth points in a drive. The percentage of correctly mitigated ground truth points was then computed in each drive as

$$Accuracy\ (\%) = \frac{100}{N} \sum_1^N I_c$$

This coarse metric does not indicate whether a ground truth data point falls in the first or last part of a mitigation, nor which stage of mitigation is active, nor whether an incorrectly mitigated ground truth point falls just before or after a period of mitigation. The accuracy metric, together with the time-in-mitigation metric, provide an idea of how parametric variations affect the output of the detection / mitigation system, and create a tradeoff between sensitivity and accuracy.

RESULTS

Ten levels of vote thresholds for the Steer RF and Pedals RF model were tested on the DrIIVE Phase I drowsiness data, which was all collected in unmitigated conditions. Both the three stage audio-visual mitigation as well as the binary haptic mitigation were run on each drive in the three conditions of that study: Day, Early Night, and Late Night. Since it was not possible to provide human interaction with the button response, an automatic button response was programmed after one second; therefore, the simulations do not account for unresponsive drivers.

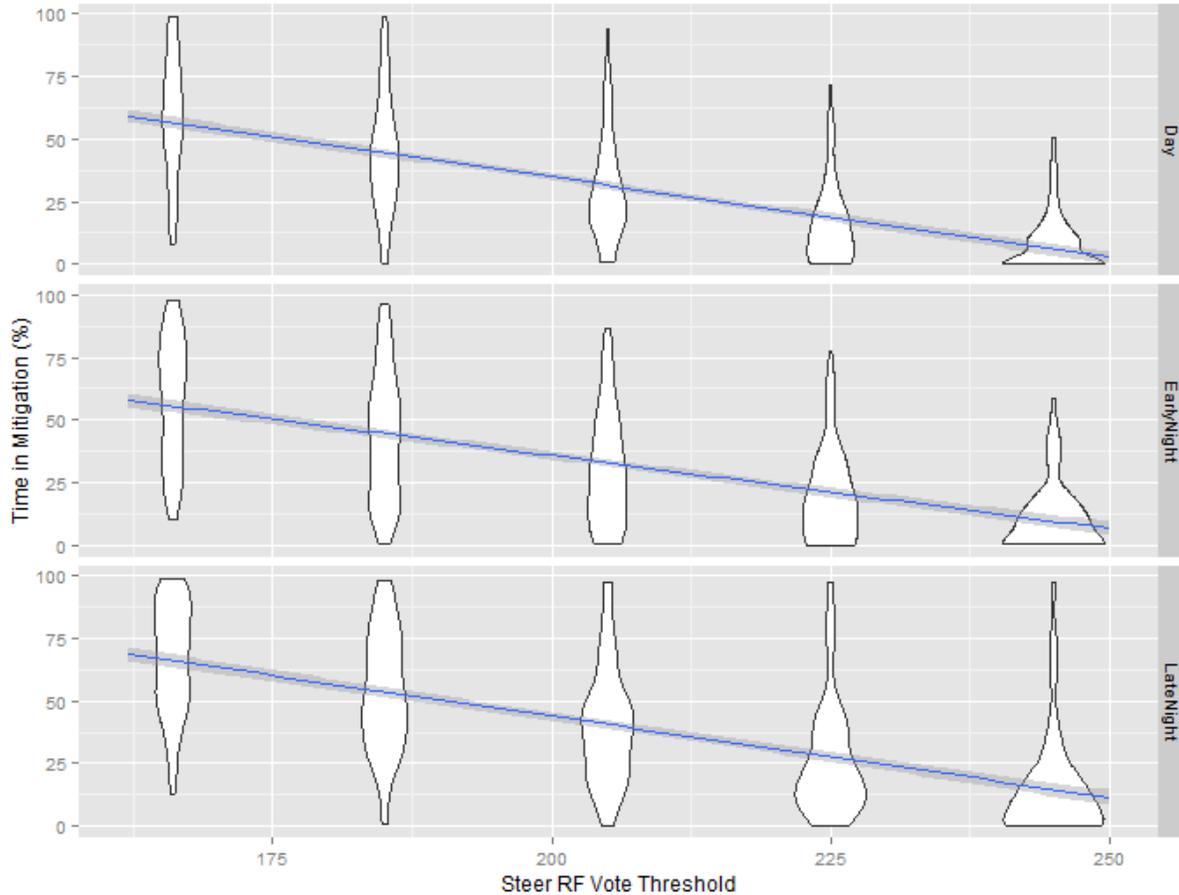


Figure 4. Violin plot and linear fit of the percentage of time in a drive that the mitigation is active for the audio/visual alert mode

The Time in Mitigation and Accuracy metrics are displayed for the three stage mitigation in Figure 4 and Figure 5, respectively. These figures show an overlay of a violin plot with a line fit, the latter with confidence intervals represented by a gray band. A violin plot shows information similar to a box plot, but shapes the sides of each ‘box’ according to the probability density of the sample points (Hintze & Nelson, 1998). The wider the shape is, the denser the points are at that location in the plot. The ggplot2 library (Wickhan, 2009) in R was used to generate the plots; and the plotting function was allowed to bin the horizontal axis from ten levels into just five bins, making the figure somewhat less dense and easier to comprehend.

The violin plot in Figure 4 shows the density of samples of the Time in Mitigation measure decreases as the vote threshold increases. This result holds across all three conditions and is completely intuitive. As more votes are required for the RF models to issue drowsy classifications, it becomes more difficult for the algorithm to transition into the drowsy state; and less time is spent in all stages of mitigation. The time in

mitigation at level ten for the Day, Early Night, and Late Night conditions is approximately 5%, 10%, and 12% respectively.

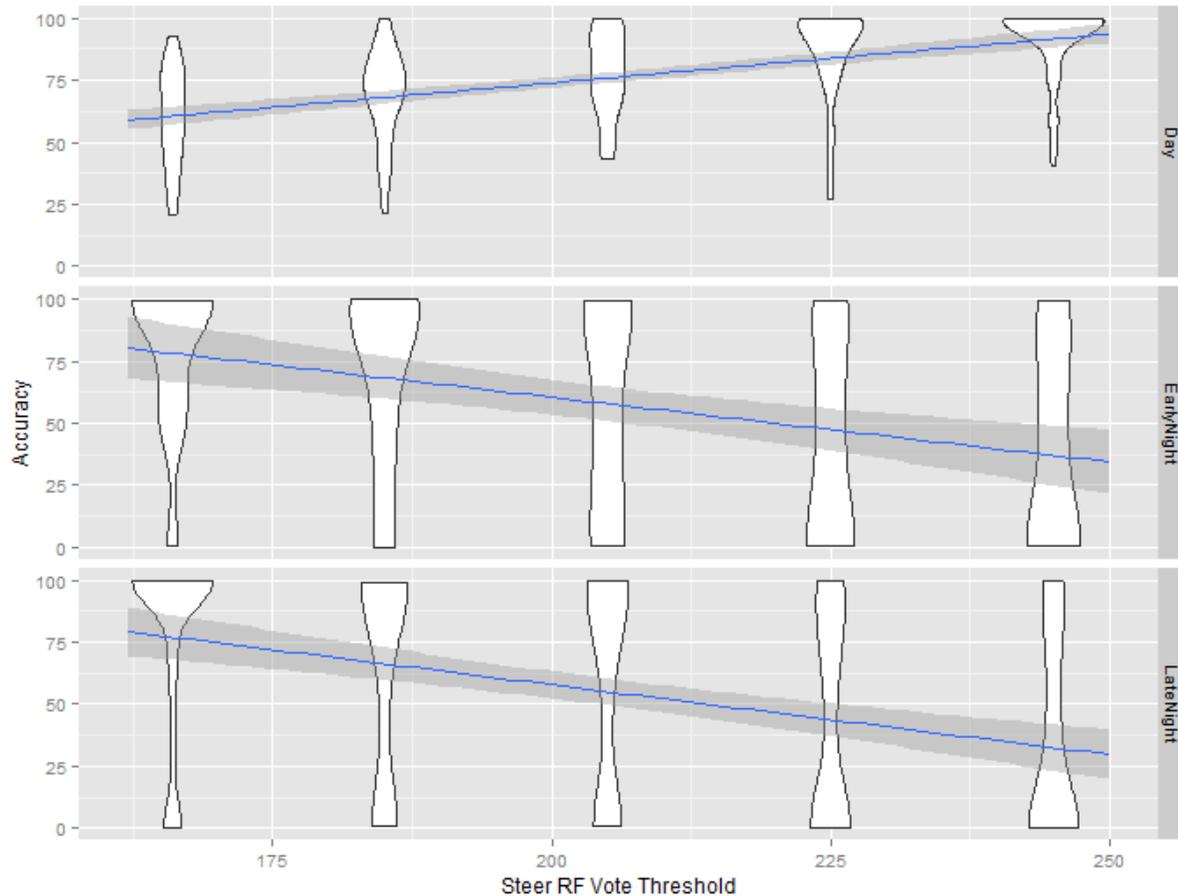


Figure 5. Violin plot and linear fit of the percentage of ground truth points that are correctly mitigated for the audio/visual alert mode

The violin plot in Figure 5 shows the density of samples of varying Accuracy as a function of the Steer RF vote threshold. The line fit serves to make clear the shift in the density of samples as the vote thresholds are increased. The accuracy in the Day condition actually rises as the RF models become more conservative. The explanation for this result is that almost all of the ground truth points in the Day conditions are awake points. Then it becomes clear that a simplistic approach of turning off the mitigation altogether would increase the accuracy in this condition to almost 100%. On the other hand, the accuracy is seen to drop for both night conditions as the vote threshold is increased, as expected. At the far end of the test, where vote thresholds for steering and pedals are 250 and 240 respectively, the estimated accuracy in the Day, Early Night, and Late Night conditions is 90%, 38%, and 30% respectively.

A similar pattern of results was obtained for the binary haptic mitigation, though the haptic system had smaller overall values for the Time in Mitigation metric. The different logic of the binary mitigation as compared to the more complex three-stage system result in less time spent in mitigation.

THREE SENSITIVITY LEVELS

The drowsiness mitigation study will have Early Night and Late Night conditions, but will not include Day drives. However three levels of sensitivity will be designed for each of the two mitigation types. Previous work resulted in trained models and an ROC curve evaluation of the models to optimize a drowsiness detection

algorithm. This ‘optimal’ model corresponds to the level one parameter set described in this paper. As discussed earlier, that optimization only considered ground truth points and classifications on in-between points were not part of the evaluation. In reality, the in-between samples make up a majority of the data in most drives and therefore contribute significantly to the number of mitigation alerts. All three conditions were in some stage of mitigation over half the time, which is especially surprising for the Day condition.

A commercial system would most likely condition the algorithm output on other factors such as time of day, driving style, traffic density, and perhaps other variables. Since we are not collecting additional Day drives, conditioning by day/night is not necessary at this time. Three mitigation models were selected for the two modalities with the purpose of obtaining a wide spread in the *timeInMitigation* and *accuracy* metrics. The starting point was to choose three target accuracy values. Those values then mapped to corresponding time in mitigation and RF model vote thresholds. The values selected for the three models are summarized in Table 1. These models are spaced far enough apart that they offer a clear distinction to the drivers who experience them.

Table 1. Three levels of mitigation selected using Late Night data for audio/visual and haptic modes

Level	Steer RF Votes		Pedal RF Votes		Time in Mitigation (%)		Accuracy (%)	
	Visual	Haptic	Visual	Haptic	Visual	Haptic	Visual	Haptic
Over sensitive	170	162	160	151	63	40	75	70
Nominal	190	175	180	165	50	38	63	63
Under-sensitive	215	195	205	185	38	30	50	50

CONCLUSIONS

Two mitigation systems, a three stage audio/visual and a binary haptic, were designed to use the output of a previously designed drowsiness detection algorithm. Additionally, three levels of each system were obtained to provide a good range of system sensitivity to drowsiness. In this way, a range of false alarm rates will be generated from the study and questions about the effectiveness of the mitigation might be differentiated from questions about the nuisance factor of the mitigation alerts.

The ultimate desired outcome for a drowsiness mitigation system is that the driver would realize their own impairment and pause the trip to rest. Such an outcome is not allowed for however in the protocol of the simulator experiment. On the other hand, a primary interest of the DrIIVE program is to study vehicle-based measures of impairment. Having determined that such measures are useful for classifying drowsiness, the data from the upcoming study may be used to test whether a mitigation system also causes detectable differences in driving performance as measured by vehicle-based sensors.

REFERENCES

- Balkin, T. J., Horrey, W. J., Graeber, R. C., Czeisler, C. A., Dinges, D. F. (2011). The challenges and opportunities of technological approaches to fatigue management. *Accident Analysis & Prevention*, 43(2), 565–72.
- Brown, T., Lee, J., Schwarz, C., Fiorentino, D., McDonald, A. (2014). *Assessing the Feasibility of Vehicle-based Sensors to Detect Drowsy Driving* (Final Report No. DOT HS 811 886). Washington, DC: NHTSA.
- Dinges, D. F., & Grace, R. (1998). *PERCLOS: A valid psychophysiological measure of alertness as assessed by psychomotor vigilance*. Federal Highway Administration. Washington D.C.
- Grace, R., Byrne, V. E., Bierman, D. M., Legrand, J.-M., Gricourt, D., Davis, B. K., Carnahan, B. (1996). A drowsy driver detection system for heavy vehicles. In *17th DASC. AIAA/IEEE/SAE. Digital Avionics Systems Conference. Proceedings*, Vol. 2, pp. I36/1–I36/8.
- Himmelmann, L. (2010). HMM: HMM - Hidden Markov Models (Version 1.0). Retrieved from <http://cran.r-project.org/web/packages/HMM/index.html>
- Hintze, J. L., & Nelson, R. D. (1998). Violin Plots: A Box Plot-Density Trace Synergism. *The American Statistician*, 52(2), 181–184.
- Ji, Q., Lan, P., & Looney, C. (2006). A probabilistic framework for modeling and real-time monitoring human fatigue. *IEEE Transactions on Systems, Man, and Cybernetics - Part A: Systems and Humans*, 36(5), 862–875.
- Ji, Q., Zhu, Z., & Lan, P. (2004). Real-Time Nonintrusive Monitoring and Prediction of Driver Fatigue. *IEEE Transactions on Vehicular Technology*, 53(4), 1052–1068.
- Klauer, S., Dingus, T., Neale, V., Sudweeks, J., Ramsey, D. (2006). *The impact of driver inattention on near-crash/crash risk: An analysis using the 100-car naturalistic driving study data. Analysis*. (Report No. DOT HS 810 594). Washington D.C.: National Highway Traffic Safety Administration.
- Krajewski, J., & Sommer, D. (2009). Steering wheel behavior based estimation of fatigue. In *Proceedings of the 5th International Driving Symposium on Human Factors in Driver Assessment and Design* (pp. 118–124).
- Kuhn, M. (2008). Building Predictive Models in R Using the caret Package. *Journal of Statistical Software*, 28(5).
- Lal, S. K. L., Craig, A., Boord, P., Kirkup, L., Nguyen, H. (2003). Development of an algorithm for an EEG-based driver fatigue countermeasure. *Journal of Safety Research*, 34(3), 321–328.
- Lin, C., Wu, R., Liang, S., Chao, W., Chen, Y., Jung, T. (2005). EEG-based drowsiness estimation for safety driving using independent component analysis. *IEEE Transactions on Circuits and Systems I: Regular Papers*, 52(12), 2726–2738.
- MacLean, A. W., Davies, D. R. ., Thiele, K. (2003). The hazards and prevention of driving while sleepy. *Sleep Medicine Reviews*, 7(6), 507–521.

- McCartt, A. T., Rohrbaugh, J. W., Hammer, M. C., Fuller, S. Z. (2000). Factors associated with falling asleep at the wheel among long-distance truck drivers. *Accident Analysis & Prevention*, 32(4), 493–504.
- McDonald, A. D., Lee, J. D., Schwarz, C., Brown, T. L. (2013a). Steering in a Random Forest: Ensemble Learning for Detecting Drowsiness-Related Lane Departures. *Human Factors: The Journal of the Human Factors and Ergonomics Society*.
- McDonald, A. D., Lee, J. D., Schwarz, C., Brown, T. L. (2013b). Steering in a Random Forest Ensemble Learning for Detecting Drowsiness-Related Lane Departures. *Human Factors: The Journal of the Human Factors and Ergonomics Society*.
- NHTSA. (2011). *Drowsy Driving* (Traffic Safety Facts No. DOT HS 811 449). Washington, DC: NHTSA.
- Powers, D. M. W. (2007). *Evaluation: From Precision, Recall and F-Factor to ROC, Informedness, Markedness & Correlation* (Technical Report No. SIE-07-001). Adelaide, Australia: Flinders University.
- R Development Core Team. (2009). *{R: A language and environment for statistical computing}*. Vienna, Austria: R Foundation for Statistical Computing.
- Sayed, R., & Eskandarian, A. (2001). Unobtrusive drowsiness detection by neural network learning of driver steering. *Proceedings of the Institution of Mechanical Engineers, Part D: Journal of Automobile Engineering*, 215(9), 969–975.
- Schwarz, C., McDonald, A., Lee, J. D., & Brown, T. L. (submitted). Time Series Classification of Drowsiness from Driving Input Signals. *IEEE Journal on Intelligent Transportation Systems*.
- Wali, M., Murugappan, M., & Ahmmad, B. (2013). Wavelet Packet Transform Based Driver Distraction Level Classification Using EEG. *Mathematical Problems in Engineering*.
- Wickhan, H. (2009). *ggplot2: elegant graphics for data analysis*. New York: Springer.
- Yang, G., Lin, Y., & Bhattacharya, P. (2010). A driver fatigue recognition model based on information fusion and dynamic Bayesian network. *Information Sciences*, 180(10), 1942–1954.
- Yang, J. H., Tijerina, L., Pilutti, T., Coughlin, J. F., Feron, E. (2009). Detection of Driver Fatigue Caused by Sleep Deprivation. *IEEE Transactions on Systems, Man, and Cybernetics - Part A: Systems and Humans*, 39(4), 694–705.

ANALYSIS OF THE ROBUSTNESS OF STEERING PATTERN BASED DROWSINESS DETECTION

Christiane Spindler
Frauke Driewer
Lorenz Schäfers
Daimler AG
Germany

Paper Number 15-0226

ABSTRACT

Several studies show that up to one in four severe traffic accidents can be attributed to drowsiness. Drivers often over-estimate their fitness level or are not aware of the danger that always accompanies drowsy driving. Since associations like the NHTSA pointed to the relevance of this topic, more and more research has been conducted and in the meantime there is also a variety of commercial systems on the market to address this risk. In this paper, we do not aim to find new methods of detecting drowsiness of a driver. Our approach is rather to choose an established method and enhance it in a way that it not only performs well in a driving simulator but also in real world drives.

The chosen drowsiness detection method is the observation of the steering wheel angle signal. It has been shown that the frequency of occurrence of a typical steering pattern, which can roughly be described as a deadband followed by a rather fast correction, is an indicator for the state of drowsiness of a driver. The advantage over other techniques like camera-based detection is that it can run in standard equipped cars. Thus it is available for the largest number of drivers and can thereby achieve the greatest effect on accident avoidance.

We investigate the chosen detection method in real world drives and discuss which other effects not related to drowsiness can evoke the described steering pattern. We focus on environmental effects like crosswind and can show that those events may lead to an increase of the amount of steering patterns. Finally, we quantify the influence on drowsiness measures. The underlying database comprises more than two million kilometers of more than one thousand drivers, all real-world drives.

Our evaluation shows that particularly on routes or in situations where those environmental influences accumulate, the drowsiness measure can be affected to an extent that leads to false triggering of the system. Therefore, we suggest measures that can be taken to reduce the influence of steering patterns that are not related to the driver's drowsiness state.

The aim of most drowsiness detection systems is to inform a driver when his state has reached a critical level and to motivate him to take appropriate measures. This presupposes confidence in the system. False warnings will negatively affect the credibility of the system.

Our purpose is to show the importance of enabling this kind of system to recognize external influences, thus making detection more robust. We consider it very important to make such systems as reliable and credible as possible, as otherwise the driver will not take the advice the system will give him. Limiting the influence of external factors is a key to achieving this goal.

INTRODUCTION

Numerous reports name drowsiness and distraction as the cause of alarming numbers of accidents. The National Highway Traffic Safety Administration (2010) reports that in 2009 16% of all fatal crashes in the United States involved distracted driving. As regards drowsiness, Horne and Reyner (1995) found that 20% of all accidents on motorways in Southwest England to which the police was present were sleep-related. According to Langwieder et al. (1994), 24% of all fatal crashes in Bavaria, Germany, in 1991 happened because the driver fell asleep. NHTSA (Royal, 2002) reports 56,000 crashes annually to be related to drowsiness as mentioned by the police, resulting in 1,550 fatalities. In the same report, NHTSA lists reasons why these numbers are presumed to be conservative. Furthermore, crashes due to drowsiness tend to have a severe outcome (Wang et al., 1996).

The focus on the topic is still increasing. NHTSA names distracted and drowsy driving as one of the traffic safety problem areas (Goodwin et al., 2013) and the Euro NCAP 2020 Roadmap aims to reward manufacturers in the area of driver state monitoring in order to bring down the numbers of vehicles departing the road (European Car Assessment Programme, 2014).

A lot of research has been conducted in the field of drowsiness recognition and in the last years several commercial systems have become available on the market, using different methods. Dong et al. (2011) and Platho et al. (2013) give an overview of driver monitoring systems and also mention the commercial products of Ford, Mercedes-Benz, Volvo and VW. All those systems aim to suggest the driver to take a rest when he has reached a critical level. Many different algorithms were developed that analyze the driving performance, e.g. based on steering behaviour or lane keeping ability. These algorithms normally detect drowsiness if the driver shows an unusual driving behaviour (e.g. leaving the lane too often) or if the driving behaviour changes significantly from the beginning (e.g. lane keeping ability decreases).

A problem of methods that use driving performance as criteria for drowsiness detection is that only the reaction of the driver can be analyzed, not the reason for certain driving manoeuvres. Attwood (2014) mentions that systems, though they work in driving simulators, may fail on real roads, as they are not able to detect what the driver is responding to, considering environmental characteristics related to road, traffic and weather.

In the present paper we discuss which environmental characteristics may have an impact on driver monitoring systems. In detail, the influence of crosswind and road disturbances is analyzed and it is estimated to what extent those events have an impact on drowsiness recognition. Finally it is shown how these external factors are taken into account in the system under consideration.

The following evaluation is based on the steering wheel angle signal as the main information source. The main advantage of this method is that no special sensor, e.g. lane detection or driver monitoring camera, is needed. The steering wheel angle signal is part of the standard equipment of present-day cars. By this means, it is possible to integrate the drowsiness detection as a standard feature and thus reach a high number of drivers.

APPROACH

Steering wheel angle based drowsiness detection

Several studies investigating the use of the steering wheel angle signal for drowsiness detection have been carried out. Dingus et al. (1987) found that the number of steering wheel velocities over 150deg/s is an indicator for drowsiness. Bouchner et al. (2006) show a positive correlation of the ratio of fast and slow steering corrections with drowsiness.

A combination of slow and fast steering velocities is also used in this study. It is based on the Mercedes-Benz Attention Assist, which is a system that detects drowsiness and long-term distraction. Both kinds of driving impairment affect the steering behavior in a similar way. The steering pattern we evaluate consists of a deadband (phase without or with very slow steering) and a subsequent fast steering correction. Friedrichs and Yang (2010) show that this pattern correlates with drowsiness.

In our experience, steering velocities differ widely between drivers. Therefore, several thresholds in the algorithm are adapted continuously during the drive and according to the behavior in the first minutes of a drive, when the driver is presumed to be rather awake.

The accumulated steering pattern is the basis for the drowsiness measure. Figure 1 shows other factors that are taken into account to make the system more robust and useable in real road environment.

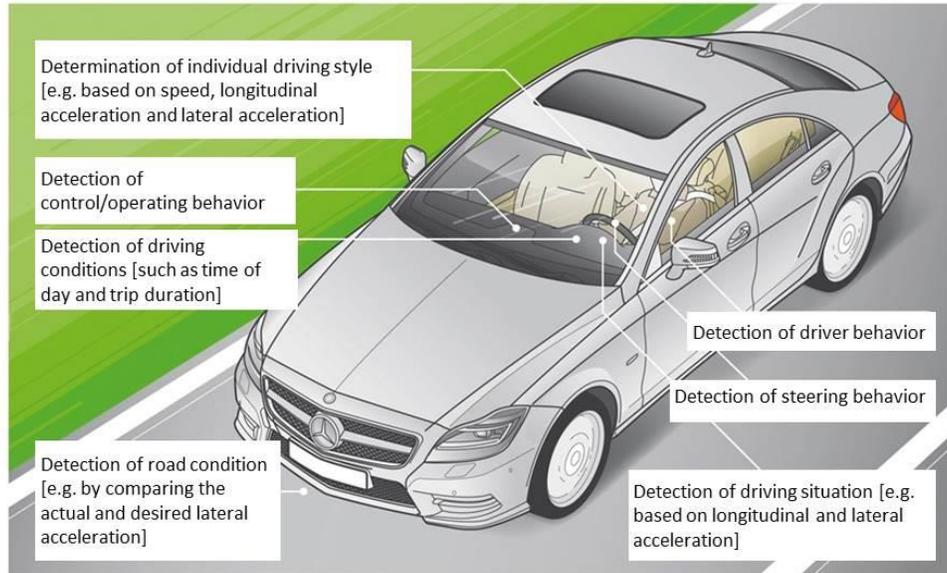


Figure 1. Features of the Mercedes-Benz Attention Assist.

Situations potentially provoking steering patterns

Friedrichs et al. (2011) identified external influences on the driving behavior. We refer to the factors listed in that study and cluster them as shown below.

Gaze direction, distraction, vehicle operation. Factors describing different kinds of driver action. These include for example eating or drinking, which can lead to abnormal steering behavior as the driver does not operate the vehicle with both hands. It may also be the driver not keeping his eyes on the road because he is attending to his children on the back seat or reading messages on his smartphone. The steering patterns arising from these actions can be classified as indicators of distraction and are thus treated by the system in the same way as steering patterns evoked by drowsiness. Vehicle operations on the other hand are part of the driving task. They can be detected by the system and the related steering patterns can be filtered out.

Vehicle type/motorization, posture. Influences that can be summarized as characteristics of the vehicle and the driver. A key issue of these factors is that they normally do not change during a drive. Hence, adaptive systems are able to minimize the influence.

Rain/fog/snow, traffic density, lane width/-number, speed, curvature. Description of the driving situation. An impact on the driving behavior is probable. Those situations are usually of longer duration. Some of these situations can easily be detected with standard sensors, e.g. speed or curvature. Others are more complicated to be analyzed online, e.g. traffic density. Nevertheless, as the factors are usually of longer duration, adaptive algorithms can react on the change in an adequate time.

Road condition, road bumps, crosswind, warping. Single, strong events with sudden occurrence that may have immediate impact on the driving behavior. Crosswind often occurs unexpectedly, laterally displaces the vehicle and thus requires a fast counter-steering. Road bumps, warping or potholes can also lead to unintentional steering corrections. Steering corrections that potentially arise from these environmental influences are neither related to drowsiness nor to distraction and should therefore not be considered for driver state monitoring.

In this study, we concentrate on the environmental events and investigate the influence of crosswind and road irregularities (road bumps, potholes) in detail. Friedrichs et al. (2011) conducted special drives for their evaluation in order to keep the dimension of the influences as small as possible. In the following evaluation real road data from naturalistic driving is used. Some restrictions were made on speed range and rated

drowsiness. Therefore, a much larger number of drives are part of the evaluation and the study of Friedrichs et al. (2011) is extended towards real driving situations.

Recognition of environmental influences

A prerequisite for all further evaluation is the ability to detect the presence of environmental influences. The detection of potholes and road bumps is based on an algorithm that looks for characteristics in the rotational speed of the wheels. The occurrence of crosswind is detected by comparing the steering angle, which provides information regarding the driver's intention, to the lateral acceleration, which supplies the actual lateral vehicle movement. This approach of crosswind detection also includes the recognition of road warping. Often a mixture of road bumps and warping occurs, which means that the detection of crosswind, warping and road bumps is not always separable. Hence, some events are recognized by both algorithms. All signals mentioned are available in standard equipped cars.

The algorithms described have been extensively proven in real world drives. For the following evaluations, the results of those algorithms have been used as labels for the presence of environmental influences.

EVALUATION

Underlying Database

All data used comes from naturalistic drives. Driving simulator data is not included. The database comprises more than two million kilometers conducted by more than one thousand drivers. Self-rating of driver drowsiness is available for each drive. This rating has been conducted according to the Karolinska Sleepiness Scale (KSS) (Åkerstedt & Gillberg, 1990). Every single drive has undergone a validation process to make sure quality standards like consistent values of the KSS-rating are fulfilled. All drives come from Mercedes-Benz cars, but have been conducted in different models from the A-Class (compact car) to the S-Class (luxury large car).

Evaluation of steering behavior

Prior to the investigation of the occurrence of steering patterns, a more general look at the steering behavior was taken. Steering velocities were explored regarding the influence of crosswind or road bumps.

As there are significant differences in steering behavior between individuals the analysis was conducted separately for single drivers. From the entire database, the ten drivers with the largest amount of recorded data were selected. Since Friedrichs et al. (2011) have shown that speed has a strong impact on the steering velocities, distributions of this signal in different speed ranges were compared for single drivers. Afterwards, the speed range under consideration was limited to velocities between 100km/h and 200km/h, as the signal values vary more at lower speeds. In addition, only parts of the drives were considered in which the driver was awake and alert.

For each drive of the ten selected drivers the ratio of the presence of crosswind to the duration of the whole drive in the considered speed range was calculated. The lower quartile Q_1 and the upper quartile Q_3 of this ratio were then used for each driver to group his measurements into rather smooth drives (*group 1*) and drives under windy conditions (*group 2*). The same was done for proportions of the presence of road bumps and warping. Accordingly, *group 1* comprises smooth drives and *group 2* drives on roads with frequent disturbances.

Subsequently, mean (*mean*) and variance (*var*) of the steering wheel velocity (*swv*) of each group of drives was calculated.

Table 1 shows the results for the crosswind comparison. The calculated ratios are defined according to Eqs. (1-2).

$$ratio\ mean = \frac{mean(swv)_{group2}}{mean(swv)_{group1}} \quad (1)$$

$$ratio\ var = \frac{var(swv)_{group2}}{var(swv)_{group1}} \quad (2)$$

From the values in Q_1 and Q_3 it can be seen how much crosswind was present in *group 1* and *group 2*. The interquartile range *IQR* shows how strong the two groups differ in their amount of crosswind.

It can be seen from the table that driver A sticks out, having the highest *ratio mean* and *ratio var*, which means that the mean value of his steering velocity is higher for drives under windy conditions while also the distribution is spread more widely. In comparison, for driver E both ratios are still greater than one but with much smaller values. Hence, this driver also has higher steering velocities with a higher variance for his drives of *group 2*, but the effect is less marked than for driver A. A look at the quartiles gives an explanation for this difference. The value of Q_3 , which is the threshold for drives under windy conditions, is much higher for driver A than for driver E, while Q_1 is the same for both drivers. Thus, data from more windy conditions is existent for driver A than it is for driver E, which results in a higher effect on the steering velocities.

In summary, for all drivers the mean value and the variance of *swv* is higher for drives of *group 2* than *group 1*. A look at the individual thresholds Q_1 and Q_3 and the *IQR* shows that this effect is stronger for individuals for which a greater difference in the ratio of crosswind occurrence is present. Taken together, these results reinforce the expectation that higher steering velocities occur with environmental disturbances.

Table1.
Comparison of steering velocity mean and variance for drives with different ratios of crosswind occurrence.

Driver	data selection	number of drives	evaluated time [min]	mean(swv) [°/s]	var(swv) [°/s]	ratio mean	ratio var	Q_1	Q_3	<i>IQR</i>
A	Group 1	79	6369	1.00	2.17	2.84	10.06	0.03	0.13	0.10
	Group 2	79	4420	2.84	21.79					
B	Group 1	44	3608	1.44	4.04	1.41	1.92	0.03	0.10	0.07
	Group 2	44	2751	2.04	7.75					
C	Group 1	43	3879	1.38	3.18	2.01	3.18	0.03	0.18	0.15
	Group 2	43	2234	2.78	10.12					
D	Group 1	87	7276	1.72	4.12	1.93	4.51	0.03	0.11	0.08
	Group 2	87	4668	3.33	18.56					
E	Group 1	37	2942	1.27	2.86	1.18	1.22	0.03	0.07	0.04
	Group 2	37	2923	1.51	3.48					
F	Group 1	39	2766	1.68	4.37	1.19	1.46	0.02	0.04	0.02
	Group 2	39	2832	2.01	6.38					
G	Group 1	44	4110	1.09	1.94	1.19	1.48	0.02	0.06	0.04
	Group 2	44	2636	1.29	2.88					
H	Group 1	47	3203	1.22	2.72	1.27	1.53	0.03	0.09	0.06
	Group 2	47	3017	1.54	4.14					
I	Group 1	29	1768	1.28	2.64	1.22	1.42	0.02	0.06	0.04
	Group 2	29	2335	1.57	3.75					
J	Group 1	36	3872	1.04	1.81	1.27	1.63	0.02	0.07	0.05
	Group 2	36	2346	1.32	2.95					

The same procedure was applied for the presence of road disturbances. The result is presented in Table 2. Though not as definitive as for crosswind, the findings are the same. For all drivers, both mean and variance of the steering velocity are higher for data of *group 2*, which are the drives with a high amount of road disturbances. The tendency of a higher portion of road irregularities leading to higher mean values and higher variance of the steering wheel velocity can also be observed: driver I has the smallest *IQR*, which means the

difference of the ratio of road disturbance occurrence between his drives in *group 1* and *group 2* is smaller than for the other drivers. This explains why the mean steering velocity differs less between *group 1* and *group 2* than it does for example for driver D, whose drives in *group 2* feature a larger ratio of road irregularities.

Table2.
Comparison of steering velocity mean and variance for drives with different ratios of road bumps occurrence.

Driver	data selection	number of drives	evaluated time [min]	mean(<i>swv</i>) [°/s]	var(<i>swv</i>) [°/s]	ratio mean	ratio var	Q_1	Q_3	IQR
A	Group 1	79	6701	1.01	2.15	2.47	8.46	0.03	0.18	0.15
	Group 2	79	5487	2.51	18.22					
B	Group 1	44	3040	1.53	4.20	1.43	2.06	0.03	0.13	0.10
	Group 2	44	2464	2.19	8.66					
C	Group 1	43	3849	1.62	3.98	1.50	2.14	0.04	0.23	0.19
	Group 2	43	2504	2.42	8.54					
D	Group 1	87	6853	1.82	4.47	1.65	3.59	0.03	0.14	0.11
	Group 2	87	5824	3.00	16.04					
E	Group 1	37	2991	1.27	2.89	1.18	1.24	0.07	0.18	0.11
	Group 2	37	2875	1.50	3.60					
F	Group 1	39	3068	1.76	5.01	1.01	1.07	0.04	0.10	0.06
	Group 2	39	2315	1.78	5.35					
G	Group 1	44	3762	1.05	1.94	1.22	1.40	0.02	0.10	0.08
	Group 2	44	2160	1.28	2.71					
H	Group 1	47	3542	1.33	3.28	1.10	1.05	0.02	0.10	0.08
	Group 2	47	3155	1.47	3.45					
I	Group 1	29	2163	1.37	3.04	1.06	1.14	0.02	0.08	0.06
	Group 2	29	1859	1.46	3.45					
J	Group 1	36	3647	1.05	1.92	1.11	1.19	0.03	0.09	0.06
	Group 2	36	3382	1.17	2.28					

Evaluation of the occurrence of steering patterns

To find out whether there are peculiarities in the number of steering patterns with the presence of crosswind, all time instances of onsets of crosswind in the speed range 60-200km/h were identified for 11,604 drives. Afterwards a time range of ten seconds before and ten seconds after those time instances was investigated for steering patterns.

Figure 2 provides the cumulated result for all time instances in which crosswind was detected. Zero on the time axis marks the beginning of crosswind. As the duration differs, the red vertical line marks the median of the end of the detected crosswind. The number of steering patterns have been counted and plotted at their instant of occurrence, relative to the beginning of crosswind and normalized with the number of crosswind events. As can be seen, in general the number of steering patterns moves around a certain level. After the onset of crosswind, a very strong rise can be observed. The subsequent lower amount is attributable to the violation of the deadband criteria caused by the counter-steering. It is apparent from this data that more steering corrections are produced under the influence of crosswind.

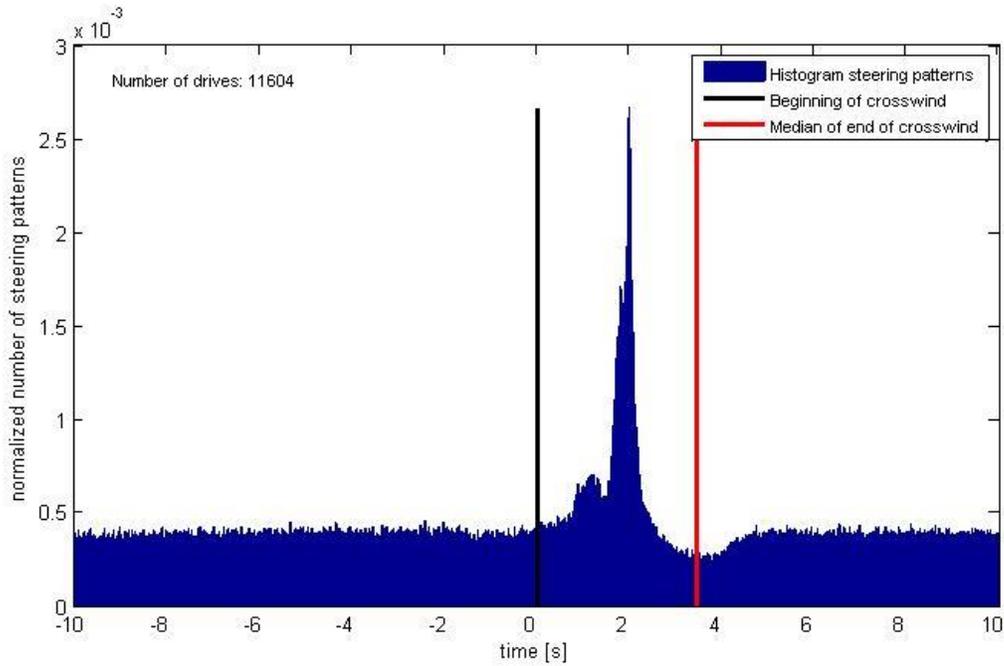


Figure2. Histogram of steering patterns around crosswind events.

The same procedure was applied for road surface irregularities. The results obtained from 11,638 drives are shown in Figure 3. The observation is the same as for crosswind. The number of steering patterns varies little around a certain level and increases strongly when road bumps occur. It is thus confirmed that road disturbances can lead to steering corrections.

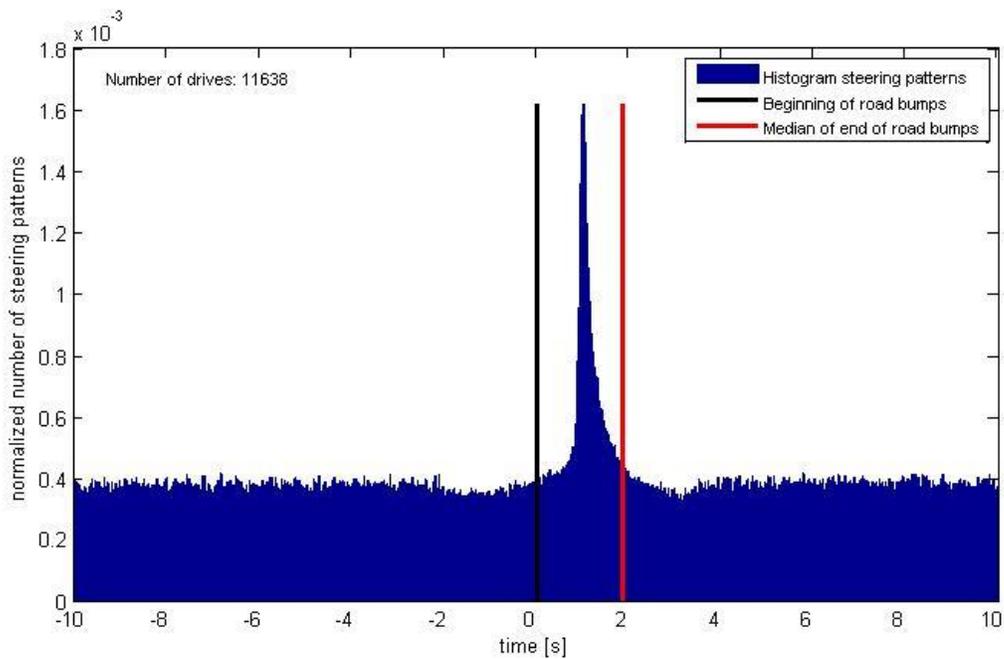


Figure3. Histogram of steering patterns around road disturbances.

Quantification of influence on drowsiness detection

The previous evaluation proved that environmental characteristics can evoke steering patterns. In the next step the dimension of the influence on a possible drowsiness measure was estimated. Based on this final evaluation it could be determined whether environmental influences present a severe problem or if effects are minor and can be neglected.

This estimation was performed by calculating the factor by which the number of steering corrections increases if environmentally influenced ones are taken into account. Only data was used, in which the driven speed lay for at least 30min in the range of 60 to 200km/h. This led to 6075 evaluable drives. For each drive, the ratio of the amount of steering corrections that were detected during the presence of crosswind to the amount of steering patterns that occurred when no environmental disturbances were present was calculated. The result shows by which factor the number of steering patterns would increase if those evoked by crosswind were ignored. It also represents an estimation of how much a drowsiness measure, based only on a summation of steering patterns, would be affected.

The same principle was applied for the computation of the increase of steering patterns as a result of road disturbances. Figure 4 presents the distributions of the results for both kinds of environmental influences in a boxplot. A factor of increase of one means that the number of steering patterns would double by taking into account the environmental disturbance-evoked ones. For a better readability, only values up to 1.5 are shown. This was done due to some striking outliers, which may occur for special driving conditions, e.g. extraordinary windy conditions.

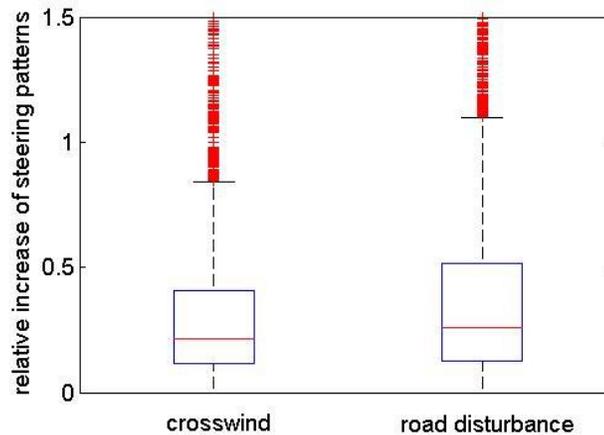


Figure4. Increase of number of steering patterns with environmental disturbances.

The median for the factor of increase on account of crosswind lies at a value of 0.215, due to road irregularities at 0.258. As explained before, the crosswind recognition and the detection of road bumps may sometimes be effective for the same events, thus it has not to be assumed that both factors of increase would add up. But, for half of the drives in the existing database, a possible drowsiness measure increases by more than 20% even regarding only one of the influences, which may indeed lead to false warnings. The problem is less severe for drives under smooth conditions and more severe if more disturbances occur. Figure 5 shows that the rise of steering patterns and the relative amount of crosswind is highly correlated, as can be expected. The same observation can be made for road irregularities, as shown in Figure 6. Especially for drives under more extreme conditions, measures have to be taken to increase the robustness of the drowsiness recognition system to prevent false alarms.

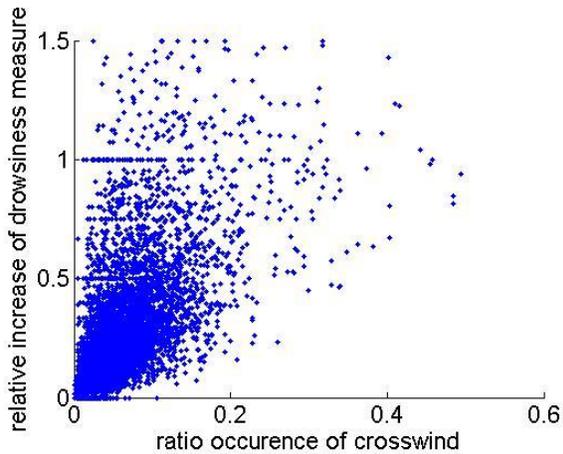


Figure5. *Relation between amount of crosswind and increase of steering patterns.*

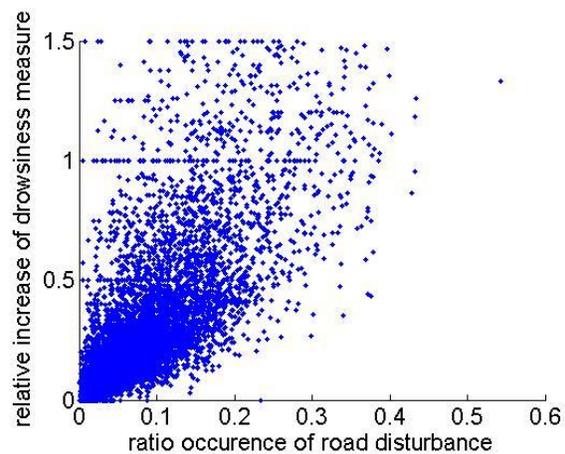


Figure6. *Relation between amount of road disturbances and increase of steering patterns.*

PROPOSAL OF MEASURES

The evaluations show that the impact of environmental influences on steering pattern based drowsiness detection systems is too strong to be neglected. In the following we propose measures that increase the robustness as they are implemented in the Mercedes-Benz Attention Assist system.

Masking

In the first step, steering corrections evoked from events like crosswind or road bumps are left out of the estimation of drowsiness. A prerequisite for this is the possibility of recognizing such disturbing influences. Not to consider those steering patterns means that the system cannot evaluate the driver's steering behavior during the presence of the environmental disturbance. This leads to some kind of system inactivity. Inactivity due to environmental influences is of short duration. Figure 2 and Figure 3 depict the median of the duration of disturbing environmental events, about 3.5s and 2s respectively. In our database, crosswind led to 4.3% of overall system inactivity, while masking due to road disturbances concerned 6.2% of all data. Both values were obtained from data in the speed range 60-200km/h. Hence, the inactive periods have only minor influence on the overall system performance.

If the system is inactive for a long time, we recommend letting the driver know that he cannot expect it to work without restrictions. This is for example the case if the system works only in a certain speed range. Transparency, such as displaying inactivity, can lead to better understanding and thus more trust in the system.

Adaption

While masking is effective for determined events, another measure is needed for all non-specific influences that cannot be detected as single environmental events can be. Increased robustness can also be achieved by making algorithms adaptive, not only to the driver but also to changes in the driving situation that cannot be attributed to special events. For example thresholds for the recognition of steering patterns should adapt during the whole drive.

CONCLUSIONS

The purpose of the current study was to determine the necessity of making driving-performance based driver state monitoring systems, especially those that rely on steering patterns, robust against environmental influences. The results of the investigations have shown that environmental influences have a significant

impact on the steering behavior and can lead to steering patterns that are not related to drowsiness or distraction.

It has also been found that the number of unwanted steering patterns cannot be disregarded. The influence on the drowsiness measure is significant, especially with higher presence of disturbances. The implementation of possibilities to detect environmental events and ignore the consequent steering corrections helps to achieve better performance of such systems in real road scenarios. The performance can be further improved by designing adaptive algorithms, e.g. by fitting certain parameters to the special driving situation.

The effectiveness of drowsiness detection systems that are limited to give advice depends on the driver's confidence. An increase in the drowsiness measure because of environmental influences can lead to false triggering of the system and thus to the driver not taking it seriously.

The presented method of evaluating influence of certain events on drowsiness detection algorithms can also be used to study the effect of other events, e.g. certain vehicle operations. It allows estimating the scale of the effect and helps deriving measures to decrease the negative consequences on the system performance. The method shows especially an efficient way to extract this information in a huge amount of existing real road data.

REFERENCES

- Åkerstedt, T., & Gillberg, M. (1990). *Subjective and objective sleepiness in the active individual*. International Journal of Neuroscience, 52, 29-37.
- Attwood, D. (2014). *Vehicle-based driver interlocks: After 50 years of research, can we now design an on-board, real-time system to detect abnormal driving behavior?* SAE Technical Papers, 1.
- Bouchner, P., Piekník, R., Novotný, S., Pěkný, J., Hajný, M., & Borzová, C. (2006). *Fatigue of car drivers - detection and classification based on the experiments on car simulators*. WSEAS Transactions on Systems, 5(12), 2789-2794
- Dingus, T., Hardee, H., & Wierwille, W. (1987). *Development of models for on-board detection of driver impairment*. Accident Analysis & Prevention, 19(4), 271-283.
- Dong, Y., Hu, Z., Uchimura, K., & Murayama, N. (2011). *Driver Inattention Monitoring System for Intelligent Vehicles: A Review*. IEEE Transactions in Intelligent Transportation Systems, 12(2), 596-614.
- European Car Assessment Programme (2014). *2020 Roadmap*. Retrieved from www.euroncap.com.
- Friedrichs, F., & Yang, B. (2010). *Drowsiness monitoring by steering and lane data based features under real driving conditions*. European Signal Processing Conference, 209-213.
- Friedrichs, F., Hermannstädter, P., & Yang, B. (2011). *Consideration of Influences on Driver State Classification*. 2nd Conference on Driver Distraction and Inattention (DDI), Gothenburg.
- Goodwin, A., Kirley, B., Sandt, L., Hall, W., Thomas, L., O'Brien, N., & Summerlin, D. (2013). *Countermeasures that work: A highway safety countermeasures guide for State Highway Safety Offices*. Washington, DC: National Highway Traffic Safety Administration.
- Horne, J., & Reyner, L. (1999). *Vehicle accidents related to sleep: A review*. Occupational and Environmental Medicine, 56(5), 289-294.
- Horne, J., & Reyner, L. (1995). *Sleep related vehicle accidents*. British Medical Journal, 310(6979), 565-567.
- Langwieder, K., Spornier, A., & Hell, W. (1994). *Struktur der Unfälle mit Getöteten auf Autobahnen im Freistaat Bayern im Jahr 1991*. München: HUK-Verband, Büro für Kfz-Technik.

- National Highway Traffic Safety Administration (2010). *Traffic Safety Facts: Distracted Driving 2009*. Washington, DC: NHTSA's National Center for Statistics and Analysis.
- Platho, C., Pietrek, A., & Kolrep, H. (2013). *Erfassung der Fahrermüdigkeit*. Berlin: Bundesanstalt für Straßenwesen.
- Royal, D. (2002). *National Survey of Distracted and Drowsy Driving Attitudes and Behavior*. Washington, DC: National Highway Traffic Safety Administration, 1.
- Wang, J.-S., Knipling, R. R., & Goodman, M. J. (1996). *The Role of Driver Inattention in Crashes; New Statistics from the 1995 Crashworthiness Data System*. 40th Annual Proceedings of the Association for the Advancement of Automotive Medicine, Vancouver.

THE POTENTIAL FOR ADAPTIVE SAFETY THROUGH IN-VEHICLE BIOMEDICAL AND BIOMETRIC MONITORING

Stephen A. Ridella

Julie J. Kang

National Highway Traffic Safety Administration
USA

Satoshi Kitazaki

University of Iowa
USA

Paper Number15-0377

ABSTRACT

A 2009 study by the National Highway Traffic Safety Administration identified certain medical conditions as contributing factors in crash causation (Hanna 2009). It was found that about 1.3% of all crashes included in the National Motor Vehicle Crash Causation Survey (NMVCCS) were precipitated by driver reported medical emergencies and 84% of the drivers in crashes precipitated by medical emergencies experienced seizures (epileptic and others), blackouts (non-diabetic), and diabetic reaction prior to the crashes. Drivers who had crashes precipitated by medical emergencies were more likely to sustain severe injury (28% for incapacitating injury and death for crashes with medical emergency; 11% for crashes without medical emergency). Thus, the premise exists that there may be benefit to identify the driver (and other occupants) of the vehicle as well as monitor their current health status through passive or active methods. This monitoring could take into account chronic conditions (such as bone mineral density) through driver input or through initial vehicle startup measurements which could be used to provide optimal comfort or safety system performance. Additional information about the driver's health or behavioral conditions could be interpreted from blood pressure, heart and respiration rate, blood glucose levels and other physiological parameters and could lead to vehicle intervention in driving and/or alert EMS or police of the impending health condition that may affect driving or cause a crash. This monitoring could be done in many ways such as the recent rapid growth in wearable technology with the ability to pair to apps.

This paper will discuss issues related to driver behavioral and health monitoring and review potential technologies for monitoring and as well as methods for biometric identification. Recent publications on driver crash risk due to chronic and acute health conditions will be summarized. Finally, applications that may be associated with the monitoring will be discussed.

INTRODUCTION

Driver state is an important factor affecting safe driving behavior. Detection and intervention of drowsiness, distraction and drunkenness have been studied by many up to now. The decline of an individual driver's health is another potential cause for a significant proportion of crashes. Researchers as early as 1967 recognized medical impairment as a possible contributor to traffic "accidents" (Waller, 1967). In his report, Waller outlined seven criteria that should be adhered to in a study of this subject, including "there must be a reasonable mechanism for identifying most of the high risk persons, and for doing so early enough to avoid a substantial portion of their accident experience". In a recently released report, the National Highway Traffic Safety Administration reviewed crashes from the National Motor Vehicle Crash Causation Survey (NMVCCS) to determine the "critical reason, which is the last event in the crash causal chain, and concluded that driver was that reason for 94% of crashes in that nationally represented survey (NHTSA, 2015). Of those 94% of crashes (representing over 2 million crashes), seven percent, or 145,000 crashes were attributed to a driver "non-performance error", which could be drowsiness or an acute medical condition. Using the same data set, Hanna (2009) found that 1.3% of all crashes involving light passenger vehicles in NMVCCS were precipitated by a driver's medical emergency and 84% of the drivers in crashes precipitated by medical emergencies had experienced seizures (related to epilepsy and other conditions), blackouts (non-diabetic), and diabetic reaction prior to the crashes (Figure 1). Hanna found that the drivers who had crashes precipitated by medical emergencies were more likely to be more severely injured (i.e. 28% suffered incapacitating injury and death in crashes with a medical emergency compared to 11% in crashes without medical

emergency). Although the ratio of the crashes related to acute medical emergencies was small in the study, there were many other disease-related cognitive and psychomotor impairments (chronic) that may have increased the risk of crash. The Federal Motor Carrier Safety Administration (FMCSA) commissioned a series of studies to determine risk of motor vehicle crash for a variety of medical disorders based on reports and medical expert panel opinions provided by) (Table 1, FMCSA, 2007-2011). The table shows a statistically significant increased risk for a crash based on a driver having one of several disorders. Obstructive sleep apnea (OSA) is shown to have the highest relative risk (1.30-5.72). Another general consideration is that some diseases such as type II diabetes, cardiovascular disease, mild cognitive impairment (MCI) and Alzheimer’s disease are associated with aging. Now that the aging population is growing rapidly in the US, crash rates of occupants with those diseases may see increases despite the recent report by the Insurance Institute for Highway Safety (IIHS) indicating that fatal crash involvement rates of older drivers were decreasing faster than those of younger drivers (IIHS Status Report, 2014). NHTSA’s 5-year plan for traffic safety for older people (NHTSA, 2013) also highlights older drivers’ risk in association with increased medical problems.

Advances in vehicle electronics have made it possible for drivers and passengers to customize their driving experience in many ways. Personalized settings for seat position, as well as heating/cooling/ventilation and entertainment (separated for drivers and passengers) have been available for many years. Recent technological progress in diversity, sensitivity, data capacity and miniaturization of biometric and biomedical instrumentation sensors and devices are enabling the general public to have more real-time access to personal health status as well as enjoy more security for their personal electronic devices. Recent development has resulted in devices that can be embedded anywhere such as clothes, wristbands, watches, vehicle interiors, etc. to detect and report medical information such as body temperature, heart, respiration and perspiration rates, blood glucose and oxygenation levels, and other physiological functions. This data combined with user identification through recognition of fingerprints, iris, facial and/or voice inputs can provide a rapid analysis of a person’s state of health. The availability and use of this information has implications in many markets and significant potential to increase driving comfort and safety when embedded into appropriate algorithms related to vehicle design and performance. For example, a vehicle that senses that the driver has an elevated body temperature and has increased his/her respiration rate significantly may automatically open windows or increase interior ventilation to improve comfort. Also, providing input to the vehicle that the driver is 75 years old and thus has reduced bone mineral density, the vehicle may adjust restraint system parameters to optimize occupant protection in the event of a crash. The same vehicle system could also forward the driver’s vital information to first responders and other health professionals even before they reached the scene of the crash.

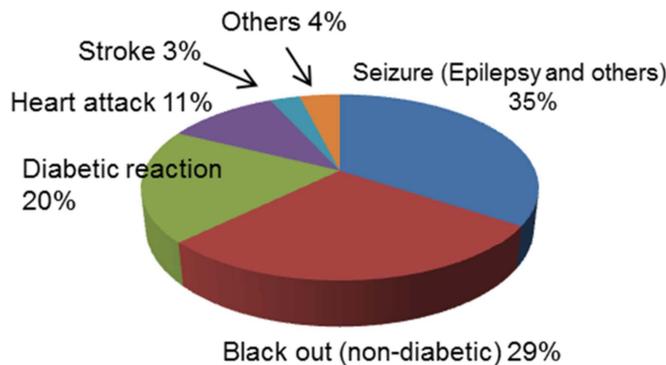


Figure 1 . Type of medical conditions which precipitated crashes (Hanna, 2009).

This paper will present the current state of passive personal identification and monitoring of a person’s health status, as well as the expected developments of such systems in the future. A discussion of how these devices could influence vehicle comfort and safety will be provided through a summary of the technology available or in development, the challenges of integrating the devices to the vehicle, the potential use, accuracy, standardization and privacy of data as well as other policy implications of this technology. Practical examples will be given to exemplify

the market readiness of technology and the potential for diversification of products and applications and their breadth and depth.

HEALTH CONDITIONS THAT MAY AFFECT DRIVING

Medical conditions are generally divided into two categories, chronic and acute problems. Some diseases have either chronic or acute problems and some have both. Effective in-vehicle interventions for these two aspects are different. The driver is at extremely high risk of crash in the case of acute health decline which shows distinctive physiological and behavioral changes. An in-vehicle system that detects a decline and controls the vehicle for an unresponsive driver may help to avoid a crash. On the other hand, chronic health problems develop slowly and degrade cognitive and behavioral performance of the driver over time, resulting in higher general risk of crash. In-vehicle systems that increase the safety margin of the vehicle based on individual's driving ability may minimize the increased risk. The following discussion describes medical conditions that may affect driving.

Table 1.
Summary of motor vehicle crash risk for various medical disorders (FMCSA, various reports 2007-2011, Rizzo, 2011)

Medical Condition	Relative Risk (95% Confidence Interval)
Cardiovascular Disease	1.43 (1.11-1.84)
Diabetes Mellitus	1.28 (1.12-1.47)
Obstructive Sleep Apnea	1.30-5.72 (pooled studies)
Seizure Disorders	1.13-2.16 (pooled studies)
Traumatic Brain Injury	1.32 (0.77-2.25)

Diabetes

The likelihood that a person has diabetes (type II) increases with age. Wild *et al.*, (2004) estimated that the number of people over 65 years old with diabetes in developed countries will increase by 80% between 2000-2030 and the total number of people over the age of 46 with diabetes will be nearly 80 million people by 2030. An acute risk factor for driving is hypoglycemia (it can lead to coma) which can be caused mainly by insulin-dependent Type 1 diabetes, whereas the chronic risk factors for driving are neuropathy (decreased sensation at feet and hands), retinopathy (vision loss), and encephalopathy (cognitive decline), all of which can be caused by both Type 1 and Type II diabetes (Rizzo,2011). However, scientific evidence for drivers with diabetes being at greater risk for crashes is not conclusive. Tregear *et al.* (2007) reviewed and conducted meta-analysis of thirteen studies, comparing crash risk among drivers with diabetes to drivers without diabetes, and found that the risk for crash among drivers with diabetes was 19% greater than the risk among drivers without diabetes (within the range published by FMCSA in Table 1). On the other hand, Tregear *et al.* found no statistically significant evidence to suggest that insulin-treated individuals are at higher risk for crash than individuals with diabetes not being treated with insulin. It seems that diabetes increases the crash risk but contributions from the acute symptoms and the chronic symptoms to the increased risk are not yet known. The American Diabetes Association (ADA) states that diabetes management and education of both patients and health care professionals is the important intervention to the driving risk due to hypoglycemia (American Diabetes Association, 2012).

Obstructive sleep apnea (OSA)

There is strong evidence that the highest relative risk of motor vehicle crash is for OSA, ranging between 1.30 and 5.72 (FMCSA, 2007, Table 1). There is evidence that OSA affects a significant portion of the population regardless whether it is diagnosed or undiagnosed. Hiestand *et al.* (2006) conducted a telephone interview of 1506 US adults using the Berlin questionnaire and found that 26% of respondents met the criteria indicating a high risk of OSA. NHTSA estimated that 1.4% of total crashes and 1.75% of fatal crashes were related to sleepiness (NHTSA, 1985). However, Leger (1994) suggested that NHTSA's estimation was underestimated and provided a new estimation of sleep-related crash rates as 41.6% of total crashes and 36.1% of fatal crashes. Because of potential under-reporting and inability to determine post-mortem that drowsiness or micro-sleep episodes contributed to the crash, there is no clear-cut estimation of crash rate related to OSA. However, considering the high risk factors which were estimated

in strictly controlled driving simulator studies, OSA should be one of the conditions that can contribute to a risk of a vehicle crash.

Another important aspect of OSA is the fact that the disease remains undiagnosed in many individuals. Hiestand et al. (2006) concluded that the prevalence of OSA in the US was estimated to be between 5% and 10%, but only 1 in 10 of those with OSA were adequately diagnosed and screened for Continuous Positive Airway Pressure (CPAP) treatment. Untreated OSA can cause daytime somnolence, cognitive impairment, loss in work productivity with a typical symptom of microsleep and increase risk of motor vehicle crashes. It should be also noted that some sleepy drivers are not aware of their impaired status, possibly because of related cognitive impairment of an altered frame of reference for fatigue (Rizzo, 2011).

Effects of OSA on driving performance have been investigated in driving simulator experiments by some researchers. It was found by Paul et al. (2005) and Boyle et al. (2008) that untreated OSA patients showed greater variation in steering, lane position and TLC (Time-to-Lane-Crossing) during micro-sleep episodes, and degree of driving performance decrement was correlated with microsleep duration, particularly on curved roads. Risser and Ware (1999) found that untreated OSA patients demonstrated increased lane position variability and road departure incident which were positively correlated with frequency and duration of attention lapses (sleeps). Drowsy driver detection and alert systems are commercially available now. However, it remains unknown to what extent the current technology detects the critical state of drivers with OSA featured by frequent occurrence of micro-sleep episodes and to what extent the technology can mitigate the risk of crash.

Other Disease Conditions

In addition to insulin dependent diabetes (typically Type I) and OSA, other diseases such as cardiovascular disease, seizure disorders and traumatic brain injury can also expose the driver to an increased crash risk due to acute symptoms (e.g.: heart attack and stroke, epileptic seizure). With cardiovascular disease, there are increasingly more middle-age and older drivers being treated for symptoms associated with atrial fibrillation and congestive heart failure. Also, conditions associated with respiratory health such as Chronic Obstructive Pulmonary Disease and asthma are becoming increasingly common in the adult population. These conditions can turn from chronic to acute without warning. Each acute health decline is associated with distinctive physiological and behavioral changes from the normal condition and it is important to have the ability to detect such declines of the occupant while in the vehicle to activate an in-vehicle intervention.

More long-term degenerative conditions include osteoporosis and mild-to-severe cognitive degeneration such as Alzheimer's disease. Ridella et al. (2012) found that osteoporosis or poor bone quality was the most significant contributing factor to injury, specifically, incidence of rib fractures, in older occupants involved in an injury-producing car crash. These crashes typically involved lower crash speeds than did crashes involving younger injured drivers and occupants. While treatable with many medications, bone quality continues to diminish with age with evidence that the pace of bone loss is more significant in women than men. Cognitive impairments are more difficult to determine in real-time without an adequate baseline or history for comparison.

REVIEW OF CURRENT TECHNOLOGIES USING DRIVER MONITORING

Occupant Identification

A first step in the process of monitoring is passive, non-invasive identification of the driver and perhaps passengers. The driver's identification would be useful in a host of different applications. Establishing identity may allow for the vehicle to create a baseline of the occupant's health status that can be used in current as well as future driving tasks. Algorithms may be developed for the vehicle to learn how it is driven in certain situations and the associated physiological measurements specific to that person. Also, there may be situations where several drivers share a single vehicle such as the use of a family car, where there may be a range of driving abilities. A teenager who is enrolled in a graduated licensing program may have certain vehicle restrictions put upon them whereas other members of the family may enjoy the full privileges of driving. Therefore, based on the driver's identity, a monitoring system would adjust the vehicle's abilities or monitor highly complex driving tasks more closely.

In addition, a history of the driver's performance would also lend insight into the longitudinal data for comparative purposes. This information could be used to determine long-standing trends in performance such as declining mental capacities (e.g. Alzheimer's or dementia), health issues, changes in driving performance, etc.

Technology to determine identity has grown exponentially in the past few years due to consumer demand for greater security of data. Much of this technology has the ability to be adapted to the vehicle environment or brought in through portable devices. Camera-based technology can be used for identification through facial recognition or iris analysis software. Fingerprint, vascular pattern and voice-recognition scanning are other methods of uniquely identifying an individual. An alternative to passive identification is the driver's userprovided information through a key card, implantable radiofrequency ID or RFID or other device that uniquely identifies an individual.

Behavioral Monitoring (DrIIVE)

Driver Monitoring of Impairments and Inattention using Vehicle Equipment (DrIIVE) is a current NHTSA project that uses driver monitoring data. DrIIVE is focused on the development of an algorithm that can accurately identify and distinguish among different forms of inattention or impaired driving including alcohol-impaired, drowsy, and distracted driving. DrIIVE determines driver behavior data from vehicle measures such as steering and pedal inputs, lane variability, and compares signatures of normal driving with impaired driving. The goal is to use the DrIIVE algorithm to identify and evaluate the effectiveness of driver monitoring countermeasures on impaired driving behavior.

Alcohol Impairment (DADDS)

In 2008, NHTSA launched a cooperation program to develop in-vehicle technology that could accurately, precisely, and reliably measure a driver's blood alcohol concentration in a non-invasive way in a very short time. (Monk, 2012). Now in its second phase, two subsystems have been developed and are being integrated into a research vehicle for further testing. One system is breath based and continually samples the area around the driver for alcohol and carbon dioxide through an infra-red sensor whose measurements can be converted into a blood alcohol concentration. A second subsystem is touch-based and can measure the absorbed near infra-red light in a person's finger and derive an alcohol concentration. These systems have the dramatic potential to reduce crashes and fatalities involving drunk drivers by denying the driver the ability to start and drive the vehicle.

Physiology/Health Monitoring

Monitoring driver health should be non-invasive and passive. Both in-vehicle and wearable technology have been developed, however, only wearable technology has been commercialized to date. As sensor technology has become smaller and less expensive, a vast array of sensor applications have been developed or published. Ford Motor Company, working with a restraint and sensor supplier, developed a prototype vehicle to measure a variety of physiological signals (Watson et al, 2011). Figure 2 below indicates that they were looking at both comfort (temperature difference) and real-time physiology (heart and respiration rate). They indicated that the signals could be integrated for use in assessing driver performance as a function of wellness, workload, and stress. Ford also has demonstrated an in-vehicle glucose monitoring system that could detect a driver's possibility for a diabetic episode (Ford, 2011).

Demonstrations of blood oxygenation measurements using a variant of typical finger-tip pulse oximeters, Meditech 2011), blood pressure and bone mineral density bring more information about driver/occupant health into the vehicle. The BOSCOS (BOneScanning for Occupant Safety, Hardy et al, 2005) project created an in-dashboard ultrasound sensor that could deduce bone strength based on measurements taken from the distal third of a finger when inserted into the device.

Wearable devices or mobile human health monitoring is a maturing area of health awareness particularly in the sports medicine market. Miniaturized electronics or MEMS technology has allowed for creation of wearable wrist bands, head bands, even undergarments that are capable of accurately measuring heart rate, respiration, sweat production, etc. These devices are usually coupled to a portable electronic device application to record daily exercise results and associated physiological responses. The applications have algorithms to detect medical issues and

performance progress with the person wearing the technology as well as determine stress or anxiety levels. The cost of these devices has dropped dramatically in the past several years with many devices selling in the \$50-\$150 range. The sophistication, accuracy and reliability of these devices is steadily improving such that some diagnosis of heart and respiratory health may be deduced from the signals rather than just the instantaneous rates that are reported by the devices.

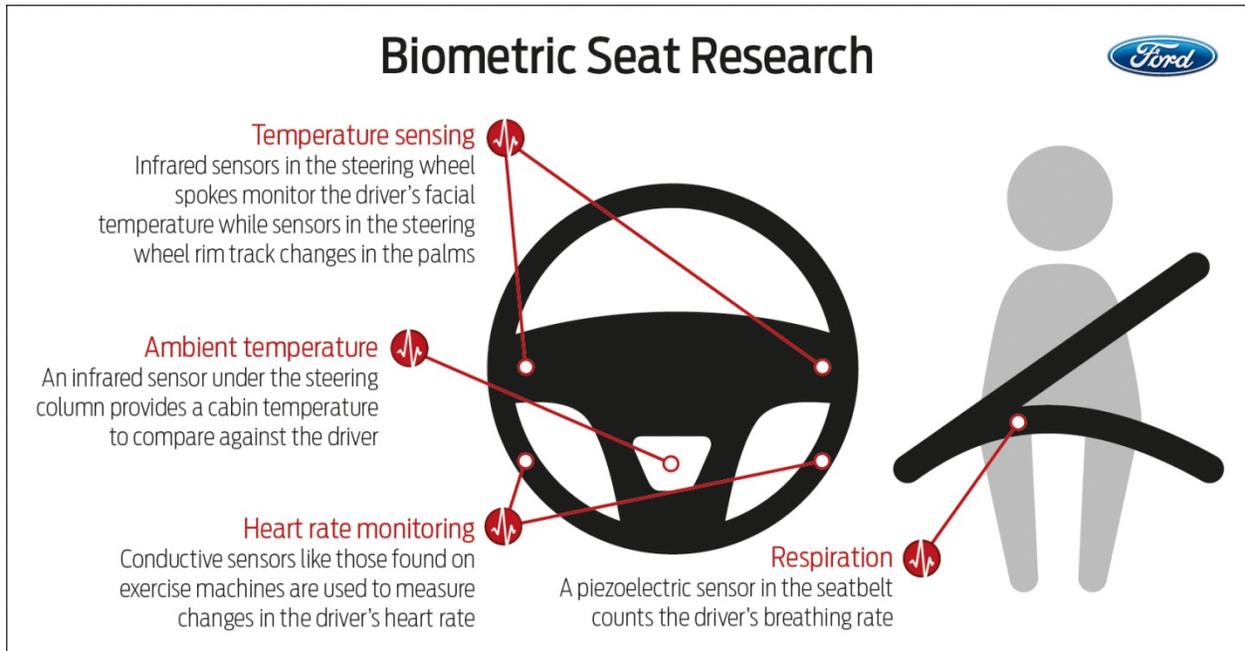


Figure 2. Examples of prototype health monitoring sensors embedded in vehicle components (Wired , 2011)

Establishing a Baseline

Another new challenge within the scope is connection and integration of monitoring the state of the driver in the home (off board) and in the vehicle (on board). Monitoring a person's life in the home through so called "Smart Home" or "Wearable sensors" is becoming important for "Aging in Place" or "Tailor-made Health Care" and it is actively being tested by research projects for its effectiveness and clarification of system and social requirements. However, there are no systems / research projects focusing on integration of on and off board driver monitoring. Early signs of an acute health problem may be detected in the home before driving and such data can be brought into the vehicle to provide the in-vehicle monitoring system with an initial parameter set to raise the sensitivity of detection. For example, lower glucose level data or poor sleep quality data measured in the home may be used in the vehicle for earlier detection of hypoglycemia and frequent micro-sleep occurrence (Table 2). Driver state monitored in the home and in the vehicle can also make use of cause-effect relationships. For example, failure to take medicines in the home may increase the possibility of occurrence of hypoglycemic episodes or epileptic seizure behind the wheel. Such behavioral data measured in the home can also be brought into the vehicle to enhance or change the in-vehicle intervention strategy (e.g. failure to take anti-seizure medicine locks the ignition). Integration of on- and off-board driver monitoring will increase accuracy of detection of drivers' health problems and strengthen the intervention strategy to avoid crash.

Another example of driving performance data being brought into the off-board network involves continuous monitoring of the driving environment. Driving includes many complex and parallel cognitive and physical tasks under certain levels of stress that may magnify certain aspects of a driver's chronic health problems (e.g. cardiovascular diseases). Therefore, a driver's state measured in the vehicle can be brought into a hospital or a smart home to be integrated with other behavioral and physiological data measured in the home to diagnose the chronic health state with better accuracy or in an earlier phase.

IMPLEMENTATION/APPLICATION

Evolution of the current driver state monitoring technology and integration with vehicle-embedded biometric sensors or wearable sensors with wireless connection to the vehicle (on-board monitoring) should be able to detect some or all of the acute health declines in real time. Candidates for on-board biometric sensors for each of the driver's acute health decline are summarized in Table 2. Most of the biometric sensors for on-board monitoring shown in Table 2 are still under research or development. Applicability of these sensors to the detection of acute symptoms needs to be investigated with consideration for the cost and the user acceptance.

If the driver-state monitoring can detect early signs of a decline while the driver is still conscious, a multilayered intervention strategy using information and assistive vehicle control could be taken. When the driver fails to take actions to avoid a crash or if the system fails to detect early signs of the decline due to too rapid decline, future autonomous vehicle technology will be the key to avoiding crashes by bringing the vehicle safely to a stop on the hard shoulder of the road for the driver who is likely to be unconscious (i.e. Autonomous Emergency Stop System or AESS, Shunk, 2009, Nissan, 2013). Stopping a vehicle in busy traffic or high speed traffic could induce additional crashes involving other vehicles. Vehicle-to-vehicle communication to broadcast the emergency signal to surrounding and following vehicles could be included in the AESS. Accuracy of detection also needs to be high. Integration of environmental sensors with driver-state monitoring could be considered so that the system activates the Autonomous Emergency Stop System when the driver is in a health decline and the risk of crash is imminent (Figure 3).

The contributing factors and co-morbidities associated with the most common injuries also point to interventions that could benefit the older occupant. While knowing age of a driver or occupant may help in some driving task assistance, it is less of an indicator of overall health. Sensing occupant bone quality can lead to real-time adaptive restraint systems that lower the loads on the poorer quality bones of older or less healthy occupants and could help reduce incidence of rib injuries. Newer technologies such as 4-point belt systems and inflatable seat belts also help to reduce or distribute chest loading (Ridella, 2012).

Driver-state data measured in the vehicle can be also brought out of the vehicle to the off-board network. Advanced Automatic Collision Notification System (AACN) is an example that automatically sends notification of a collision event together with vehicle and driver-state data from the pre-crash phase to a hospital so that the ER will have sufficient preparation time prior to arrival of the casualties.

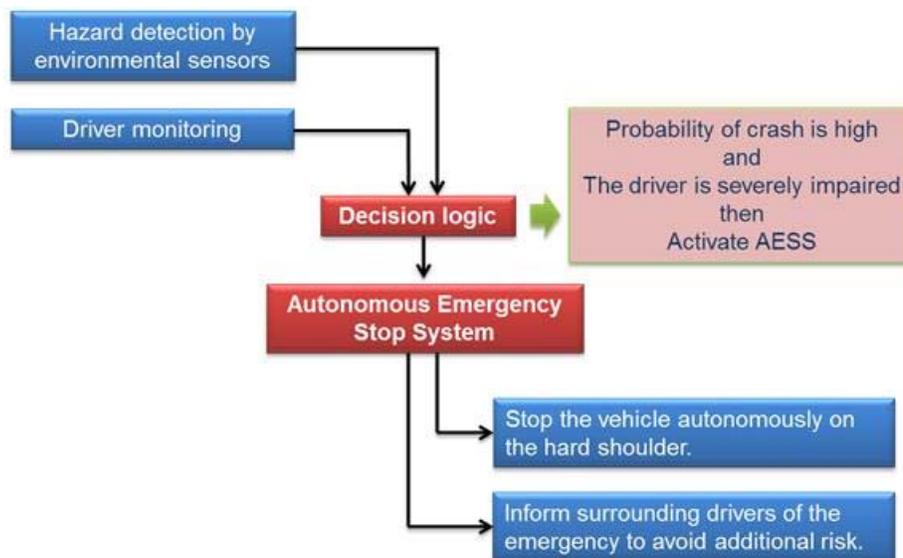


Figure 3 . Functional flow of Autonomous Emergency Stop System.

ISSUES

Privacy, protection and malicious intent

Since the devices and other instrumentation in the vehicle or on the occupant are measuring, monitoring, transmitting, and/or recording health status information, concerns arise regarding the protection and privacy of this data. The Health Insurance Portability and Accountability Act of 1996 (HIPAA) sets rights for an individual's health information and prescribes rules and limits over who can review and receive that information. This rule applies to any form of the information whether it is oral, written or electronic. The information that is envisioned to be collected may not necessarily be covered by HIPAA depending on how it is collected and used. For example, it may be that the data is only used in a non-identifiable, real-time manner to inform the vehicle that the occupant is drowsy or has an abnormal condition. In other cases, when linked to driver/occupant identification technology, the data is personal and possibly subject to HIPAA. It could also be recorded and/or transmitted depending on application if there is a need to interpret the data in real-time by a health professional or to review the data by law enforcement at a later date should an event occur. HIPAA rules are very specific about cases such as this, however, it is likely an interpretation of HIPAA in the vehicle environment may need further review as this technology develops.

Another consideration is the malicious use of the health information data. Security of the data, whether stored in the vehicle, in a portable device or other form needs to be protected from other sources that could use or manipulate the data in harmful ways. For example, a baseline health information data set could be altered either through direct intervention or electronic methods such that an abnormal event would not be detected by the vehicle or software. An intervention or denial of service could occur during driving, thus making the system inoperative. There are endless scenarios that could occur depending on sophistication of the system. Current cybersecurity research regarding other vehicle communications should include the possibility of how the health and driver monitoring activities can be protected from mischievous intent.

Performance

Performance of such a system would require extensive testing for reliability, repeatability and reproducibility. In his short article nearly fifty years ago, one of Waller's criteria for monitoring for physical impairment was that "few persons of low accident risk should be falsely categorized in the high accident risk group". That is, the number of false-positive identifications should be minimized and conversely, the number of true-positives should be maximized. It is imperative that these devices are calibrated properly or can be re-calibrated based on manual feedback from the driver/occupant or perhaps periodically from a software "push" through an application or other connected technology whether portable or vehicle-based.

Cancelable

Finally, the driver and/or other occupants may reserve the right to cancel or not participate in real-time identification and monitoring. This may apply when a vehicle is not driven by an owner or designated driver for whom baseline data exists. Also, rental cars may either not be equipped with such devices or a driver may opt out from participation.

SUMMARY

The paper discussed the premise that identification of a vehicle driver (and/or occupants) as well as monitoring their health, mood or behavioral status while driving, may have significant value for safety. It is documented that health conditions may contribute to increased crash risk and that those with conditions have poorer outcomes should a crash occur. By monitoring a driver's health status in real time, possibly comparing to a baseline value, acute conditions may be detected and a warning, intervention, or other countermeasure may be applied. There is abundant technology in development as evidenced by manufacturers' documented research. Also, both traditional and non-traditional automotive suppliers have been involved in the early vehicle-based technology research, however, the sports market has dominated the wearable monitoring device development and production. Research projects in driver behavior, alcohol detection and a host of other technologies in development may lead to new advances in safety as the population of driver's age and people are more aware of their own health status.

REFERENCES

1. American Diabetes Association. (2012). Diabetes and Driving. *Diabetes Care*, Vol.35, *Supplement 1*, ppS81-S86.
2. Badugu,R., Lakowicz,J.R., and Geddes,C.D. (2004). Noninvasive continuous monitoring of physiological glucose using a monosaccharide-sensing contact lens. *Analytical Chemistry*, Vol.76, No.3, pp610-618.
3. Baig,M.M., Gholamhosseini,H., and Connolly,M.J. (2013). A comprehensive survey of wearable and wireless ECG monitoring systems for older adults. *Medical & Biological Engineering & Computing* Vol. 51, No. 5, pp485-495.
4. Balanou,E., Gils,M.V., and Vanhala,T. (2013). State-of-the-art of wearable EEG for personalized health applications. Presentation material at 10th International Conference on Wearable Micro and Nano Technologies for Personalized Health. Retrieved February 20, 2014 from <http://phealth2013.eu/uploads/keepitsimple/E.Balanou.pdf>
5. Boyle,L.N., Tippin,J., Paul,A., and Rizzo,M. (2008). Driver performance in the moments surrounding a microsleep. *Transportation Research Part F* 11, pp126–136.
6. Casson,A.J., Yates,D.C., Smith,S.J.M., Duncan,J.S., and Rodriguez-Villegas,E. (2010). Wearable Electroencephalography. *IEEE engineering in medicine and biology magazine* (2010, May/June), pp44-56.
7. FMCSA. (2007) Expert Panel Recommendations: Cardiovascular Disease and Commercial Motor Vehicle Driver Safety. Evidence Report. Presented to FMCSA, April 2007.
8. FMCSA. (2007). Expert Panel Recommendations: Seizure Disorders and Commercial Motor Vehicle Driver Safety. Evidence Report. Presented to FMCSA, November, 2007.
9. FMCSA. (2007) Expert Panel Recommendations: Obstructive Sleep Apnea and Commercial Motor Vehicle Driver Safety. Evidence Report. Presented to FMCSA, November, 2007.
10. FMCSA. (2009) Expert Panel Recommendations: Traumatic Brain Injury and Commercial Motor Vehicle Driver Safety. Evidence Report. Presented to FMCSA, March, 2009.
11. FMCSA. (2011) Expert Panel Recommendations: 2010 Update Diabetes and and Commercial Motor Vehicle Driver Safety.Evidence Report. Presented to FMCSA, May ,2007.
12. Ford (2011, May). *Ford In-Car Health and Wellness Solutions*. Retrieved February 20, 2014 from <http://medicaldesign.com/site-files/medicaldesign.com/files/archive/medicaldesign.com/Medtronic-fact-sheet-HR.pdf>
13. Hanna, R. (2009). The contribution of medical conditions to passenger vehicle crashes. *DOT Report*, DOT HS 811 219.
14. Hardy, R.N., Watson, J.W., Cook, R., Zioupos, P., Forrester, B., Frampton, R., Page, M., Kennedy, A., Peach, S., Sproston, P. Development and Assessment of a Bone Scanning Device to Enhance Restraint Performance. In: Proceedings of the 19th International Conference on the Enhanced Safety of Vehicles. Washington, D.C.
15. Hiestand,D.M., Britz,P., Goldman,M. and Phillips,B. (2006). Prevalence of symptoms and risk of sleep apnea in the US population. *CHEST Journal*, Vol.130., No.3, pp780-786.

16. Insurance Institute for Highway Safety. (2014). Status Report, Vol. 49, No. 1.
17. Iowa State University (2005). *Smart Home Lab*. Retrieved February 20, 2014 from <http://smarhome.cs.iastate.edu/Demos.html>
18. Leger,D. (1994). The cost of sleep-related accidents: a report for national commission on sleep disorders research. *Sleep*, Vol.17, No.1, pp84-93.
19. Meditech (2011). *Finger Pulse Oximeter*. Retrieved February 20, 2014 from <http://www.meditech.com.cn/Finger-Pulse-Oximeter>
20. Misugi,K., Kanamori,H., and Atsumi,B. (2003). Toyota's program for universal design in vehicle development. Universal design for the Toyota "RAUM", *Toyota Motor Corporation, "Raum" press information*
21. Monk,C (2012, August). *Driver Alcohol Detection System for Safety (DADSS)*. Retrieved February 20, 2014 from <https://secure.hosting.vt.edu/www.apps.vtti.vt.edu/PDFs/NDRS-2012-presentations/DADSS%20VTI%20Naturalistic%20Conference%20Presentation%202012%20v3.pdf>
22. NHTSA (1985). Transportation-related sleep research. CAR files study. *Report to the Senate Committee on Appropriations and House Committee on Appropriations*. Washington,DC., US Department of Transportation.
23. NHTSA (2013). Traffic safety for older people - 5-year plan. DOT HS 811 837.
24. NHTSA (2015). Critical Reasons for Crashes Investigated in the National Motor Vehicle Crash Causation Survey. DOT HS 812115.
25. Nijssen,T.M.E., Arends,J.B.A.M., Griep,P.A.M., and Cluitmans,P.J.M. (2005). The potential value of three-dimensional accelerometry for detection of motor seizures in severe epilepsy. *Epilepsy & Behavior*, Vol. 7, pp74-84.
26. Nissan (2013, August). *Nissan Autonomous Drive: Emergency Stop*. Retrieved February 20, 2014 from <http://nissannews.com/en-US/nissan/usa/releases/nissan-announces-unprecedented-autonomous-drive-benchmarks/videos/nissan-autonomous-drive-emergency-stop>
27. Paul,A., Boyle,L.N., Tippin,J., and Rizzo,M. (2005). Variability of driving performance during microsleeps. *Proceeding of the third International driving Symposium on Human Factors in Driver Assessment, Training and Vehicle Design*. pp18-24.
28. Pautan.org (2011, May). *Ford Develops Heart Rate Monitoring Car Seat*. Retrieved February 20, 2014 from <http://paultan.org/2011/05/26/ford-develops-heart-rate-monitoring-car-seat/>
29. Positive ID (2013). *Non-Invasive Breath Glucose Detection Device*. Retrieved February 20, 2014 from http://www.positiveidcorp.com/products_breathglucosedetection.html
30. Quick,D. (2009, January). *Implantable Sensor Simplifies Blood Pressure Readings*. Retrieved February 20, 2014 from <http://www.gizmag.com/tiny-sensor-simplifies-blood-pressure-readings/10846/>
31. Ridella, S.A., Rupp, J.R., and Poland, K. Age-Related Differences in AIS 3+ Crash Injury Risk, Types, Causation and Mechanisms. In: *Proceedings of the 2012 International IRCOBI Conference on the Biomechanics of Injury*, pp. 43-59, Dublin, Ireland, September 2012.

32. Ridelust (2011, January). *Ferrari Working On Driver Monitoring Technology*. Retrieved February 20, 2014 from <http://www.ridelust.com/ferrari-working-on-driver-monitoring-technology/>
33. Risser,M.R. and Ware,J.C. (1999). Driving simulation with EEG monitoring in normals and obstructive sleep apnea patients. *Sleep*, Vol. 23, No.3, pp393-398.
34. Rizzo,M. (2011). Impaired driving from medical conditions: A 70-year-old man trying to decide if he should continue driving. *JAMA*, March 9, 2011, Vol.305, No.10, pp1018-1026.
35. Shunk,C. (2009, June). *BMW Working On System That Stops Vehicles During A Medical Emergency*. Retrieved February 20, 2014 from <http://www.autoblog.com/2009/06/05/bmw-working-on-system-that-stops-vehicles-during-a-medical-emerg/>
36. Skubic,M., Alexander,G., Poppescu,M., Rantz,M., and Keller,J. (2009). A smart home application to eldercare: Current status and lessons learned. *Technology and Health Care* 17, pp183-201.
37. Tregear,S.J., Rizzo,M., Tiller,M., Schoelles,K., Hegmann,K.T., Greenberg,M.I., Phillips,B. and Anderson,G. (2007). Diabetes and motor vehicle crashes: a systematic evidence-based review and meta-analysis. *Proceedings of the Fourth International Driving Symposium on Human Factors in Driver Assessment, Training and Vehicle Design*. pp343-350.
38. University of Florida. *The Gator Tech Smart House (GTSH)*. Retrieved February 20, 2014 from <http://www.icta.ufl.edu/gt.htm#1>
39. University of Rochester Medical Center (2014). *University Creates Medical "Smart Home" To Study Future Health Technology*. Retrieved February 20, 2014 from <http://www.urmc.rochester.edu/news/story/index.cfm?id=-103>
40. Waller, J.A. (1967). The Contribution of Medical Impairment to Traffic Accidents. In: *The Prevention of Highway Injury*, Selzer, M.L., Gikas, P.W. and Huelke, D.F. (eds). , Highway Safety Research Institute, Ann Arbor, MI.
41. Watson, T. , Krause J., Le, J. and Rao, M.K. (2011). Vehicle Integrated Non-Intrusive Monitoring of Driver Biological Signals. SAE Paper No. 2011-01-1095., SAE International, Warrendale, PA.
42. Wild,S., Roglic,G., Green,A., Sicree,R. and King,H. (2004). Global prevalence of diabetes; Estimates for the year 2000 and projections for 2030., *Diabetes Care*, Vol.27, No.5.
43. Wired. (2012). <http://www.wired.com/2012/06/ford-workload/>

Table 2.
Monitoring technologies and application to driver's state

Monitoring modalities and sensors		Crash related driver state							Other driver states
		Acute health problems							
Drowsiness Distraction		Diabetic hypoglycemia	Frequent microsleeps with OSA	Heart attack	Stroke	Epileptic seizure			
On-board monitoring	Vehicle behavioral signals	←	←	←	←	←	←	←	
	Driver behavioral signals	←	←	←	←	←	←	←	
Off-board monitoring	Biometric sensors	←	←	←	←	←	←	←	
	Driver monitoring camera	←	←	←	←	←	←	←	
Smart Home: Home embedded sensors (No connection to in-vehicle monitoring yet.)	Vehicle embedded biometric sensors	←	←	←	←	←	←	←	
	Wearable sensors	←	←	←	←	←	←	←	
		<ul style="list-style-type: none"> -Breath Alcohol -Ignition Interlocks* -Spectrometry by Ignition touch* -Spectrometry by IR sensor* (Morc, 2012) 	<ul style="list-style-type: none"> -Breath glucose monitor (Positive ID, 2013) 	<ul style="list-style-type: none"> -ECG on seat* (Pautan, 2011) 	<ul style="list-style-type: none"> -Wearable EEG monitor -Wearable blood pressure sensor (Quick, 2009) 	<ul style="list-style-type: none"> -Wearable EEG monitor -Wearable ECG monitor (Baig, et al., 2013, Meditech, 2011) -Wearable blood pressure sensor 	<ul style="list-style-type: none"> -Wearable EEG monitor -Body acceleration monitor (Nijssen et al., 2005) 	<ul style="list-style-type: none"> -Heart rate: temperature on steering wheel for stress* -Respiration on seat belt for stress* (Ford, 2012) -EEG on headrest for stress* (Rudelust, 2011) -Skin temperature by IR for climate control* (Misugi et al., 2003) 	

Application of Biologically Inspired Visual Information Processing in Affective Driver Status Monitoring

IL SONG HAN
WOO-SUP HAN

Korea Advanced Institute of Science and Technology
Korea, Republic of

Paper Number 15-0150

ABSTRACT

Recent times have seen an increased interest in technologies of driver assistance. Understanding the driver's current status is crucial for the implementation of Advanced Driver Assistance System (ADAS) and Driver Status Monitoring (DSM). Emotional factors such as anger have been long attributed to aggressive driving behaviours and increased likelihood of road accidents. Therefore, being able to accurately detect the affective states of the vehicle occupant will be critical for enhanced safety and comfort.

In this paper, we present a methodology for the evaluation of the emotional states of vehicle drivers. The proposed approach performs an assessment of the emotional states by using combination of biologically inspired visual information processing and neural networks coupled with feedback mechanisms. The system consists of the following stages: (1) biologically inspired image pre-processing; (2) facial feature extraction; (3) multilayer perceptron for classification; and (4) feedback mechanism. The system has been preliminary validated by using data available from Japanese Female Facial Expression (JAFFE) database. Four affective states were identified and tested, which includes anger, sadness, and happiness. Subsequent tests have shown the successful detection rate of 91.3% with test images, and over 70% correct classification in images with Gaussian noises, respectively.

INTRODUCTION

Recently there has been increasing attention towards automobile safety. According to the recent report by the UK government, road deaths have increased by 4% in the first quarter of 2014 compared to the previous year [1]. While the overall figure of fatalities has been under decline for the past decade, thanks to the improvement in medical treatment as well as the greater attention shown to the vulnerable road users such as the pedestrians and motor cyclists, the issue of road safety remains a key area for almost all nations in the world regardless of their wealth, with the fatality figure reaching over 1.2 million in the year 2010. Yet, only 28 nations in the entire world, which account for less than 450 million peoples of 7 billion world populations, are deemed to have sufficient legal measures to protect road users [2].

On the other hand, recent trends have shown that the developed nations are devoting a substantial amount of resources to improve road safety. In Britain, the government has spent £15.1 billion on the prevention of road accidents alone [3]. On a continental level, European Commission has launched a new EU Road Safety programme in 2010 as part of drive to cut the number of road fatalities by half between 2011-2020. All these findings show that the issue of road safety goes beyond the scope of industry.

Improving road safety comes in many different forms. The EU Road Safety programme has been divided into seven sub-sections: improving the education and training of road user; better enforcement of road rules; providing safer transport infrastructures; developing safer vehicles; promoting the use of modern technology to increase road safety; improving emergency and post-injuries services; and increased effort to protect vulnerable road users.

Within this paper, we have identified human emotions as one of the key influencers of driving performance. While most of existing Driver Monitoring System (DMS) focuses on drivers' concentration or drowsiness, another aspect of driver's status that deserves our attention is in emotion. Certain states of emotion, such as anger, and frustration can lead to aggressive and risky driving, which could result in accidents and fatalities. We believe that the accurate detection of emotional states from facial expressions is an important method to analyse and provide feedback to the vehicle driver's concentration level. Through successful recognition of emotional status, such as distraction, anger or frustration, it will be possible to improve the driving conditions and thereby safety of the vulnerable road users such as the pedestrians and cyclists.

We have applied image processing techniques and neural networks in the measurement of affective states based on still facial images in this study. The result of the finding can be used as a basis for future research in the application of affective computing in DMS.

There have been a number of studies which looked to present the feasibility of affective computing for improving road safety, and testing various methods of image processing in affective computing. It is notable that, despite numerous papers in the literature, there have been a minimal number of studies carried out on the use of image processing techniques coupled to neural networks in order to create an intelligent vision system which could evaluate human emotions.

METHODS

Biologically Inspired Orientation Filters

We claim that biologically inspired visual information processing offers a robust method for mimicking the robustness and flexibility of the primary visual cortex. One major strand of knowledge behind our current understanding of the behaviour of the primary visual cortex, an important brain area for vision, comes from the set of experiments by Torsten Hubel and David H. Wiesel [4]. Their experiment consisted of inserting microscopic electrodes into the visual cortex of experimental animals. This was used to read the activity of single cells in the visual cortex while presenting various stimuli to the animal's eyes.

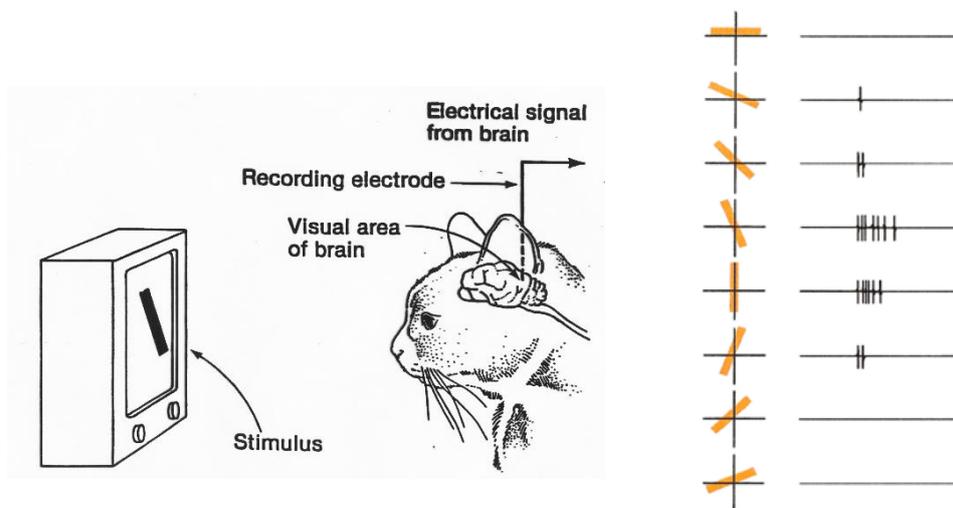


Figure 1. Hubel & Wiesel's Experiment Setup (left), and the responses of the cat's cortex when a rectangular slit of light of different orientations is shown (right)

They discovered that a topographical mapping in the cortex, i.e. that nearby cells in the cortex represented nearby regions in the visual field, i.e. that the visual cortex represents a spatial map of the visual field. Individual cells in the cortex, they found responded to the presence of edges in their region of the visual field. Furthermore, cells were found which would fire only in the presence of a vertical edge at a particular location in the visual field, while other nearby cells responded to edges of other orientations in that same region of the visual field. These orientation-sensitive cells were called "simple cells", and were found all over the primary visual cortex [5][6]

Based on Hubel and Wiesel's experimentation, the biologically inspired visual information processing incorporates the orientation selectivity of simple cell neurons to extract the features of facial images. The behaviour of the simple cells suggest that these cells possessed a patterned receptive field, with excitatory and inhibitory regions so that the cell would activate only if it received input (due to light) in the excitatory portion of its receptive field in the absence of input from the inhibitory portion. This operation is comparable to the operation of edge detection in image processing, which would process an image by spatial convolution with an edge filter [7].

We have simulated these biological operations through the application of multiple orientation filters for the feature processing of the facial expression. Our aim was to develop a filter which could accurately recreate the characteristics of the simple cells in mammalian visual cortex. These filters consist of six filters, each with specific orientation selectivity, yielding outputs of six orientation images for each input image (test image and reference image). Each set of six orientation filters provided orientation-selectivity for 0, 30, 45, 60, 90, and 135, degrees, and represents the receptive field properties of simple (linear) cells in V1.

Thus, the facial images were processed through this simulated simple cell function of the orientation selectivity and the six orientation images would then be combined to provide a resultant orientated image.

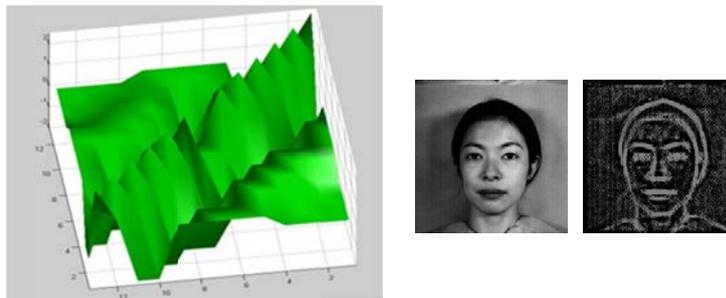


Figure 2. Image filter for 45 degrees orientation selectivity (left), and how the input image (middle) compares to the image processed by the image filter (right)

Affective Computing and Facial Action Coding System (FACS)

The resulting images have been then further processed to focus on the key 'regions of interest' (ROI). For determining the ROI, we have referenced results from contemporary researches in facial affect detection in the fields of affective computing. The detection and processing of facial expression has already been done through various methods such as, optical flow analysis, neural network, and active appearance model [8][9]. Facial affect detection can be done as a stand-alone affective computing system, and it can also be combined or fused alongside other modality (multimodal recognition, such as facial expression + speech pattern analysis or facial expression + hand gestures) to deliver a more robust analysis of the subject's emotional status. In this case, we have made use of Facial Action Coding System (FACS) for the purpose of determining ROI. FACS is one of the products of efforts to taxonomies human facial movements, and can be understood as an index list of facial expressions. Its foundation lies in Paul Ekman's cross-cultural research from the 1960s. His work suggested that affective expressions in face are not culturally determinant, but rather universal [10]. To support his points, Ekman carried out tests on Fore tribesmen in Papua New Guinea. The members of this tribe were not exposed to any contemporary emotional expressions, yet certain emotions – in particular six basic emotions: *anger, disgust, fear, happiness, sadness, and surprise*, were classified in the similar manner [11][12]. Thus, Ekman's study implied that certain facial expressions are biological in origin, and can be categorised with the same criteria.

While types of expressions which could be classified as the basic emotion have been under debate, the idea behind the various theories of emotion is the same. It suggests that the excluding core emotions, which could range between six to ten, all the other emotional states are simply synonyms of these core sets [13]. For our study, we have limited the emotive categories to be tested to four affective states of *happiness, sadness, anger* and *neutral*.

Thus, it is possible to demonstrate almost all facial expressions in terms of the fundamental actions of muscles, otherwise known as Action Units (AU) [14]. Every AU is independent of each other, and as such, a collection of AUs can be used to describe facial muscle movements including description of basic emotions. Ekman have incorporated FACS (Facial Action Coding System) into his research and provided corresponding AUs for each of the basic emotions, in addition to his own interpretation of their meaning. The six basic emotions of Ekman

would be described as shown in Table 1. In the context of affective computing, identification of emotional states using FACS code is known as EMFACS (Emotional facial Action Coding System), and they would concentrate on AU combinations related to emotional states. [15]

Table 1.
Basic Emotions and Corresponding Action Units

Emotion	Action Units (AU)
Happiness	6 (‘cheek raiser’) + 12 (‘lip corner puller’)
Sadness	1 (‘inner brow raiser’) + 4 (‘brow lowerer’) + 15 (‘Lip Corner Depressor’)
Surprise	1 (‘inner brow raiser’) + 2 (‘outer brow raiser’) + 5B (‘Upper Lid Raiser*’) + 26 (‘Jaw Drop’)
Fear	1 (‘inner brow raiser’) + 2 (‘outer brow raiser’) + 4 (‘brow lowerer’) + 5 (‘Upper Lid Raiser’) + 7 (‘Lid Tightener’) + 20 (‘Lip Stretcher’) + 26 (‘Jaw Drop’)
Anger	4 (‘brow lowerer’) + 5 (‘Upper Lid Raiser’) + 7 (‘Lid Tightener’) + 23 (‘Lip Tightener’)
Disgust	9 (‘Nose Wrinkler’) + 15 (‘Lip Corner Depressor’) + 16 (‘Lower Lip Depressor’)

Determining Regions of Interest (ROI)

The actual process of calculating ROI has been done through detection of eyes and subsequent establishment of centre points between the eyes as the reference point through template based object recognition of eyes. From here, relevant AUs can be analysed for both the input image and neutral ‘reference’ image to which it would be compared against, as shown below. (figure!)



Figure 3. (from left) Input image, orientation image, detection of central point between two eyes as the reference location of calculating the ROIs, and how the calculated ROIs appear in the original input image.

Tests show that eye locations were detected successfully in both frontal and side facial image, demonstrating the effectiveness of combination of orientation selective feature extraction and feature based object recognition. This would, in turn, allow for measuring the level of driver attentiveness in road driving

situations. Subsequently, the detection of emotion (happiness, anger, or etc.) was based on the evaluation of processed facial features of selected areas.

Neutral Network

One of the requirements in developing a reliable warning system would be to provide feedback to the driver, whenever the driver would go into an extreme emotional state. To achieve this aim, it was important to detect which emotional state was the subject under, as certain emotive states are more counterproductive to safe driving than the other. However, more crucial aspect was in ensuring that neutral states of the subject was not being misinterpreted as an emotional state of any category. In this regards, an output of 1 (safe / neutral state) or 0 (unsafe / emotional state) would provide a sufficient result. The proposed system is designed to improve the successful detection of safe or unsafe emotional status of a driver, by tuning the ROI salience data encoding with the weight to the neural network input for improving the ordinary neural networks of Multi-Layer Perceptron (MLP).

Our input data would consist of analysed ROIs from the test and the reference images. This would be yielded by the data set of 90 images processed by the multiple orientation filters. The six data extracted from six ROI were applied to the MLP to be classified into four emotional status, i.e. 'Neutral', 'Angry', 'Happy' and 'Sad'. The MLP of two hidden layers is designed to produce 4 outputs, which was implemented using the MATLAB neural network tool box.

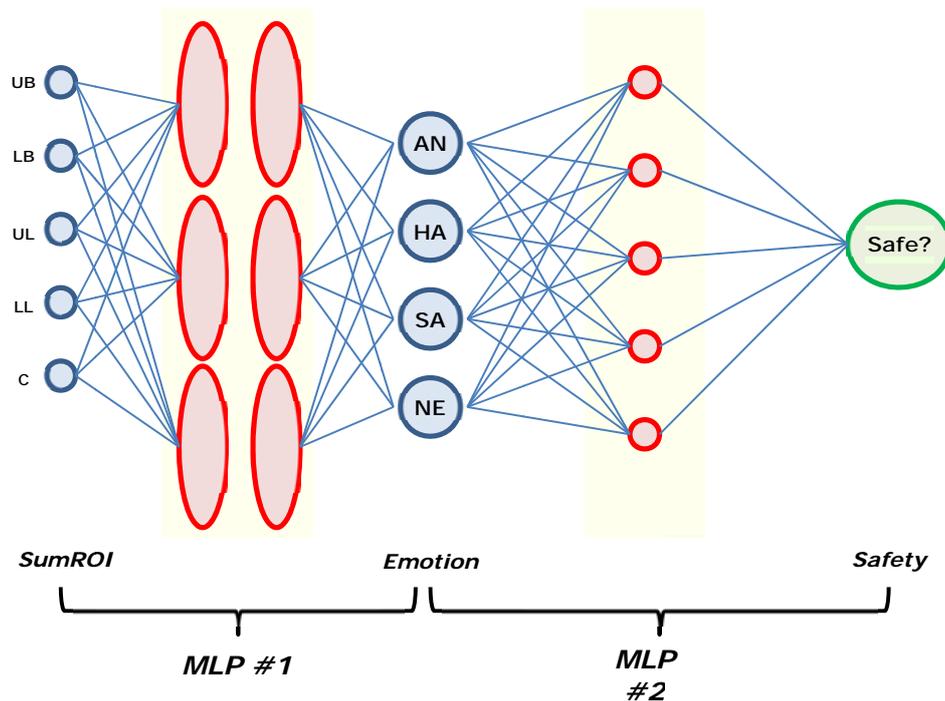


Figure 4: Design of proposed neural network system

During the design phase, the optimum network structure of 5x10x10x4 was determined by iterative trials of various network architectures with the observation of weight value distribution. The image processed by multiple orientation filters is assumed to be consistent in facial direction, size or scale.

The purpose of monitoring the driver's emotional status is attained by detecting the neutral emotional status as *safe* and the other status as *unsafe*, using the MLP of 4x10x5x1. The upper part of networks (4x10x5x1) is trained by an ideal situation, while lower part of networks (5x10x10x4) is trained by the back propagation learning using processed image data set.

Feedback Mechanism

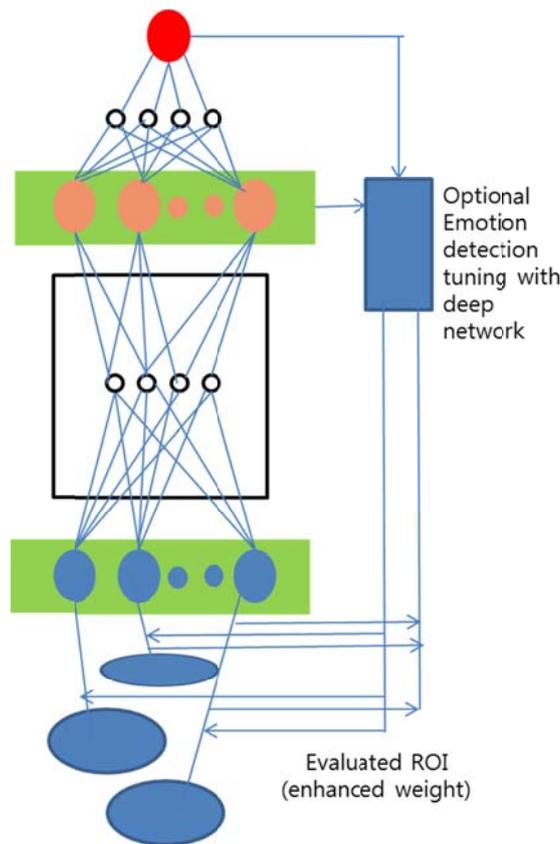


Figure 5. Proposed emotion classification system with weight tuning or learning at an early stage.

We stated in the introduction that our initial purpose was to develop a warning system which could provide an alert to the driver, whenever the driver would go into an extreme emotional state. To achieve this aim, it was important to detect which emotional state was the subject under, as certain emotive states are more counterproductive to safe driving than the other. However, more crucial aspect was in ensuring that neutral states of the subject was not being misinterpreted as an emotional state of any category. In this regards, an output of 1 (safe / neutral state) or 0 (unsafe / emotional state) would provide a sufficient result. The proposed system is designed to improve the successful detection of safe or unsafe emotional status of a driver, by tuning the ROI salience data encoding with the weight to the neural network input for improving the ordinary neural networks of MLP.

In addition to cascaded neural networks of MLP explained previously, the feedback to ROI is applied by the trained upper and lower networks of MLP. In case of failed detection of safety monitoring output, the encoding of ROI values to neural input is subject to the feedback with the lower weight, for the pre-analysed ROI for the largest value of emotional status – which is the main contributor to the detection failure. The feedback rate to ROI applied to neural network is a fixed value and the selection of ROI is controlled by the ROI value and the input status of upper network when feedback is required, for this framework. The method of deep network is scope of investigation by evaluating the effectiveness and relevance to the performance of the target function.

DATA SOURCES

The accuracy of the emotion detection was measured by tests carried out on the JAFFE database, through processing with two feedforward neural networks (multi-layer perceptrons) and the feedback mechanism based on a deep learning concept. The biologically inspired visual information processing showed a significantly high accuracy to emotion recognition without the need for precise matching or complex computation. This was superseded by mimicking the primary function of the simple cell of visual cortex, which provided a degree of robustness by maintaining accuracy even in test images with Gaussian noises.

RESULTS

Based on the data set of 360 images, the overall rate of detection returned over 91.3% accuracy, as shown in the Table 2 below. The proposed method of combining biologically inspired visual information processing with multilayer perceptron also gave a satisfactory performance, maintaining a 70+% accuracy level up to 5% Gaussian noise level. Investigation of deep-learning inspired feedback mechanism also yielded an improved performance, resulting in 27.2% reduction in the occurrence of false-positives.

Table 2.
Result of Biologically Inspired Visual Information Processing & Multilayer Perceptron

Input Class	Output Classification Result				Accurate Classification Rate
	Happy	Angry	Sadness	Neutral	
Happy	81	1	7	1	90.0 % (81 / 90)
Angry	0	90	0	0	100.0 % (90 / 90)
Sad	2	2	78	8	86.7 % (78 / 90)
Neutral	2	0	8	80	88.9 % (80 / 90)

Compared to the comparative studies carried out in affective computing, combination of biologically inspired visual information processing and MLP have outperformed methods based on artificial neural networks (ANNs; 84% accuracy) and combination of Sequential Floating Forward Search (SFFS), Fisher Projection (FP), and K-nearest neighbour algorithm (KNN; 81% accuracy) [16].

In depth review of biologically inspired visual information processing approach revealed certain patterns of interest. It was demonstrated that four emotional states (happiness, anger, sadness, and neutral) can be successfully classified with a small number of variables. Yet there were some level of differences as to how accurately each emotion was classified by biologically inspired visual information processing approach.

Angry emotional state was most likely to be detected with accuracy, with 100.0% successful detection rate, followed by happy emotional state with 90.0% classification accuracy. Sad emotional state had the most inaccurate incidents of classifications, which has also been documented in previous researches [17]. Reason for this difficulty could be attributed in parts to the comparative lack of facial expression in sad emotional state.

Compared to anger, which involves composite movements of eyebrows, competitive movements of eyelids, and major movement in lip areas, sadness would be expressed by movement of eyebrows and small movement to the corner of the lips. Even in real-life situations, detecting expression of subtle sadness is a challenge, and it is possible to infer such factors have been represented in the test results as well.

Further tests were carried out to determine the robustness of biologically inspired visual information processing approach, through introducing Gaussian noise to the images. This was done in order to examine its applicability in real life situations, where the image data is often corrupted. The results of classification with noisy test images showed a reduction of 16.01% in detection rate at 1% noise level, and maintained over 70% level of accuracy even at 10% noise level.

Table 3.
Detection rate (%) of biologically inspired visual information processing method with different level of Gaussian noise

Noise Level	SumROI(G)	SumROI(M)
0.0%	78.7	91.7
1.0%	76.4	79.1
2.5%	73.1	75.7
5.0%	68.8	73.5
10.0%	64.1	65.9

DISCUSSIONS AND LIMITATIONS

In this study, our proposed model for identifying affective states required the presence of a neutral emotive state, which would be used as a reference. While it is theoretically possible to process an individual's emotional status from a single image without any reference, in reality, there is a wide variability to the strength of human expression even within same emotive state. This variability is deemed dependent on the cultural, national, regional and gender difference of the subject [18]. This had also been reinforced during the analysis of datasets: that absolute values of expression within an individual cannot be used as a reliable measure. Further tests and statistical analysis could be carried out in the future as a means to develop an accurate and reliable indicator of emotive states based on input image alone.

Another challenge we have recognised during the experiments was that there are at least two kinds to the expression of emotions – genuine and fake. A fake expression often differs from a genuine expression. For example, in case of happiness, only the zygomatic major muscle, which runs from the cheekbone to the corner of lips, moves in case of fake emotional status. On the other hand, a genuine expression of happiness would involve movement of orbicularis oculi and pars lateralis (eyebrows) as well as zygomatic major. In addition to limited facial muscle movements, fake emotive states are also known to contain a certain degree of asymmetry [19],[20].

On the other hand, for certain emotional states in JAFFE dataset, there were very little expressions present. In comparison to the neutral image, even human vision had difficulty in determining the emotional state without the label.

Further investigation of other facial affection database has shown that this issue isn't unique to JAFFE database, as demonstrated in the figures below.

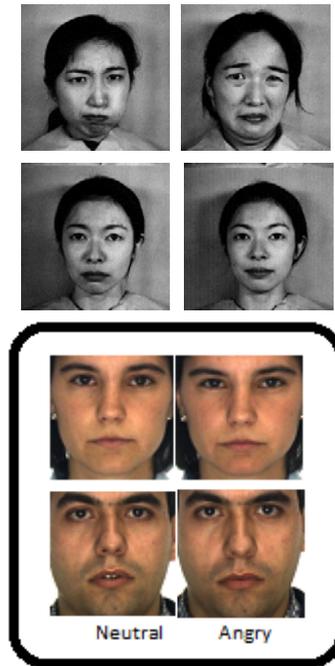


Figure 6. Limitations of dataset

Top Row: Examples of 'fake' emotion – asymmetry is clearly visible in the fake facial expression

Middle Row: Images displaying supposed expression of sadness (left) and neutral (right).

Notice the lack of difference between this image and that of the neutral state (right).

Black Box: Example of angry and neutral emotion from an alternative database.

Without the labels, it would be challenging to determine that left image is demonstrating anger and the right one is of the neutral emotional state.

CONCLUSIONS AND RELEVANCE TO THE SESSION SUBMITTED

We believe this research has produced a foundation from which further studies could be carried out. For example, would be in refining the biologically visual information processing approach. We have recognised that there are at least two kinds to the expression of emotions – genuine and fake. Within the context of driver monitoring system, fake emotions are unnecessary and should be ignored. Further studies carried out with images of emotions in real-life situations could provide a better training for the monitoring system capable of demarcating two categories.

Secondly, we have worked with the assumption of only neutral emotional states allowing for ideal driving situations. This hypothesis is under discussion within affective computing research, and will be given further updates in the coming time.

ACKNOWLEDGEMENT

The research was partly promoted as part of 'Technology Development for Behavior Improvement and Violation Control of High Risk Drivers' that was carried out by Ministry of Land, Infrastructure and Transport in Korea.

REFERENCES

- [1] Reported Road Casualties in Great Britain: Quarterly Provisional Estimates Q1 2014. (2014). London: Department for Transport.
- [2] World Health Organization. (2013). *WHO global status report on road safety 2013: supporting a decade of action*. World Health Organization.
- [3] Reported Road Casualties in Great Britain: 2012 Annual Report. (2014). *Statistical Releases*. London: Department for Transport.
- [4] Hubel, D. H., & Wiesel, T. N. (1962). Receptive fields, binocular interaction and functional architecture in the cat's visual cortex, *The Journal of Physiology*, 160(1), 106.
- [5] Han, W. S., & Han, I. S. (2014). All Weather Human Detection Using Neuromorphic Visual Processing. In *Intelligent Systems for Science and Information* (pp. 25-44). Springer International Publishing.
- [6] Han, I. S. and Han, W. S. (2014). Robust Detection of Pedestrian/Cyclist by Neuromorphic Visual Processing, In *FISITA 2014 World Automotive Congress, 2-6 June 2014*.
- [7] Jhuang, H., Serre, T., Wolf, L., & Poggio, T. (2007, October). A biologically inspired system for action recognition, *IEEE 11th International Conference on Computer Vision, 2007. ICCV 2007*. (pp. 1-8). IEEE.
- [8] Picard, R. W. (2010). Affective computing: from laughter to IEEE, *IEEE Transactions on Affective Computing*, 1(1), 11-17.
- [9] Hudlicka, E. (2003). To feel or not to feel: The role of affect in human-computer interaction. *International Journal of Human-Computer Studies*, 59(1), 1-32.
- [10] Ekman, P. (1992). An argument for basic emotions, *Cognition and Emotion*, 6(3-4), 169-200.
- [11] Ekman, P. E., & Davidson, R. J. (1994). *The Nature of Emotion: Fundamental Questions*, Oxford University Press.
- [12] Izard, C. E. (1992). Basic emotions, relations among emotions, and emotion-cognition relations, *Psychological Review*. 99(3):561-5
- [13] Beck, J. S. (1995). *Cognitive Therapy*. John Wiley & Sons, Inc.
- [14] Ekman, P., & Rosenberg, E. L. (eds.). (1997). *What the Face Reveals: Basic and Applied Studies of Spontaneous Expression Using the Facial Action Coding System (FACS)*. Oxford University Press.
- [15] Kanade, T., Cohn, J. F., & Tian, Y. (2000). Comprehensive database for facial expression analysis, *Proceedings of Fourth IEEE International Conference on Automatic Face and Gesture Recognition, 2000*. (pp. 46-53). IEEE.
- [16] Kolodyazhnyy, V., Kreibig, S. D., Gross, J. J., Roth, W. T., & Wilhelm, F. H. (2011). An affective computing approach to physiological emotion specificity: Toward subject-independent and stimulus-independent classification of film-induced emotions, *Psychophysiology*, 48(7), 908-922.
- [17] Picard, R. W., Vyzas, E., & Healey, J. (2001). Toward machine emotional intelligence: Analysis of affective physiological state, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 23(10), 1175-1191.
- [18] Ekman, P., & Rosenberg, E. L. (eds.). (1997). *What the Face Reveals: Basic and Applied Studies of Spontaneous Expression Using the Facial Action Coding System (FACS)*. Oxford University Press.
- [19] Ekman, P., Davidson, R. J., & Friesen, W. V. (1990). The Duchenne smile: Emotional expression and brain physiology: II. *Journal of Personality and Social Psychology*, 58(2), 342.
- [20] Shackelford, T. K., & Larsen, R. J. (1997). Facial asymmetry as an indicator of psychological, emotional, and physiological distress. *Journal of personality and social psychology*, 72(2), 456.
- [21] W. Han (2014). Multiple Direction Filter as the Feature Map for Image Analysis of Affective Computing. MSc thesis. Imperial College London