Task Demand of Sequentially Presented Information While Moving in Augmented Reality*

Takahiko Kimura
Kansai University of Welfare Sciences,
Osaka, Japan

Naoyuki Midorikawa, Kazumitsu Shinohara,
Toshiaki Miura
Osaka University, Osaka, Japan

Recently, a large number of devices using AR (augmented reality) have been developed. AR devices are useful tools for acquiring information from the external world as long as allocation of attention to the space operates without difficulty. For example, a temporal limitation of attention deteriorates response performance when information is sequentially presented. In this study, we investigated temporal characteristics of attention in a moving context using AR information in near and far spaces. Single and dual task conditions were compared. In the dual task condition, the observer was required to respond to detect and discriminate. Results indicated that information from AR devices with high task demand would reduce response performance if the observer focused attention to the space regardless of near and far conditions. Sequential presentation of AR images produces high cognitive loads, and therefore, more attention resources are required under these conditions.

Keywords: augmented reality, visual attention, dual task, task demand, optic flow

Introduction

Although the rapid development information technology enables a variety of methods for presenting information, there remain several problems to be examined to ensure safe use. For example, HUD (head-up displays) and HMD (head-mounted displays) are valuable devices for the practical application of flight controls in aircraft and display information when driving a car. These types of display devices are expected to facilitate perceptual and cognitive tasks for observers by reducing head and eye movements. In general, because information using HUD and HMD is presented at a fixed distance from observers, this can potentially interfere with observer cognitive processing. For example, when driving a car, it is necessary for a driver to process external information in the foreground appropriately in order to maintain safe driving conditions. Depending on the driving situation, it has been shown that information presented at fixed distances can be troublesome (Morita, Mashiko, & Okada, 1997).

Recently, driving technology has been developed to usefully superimpose virtual images onto real world objects. This type of assistive technology is considered to be an application of MR (Mixed Reality) or AR.

*Acknowledgements: The authors gratefully acknowledge Koji Nakamura, Toshio Yamamoto, and Toshiro Kogure who were involved in a collaboration that took place prior to conducting this work. This work was supported in part by a grant from the Japan Society for the Promotion of Science.

Takahiko Kimura, Ph.D., associate professor, Faculty of Health Sciences for Welfare, Kansai University of Welfare Sciences.
Naoyuki Midorikawa, Graduate School of Human Sciences, Osaka University.
Kazumitsu Shinohara, Graduate School of Human Sciences, Osaka University.
Toshiaki Miura, Graduate School of Human Sciences, Osaka University.
(Augmented Reality) in which virtual images are generated by computer to present information. These MR or AR technologies have potential merit in that driver distraction can be reduced, because there is less head and eye movement. Also, this technology can make accessible real-time information related to the nature and position of driving hazards (Rusch, Schall, Gavin, Lee, Dawson, Vecera, & Rizzo, 2013). A novel idea with MR and AR technology is to superimpose information on an object of interest in the real world such as traffic signs and pedestrians. In these cases, to achieve synchrony between the real world objects of interest and superimposed information, it is important to recognize information in the external environment. For example, superimposed virtual images associated with real objects have been presented while walking and driving continuously. Thus far, the virtual images have been produced by retinal disparities (Nakamura, Ando, & Kawahara, 2005), and then superimposed information is made larger while the observer is moving forward (Suzuki, Uehira, Abekawa, & Tokumasu, 2007). These methods to acquire information are advantageous, because the virtual images are presented in necessary and suitable places.

However, some unresolved problems still remain in relation to human information processing, which is well known to have limited capacity. In regards to this limitation, attention is an important cognitive function that serves as a filter to select necessary information from the environment. The limited capacity of attention was a very popular and important finding (Wickens & McCarley, 2008). Because of this limitation, the human cognitive system sometimes makes mistakes when a dual task condition is performed like driving and using a cellphone. This indicates that the dual task situation has a higher cognitive load that can adversely affect safe driving behaviour.

More realistic situations were investigated by Rusch et al. (2013) who indicated that allocation of attention by AR spatial cues increases the detection of potential hazards such as pedestrians. In their study, an AR spatial cue was presented between 11 and 13 s. Because of the relatively long cue duration, the observers might be able to allocate attention continuously. Concerning characteristics of attention system, it is noteworthy that attention has a “bottleneck” which represents a limitation in human information processing (Pashler & Johnston, 1998). When a large amount of information is presented at the same time or sequentially, the capacity of human cognitive processing fails to capture it all because of the bottleneck limitation in the cognitive mechanism. Evidence of the cognitive bottleneck comes from a phenomenon known as the “attentional blink” (Shapiro, Arnell, & Raymond, 1997). An attentional blink occurs when a visual stimulus is presented sequentially. Typically, when a pair of target stimuli are presented, the later stimulus can be overlooked, because attention is occupied by processing the first stimulus. This problem would be related to the application of AR using virtual images, for example, if AR information was presented sequentially and briefly. Depending on the task demands, critical problems may arise in the use of AR technology for sequential information presentation.

Another unresolved problem is self-motion. In general, we use smartphones or other devices in connection with AR or MR while moving forward during activities such as walking and driving. Kimura, Miura, Doi, and Yamamoto (2009) reported that the characteristics of attention depended on whether a visual target was presented in near or far space. They found that attention shifted to an unexpected event in near space was faster than it did in far space, and that this effect was more remarkable when observers were moving forward. This revealed a possibility that peripheral information in the visual environment, called optic flow (Gibson, 1950), plays an important role in cognitive tasks involving central vision. All devices which make use of virtual images with AR technology should manipulate their images to accommodate forward movement by the user.
Rusch et al. (2013), using a driving simulator, showed AR cues that controlled attentional shift had a benefit on the detection of hazards while driving. These studies, therefore, suggest that an understanding of human cognitive function while moving is essential to further development of AR devices and the assessment of their safety.

In the present study, we conducted an experiment using a haploscope to simulate an AR environment in which virtual information was presented sequentially. The main aim of this study was to clarify characteristics of attention while using AR devices like HUD. Specifically, the effect of task demand on discrimination task was investigated using single or dual task conditions. According to the previous studies (Lavie, 1995; Kimura, Miura, & Shinohara, 2009), the higher task demand needs more attentional resource in both 2D and real 3D space. Therefore, the main hypothesis in this study is that the discrimination task would be deteriorated when more attentional resource is necessary because different AR information was presented sequentially. In addition, the effects of simulated self-motion by optic flow in the peripheral vision were examined.

**Method**

**Participants**

Eleven volunteers (five males and six females) with a mean age of 23 years ($SD = 3.0$) participated in this experiment. Each volunteer had normal or corrected-to-normal vision.

**Stimulus and Apparatus**

A haploscope was developed by using two 17-inch LCD (Liquid Crystal Displays: Mitsubishi RDT1713S) and two semi-transparent mirrors. The haploscope was positioned as in Figure 1 to present virtual images in a real environment. All stimuli were controlled by a computer (DELL Precision 360 2.8GHz) with graphic accelerator (Matrox Millennium G550), and a keyboard was used to collect responses. The focal distance for a virtual image was 70 cm from a participant. A circle of 4 mm diameter and an L or T (each line was 4 mm) were used for visual stimuli. These were presented as virtual images at either 847 cm (far space condition) or 287 cm (near space condition) from a participant. Furthermore, moving forward was simulated by optic flow, caused by a number of similar squares expanding from the centre of a display.

![Figure 1. Apparatus for the experiment.](image-url)
Experimental Design

The experimental design was as follows: presentation distance condition (near and far space) × optic flow condition (static and flow) × task (single and dual task). Therefore, we conducted the experiment using a three within-observer factors design.

Procedure

Each participant observed virtual images through the semi-transparent mirrors with binocular vision. Before beginning the experiment, each participant was required to confirm fused right and left eye images (fusion phase). A fusion phase preceded every trial for confirmation of preventing double vision. After every fusion phase, each participant started a trial by pressing a button (Figure 2).

![Figure 2. Procedure of dual task condition.](image)

After starting each trial, a green cross fixation point was presented for 1,000 ms, and the observers were required to gaze at it. At the moment when the fixation point was presented, optic flow by rectangles in the peripheral visual field was also presented until the end of a trial under the flow condition. In the dual task (2-task) condition, 1,000 ms after at the moment when the fixation point disappeared, a red circle representing a detection stimulus was presented. In every trial, the colour changed from red to green in any of 1 or 2 times, with a no colour change case in addition. The colour change occurred 1,000, 1,500, or 2,000 ms after the moment when the red circle was presented. These three durations were used to blind the participants to the timing of colour changes. It could be considered that this change detection task functioned as a cognitive load because the participant maintained their attention to specific space as a vigilance task (Mackworth, 1948). After 100 ms the colour of the circle returned from green to red. In the detection task, the participant was required to detect each colour change and to press the R key on a keyboard using their left index finger.

After disappearance of the circle for the detection stimulus, an L or T was presented in the same spatial position of the circle for a discrimination task until participant’s response. In the discrimination task, the participants were required to discriminate the shapes of letters and to press the assigned buttons (7 or 8 keys on a numerical keypad) using their right middle and index fingers. The assigned key for the shape of the letters was counter-balanced as much as possible between the participants to eliminate order effects.

In the single task condition, there was no detection task described as above. Therefore, only the letter discrimination task was conducted.
Results and Discussion

In order to examine the effect of a continuously presented virtual image on the letter discrimination task, error ratio and reaction times were analysed with and without a colour detection task and optic flow.

First, Table 1 shows error ratios for the letter discrimination task. Three-way, repeated measures ANOVA was conducted (presentation distance: near or far × with or without optic flow × with or without a detection task). There were no significant main effects or interactions, indicating that accuracies for the letter discrimination task were similar in the present experiment. Second, mean reaction times for the discrimination task are shown in Figure 3. Three-way, repeated measures ANOVA was conducted (presentation distance: near or far × with or without optic flow × with or without a detection task). The main effect of with or without a detection task was significant ($F(1,10) = 13.283, p < 0.01$). No other main effects and interactions were significant, indicating that reaction times to the single task condition were faster than those to the dual task condition in both near and far space regardless of whether optic flow was presented or not.

Table 1

<table>
<thead>
<tr>
<th>Error Ratios for the Letter Discrimination Task</th>
</tr>
</thead>
<tbody>
<tr>
<td>2-Tasks</td>
</tr>
<tr>
<td>Flow</td>
</tr>
<tr>
<td>Far (%)</td>
</tr>
<tr>
<td>Near (%)</td>
</tr>
</tbody>
</table>

Figure 3. Mean reaction times for a second stimulus (discrimination task) as a function of task condition (left: far presentation condition, right: near presentation condition). Error bars indicate standard errors.

These results indicated that the sequentially presented virtual images in the AR environment in near and far spaces could be a potential factor to increase reaction times, because presentation of virtual images in this experiment distracted attention from the letter discrimination task. Furthermore, because of the lack of differences in the error ratios, this result revealed no trade-off between speed and accuracy. From a human cognitive perspective, it is possible that PRP (the psychological refractory period) was related to this delay of the reaction time in the dual task condition (Pashler, 1998). PRP is described as a temporal feature of the attention mechanism. PRP is a delay in the processing of a second stimulus after finishing the processing of a first stimulus, because attentional resources are limited. Furthermore, in more realistic situation, Dixon, Daly,
Chan, Vescan, Witterick, and Irish (2013) indicated that the limitation of attention capacity caused the difficulty in awareness of the unexpected event when AR information was presented in surgical procedure. In the present study, because the first detection task required attentional resources, a bottleneck or limitation in processing could have occurred, and therefore, the discrimination task following the detection task was subject to interference.

Another possible explanation is that efficiency in task switching from one task to another task sequentially presented was related to the delay in the reaction times. Shinohara and Miura (2005) suggested that when using a car navigation system, processing demand of a task performed before task switching influenced attention distraction in the following task. Therefore, in the case where different types of AR information were sequentially presented, cognitive load may be increased and it may cause the interference of attention switching.

Previous study of the characteristics of attention while using AR devices indicated that, in the traffic situation, reaction times were longer when a virtual image of both intersection and moving direction (arrow) was presented using conventional HUD than when only a virtual image of moving direction was superimposed over an intersection in the real world (Fukano, Okabayashi, Sakata, & Hatada, 1994). The results in the present study revealed that there is a possibility of deterioration in reaction times when virtual images were presented sequentially while moving, indicating that further research is needed in order to make a guideline for establishing effective and safe use of AR devices.

In conclusion, we clarified an attention limitation to the information presented by virtual images when they are presented sequentially in both near and far spaces, indicating that the information presented by AR devices with high task demand can reduce response performance if the observer’s attention is focused on the space. Further studies are needed to reveal human cognitive processing while using advanced techniques such as AR.

References


