Transportation Research Part F 32 (2015) 46–55

Contents lists available at ScienceDirect

Transportation Research Part F

journal homepage: www.elsevier.com/locate/trf

Allocation of visual attention while driving with simulated augmented reality

Robin Eyraud [a,b,*], Elisabetta Zibetti [b], Thierry Baccino [b]

[a] BERTIN ERGONOMIE (Ergonomics & Human Factors Department of Bertin Technologies), 10 bis avenue Ampère, 78180 Montigny le Bretonneux, France
[b] Laboratoire CHART-LUTIN (Cognition Humaine & Artificielle EA 4004, Université Paris 8, EPHE Paris, Cité des Sciences et de l'Industrie, Paris), Université Paris 8, 2 rue de la liberté, 93526 Saint Denis – Cedex 02, France

A R T I C L E I N F O

Article history:
Received 16 May 2014
Received in revised form 16 April 2015
Accepted 27 April 2015
Available online 23 May 2015

Keywords:
Allocation of visual attention
Driving
Augmented reality

A B S T R A C T

Augmented reality (AR) technologies aim to optimize the visual attention of the driver by increasing the salience of high value elements. In such systems, 'value' is typically seen as linked to the general activity of driving, but not manoeuvres. However, several studies have shown that during activity, eye movements are specific to the immediate goal. In our experiment, 48 participants watched videos of automobile driving situations, during which they had to make decisions. In these videos, some situational cues were graphically highlighted. Depending on the experimental group, highlighted cues related to either the general driving task (e.g. road signs, pedestrians) or to a specific manoeuvre (e.g. look for overtaking cars before changing the lane). The results show that AR impacts the allocation of visual attention more strongly during the decision-making phase. In all AR conditions, the ability to extract information is less efficient. In particular, highlighting (by AR) general cues does not affect the detection of cues related to a manoeuvre, but it does change the allocation of visual attention: fixations are more numerous and less task-specific.

1. Introduction

The use of in-vehicle technologies (radio, GPS, cellular applications, etc.) multiplies the number of sources demanding the attention of drivers, and consequently totally changes their visual attention (Minin, Benedetto, Pedrotti, Re, & Tesauri, 2012; Benedetto et al., 2011; McGehee, 2001). When using such systems, visual attention is given to both the road environment and what is going on inside the car. The recent development of augmented reality (AR) systems hints at further major changes. These systems will make it possible to integrate information directly into the external environment, and thereby alter the driver's attention in yet another way. To better understand the impact of such systems, we refer to the general model of the allocation of visual attention (SEEV) developed by Wickens, Helleberg, Xu, and Horrey (2001).

The SEEV model (Wickens et al., 2001) explains the allocation of visual attention during activity in terms of top-down and bottom-up processes. It was originally applied to the activity of aircraft piloting (Wickens, Goh, Helleberg, Horrey, & Talleur, 2003) before being applied to automobile driving (Horrey, Wickens, & Consalus, 2006; Benedetto, Pedrotti, Bremond & Baccino, 2013). In the model, visual scanning is guided by four factors: Salience, Effort, Expectancy and Value. Salience is the main bottom-up factor of visual attention (Tattler, Hayhoe, Land, & Ballard, 2011) and several factors have been

* Corresponding author at: EA 4004 CHArt – Cognitions Humaine & ARTificielle, Université Paris 8, 2 rue de la liberté, 93526 Saint Denis – Cedex 02, France. Tel.: +33 (0)1 49 40 64 79; fax: +33 (0)1 49 40 67 54.

E-mail addresses: robin.eyraud@gmail.com (R. Eyraud), ezibetti@univ-paris8.fr (E. Zibetti), Thierry.baccino@univ-paris8.fr (T. Baccino).

http://dx.doi.org/10.1016/j.trf.2015.04.011
1369-8478/

identified, such as color, shape or motion (Itti & Koch, 2001). Effort is the second bottom-up factor, it corresponds to the visual angle between different pieces of information (i.e. the distance that the eye travels to reach a zone). This distance, if it is too far, can inhibit the intake of information (Kvalseth, 1977; Sheridan, 1970; Wickens et al., 2001). Expectation is a top-down factor and corresponds to the probability of seeing an event in a zone. Finally Value, another top-down factor, is the importance of a visual element in carrying out a task, and the importance of the task itself. This factor was included in the earlier Carbonell (1966) model, where the value of a zone depends on the benefit of looking at it, and the cost if it is ignored. To take the example given in Horrey et al. (2006), seeing lines painted on the road has many benefits during a lane change, while ignoring dashboard information does not have a high cost.

It is not possible to process all the visual information in a scene (Ullman, 1984) and any information management strategy must be specific to the task if it is to be efficient (Shinoda, Hayhoe, & Shrivastava, 2001). In the SEEV model, although visual attention is guided by all four factors it must be guided primarily by top-down factors (Expectation and Value) in order to be optimal. Conversely, bottom-up factors (Salience and Effort) must either be minimized or made use of (Wickens et al., 2003). While a salient element can guide the eye, if it is not relevant, it will simply be a visual distraction. At the same time, a relevant element that requires too much effort to be consulted is likely to be overlooked. An optimal allocation of visual attention is therefore highly dependent on how the device is designed. Interfaces can be designed in such a way that important information is highlighted, for example by a flashing fault indicator (increased Salience). Similarly, important information can be placed in the driver's field of vision – for example a blind spot sensor in the rear view mirror (reduced Effort). However, unlike the car cockpit, the road environment is uncontrollable, dynamic and relatively random. Consequently, only the Salience parameter can be used to support the optimal allocation of the driver's attention.

Highlighting cues helps attract attention (Jonides, 1980), and thus to detect a target more quickly (Fisher & Tan, 1989). In particular, in the field of aircraft piloting, it has been shown that this improves the accuracy of detecting an inconspicuous target (Yeh & Wickens, 2001). Head-Up Display (HUD) technology allows to draw attention on cued elements. But HUD is most effective when it provides conformal symbology, i.e. when it links elements of the display image to elements in the far domain (Caird, Horrey, & Edwards, 2001; Wickens & Long, 1995). Augmented reality (AR) is based on this principle of conformity. An AR system completes the real world with virtual elements that appear to coexist in the same space (Azuma et al., 2001). In its current form, the application of AR to automobile driving consists of the overlay of virtual elements on a Head-Up Display (HUD) or the windscreen (Krevelen & Poelman, 2010; Narzt et al., 2003). There are two AR modalities. The first adds information (such as a map, a direction or a point of interest) into the environment (e.g. Kim & Dey, 2009). The second highlights elements already present in the environment such as a road sign, line or pedestrian (e.g. Narzt et al., 2006). In this second modality, AR enhances the visibility of elements that have high value to the driver, in order to optimize their visual attention. Within this second AR context, simulator studies have shown that AR helps improve the detection of roadside hazards such as pedestrians or warning signs (Rusch et al., 2013). The same positive impact of AR have been observed for elderly drivers (Schall et al., 2013; Rusch, Schall, Lee, Dawson, & Rizzo, 2014). But while AR has the potential to optimize visual attention, risks of distraction remain (Ververs & Wickens, 1998). For instance, it has been observed that AR elements may mask or distract drivers from other relevant information (Schall, Rusch, Lee, Vecera, & Rizzo, 2010).

As currently conceptualized in AR systems, value relates to the environment and the general task of driving, but not manoeuvres. In Michon's (1985) classical model of driving, the manoeuvre is at an intermediate hierarchical level: between the strategic level (planning the activity of driving, route management) and the operational level (e.g. controlling the brake pedal). However, some authors suggest that during activity, the visual pathway is very specific to the immediate task. Studies of eye movements in daily tasks show that fixations extract very specific information related to the purpose of the current task (Rothkopf, Ballard, & Hayhoe, 2007). For example, in a task involving the preparation of tea, Land, Mennie, and Rusted (1999) noted that objects in the visual scene that were irrelevant to the action were rarely fixated. Echoing this, Hayhoe, Shrivastrava, Myruczek, and Pelz (2003) studied the eye movements of participants in a task involving sandwich preparation. Fixations were very specific to each stage of the achievement of the goal, and only 2% involved irrelevant objects (ingredients or unused tools). While visual attention seems to be focused on elements that are relevant to the purpose of the task, research based on the paradigm of attentional blindness shows that other, visually salient elements tend to be ignored (Mack & Rock, 1998; Simons & Chabris, 1999). In the same vein, Triesch, Ballard, Hayhoe, and Sullivan (2003) used the paradigm of change in an experiment where participants were asked to classify bricks in a virtual environment. Participants detected the change more often when it involved a property that was relevant to the purpose of the task.

Ultimately, the results reported in the literature emphasize that eye movements during activity may be influenced by the purpose of the task. Hence, eye movements seem to be very specific to the goal in progress, and tend to ignore irrelevant objects at time *t*. Therefore, it appears to be important that an AR system is capable of guiding the eye in ways that respect the process of information capture, in order to limit distractions problems for the driver. Our research examines changes to the allocation of visual attention created by a simulated AR system. We set up an experiment in which participants watched video footage of car driving situations. During the video, a voice instructs the driver to make a manoeuvre and the participant has to decide if the situation allows them to carry it out. Various cues in the visual scene were graphically highlighted: general driving cues; those related to a particular manoeuvre; all cues; or no cues. In addition, we controlled the congruence of cues with the manoeuvre, i.e. whether the cues made the manoeuvre possible or not. Finally, we looked at two phases of driving: the phase where the driver drives straight ahead and has no particular goal; and the phase where the driver intends to make a manoeuvre and must decide if the situation allows it.