Real Time Implementation of an On-Road Video Driver Drowsiness Detector: Two-Camera Profile Inputs for Improved Accuracy

Morris Steffin, M.D. 1, Keith Wahl, M.D., F.A.C.S. 2
1 Chief Science Officer, VRNEUROTECH, San Diego, California
2 Chief Operating Officer, VRNEUROTECH, San Diego, California

Abstract: Drowsiness in drivers and pilots is a major cause of injuries and accidents. A real-time on-road alertness monitor is described whose output is derived from two laterally mounted video cameras outside the driver’s field-of-view. Output from the system is a 12-channel scalar set that measures the frequency and duration characteristics of the driver’s head position in relation to a standard driving position correlated behaviorally with alertness. The two-camera enhancement results in greater accuracy and stability allowing for validation studies with other parameters of alertness that are useful (such as electroencephalography) but not readily adaptable to the on-road environment. The processing methods described can be implemented in an embedded-processor configuration suitable for in-vehicle deployment.

INTRODUCTION

Driver and pilot fatigue, as well as pilot impairment resulting from drowsiness or environmental factors of incapacitation, are major sources of accidents and injuries. An approach to continuous, practical, noninvasive, real-time, on-board video behavioral monitoring has been described by Steffin and Wahl. 1 That method provides the mechanism for extracting behavioral corollaries of drowsiness from salient changes in feature complexes involving stable facial regions (fiducials) including eyebrow-palpebral fissure complex (EPFC), the mouth region, and the facial boundaries. As a result of the video-to-scalar operation of that approach, 12 scalar data channels are generated for each of the video analysis subregions (SRs) in each of three major regions of interest (ROIs), head, EPFC, and mouth. The method previously described was limited by processing video data from a single centrally mounted windshield camera. A shape detection algorithm that filtered video data according to shape expectations of the respective SRs resulted in 12 scalar channels whose values were representative of the precision of alignment of the subject facial features to the standard facial position correlated with alertness.

However, this configuration partially impaired the driver’s view, and was less than optimally sensitive and specific. Improvements in the technique include simultaneous input from two laterally mounted cameras that capture profile facial views rather than a frontal view. The result is increased resolution of facial features that are relevant to drowsiness.

Figure 1. Schema for data collection with two cameras. Outputs from laterally placed cameras are transferred to a frame buffer for video-to-scalar image processing. Levels of processing include intensity discrimination of facial fiducials (left panel) and shape discrimination (right panel).
Figure 1 shows the schema of this approach. Real-time video from the two cameras is transferred to a frame buffer for analysis. Two levels of processing, intensity discrimination (upper left panel), and shape discrimination (center panel) allow derivation of 12 channels of data corresponding to the alignment of the facial fiducials (eye, head, and mouth) with the standard, straight-ahead position. Further, the profile schema generates comparison of the fiducial shapes in real time against shape references derived from the actual face in standard position, which may be grabbed as the driver first approaches cruising speed. This procedure enhances accuracy, as compared to the somewhat arbitrary shape discrimination functions employed in the previous method, and will simplify calibration of the system to the individual driver.

RESULTS

Figure 2 demonstrates the increased resolution of facial fiducials resulting from the improvements in facial video acquisition described. For clarity, identical video input is fed to both channels, and analysis is performed on the left channel. The active scalar outputs shown represent the mouth and chin regions for the left channels. In A, there is a clear downward deflection, with good signal-to-noise level in both active channels, representing changes in configuration of the labial region (upper trace) and mental region (lower trace) at the time of mouth opening. The movements were on command, rather than as a result of a natural yawn, so the deflections are abrupt and stable, and give a good estimate of the frequency response and noise characteristics of the system. Similar deflections resulting from mouth closure, also on command, are shown in B. In C, a series of openings and closings are shown over a 10-second sweep.

Figure 3 shows a similar level of resolution and favorable signal-to-noise characteristics for eye blinks, also, in this instance, on command. The upper channel records eyebrow movement and the lower channel records the time characteristics of the palpebral fissure. In A, eye opening is followed by eye closure, while in B eye opening is the final state, as indicated in both the video and the scalar data. A series of eye blinks (on command) is indicated in C.
DISCUSSION

The approach previously described (Steffin and Wahl\(^1\)) represents methodology for extracting behavioral correlates of drowsiness from facial features with practicable hardware that can be reduced to a self-contained embedded-processor system. However, ergonomic factors of frontal camera placement cause driver inconvenience and tend to reduce the accuracy of the technique. The reliability of the measurements with this new configuration have increased, and the intrusion into the driver's field-of-view has been reduced by the dual camera monitoring techniques in the described implementation. As a result, the reliability of the technique appears to be significantly increased.

Future directions include enhanced compensation for ambient light variation. Initial considerations indicate that near infrared monitoring, as opposed to visible light, will provide further improvements in the stability of the system in daylight and will be required for this measurement system at night. Research is in progress regarding data collection with limited infrared.

It is anticipated that the same attributes of efficiency in processing and coupling of system output to relevant driver behavior will be realized in the described configuration, so that a reliable system of drowsiness detection, independent of ambient visible light, will be achievable. Behavioral studies are underway to determine correlation of the scalars with more generalized measures of performance (such as reaction time and accuracy of obstacle avoidance) and with electroencephalographic monitoring of alertness as a prelude to on-road testing.

REFERENCE


Submitted: November 7, 2003
Accepted: April 5, 2004