Comparison between Binocular and Monocular Augmented Reality Presentation in a Tracing Task

Akihiko Kitamura†, Hiroshi Naito†, Takahiko Kimura††, Kazumitsu Shinohara†, Takashi Sasaki (member)†††, Haruhiko Okumura (member)†††

Abstract To establish the superiority of monocular presentation in augmented reality (AR) over binocular presentation, we report two experiments involving a tracing task with an AR image covering a visual field observed in both the monocular and binocular conditions. In Experiment 1, the subjective visibility of the AR image and the accuracy of the tracing task were measured. In Experiment 2, participants identified a character in the AR image while tracing. The AR image was less visible in the monocular condition than in the binocular condition. Hence, participants could observe the real world more easily in the monocular condition, and the resulting performance on the tracing task was better in the monocular condition. Information acquisition from the AR image in the monocular condition was equivalent to that in the binocular condition. These results exhibit the superiority of monocular AR presentation for the performance of manual tasks.

Key words: augmented reality, monocular observation, optical see-through, manual task, visual attention, binocular rivalry

1. Introduction

Augmented reality (AR) is a technology used to superimpose information in the real world. Using AR, information can be allocated close to objects, and observers can thus simultaneously view the real world as well as visual information without needing to move their gazes. This characteristic is desirable for an in-vehicle information system and for safe driving. Hence, AR systems have been applied to various situations. Moreover, the market for head-up displays (HUDs) used for AR is expanding. Despite their significant advantages, there are some problems with AR systems. AR systems of certain kinds involve widespread AR images covering a wide range of the visual field, thus rendering it difficult for the observer to simultaneously view the real world. If the AR image is sufficiently large to cover the entire field of view (FOV), it becomes impossible for an observer to perform any task in the background.

Moreover, it is important for the human visual system to appropriately direct its attention to the relevant object in order to detect, acquire, and process visual information. With regard to the difficulty of viewing the real world in AR, it is a kind of problem in visual attention because it affects the selection of visual information. It is known that visual attention has two components: voluntary and automatic. The voluntary component of visual attention is endogenous, and is related to human intention and experience. The automatic component of attention is exogenous, and is related to the change in stimuli in the external world (e.g., abrupt illumination changes). With regard to the usability of AR interfaces, these components are crucial to safety and efficiency. Therefore, an investigation into the operation of visual attention should be valuable in designing an AR user interface.

In order to address the above problem, we have in past research proposed a "monocular AR presentation" where AR images are presented only to one of an observer’s eyes. We highlighted the advantage of monocular AR presentation. In our experiments, when an AR image was presented using binocular vision, it interfered with the detection of luminance changes in the background in the real world. On the contrary, when the image was presented using monocular vision, an observer was able to detect the changes in luminance across a wider area in the peripheral visual field. This result indicated that visual attention is distributed more widely in the real world when monocular
vision is used to present AR images.

However, the AR image used in the above experiment was a small character and, hence, the background was not completely covered. In an actual scene, AR images may be expanded to a wider area\(^2\), for example, the assembly guidance presented as the AR image is overlaid on the real image of assembled product. Characteristics of visual attention in this situation have not yet been examined. Therefore, the effects of visual attention in cases involving AR images covering a wider area of the FOV should be investigated.

When a wide AR image is presented monocularly, the input to the right eye is significantly different from that to the left eye. This situation leads to spontaneous alternations between one monocular image and another, which is known as binocular rivalry\(^8\),\(^9\),\(^10\). If one wants to maintain a percept when binocular rivalry occurs, voluntary attention needs to be directed to the relevant stimulus\(^8\),\(^9\). The dominant percept can hence be prolonged by voluntary attention\(^10\). Therefore, if the observers selectively direct their attention to AR image or real world, the AR image interferes with information acquisition less in the monocular presentation than in the binocular presentation.

We examined the superiority of the monocular presentation of a wide AR image in conveying real-world information over its binocular presentation. In Experiment 1, we compared the monocular condition with the binocular condition of AR image presentation for tracing accuracy and the perception of an AR image. In Experiment 2, we investigated information acquisition from the AR image during the tracing task. In the monocular condition where binocular rivalry should occur, we anticipated that observers could perform a manual tracing task in the real world accurately because they could acquire information from the real world.

2. Experiment 1

To examine the effect of the visibility of the AR image on the accuracy of a tracing task in Experiment 1, we controlled two pairs of contrast ratios (CR): the ratio of the AR image luminance to the pen-tablet monitor background (CR\(_{BKG/AR}\)), and the ratio of the luminance of the pen-tablet monitor background to the lines of the star-shaped figure (CR\(_{STR/BKG}\)).

2.1 Participants

Eight students participated in Experiment 1 (two males, six females). Their mean age was 22.1 (SD = 1.0). All participants had normal or corrected-to-normal vision (at least 0.8 as decimal visual acuity).

2.2 Stimulus and apparatus

A semi-transparent mirror, polarized filters, a liquid crystal display (LCD, Toshiba, REGZA 32C8000, 32 inches, 1366 × 768 dots), and a pen-tablet LCD monitor (WACOM, cintiq22HD, 21.5 inches, 1920 × 1080 dots) were positioned as shown in Fig. 1.

A star-shaped figure within a 27.5° × 27.5° (in visual angles) gray square was shown on the pen-tablet monitor (Fig. 2). The AR image was a translucent square, and covered the entire star-shaped figure. The luminance of the AR image was 16.2 cd/m\(^2\). The participants observed the AR image while observing the star-shaped figure. In the binocular condition, the AR image was shown to both a participant’s eyes. In the monocular condition, the AR image was only shown to the participant’s dominant eye. Since the normal LCD image was polarized, the participants observed the AR image through polarized filters which, in order to change the observation condition, could be individually rotated to intensify or diminish the AR image.

CR\(_{BKG/AR}\) was set to 1/8, 1/4, and 1/2, and CR\(_{STR/BKG}\) to 1/7, 1/4, or 1/2. When CR\(_{BKG/AR}\) increased, it became more difficult for participants to observe the background; and when CR\(_{STR/BKG}\) decreased, it became more difficult to observe the star-shaped figure. CR\(_{BKG/AR}\) = 1/8 and CR\(_{STR/BKG}\) = 1/2 was the most difficult to observe the star-shaped figure. CR\(_{BKG/AR}\) = 1/2 and CR\(_{STR/BKG}\) = 1/7 was the easiest to observe it.

To present stimulus and measure tracing time and the point of contact between the stylus pen and pen-tablet monitor, we made an application using Microsoft Visual Basic 2010 Express.
2.3 Procedure

The participants were asked to trace the star-shaped frame (Fig. 2). After the trial, participants rated the visibility of the AR image. The participants first touched a green circle on the star-shaped figure with a stylus pen. The AR image was then shown to them, and covered the star-shaped figure. The participants then traced within the frame. When the participants’ trace returned to the starting point (◦ in Fig. 2), the star-shaped figure disappeared and the next trial set began.

We defined tracing time as the time taken to trace the star-shaped figure, and the trace error as the distance between the center of the contour of the star-shaped figure and the location of the stylus pen (see Fig. 3). Given that the trace error increased with the increase in the distance between the center of the frame and the stylus pen, the center of the frame was defined as having trace error 0. So long as the trace error was less than 4 pixels, participants were able to successfully trace between the two lines of the star-shaped figure. The subjective visibility of the AR image was measured as the dependent variable (range = 0 – 100: 0 = full perception of the AR image, 100 = complete invisibility of the AR image). If binocular rivalry occurred, the AR image was perceived as somewhat invisible. Participants rated the visibility of the image following the trial.

2.4 Results

We analyzed tracing times using a 2 (observation condition: binocular, monocular) × 3 (CR$_{BKG/AR}$: 1/8, 1/4, 1/2) × 3 (CR$_{STR/BKG}$: 1/7, 1/4, 1/2) analysis of variance (ANOVA). Only the main effect of CR$_{BKG/AR}$ was significant [$F$ (2, 14) = 4.23, $p < .05$]. In accordance with Ryan’s method, the tracing time with CR$_{BKG/AR}$ = 1/8 was significantly longer than that at 1/2, and was marginally longer than at 1/4.

Fig. 4 shows the trace error as a function of the observation conditions, CR$_{BKG/AR}$ and CR$_{STR/BKG}$. Trace errors were analyzed in the same manner. The main effects of the observation condition [$F$ (1, 7) = 15.70, $p < .01$], CR$_{BKG/AR}$ [$F$ (2, 14) = 16.46, $p < .001$], CR$_{STR/BKG}$ [$F$ (2, 14) = 11.04, $p < .005$], and the interaction between the observation condition and CR$_{BKG/AR}$ × CR$_{STR/BKG}$ [$F$ (4, 28) = 3.36, $p < .05$] were all significant.

The results showed that at CR$_{BKG/AR}$ = 1/8, the trace error was higher in the binocular condition than in the monocular condition ($p < .001$). Moreover, in the binocular condition, the trace error was higher when CR$_{BKG/AR}$ = 1/8 than when CR$_{BKG/AR}$ = 1/4 and 1/2 ($p < .05$). This result indicates that as it became more difficult to observe the star-shaped figure, the trace error increased in the binocular condition but
Subjective AR image visibility

Error bars indicate standard errors. A subjective AR image visibility rating of 0 indicates full visibility and 100 indicates complete invisibility. Asterisks indicate significant differences between the observation conditions. Those in CR\textsubscript{BKG}/AR and CR\textsubscript{STR}/BKG are not depicted.

remained constant in the monocular condition.

Subjective visibility was analyzed using three (observation condition, CR\textsubscript{BKG}/AR and CR\textsubscript{STR}/BKG) factorial ANOVA (Fig. 5). The main effects of the observation condition \([F(1, 7) = 25.26, p < .005]\), CR\textsubscript{BKG}/AR \([F(2, 14) = 31.65, p < .001]\) and CR\textsubscript{STR}/BKG \([F(2, 14) = 16.04, p < .001]\) were significant. The AR image was determined to be less visible in the monocular condition than in the binocular condition. Further, as CR\textsubscript{BKG}/AR increased or CR\textsubscript{STR}/BKG decreased, the AR image was more clearly perceived.

2.5 Discussion

In Experiment 1, with binocular vision and in difficult (CR\textsubscript{BKG}/AR was high and CR\textsubscript{STR}/BKG was low) conditions, the accuracy of the tracing task was impaired. On the contrary, with monocular vision, the accuracy remained high in all contrast conditions (Fig. 4). With monocular vision, the AR image was determined to be less visible in the monocular condition than in the binocular condition. Further, as CR\textsubscript{BKG}/AR increased or CR\textsubscript{STR}/BKG decreased, the AR image was more clearly perceived.

3. Experiment 2

While the tracing tasks of Experiment 1 and 2 were similar, in Experiment 2 participants had to discriminate an alphanumeric character (a letter or number) appeared in the AR image, and we compared monocular and binocular observation.

3.1 Objectives

During the tracing task, the AR image was less visible in Experiment 1 in the monocular condition. This result may imply that observers cannot acquire information from the AR image in the monocular condition. If participants cannot see the AR image, it is useless for information presentation systems.

In contrast to Experiment 1, the participants had to acquire information from the AR image in Experiment 2. The discrimination task was to imitate an actual AR situation, such as inspection or assembly. If, in the monocular AR condition, participants can acquire information from the AR image (or visual attention is distributed to the AR image efficiently just as well as in the binocular condition), their reaction times to the information in the AR image should be equivalent to that the binocular condition.

A non-dominant condition (the AR image was presented to the non-dominant eye) was included as part of the monocular observation in order to examine the possibility that the dominant and non-dominant eyes operate according to different processes.

3.2 Participants

Nine students took part in Experiment 2 (all males). Their mean age was 24.4 (SD = 2.0). All participants had normal or corrected-to-normal vision (at least 0.9 as decimal visual acuity).

3.3 Stimulus and apparatus

The arrangement of the apparatus was the same as in Experiment 1. The luminance of the AR image was set at 16.2 cd/m\textsuperscript{2}. The contrast ratio was fixed. The value of CR\textsubscript{BKG}/AR was set to 1/4 and that of the CR\textsubscript{STR}/BKG to 1/2. Moreover, an alphanumeric character, 1.2° in visual angle squared, in the AR image was used for the discrimination task. The character was pre-
presented at one of nine locations: the point touched by
the stylus pen and eight points deviating from the loca-
tion touched by the pen by 1°, 2°, 4°, or 6° to either
the center or outside the star-shaped figure.

3.4 Procedure
During the tracing task, participants had to press
the correct key depending on an alphanumeric charac-
ter. Reaction time was used as the indicator of infor-
mation acquisition, and was defined as time it took a
participant to press the key after presentation of the
alphanumeric character. If participants could acquire
information in the monocular condition just as well as
in the binocular condition, there would be no differ-
ence in reaction times. Reaction time and trace error
were measured in the same way as in Experiment 1. In
Experiment 2, there were three observation conditions:
binocular, dominant, and non-dominant.

3.5 Results
We analyzed the trace error using a 3 (observation
conditions: binocular, dominant, non-dominant) × 9
(alphanumeric stimulus position) ANOVA. The anal-
yses indicated a significant main effect of observation
condition [F (2, 16) = 15.19, p < .001]. As shown
through Ryan’s method, the trace error was signifi-
cantly higher in the binocular condition than in both
the dominant and non-dominant monocular conditions
(p < .05). Thus, performance on the tracing task was
better during monocular observation than during binoc-
ular observation. There was no interaction.

Reaction times for discrimination of the AR alphanu-
meric character were analyzed in the same manner as
above (Fig. 6). Only the main effect of alphanumeric
stimulus position was significant [F (8, 64) = 28.46,
p < .001]. In all observation conditions, reaction time
increased as the distance between the location of the
alphanumeric character and stylus pen increased.

3.6 Discussion
In Experiment 2, the trace error was higher in the
binocular condition than in both the dominant and non-
dominant monocular conditions as well as the results of
Experiment 1. Thus, the advantage of the monocular
condition persisted even when participants had to ac-
quire information from the AR image.

Furthermore, there were no significant differences in
reaction times for discrimination in the observation con-
ditions, regardless of the position of the character (Fig.
6). This result implies that monocular condition is
equivalent to binocular condition in terms of informa-
tion acquisition from AR images.

In summary, monocular presentation is superior for
real-world tasks than binocular presentation, even when
observers need to pay attention to an AR image and the
image is less visible in the monocular presentation.

4. General discussion
We conducted two experiments to show the superior-
ity of the monocular AR presentation for manual tasks
as a wide AR image covered the foreground.

In Experiment 1, the AR image was rated to be less
visible in the monocular condition. This result is con-
gruent with our hypothesis that participants can ob-
serv e the real world more easily in the monocular con-
dition than in the binocular condition. For the tracing
task, participants were required to direct their attention
to the star-shaped figure. Voluntary attention prolongs
the dominant percept, and hence participants were
able to suppress the percept of the AR image and sus-
tain that of the star-shaped figure. As a result, the
AR image was less visible in the monocular condition,
which made the tracing task easier.

In Experiment 2, participants were asked to acquire
and distinguish information during the tracing task in
the monocular condition as well as the binocular condi-
tion. The results indicate that attention could operate
in the monocular condition equivalently to the binocu-
lar condition. When a novel stimulus is presented to the
FOV and includes a change in brightness, form, or color,
attention is automatically directed to the stimulus,
and when the stimulus is presented to the fellow eye,
the stimulus has the potential to change the percept
due to binocular rivalry. Based on these findings,
the alphanumeric character in the AR image presented
to the fellow eye was involuntarily attended to. This
might have enabled the rapid information acquisition
in the monocular condition equivalent to that in the
5. Conclusion

In this study, we established the superiority of monocular AR presentation over binocular presentation when a wide AR image, which is suitable for inspection or assembly tasks, covered the background. In the monocular condition, the AR image was less visible, resulting in more accurate performance in tracing task than in the binocular condition. When participants were required to acquire information from the AR image, as well, their information extraction performance in the monocular condition was equivalent to that in the binocular condition. These results show the advantages of monocular AR presentation for wide AR images. Previously, it has been shown that the monocular AR presentation is feasible with use of a concave reflector to focus light to one ular AR presentation is feasible with use of a concave reflector to focus light to one ular AR presentation is feasible with use of a concave reflector to focus light to one ular AR presentation is feasible with use of a concave reflector to focus light to one

Results obtained in this study show the advantage of monocular AR over binocular AR, particularly for presenting wide AR images.

References

1) M. L. Rusch, M. C. Schall Jr., P. Gavin, J. D. Lee, J. D. Dawson, S. Vecerra and M. Rizzo: "Directing driver attention with augmented reality cues." Transportation Research Part F: Traffic Psychology and Behaviour, 16, pp. 127 - 137. (Jan. 2013)
2) K. Kiyokawa: "HMD Technologies for AR." Proceedings of The international Display Workshops, 21, pp. 1300-1303. (Dec. 2014)
3) C. Kusaka and J. Kimura: "Projection Technology Will Take Automatic Head-up Display to the Next Level." Proceedings of The International Display Workshops, 21, pp. 1076-1078 (Dec. 2014)
4) R.D. Wright and L.M. Ward: "Orienting of attention." Oxford University Press, New York (2008)
5) T. Sasaki, A. Hotta, A. Moriya, T. Murata, H. Okumura, K. Horinuchi, N. Okada, M. Ogawa and O. Nagahara: "Hyperrealistic Display for Automotive Application." SID Symposium Digest of Technical Papers, 41, 1, pp. 953-956. (May 2010)
6) A. Kitamura, H. Naito, T. Kimura, K. Shinohara, T. Sasaki and H. Okumura: "Distribution of Attention in Augmented Reality: Comparison between Binocular and Monocular Presentation." IEICE Transactions on Electronics, E97-C, 11, pp. 1081-1088. (Nov. 2014)
7) A. Kitamura, H. Naito, T. Kimura, K. Shinohara, T. Sasaki and H. Okumura: "The Superiority of Widespread Monocular Augmented Reality Presentation in a Manual Tracing Task." Proceedings of The International Display Workshops, 21, pp. 1337-1340 (Dec. 2014)
8) P. Zhang, K. Jamison, S. Engel, B. He and S. He: "Binocular rivalry requires visual attention." Neuron, 71, 2, pp. 362-369 (July 2011)
9) P. Zhang, Y. Jiang and S. He: "Voluntary Attention Modulates Processing of Eye-Specific Visual Information." Psychological Science, 23, 3, pp. 254-260 (Feb. 2012)
10) S.C. Chong, D. Tadin and R. Blake: "Endogenous attention prolongs dominance durations in binocular rivalry." Journal of Vision, 5, 6, pp. 1004-1012 (Dec. 2005)
11) J. Theeuwes: "Cross-dimensional perceptual selectivity." Perception & Psychophysics, 50, 2, pp. 184-193 (Aug. 1991)
12) T.L. Ooi and Z. J. He: "Binocular rivalry and visual awareness: the role of attention." Perception, 28, 5, pp. 551-574 (1999)