

UMTRI-2005-23

**PEDESTRIAN DETECTION
WITH NIGHT VISION SYSTEMS
ENHANCED BY AUTOMATIC WARNINGS**

**Omer Tsimhoni
Michael J. Flannagan
Takako Minoda**

September 2005

PEDESTRIAN DETECTION WITH NIGHT VISION SYSTEMS
ENHANCED BY AUTOMATIC WARNINGS

Omer Tsimhoni
Michael J. Flannagan
Takako Minoda

The University of Michigan
Transportation Research Institute
Ann Arbor, Michigan 48109-2150
U.S.A.

Report No. UMTRI-2005-23
September 2005

Technical Report Documentation Page

1. Report No. UMTRI-2005-23		2. Government Accession No.		3. Recipient's Catalog No.	
4. Title and Subtitle Pedestrian Detection with Night Vision Systems Enhanced by Automatic Warnings				5. Report Date September 2005	
				6. Performing Organization Code 302753	
7. Author(s) Tsimhoni, O., Flannagan, M.J., and Minoda, T.				8. Performing Organization Report No. UMTRI-2005-23	
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 The Affiliation Program currently includes Autoliv, Avery Dennison, Bendix, BMW, Bosch, Com-Corp Industries, DaimlerChrysler, DBM Reflex, Decoma Autosystems, Denso, Federal-Mogul, Ford, GE, General Motors, Gentex, Grote Industries, Guide Corporation, Hella, Honda, Ichikoh Industries, Koito Manufacturing, Lang-Mekra North America, Magna Donnelly, Muth, Nissan, North American Lighting, OLSA, OSRAM Sylvania, Philips Lighting, PPG Industries, Renault, Schefenacker International, Sisecam, SL Corporation, Solutia Performance Films, Stanley Electric, Toyoda Gosei North America, Toyota Technical Center USA, Truck-Lite, Valeo, Visteon, 3M Personal Safety Products, and 3M Traffic Safety Systems. Information about the Affiliation Program is available at: http://www.umich.edu/~industry/					
16. Abstract <p>This experiment compared pedestrian detection using far-infrared (FIR) and near-infrared (NIR) night vision systems, combined with automatic warnings at one of two distances or no warning at all. Sixteen subjects (eight younger than 30 years and eight older than 60 years) pressed a button as soon as they saw a pedestrian on a night vision system in the center console of a vehicle simulator. In addition, they performed a concurrent simulated steering task that required almost continuous viewing of the forward scene, similar to real driving.</p> <p>As in a previous experiment (Tsimhoni, Bärghman, Minoda, and Flannagan, 2004), detection distances with FIR systems were substantially greater than with NIR systems. Detection distances with both systems were shorter than in the previous experiment by about 20 m, probably because of the addition of simulated steering in the present experiment. The automatic visual warning was a blue rectangle that zoomed in on the pedestrian in the video display. In the long-distance condition, it was presented when the pedestrian was 150 m away. Detection distance and accuracy for both night vision systems increased, but the effects were more prominent for the NIR system. Automatic warnings at 75 m improved performance with NIR but worsened performance with FIR, perhaps because in some trials subjects waited for the automatic warning before responding. Subjective ratings of mental workload and of effort were higher for NIR than for FIR, but the addition of automatic warnings did not decrease perceived workload significantly. Overall, automatic visual warnings based on image processing were effective in increasing accuracy and detection distance for pedestrians except when short-distance warnings were used with the FIR system.</p>					
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 30	22. Price

ACKNOWLEDGMENTS

Appreciation is extended to the members of the University of Michigan Industry Affiliation Program for Human Factors in Transportation Safety for support of this research. The current members of the Program are:

Autoliv	Magna Donnelly
Avery Dennison	Muth
Bendix	Nissan
BMW	North American Lighting
Bosch	OLSA
Com-Corp Industries	OSRAM Sylvania
DaimlerChrysler	Philips Lighting
DBM Reflex	PPG Industries
Decoma Autosystems	Renault
Denso	Schefenacker International
Federal-Mogul	Sisecam
Ford	SL Corporation
GE	Solutia Performance Films
General Motors	Stanley Electric
Gentex	Toyoda Gosei North America
Grote Industries	Toyota Technical Center, USA
Guide Corporation	Truck-Lite
Hella	Valeo
Honda	Visteon
Ichikoh Industries	3M Personal Safety Products
Koito Manufacturing	3M Traffic Safety Systems
Lang-Mekra North America	

CONTENTS

INTRODUCTION	1
METHOD	4
Subjects	4
Apparatus	4
Procedure	10
Experimental Design and Data Analysis	11
RESULTS	13
Detection Distance	13
Detection Accuracy.....	17
Glance Behavior.....	19
Subjective Rating of System Effectiveness	20
Subjective Rating of Workload.....	21
DISCUSSION AND CONCLUSIONS	22
REFERENCES	26

INTRODUCTION

This experiment is a follow-up to a previous experiment we conducted on night vision systems (Tsimhoni, Bärghman, Minoda, and Flannagan, 2004). In that experiment, two major sensing technologies for automotive night vision systems were compared in a laboratory experiment: 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. We argued that pedestrian detection should be the primary objective of any night vision system (Rumar, 2002; Sullivan and Flannagan, 2001). We compared detection distances for pedestrians in both systems using matched stimuli and found that detection distances with FIR were overwhelmingly greater than with NIR.

The present experiment extends the previous experiment in two ways. First, subjects viewed the night vision display intermittently as they had to perform a concurrent simulated driving task that required viewing the forward scene almost continuously. Consequently, their overall performance was expected to degrade, but it was not clear by how much and whether the degradation would be proportional for both night vision systems. Second, some of the night vision scenes were artificially enhanced by automatic pedestrian detection at two set distances from the pedestrian. As discussed below, automatic pedestrian warnings are receiving great development interest and are beginning to become available on new vehicles. It was expected that pedestrian warnings would improve performance but it was not clear how they would change the drivers' scanning pattern of the display and whether the distance at which warnings appear would affect detection distance and detection probability.

Pedestrian Detection Systems

In the last decade or so, pedestrian detection systems for automotive applications have been a focus of engineering research. For example, a search for the term "pedestrian detection" was performed in the online library of the IEEE (ieeexplore.ieee.org), which indexes over 1.2 million publications. Of the 61 publications from 1990 to 2004 on pedestrian detection, more than three quarters (46 of 61) were published in the last three years and more than a third (26 of 61) were published in the last year (2004). Most of these publications discuss new algorithms for the detection of pedestrians from a moving vehicle.

The objective of pedestrian warning algorithms is to accurately detect pedestrians and provide the driver with informative warnings. In the eyes of the driver, the end product of a good system provides a timely warning and, possibly, additional information such as the position of the pedestrian or an overlaid icon on a night vision display. Although generic image processing algorithms have been addressing similar goals for many years, there are several problems that are unique to image processing in automotive applications. For example, it is difficult to distinguish between objects in the foreground and the background of the image because the entire image is continuously changing and because pedestrians vary in scale based on their distance to the viewer (Xu, Liu, and Fujimura, 2005). Various algorithms have been suggested in an effort to solve these issues. For example, stereo-based detection allows explicit occlusion analysis, is robust to illumination changes, and helps approximate the size of the target based on distance estimation (Zhao and Thorpe, 2000). Other methods, often integrating multiple approaches such as background elimination and analysis, periodic motion, symmetry, and silhouette shape analysis, have been proposed (see Xu, Liu, and Fujimura, 2005 for the latest review).

Although the objective of detecting pedestrians from visual inputs is achieved fairly well by most of the proposed systems, there is still much room for improvement. Typical accuracy values for recently-published detection systems are still far from optimal (percent hit / percent false alarm): 75-90% / 100% (Nanda and Davis, 2002), 85% / 3% (Zhao and Thorpe, 1999), 84% / 19% and 92% / 3% per frame (Fang et al., 2003), 93% / 5% (Xu, Liu, and Fujimura, 2005), 70% / 20% (Bertozzi et al., 2004), and 85-93% / 0.01-0.1% (Shashua, Gdalyahu, and Hayun, 2004). (Methods of calculation differ among experiments. These values are presented here for general impression only.)

As pointed out by Parasuraman, Hancock, and Olofinboba (1997), if the probability of true warnings (i.e., when the driver is about to hit a pedestrian) is low, as it often is in reality, then the odds of a true alarm can be quite low even for very sensitive warning systems with very high hit rates and low false alarm rates. For example, if the probability of a true event to occur is 0.001, a system with a hit rate of 99% and a false alarm rate of 1% will produce 11 alarms for every true event it detects. Because the probability of hitting a pedestrian in reality is substantially less than 0.001, as in this example, the odds of a true alarm are substantially less than 10%. One way around this problem is to alert the driver to more frequent events. For example, Farber and Paley (1993) discussed systems that give an alarm in collision-possible conditions. Because collision-possible events occur more often than actual collisions, the system will be correct more often, thus gaining the driver's credibility. A side benefit of providing alerts

more often is that when a lifesaving alarm occurs, drivers are less surprised because they are familiar with the alarm. Another solution is to present graded alarms which inform the driver of the severity level of the event (Sorkin, Kantowitz, and Kantowitz, 1988).

The accuracy values described above present the accuracy of the system without taking into account the accuracy of the driver. To describe the overall value of the system, system accuracy must include the driver as an integral part of the system. If the driver misses, or chooses to ignore, a pedestrian warning on the night vision system, the overall accuracy of the system will decrease. On the other hand, if the night vision scene is projected on an in-vehicle display or a HUD in addition to the warning, the driver may detect pedestrians on the display before the system provides a warning, or perhaps even detect pedestrians that are missed by the system. In that case, overall system accuracy has the potential to be higher than that of the automatic detection system alone.

In the present experiment, we chose to focus on detection by the driver and assume perfect detection by the system. Two detection distances (150 m and 75 m) were used for both night vision systems. At 45 mph (72.5 km/hr), 150 m is thought to be the minimum visibility distance that gives the driver an opportunity to detect and identify objects and to choose an appropriate action (Rumar, 2001, p. 6). The short detection distance (75 m) is just above the minimum visibility distance to allow an emergency stop at 45 mph (based on a response time of 1.5 s and a brake force of 0.5 g). These two detection distances (150 m and 75 m) happen to be approximately representative of the distances at which automated warnings could be calculated and generated in current FIR and NIR systems, respectively. In some cases, the pedestrian might be detected by the driver on the in-vehicle display before the warning appears. In other cases, the driver might still need to confirm the presence of a pedestrian after the warning appears because the pedestrian might be difficult to see or not yet visible.

In the existing human factors research on night vision systems, several studies have addressed the issue of detection distance and accuracy (see Tsimhoni and Green, 2001 for a review). None of those studies, however, have addressed the effects of automatic warnings on detection distance and accuracy.

In light of the above, this experiment addressed two main issues:

1. The effect of automatic pedestrian warnings on a driver's ability to detect pedestrians and on subjective workload using two types of night vision systems.
2. The effect of adding a steering task that requires intermittent glances away from the night vision display on pedestrian detection. This issue would be addressed by comparing results to a previous experiment in which a steering task was not used.

METHOD

Subjects

Sixteen licensed drivers participated in this experiment—eight younger (ages 21 to 28 years, mean of 23) and eight older (ages 62 to 78 years, mean of 71), with equal numbers of men and women in each age group.

All subjects' corrected vision (tested with an Optec 2000 Stereo Optical Vision Tester) was 20/40 or better. Midrange acuity (80 cm) was 20/40 or better for younger subjects, and 20/100 or better for older subjects (mean of 20/45). None of the subjects had driven with a night vision enhancement system before but three subjects (a young man, a young woman, and an older woman) reported that they were familiar with the concept.

Apparatus

Instrumented Vehicle

To collect night vision system video footage for later presentation in a laboratory setting, a 1993 Honda Accord was instrumented with video equipment that recorded output from two night vision cameras installed near the vehicle grille, a camcorder mounted inside the vehicle, 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.

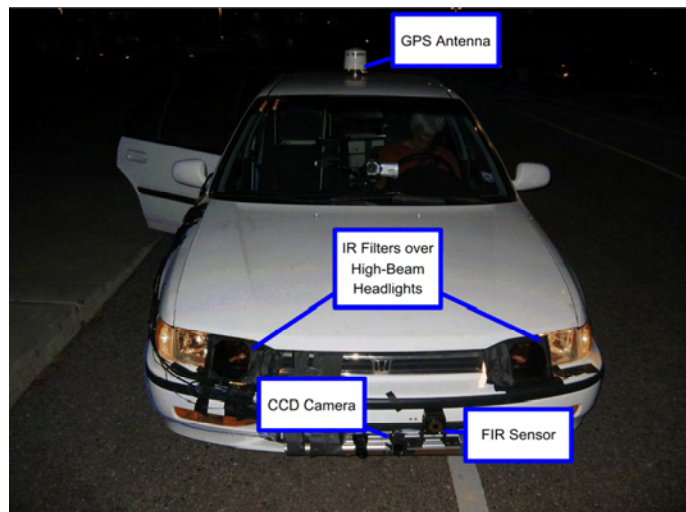


Figure 1. Instrumented vehicle.

Night Vision Enhancement System - FIR

An Autocam Autoliv passive IR sensor, installed near the grille 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 than 100 mK. The sensor array size was 320×240 pixels and the camera provided a field of view (H) \times 18 (V) deg (which was later cropped to 24×12.5 deg).

Custom-Built Night Vision Enhancement System - NIR

The NIR night vision system 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) \times 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.

Driver's View Camera

A Sony Camcorder (DCR-TRV30) was mounted on the dashboard inside the vehicle to record the forward scene from next to the driver's eye position. The camcorder lens was covered by a filter that blocked the NIR light produced by the custom IR headlights. The images produced by this camcorder were a rough representation of an unaided night view of the road as seen by the driver.

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 and Flannagan, 2001). They consisted of urban streets, main arterials, and rural roads.

The instrumented vehicle was driven on these 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 four routes used in this experiment. Five pedestrians (three men and two women) participated in the image-collection session. They were positioned on each of the four routes driven, for a total of 20 pedestrian targets. Figure 2 shows a “lineup” view of the pedestrians with the FIR camera (top) and NIR camera (bottom).

Table 1
Route description.

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

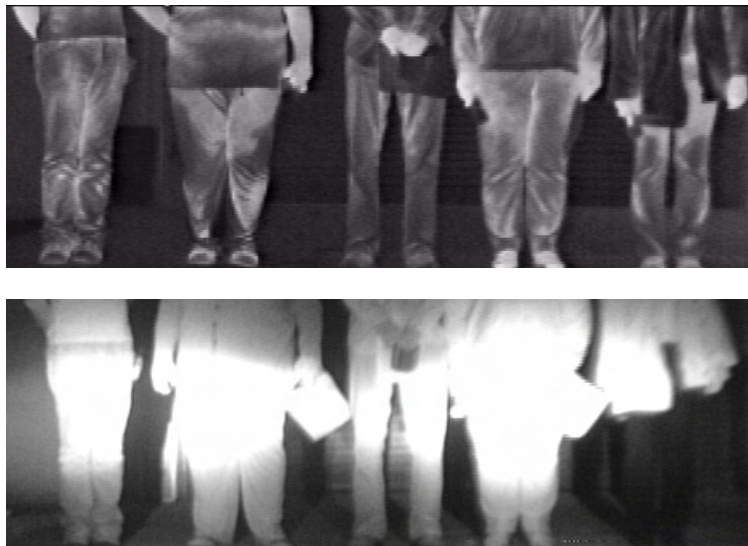


Figure 2. Lineup images of five pedestrians participating in the image-collection session. FIR (top) and NIR (bottom).

Video Manipulation

Video clips from the FIR and NIR cameras and from the camcorder were recorded simultaneously by three DVCAM digital videocassette recorders (Sony DSR-20). Simultaneous recording allowed a direct comparison between the night vision cameras, as the video clips produced 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 all 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).

A simulation of pedestrian warnings was generated by overlaying rectangle shapes on pedestrians in the existing movie clips. Short-distance warnings appeared at a distance of 75 m from the pedestrian and long-distance warnings appeared at a distance of 150 m from the pedestrian. Figure 3 shows a side-by-side comparison of movie clips with 150 m warnings in FIR and NIR.

The warnings were designed so that they would be visible in the subject's peripheral vision when displayed on the center console. A blue and white rectangle first spanned the entire screen and then zoomed in on the pedestrian. The large movement on the screen helped draw the subject's attention and assisted the driver in detecting the position of the pedestrian on the screen.

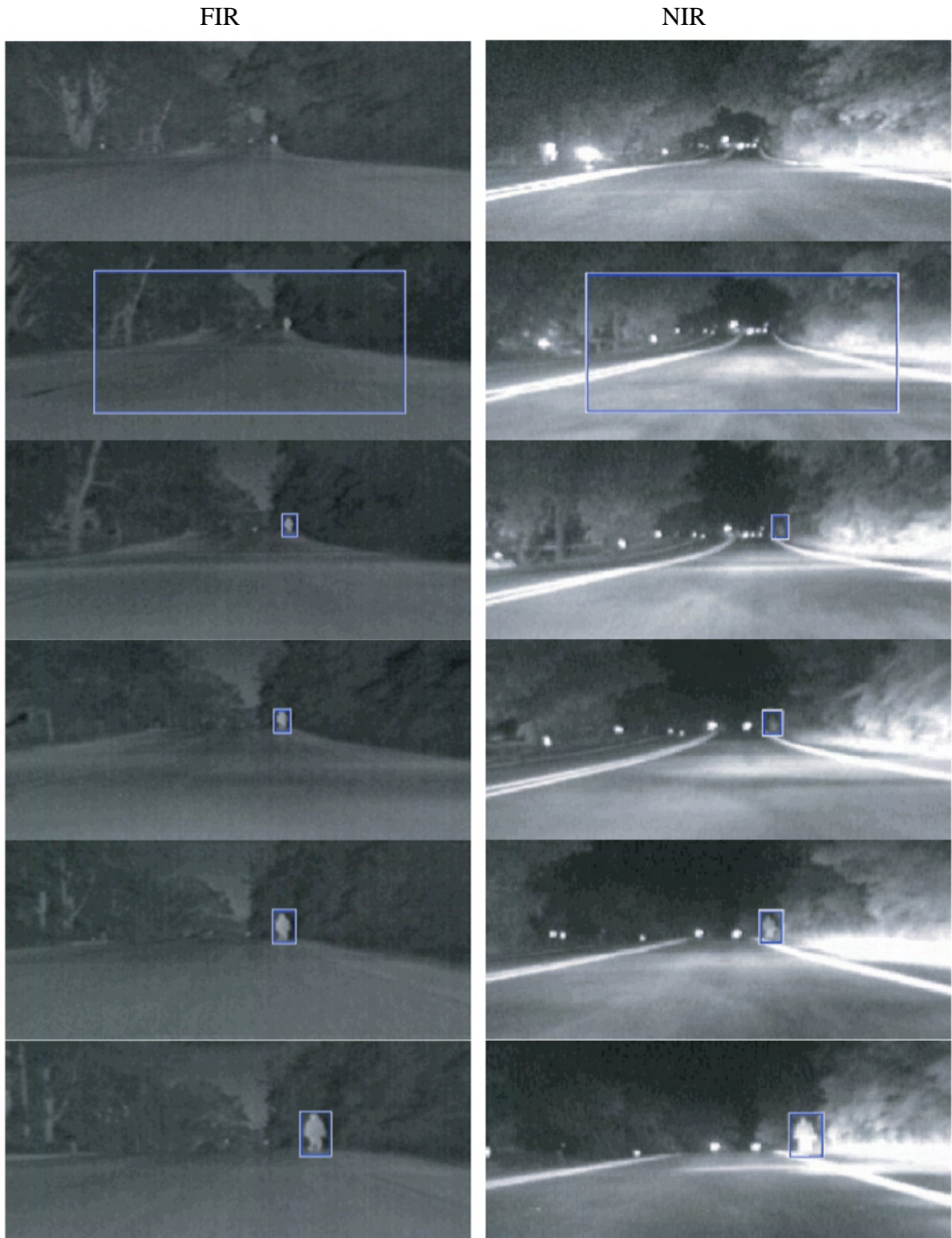


Figure 3. Sample images for FIR (left) and NIR (right) at about 30 m intervals approaching the pedestrian. At 150 m from the pedestrian, a blue rectangle warning appeared on the entire screen and then zoomed in on the pedestrian.

Experimental Setting

The experiment was conducted in the UMTRI Driving Simulator Laboratory. Movie clips were displayed on a 12.1" LCD (SlimAge 400A) covered by a black frame with an opening slightly larger than the viewable image. Movie clips, which were 518×242 pixels, appeared on the screen at a size of 15.2×7.1 cm. The size of the display was chosen to resemble six-inch diagonal in-vehicle displays. The display was positioned in the center console, 60-65 cm forward of, 25-30 cm below, and 35 cm to the right of the subject's horizontal line of sight, varying to some extent with the height and posture of the subject. Direct distance from between the subject's eyes to the center of the display was within 74-80 cm for all subjects. 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.

Tracking Task

As the night vision system movie clip was shown on the center console of the vehicle, a synchronized movie clip that had been recorded by the camcorder in the experimental vehicle was projected on the forward screen of the simulator laboratory, which was about 4.5 m in front of the subject (Figure 4). The subject was required to move the steering wheel so that a red dot, produced by a laser pointer on the steering wheel, would remain in the center of the lane. The task was chosen to resemble the visual and motor demands of an actual steering task while fitting naturally in the simulated scenario.

Detection Task

Subjects were instructed to press a button as soon as they saw a pedestrian in the in-vehicle display. To allow subjects to keep both hands on the steering wheel while pressing the button and to avoid long and uneven delays before key presses, the button was mounted on the subject's right index finger so that it could be depressed against the steering wheel without the subject having to change his or her grip of the steering wheel.



Figure 4. Image of the experimental setup.

Procedure

Subjects were divided randomly into four groups to balance the order of presentation of the NIR and FIR views and the order of trials with or without warnings. After filling out consent and biographical forms, and before discussing anything about night vision systems, subjects viewed two short clips of night driving: a clip of system “A” and a clip of system “B”. They then rated the effectiveness of each system on a seven-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 finger-mounted button “as soon as you see a pedestrian in the in-vehicle display.” A confirmation tone indicated that their key press had been registered. After about 10 minutes of practice, the test trials began. Each subject viewed 16 clips of four to six minutes each, with breaks after every four clips (about 20 minutes). Before each break, subjects completed a NASA TLX workload evaluation form for the four clips they had just viewed.

The order of clip presentation followed an ABBA pattern to balance practice effects. The first four clips were of system A, followed by eight clips of system B, and ending with another four clips of system A. For each system, one block of four clips

included two short and two long pedestrian warnings. The order of short and long warnings within the block was balanced across subjects, as was the order of blocks with or without warnings.

Each block of four clips included all four road routes (Table 1) in an order that was balanced across 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 again rated system effectiveness.

Experimental Design and Data Analysis

Each subject viewed 80 pedestrians on 16 video clips. Of those pedestrians, 40 were viewed with each system type (NIR and FIR). Of the 40 pedestrians viewed in each system, 20 were without a warning, 10 were with a short warning, and 10 were with a long warning.

The experimental design examined the effect of two night vision systems (NIR and FIR) and three levels of warnings (150 m warning, 75 m warning, or no warning) nested within four clips of five pedestrians nested within subject, and with age (younger and older) as a between-subject variable. To avoid showing a subject the same clip more than twice with the same night vision system, only two of three levels of the warning conditions were shown for each road. This fractional factorial approach allowed a within-subject and within-night-vision-system comparison of long warning to no warning and of short warning to no warning. Comparison between long and short warnings was between subjects. Table 2 shows an example of 16 movie clips seen by one subject. The choice of roads for each warning was balanced across subjects.

Table 2
Example of roads shown with each system and warning type on 16 movie clips for one subject. Each movie clip contained 5 pedestrians.

Road	System and Warning					
	FIR			NIR		
	None	Short	Long	None	Short	Long
A	✓		✓	✓		✓
B	✓		✓	✓		✓
C	✓	✓		✓	✓	
D	✓	✓		✓	✓	

A repeated measures analysis was performed for the following dependent variables:

- (1) Detection distance – defined as the straight-line distance between the vehicle and the pedestrian when detection was reported.
- (2) Detection accuracy – defined as the percentage of pedestrians detected at or before passing them (hits) and the number of times per trial that pedestrians were reported when none were present (false alarms).
- (3) Glance frequency and duration – calculated as the time interval between glances and the duration of the last two glances before a pedestrian was detected. These measures were based on a subset of the data (eight trials per subject).
- (4) Subjective evaluation of the appearance of the system, workload experienced by the subjects, and system effectiveness.

RESULTS

Detection Distance

A repeated measures ANOVA of detection distances revealed a significant main effect of system type, $F(1,14) = 80.1, p < .001$ (Figure 5). Subjects detected pedestrians with FIR at three times greater distances than with NIR. Although detection distances of younger subjects tended to be greater than those of older subjects, the main effect of subject age group, $F(1,14) = 0.34$, and the interaction between system and age group, $F(1,14) = 1.8$, were not statistically significant.

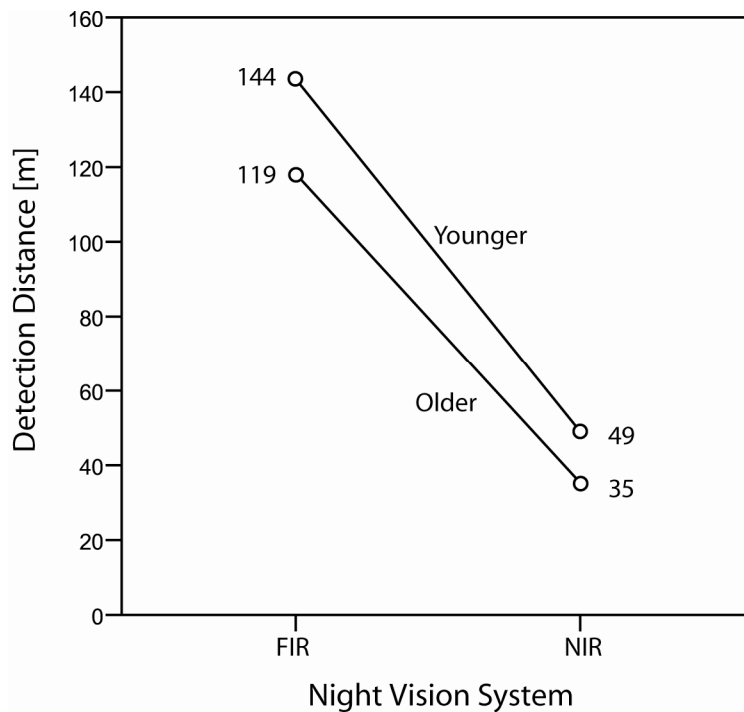


Figure 5. Detection distances with FIR were about three times longer than with NIR but there was not a statistically significant age difference.

When a long pedestrian warning (150 m) appeared, detection distance improved significantly versus when there was no warning $F(1,14) = 63.2, p < .001$. This effect was more pronounced for the NIR system than for the FIR system $F(1,14) = 22.4, p < .001$ (Figure 6, left pane). The mean detection distance with long warnings was shorter with NIR than with FIR, partly because the pedestrian was usually visible in the FIR scene before the warning appeared. With NIR, however, the pedestrian was usually not clearly visible in the scene until shorter distances. Some subjects may have been reluctant to indicate they saw a pedestrian based on the warning alone.

When a short pedestrian warning (75 m) appeared, detection distance improved for NIR but degraded for FIR. The interaction between system type and warning was significant, $F(1,14) = 32.9, p < .001$ (Figure 6, right pane). Because detection without warning with NIR was 42 m, a warning at 75 m helped increase detection distance. On the other hand, because detection without warning with FIR was 131 m, a short warning at 75 m did not help. In fact, it shortened detection distance, probably because some subjects waited to see the warning before they responded or just monitored the displayed less. Subjects knew that if there were a warning it would appear at a short distance before the pedestrian because trials were blocked by system and warning type. It is in light of this knowledge that they waited for the warning to appear before they responded.

Figure 7 shows the distribution of detection distances for each of the system/warning combinations tested. It helps explain the effects of system and warning type on detection distance. First, the variability between trials is illustrated. For example, detection distance with FIR and a long warning was on average 150 m. Examination of the bottom left pane in Figure 7 shows that 12.5% of responses occurred immediately after the warning appeared at 150 m, but there were also many responses before and after the warning appeared. Second, it is seen that the difference between FIR and NIR without warnings was not only in the mean value but also in the shape of the distribution. FIR without a warning was skewed to the right and extended to long distances (some in the 300-400 m range). In contrast, NIR without a warning was more concentrated around the mean with an upper bound of 175 m, as expected from visibility limitations of the system. Third, the effect of short warning is seen by comparison to the corresponding no warning condition with the same system. In the case of NIR, short warnings increased detection distance, whereas long warnings reduced the portion of responses both below 75 m and above 75 m, with a combined effect of reducing the mean detection distance from the case of no warnings. Finally, long warnings increased detection distance for both systems.

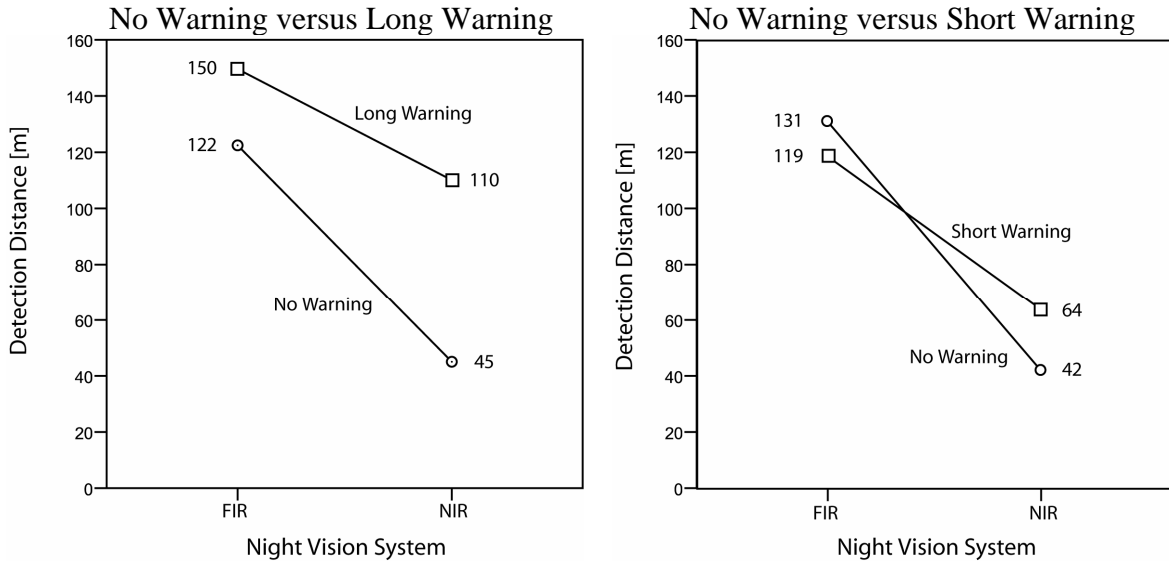


Figure 6. Detection distance by system and warning type for long warnings (left pane) and short warnings (right pane). Note that because of the fractional factorial design the roads used for the within-subject comparison were different for each of the warnings. Thus, the no-warning condition differs between panes.

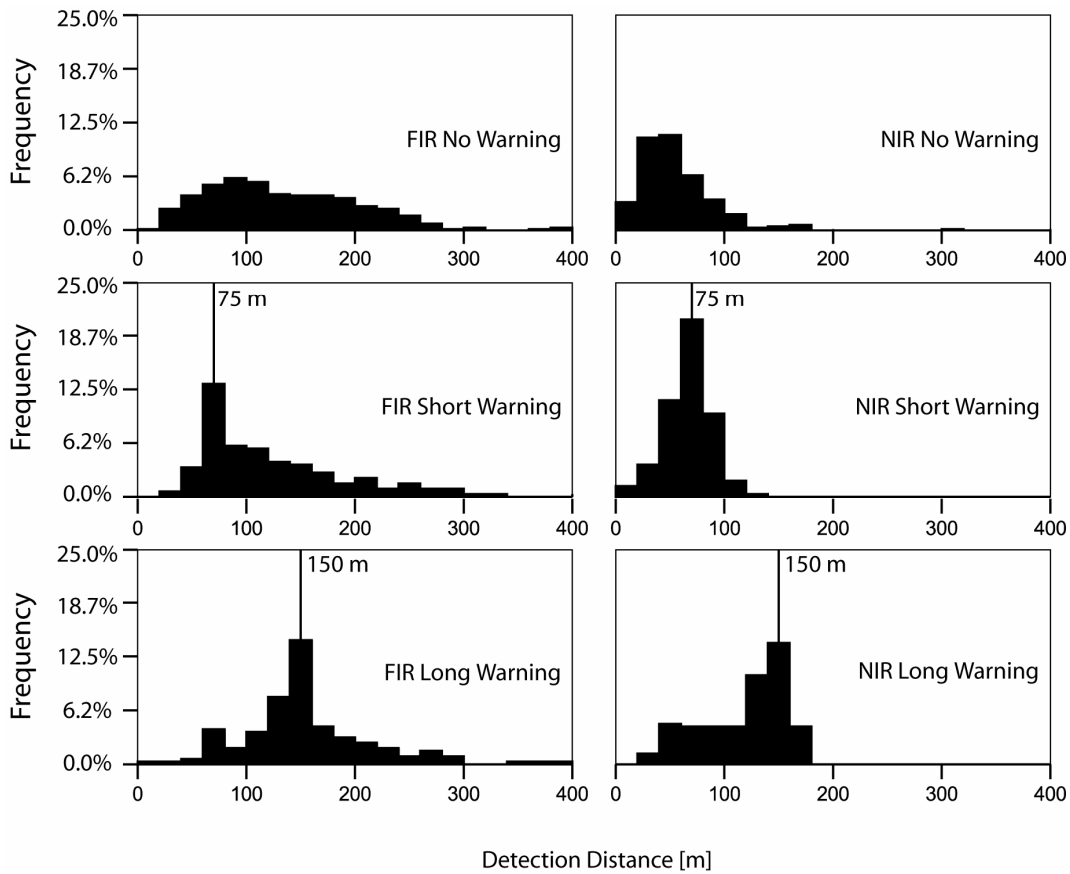


Figure 7. Distributions of detection distance by system type and warning type. Warning distances are indicated within the bottom four panels.

To evaluate the effect of the steering task on detection distances, results from this experiment were compared to results from the previous experiment (Tsimhoni, Bärghman, Minoda, and Flannagan, 2004), in which subjects looked continuously (without performing a tracking task) at an in-vehicle display identical to the one in the present experiment. A subset of results from the previous experiment (results from the four out of six roads that were used in the present experiment) was compared to results from the present experiment with the effect of driving as a between-subject (and between experiment) variable. Overall, detection distance decreased by 23 m for FIR and 16 m for NIR (Figure 8). The effect of driving was significant, $F(1,28) = 4.6, p < .05$ but the interaction between driving and system was not significant, $F(1,28) = 0.3$, suggesting that the loading task had a similar effect on performance in both systems. Alternatively, as a proportion of detection distance in the no driving condition, the reduction in detection distance with the NIR system was substantially greater; the reduction from 58 to 42 m is a reduction of 28% while the reduction from 154 to 131 m is only 15%. In either case, the decreases in mean detection distance correspond to about a one-second delay (at 70 km/h), which may be explained by the need to sample the IR display intermittently in the present experiment but not in the previous experiment.

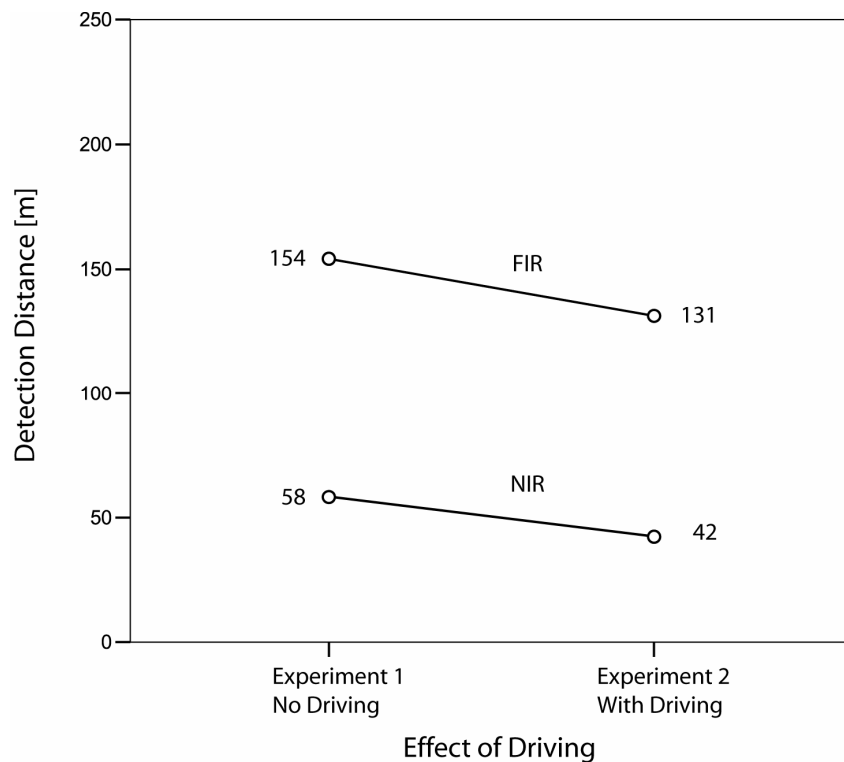


Figure 8. In the present experiment, detection distance with both systems decreased from the previous experiment, suggesting that the cost of having to look away from the display translated to about a one-second delay.

Detection Accuracy

Overall, pedestrians were missed in 99 out of 1,280 trials (7.7%). A majority of the misses (70 of 99) were made with NIR without a warning (21.9% of the 320 trials in that condition). The percentage of misses with a pedestrian warning (FIR: 3.1% and NIR: 3.4%) or with an FIR system without a warning (2.5%) were considerably lower than NIR without a warning (Table 3).

Table 3
Number of missed pedestrians as a function of system and warning type.

Warning Type	Number of Misses (percent of cases)	
	FIR	NIR
No Warning	8 (2.5%)	70 (21.9%)
Short Warning	8 (5%)	4 (2.5%)
Long Warning	2 (1.2%)	7 (4.4%)

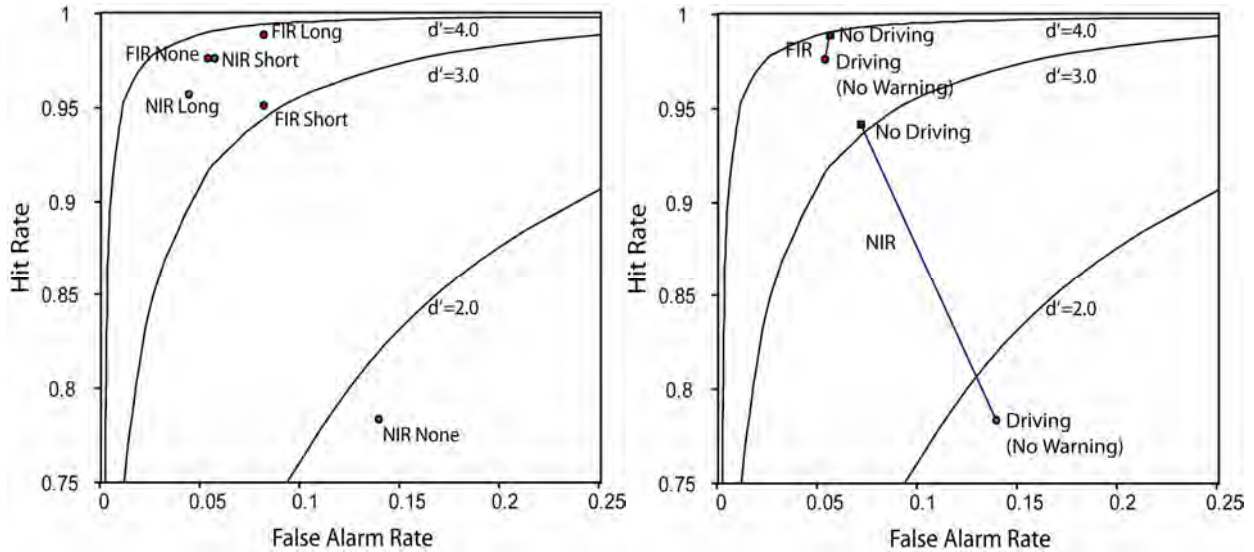
Subjects indicated that a pedestrian was present when there was no pedestrian potentially visible in the scene (a false alarm) 103 times during the experiment. Whether a pedestrian was potentially visible was determined by the maximum distance at which any indication could be discerned in a careful frame-by-frame inspection of the FIR video. This was normally the first frame in which there was an unobstructed line of sight to the pedestrian, for example when the pedestrian appeared around a curve in the road. Most false alarms (44) were made with NIR and no warning. Table 4 shows the number of false alarms as a function of system type and warning. Percent values indicate the percent of trials in which false alarms occurred. Trials were defined as the five intervals in each road video clip during which a pedestrian was not discernible in the FIR video.

Table 4
Number of false alarms as a function of system and warning type.

Warning Type	Number of False Alarms (as percent of trials)	
	FIR	NIR
No Warning	17 (5.3%)	44 (13.8%)
Short Warning	13 (8.1%)	9 (5.6%)
Long Warning	13 (8.1%)	7 (4.4%)

Figure 9 shows the effect of system type on hit rates and false alarm rates and the resulting measure of sensitivity d' as defined in signal detection theory (see Macmillan and Creelman, 2005). Hit rate represents the portion of trials in which the pedestrian was detected (at any distance before the pedestrian). False alarm rate is the number of false detections of a pedestrian divided by the number of trials. The lowest system sensitivity (the lowest d') was with NIR without a warning. Of the other conditions, FIR with a long warning had the highest sensitivity, but was not much better than FIR without a warning.

A comparison with results from the previous experiment (Figure 10) indicates the effect of the driving task on detection. Sensitivity for NIR with no warning decreased considerably with the addition of the driving task. Sensitivity for FIR with no warning decreased too, but the magnitude of the effect was small.



System	Warning	d'
NIR	None	1.87
	Short	3.55
	Long	3.42
FIR	None	3.58
	Short	3.04
	long	3.64

System	Experiment	d'
NIR	1	3.04
	2	1.87
FIR	1	3.90
	2	3.58

Figure 9. Hit rates and false alarms by system and warning type.

Figure 10. The effect of driving on detection accuracy. Accuracy with NIR decreased significantly as a result of the added driving task.

Glance Behavior

A sample of trials was analyzed for measures of eye glances using video clips of the subjects' eyes. Eight trials per subject were analyzed for each subject (2 trials \times 2 systems \times 2 warning levels). Two measures of eye glances were collected: (1) glance duration—defined as the duration of a glance from the moment the eyes leave the forward scene to the moment they leave the in-vehicle display, and (2) glance frequency—defined as the time interval between consecutive glances.

Figure 11 shows glance duration as a function of system type and warning level. There was a trend for glances to be longest when no warning was present and shortest with short warnings, but the differences among the conditions were not statistically significant.

On average, subjects glanced at the display every 3.0 s when there was no warning and every 3.3 s when there was a warning, but the differences among conditions were not statistically significant.

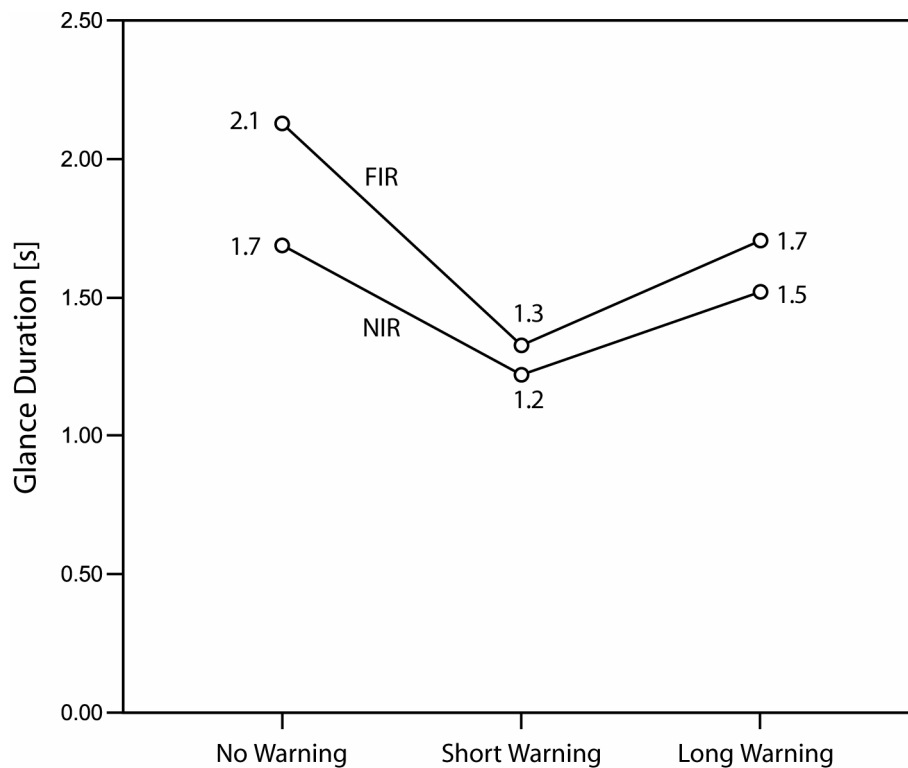


Figure 11. Glance duration by system type and warning level.

Subjective Rating of System Effectiveness

Subjects were asked to rate the effectiveness of the two experimental night vision systems after viewing a short introductory clip of each system and before any other exposure to the systems. 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?”).

Figure 12 shows the effectiveness ratings before and after the experimental trials. FIR was rated as more effective than NIR, $F(1,14) = 26.1, p < .001$. There was a significant decrease in the perceived effectiveness of both systems after the experimental trials $F(1,14) = 8.5, p < .01$. The main effect of age and higher order interactions with age were not statistically significant.

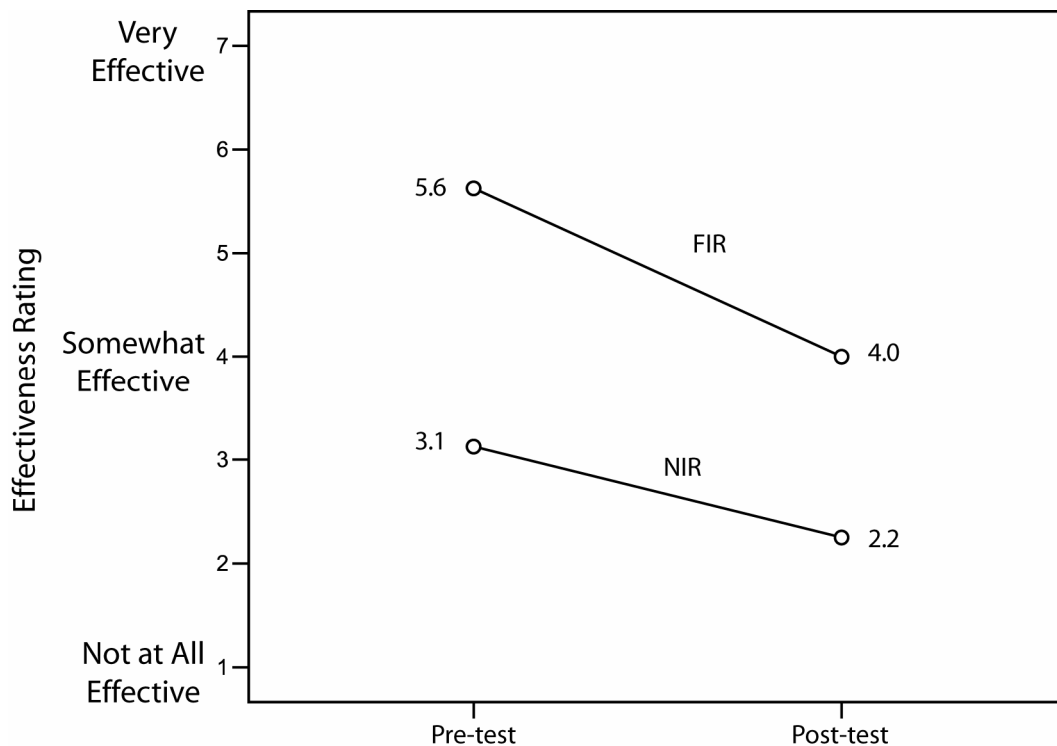


Figure 12. Rating of effectiveness of two night vision systems before and after the experiment.

Subjective Rating of Workload

Subjective ratings of workload were collected using NASA TLX (Hart and Staveland, 1988). NASA TLX is a multi-dimensional subjective workload rating technique that is commonly used to evaluate workload. Subjects rate their workload on six scales representing mental demand, physical demand, temporal demand, performance, effort, and frustration level. Their ratings are then combined, using weights or by simple addition, into a TLX score that represents their overall workload. The results reported here are based on unweighted addition of the ratings on each scale (Moroney et al., 1992). Figure 13 shows the effect of system and warning on TLX workload scores. Workload was rated higher with NIR than with FIR, $F(1,14) = 28.0, p < .001$. The effect of warning was nearly significant $F(1,14) = 2.9, p = .11$, and the interaction between system and warning was nearly significant, $F(1,14) = 2.3, p = .15$. Workload was reduced by the addition of warnings to NIR, but remained almost unchanged with the addition of warnings to FIR.

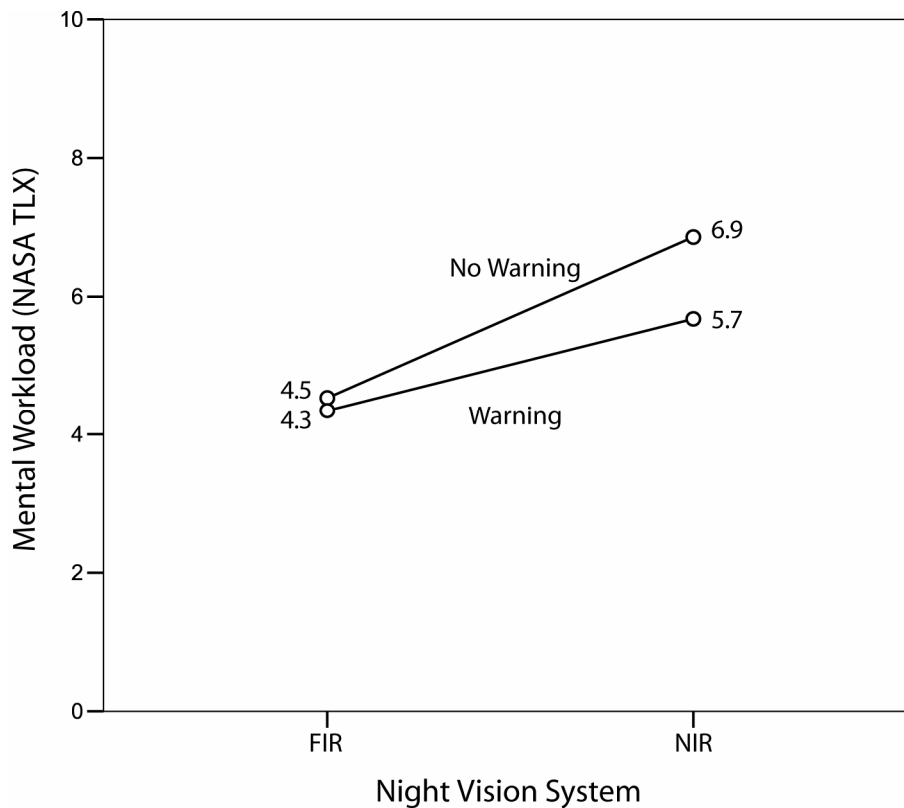


Figure 13. Workload rating (NASA TLX) was higher with NIR than with FIR. The addition of automatic warnings slightly decreased workload using NIR but not using FIR.

SUMMARY AND CONCLUSIONS

Overview

Underlying this experiment is the premise that the primary purpose of night vision systems should be to enhance the visibility of pedestrians. Support for this premise comes from evidence that the major safety problem caused by darkness (independent of other factors related to the time of day such as fatigue and alcohol use) is the increased risk of pedestrian collisions (Sullivan and Flannagan, 2001). Furthermore, expert analysis suggests that the minimum visibility distance that gives a driver on a rural road an opportunity to detect and identify an object and to choose an appropriate action is about 150 m (Rumar, 2002). We showed in a previous experiment (Tsimhoni, Bärghman, Minoda, and Flannagan, 2004) that FIR night vision systems may provide detection distances on the order of 150 m but that NIR systems provide detection distances that are much shorter. A possible solution to that problem is to enhance night vision systems with automatic pedestrian warnings that may be generated with image processing algorithms.

Implementations of automatic pedestrian warnings may range in complexity from a simple notification for a pedestrian ahead (using a warning lamp or auditory tone) up to a localized icon overlaid on the display of a night vision system inside the vehicle, or even overlaid on the road via a conformal head up display or head mounted display.

Regardless of the implementation details, automatic pedestrian warnings in night vision systems have the potential to increase pedestrian detection distance and accuracy. The extent to which warning systems might be effective depends on, among other things, the ability of the driver to detect pedestrians in a timely manner, and the driver's inclination to respond to warning systems. Warnings that use multiple modalities with strong signals are likely to result in short response times, but because of their intrusive nature they are also prone to being considered an annoyance by some drivers unless they are always correct. In the present experiment, we chose visual warnings that were not extremely distracting. We were interested in a warning that would be strong enough to be detected most of the time but that would not overwhelm the driver. If we had used strong multimodal warnings in our experiment, we would probably have seen immediate responses to all warnings.

The Effect of Automatic Pedestrian Warning

As expected, warnings at 150 m increased the mean detection distance of pedestrians. Detection distance with an NIR night vision system increased from 45 m when no warning was present to 110 m, and the hit rate increased from 78% to 96%. Results with an FIR system followed the same trends but were less pronounced. Detection distance increased from 122 m when no warning was present to 150 m, and the hit rate increased from 98% to 99%.

Warnings at 75 m were effective for NIR. They increased the mean detection distance from 42 m to 64 m, and the hit rate increased from 78% to 98%. Results with the FIR system, however, did not follow the same pattern. Detection distance decreased from 131 m to 119 m, and the hit rate decreased from 98% to 95%. Because pedestrians were visible on the FIR night vision display before the automatic pedestrian warning appeared, some subjects may have chosen not to wait for the warning and to respond as soon as they saw the pedestrian. Others, on the other hand, may have waited for the automatic warning before they responded or they may have relied on the warning function and decreased their monitoring of the screen images themselves.

The Effect of Simulated Steering on Pedestrian Detection

A comparison of the present results with those from our previous experiment provided insights on the effect of steering on pedestrian detection. Detection distance decreased with the addition of the loading task by 23 m for FIR and by 16 m for NIR. The degradation in performance corresponds to about a one-second delay (at 70 km/hr), which may be explained by the need to sample the display intermittently. We hypothesize that the delay would increase further if drivers sampled the display less frequently. The effect of the loading task on hit rate differed considerably between night vision systems. Whereas the hit rate for FIR decreased from 99% in the previous experiment to 98% in the present one, the hit rate for NIR decreased from 94% in the previous experiment to 78% in the present one. Similarly, whereas false alarms for FIR remained approximately unchanged at 6%, false alarms for NIR doubled from 7% to 14%.

The degradation in detection accuracy using the NIR system, but not the FIR system, may be related to problems in extracting information from the NIR display but not in the FIR display. It may be easier to miss a pedestrian with the NIR system than with the FIR system when sampling them intermittently because pedestrians in the NIR system appear less distinctly and there is more clutter around them (Tsimhoni and

Flanagan, 2005). Because detection distance was already low, a decrease of 16 m in mean detection distance for NIR might have been sufficient to shift near detections into misses.

The Effect of Age

Detection distances by older subjects tended to be about 20-30% shorter than those by younger subjects, but the difference was not statistically significant. Age differences in detection accuracy were more pronounced. The missed rate with NIR without a warning was 8% for older subjects and 2% for younger subjects. The effect of age on hit rates of other conditions was small. The false alarm rate with NIR without a warning was 5.2% for older drivers and 1.7% for younger subjects. Degradation in hit rate as well as false alarm rate suggests an overall degradation in older subjects' ability to detect pedestrians rather than bias to saying there is or there is not a pedestrian. When warnings were shown, younger subjects made few false alarms (0.1%), but older subjects made considerably more false alarms (1.7%). The magnitude of these age differences is consistent with other literature on the effect of age on performance with telematics (see Green, 2001).

Conclusions

In our previous study, in which subjects viewed the night vision displays continuously, we concluded that to the extent that the two systems used in the experiment reasonably represented their respective technologies, the results support the expected enhancement of pedestrian detection in FIR systems relative to NIR systems. Results from the present experiment reinforce that conclusion. Detection distances with NIR with no automatic warning were substantially inferior to FIR under similar conditions. Furthermore, detection accuracy with NIR without automatic warning degraded so much (by the need to do the steering task) that subjects missed 22% of pedestrians. However, NIR systems may be enhanced to improve pedestrian detection. Such improvement was found in our simulated automatic visual warning.

We conclude that automatic pedestrian warning, in the form of highlighting pedestrians on the night vision display, is generally helpful in increasing detection distances and accuracy. However, warnings might not be effective if they do not appear far enough in advance of the pedestrian. As demonstrated in this experiment, some subjects detected the pedestrians with the short (75 m) warning later than they did

without the warning. If warning systems cannot exceed human performance, there may be instances in which they will hurt rather than help drivers to detect pedestrians.

One of the possible risks of introducing night vision systems in vehicles is that they will increase the workload imposed on drivers. Automatic warnings have the potential of reducing workload by reducing the need to constantly sample the display. In our experiment, however, automatic warnings did not significantly reduce perceived workload, as subjects continued to sample the display frequently (about every 3.0 s). It is likely that a stronger warning with the same reliability (of 100% or close) would have reduced the need to sample the display so frequently, thus reducing workload.

REFERENCES

- Bertozzi, M., Broggi A., Fascioli, A., Graf, T., & Meinecke, M.M (2004). Pedestrian detection for driver assistance using multiresolution infrared vision. *IEEE Transactions on Vehicular Technology* 53(6), 1666-1678.
- Fang, Y., Yamada, K., Ninomiya, Y., Horn, B., & Masaki, I. (2003). Comparison between infrared-image-based and visible-image-based approaches for pedestrian detection. *Proceedings of IEEE Intelligent Vehicle Symposium*, 505-510.
- Green, P. (2001). *Variations in Task performance between younger and older drivers: UMTRI research on telematics*. Paper presented at the Association for the Advancement of Automotive Medicine Conference on Aging and Driving, Southfield, MI.
- Hart, S.G. & Staveland, L.E. (1988). Development of the NASA TLX, Task Load Index: Results of empiric and theoretical research In. P.A. Hancock & N. Meshkati (Eds.) *Human mental workload*. New York: North Holland.
- Macmillan, N. A., & Creelman, C.D. (2005). *Detection theory: A user's guide* (2nd ed.). Mahwah, New Jersey: Lawrence Erlbaum Associates.
- Moroney, W.F., Biers, D.W., Eggemeier, F.T., and Mitchell, J.A., (1992). A comparison of two scoring procedures with the NASA Task Load Index in a simulated flight task. *Proceedings of IEEE National Aerospace and Electronics Conference*, 734-740.
- Nanda, H. & Davis, L. (2002). Probabilistic template based pedestrian detection in infrared videos. *Proceedings of IEEE Intelligent Vehicle Symposium*, 15-20.
- Parasuraman, R., Hancock, P.A., and Olofinboba, O. (1997). Alarm effectiveness in driver-centred collision-warning systems. *Ergonomics*, 40(3), 390-399.
- 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.
- Shashua, A., Gdalyahu, Y., & Hayun G. (2004). Pedestrian detection for driving assistance systems: single-frame classification and system level performance. *Proceedings of IEEE Intelligent Vehicle Symposium*, 1-6.

- Sullivan, J.M., & 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., Bärghman, 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., Bärghman, J., Minoda, T., & Flannagan, M.J. (2004). *Pedestrian Detection with Near and Far Infrared Night Vision Enhancement* (Technical Report UMTRI-2004-38). Ann Arbor, MI: The University of Michigan Transportation Research Institute.
- 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.
- Tsimhoni, O., & 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.
- Tsimhoni, O., & Flannagan, M.J. (2005). Visual clutter in active night vision systems reduces detection distance. *Proceedings of the International Symposium on Automobile Lighting*, Darmstadt, Germany.
- Xu, F., Liu, X., & Fujimura, K. (2005). Pedestrian detection and tracking with night vision. *IEEE Transactions on Intelligent Transportation Systems*, 6(1), 63-71.
- Zhao, L. & Thorpe C. E. (2000). Stereo- and neural network-based pedestrian detection. *IEEE Transactions on Intelligent Transportation Systems*, 1(3), 148-154.