The Driver Monitor System: A Means of Assessing Driver Performance

Kevin C. Baldwin, Donald D. Duncan, and Sheila K. West

With the “graying of America,” the need for the appropriate assessment of driving performance in the elderly is increasing. Three general approaches may be taken: self-report based on questionnaires, driving simulators, and driving monitors. Each approach has its own advantages and drawbacks. Self-report is subjective and of questionable reliability. Driving simulators provide strict control over the driving environment but are complex and costly and may not represent true driving situations. In addition, the driver is cognizant of the evaluation, which may alter driving behavior. A driver monitor, the approach we have chosen, is less costly and may provide more objective measures. Our driver monitor system is composed of a GPS receiver, a two-axis accelerometer, and three video cameras. We detail the design considerations that led to this architecture, discuss the software for control and data acquisition, and present some of our preliminary data screening and feature extraction algorithms. Although our system is intended more as a research tool, it also has obvious clinical application for testing driving ability.

INTRODUCTION

Driving an automobile is a major component of independence and mobility for older people.\textsuperscript{1–3} Loss of driving privileges and resultant dependencies are greatly feared, and this older population may continue to drive even if they are at high risk of crashes.\textsuperscript{4,5} As the elderly become aware of their growing limitations, they may compensate by adopting behaviors that avoid high-risk situations such as nighttime or rush hour driving. The absence of such awareness, or denial of limitations, coupled with the need for driving independence, may lead to a delay in changing driving behaviors that put this population (and others) in danger.

Some evidence exists that crashes among older people may be the result of exposure to high-risk situations along with limitations in actual driving performance. Avoidance of high-risk situations, as well as attention to driving performance and other compensatory strategies, are behaviors that allow older drivers with limitations to continue driving safely under prescribed circumstances. Such voluntary behaviors are felt to contribute to risk reduction.\textsuperscript{6–9} Thus, there is much interest in determining actual driving behaviors and factors that are important in an elderly person’s decision to adopt compensatory strategies or to cease driving altogether.
Some data suggest that physical, cognitive, and sensory limitations, as well as awareness of those limitations, are predisposing factors that lead to the initiation of compensatory strategies.\textsuperscript{7,10–18} However, the possible interaction of these factors and the time frame to produce alterations in driving behavior are unknown. Moreover, current techniques for assessing driving practices using traditional questionnaires, driving simulators, or examiners bear an unknown relationship to actual practices in daily life.

THE DRIVER MONITOR SYSTEM
Our research dictated several key features for the driver monitor system (DMS):

- It had to be both robust and small enough to fit inside a standard vehicle without obscuring the driver’s view.
- System components had to be relatively inexpensive.
- It had to record key features under specific conditions such as night driving or driving in bad weather.
- It had to be able to store a week’s worth of data on the driver so that reasonable statistics could be constructed.

The intent was to evaluate the driving performance of 1500 drivers, each for a 1-week period, several times over the course of a multi-year study. Logistics dictated that 50 separate DMS units be constructed, so cost was an important issue.

Evolution of Function and Form
Over the course of perhaps 1.5 years, the basic design of the DMS evolved, in terms of both required functionality and the hardware used to achieve it, through a series of prototypes. During the initial concept stage, we believed that the logging of time-tagged GPS data would provide sufficient information for the study. Shortly thereafter we realized that acceleration data would prove useful in assessing sudden movements not captured by the GPS. Questions about traffic conditions, meteorological conditions, and the spatial relationship between the driver’s car and other vehicles led to the decision to use video. Various lighting conditions necessitated two cameras, one for observing driver color cues such as red lights and one for low-light conditions. Finally, the issues of driver identity (Who is behind the wheel?) and driver activity (At what frequency does the driver check the rear view mirrors?) arose. The solution was to use a third camera that viewed the driver. At each step in the evolution of the system’s functionality we realized that, with the addition of just one more feature, we could achieve greater and greater capability.

As expected, the various additions placed increasing demands on the computer for control and data storage. For instance, a single USB webcam was initially used to record traffic conditions, snapping frames at a rate of one per second. (This rate was dictated by a combination of CPU load and hard-drive capacity considerations.) To record an image of the driver, a second USB webcam was added to the system. A number of limitations became apparent at this point. For example, the driver webcam images were typically low contrast because of backlighting of the driver, and the forward-looking cameras had poor interscene dynamic range (nighttime images were mostly black). The solution was to switch to CCD-based cameras that were usually used for security applications. A set of typical nighttime images is shown in Fig. 1. USB video capture was a significant bottleneck for image acquisition, with USB 1.0 being the only feasible high-bandwidth port available at the time. To overcome the problems of low image acquisition rate and the time sample skewing that would result from multiplexing the video signals, a quad processor was used. The quad

Figure 1. Images captured on a street with no lighting at about midnight. Traffic conditions were recorded using a low-light-level monochrome camera (left). An image of the vehicle operator was recorded with a monochrome camera with an integrated infrared light emitting diode illuminator (right).
Processor is a piece of video security equipment that "stitches" four separate video signals into a single image. Thus, simultaneous image capture could be achieved on all three cameras at the frame rate associated with a single USB acquisition device.

We anticipate that the system will continue to evolve, but for the most part, it has the required features for a thorough and economical characterization of driving performance.

As in any engineering design, the system represents a compromise between cost and functionality. For example, a desirable adjunct to a GPS would be an electronic compass. For ease of installation in various types of vehicles, however, a tilt-compensated device would be required, but the cost of such a device was prohibitive. Another desirable feature would be an active device such as a laser radar for assessing following and oncoming distances. Again, cost was an issue, but so too was the difficulty of installing such a device on a variety of vehicles. Finally, a desirable feature for video capture and storage would be a digital video recorder with integrated quad processor. This would provide dedicated video storage as well as image acquisition at a rate of 30 frames per second.

Pilot Study Prototype

To evaluate the performance of our prototype system, we enlisted a number of volunteers at our Salisbury Eye Evaluation Clinic in Salisbury, Maryland. The protocol was approved by the Joint Committee on Clinical Investigation of the Johns Hopkins Medical Institutions. The protocol followed the tenets of the Declaration of Helsinki, and informed consent was obtained after the nature of the study was explained.

Figure 2 shows one of the DMS prototype units used to collect the pilot study data. Data were logged using a compact 700-MHz PC, which runs a data collection application developed using LabVIEW. The hard drive used for the pilot study was a 40-GB EIDE laptop drive, which could store over 20 h of continuous data. Based on our initial self-report of driving behaviors, none of the group aged 70 and older reported driving more than 15 h over 5 days, so we had enough storage for the 5-day use. The PC provided a sufficient number of USB and RS232 ports to support the desired instruments. Power for the computer and all peripherals was derived from a regulated 12-V power supply that, in turn, was plugged into the vehicle cigarette lighter (or receptacle used for recharging cell phones).

Accelerations parallel and perpendicular to the automobile were measured by a single board unit. Axial (parallel) accelerations are associated with the use of the automobile accelerator and brakes. Lateral (perpendicular) accelerations are produced principally by steering maneuvers. In this first prototype, the acceleration sensor was mounted to the USB video capture device to conserve space. The GPS unit, indicated with a white upward arrow, is also shown in Fig. 2.

The video camera system comprised a color CCD camera, a low-light-level monochrome CCD camera with auto-iris lens, and a compact monochrome CCD camera with integrated infrared illumination. The video signals were fed into the quad video processor, and the resulting "patchwork" video was captured by a USB device.

Steps were taken to make the DMS inconspicuous to enable an objective driving assessment. This particular prototype was designed to fit under the passenger seat. Future systems will be integrated into a unit that can easily be strapped to the back of the passenger seat, with the exception of the small two-axis accelerometer unit and cameras. The camera/GPS assemblies were attached to the inside upper right-hand corner (passenger side) of the windshield with suction cups. Although at first glance the video system appeared obvious to the driver, it was quickly forgotten. This may be due in part to distributing the pieces (rather than clumping them) as far from the driver as possible, as well as using dark-colored components located in the shadow of the far roof support.

Pre-Collection Calibration

The DMS was installed in the driver’s car by a technician who calibrated the unit. Calibration provides standardization of the line of sight (LOS) of the camera at the outset, with prescribed distances.
to known objects. A standard target was placed at a
known height at the front bumper and again 30 ft from
the front bumper. The concern that this image may be
displaced because of vibration can be assessed and com-
penated for automatically because one of the forward-
looking cameras has a wide enough field of view (FOV)
that the forward portion of the car (hood) was visible
at the lower portion of the image. Simple correlation-
matching techniques could detect the hood of the car
in the calibration images relative to the hood of the car
in images throughout the measurement period; shifts of
greater than 2% resulted in automatic recalibration at
the processing step. In this way, standard distances were
specific for each driver in the study.

DATA PREPROCESSING

Data from each volunteer were subjected to several
stages of analysis. The first was a “quick look” to verify
driver identity and the quality of the data. Next was an
autonomous extraction of various features that in and of
themselves were of interest in quantifying the driving
behavior. This autonomous step also extracted cues for
subsequent manual inspection by the analyst. We dis-
cuss these analysis stages below.

Post-Collection Data Quick Look

An important software tool for quickly reviewing
data from the DMS is the Driver Monitor Reader. The
version of the application presented here (for an earlier
DMS prototype that used two USB webcams, one for
the driver and one for the forward
view) displays the GPS, video, and
acceleration data in a coordinated
fashion. This interface has many
display options. We will mention a
few of the major features.

Figure 3 is a screenshot of a short
driving segment. The entire travel
route (green) from the GPS receiver
is on the left, the forward-looking
video at the upper right, and accel-
eration data at the lower right.
Immediately below the video frame
are the control buttons, much like
those on a VCR, for playing the
video forward or back or for advanc-
ing or backspacing one frame at a
time. The particular video frame
corresponds to the center of the
period indicated by the highlighted
white squares in the GPS route.
Acceleration data (red is axial,
white is lateral) are displayed in the
form of an oscilloscope trace whose
time extent corresponds to the same

Automated Processing for Cues
and Feature Extraction

Since the different sensors in the DMS acquire an
enormous amount of data, it would be impractical for an
analyst to visually inspect any more than a small portion
of the data. As a result, we use a combination of autono-
mous and semi-autonomous data processing approaches.
Autonomous algorithms are used to extract features
from the data related to driving performance (total miles
driven, miles driven at night, geographical extent, speed
and acceleration statistics, etc.). These features are sub-
sequently subjected to statistical analysis. In addition, we
use a series of preliminary screening algorithms to cull
portions of the data worthy of closer examination by the
In the following paragraphs, we discuss several notions of each type of algorithm. We demonstrate that quantitative measures of driving performance can be extracted from the data autonomously and that cues can be derived from one or more of the sensors. For example, data from both the GPS receiver and the accelerometers can be used to designate the portion of the video data that should be analyzed manually. The various strategies outlined herein illustrate how data from the DMS can be used to quantitatively assess driving performance with minimum manual scrutiny.

Position

Figure 4 shows an example travel segment between Silver Spring and Annapolis, Maryland. The GPS data are analyzed at this level using ArcView. This is a Geographic Information Systems (GIS) application that uses geopolitical and geophysical data down to the detail of local streets, buildings, parks, etc. ArcView allows the use of existing databases containing information, for instance, on street centerline, type, and condition, as well as analysis of our GPS data within this context. As an example, one of the “layers” in the GIS analysis would contain street data, including the attribute of posted speed limit. GIS databases at the Maryland State Highway Administration and Salisbury State University provide a context for analyzing our GPS data, e.g., we can easily condition our left-hand turn cues on the existence of an intersection. Thus false alarms, such as might be caused by road curves, are obviated.

Velocity

GPS time is acquired concomitantly with position fixes, so velocities are easily determined. Figures 5a and 5b show the velocity history and velocity histogram, respectively, for the route from Silver Spring to Annapolis shown in Fig. 4. These results show that the majority of time for this travel segment was spent on freeways. However, some portions of the segment have zero velocity (stoplights, stop signs) and minor maxima at 35 and 45 mph. Statistics of this kind are useful for characterizing the driving environment during a particular segment. Conversely, with knowledge of the particular street or road (and the posted speed limit), statistics such as these are useful for assessing compliance.

Heading

We now discuss the method of calculating compass heading, measured clockwise from north. In and of itself, compass heading is of little interest, but is useful for autonomously detecting turns. Specifically, we are interested in left-hand turns made from a stop.

Determination of heading requires at least two GPS fixes. For accurate estimates, the distance between successive fixes must be at least 20 ft, little more than the length of the average car. This criterion obviates inaccurate estimates when the car is stationary. Naturally, the heading is “wrapped” to the fundamental (0°, 360°) interval. (If one’s heading is 355° and a 15° turn to the right is made, the new heading is 10° instead of 370°.) Unwrapping is done by detecting discontinuities of 360°. Left-hand turns display angles that decrease by 90° to within some tolerance (e.g., 90° ± 10°). Similarly, right-hand turns correspond to angles that increase by 90°. A right-angle turn by itself is not the only criterion used to derive a cue to further inspect the video. One can easily imagine that a road may curve by such an angle. Thus, we place an upper limit on the velocity during the turn; specifically, a zero velocity followed by a turn is of interest.

An example of these calculations is shown in Fig. 6 in which two turns of about 90° are indicated. Turn B is associated with a decrease in heading angle, so it is a left-hand turn. Inspection of the velocity record for this time (Fig. 5a), however, indicates that this is a high-speed turn and so is likely made on a highway. In fact, inspection of the video record showed that this turn is the
ramp from I-495 southbound to Route 50 eastbound. Turn A is associated with an increase in the heading angle of about 90°, i.e., a right-hand turn. Inspection of the velocity record for this point in the travel segment shows that, immediately before the turn, the velocity was essentially zero. Hence, this turn was likely made from a stoplight.

In the manner described, it is straightforward to autonomously detect left-hand turns made from a stop. This cue is then used to designate the corresponding video frames that are to be inspected for oncoming traffic and for head motion.

**Acceleration**

Acceleration, both axial and transverse, is recorded at a 50-Hz rate. Exceedences of prescribed absolute acceleration and deceleration thresholds are used as cues to inspect the corresponding velocity and video data. Similarly, the transverse acceleration record is inspected for threshold exceedences associated with sudden steering maneuvers. As we demonstrate below, variations in acceleration are also characteristic of certain driving environments.

We begin our discussion with the axial acceleration record for our example travel segment (Fig. 7). This record clearly shows the periods of larger acceleration (positive) and braking (negative) that are associated with city and suburban driving. In the middle of the segment (5–30 min), variations are much smaller as this is the freeway portion of the travel segment. We derive cues from such a record for closer analysis of the video data. These cues are the exceedences of absolute acceleration thresholds (green lines at ±150 milli-g) and their associated times of occurrence. The corresponding geographic locations of these cues are denoted in Fig. 8 by the black asterisks.

Similarly, the transverse acceleration record is characterized in the same fashion (Fig. 9). In such a record one observes accelerations associated with maneuvering. For example, consider the large spike in the acceleration record at about 15 min. From the velocity record (Fig. 5a), we see that the speed around this point in time is high. Thus it is likely that this is a curve in a highway. In fact, it is the ramp from southbound I-495 to eastbound Route 50 that we discussed earlier in the context of turns.

**Traffic Lights**

Red traffic lights are detected by autonomous inspection of the hue of the video frame. This algorithm has a number of steps. From the color images, the red/gray-scale ratio is calculated. This is thresholded and a morphologic opening is performed to

![Figure 5. Velocity history (a) and histogram (b) for the travel segment from Silver Spring to Annapolis shown in Fig. 4.](image)

![Figure 6. Compass heading history with example 90° turns indicated along the Silver Spring to Annapolis segment shown in Fig. 4.](image)
eliminate noise. The approach very effectively detects red traffic lights. Unfortunately, it also produces false alarms caused by red taillights.

To discriminate against these false alarms, we use the apparent relative velocity of objects within the FOV. In the machine vision community this concept is known as optical flow. Optical flow depends on position within the FOV as well as relative motion with respect to the camera. For a sequence of images as viewed from a moving vehicle, stationary objects in the FOV have an apparent velocity that is directed radially from the camera’s LOS. Traffic lights, which are typically in the upper half of the image, display an optical flow that is quite different from that of taillights in the lower half of the image. Flow associated with taillights is in a different radial direction and is much reduced because of the lower relative velocity. Difficulties remain, however, if the taillights also appear higher up on the car, as on some SUVs and trucks.

Lane Changes

Safe lane changes are of obvious importance. Autonomous detection of lane changes also depends on computing the optical flow within an image sequence. Optical flow can be computed in a number of ways. When the frame rate is low and there is substantial translation of spatial structures from one image frame to the next, correlation-matching techniques work well. For higher frame rates or very low velocities, gradient techniques are appropriate. For detection of pavement stripes, we use a combination of these techniques to compute the optical flow within the lower half of the image plane. The position of these high-velocity image components relative to the centerline of the vehicle allows us to detect lane changes. Times for these incidences are cues for inspection of the video record for the driver camera.

ANALYSIS

One of the cues autonomously extracted from the GPS data is incidence of left-hand turns. These tabulated times are used to evaluate the corresponding video data, with the performance metric being the distance to oncoming vehicles. Such an estimation is illustrated below. Although this example is for calculating following distance, estimation of distance to an oncoming vehicle is identical.

Distance to Oncoming Cars

Figure 10 (upper left) shows a single frame of imagery with a manually drawn red line through the license plate. The lower left panel is a plot of the pixel gray level along the indicated trace. The extent of the license plate image is clear. Because the angular extent of the camera FOV is established from the calibration step, the apparent width of the license plate (a 12-in. standard)
provides a means of inferring the distance to the vehicle.

Other factors affect the estimate of range with this method. For example, license plate brackets partially occlude the image of the plate, and similarly colored plates and vehicles present difficult contrast conditions. Fortunately, the great majority of license plate brackets occlude a small fraction of the width of the plate. Occlusion in the vertical direction has significantly more variability, which precludes a more accurate range estimate based on license plate area. Clearly, further work is required for the automation of range estimates using this method.

**Head Motion**

We take head motion as an indication of awareness of the spatial relationship between the driver's vehicle and surrounding traffic and obstacles. Specifically, we attach significance to the driver's use of the rearview mirrors prior to lane changes. As previously discussed, the data preprocessing software extracts times for these lane changes, thus providing times for the analyst to inspect the video record. Of course gross head motion is easily computed with simple frame differencing or flow computations, but the driver's quick glance in the mirror is often more subtle than head motion. Without more involved monitoring approaches such as the use of eye tracking, manual inspection of the video record is necessary.

**DISCUSSION**

Recent tragic events involving licensed older drivers point to the urgent public health need to determine the driving habits of this population, instances of poor driving performance, and factors that predict increasingly poor performance. Older drivers may compensate for perceived disabilities by altering their driving behaviors as well, and there is a need to understand such compensatory strategies and to determine if they help reduce the risk of crashes. To undertake such research, an objective determination of actual driving performance under conditions in which older drivers actually drive is urgently needed. Until now, no device or analytical framework existed to extract such data for population studies. Our ongoing research efforts include measurement of two significant components of driving ability: vision and cognitive function. The DMS provides the means of linking these pathophysiological factors to actual driving performance.

In addition, we foresee significant clinical applications for such a system. Clinical assessment of driving performance relies on a battery of tests with an on-road observer, an expensive and time-consuming undertaking whose outcome rests on a small sample of driving performance. With a fully operational DMS, the data could be fed through our analysis package and deviations from norms determined. The development of a small, unobtrusive, and objective device for recording and processing key data on driving behaviors is necessary to take driving research to the next level.

**REFERENCES**

Driver Monitor System

Kevin C. Baldwin received B.S. and M.S. degrees in optics in 1990 and 1991, respectively, from the University of Rochester, an M.S. in electrical engineering from the Rochester Institute of Technology, and an M.S.E. in 1994 and Ph.D. in 1998, both in electrical engineering, from JHU. Dr. Baldwin is responsible for the development of the DMS hardware and software.

Donald D. Duncan received a B.S. degree in 1970 from the University of Kentucky and M.S. and Ph.D. degrees in 1973 and 1977, respectively, from the Ohio State University, all in electrical engineering. Dr. Duncan joined APL in 1983 as a Senior Staff engineer, specializing in physical optics. A member of the APL Principal Professional Staff since 1989, he is currently the Supervisor of the Phenomenology and Measurements Section of the Electrooptics Group in the Air Defense Systems Department. Since 1995, he has held an appointment as Associate Professor of Ophthalmology in the JHU Department of Medicine, where he is involved in researching the epidemiology of cataract. The driver monitoring research effort is led by Sheila K. West, who received her Ph.D. in epidemiology from the JHU School of Public Health in 1980. Since then Dr. West has worked on the epidemiology of major public health problems, including the epidemiology of Hepatitis B in the Philippines in collaboration with The University of the Philippines and the epidemiology of major eye diseases at The Wilmer Eye Institute of the Hopkins School of Medicine. She is currently the El Maghraby Professor of Preventive Ophthalmology. Dr. West’s grants include a major effort involving a population on Maryland’s Eastern Shore to examine the impact on function (notably driving) of deficits in multiple measures of vision. The DMS Team can be contacted through Dr. Baldwin. His e-mail address is kevin.baldwin@jhuapl.edu.

Sheila K. West

THE AUTHORS

Kevin C. Baldwin

Donald D. Duncan

Sheila K. West