Evaluation of night vision enhancement systems: Driver needs and acceptance

Diana Rösler[1], Josef F. Krems[1], Sascha Mahlke[2], Manfred Thüring[2], and Katharina Seifert[3]

Abstract: Night vision enhancement systems (NVESs) have the potential to assist the driver by presenting additional information about the road as well as not yet visible critical obstacles ahead of the vehicle [e.g., 1]. Although a broad discussion exists about the capabilities as well as the potential risks for traffic safety, little is known about the usability of NVESs, and in particular, the consequences of design decisions on driver behaviour, cognitive load, and user acceptance.

To learn more about these issues, different NVESs were assessed under real traffic conditions. Systems differed in type and position of the displays used to present either analogue far or near infrared sensor information. In addition, automatic pedestrian recognition software was applied. To summarize, results indicate that NVESs have a promising and worthwhile potential to increase traffic safety at night, but further substantial advancement in information presentation and selection is essential to reduce driver workload and support information processing.

Keywords: vehicle design, driver assistance, usability, heuristic evaluation

1 Introduction

Driving depends essentially on visual information processing. Therefore, it is of no surprise that under restricted visibility conditions such as at night, under heavy rain, or in fog, many severe and fatal crashes occur [2]. Because of technical and legal regulations the problem cannot be solved simply by increasing the light output of low beams. Therefore, infrared-based night vision enhancement systems (NVESs) were developed which try to overcome the limits of humans` sensory and cognitive abilities, as well as limits in reaction time. Infrared sensitive camera based night vision systems can enhance the visibility of objects emitting or reflecting heat waves, making visibility comparable to high beam conditions. In our evaluation, we compared different night vision systems to learn about the effects of these different systems' properties on nighttime driving, and to gather suggestions for the further development of such systems. The objective of the study was to answer the following questions: Do the systems improve object recognition by the driver? Do the systems differ in their ability to support the driver in detecting objects? Do the systems increase driver workload? Would drivers like to use such systems?

[1] Chemnitz University of Technology, Chemnitz, Germany (e-mail for correspondence: diana.roesler@phil.tu-chemnitz.de)
[2] Berlin University of Technology, Berlin, Germany
[3] Volkswagen AG Group Research, Wolfsburg, Germany

Our study was divided into a heuristic evaluation with eight experts from ergonomics and psychology, and an experimental study with 15 male and female drivers. In this paper we will present results from the heuristic evaluation; further results from the experimental study with ordinary drivers can be found in [3].

2 Method

Experts assessed six NVESs in a free evaluation according to Nielsen's proposals about the heuristic evaluation process [4; 5]. The systems mainly differed in the sensor technology (near versus far infrared), the type of image processing (analogous display versus object recognition), and the type of display (head-down versus head-up). We chose two methods for data collection: The first method comprised free verbal reports [6; 7] during system use under real traffic conditions at night. For the second part of the study, a semi-structured interview with heuristic guidelines was used.

2.1 Participants

The eight male experts were on average 54 years old. Each of them is an acknowledged specialist and has been working on psychological and/or human factors issues for more than ten years. Two experts focus their research on perceptual problems, five concentrate on the evaluation and design of telematics systems, and one has his main focus on traffic safety. None of the experts had been directly engaged in night vision research prior to our study. All experts were remunerated.

2.2 Apparatus

Two instrumented cars equipped with a far and a near infrared sensor were used. The infrared sensors' output was presented on different types of displays (see examples in *Figure 1*). These were: (1) a head-down display integrated in the instrument panel behind the steering wheel in front of the driver, (2) a virtual head-up display, (3) a real head-up display where the image was projected on a narrow mirror at the bottom of the windshield, and (4) an abstract LED display that, in combination with an active pedestrian detection system, signalized the presence of an object. A flashing LED indicated the event "pedestrian detected" and the direction where the object was localized.

Figure 1: Display technologies (PR LED and PR MHUD)

The following combination of sensor, image processing, and display types were evaluated:

- Near infrared sensor with head-down display (NIS HDD)
- Near infrared sensor with mirror based head-up display (NIS MHUD)
- Far infrared sensor with head-down display (FIS HDD)
- Far infrared sensor with virtual head-up display (FIS VHUD)
- Far infrared sensor with automatic pedestrian recognition and an analogue mirror based head-up display (PR MHUD)
- Far infrared sensor with automatic pedestrian recognition and an event based LED display (PR LED)

2.3 Test route

The route had a total length of 16.3 km and was chosen in the surrounding area of Chemnitz, Saxony, Germany. It consisted of public street sections that differed in their demands. The route comprised straight and well constructed one way and two way highways, as well as winding and mountainous roads. In order to elicit enough critical events in comparable driving situations, five persons took a position along the roadside. Their task was to move, to stand still, or to crouch while the test car passed. A test trial with one NVES lasted about 25 min.

2.4 Procedure and data collection

The study consisted of two parts, a night session where the experts tested the system under real traffic conditions and a structured interview the following day.

The first part of the study started with a short introduction about the goals of the study, the overall procedure, and the experts' task. Verbal reports (think aloud protocols) were recorded during the approximately 25 min active system use under real traffic conditions at night. During the drive, the experts were encouraged to articulate their impressions, thoughts, and opinions of the design and system usability. The task for the expert was to detect the persons along the route. After each test trial with one of the systems the expert had to give his first impression and to fill in workload questionnaires. Workload was assessed with (a) the SEA-Scale measuring general demand [8, one item], and (b) a modified version of the NASA Task Load Index [NASA-TLX, five items, 9]. With this information different dimensions of demand were assessed (e.g., visual demand, mental demand, temporal demand, and frustration).

After a short break the next trial started with the next system. The procedure was the same for all systems. The sequence, in which systems were presented at night under real traffic conditions, was randomized for each expert. The entire evaluation of the six systems lasted between four and five hours and started with nightfall.

The second part of the study took place on the following morning. A semi-structured interview (49 questions) was carried out with the experts. First, a short 15 min video of the six NVES was presented, containing posed sequences of persons who walked, stood still, lay down on the ground, or crouched. Afterwards, we interviewed for approximately two hours. Questions were asked about the different display locations and system technologies, acceptance, usability, attraction, learnability, potential impacts on and risks for traffic safety, as well as desirable future developments.

3 Results

First, all experts' statements from the video recordings and the interviews were transcribed. Next, the data was coded [10]. It was sorted and duplications were excluded, so that each expert made only one statement about a specific aspect of the evaluation. Subsequently, the information was categorized and condensed according to different main patterns that were extracted during content analysis [11]. The information was clustered into seven main categories that partly depended on each other: (1) potential of the night vision enhancement approach, (2) display characteristics like position, contrast, or picture quality, (3) camera arrangement and matching of display information and real world perception, (4) alert announcement of hazards, (5) workload, (6) learnability, and (7) expected user acceptance and satisfaction. Due to the nature of a free heuristic evaluation, not all experts commented on all evaluation aspects. Some aspects were mentioned only by two or three experts. Relative frequencies are reported for better comparison, but due to the small number of experts, these frequencies should only be used as a rough orientation.

Potential. All the experts regarded the development of NVESs as very promising and worthwhile. A further development was supported. The experts suggested that in doing so, the main focus should be on software-supported detection of objects and especially of pedestrians. Systems that assist only in driving and prevision in general were considered as less useful [see also 1]. None of the systems was perceived as mature enough for bringing to the market at the time of evaluation. At least 83 % of the experts preferred the system PR LED (for abbreviations see *Apparatus*). This system was perceived by 63% of the experts as being sufficiently capable of capturing the drivers' attention and, therefore, of having the potential to warn against obstacles. At the same time, the experts did not expect negative effects on vehicle control and ultimately traffic safety (67% of experts).

Display characteristics. The positions of PR LED, MHUD, and VHUD obtained positive assessments (mentioned by 83%, 63%, and 50% of experts). These displays were characterized as facilitating the recognition of information with peripheral vision. A majority of the experts suggested not using the HDD position to present night vision information (63% of experts), since in this location the information cannot be perceived peripherally (83% of experts). Features like contrast and the quality of the system image were assessed most positively for the NIS HDD (63% of experts), compared to the far infrared sensor systems, which received negative assessments for lack of contrast (for FIS HUD 75% of experts, PR MHUD 67%, and FIS HDD 57% of experts). These positive considerations of NIS HDD were extenuated by the effect of oncoming headlights or retro-reflective road signs that created glare (blooming) on the display (88% of experts). Thus, relevant information like pedestrians or animals could become cross-faded and would not be detectable (50% of experts).

Camera arrangement. All systems presented the environmental information in a perspective different from the drivers' point of view. Especially for the far infrared sensor systems, the position of the camera in the radiator cowling seemed to be too low (63% of experts). With both far and near infrared sensor systems, there is no chance to use the displays in curves because the perspectives of the cameras and the driver differed highly (50% of experts with respect to far infrared sensor systems and 75% of experts regarding near infrared sensor systems). The experts suggested mounting the night vision cameras directly on the axis of the drivers' field of view. For instance, the position of the sun visor was suggested as a possible solution. Another solution suggested by the experts to overcome that problem was the enlargement of the depicted environmental section (50% of experts), mainly in the horizontal dimension. That way the close-up range of the car could also be displayed.

Matching. A further aspect of night vision system usage is the ability of the drivers to relate the system image to the reality outside of the vehicle. The NIS HDD provided the most useful cues and caused the least problems (75% of experts). All other systems (NIS MHUD, far infrared sensor as well as PR LED, frequencies between 50 and 83% of experts) had deficiencies in allocation. Representations of relevant information, like road signs and markings, were missed in the far infrared sensor images and thus made it very difficult for the drivers to locate, for example, the position of a pedestrian on the real road. In some nights, mainly rainy ones, differences in temperature were missed so that the interpretation processes were obstructed.

Alert announcement of hazards. The intensity of critical stimuli (pedestrians) as well as the attention capturing effect was most preferred in the PR LED system (67% of experts). The PR LED provided simple and obvious signals of potential obstacles that could be perceived peripherally. This capability was also considered for the far infrared sensor systems, especially for supporting detection of hazards (between 63 and 100% of experts). Attention is captured by the far infrared sensor systems like the PR LED because warm objects were presented much more brightly compared to the surrounding area. However, the contrast needs to be intensified (75% of experts). Problems were seen in the near infrared sensor presentations. More or less all objects were presented quite similarly. Distraction can occur and there is the possibility that a pedestrian is mistaken for a reflective post on the roadside (63% of experts). Hence, the experts suggested that only relevant information should be presented in the foreground, making this information detectable with peripheral vision. The intensity of irrelevant information should be perceptibly reduced. For example, clouds were presented very clearly with the far infrared sensor, possibly distracting the driver.

Workload. Items of the workload questionnaires had to be rated by the experts for normal users' purposes. The general demand assessed with the SEA-Scale was lowest for PR LED; it was evaluated as "hardly demanding." The PR MHUD system was seen as "a little demanding". All other systems were judged to be in "some degree demanding." According to the experts, none of the systems was demanding in such a way that it would be critical for traffic safety. Results of the NASA-TLX showed no differences in general system appraisal. In comparison with all NVESs, PR LED was again rated as having the lowest demand value.

Learnability. In the experts' opinion the learnability of usage varies dramatically for the different NVESs. The PR LED system was rated as having the fewest learning costs because of its event correlation and its frugal and intuitive PR LED display. All other systems, including PR MHUD, were rated as requiring much more learning. The images, especially of the far infrared sensor systems, needed to be interpreted and matched with reality. In addition, the experts voiced that system usage and glance behavior would have to be learned and would need to be integrated into normal driving processes.

Acceptance. Finally, experts appraised the regular application of the systems. This criterion acted in certain ways like a general acceptance assessment and could therefore be seen as a summary statement. The experts could only imagine the PR LED presentation for a system used regularly for several hours a night. For all other systems, the high visual demand and therefore enhanced visual glance times for display scanning were mentioned as fatiguing and adverse to traffic safety. Experts emphasized again that only necessary and relevant information that would help to detect hazards as well as to allocate them in the environment, should be presented. There would not be any reason to use the system when the system does not actively inform the driver when to look at it. A system that only increases general visibility was not considered as effective.

4 Summary and discussion

All experts considered night vision enhancement as a very promising and worthwhile approach to assist drivers and to increase safety under impaired visual conditions. Consequently, all experts endorsed further development. The experts recommended that the main focus should be on an automatic, software supported detection of critical events. Such a support has the advantage that drivers can hold their eyes on the road at all times. Peripherally presented stimuli of obstacles provides cues about potential hazards and increases the alertness of the driver. In conclusion, such an approach was assessed as the most effective one. Analogue displays were considered less effective since they require additional visual attention and may lead to symptoms of fatigue. Therefore, the use of analogue displays was not recommended as the solution for future applications.

Against this background, the PR LED system was preferred because of its high efficiency. Its simple and intuitive configuration holds no risk of distraction as well as no additional visual demand, and, therefore, no constraints on traffic safety. Conspicuous signals refer to the critical obstacles in a fast and directional way. The position of the display was perceived as well chosen and as being able to be perceived peripherally.
Although many features of the NIS HDD image were assessed positively, the warning function against potential risks was rated as very low. Different objects in the scene were presented very similarly (e.g., reflective post and pedestrian), so that attention capturing of critical targets was hard to achieve. In addition, critical events were overlooked because of blooming effects. Far infrared sensor systems hold the potential to show additional information that the driver cannot see under normal light conditions. Like the PR LED, a far infrared sensor is useful as a warning signal because of the accentuation of warm objects that appear very bright in a colder surrounding. Extensive problems are anticipated with respect to learnability because of temperature-dependent variances in the resulting image, for example, with cooling of the surroundings during longer nighttime driving. For both far and near infrared sensor systems based on video display, the problem was that no active system component tells the user when to look at the display. The issues mentioned for both system variants limit their efficiency to a considerable degree.

5 References

[1] Rumar, K. (2002). *Night vision enhancement systems: What should they do and what more do we need to know?* (UMTRI - 2002 - 12). Ann Arbor, Michigan: University of Michigan, Transportation Research Institute.

[2] Sullivan, J. M., & Flannagan, M. J. (2001). *Characteristics of pedestrian risk in darkness.* (UMTRI – 2001 – 33). Ann Arbor, Michigan: University of Michigan, Transportation Research Institute.

[3] Mahlke, S., Rösler, D., Seifert, K., Krems, J.F., & Thüring, M. (in press). Evaluation of six night vision enhancement systems: Qualitative and quantitative support for intelligent image processing systems. *Human Factors.*

[4] Nielsen, J. (1993). *Usability Engineering*. Boston: Academic Press.

[5] Nielsen, J. (1994). Heuristic evaluation. In J. Nielsen, & R. L. Mack (Eds.), *Usability Inspection Methods,* (pp. 25-62), New York: John Wiley & Sons.

[6] Ericsson, K. A., & Simon, H. A. (1993). *Protocol analysis; Verbal reports as data (revised edition).* Cambridge, MA: Bradford books/MIT Press.

[7] Ericsson, K. A., & Simon, H. A. (1998). How to study thinking in everyday life: Contrasting think-aloud protocols with descriptions and explanations of thinking. *Mind, Culture, & Activity*, 5(3), 178-186.

[8] Eilers, K., Nachreiner, F., & Hänecke, K. (1986). Entwicklung und Überprüfung einer Skala zur Erfassung subjektiv erlebter Anstrengung. *Zeitschrift für Arbeitswissenschaft*, 40(4), 215-224.

[9] Hart, S. G., & Staveland., L. E. (1988). Development of a multi-dimensional workload rating scale: Results of empirical and theoretical research. In P. A. Hancock & N. Meshkati (Eds.), *Human Mental Workload.* Amsterdam: Elsevier.
[10] Miles, M. B., & Huberman, A. M. (1994). *Qualitative data analysis: An expanded sourcebook (2nd ed.).* Thousand Oaks: Sage Publications.
[11] Patton, M. Q. (2002). *Qualitative research and evaluation methods (3rd ed.).* Thousand Oaks: Sage Publications.