OBJECT DETECTION AND IDENTIFICATION USING ENHANCED CAMERA/VIDEO IMAGING SYSTEMS (E-C/VISs) ON HEAVY TRUCKS

William A. Schaudt, Walter W. Wierwille, and Richard J. Hanowski
Virginia Tech Transportation Institute
Blacksburg, Virginia, USA
Email: aschaudt@vtti.vt.edu

Summary: Tests were performed to determine the feasibility of developing an Enhanced Camera/Video Imaging System (E-C/VIS) to provide heavy vehicle drivers with better situation awareness to the sides and rear of their vehicles. It is well known that large blind spots currently exist in these areas and that sideswipe crashes can occur as a result. An additional goal was to extend the operating envelope of conventional video to nighttime and to inclement weather. A three channel system was envisioned in which there would be a camera at each (front) fender of the tractor looking backward along the sides of the tractor trailer. The third channel would be aimed rearward from the back of the trailer. Indoor tests involved selection of components having the best capabilities, while early outdoor tests used the selected components in a single-channel side mounted system. Once developed, the heavy vehicle three-channel system was tested in a static object detection and identification experiment, as well as a dynamic on-road experiment. The current document describes the static object detection and identification experiment methodology and results. In regard to object detection and identification, objects were correctly detected and identified significantly more often with the E-C/VIS than with mirrors alone. Objects directly behind the heavy vehicle could be detected with the rear wide-angle look-down camera of the E-C/VIS whereas such objects could not be detected with conventional side mirrors.

INTRODUCTION

The Virginia Tech Transportation Institute (VTTI) has previously investigated the use of Camera/Video Imaging Systems (C/VISs) in heavy vehicles through research funded by the National Highway Traffic Safety Administration (NHTSA) and the Federal Motor Carrier Safety Administration (FMCSA) (Wierwille, Schaudt, Spaulding, Gupta, Fitch, Wiegand, and Hanowski, 2008a; Wierwille, Schaudt, Gupta, Spaulding, and Hanowski, 2008b). This earlier research involved devising, developing, and testing a final set of 11 concept systems so that recommendations could be made and specifications written. These systems were tested in daytime conditions with good visibility. The current research further investigated the use of C/VISs in less favorable environmental conditions, namely nighttime and inclement weather. A three-camera system was developed: one camera on each side of the heavy vehicle and one at the rear. The intended purpose of the work was to reduce or eliminate blind spots and other uncertainties that might occur along the sides and to the rear of the heavy vehicle. The objective in so doing was that the driver of the heavy vehicle would be better informed regarding the environment around the sides and to the rear of the heavy vehicle; that is, he or she would have better situation awareness. This system was developed as an enhancement, that is, the side

mirrors would remain on the vehicle even though the video systems would be added. This was done so that if malfunction of either side video chain occurred, the heavy vehicle could still be driven in a conventional manner. There were two phases to this project. Phase 1 consisted of initial laboratory testing of the "best available technologies", development of the passenger side Enhanced C/VIS (E-C/VIS), and finally formal human factors stationary testing in various conditions, e.g. nighttime, daytime, rain, and wind. Phase 2 consisted of developing the final system and performing an object detection and identification static experiment as well as formal human factors dynamic testing on the Virginia Smart Road. The purpose of this paper is to discuss the object detection and identification static experiment used in Phase 2, and results obtained.

METHOD

As stated earlier, the decision was made to perform an object detection and identification task, but to do so using a stationary situation. This would allow participants to become familiar with the E-C/VIS prior to driving on the Smart Road. At the same time, object detection would become a safe exercise, while useful information regarding ability to detect targets would be obtained. Tests were performed in a large blacktopped area at VTTI using a variation of the paradigm developed by Jenness, Llaneras, and Huey (2005) under subcontract in the original C/VIS project.

Participants

Tests were run with CDL drivers. Participants were recruited from a volunteer database. Recruiting was carried out without regard to gender, but participants had to have at least two years of full time experience as a heavy vehicle driver. They were given visual acuity tests to ensure that their vision was adequate (20/30 or better, corrected if necessary). As it turned out, eight males agreed to participate and were found to be qualified.

Equipment & Materials

A tractor-trailer was equipped with the three-channel E-C/VIS, as previously described. This system was composed of the three cameras in weatherproof housings and three monitors, one at each A-pillar and one at the equivalent center rearview mirror position. Eight IR LED illuminators were mounted around the tractor-trailer. Three laptop computers were on board to perform the image processing that emphasized outlines of objects in each video view. Figure 1 shows the passenger side fender-mounted camera position and an example of a passenger side nighttime processed image of a vehicle in the adjacent lane. Data gathering instrumentation for testing was handled totally separately from the E-C/VIS instrumentation. Two cameras were added to the tractor-trailer, one just above each A-pillar monitor. These were used to determine a participant's eye glance position. Recorded video allowed for distinguishing between driver glances at the side mirrors and A-pillar monitors. Each face camera included its own IR illumination, which of course could not be seen by the participant. This illumination was necessary because of the need to record video at night. The two camera outputs were recorded digitally. This image contained a time stamp and an audio track with two microphones as inputs (one positioned near the participant, and one positioned near the experimenter). This audio track

allowed search times to be measured and correctness of identifications could be checked after the experiment was run. The experimenter sat behind the participant, but in a centered position. The experimenter could look over a participant's right shoulder and also see all three E-C/VIS images directly to ensure that all elements of the E-C/VIS were operating properly.

Figure 1. Photo on Left: Passenger Side Fender-mounted Camera.
Photo on Right: Example of a Passenger Side Nighttime Processed Image of a Vehicle in the Adjacent Lane

Procedure

After preliminary qualifications were completed in one of the buildings at VTTI (including signing the informed consent, performing simple hearing and vision tests, and filling out necessary payment forms) the participant was taken to the parked, equipped tractor-trailer. There the participant was seated in the driver's seat and instructed to look down (noise cancelling headphones were used to shield the participant from sounds of the exterior maneuvering of objects). On command, the participant looked up and searched for the object. Once found, the participant so indicated, identified the object, and gave an approximate location. Participants who could not find an object simply said it was undetectable. Half of the participants performed with Baseline (E-C/VIS not operating) for the first 15 objects/locations while these same participants performed with the E-C/VIS operating for the second set of 15 objects/locations. The other half of participants performed in reverse order for E-C/VIS versus Baseline. Actually, the two sets of objects/locations were identical, but were presented in two pseudo-random orders; that is, one order for the first 15 objects and the other order for the second 15 objects. The orders were changed in such a way that the participant was unable to predict where the next object might be, what type of object it might be, and what its location was. On the other hand, because every object was presented twice, once in Baseline and once with the E-C/VIS, within-subject comparisons could be made as a function of object/location and E-C/VIS operating or Baseline. Results were scored in terms of correctness of object location, correctness of object location *and* identification (combined), and response time.

Figure 2 shows the layout for the object detection and identification static experiment. Five lanes on a large asphalt pad were used, with the tractor-trailer in the center lane. The figure shows the approximate planar coverage of the three video cameras. Four objects were used for the tests: an orange barrel with two horizontal silver stripes (shown in the figure as a circle), a pedestrian

(shown as a small circle with a surrounding ellipse), a yellow adult tricycle (shown as a pair of handlebars with a line segment connected to the representation of the two rear wheels), and a red stop sign placed on a standard (shown as a hexagon surrounding an x in the figure). The 30 test conditions were ordered as shown by the numbers just to the left of the object in the figure. For example, the tenth presentation was the adult tricycle located in the D1 (driver's side adjacent) lane at distance position 3. This object and location also corresponded to the thirtieth presentation. A photograph of the four objects is shown in Figure 3. This experiment was performed with the IR illumination on at all times. The reason for this was that the IR illumination was not helpful to the participant in Baseline, because it could not be detected visually. It was only helpful when viewed with the E-C/VIS.
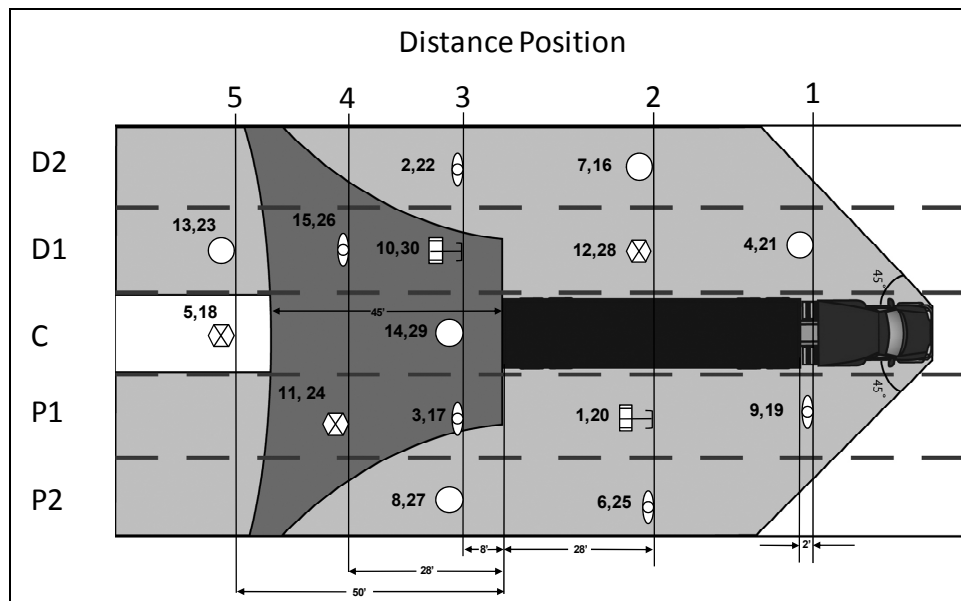

Figure 2. Layout for the Object Detection and Identification Task

**Figure 3. Daytime Photograph of the Four "Objects" Used in the Experiment**

RESULTS

To begin the analysis, two pairs of objects/locations were removed from the larger data set because they represented special conditions. Referring back to Figure 2, note that one pair (14, 29) could only be seen in the center monitor of the E-C/VIS. It could not be seen in the other E-C/VIS monitors or with the mirrors. Consequently, the correct response for this pair would be "undetectable" for the Baseline condition, but detectable for the E-C/VIS condition with the participant indicating that the object (a barrel) was directly behind the trailer. Results for this pair showed that, indeed, all eight participants found the target to be "undetectable" in Baseline and seven of the eight participants correctly located the target and identified it when the E-C/VIS condition was presented. The results were tested using a Fisher Exact test and were found to be significant in the two-sided test ($p = 0.0014$). These results indicate that the E-C/VIS condition resulted in better coverage behind the trailer. Another pair (5, 18) was also separated from the main data set. In this case, the stop sign could not be seen in either the Baseline mirrors or in the E-C/VIS monitors. This stop sign was behind the trailer, but sufficiently far back that it was out of the field of view of the rear wide-angle look-down camera. Consequently, all responses should have been found "undetectable" to be correct. Indeed, results showed that all eight participants responded with "undetectable" in both cases and, of course, there was then no statistically significant difference. In regard to response time, differences between the E-C/VIS and Baseline conditions were tested, with the result that response times were found to not significantly differ using a t-test: $t(7) = -0.59$, $p = 0.57$.

Thirteen pairs remained after (14, 29) and (5, 18) were removed. In these remaining pairs, as Figure 2 shows, all objects should have been detectable in both Baseline and in E-C/VIS conditions, neglecting the fact that data were gathered at night without added *visible* light from the heavy vehicle (note that normal heavy vehicle lighting was present). This represents a nighttime comparison, which is where the E-C/VIS was intended to extend the range of vision. The first analysis was directed at correctness of object location. There were 13 pairs for each of eight participants, resulting in a total of 104 trials for Baseline and 104 trials for E-C/VIS. A Chi-Square analysis demonstrated a significant difference between the Baseline and E-C/VIS conditions: $\chi^2(1) = 22.5$, $p < 0.0001$. In terms of comparison, participants provided the correct location of objects 96 percent of the time with the E-C/VIS operating, and 72 percent of the time with Baseline. The data were also subjected to a correctness of location *and* correctness of object identification analysis. In a few cases, objects were correctly located and then incorrectly identified. Once again there were 13 pairs for each of the eight participants, again resulting in a total of 104 trials for Baseline and 104 trials for E-C/VIS. A Chi-Square analysis again demonstrated a significant difference between Baseline and E-C/VIS conditions: $\chi^2(1) = 15.0$, $p = 0.0001$. Participants provided correct location *and* correct identification of objects 91 percent of the time with the E-C/VIS and 70 percent of the time with Baseline.

Participant response times were analyzed for the 13 pairs of locations/objects. Two types of analyses were performed in an attempt to handle incorrect location responses. In the first of these analyses, incorrect location responses were handled by adding 1 s to the longest correct response time for each given participant. Here, the reasoning was that the participant would very likely take at least that long to provide the correct response. This type of analysis allowed completion of the data matrix of values so that an equal-N's analysis could be carried out. In the second of

these analyses, response times were analyzed regardless of correctness. In other words, this analysis was simply response time from the cue to begin the search until the participant responded with a location, whether correct or incorrect. This approach also resulted in a complete data set. In each case, a 2 (Baseline versus E-C/VIS) by 13 (Object location-Type) within-subject ANOVA was performed, using the data from the eight participants. For the first analysis, the Baseline versus E-C/VIS main effect was not significant, but wasn't far from it $F(1, 7) = 4.59$, $p = 0.0694$. The main effect of Object location-Type was significant: $F(12, 84) = 6.82$, $p < 0.0001$, and so was the interaction of Baseline versus E-C/VIS with Object location-Type $F(12, 84) = 5.53$, $p < 0.0001$. Results for the second analysis in which responses were included regardless of correctness provided similar results. The Baseline versus E-C/VIS main effect was not significant, but was quite close: $F(1,7) = 5.12$, $p = 0.0582$. The main effect of location type was significant with $F(12, 84) = 9.17$, $p < 0.0001$, and the interaction was significant with $F(12, 84) = 5.94$, $p < 0.0001$. For information purposes, the main effect of Baseline versus E-C/VIS with "corrections" made for incorrect responses and the main effect of Baseline versus E-C/VIS on response time, regardless of correctness, are plotted in Figure 4.
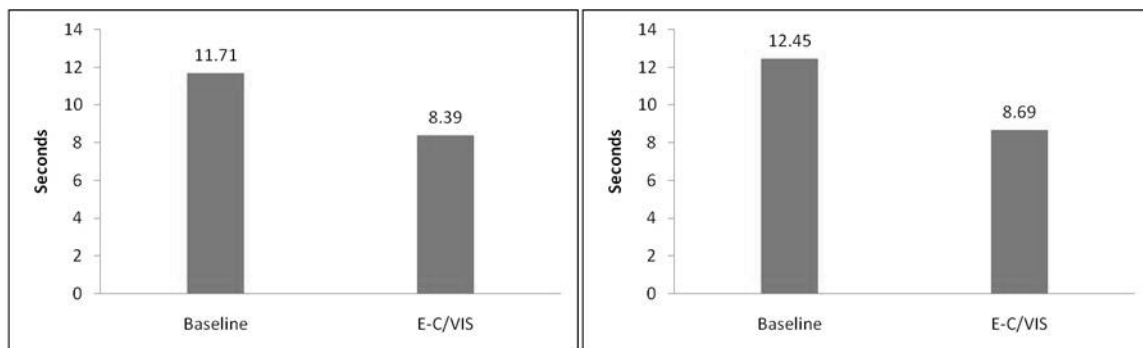
Figure 4. Graph on Left: Response Times as a Function of Baseline versus E-C/VIS, with Corrections Made for Incorrect Responses (note that *p* = 0.0694). Graph on Right: Response Times as a Function of Baseline versus E-C/VIS Regardless of Correctness (note that *p* = 0.0582)

SUMMARY AND CONCLUSIONS

All results were encouraging and indicated that participant performance was better with the E-C/VIS. In regard to object detection and identification, objects were correctly detected and identified significantly more often with the E-C/VIS than with mirrors alone. Objects directly behind the heavy vehicle could be detected with the rear wide-angle look-down camera of the E-C/VIS whereas such objects could not be detected with conventional side mirrors. The test conditions were relatively dark, except for the situations involving IR illumination (with which the E-C/VIS functioned effectively). The rear channel of the E-C/VIS provides a great advantage of detection and identification of objects to the rear.

An additional finding in regard to object detection and identification was that the E-C/VIS did not generally add to response time. In fact, response times were generally shorter when the system was activated, even though this produced several more sources for the participant to scan. This finding indicates that the improved detection *did not* come at the expense of increased response time.

It will be recalled that participants first performed in the object detection and identification tests and then in formal human factors dynamic testing on the Virginia Smart Road. Once all of these tests were complete, participants were asked to rate on three additional scales prior to being dismissed. They rated on overall usefulness, whether or not they would like to have an integrated E-C/VIS on their own rig, and whether or not the E-C/VIS improved situation awareness (defined as being aware of the situation along the sides and to the rear of the heavy vehicle). In all three cases, ratings were very high indicating high acceptance of the E-C/VIS by CDL drivers. Of course, these results included the Smart Road experience as well as the object detection and identification tasks.

Overall, this project resulted in a successful development effort that met performance objectives and also demonstrated good driver opinion ratings. An E-C/VIS is believed to represent a distinct step forward in heavy-vehicle design and safety.

ACKNOWLEDGEMENTS

REFERENCES

Jenness, J.W., Llaneras, J.W., and Huey, R.E. (2005). *Static testing method for assessing quality of indirect visibility on heavy trucks; Final report*. Phase 1, Track 1, Task Order 18, NHTSA Contract DTNH22-00-C-07007. Rockville, MD: Westat (April).

Wierwille, W.W., Schaudt, W.A., Gupta, S.K, Spaulding, J.M., Fitch, G.M., and Hanowski, R.J. (July, 2008a). *Development of a performance specification for indirect visibility systems in heavy vehicles final report supporting research*. Report No. DOT HS 810 960. Washington, DC: U.S. Department of Transportation, National Highway Traffic Safety Administration.

Wierwille, W.W., Schaudt, W.A., Gupta, S.K, Spaulding, J.M., and Hanowski, R.J. (July, 2008b). *Development of a performance specification for indirect visibility systems in heavy vehicles final report specifications*. Report No. DOT HS 810 958. Washington, DC: U.S. Department of Transportation, National Highway Traffic Safety Administration.