PEDESTRIAN DETECTION WITH NEAR AND FAR INFRARED NIGHT VISION ENHANCEMENT

Omer Tsimhoni
Jonas Bärgman
Takako Minoda
Michael J. Flannagan

December 2004

PEDESTRIAN DETECTION WITH NEAR AND FAR
INFRARED NIGHT VISION ENHANCEMENT

Omer Tsimhoni
Jonas Bärgman
Takako Minoda
Michael J. Flannagan

1. Report No. UMTRI-2004-38	2. Government Accession No.	3. Recipient's Catalog No.
4. Title and Subtitle Pedestrian Detection with Near and Far Infrared Night Vision Enhancement		5. Report Date December 2004
		6. Performing Organization Code 302753
7. Author(s) Tsimhoni, O., Bärgman, J., Minoda, T., and Flannagan, M.J.		8. Performing Organization Report No. UMTRI-2004-38
9. Performing Organization Name and Address The University of Michigan Transportation Research Institute 2901 Baxter Road Ann Arbor, Michigan 48109-2150 U.S.A.		10. Work Unit no. (TRAIS)
		11. Contract or Grant No.
12. Sponsoring Agency Name and Address The University of Michigan Industry Affiliation Program for Human Factors in Transportation Safety		13. Type of Report and Period Covered
		14. Sponsoring Agency Code

15. Supplementary Notes

16. Abstract

Current headlighting and road lighting are only partly effective in reducing the risk of driving at night. Various forms of night vision enhancement systems, using a variety of sensing technologies, are being developed to further reduce this risk. Two major sensing technologies are receiving particular development interest and are both currently available on new vehicles: far infrared (FIR) systems, which generate images by passively detecting thermal emissions from objects and surfaces in the road scene, and near infrared (NIR) systems, which actively illuminate the scene in the near infrared spectrum and capture the reflected radiation. The images generated by these systems, and the ways they are used by drivers, are expected to differ. There is evidence that the major safety problem caused by darkness is increased risk of pedestrian collisions. Because pedestrians are usually prominent among far infrared sources in roadway scenes, their detection may be especially enhanced in FIR views.

To compare pedestrian detection in NIR and FIR views, a test vehicle equipped with each type of system was driven at night on several roads with pedestrians standing along the route. Video clips, recorded from both systems simultaneously, were later shown in a laboratory setting to 16 subjects (eight younger than 30 years and eight older than 60 years). Subjects pressed a button as soon as they saw each pedestrian. Detection distances with FIR were significantly greater than with NIR. Younger subjects had greater detection distances than did older subjects, and both age groups had greater detection distances with FIR. The effectiveness of NIR and FIR systems can be expected to depend on the details of implementation as well as any inherent advantages of either technology. To the extent that the two systems used in this experiment reasonably represent the respective technologies, the results support the expected enhancement of pedestrian detection in FIR systems.

17. Key Words Night vision, infrared, FIR, NIR, pedestrian detection		18. Distribution Statement Unlimited	
19. Security Classification (of this report) None	20. Security Classification (of this page) None	21. No. of Pages 25	22. Price

i

ACKNOWLEDGMENTS

CONTENTS

INTRODUCTION

Current headlighting and road lighting are only partly effective in reducing the risk of driving at night. Various forms of night vision enhancement systems (NVESs), using a variety of sensing technologies, are being developed to further reduce this risk. Two major sensing technologies are receiving particular development interest and are both currently available on new vehicles: near infrared (NIR) systems, which actively illuminate the scene in the near infrared spectrum and capture the reflected radiation and far infrared (FIR) systems, which generate images by passively detecting thermal emissions from objects and surfaces in the road scene. The images generated by these systems, and the ways they are used by drivers, are expected to differ.

NIR-based systems are often portrayed as the natural and best solution for night vision enhancement. The technology underlying NIR systems is not essentially different from conventional CCD cameras with a source of illumination. At the simplest level, the only difference is in the range of the spectrum used and the fact that NIR is not visible to an unaided eye. In-vehicle systems using NIR are therefore often described as having natural looking images, which resemble conventional pictures and need minimal interpretation (e.g., Küpper & Schug, 2002). The type of information that is visible with NIR systems matches most of what is seen with an unaided eye, although the relative brightness of some objects may differ from that in visible light. The color of clothes in NIR systems is an important example. Most clothes appear very bright in NIR systems, even if they are black in the visible spectrum, depending on cloth material (e.g., Küpper & Schug). Similarly, lane markers appear bright because they are reflective in the NIR range. NIR systems can be devised with a mature technology whose current costs are within a reasonable price for installation on vehicles.

The primary drawbacks associated with NIR systems include their susceptibility to glare, blooming, and streaking from active sources of light such as oncoming traffic, traffic lights, and streetlights, and from reflective objects such as road signs. NIR illuminators may cause glare to other drivers using the same type of system, and may cause damage to eyes at short distances (<1 m) if the illuminators are very powerful, as they usually are (Yagi et al., 2003).

Much engineering work is being done to address these drawbacks and reduce their effects on the performance of drivers with such systems. For example, Holz & Weidel (1998) reported the development of laser-diode based IR illuminators. Laser-diode technology may effectively increase the power of the illuminators and reduce the susceptibility of the camera to glare. Time synchronization of the camera with the laser pulses opens additional options for dealing with the drawbacks of current NIR systems. As a result, detection distances with such systems are expected to be higher than with conventional incandescent illuminators. Image processing at

different levels is another promising technology that may improve the quality of the image delivered to the driver and highlight important information.

FIR systems differ from NIR systems in almost all of the aforementioned aspects. The imagery they produce is often portrayed as unnatural and difficult to understand (although recent advances have contributed to more natural-looking FIR images). Another common concern is that FIR systems provide only minimal information about the road, such as what would be necessary for vehicle control and for staying in the lane if the driver were to use the system solely for driving. FIR technology has been used extensively in the military, but only minimally in civilian transportation (e.g., Martinelli & Boulanger, 2000) and is considered more costly than NIR. Finally, limitations of transmission through glass preclude the camera from being positioned behind the windshield in the cabin, which adds a level of complexity to the use of such systems.

On the other hand, FIR systems do not share some of the drawbacks associated with NIR. They are not susceptible to glare from active or reflective sources and, because they do not employ an active source of illumination, they do not cause glare or safety concerns to other users of the road. FIR images show pedestrians and other warm objects very clearly. As we discuss below, the information that they do not show may be of limited value to the driver.

It is not surprising, then, that scientists and engineers in the automotive industry do not share a single opinion about a preferred system. Preference for either system is probably based primarily on beliefs about how the system should be used by the driver. In military applications, for example, there is clear preference for FIR for the detection of targets, and preference for NIR as a primary source for night driving under good illumination conditions. Collins et al. (1998) found that night vision goggles using NIR technology allowed for faster driving with greater reported confidence on off-road terrain. In a comparison study of civilian systems, Blanco et al. (2001) found the detection of most objects with FIR to be at a greater distance than with conventional headlamps but noted that objects such as tire treads and a child's bicycle were not detected well with FIR. (For additional discussion on the literature about night-vision enhancement systems, see Tsimhoni & Green, 2001 and Rumar, 2002.)

The critical question is therefore: "What should be the main purpose of a night vision enhancement system?" Sullivan & Flannagan (2001) provided evidence that the major safety problem caused by darkness is the increased risk of pedestrian collisions. They found that fatalities due to collisions with pedestrians increased by a factor of four in darkness versus daylight, independent of other factors related to the time of day such as fatigue and alcohol use. Considering these results along with other evidence, Rumar (2002) proposed the following answer to the above question:

"Based on the characteristics of night traffic and night crashes, the NVES should primarily enhance the visibility of pedestrians, cyclists, and animals. Rear-end crashes should be solved by other means, and single vehicle crashes appear to be more dependent on alcohol and fatigue." (p. 5)

If this recommendation is correct, a clear criterion for comparison between candidate systems can be expressed. Namely, a system that helps drivers avoid hitting pedestrians is the system of choice. The purpose of this experiment was, therefore, to directly compare FIR and NIR systems based on how well they help drivers see pedestrians at night.

Although several previous experiments have compared FIR systems to NIR systems (e.g., Collins et al., 1998; Meitzler et al. 2000; and Piccione et al., 1997) some elements, important for our present purposes, are missing from those comparisons. First, most of those comparisons were made in the context of military driving and did not focus on the detection of pedestrians. Second, the comparisons between FIR and NIR were not based on the same stimuli and did not use the same subjects. Third, it is not clear whether performance differences were due to the information available from the display or the way drivers use that information by varying their scanning patterns.

The most important assertions, assumptions, and limitations that were considered in designing the experiment are summarized below:

(1) Full emphasis will be put on the task of seeing pedestrians.

(2) Two systems will be tested on a single vehicle, one of each system type. The systems will be assembled from available components and at least one will allow full control over resolution and field of view so that those variables can be equated.

(3) The comparison will be independent of details of implementation as much as possible. The systems will be implemented in a reasonably balanced way without bias. Where feasible, identical elements will be used by both systems. For example, images will be recorded simultaneously so that factors related to pedestrians, traffic, and the environment will be the same on both systems. The display used will be the same.

(4) The comparison will be made within-subjects to avoid noise in the data due to between-subject differences.

(5) Routes and the location of pedestrians will be chosen to represent a reasonably wide range of driving scenarios, but extremes will be avoided. Preferred routes are those with the most potential for reducing collisions with pedestrians: arterial roads (rural and urban principal and minor arterials) (Sullivan & Flannagan, 2001).

(6) The locations of all pedestrians will be near the right edge of the road to minimize glare from traffic in the NIR system. Pedestrians will stand still, facing the vehicle. For simplicity, there will not be special cases such as a pedestrian changing tires on a car or a pedestrian in the middle of the road.

(7) Oncoming headlamp glare will be minimized by avoiding oncoming traffic overlapping with pedestrian positions so as not to favor the FIR system with regard to that aspect of performance.

(8) The comparison will be made in a laboratory setting based on continuously viewing the displayed information. The effects of concurrent driving will not be tested. This will allow us to measure how well subjects can extract information from each display type under nearly ideal conditions. How well subjects can use the displays under driving conditions—including, for example, the need to strategically allocate glances to the displays—is also important, but is beyond the scope of this study.

METHOD

Subjects

Sixteen licensed drivers participated in this study – eight younger (ages 21 to 30 years, mean of 24) and eight older (ages 64 to 79 years, mean of 71), with equal numbers of men and women in each age group.

Subjects' corrected vision (tested with an Optec 2000 Stereo Optical Vision Tester) was 20/30 or better for younger subjects, and 20/35 or better for older subjects. Midrange acuity (100 cm) was 20/22 or better for younger subjects, and 20/70 or better for older subjects (mean of 20/40). None of the subjects had driven with a night vision enhancement system before but a few subjects (four young men and one older man) reported they were familiar with the concept.

Apparatus

Instrumented Vehicle

A 1993 Honda Accord was instrumented with video equipment that recorded output from two night vision cameras installed near the vehicle grill, and a data collection computer that recorded vehicle position using a differential GPS synchronized to one of the video recorders. Figure 1 shows the primary instruments installed on the vehicle.
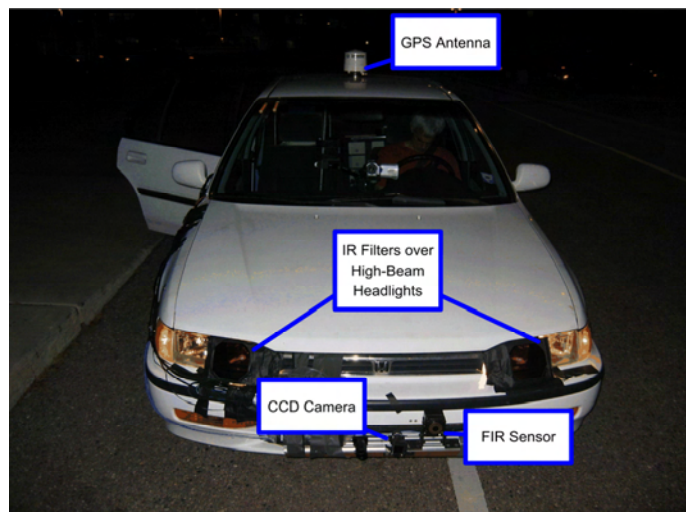

Figure 1. Instrumented vehicle.

Night Vision Enhancement System - FIR

An Autocam Autoliv passive IR sensor, installed near the grill of the vehicle, was used to collect white-hot thermal video images via a composite monochrome video connection. This long IR wavelength (8-14 microns) thermal sensor used an uncooled microbolometer with IR sensitivity of less than100 mK. The sensor array size was 320×240 pixels and the camera provided a field of view (H × V) of 36 × 18 deg (which was later cropped to 24 × 12.5 deg).

Custom-Built Night Vision Enhancement System - NIR

The NIR NVES consisted of a CCD camera (Supercircuits PC164C) using a Sony SuperHAD 1/3" monochrome CCD, a 6 mm lens, and an NIR-pass filter. High beam headlamps, covered with two layers of NIR-pass filters, were constantly on. The filters (Edmund Industrial Optics, model NT43-951) were optical cast-plastic filters with transmission above 90% over 700 nm, and below 1-2% under 660 nm. Positioning two filters on top of each other reduced the visibility of illuminators in the visible range, while keeping the transmission in the NIR range above 85%. The camera provided a resolution of 510 (H) × 492 (V) pixels and a horizontal field of view of about 48 deg (which was later cropped to 24 × 12.5 deg). Its high sensitivity to light level (0.0003 lux minimum illumination) and automatic shutter for gain control (1/60 – 1/100,000) provided good images on dark roads and quick response to changes in illumination. No additional real-time image processing or filtering to improve the image and reduce glare from oncoming vehicles was performed.

The experimental NIR night vision enhancement system was compared to an off-the-shelf system that was available to the experimenters after the experiment had been completed. The purpose of this comparison was to highlight the differences between the experimental system and a single sample of a production-level system in order to further validate the applicability of the results of this experiment.

Differences between the systems were expected in terms of camera characteristics (camera resolution, sensitivity to light, sensitivity to glare, etc.) and illuminator characteristics (mainly power). Figure 2 shows a comparison of an off-the-shelf NIR system (left) with the NIR system used in the experiment (right). The main differences between the NIR systems appeared to be in their sensitivity to glare. For example, compare the shape and size of the halo around the streetlight in the two top images and the brightness of the road in front of the vehicle. However, the effect of glare from oncoming vehicles and streetlights is not likely to affect the detection of pedestrians who are on the right side of the street. The only source of glare that is likely to affect detection of pedestrians is glare from any reflective signs located near the pedestrian. Both systems seemed equally likely to be affected by glare from reflective signs. The main difference

was in glare from active light sources. In summary, we believe the custom NIR system used in this experiment, although more sensitive to glare than an off-the-shelf system, was representative enough of NIR systems to make reasonable inferences about the differences between the NIR and FIR systems with respect to detecting pedestrians.

Figure 2. Visual comparison of image quality. Left: NIR off-the-shelf system; Right: NIR custom system similar to NIR from experiment.

Route and Pedestrians

Routes were chosen to represent a range of road types where pedestrian fatalities attributable to darkness are most likely to happen (Sullivan & Flannagan, 2001). They consisted of urban streets, main arterials, and rural roads.

The instrumented vehicle was driven on six routes in Ann Arbor on July 8, 2004, from 1:00 to 3:00 am. Five pedestrians were positioned in predetermined locations on each route on the right side of the road. Pedestrians stood still, facing the instrumented vehicle, in positions similar to where real pedestrians might stand or walk at night. Table 1 describes the routes driven.

Table 1
Route description.

Route description	Road type and speed limit
East on Baxter Rd. North on Green Rd. East on Plymouth Rd.	Combination of road types beginning with a 2-lane road turning into a 4-lane arterial
West on Plymouth Rd. South on Green Rd. West on Baxter Rd.	Speed limits: 35-50 mph
South on Dixboro Rd.	2-lane rural road, no street lights
North on Dixboro Rd.	Speed limit: 45 mph
West on Washtenaw Ave. Merge to west on Stadium Blvd.	4-lane main arterial, many street lights, other light sources from stores and gas stations on side of road
East on Stadium blvd. Merge to east on Washtenaw Ave.	Speed limit: 45 mph

Five pedestrians (three men and two women) participated in the image-collection session. They were positioned on each of the six routes driven for a total of 30 pedestrian targets. Figure 3 shows a "lineup" view of the pedestrians with the FIR camera (top) and NIR camera (bottom).

Prior to the image-collection session, a few of the subjects where asked to bring several sets of clothing of different materials and colors. These clothes where then tested for reflectance in the NIR spectrum. All but one out of about 30 pieces of clothing appeared bright in the NIR system (the pants of the rightmost pedestrian in the bottom frame of Figure 3). All other clothes appeared bright, although some of them were in fact black in the visible spectrum.

Line of Sight

We inspected FIR and NIR tapes to ascertain the longest distance at which there was a clear line of sight to each of the pedestrians. Line of sight was less than 500 m in all except two cases. We believe that this method of determining line of sight is representative of the actual line of sight in the cases under 500 m. For the two cases in which line of sight was greater than 500 m, it is possible that clear line of sight to the pedestrian is underestimated because of resolution limits of the FIR system.
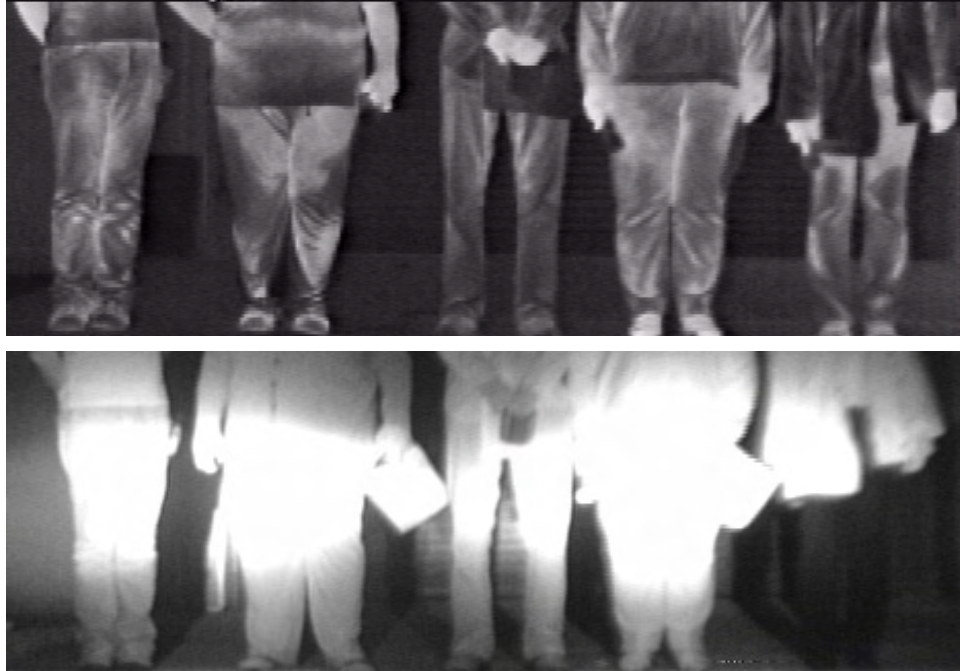
Figure 3. Lineup images of five pedestrians participating in the image-collection session. FIR (top) and NIR (bottom).

Video Manipulation

Video clips from the FIR and NIR cameras were recorded simultaneously by two DVCAM digital videocassette recorders (Sony DSR-20). Simultaneous recording allowed a direct comparison between the cameras, as the produced video clips differed only by the type of camera used. Dynamic events, mainly related to traffic and the speed of the instrumented vehicle, were the same for both cameras.

The process of extracting video clips was intended to convert the collected video footage to digital clips of manageable size and to achieve similar screen size and frame rates between the camera types while minimizing image degradation. Digital video from DVCAM tapes was transferred digitally to a PC and saved in raw audio video interleave format (avi). Compressed clips were then generated using video processing software utilities including AVISynth, VirtualDub, and DirectShow. Filters that were applied to the source video included: (1) deinterlace (LEADtools), (2) crop, and (3) compress (3ivx D4 4.5.1). The FIR video was cropped from 720×480 to 466×240 pixels. The NIR video was cropped from 720×480 to 358×184 pixels and then resized on the screen to 466×240. (The actual sensor resolution for each camera is as described earlier. The image size of 720×480 represents the image size after capturing to a computer.) Figure 4 shows a side-by-side comparison of a few sample frames in FIR and NIR for one pedestrian.
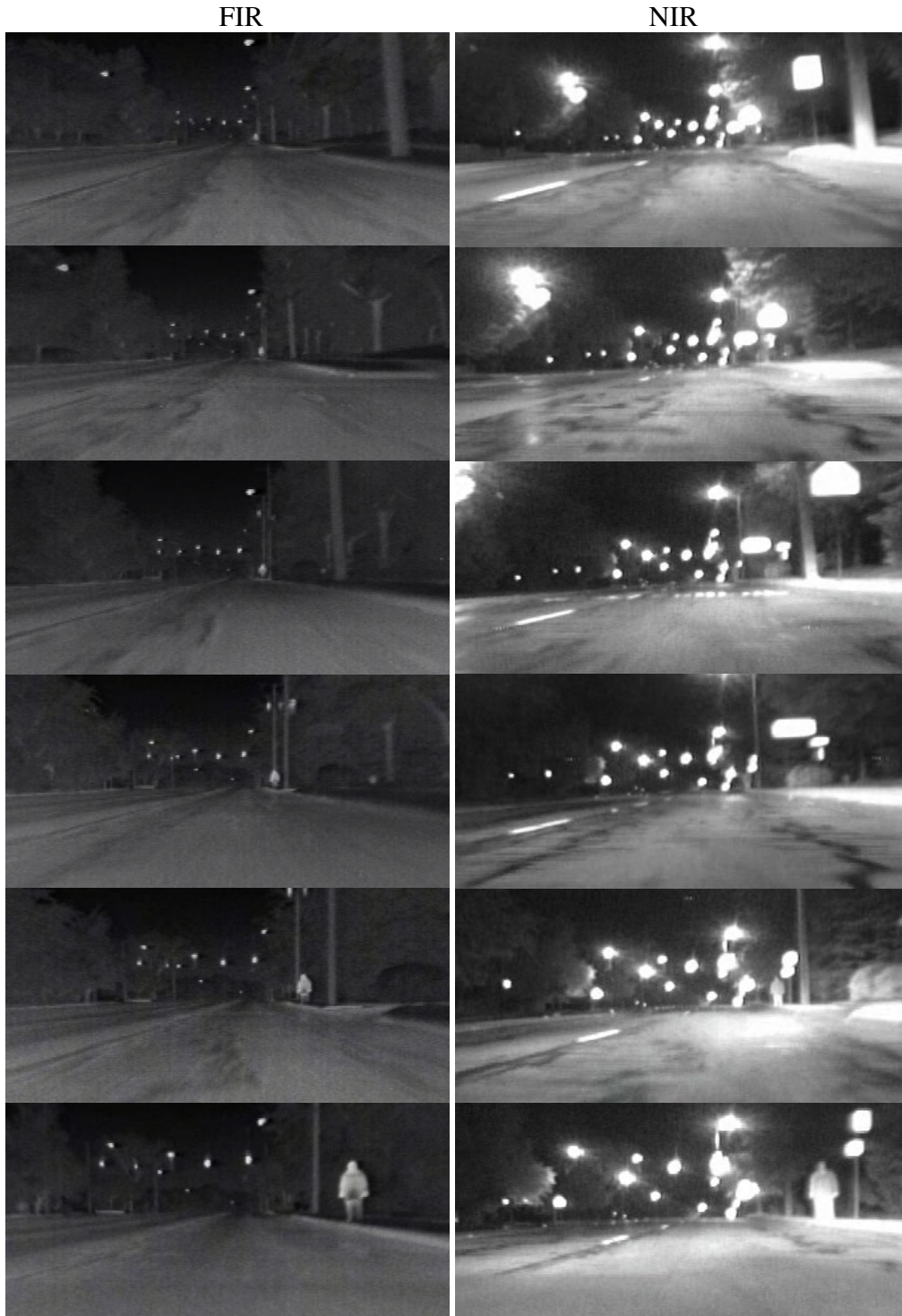
9

FIR   NIR

Figure 4.   Sample images for FIR (left) and NIR (right) at 2 s (~40 m) intervals approaching the pedestrian.

Experimental Setting

The experiment was conducted in the laboratory setting shown in Figure 5. Movie clips were displayed on a Dell UltraSharp 19" LCD covered by a black frame with an opening slightly larger than the viewable image. Movie clips, which were 466×240 pixels, appeared on the screen at a size of 13.7×7.1 cm. The size of the display was chosen to resemble 6-inch diagonal in-vehicle displays. The display was positioned directly in front of the subject's seat, 70-80 cm forward of and 0-20 cm below their horizontal line of sight, depending on their height and posture.

The magnification of the movie, as viewed by the subject, was 1:2.3 (minification of 2.3). The camera view, which covered a field of view of 24×12.5 deg, spanned a field of view of about 10.4×5.4 deg at the subject's eye.

Figure 5. Diagram of the experimental setup.

Procedure

Subjects were divided randomly into two groups to balance the order of presentation of the NIR and FIR views. After filling out consent and biographical forms, and before discussing anything about NVES, subjects viewed three short clips of night driving: a clip of system A, a clip of system B, and a third clip showing system A above system B. They then rated the effectiveness of each system on a 7-point scale. Systems A and B represented the FIR and NIR systems, respectively, for half of the subjects. For the other half, they represented NIR and FIR, respectively.

The experimental session began with three practice clips from system A followed by three practice clips from system B. Subjects were asked to tap the spacebar on a keyboard in front of them "as soon as you see a pedestrian." A confirmation tone indicated that their key press had been registered. After about 10 minutes of practice, the test trials began. Each subject viewed 12 clips of 5-7 minutes each, with breaks after every three clips (about 20 minutes).

The order of clip presentation followed an ABBA pattern to balance practice effects. The first three clips were of system A, followed by six clips of system B, and ending with another three clips of system A. Within each group of three clips, all three routes (Table 1) were included but were presented in an order that was randomized between subjects. To reduce memorization of the location of pedestrians, clips of NIR and FIR taken on the same route were not shown in succession.

After completing the assignment, subjects filled out a post-test form in which they repeated the rating of system effectiveness and answered a few additional questions.

Experimental Design and Data Analysis

The experimental design examined the effect of two NVES systems (NIR and FIR) with six blocks (route driven) nested within subject, and with age (younger and older) and gender (female and male) as between-subject variables. A repeated measures analysis was performed for the following dependent variables and, in addition, paired t-tests were run for detection distances for each of the 30 pedestrian stimuli.

(1) Detection distance – defined as the straight-line distance between the vehicle and the target object when detection was reported.

(2) Detection accuracy – defined as the percentage of targets detected at or before passing the target.

(3) Subjective evaluation of the appearance of the system and its potential usefulness.

RESULTS

Detection Distance

A repeated measures ANOVA of detection distances revealed a significant main effect of system type, $F(1,12) = 174.7$, $p < .001$, subject age group $F(1,12) = 31.6$, $p < .001$, and the interaction between them, $F(1,12) = 13.4$, $p < .01$. On average, younger subjects detected pedestrians at 70% longer distances than did older subjects (Figure 6). The interaction between age and system type appears to involve increases in seeing distance that are proportionally the same for the younger and older groups. If the data are log transformed, the main effects remain significant but the interaction between them is eliminated.
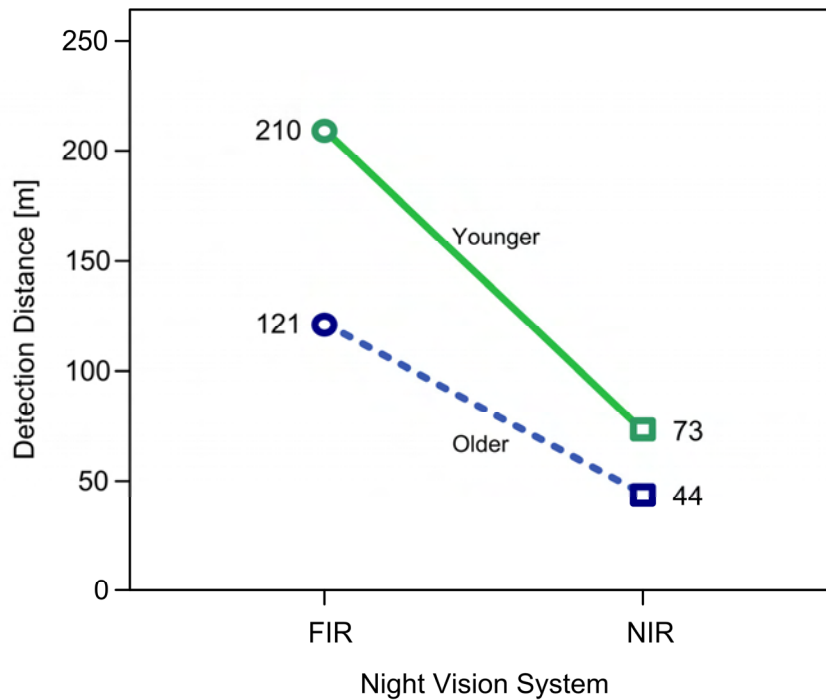

Figure 6. Detection distance by system type and subject age group.

Figure 7 shows the mean detection distances for each of the 30 pedestrian targets with the FIR and NIR systems, sorted by the detection distance with the FIR system. The mean difference between systems for individual targets ranged from 25 m to 206 m. All but three paired t-tests yielded highly significant differences $t(15) > 4.67$, $p < 0.001$. For three pedestrians, t-tests were not as significant, $t(15) = 2.56$, $p < 0.05$; $t(15) = 2.98$, $p < 0.01$; and $t(15) = 3.44$, $p < 0.01$. For these three pedestrian targets, the line of sight to the pedestrian was

physically limited to about 150 m or less, thus reducing the maximum possible detection distances and the relative advantage of the FIR system.
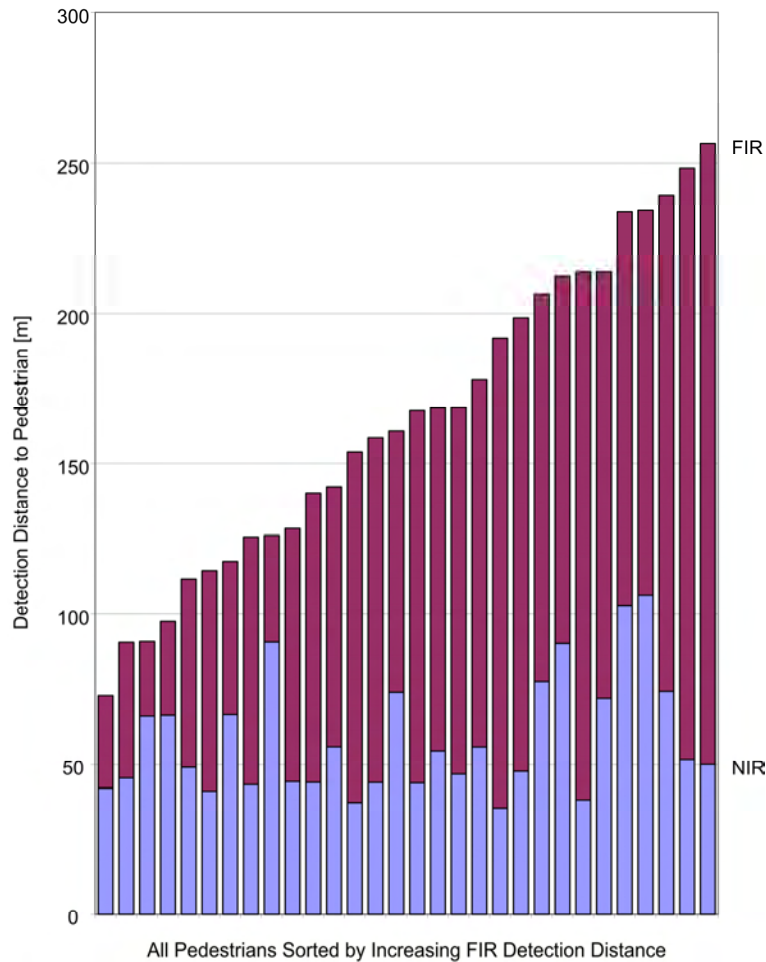

Figure 7. Mean detection distance for the individual pedestrians with the two system types.

Further analysis of detection distances was performed with respect to the clear line of sight that we estimated for each pedestrian encounter. A linear regression of mean FIR detection distance on clear line of sight accounted for 61% of the variability in detection distance (Figure 8). Detection distance tended to increase with the line of sight to pedestrians. A similar regression for the NIR detection distance did not account for a substantial amount of the variance (less than 1%), suggesting that performance with the NIR system was not affected by line of sight limitations but by other limitations.

We considered visual clutter in the images as an alternative explanation for the variation in detection distances with NIR. Visual clutter of sample images for each pedestrian was rated subjectively by 5 people naïve to the purpose of the experiment. It was expected that cluttered images (resulting from proximity of the pedestrian to distracters and other bright objects in the scene) would reduce the ability of subjects to see pedestrians early.

14

A linear regression of mean NIR detection distance on clutter rating accounted for 25% of the variability in detection distance (Figure 9). Detection distance decreased with increase in clutter in the scene. A similar regression of FIR detection distance accounted for 24% of the variability in detection distance but had a steeper slope.
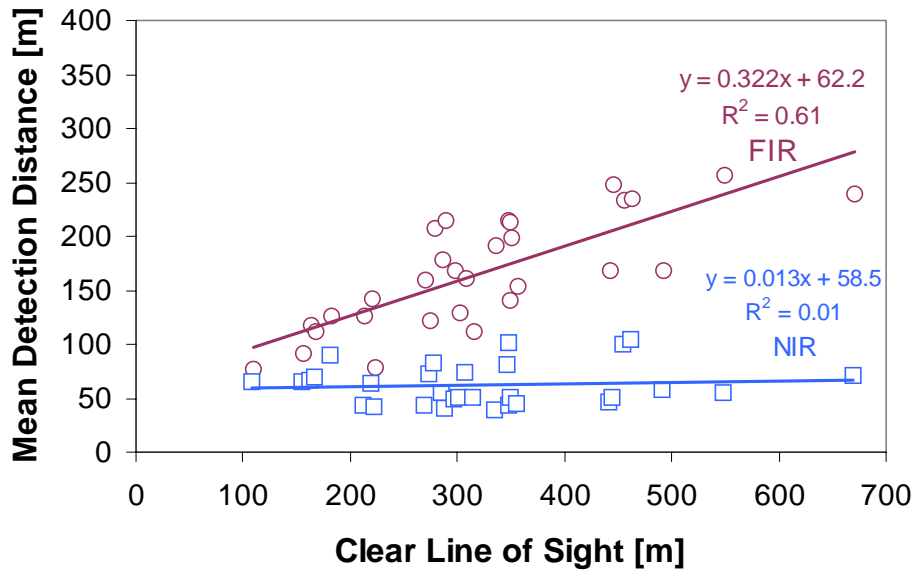

Figure 8. Regression of NIR and FIR detection distances on maximum viewing distance (clear line of sight).
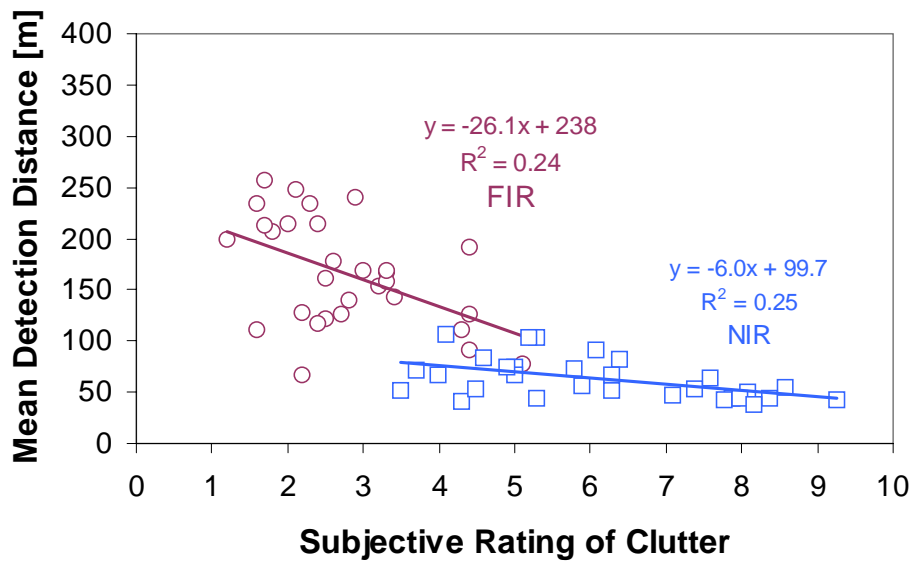

Figure 9. Regression of NIR and FIR detection distances on subjective rating of clutter.

Detection Accuracy

Overall, pedestrians were missed in 34 out of 960 trials (3.5%). The percentage of misses with FIR (1.0%) was lower than with NIR (6.0%) (Table 2). With FIR, the misses were distributed evenly between the two age groups. With NIR, however, older subjects had substantially more misses than did younger subjects.

Table 2
Number of missed pedestrians as a function of system type and age-gender group.

Age group	Number of misses (percent)	
	FIR	NIR
Younger female	1 (0.8%)	2 (1.7%)
Younger male	1 (0.8%)	0 (0%)
Older female	2 (1.7%)	17 (14.2%)
Older male	1 (0.8%)	10 (8.3%)

Subjects indicated that a pedestrian was present when there was no pedestrian potentially visible in the scene (a false alarm) 60 times during the entire experiment. Whether or not a pedestrian was potentially visible was based on the maximum viewing distances to each pedestrian. Older subjects made more false alarms than did younger subjects, but there was no difference between the systems and no interaction between system and age. Table 3 shows the number of false alarms for each system by age group. The rate of false alarms per minute was based on the amount of time when no pedestrian was visible for each system.

Table 3
Number of false alarms as a function of system type and age-gender group.

Age group	Number of false alarms (rate per minute)	
	FIR	NIR
Younger female	7 (0.27)	4 (0.12)
Younger male	3 (0.12)	7 (0.22)
Older female	8 (0.31)	13 (0.40)
Older male	9 (0.35)	9 (0.28)

Subjective Rating of Preference between Systems

Subjects were asked to rate the effectiveness of the two NVES after viewing a short introductory clip of each system and a side-by-side clip from both systems. These ratings were obtained before any other exposure to the experimental NVES scenes. It was expected that some

subjects would prefer the NIR system to the FIR system because NIR scenes are considered to be more realistic than FIR scenes, and drivers tend to assign more weight to the importance of seeing objects close to the vehicle over the importance of seeing pedestrians far away. For the latter reason, the clips were shown to the subjects without any reference to pedestrians or to the specific purpose of the system. For each system, subjects had to answer the question: "How effective would supplementary view A [or B] be in helping you drive safely at night?" Their answer was given on a seven point rating scale with 1 = "not at all effective," 4 = "somewhat effective," and 7 = "very effective." After the experimental trials were concluded, the subjects were presented with the same question, and two additional clarification questions regarding the ease of detecting pedestrians ("How easy is it to detect a pedestrian in view A [or B]?") and the expected frequency of using each system ("While driving at night, how often would you look at a display like view A [or B] if it were installed in your car?")

       Subjective ratings of the effectiveness of the NVES revealed a three-way interaction among the system type, the subject's age group, and whether the rating was given before the trials (pre-test) or after them (post-test), $F(1,12) = 11.8$, $p < .01$ (see Figure 10). In pre-test ratings, younger subjects rated the FIR system as more effective than the NIR system (6.0 vs. 2.9, respectively). Their ratings remained essentially the same after completing the experiment (5.8 vs. 2.9, respectively). Older subjects did not show an overall preference for either of the systems before the trials (4.9 vs. 5.1, respectively). After the experiment, however, older subjects rated FIR as more effective than NIR (5.2 vs. 4.0, respectively).
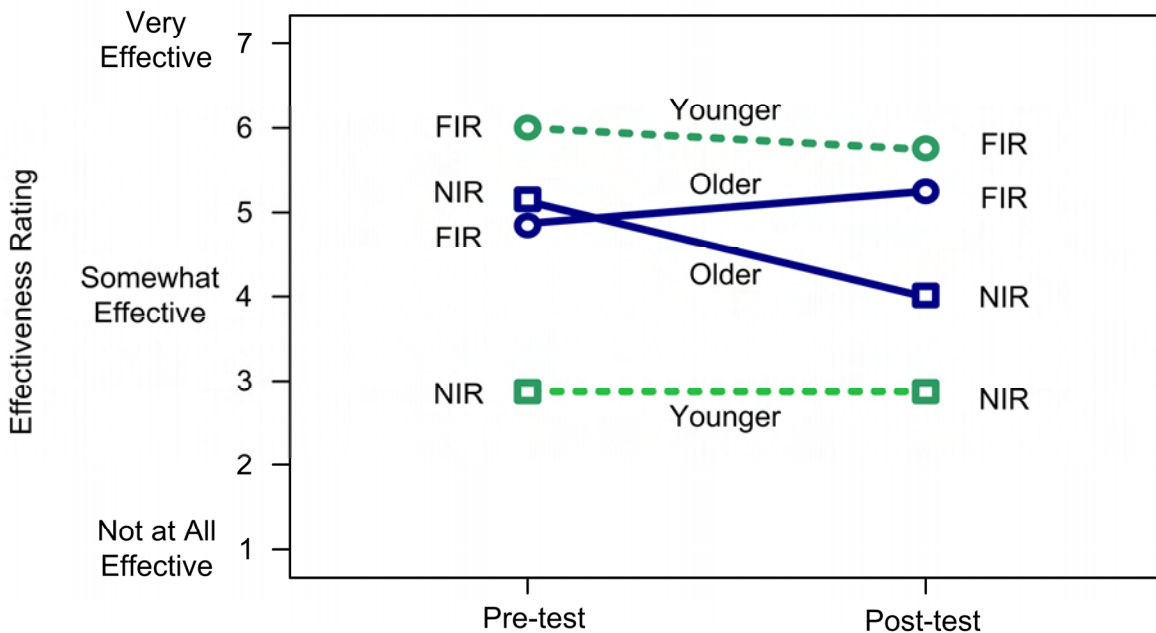

Figure 10. Rating of effectiveness of two NVES systems before and after the experiment.

Post-test, subjects rated detecting pedestrians with the FIR system as "very easy" (mean of 6.4 on a 7 point scale, where 7 = "very easy"). Ten of 16 subjects rated the FIR system as "very easy" (assigning the maximum scale value, 7). In contrast to the ratings of the FIR system, subjects rated detecting pedestrians with the NIR system between "not at all easy" and "somewhat easy" (mean of 3.1).

Subjects were asked to rate on a seven-point scale how often they would look at the display of an NVES in their car. Three anchors were given: 1 = "much less than I use a rearview mirror," 4 = "as much as I use a rearview mirror," and 7 = "much more than I use a rearview mirror." Most subjects said they would use the FIR system about as much as they use their rearview mirror, and only two subjects said they would use it less. The mean rating for FIR was 4.5. In contrast, half of the subjects said they would use the NIR system less, or much less, than they use a rearview mirror. The mean rating for NIR was 2.9

DISCUSSION AND CONCLUSIONS

The purpose of this experiment was to make a direct comparison of FIR and NIR systems based on how well they can help drivers detect pedestrians at night. Underlying this approach was evidence that detecting pedestrians at night is the most important task with which drivers need help in darkness.

The results favored FIR unambiguously. Detection distances with FIR were substantially greater than with NIR for all the pedestrian targets tested. On average, detection distances were 165 m with FIR and 59 m with NIR. The advantage of FIR was present for both younger and older subjects. Detection distances of younger subjects were about 70% greater than those of older subjects.

The effectiveness of NIR and FIR systems can be expected to depend on details of implementation. For example, camera resolution, different display sizes, and display location are likely to affect driver performance. To minimize the effect of details of implementation on the comparison between systems, images from the systems compared in this experiment were shown on the same display, and their field of view and resolution were matched. Furthermore, the scenes on both systems were taken on the road at the same time, with the cameras positioned side by side. Thus, the comparison was designed so that the results would identify differences inherent in each system type.

To the extent that the two systems used in this experiment reasonably represent their respective technologies, the results support the expected enhancement of pedestrian detection in FIR systems relative to NIR systems. Although a single NVES was used in this experiment to represent each system type, other existing systems using the same technology are not expected to change the results of this experiment substantially. Nevertheless, technology advancement and innovation in the future may introduce new NIR systems that are capable of better detection of pedestrians. This experiment can serve as a benchmark for their expected performance. Manufacturers of future NIR systems should address the issue of pedestrian detection and show that detection distances with their systems compare favorably to the values reported here.

In an analysis using subjectively rated image clutter, a negative correlation was found between image clutter and detection distance. When image clutter was low, detection distances were high and when image clutter was high, detection distances were low (see similar results by Aviram & Rotman, 2000). These results suggest that an important issue with night vision images is extra information that clutters the display and interferes with the detection of pedestrians. Interestingly, the more natural images appear, the more clutter is in those images. This counterintuitive finding suggests that FIR systems have an important advantage over NIR systems in that there is less clutter to delay the decision about the presence of a pedestrian in the image.

Future Work

In the current experiment, subjects continuously monitored an NVES display for the presence of pedestrians. In real implementations of NVES, however, drivers glance at the display for short durations that are continuously interrupted by longer glances at the road. Thus, the current experiment compared performance between two types of NVES in best-case conditions. Subjects allocated the entire time to viewing the display, they had no additional workload associated with a driving task, and they did not have to deal with switching costs (associated with glance allocation strategies).

Combining the results of this study with the effect of intermittent glances at the display is a possible next step. Now that the effect of system type with regard to the detection of pedestrians has been quantified, the interaction of system type with temporal limitations and mental workload can be addressed. An experiment in which subjects drive a vehicle or a simulated vehicle while performing the pedestrian detection task would be valuable in revealing those interactions. Another option is to simply manipulate the duration of glances at the display using the task occlusion method (see Tsimhoni & Green, 2003) in which the display is intermittently covered to control the duration of consecutive glances, or simply by briefly displaying single frames from the display and using signal detection theory to characterize the responses.

We speculate that limiting the total glance time at the display would further bias the results in favor of the FIR system. Because the interpretation of the scene would be limited in time, and because the scene would be dynamic, there should be an increased cost associated with a higher number of distracters in the NIR view. As shown in the subjective evaluation of clutter in the current experiment, FIR was less cluttered (had less distracters) than NIR. We expect that visual clutter will be a major source of mistakes and missed signals when viewing time is limited.

An additional possible direction of research involves more detailed prediction of performance based on objective analysis of descriptive parameters of the scene, such as visual clutter. If objective measures of visual clutter (e.g., Aviram & Rotman, 2000) are found to be good predictors of performance in the detection of pedestrians relative to the subjective clutter rating used in the current experiment, they can be used to make predictions of early detection of pedestrians.

REFERENCES

Aviram, G., & Rotman, S. R. (2000). Evaluating human detection performance of targets and false alarms, using a statistical texture image metric. *Optical Engineering, 39* (pp. 2285-2295).

Blanco, M., Hankey, J. M., & Dingus, T. A. (2001). Evaluating new technologies to enhance night vision by looking at detection and recognition distances of non-motorists and objects. *Proceedings of the Human Factors and Ergonomics Society 45th Annual Meeting* (pp. 1612-1616).

Collins, D. J., Piccione, D., & Best, P. S. (1998a). *Driver detection of drop-offs when using thermal and intensified imaging night vision devices*. Alexandria, VA: DCS Corporation.

Holz, M., & Weidel, E. (1998). *Night vision enhancement system using diode laser headlights: Electronics for trucks and buses* (SAE Technical Paper Series No. 98-2778). Warrendale, PA: Society of Automotive Engineering.

Küpper, L., & Schug, J. (2002). *Active night vision systems.* (SAE Technical Paper Series No. 2002-01-0013). Warrendale, PA: Society of Automotive Engineering

Martinelli, N. S., & Boulanger, S. A. (2000). *Cadillac DeVille thermal imaging night vision system.* (SAE Technical Paper Series No. 2000-01-0323). Warrendale, PA: Society of Automotive Engineering.

Meitzler, T., Bryk, D., Sohn, E., Lane, K., Bednarz, D., Jusela, D., Ebenstein, S., Smith G.H., Rodin, Y., Rankin, J.S., & Samman, A.M. (2000). Noise and contrast comparison of visual and infrared images of hazards as seen inside an automobile. *Proceedings of SPIE: Vol. 4023. Enhanced and Synthetic Vision 2000* (pp. 139-153).

Piccione, D., Best, S., Collins, D., & Barnes, J. (1997). *Concept experiment program test of AN/VAS-5 driver's vision enhancer* (Report No. TRADOC 97-CEP-466). Fort McClellan, Alabama: U.S. Army Military Police School.

Rumar, K., (2002). *Night vision enhancement systems: what should they do and what more do we need to know?* (Report No. UMTRI-2002-12). Ann Arbor, MI: The University of Michigan Transportation Research Institute.

Sullivan, J.M., and Flannagan, M.J. (2001). *Characteristics of pedestrian risk in darkness.* (Report No. UMTRI-2001-33). Ann Arbor, MI: The University of Michigan Transportation Research Institute.

Sullivan, J.M., Bärgman, J., Adachi, G., & Schoettle, B. (2004). *Driver performance and workload using a night vision system.* (Report No. UMTRI-2004-8). Ann Arbor, MI: The University of Michigan Transportation Research Institute.

Tsimhoni, O., and Green, P. (2003). Time-Sharing of a Visual In-Vehicle Task while Driving: Effects of Four Key Constructs. *Proceedings of the 2nd international driving symposium on human factors in driver assessment, training, and vehicle design*, Park City, Utah, pp. 113-118.

Toyofuku, K., Iwata, Y., Hagisato, Y., & Kumasaka, T. (2003). *The "night view system" using Near-Infrared light: Intelligent vehicle initiative* (SAE 2003-01-0018) Warrendale, PA: Society of Automotive Engineering.

Tsimhoni, O., & Green, P. (2002). *Night vision enhancement systems for ground vehicles: the human factors literature* (Report No. UMTRI-2002-05). Ann Arbor, MI: The University of Michigan Transportation Research Institute.

Yagi, S., Kobayashi, S., Inoue, T., Hori, T., Michiba, N., & Okui, K. (2003). *The development of an Infra red projector: Lighting Technology* (SAE Technical Paper Series No. 2003-01-0987). Warrendale, PA: Society of Automotive Engineering.