Peripheral Vision Annotation: Noninterference Information Presentation Method for Mobile Augmented Reality

Yoshio Ishiguro
Graduate School of Interdisciplinary Information Studies,
The University of Tokyo.
JSPS Research Fellow
7-3-1 Hongo, Bunkyo
Tokyo, JAPAN
Ishiy@acm.org

Jun Rekimoto
Interfaculty Initiative in Information Studies,
The University of Tokyo.
Sony Computer Science Laboratories, Inc.
7-3-1 Hongo, Bunkyo
Tokyo, Japan
rekimoto@acm.org

ABSTRACT
Augmented-reality (AR) systems present information about a user's surrounding environment by overlaying it on the user's real-world view. However, such overlaid information tends to obscure a user's field of view and thus impedes a user's real-world activities. This problem is especially critical when a user is wearing a head-mounted display. In this paper, we propose an information presentation mechanism for mobile AR systems by focusing on the user's gaze information and peripheral vision field. The gaze information is used to control the positions and the level-of-detail of the information overlaid on the user's field of view. We also propose a method for switching displayed information based on the difference in human visual perception between the peripheral and central visual fields. We develop a mobile AR system to test our proposed method consisting of a gaze-tracking system and a retinal imaging display. The eye-tracking system estimates whether the user's visual focus is on the information display area or not, and changes the information type from simple to detailed information accordingly.

Categories and Subject Descriptors
H5.2. Information interfaces and presentation (e.g., HCI): User Interfaces.

General Terms
Human Factors

Keywords

Wearable AR, retinal imaging display, eye gaze.

1. INTRODUCTION
Humans do not perceive elements of their visual field equally fully: their gaze point, central vision, and peripheral vision differ in shape, color, and other object perceptions.

In augmented-reality (AR) environments, the annotation (virtual) object is superimposed over the real environment. In this environment, users are able to visualize computer-generated objects just as they perceive real objects. The annotation objects overlap each other when there is a lot of information. Much research has been conducted on how to present information effectively by using methods for annotating object layout and selecting information [1-3].

Mobile AR systems are also being actively researched. These systems are expected to support people's daily lives through their use of mobile devices. Mobile handheld AR systems based on smartphones, such as Layar [4] and Wikitude [5], are gaining attention with the trend toward mobile AR. These systems combine camera-captured images with overlaid information annotation and help users to understand the real-world environment. For the mobile environment, the freehand usage scenario is important, and for this reason, head-mounted displays (HMDs) based on AR are considered to be convenient. Additionally, gaze direction is measured for freehand information manipulation such as pointing.

Previous information presentation methods [1, 3] do not consider the mobile usage scenario. As a result, many annotation objects are displayed using HMDs, causing occlusion and blind spots in the user's visual field and ultimately distracting the user's attention (Figure 1). Key considerations for safe mobile interaction are

- Not disturbing the user's behavior with virtual objects.
- Freehand interaction.

To solve these problems, we propose a method for controlling the

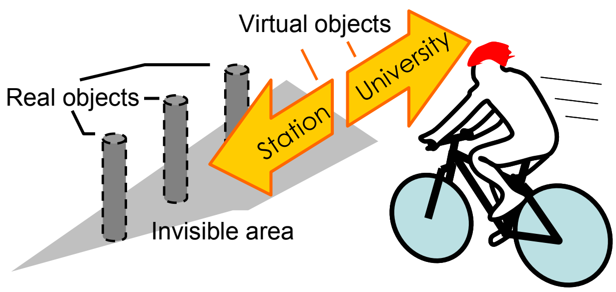

Figure 1. The overlaid information annotations create blind spots in the user's visual field, and ultimately disturb the user's behavior.

information that is present by sensing gaze direction and further examining the differences in human visual perception. Our proposed method can distinguish the peripheral and central vision areas by measuring gaze direction, and then the system changes the level-of-detail for annotated information (Figure 2).

We propose a mobile AR system (Figure 3) that displays detailed information of people, texts, and other gazed-at objects that the system has automatically extracted [6]. This system uses HMD, without disturbing the user's activities.

2. RELATED WORK

The annotation layout and/or control methods have been examined in AR research [1-3], such as eliminating crossover virtual annotations and changing the amount of annotation by background objects. Leykin and Tuceryan proposed the method for determining the readability of text labels by using a pattern-recognition approach [3]. It used textural properties and other visual features to determine readability. Nakamura proposed a method that controlled the amount of annotation by measuring glabellas [2]. These methods are useful in the mobile environment because the user's hands are freed for other uses.

For controlling the amount of information, Mark Weiser uses "periphery" to name what we are attuned to without attending to it explicitly [7]. For information overload, calm technology engages both the center and the periphery of our attention, and in fact moves back and forth between the two. Jilter proposed a region-based information filtering algorithm [8]. This algorithm takes the state of the user and the state of individual objects about which information can be presented. It can dynamically respond to change in the environment and the user's state; however, without view management, presenting the viewer with information from just a single annotation object can blind the user's view. View management is also important for the safe use of mobile AR systems.

Eye movements used to control the information presentation of HMDs. It have not been studied in terms of avoiding interfering with people's activities. However, using gaze direction for selection purposes requires consideration of the "Midas touch problem" [9]. This problem occurs because the eye movement cannot replace the computer mouse entirely, because eye movement is not able to "click" like a mouse. Thus, a target such as an icon may be involuntarily selected when the viewer looks at it to get detailed information. To solve this problem, dwell-time selection is usually used for gaze interaction [10].

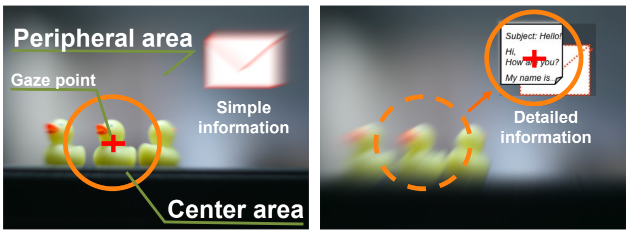

Figure 2. This illustration shows our proposed method. Icons are presented in the user's peripheral area, which automatically change to detailed information when the user gazes at them.

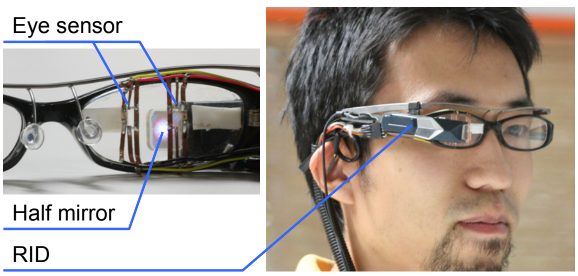

Figure 3. The eye tracker [6] and information display consisted of eye-sensing glass and a retinal imaging display (RID). This device can detect eye activity (gaze direction, blinks), and it can show the image by using a laser-light source.

3. GAZE OPERATED INFORMATION PRESENTATION

Humans can control the amount of information about their environment by creatively using the difference between the central and peripheral visual fields. The central visual field can recognize smaller objects and capture more detailed information than can the peripheral visual field. Figure 4 shows our visual perception [11]. Our eye can easily recognize a simple icon from the center of view to 60 degrees in the visual field; however, letters are not recognizable in that same range. In addition, our eye cannot recognize texts outside of an area within 20 degrees of the gaze point. Thus, the user cannot read detailed textual content in peripheral visual areas. However, the user can be made aware that such content is present in the display.

In this research, the user's eye direction is used as the method of operating the presentation of information. Normally, our eyes watch in the frontal and horizontal directions of our body (-5 degree in the case of standing up straight) [11]. Clearing this field of view is important during activities such as walking; thus, information is not presented in the center of view, and the display area is set within the other field of view. Consequently, in this study, we propose following two information presentation methods by using these human visual perceptions.

3.1 Cursor see-through method

When the system recognizes a real object, the real object is highlighted visually by using AR technology and object-recognition methods [6]. The user sees a highlighted object through the cursor area. In this method, the cursor area is placed at the edge of the view angle and is displayed visually. The user can select real objects by capturing them in the cursor area.

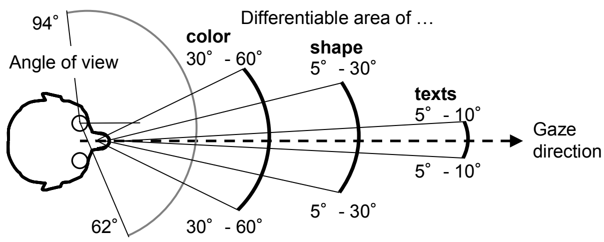

Figure 4. This illustration shows recognizable angle for text, shapes, and color of human [11].

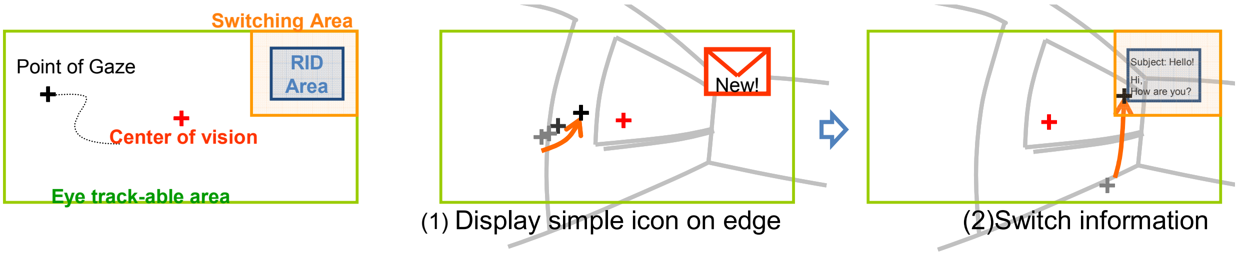

Figure 5. Peripheral vision annotation method. The display area is placed at the edge of the field of view. New information becomes available about the objects that are gazed at [6]: arriving email and similar kinds of events. (1) The simple icon will be shown in the display area. (2) The user gazes at this icon when interested in its information; then the display provides detailed information. If the user loses interest and ceases the gaze, it is just ignored.

Normally, users do not regularly gaze at this edge area. Thus, these specific types of gaze actions must be done intentionally, and the system can recognize the user's target-selecting actions by using eye-gaze direction and its cursor area. This method uses intentional eye movements, without any need for predetermined dwell times, to directly select real objects.

3.2 Peripheral vision annotation method

The cursor see-through method uses real target (choice) objects for notification and selecting. In the peripheral vision field, information is presented by using the following two processes.

1. Noticing simple icons in the user's peripheral vision field when the information is available.
2. Displaying detailed information once the simple icons are noticed and gazed upon in the display near the icon area.

These processes allow users to get detailed information only when it is needed. Moreover, since this switching does not require physical tools, it can be operated smoothly. Since detailed information is shown only when the display area is gazed at, the display area can be set at the periphery of the visual field, and the user's activity will not be disturbed (Figure 5). The "switching area" is placed at the same site as the "RID area." The RID area presents the annotations. Our proposed method takes advantage of the "Midas touch problem." Thus, it is possible to see information whenever the user wants to do so.

4. HARDWARE DESIGN

To confirm the method's validity, a mobile AR display system is built using a combination of a mobile, spectacle-type, wearable RID and an eye-tracker system.

4.1 Retinal Imaging Display (RID)

The retinal imaging display (RID) is made by the Brother Industries, Ltd. Images projected onto the retina appear as if they existed in front of the user. The transparent optics enable users to see RID images without affecting their visual field [12]. This device uses micro-electromechanical system (MEMS) technologies. A laser is used for the light source (Figure 6). The resolution is 800 x 600 (SVGA), and the view angle of the RID is 18 x 13.5 degrees. Our system allocates the display at the periphery of the visual field, so it does not disturb human daily activities such as walking, running, and other activities, because the display's view angle is very narrow compared with a human's full field of vision. This display uses a small half-mirror in front of the view; thus, other fields of view are not blocked. Moreover, wherever the user is looking in the real world, the displayed CG objects always come into focus. On the other hand, the LCD-based HMD can be seen with binocular stereopsis, but it is difficult to express a 3D scene because the focal length is limited by the optics system. Since user does not need to change the focus by using an RID-based HMD, this system is adapted more for mobile AR systems than LCD-based HMDs.

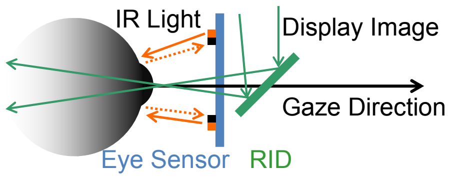

Figure 6. Combination with eye sensor and RID. These devices do not interference each other function.

4.2 Aided-Eyes: mobile eye-sensing glasses

The eye-activity sensing glasses that use the corneal limbus tracker method can measure gaze direction by using an infrared LED and a phototransistor [6]. Two infrared LEDs and four phototransistors are mounted on the glasses (at the side of the eye). The infrared light is reflected by the eye surface and is received by phototransistors. Because this device uses only a small light source and a phototransistor, the glasses do not block RID light passage, and these glasses have high compatibility with RIDs.

In order to be robust for environmental and display light, the time delta sample-and-difference measurement circuit, used as a modulated IR light source for the eye tracker, is added to the system in this study. This circuit can rapidly turn on and off in reference to the IR light source. One of the band pass filter (BPF) methods to remove noise is used. In this study, this circuit works at approximately 49 kHz BPF for higher robustness. The eye-tracker specifications are a sampling rate of 600 Hz and an accuracy of approximately 5 degrees.

4.3 Combination with Aided-Eyes and RID

Before the eye-gaze direction measurement, the head position and the display are fixed for calibration. Then, the display shows the target to be gazed at for the calibration. After the calibration, the user sets the RID display's location. These processes make it

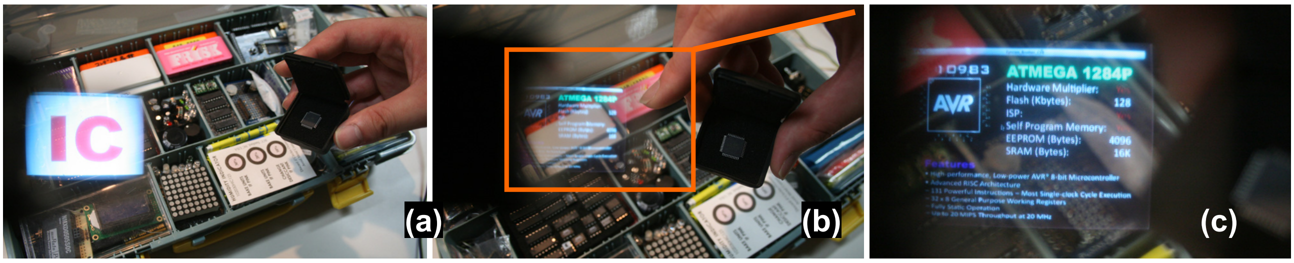

Figure 7. These photographs show the displayed icon and detailed information (a simulated image that is captured by the camera). (a) The simple icon appears in the display area. The icon shape is changed by the available information. Humans can perceive these simple icons that appear in the periphery of the visual field. (b) The displayed information can be switched to detailed information by gazing at a simple icon. This photograph captures a point in time after the simple icon was gazed at and it switched to the details of this chip. (c) The user can easily read text information by gazing at and concentrating on the display area (which is just a photographic enlargement).

possible for the system to record RID positional information in the field of view.

5. RESULTS

The RID does not disturb the field of view when it does not present information (Figure 7). However, the icon is recognized quickly once information for presentation appears. In that case, the user sees the icon naturally, and the system automatically switches from the icon to detailed text information. The response time from the user's gazing at the icon is less than 50 ms (because the GUI system works at 24 FPS). The user can easily read text information by gazing at the display area, because the display has high (800 x 600) resolution. This method can be changed smoothly from the icon to detailed information by recognizing that the user's interest has moved from the central view to the peripheral view. In addition, the gaze-tracking system is not affected by the laser light source from the RID, and the presentation image is not disturbed by the gaze tracker either.

We confirmed these methods of using the eye tracker and RID. It has a narrow angle of display, compared to the human field-of-view. Thus, we did not consider displaying multiple icons under the existing conditions. In the future, our proposed method will effectively work by setting the display area so the HMD will cover the entire field of view.

6. CONCLUSION

In this paper, a gaze-operated information-presentation method for head-mounted retinal-imaging displays was studied. This method used gaze direction for operating the presentation information. The user could detect an icon that was placed in a peripheral area, and could choose the icon with their gaze. Additionally, the icon did not disturb the user's central field of view; thus, the user could also ignore the icon. The system with a mobile eye tracker and an RID was built, and it confirmed our proposed method.

We also studied an eye tracker and gazed-at target object-extraction system. Previously, we already succeeded with a prototype system. Thus, for future work, we will carry out additional evaluation of the information-presentation system and the information-extraction system with long-term, real-life experiments. We consider that mobile AR systems using our proposed information presentation method will be able to support humans in their daily lives.

7. ACKNOWLEDGMENTS

We would like to thank Brother Industries, Ltd for providing valuable devices. This research was partially supported by the Ministry of Education, Science, Sports and Culture, Grant-in-Aid for JSPS Fellows, 21-8596, 2009.

8. REFERENCES

1. Bell, B., Feiner, S. and Höllerer, T. View management for virtual and augmented reality. *Proc. UIST '01*. ACM Press (2001), 101-110.
2. Nakamura, H. and Miyashita, H. Control of augmented reality information volume by glabellar fader. *Proc. Augmented Human 2010,* ACM Press, (2010), 1-3.
3. Leykin, A. and Tuceryan, M. Automatic Determination of Text Readability over Textured Backgrounds for Augmented Reality Systems. Proc. ISMAR 2004. IEEE Computer Society, (2004), 224-230.
4. Layar Reality Browser, http://www.layar.com/
5. Wikitude, http://www.wikitude.org/
6. Ishiguro, Y., Mujibiya, A., Miyaki, T., and Rekimoto, J. Aided eyes: eye activity sensing for daily life. *Proc. Augmented Human 2010*, ACM Press, (2010), 1-7.
7. Weiser, M. and Brown, J. S. Designing calm technology, PowerGrid Journal, (1996).
8. Julier, S. Feiner, S. Sestito, S. Information Filtering for Mobile Augmented Reality, Computer Graphics and Applications, vol. 22, issue 5, 2002
9. R. J. K. Jacob. Eye movement-based human-computer interaction techniques: Toward non- command interfaces. *In Advances in Human- Computer Interaction*, Ablex Publishing Co, (1993), 151-190.
10. Park, H, M. Lee, S, H. Choi, J, S. Wearable Augmented Reality System using Gaze Interaction, *Proc*. ISMAR 2008
11. Komatsubara, A. Human error, Maruzen co. ltd. 2008. (In japanese)
12. Mobile Spectacle-type Wearable Retinal Imaging, Display http://www.brother.com/en/news/2009/rid/