Federal Aviation Administration

DOT/FAA/AM-17/4
Office of Aerospace Medicine
Washington, DC 20591

Current Status of Gaze Control Research and Technology Literature Review

Matthew J. Hollomon
Daniela Kratchounova
David C. Newton
Kevin Gildea
William R. Knecht

Civil Aerospace Medical Institute
Federal Aviation Administration
Oklahoma City, OK 73125

January 2017

Final Report

NOTICE

This document is disseminated under the sponsorship of the U.S. Department of Transportation in the interest of information exchange. The United States Government assumes no liability for the contents thereof.

This publication and all Office of Aerospace Medicine technical reports are available in full-text from the Civil Aerospace Medical Institute's publications website: http://www.faa.gov/go/oamtechreports

Technical Report Documentation Page

1. Report No.	2. Government Accession No.	3. Recipient's Catalog No.
DOT/FAA/AM-17/4		

4. Title and Subtitle	5. Report Date
Current Status of Gaze Control Research and Technology Literature Review	January 2017
	6. Performing Organization Code
7. Author(s) Hollomon MJ, Kratchounova D, Newton DC, Gildea K, Knecht WR	8. Performing Organization Report No.
9. Performing Organization Name and Address FAA Civil Aerospace Medical Institute P.O. Box 25082 Oklahoma City, OK 73125	10. Work Unit No. (TRAIS) 11. Contract or Grant No.
12. Sponsoring Agency name and Address Office of Aerospace Medicine Federal Aviation Administration 800 Independence Ave., S.W. Washington, DC 20591	13. Type of Report and Period Covered 14. Sponsoring Agency Code

15. Supplemental Notes

16. Abstract

Gaze control can take different forms throughout Majaranta and Bulling's (2014) explicit-implicit continuum. Specifically, it can range from explicit eye command input and control to passive eye monitoring diagnostic applications. Gaze control can also be used either as a stand-alone control modality or as part of a larger multi-modal interface. In fact, employing eye tracking and eye movements in control input user interfaces in aviation, and emulating the naturally occurring modality communication and cooperation patterns in the design of multimodal interfaces, are not new. The advantages of using gaze control as part of a larger multimodal interface include intuitiveness, mimicking natural behavior, and reduction of uncertainty in interpreting user intent. The current level of gaze control technology is at a point where it is viable for research in a flight deck environment.

17. Key Words	18. Distribution Statement
gaze control, eye tracking, multimodal control input interfaces	Document is available to the public through the Internet: http://www.faa.gov/go/oamtechreports/

19. Security Classif. (of this report)	20. Security Classif. (of this page)	21. No. of Pages	22. Price
Unclassified	Unclassified	36	

Form DOT F 1700.7 (8-72) Reproduction of completed page authorized

ACKNOWLEDGMENTS

The authors wish to thank Janine King and Debbie Perry for their assistance in quality assurance, source identification, and research documentation, and Roni Anderson for tracking down an extensive list of references used in this report.

Contents

CURRENT STATUS OF GAZE CONTROL RESEARCH AND TECHNOLOGY LITERATURE REVIEW

List of Figures

CURRENT STATUS OF GAZE CONTROL RESEARCH AND TECHNOLOGY LITERATURE REVIEW

INTRODUCTION

Eye gaze occurs when the eyes look steadily and intently at an object. This occurs naturally and requires little conscious effort. It not only shows where visual attention is directed, but it often precedes actions—that is, people look at things before acting on them. Eye gaze as a control input device for user interfaces, such as computers, is impossible without eye-tracking technology. Consequently, eye tracking and gaze control are terms often used interchangeably. While the main focus here is on gaze as a control input device, the two terms are addressed individually in the context of the history, research, development, and technology implementation.

Gaze control is defined as any control input method that uses eye movements, either in part or completely, to interact with a user interface (Majaranta, Ahola, & Špakov 2009; Majaranta & Bulling, 2014). Current gaze control systems usually contain an eye tracker as a system component. It detects the eye movements of the operator using the most common method of eye tracking—the Pupil Center Corneal Reflection.[1] For example, one method for gaze control input uses visual fixation (i.e., dwell) on a target object or a display area for a specific predetermined length of time (Hansen, Engell-Nielsen, & Glenstrup, 1998).

A review of the research literature on gaze control is provided. The review includes an overview of gaze as (a) a unimodal control input device,[2] and (b) an integrated input device within a multimodal control input interface. Eye tracking and gaze control in aviation research literature and major human factors considerations are addressed. A comprehensive history of eye tracking and the current state of the art of gaze control are provided in the appendices.

UNIMODAL GAZE CONTROL INPUT TECHNOLOGY CONTINUUM, METHODS, AND APPLICATIONS

Single modality (unimodal) user interfaces employ one human sensory system (e.g., vision, hearing, or touch) as a method of user interaction. Unimodal methods of gaze control input may include dwell time, blinks and winks, glances, gaze gestures (predefined, unique patterns of eye movements used in the design of user interface), and pupil dilation. For example, gaze pointing and selecting an object or target on a display could be implemented by employing dwell time, blinks and winks, etc. While eye gaze may not be as accurate an input device as, for example, a mouse, if the target objects are large enough, gaze can be much faster at pointing at targets than pointing at targets with a mouse (Sibert & Jacob, 2000). Interactive applications that make use of eye movements have traditionally been focused on command-and-control interactions. In gaze command-and-control interfaces, the user makes inputs with conscious, explicit eye movements. Examples of gaze control interfaces include cursor control, text entry, and item selection via dwelling, winking, blinking, or saccading. Explicit eye input can be beneficial in situations where (a) hands and/or feet are occupied with other tasks, (b) there is limited space for additional manual controls on the interface, or (c) hand or foot use is impaired (Majaranta & Bulling, 2014). However, applications that make more subtle use of eye movements have recently become increasingly common (Gehm, 2013; Hansen, Alapetite, MacKenzie, & Møllenbach, 2014; Reiser, 2015). The following sections provide an overview of some existing methods of gaze control input as well as specific applications of these methods, all in the context of a well-defined continuum of gaze control input user interactions (Majaranta & Bulling, 2014).

[1] An eye tracking technique where the basic concept is to use a light source to illuminate the eye causing highly visible reflections, and a camera to capture an image of the eye showing these reflections.

[2] A device using a single independent sensory input channel (modality) in human-computer interaction

Gaze Control Input: The Explicit–Implicit Continuum

Majaranta and Bulling (2014) underscored the potential of using gaze in human-computer interfaces either as a control input method or as an information source for intelligent user interfaces. Furthermore, they proposed a gaze control interface continuum with fully explicit (active) input on one end and implicit (passive), unintentional input on the other (Figure 1). For the purposes of this literature review the continuum serves as a conceptual framework, providing insight to the potential applications of different eye movements as control input devices.

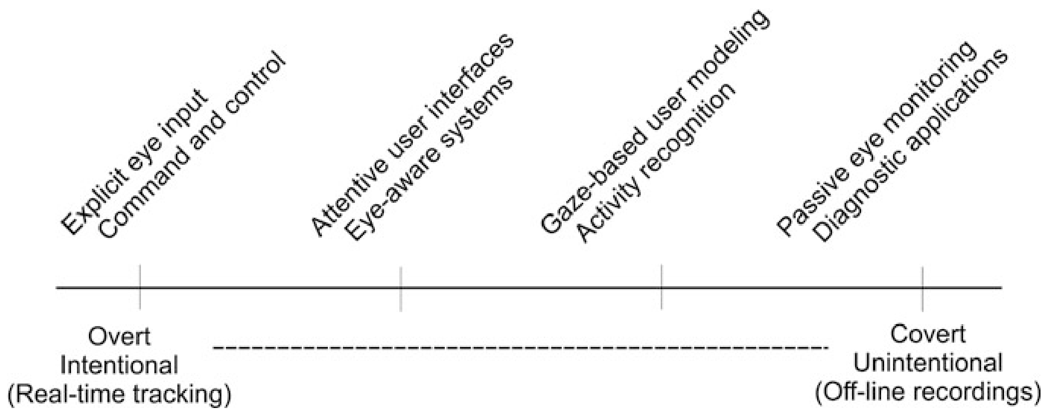

Figure 1. Explicit–Implicit Spectrum of Gaze Control (Majaranta & Bulling, 2014)

Explicit (Command and Control) Gaze Control Input.

Explicit gaze control inputs require the user to consciously control their eye movements for the purposes of inputting commands. Five main types of explicit gaze control inputs are covered in the following section and these include (a) fixation/dwell time, (b) blinks and winks, (c) quick glances, (d) glance-off-screen/snap clutch, and (e) pupil dilation. A variety of studies have been conducted that examine the performance benefits of such input methods, with evidence indicating that some explicit eye gaze input methods are more promising than others.

Fixation/dwell time gaze control methods.

Time-based dwell control occurs when the user selects an object on a screen by fixating eye gaze on it for a certain period of time. Ware and Mikaelian (1987) conducted two experiments to evaluate gaze control as a method for computer input and specifically for selecting an object on a computer display. The first experiment measured selection time as a function of distance to the target from the initial fixation point using three selection methods: (a) prolonged eye-fixation on the display only, (b) on-screen select button in combination with eye fixation on the display, and (c) physical button pressed by the user's hand in combination with eye fixation on the display. Results from the first experiment indicated that selection times were always less than one second, where selection by prolonged fixation combined with the physical button were equally fast, and faster than using prolonged fixation combined with an on-screen button for selection. The differences between the error rates in the different conditions were not statistically significant.

The second Ware and Mikaelian experiment measured how target size affected response speeds and error rates for physical button and dwell selection methods. The experiment initiated a random sequence of 64 trials, with each trial immediately self-initiating after the conclusion of the previous one. Five item-sized squares were presented for selection; the sides had approximate viewing angles of 3.0, 2.25, 1.5, 0.75, and 0.45 degrees of visual angle. Dwell selection time for the target was set to 400ms, which was slower than the time required to make a selection via physical button for all sizes. Item selection speed decreased when moving from the 1.5 degree-sized item to the 0.75 degree-sized item, and again from the 0.75 degree-sized item to the .45 degree-sized item. For all target sizes, there were fewer errors when using dwell-time selection than for physical button selection. Comparing the error rate and response time between dwell-time selection and physical button selection, there is a speed-accuracy tradeoff; dwell-time selection may be slower than physical button selection but results in a lower error rate.

Jacob (1993) examined the dwell time method for selection of objects on a screen. Two techniques were implemented simultaneously. First, if the user continued to look at the object for a sufficiently long time, it was selected without further operations. Second, an optional button press could be used to avoid waiting for the dwell time to expire. The idea was that the user could trade between speed and a free hand. The approach of using a 150-250 ms dwell time resulted in improved user perception of system performance. The lag between eye movement and system response (required to reach the dwell time) was not perceptible by the user. However, the lag was long enough to accumulate sufficient data for the system to recognize the fixation and process it. The subjective feedback collected during the experiments was of a highly responsive system, almost as though the system was executing the users' intentions before they expressed them. For situations where selecting an object was more difficult to undo, rather than a longer dwell time, a button confirmation was used. The research found no case where a dwell time longer than 750 ms was useful, possibly because people do not normally fixate one spot for that long, and when a dwell time longer than 750 ms was used, it created a perception that the system was unresponsive.

Blinks and winks gaze control methods.

Shaviv (1993) investigated the use of winks as a method of gaze control based on the premise that winking was a deliberate action. Specifically, when a user looked at an object on the screen the object became highlighted, but no action was performed until the user winked; then the command was executed. The author reported that the results from several experiments showed gaze controlled interactions such as winks were especially suitable for large screens and systems with multiple screens. Because the time it took to move gaze focus was nearly constant, the benefit of gaze control became greater when the distance was larger. The results of this study also suggested that the participants' feedback was generally positive to gaze control. However, there were very different opinions expressed about what kind of deliberate eye movements were most appropriate, and for what tasks. Shaviv (1993) asserted that further research was needed in order for the interaction methods to be properly developed. He further indicated that when developing new and novel user interfaces, the focus should be on multimodality, in general, and multimodal user interfaces and interaction methods, in particular.

Quick glances gaze control methods.

Ohno (1998) proposed using quick glances as a method of command input in a window, icon, menu, and pointer (WIMP) system to decrease the difference in time between natural eye movement and the gaze action in the context of how eye movements were interpreted as control inputs. Ohno was concerned that if eye movement had a specific function instead of just looking at the display, the operation time may become longer and users may not be able to maintain effective control. The concept was used in the Quick Glance Selection Method and examined in an experiment that involved two possible variants of target selection. The first was to glance at a command from a main menu (with sub menus) and then at the item that was to be selected. The second was to simply look at the item when the command was already active. It was expected that the second approach would eliminate the extended dwell time related to dwell time-based gaze control. The results showed that (a) the Quick Glance Selection Method was faster to control than using a mouse to click on the main menu and then again on the submenu when performing a selection task in a hierarchical menu, (b) making a selection by means of an eye focus point took longer than just reading menu commands, and (c) most errors, which consisted of wrong menu selection, unconscious glances at the selection mark, and wrong position of the eye mark, were induced by lack of adequate visual feedback. The author recommended methods for reducing selection error, improving performance, and usability that included providing an effective "undo" function, use of a gaze cursor, and active eye-tracking calibration. These recommendations were similar to the general recommendations for other input methods (Nielsen, 1994).

Glance-off-screen/snap clutch gaze control methods.

Istance, Bates, Hyrskykari, and Vickers (2008) identified three main categories of challenges associated with gaze control technology: eye-tracking usability, accuracy, and inadvertent activation (also known as the "Midas touch"). The authors suggested that the glance-off-screen/snap clutch method of gaze control for both active/explicit (i.e., gaze is used as a primary and command-based input modality) and passive/implicit (i.e., gaze-aware attentive interfaces) gaze control interfaces could overcome that problem. The notion behind the method was to provide a fast way of disengaging gaze control when it was not needed, and that a quick gaze gesture, such as a glance or a snap in one direction just past the edge of the screen, would "clutch" the system into treating visual dwells in a certain way, depending on the system mode enabled with the clutch. These system modes included "Eye Control Off," "Dwell Click," "Park it Here," and "Drag from Here." Visual dwells during the "Eye Control Off" mode had no effect. During the "Dwell Click" mode, dwelling replicated the function of a clicking a conventional mouse. "Park it Here" mode locked the pointer on location, and the resulting dwell caused the pointer to be "clicked." A second look at the pointer unlocked its position, allowing it to be moved and dropped to a different point on the screen. "Drag from Here" mode initiated a dragging action across the screen, such as when a file is dragged between folders on a conventional computer display, or when a change of view occurs in computer games.

Two expert level players who were familiar with the 3D virtual community environment participated in user trials and performed four tasks: (a) locomotion (navigate an avatar along a predefined path); (b) camera movement where the subjects were required to zoom in, rotate, and zoom out the camera around an avatar; (c) object manipulation (create a cube and scale it to a given size); and (d) application control (navigate through a series of menus and change the appearance of an avatar). The trials employed gaze as an emulation of a cursor control device, placing the cursor where the user was gazing onto the screen. This eye-mouse had four modes: "off," "click," "park," and "drag." The researchers collected performance times for each task and analyzed errors such as path deviation and activation of a control object other than intended. The results showed that the snap clutch device facilitated performance on two tasks: (a) locomotion while using the "Drag from Here" mode and (b) the camera rotation task using the "Park it Here" mode, which had been difficult to perform with standard dwell click emulation. The authors noted that further research was needed to determine the practical limit of using different modes to mitigate inadvertent activation.

Pupil dilation gaze control methods.

Ekman, Poikola, and Mäkäräinen (2008) investigated consciously controlled pupil size as an input modality. They hypothesized that given continuous feedback, users could learn to control pupil size via physical and psychological self-regulation, and they measured the magnitude of self-evoked pupil size changes. The participants had to follow seven different instructions. Real-time graphical feedback on pupil size was provided. Results indicated that some types of voluntary effort affect pupil size on a statistically significant level. Furthermore, the results showed that certain physical activity, positive emotions, and the switching focus between near and far points resulted in statistically significant pupil size changes, and thus demonstrated some potential for pupil dilation to be used as an input method for gaze control technology.

Applications of explicit gaze control input.

Pomplun, Ivanovic, Reingold, and Shen (2001) presented a gaze control interface that allowed users to magnify and examine any part of an image by looking at the area of interest on one screen and shifting their gaze to the magnified second screen. No manual input was required to control the interactions. The interface was empirically evaluated using comparative visual search tasks that required several steps of "zooming" in and out of a search display. Performance was assessed separately for input devices (using gaze control or using a mouse). The results showed that performance with the gaze control interface was comparable with a standard mouse input device and that using the gaze control interface could be learned very quickly. The authors asserted that appropriately designed gaze control interfaces could replace conventional interfaces in areas of application where hands-free interaction with machines was desirable and recommended further empirical research focused on developing reliable algorithms that can determine the users' intent from their eye movement patterns.

Ashmore, Duchowski, and Shoemaker (2005) evaluated how real-time analysis of users' eye movements could be used for refining existing pointing methods involving a fisheye lens.[3] Their approach involved hiding the fisheye lens during visual search and morphing the fisheye into view as soon as the user completed a saccadic eye movement[4] and began fixating on a target. This method allowed the user to maintain an overview of the desktop during search while selectively zooming in on the foveal region of interest. The results from comparing this interaction method with methods where the fisheye lens is either continuously in view or is not used at all showed performance benefits in terms of speed and accuracy. More specifically, the new method was the fastest for item selection.

Attentive (Eye-aware) Gaze Control Input.

Attentive gaze control input methods are a part of the explicit-implicit continuum and involve tracking eye movements for the purpose of changing background aspects of a system's function (e.g. pixel resolution). While they can be controlled using explicit eye movements, the intent of the input is to allow background changes without requiring overt manipulation. Duchowski, Cournia, and Murphy (2004) defined attentive user interfaces where the user was aware of the system's function, but eye movements caused only subtle background changes such as reducing pixel resolution for areas adjacent to where the user is looking, rather than execute commands. The authors suggested that such interfaces could be integrated into systems in order to improve systems' efficiency (e.g., reducing screen resolution of displays not being attended to in order to reduce computational demand). These systems could also include explicit eye input capabilities in addition to eye-aware interfaces; however, the attentive element may not require explicit commands on the user's part in order to execute its programmed function.

Gaze-based (Activity Recognition) User Modeling Input Methods

Gaze-based user modeling systems (Figure 1) track eye movements over time in an implicit fashion (the user is not aware that the system is operating and is not able to consciously control its function) in order to collect information on the user's visual behavior, cognition, and attention and quantify the similarity of eye movements. These models typically rely on computational methods from machine learning and pattern recognition, and the information is then used to refine the system's functionality using computational models (Majaranta & Bulling, 2014).

Jacob (1993, 1995) pointed out that eye movement-based interaction was an emerging style of user-computer interaction that had evolved toward more implicit (rather than explicit) commands and interactions that are simultaneous and parallel in nature. Rather than respond to explicit inputs, such interfaces could give the impression of responding to users' intentions. Instead of requiring the user to make specific eye movements to activate features and functions, the system uses information collected from users' natural eye movements while interacting with it. For example, the researchers experimented and later rejected long gazes as a method of interaction with a system because the gazes were not natural eye movements and used them only as long as naturally occurring fixations. The results showed that designing interactions around a philosophy that emphasized natural eye movements worked well with environments based on natural spatial navigational commands (e.g., virtual environments) that are inherently non-command-based.

Hansen, Engell-Nielsen, and Glenstrup (1998) used the phrase "interest and emotion sensitive media" to describe a non-command gaze interaction that aggregated ocular user behavioral data over time. They measured the interest of the observer by continuous identification of the areas of attention and affective reactions to various types of events revealed by measuring the blink rate and changes in pupil size.

[3] An ultra wide-angle lens

[4] A saccade is a quick, simultaneous movement of both eyes between two phases of fixation in the same direction.

Consequently, a model of the gaze data that revealed the user's interest would modify the further interactions either immediately following or with a delay to the users' behavior.

Horrey, Wickens, and Consalus (2006) examined how characteristics of a simulated traffic environment and in-vehicle tasks impact driver performance and visual scanning. The goal was to determine the extent to which a computational model of visual attention based on objective task characteristics could predict scanning behavior and the extent to which scanning patterns impacted different aspects of driving performance such as hazard monitoring and lane keeping. The results indicated that the model was a plausible and effective predictor of scanning in driving. The researchers suggested that when coupled with further validation, designers may be able to use this model to predict the allocation of visual attention in different highway conditions and thereby predict vulnerability to missing roadway hazards.

Passive (Diagnostic) Eye Monitoring Input Methods

Majaranta and Bulling (2014) defined passive long-term eye movement monitoring as a means for computing systems to better understand the user behavioral patterns outside laboratory settings. The user could be completely unaware of the system being functional and operating, and their eye movements would not influence any aspect of the system in real-time. As a diagnostic tool, eye monitoring provides objective and quantitative evidence of the user's visual and attentional processes where eye movements are generally recorded to establish the user's attentional patterns over a given stimulus. The recorded eye movements would be used for post-trial, off-line assessment of the participants' gaze in order to further develop and refine particular applications (Duchowski, 2002).

Summary

Using eye gaze in user interfaces either as an explicit input method or as a more implicit "behind the scenes" information source is feasible and has potential. Eye movements are largely unconscious and automatic; however, when needed, people can control gaze at will, making deliberate eye control possible. Additionally, employing gaze as a sole control input method could be challenging. Gaze is easily distracted by movement in the peripheral vision, resulting in unwanted glances away from the object of interest. Therefore, the interface system needs to be able to distinguish casual viewing from intentional control in order to prevent the inadvertent activation problem (Jacob, 1991). Furthermore, using gaze as an input control device may be (a) effortful, because many eye movements are mostly subconscious, thus requiring effort to bring them under conscious control; and (b) difficult, because the eyes are being used for both input and output, thus requiring careful coordination of reading for input by putting eye selection output aside. It has become increasingly evident that the difficulties encountered in the use and interpretation of individual modalities control inputs such as gaze control may be overcome by integrating them into a multimodal user interface (Jacob, 1993; Zhai, Morimoto & Ihde, 1999; Oviatt & Cohen, 2000; Yamato, Monden, Matsumoto, Inoue & Torii, 2000; Song, Cho, Baek, Lee, & Bang, 2014).

MULTIMODAL CONTROL INPUT METHODS INCLUDING GAZE CONTROL

The drawbacks associated with eye gaze described above make it challenging to implement as a sole means of input control. However, when considering the current maturity level of gaze control technology, there is a potential for it to benefit the field of human-computer interaction in conjunction with other technological advancements such as multimodal interfaces, artificial intelligence, etc. Multimodal control input interfaces that include gaze as one of the input control modalities are discussed in the sections that follow.

Gaze and Traditional Physical Input Devices

Jacob (1993) investigated eye movement-based interfaces, observing that short dwell (i.e., 150-250 ms) for simple object selection, a physical button/short dwell combination for complex selection, and smooth eye movements for object movement were effective methods of implementing natural eye movements and did not require the user to make unnaturally long dwells. Jacob (1993) concluded that non-command, eye movement-based interfaces could improve human-computer interaction by transforming explicit, serial interactions to implicit, parallel interactions. Jacob (1993) also discussed non-command, eye movement-based interfaces, related human factors considerations, preliminary investigations into successful implementation, as well as future applications in computer interfaces. He further emphasized that eye movements may be an ideal modality for implicit input control interfaces due to the low cognitive load and mostly subconscious operations.

Yamato, et al. (2000) proposed two techniques to reduce errors and increase efficiency of gaze control systems. Both techniques used the same graphical user interface, which combined eye and traditional input devices into a multimodal interface. The basic notion of the approach was to use the eye for moving a cursor and to use the hand for identifying a graphical user interface (GUI) button to be selected. To improve the accuracy and efficiency of cursor movement via eye gaze, the proposed interface provided two types of cursor adjustment techniques: automatic and manual. The first technique was a combination technique, which used the eye to move the cursor to the GUI button, and then the hand was used to activate that button. The second technique was an adjustment technique. This technique included two different methods of adjustment: automatic and manual. Both adjustment methods used the eye and hand similarly to the combination technique, except for the cursor movement behavior near the activation button. With the automatic adjustment method the cursor was automatically moved to the closest GUI button when the mouse button was pushed, while the manual adjustment method allowed the user to move the cursor to the button area with the eye, and then the cursor was moved onto the GUI button manually with the mouse. Button sizes varied between 1, 2, 3, 4, and 5 $cm,^2$ and spaces between buttons were either 1 or 5 cm. There were very few misses for both the automatic and manual adjustment techniques when buttons were separated by 5cm. However when the buttons were separated by 1cm and the automatic adjustment technique was used, many misses were reported. In the 1cm separation condition when the manual adjustment technique was used, almost no misses were reported. When measuring completion times using 5 cm^2 buttons, both the automatic and manual techniques were faster than when using a mouse. However, when using buttons under 4 $cm,^2$ the automatic adjustment method was slower than using a mouse and nearly impossible to use with 1 cm^2 buttons. The manual adjustment technique worked well no matter what the button size.

Zhai, Morimoto, and Ihde (1999) proposed a multimodal technique called Manual And Gaze Input Cascaded (MAGIC) in order to avoid what was considered unnatural overload of a visual perceptual channel with a motor control task (such as moving a cursor or selecting an icon on a computer screen). The general concept behind MAGIC pointing techniques was to use gaze to move a cursor to the general area of a target. The user would then manually move the cursor the last, small distance to the target and initiate the command to activate the target as desired. Two different MAGIC methods of multimodal control techniques were tested where participants were asked to point at targets as they appeared in a random order, and the results showed that the methods were usable and required little training to achieve proficiency.

Gaze and Touchless Gestures

Song, et al. (2014) developed a multimodal control interface called GaFinC (Gaze and Finger Control interface) for use in computer-aided design applications. The interface used gaze tracking as a pointing mechanism, and touchless gestures for control inputs. GaFinC could track precise hand positions, recognize several finger gestures, and utilize an independent gaze-pointing interface for locating the point of interest. The participants performed three types of tasks: (a) translation task (one hand pinch gesture and movement), (b) rotation task (two hands pinch rotation gesture), and (c) zooming task (two hands pinch zoom gesture). Tests of manipulation accuracy and time were conducted, and the results were compared with those of a conventional mouse. The comfort and intuitiveness level were also scored by means of user interviews. The results suggested that although the GaFinC interface displayed inadequate performance in terms of accuracy and times as compared to a mouse, users found it to be more intuitive than a mouse interface while maintaining a usable level of comfort. The results from objective (as opposed to the subjective) measures provide more prudent justification for design implementation, especially when the subjective perception about the intuitiveness of the software interface did not support the objective measures of performance.

Intelligent Gaze

Salvucci and Anderson (2000) examined a multimodal application of gaze-control technology, which they called an "intelligent gaze-added" interface. The idea was to use gaze-control to augment an existing system rather than replace already existing functions. Incorporating "intelligence" into this system meant using a probabilistic algorithm and user model to interpret gaze focus with eye-tracking data to anticipate the user's intentions to alleviate typical problems (e.g., poor cursor accuracy and unintentional icon selection). The final product was based on a standard "window, icon, menu, and pointer" interface. Users were reportedly able to quickly adapt to the gaze-added interface and use the functions along with other inputs, although no data was provided to support this observation. The results also showed that there were times when the user's intentions were ambiguous. To reduce the ambiguity, inputs from other modes such as pupil dilation, models of interest, and emotions were suggested.

Gaze and Electromyography (EMG)

Agustin, Mateo, Hansen, and Villanueva (2009) evaluated the potential of gaze control input for target acquisition and target tracking tasks. The researchers compared the target acquisition and target tracking performance of two eye tracker systems (i.e., Tobii, Quick Glance 3) with four other input devices (i.e., mouse, touch screen, head tracker, and joystick) with two experiments. Target acquisition performance was compared using either gaze or mouse for pointing, and either a mouse button or an EMG switch for system activation. The results from Experiment 1 indicated that both eye-tracking systems elicited superior throughput compared to the head tracker and joystick and comparable throughput to the mouse and touch screen, but only for large-diameter (i.e., 150-pixel) targets. Additionally, with the large-diameter targets, task completion time was marginally better with the eye-tracking systems than with all other input devices. However, time on target was best for small diameter (i.e., 75-pixel) targets, especially for eye-tracking systems—this measure did not differ between input devices for large-diameter targets. In Experiment 2, it was observed that the gaze interface offered significantly increased mean throughput over a joystick and that gaze input in combination with both the button and EMG elicited reduced task completion time by 21 ms and elicited comparable error rate with the mouse combined with a button or EMG. These observations suggested that, given a sufficiently accurate and responsive eye tracker combined with a well-designed interface, using gaze in target acquisition and tracking tasks would be beneficial, especially with larger targets.

Chin, Barreto, Cremades, and Adjouadi (2008) compared a multimodal control system (eye gaze and EMG) to eye-gaze control only and mouse control interfaces. The multimodal system inputs consisted of electromyogram (EMG) signals from facial muscles and point-of-gaze coordinates produced by an eye-gaze tracking system. The target end-users for this system were individuals who were unable to use their hands because of spinal dysfunction or other conditions. Results of two experiments indicated that although the multimodal control system was slower than eye gaze *only* and mouse control, it effectively controlled the cursor without spatial accuracy limitations and facilitated reliable click operations. Mean errors/total session trials for comparing EMG and eye-gaze control technology and eye-gaze control technology only were 0.017 and 0.396, respectively. Comparisons between task times comparing EMG/eye gaze control, eye gaze control only, and traditional mouse use have mean trial times of 4.7 s for EMG/eye gaze control, and 3.1 for eye gaze control only. Finally, mean error rates for the three techniques were 0.01 errors/trial for the mouse, 0.14 for the EMG/eye gaze control system, and 3.98 for the eye gaze control only system.

Gaze and Voice

Wilcox, Evans, Pearce, Pollard, and Sundstedt (2008) conducted a focus group to investigate the potential for a multimodal interface including gaze and voice controls and to identify potential issues from this combination of modalities. The authors developed a computer game that incorporated eye tracking and voice commands, and the result was a hands-free system that enabled the vocal commands to be used as an alternative when eye control failed. The system also allowed blinking or winking to command an in-game catapult, and to fire on an object located in the user's crosshairs. They found that when designing a point-and-click-style game for use with an eye tracker, the aesthetics became a performance issue because the scenes and the selectable elements of the layout needed to be scaled for optimal usage. Also, the button and target sizes had to be larger, and target positions had to be carefully crafted in order to produce an effective and efficient system.

Glenn et al. (1986) developed an Ocular Attention-Sensing Interface System (OASIS), where eye movements and voice commands were used as control input devices. The system consisted of an oculometer, a voice recognition/interpretation mechanism, and a central processor. OASIS was examined in a study to research visual attention based on eye position data and the effects of providing the operator with informative, non-distracting feedback. No traditional physical controls were utilized. The results suggested an opportunity for flexibility in design of applications as the feedback cursor appeared to neither enhance nor degrade performance in the experimental tasks in general. The researchers concluded that a feedback cursor could be employed if the application warranted it or in case more flexibility was desired for the users. Overall, the authors reported that OASIS was an interface with potentially broad application potential.

Gaze and Head Tracking

Borah (1995) investigated the feasibility of using eye point-of-gaze (POG) and head control (HC) of a cursor in lieu of, or to supplement, manual control for cursor positioning tasks. The main concern was the issue of positioning a cursor with respect to targets that may need to be "small" (regarding eye tracker accuracy and precision). Therefore, a second control modality (other than eye movement) was included for closed loop error correction. Another objective was to investigate the applicability of Fitts' law[5] and related movement time prediction models to eye control. Several techniques were subjectively evaluated. This allowed users to "close the loop" and fine tune point-of-gaze controlled cursor positions. A technique that allowed the user to switch from point-of-gaze control to a low gain head position control when near the target was deemed very effective during a pilot study conducted before the formal evaluation. This technique was used during the formal experiments in addition to "POG control only," "head control (HC) only," "POG with a user controlled switch to head control for fine positioning," and "standard mouse control only" conditions using two serial, cursor positioning tasks. One of the tasks included a search component and was designed to represent a realistic interaction task. The other did not have a search component and was designed to facilitate analysis of motion time with Fitts' law and related models. The results showed that "mouse only" control produced the fastest median task completion times across the entire range of target sizes and distances. POG control only had significantly faster median times when the accuracy and precision limitations of the eye tracker were not exceeded, but these limits restricted the technique to relatively large target sizes. When targets were smaller than eye tracker performance limits, POG control becomes impossible. Both "HC only" and "POG with head controlled fine tuning" proved viable. It was possible to reliably select targets across the entire range of target sizes and motion distances tested. Both conditions produced significantly slower median motion times than "mouse only" control or "POG control only" with large targets. The researchers concluded that "POG with head controlled fine tuning" control paradigm was viable, but based on task completion time measures it did not show an advantage over the "HC only" condition data.

[5] Fitts' Law is a model for predicting movement times calculated by the following formula: $MT = a + b * ID$. *MT = Movement Time, a* & *b* are constants, and *ID = the index of difficulty,* where $ID = log_2(2D/W)$. *D* = Distance to target and *W* = target width.

Gaze and Touchless Gesture and Bio-signals

Heo, Lee, Park, Kim, and Whang (2010) generated a multimodal user interface gaming system that combined a head-mounted display (HMD), a data glove, and a bio-signal analyzer that measured signals including photoplethysmogram, galvanic skin response, and skin temperature. The HMD contained a gaze tracker that was used as an aiming cursor. The data glove was used to register grabbing and throwing gestures. The level of difficulty in the game was adaptively controlled according to the user's bio-signals. The experimental results showed that for gaming applications, the multimodal interface design improved the experience of immersion and interest more than conventional devices such as a keyboard and/or a mouse.

AVIATION-SPECIFIC EYE TRACKING AND GAZE CONTROL LITERATURE REVIEW

The notion of using eye tracking for detecting pilots' instrument scanning patterns and eye gaze as an alternative or nonconventional control device on a flight deck is more than five decades old. The following review of the pertinent literature highlights the nonlinear nature of the evolution and maturation of both eye-tracking technology and its methods of application as a control device. The section is divided into two subsections covering the aviation-specific literature on eye tracking and gaze control, respectively. In the context of the explicit-implicit continuum (Majaranta & Bulling, 2014), the studies included in this section fall toward the implicit end of the continuum. Furthermore, the reported results provide valuable knowledge and insight into pilot visual behaviors and performance. Such knowledge and insight are essential to the research and development of nonconventional control input devices that include eye gaze as either a single control input modality (unimodal) or as part of a multimodal control interface for use in the domain of aviation.

Aviation-Specific Eye-Tracking Literature Review

Fitts, Jones, and Milton (1950) addressed the importance of knowing how pilots use their eyes while flying "on instruments" and pointed out that collecting such information was essential to a basic understanding of the functions aircraft instruments serve and the human factors associated with controlling the aircraft. The authors suggested that it was reasonable to assume that frequency of eye fixations on any given instrument is an indication of the relative importance of that instrument. In contrast, the length of fixations may be considered as an indication of the relative difficulty of interpreting these instruments. The authors considered the patterns of eye movements to be a direct indication of the goodness of different panel arrangements. They concluded that eye movements vary from pilot to pilot, from maneuver to maneuver, and from instrument to instrument, and that it was possible that similar differences would be found for different aircraft, for different panel arrangements, for day vs. night flying, and for different designs of the same instrument. The findings reported in the paper included

- An average of 106 eye fixations per minute were made during instrument landing system approaches (0.59 seconds per fixation) and an average of 93 eye fixations per minute during ground control approaches (0.67 seconds per fixation).
- Different instruments were checked with very different frequencies. These differences were greater than 10%.
- Different instruments required very different lengths of time to check. These differences were greater than 50%.
- There was noticeable variability between pilots in eye-movement characteristics.
- There was a tendency for more experienced pilots to make shorter fixations.

While the study was not aimed at examining gaze control technology per se, the findings advanced our knowledge about pilot scanning patterns, which offers a valuable insight with respect to control input interface arrangement and user modeling.

Karsten, Goldberg, Rood, and Sulzer (1975) measured the visual fixations of seven air traffic controllers using an oculometer (head mounted camera with a built in eye tracker). The oculometer data were compared to manual recordings made by a human observer at a controlled workstation. Six controllers performed simulated radar control functions that included monitoring the radar display, updating flight information marked on flight progress strips, and making simulated handoffs to adjacent sectors and acknowledging handoffs from other sectors using a keyboard and slew ball. A seventh control task consisted of a 15-minute radar problem. The researchers determined that an oculometer produced an accurate assessment of visual performance that provided advantages over the conventional manual method due to its ability to detect brief glances compared to a human observer who could not.

Ottati, Hickox, and Richter (1999) used eye movements to evaluate the usability of specific instruments such as newly developed electronic moving maps. The researchers compared eye movement patterns on different terrain features between experienced and novice pilots during a visual flight rules (VFR) simulation. The authors hypothesized that experienced pilots would spend less time finding and fixating on their navigational landmarks, while novices would have greater dwell times on the map due to difficulty in finding landmarks and extracting useful data from them. As anticipated, novice pilots tended to devote more visual attention outside the cockpit than experienced pilots.

Graeber and Andre (1999) used eye tracking to study how pilots interacted with electronic moving maps for taxiing, and suggested that specific training was necessary to assure proper usage of and optimal visual attention interaction with electronic moving maps. Results indicate that atmospheric visibility significantly affected the amount of time pilots dwelled on the moving maps, with dwell time significantly higher under high-visibility conditions. Specifically, as visibility degraded, pilots spent more time looking "out the window" and less time dwelling on the moving maps with no detriment in taxi performance. This result is not surprising because pilots are required and trained to be heads-up and eyes-out scanning for hazards and maintaining information gathering out the window.

Anders (2001) used eye-tracking technology in a flight simulator where eye and head movements of professional pilots were recorded under realistic flight conditions in an investigation of human-computer interactions relevant to information management, as well as mode awareness in a modern glass cockpit. The analysis of eye movements demonstrated the importance of the primary flight display as the primary source of information during flight. The study was a proof of concept that identified a potential use of eye movement data in assessing pilot performance and identifying specific training topics for novice pilots.

Kasarskis, Stehwien, Hickox, Aretz, and Wickens (2001) compared expert and novice scan behaviors during VFR flight simulations and concluded that experts had significantly shorter dwells, more total fixations, and more target and airspeed fixations as well as fewer altimeter fixations than novices. In addition, expert pilots revealed a stronger and more defined scan pattern than novice pilots. This type of active scan pattern corresponded to better maintenance of airspeed and better landing performance.

Bjorklund, Alfredson, and Dekker (2006) evaluated pilot behaviors on simulated commercial flights in terms of eye point of gaze, mode awareness, call-outs, and tracking aircraft performance. The researchers found that aircrews often deviated from call-out and mode-monitoring procedures as defined by the air carrier and manufacturer. Additionally, flight crews were using numerous strategies to monitor the status of the aircraft utilizing resources other than the flight displays and annunciations. However, these diverse strategies did not appear to negatively influence the safety of the flight and control of the aircraft. The data identified the limitations of current flight mode annunciator designs and suggested that mode awareness is a more complex phenomenon than what can be captured by measuring eye point of gaze and communication alone.

Chuang, Nieuwenhuizen, and Bulthoff (2013) conducted a study addressing the relationship between flight control performance and instrument scanning behavior. The study was conducted on a fixed-based flight simulator and targeted the ability of untrained novices to pilot a lightweight rotorcraft in a flight scenario that consisted of mission task elements such as speed and altitude changes. The results showed that novices generally prioritized the airspeed indicator followed by the altitude indicator, suggesting that novice pilots intuitively react to changes in their airspeed and altitude and compensate accordingly. The authors recommended that training practices focus on instructing novices to behave in a more anticipatory and less reactive fashion in order to achieve better control performance and optimized workload. Furthermore, they made some suggestions regarding the spatial layout of visual instruments on a flight deck. First, the layout should be designed to encourage optimal scanning strategies, including facilitating selective attention to relevant information by reducing the clutter of non-relevant instruments. Second, the emphasis should be on the importance of a primary instrument by placing it in a central position. This could encourage users to return to it after referencing other instruments, rather than treating every instrument equally.

Aviation-Specific Gaze Control Literature Review

Flight deck visual tasks include continuous tracking, searching for a target, and visual monitoring of displays, among others. The first study of explicit gaze control on the flight deck was conducted by Calhoun, Arbak, and Boff (1984), who evaluated the feasibility of a dwell-time application in which the pilot's line of sight (gaze) was directed to a switch on a display. Two trained participants completed 252 switching actions in a 45 min session each. The participant's tasks involved (a) tracking a visual target with a force stick mounted on the right console and (b) selecting switches on the front panel of a cockpit simulator. In the switch selection task, an auditory cue was used to direct participants' gaze. The switch was illuminated by either a color or brightness change to provide positive visual feedback to the pilot when the system detected that the participant's gaze was directed at a switch. The participants then manually closed a consent switch located on the joystick or provided a verbal response into a speech system. Each switch was illuminated from the time it was se1ected by the line of sight (LOS) until either the consent switch was closed and another switch was selected, or a 5 s timeout interval expired. The eye-contro1 algorithm parameters for the sessions were such that the eye LOS was required to be within 2 cm of the center of the switch for 3 out of 5 consecutive samples. The eye-controlled switching response times were analyzed in three phases: (a) the point at which the gaze 1eft the tracking task on the upper monitor, (b) the point at which the switch was selected by the participant's gaze (i.e., illuminated), and (c) the point at which the consent switch was closed. The results suggested that eye-controlled switching was a feasible alternative to manual switching when a hands-off control mechanism was desirable. In addition to a switch selection, the researchers proposed that gaze control could be a natural interface for designating or updating symbology or data on a display. A suggested long-term goal for future research was to evaluate the effect of each type of concurrent task on task performance with eye-controlled switching and other control interfaces. The authors concluded that a better assessment of the utility of gaze control would continue to be complex until substantial experience was gained through actual use of the gaze control systems. They further pointed out that challenges notwithstanding, the benefits from such research cou1d substantially increase the efficiency of a large number of the pilot's tasks and suggested that gaze control benefits may not be immediately obvious when considering the task-saturated conditions on a flight deck, especially when many pilot tasks require visual attention. In summary, Calhoun, Arbak, and Boff (1984) reiterated the notion that if eye-controlled switching was to be seriously considered for application in future flight decks, then its efficiency relative to other control mechanisms needed to be quantified in visual workload environments analogous to that of single-pilot cockpits.

There have been several proof-of-concept demonstrations of gaze control in the aviation domain. The goal of these demonstrations has been to show that gaze control is feasible, but *not to evaluate* the use of gaze control. Tong and Fisher (1984) designed a proof-of-concept flight simulator that attempted to leverage the notion of an eye-slaved "area of interest" display. The system identified the area of interest and then provided a high-resolution image in that area, while the background within the rest of the user's field of view was presented in a lower-resolution. The project was mostly focused on system development and improvement rather than on conducting psychophysics experiments. Nonetheless, the idea of a high-resolution inset image within the field of view that takes advantage of pilot line-of-sight slaved displays for training was a significant technological development at that time.

Calhoun, Johnson, and Arbak (1986) conducted a gaze control experiment where six pilots were asked to select discrete switches on the front panel of a flight simulator while manually tracking a target. In two of the gaze-controlled methods, the pilots directed their gaze at the switch designated by an auditory cue and then made either a manual response or a verbal response input. In the conventional manual control condition, the pilots selected the switches with their left hand. The results showed that gaze control was a feasible alternative when hands-off control was needed. Tracking performance was found to differ significantly among switching conditions, indicating the importance of quantifying the efficiency of potential candidate control methods in visual workload environments equivalent to that of the application environment.

Hatfield, Jenkins, and Jennings (1995) proposed the Eye/Voice Mission Planning Interface (EVMPI), which integrated voice recognition and eye tracking into aircraft displays. The goal of this conceptual paper was to present general guidelines and principles for integrating eye-gaze and voice input in a flight deck interaction design. More specifically, system design would (a) allow the pilots to look at items on the aircraft display and speak commands to the system and (b) reduce the pilot's cognitive and manual workload. The EVMPI interface, although noted as being imperfect by the authors, yielded measureable performance gains over manual-only input in a preliminary task analysis. Additionally, eye/voice interaction significantly reduced the number of operations required of the user to complete tasks. The authors concluded that such an interface would benefit users of systems that require a heavy amount of cognitive and manual workload.

Borah (1998) reviewed the potential use of gaze control as a means of human-systems interaction. Gaze control was thought to be beneficial because it exploited the naturalness, speed, and accuracy of visual fixation. In 1998, gaze control technology had not reached a level of practicality that made it viable for aviation applications; however, it showed promise for future implementation. The gaze tracking technology included electro-oculography (measurement of change in electrical charge of the retina as it changes position relative to the skin), scleral coil (attaching a magnetic coil to the sclera of the eye and measuring change in the magnetic field via external coils), and optical eye tracking (photo sensors to measure changes in the eye's visual properties with movement) techniques. Borah suggested that the optical eye-tracking systems had the best chance of being used in operational flight. He concluded by stating the importance of assessing tradeoffs between performance requirements, simplicity, and ease of use.

Leger (1998) identified a need for alternative control input technologies (e.g., voice, gaze, gesture controls) as these improvements would allow the restructuring of the interface such that features are integrated into flight deck without introducing an information overload while the need to keep the attention of the flight crew focused on flying remains high. He further highlighted multimodal integration of two or more of these novel technologies as a method that would allow the user to operate the system under natural logic (as opposed to system-imposed logic). Implementing modality redundancy and complementarity would reduce the risk of error due to short-term memory failure and would also simplify the flight deck layout. This would offer more space for displays and important information. Finally, the author concluded that capitalizing on multimodal interactions with aircraft systems may require less training than complex conventional physical input devices.

Rood (1998) introduced some general considerations related to the integration of alternative control technologies into flight decks from both a human factors and an engineering perspective. The author purported that from a human factors perspective, alternative control technologies (such as gaze control) could help system designers avoid a bottleneck between human input and system capability, especially if the use of multiple modalities is considered. He also highlighted the importance of task modeling, prototyping, context of task, and task loading, error modeling, and analysis considerations in the design of such interfaces.

Merchant and Schnell (2000) reported on the development of a simulator that combined voice and gaze control. This combined approach was selected in order to overcome some of the limitations of gaze control and voice control when either method was used alone. Specifically, based on the literature review conducted by the researchers, when used as sole means of control input, both gaze and voice had been associated with certain limitations in accuracy, efficiency, and effectiveness particularly during high workload events. Combining the two modalities was deemed by the researchers as the most appropriate way to alleviate those limitations. Merchant and Schnell selected the pupil-corneal reflection[6] vector difference method of eye tracking for implementation in a Boeing 777-300 simulator. This alternative activation of controls was proposed for use in high workload situations such as those encountered during takeoffs and approaches in instrument meteorological conditions. Of major concern to the authors were data entry errors, since these data entry can account for as much as 2-3% of inputs during high workload phases of flight. The approach suggested by this team was to use eye gaze to select a control and then activate the control using eye gaze dwell time (implicit control) or explicit control using buttons or voice control.

In 2013, preliminary investigations into the viability of gaze control for air traffic controllers were promising. Using an air traffic control simulator, Alonso et al. (2013) compared a dwell time-based gaze control system to a manual key selection-based system, with and without feedback as to where the cursor was located on the screen, with varying degrees of difficulty (i.e., target size and separation). The gaze control system incorporating a dwell time-based selection system with continuous cursor location feedback produced the highest success rate (>90%). Based on these findings, a gaze control system such as this was determined to be potentially viable for air traffic controllers. However, more research was needed to determine how such a system compares in speed and reliability to current manual input systems.

Hansen et al. (2014) conducted an experimental investigation of gaze-based control modes for unmanned aerial systems. Ten participants (experienced gamers) performed three flying tasks, namely change in altitude, rotation, and target acquisition. The empirical measures included task completion time and examined the user experience for difficulty and reliability, as well as how enjoyable was the experience. Specifically, task completion time and keystrokes were logged from the onset of user control to the activation of a landing. The goal was to investigate the possible mapping between pilot tasks and gaze input control in the direction of x and y axes. More specifically, how control of speed, rotation, translation, and altitude could be best assigned to gaze in either *x*- or *y*- axis was considered. Two potential options for *x*-axis gaze control (lateral movement, rotation) and two for *y*- axis gaze control (longitudinal motion, altitude control) were proposed as follows:

M1–Translation and altitude by gaze; rotation and speed by keyboard;
M2–Rotation and altitude by gaze; translation and speed by keyboard;
M3–Translation and speed by gaze; rotation and altitude by keyboard; and
M4–Rotation and speed by gaze; translation and altitude by keyboard.

According to the researchers, these control models were best suited to unmanned aerial vehicles (UAVs) with an automatic control of attitude (i.e., position in the air) and would not work well for airplanes with a different set of dynamics. "Rotation and speed by gaze; translation and altitude by keyboard" mode was deemed significantly more reliable that the other modes. The comparison (based on subjective questionnaires) of mouse and gaze in terms of reliability and ease of use were significantly in favor of the mouse. These results were attributed to the natural mapping from gaze movements to rotations, similar to what would be the consequences of a lateral turning of the eye. It was also the control mode most similar to the one gamers use in three-dimensional games, where the mouse was commonly used to turn the viewpoint. The main observation made by the researchers was that that the gamers could actually control the drone by gaze, independently of control mode, and with only a small amount of practice (indicated by the low error rate - 2 crashes out of 40 trials).

Lastly, Thomas, Biswas, and Langdon (2015) created a multimodal experimental setup and tested different types of cursor control devices. They tested (a) a hands-on throttle and stick pointer, (b) an eye

[6] An eye tracking technique where the basic concept is to use a light source to illuminate the eye causing highly visible reflections, and a camera to capture an image of the eye showing these reflections.

tracker (no head movement allowed), (c) a head tracker allowing both eye and head movements, and (d) a hand tracker. They compared the task completion times when having the participant complete a point-and-select task on a multi-function display while monitoring the aircraft's changing heading on a HUD. The results indicated that participants using the head tracker performed the best, being significantly better than a hands-on throttle and stick. However, participants' initial cursor movements when using the eye tracker were made faster than other methods, but accuracy and drift issues caused longer completion times. Participants experienced the same problem as those encountered by de Reus, Zon, and Ouwerkerk (2012), which include loss of calibration due to helmet shift. The authors suggested that coupled eye and head tracking integrated into head-mounted displays solutions could not only provide pilots with a significant improvement in off-boresighting (looking away, either horizontally or vertically, from the direction the boresight/nose is pointing) but could also afford an effective interface with non-traditional flight deck controls in a wholly heads-up cockpit.

Summary

A range of eye-tracking, gaze-only, and multimodal control input methods that include gaze as an input modality has been researched over a broad spectrum of aviation applications. The most promising methods, in the context of potential implementation, are those that take advantage of the mutual disambiguation between modalities and mimic the more intuitive and natural collaboration in human-to-human interactions. The aviation research findings have identified a couple of factors essential to the successful implementation of gaze as a control input modality (Kratchounova, 2016). First, gaze as a sole means of control input device may be challenging; therefore, including eye gaze as one of multiple control input modalities is an approach that is clearly highlighted by research findings in the aviation domain. Second, if and when considered as one of the control modalities within a multimodal control input interface, gaze would be most efficient and effective as a supporting modality that is implemented as a more implicit (Majaranta & Bulling, 2014) control input device. The next section addresses the major gaze control human factors considerations.

KEY GAZE CONTROL HUMAN FACTORS CHALLENGES

At its current maturity level, gaze control technology affords a number of advanced research, development, and design tools as well as a number of technical and human factors challenges. Based on the literature review, the following subsections highlight the three most important challenges associated with utilizing gaze control technology,

Inadvertent Activation/Midas Touch

Inadvertent activation or Midas touch is the most commonly identified concern with gaze control technology. It occurs when gaze technology treats an unintentional extended fixation at an object on the screen as a deliberate one, and a feature on the screen activates. Several methods of minimizing inadvertent activation have been developed over the years and in the context of gaze technology applications. Tanriverdi and Jacob (2000) addressed it by generating an interest-based item selection method for reducing the sensitivity of the input apparatus to eye movements. Furthermore, Majaranta, Bates, and Donegan, (2009) suggested that combining gaze control with an additional input modality would be a reliable method of addressing inadvertent activations. For example, confirming a gaze-input command through a manual button, hand gesture, or an eye wink could reduce the likelihood of a false positive occurring. They also identified feedback as an important aspect of minimizing inadvertent activations. More specifically, in dwell time-based systems, a progress bar/pie indication the progression of the dwell time was shown to be successful in reducing this undesired effect. The researchers further suggested several methods of inadvertent activation mitigation such as fixation algorithms, context-aware fixation detection, drift correction, and smoothing/filtering. In summary, addressing inadvertent activations early in the research, design, and development process, as well as integrating gaze into a multimodal control input interface would both be essential strategies to successfully minimizing the potential for it to occur.

Accuracy and Repeatability of Eye Movements

There were several approaches proposed over the years to mitigate the natural physiological limitations to the accuracy and repeatability of eye movements (McMillan, Eggleston, & Anderson, 1997; Van der Lans, Wedel, & Pieters, 2011). These approaches include using (a) multiple modalities (i.e., eye gaze in combination with manual input, Yamato et al., 2000), (b) increased object/target size (Istance et al., 2010), (c) two- and three-part gestures (Ashmore et al., 2005), and (d) magnification/zoom (Skovsgaard, Mateo, & Hansen, 2011). While it is not possible to compare these four approaches across studies, it is apparent that they are all effective at mitigating performance issues while using eye gaze input. Specifically, each approach reduced response time, error rate, and/or increased accuracy in the task implemented in each study.

Level of Expertise

While level of expertise is not a factor specific only to eye gaze as a control input device, the implications associated with it are especially important due to the nature of eye movements and their association with visual attention and visual control. Law, Atkins, Kirkpatrick, and Lomax (2004) examined eye movements of expert surgeons and novices in a virtual aiming task to identify any important differences in eye movement patterns that could be used in training. The results from eye gaze analysis showed that novices required more visual feedback of the surgical tool position to complete the task than did experts. Furthermore, the experts were likely to maintain eye gaze on the target while manipulating the tool, whereas novices were more varied in their behaviors.

Jarodzka, Scheiter, Gerjets, and van Gog (2010) examined expertise differences in perceiving and interpreting complex, dynamic visual stimuli on a performance and on a process level. Performance, eye movement, and verbal report data were obtained from experts and novices. Results showed that experts, as compared to novices, attend more to relevant aspects of the stimulus, use more diverse task approaches, and use knowledge-based shortcuts.

Based on these findings, it can be concluded that gaze control user interface research, development, and design—especially in the context of complex and rich in dynamic visual stimuli environments—needs to take into consideration all levels of expertise among pilots.

Summary

The successful implementation of gaze input control devices into a unimodal or multimodal user interface is contingent on a thorough examination of the impact of using gaze control technology on human performance, fatigue, and workload in all operations conditions and for each specific user's task. The final decision about what interface type (unimodal or multimodal) or where on the explicit-implicit spectrum it should be placed needs to be driven by the prudent implementation of the results of such examination.

DISCUSSION AND CONCLUSION

When people communicate with each other, they speak, gesture, shift eye gaze, and move in an interaction flow that affords little or no resemblance to the discrete keyboard strokes and mouse clicks entered sequentially when using graphical user interfaces. Recent advances in technology maturation have prompted a profound shift in the design of user interface philosophy. Rather than imposing restrictive and unnatural user behaviors that stem from technology or system limitations, designers are embracing users' natural behavior, ultimately arriving at increasingly sophisticated, multimodal interface designs (Oviatt & Cohen, 2000). For a successful shift toward rebalancing the capabilities and limitations of the input and output modalities in user interfaces, it is essential to take advantage of a) the naturally occurring mutual disambiguation between modalities, and b) modality combinations that reduce the uncertainty in interpreting user intent.

Controls on today's aircraft flight decks are generally activated using a multitude of rotary knobs, push-buttons, cursor control devices, etc. Recently, the desire to find the optimal pilot-aircraft interface, where the interactions flow as naturally as possible, motivated the research community to focus on alternative control technologies (Leger, 1998; Merchant, & Schnell, 2000; Rood, 1998). Considerable research has been conducted for the military and civilian agencies regarding eye-tracking and gaze control systems, speech recognition and voice control, touch screens, touchless gesture recognizers and controllers, and other nonconventional controls as both standalone alternative control input methods and in combination with conventional physical interfaces. The application of new control technologies may simplify operations on flight decks by allowing more direct and intuitive interactions with aircraft systems and reducing the effort of interaction with systems when selecting the most contextually suitable control modality or combination of modalities.

An important benefit from integrating eye gaze into user interfaces is that eye position implies the area of the user's attention. People easily gaze at the world around them while performing other tasks. Therefore, with certain types of tasks, eye gaze combined with other input modalities requires little additional effort. Yet, several factors indicate that gaze control technology is not mature enough to serve as a primary command input method. These factors include the continual complexities of addressing the inadvertent activation or Midas touch, performance stability and robustness, and efficiency and effectiveness in interpreting user intent. Jacob (1993, 1995) pointed out that eye movement-based interfaces had evolved toward more implicit commands where, rather than responding to their explicit inputs, such interfaces give the impression of responding to the users' intentions. Instead of requiring the user to make specific eye movements to activate features and functions, the system uses information collected from user's natural eye movements while interacting with it. This new style of designing user interfaces may be well suited for implementation in complex and dynamic visual stimuli environments such as flight decks.

REFERENCES

Agustin, J. S., Mateo, J., Hansen, J. P., & Villanueva, A. (2009). Evaluation of the potential of gaze input for game interaction. *PsychNology Journal, 7*(2), 213-235.

Alonso, R., Causse, M., Vachon, F., Parise, R., Dehais, F., & Terrier, P. (2013). Evaluation of head-free eye tracking as an input device for air traffic control. *Ergonomics, 56*(2), 246-255.

Anders, G. (2001). Pilot's attention allocation during approach and landing–eye- and head-tracking research in an A330 full flight simulator. In *International Symposium on Aviation Psychology* (ISAP). Columbus, OH. Retrieved from http://www.geerdanders.de/literatur/2001 ohio.html

Ashmore, M., Duchowski, Andrew T., & Shoemaker, G. (2005). *Efficient eye pointing with a fisheye lens*. Department of Computer Science, Clemson University, SC.

Automotive HMI research project (2015, August). Retrieved from http://www.eyetechds.com/automotive-eye-tracking-research-project.html

Bjorklund, C.M., Alfredson, J., & Dekker, S.W.A. (2006). Mode monitoring and call-outs: An eye-tracking study of two-crew automated flight deck operations. *Int. J. Aviat. Psychol. 16*(3), 257–269 (2006).

Bolt, R. A. (1982). Eyes at the interface. In *Proceedings of the 1982 conference on Human Factors in Computing Systems*. Gaithersburg, Maryland: ACM, 360-362. doi:10.1145/800049.801811

Borah, J. (1995, September). *Investigation of eye and head controlled cursor positioning techniques*. Air Force Materiel Command Report AL/CF-SR-1995-0018

Borah, J. (1998). Technology and application of gaze based control. *Alternative Control Technologies: Human Factors Issues* (Section 3-1). Brétigny, France: NATO.

Brandt, H. F. (1937). A bidimensional eye-movement camera. *The American Journal of Psychology, 49*(4), 666-669.

Buswell, G. (1935). *How people look at pictures*. Chicago, IL: The University of Chicago Press.

Calhoun, G. L., Janson, W. P., & Arbak, C. J. (1986). Use of eye control to select switches. *Proceedings of the Human Factors and Ergonomics Society 30th Annual Meeting, 30*, 154–158.

Calhoun, G. L., Arbak, C. J., & Boff, K. R. (1984). Eye-controlled switching for crew station design. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting, 28*(3), 258–262.

Chin, C. A., Barreto, A., Cremades, J.G. & Adjouadi, M. (2008). Integrated electromyogram and eye-gaze tracking cursor control system for computer users with motor disabilities. *Journal of Rehabilitation Research & Development, 45*(1), 161-174.

Chuang, L. L., Nieuwenhuizen, F. M., & Bulthoff, H. H. (2013). A fixed-based flight simulator study: The interdependence of flight control performance and gaze efficiency. *Lecture Notes in Computer Science, 8020*, 95-104.

COGAIN. (2008). *Exploration of safety issues in eyetracking*. (Project Report IST-2003-511598). Network of Excellence Information Society Technologies.

Cornsweet, T. N., & Crane, H. D. (1973). Accurate two-dimensional eye tracker using first and fourth purkinje images. *Journal of the Optical Society of America, 63*(8), 921-928.

Delabarre, E. B. (1898). A method of recording eye-movements. *The American Journal of Psychology, 9*(4), 572-574.

de Reus, A.J.C., Zon, R., Ouwerkerk, R. (2012). *Exploring the use of an eye tracker in a helmet mounted display*. (NLRTP-2012-001) Amsterdam, The Netherlands: NLR National Aerospace Laboratory.

Dodge, R., & Cline, T. (1901). The angle velocity of eye movements. *Psychological Review, 8*(2), 145-157.

Duchowski, A. T. (2002). A breadth-first survey of eye-tracking applications. *Behav Res Meth 34*(4), 455–470.

Duchowski, A. T., Cournia, N. A., & Murphy, H. A. (2004). Gaze-contingent displays: A review. *CyberPsychol Behav, 7*(6), 621–634.

Ekman, I. M., Poikola, A. W., & Mäkäräinen, M. K. (2008). Invisible eni: Using gaze and pupil size to control a game. In *CHI '08 Proceedings of the SIGCHI conference on Human Factors in Computing Systems: Extended Abstracts*. Florence, Italy: ACM, 3135-3140.

Fisher, D. F., Monty, R. A., & Senders, J. W. (Eds.) (1981). *Eye movements: Cognition and visual perception*. Hillsdale, NJ: Lawrence Erlbaum.

Fitts, P. M., Jones, R. E., & Milton, J. L. (1950). Eye movements of aircraft pilots during instrument-landing approaches. *Aeronautical Engineering Review, 9*(2), 24-29.

Fujii, K., Salerno, A., Sriskandarajah, K., Kwok, K. W., Shetty, K. & Yang, G. Z. (2013). Gaze contingent cartesian control of a robotic arm for laparoscopic surgery. *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 3582-3589.

Gehm, R. (2013, January). *Hyundai HCD-14 Genesis concept previews eye-tracking, gesture recognition technologies*. SAE International. Retrieved from http://articles.sae.org/11727/

Glenn, F. A., Iavecchia, H. P., Ross, L. V., Stokes, J. M., Weiland, W. J., Weiss, D., & Zaklad, A. L. (1986). Eye-voice-controlled interface. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting, 30*(4), 322-326.

Graeber, D. A., & Andre, A. D. (1999). Assessing visual attention of pilots while using electronic moving maps for taxiing. In R.S. Jensen, B. Cox, J. D. Callister, & R. Lavis (Eds.), *Proceedings of the Tenth Symposium on Aviation Psychology, 2*, 791-796. Columbus, OH: Ohio State University.

Hansen, J. P., Alapetite, A., MacKenzie, I. S., & Møllenbach, E. (2014). The use of gaze to control drones. *Proceedings of the ACM Symposium on Eye Tracking Research and Applications - ETRA 2014*, pp. 27-34. New York, NY: ACM. doi:10.1145/2578153.2578156

Hansen, J. P., Engell-Nielsen, T. & Glenstrup, A. J. (1998). *Eye-gaze interaction: A new media –not just a fast mouse.* Retrieved from http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.97.4792&rep=rep1&type=pdf

Hatfield, F., Jenkins, E.A., & Jennings, M.W. (1995). *Eye/voice mission planning interface* (EVMPI) (Technical report AL/CF-TR-1995-0204), Wright-Patterson AFB, OH: Air Force Materiel Command.

Hartridge, H., & Thomson, L. C. (1948). Methods of investigating eye movements. *British Journal of Ophthalmology, 32*(9), 581-591.

Heo, H., Lee, E., Park, K., Kim, C., & Whang, M. (2010). A realistic game system using multi-modal user interfaces. *IEEE Transactions on Consumer Electronics, 56*(3), 1364-1372.

Horrey, W. J., Wickens, C. D. & Consalus, K.P. (2006). Modeling drivers' visual attention allocation while interacting with in-vehicle technologies. *Journal of Experimental Psychology: Applied, 12*(2), 67-78.

Huey, E. B. (1918). *The psychology and pedagogy of reading with a review of the history of reading and writing, and of methods, texts, and hygiene in reading.* New York, NY: The Macmillan Company.

Istance, H., Aulikki, H., Lauri, I., Santtu, M., & Stephen, V. (2010). *Designing gaze gestures for gaming: An investigation of performance.* In proceedings of the 2010 Symposium on Eye-Tracking Research & Applications, 22-24 March, 2010, Austin, Texas. doi:10.1145/1743666.1743740

Istance, H. O., Bates R., Hyrskykari, A., & Vickers S. (2008). Snap clutch, a moded approach to solving the Midas Touch problem. In *proceedings of the 2008 symposium on Eye tracking research & applications ETRA '08* (pp 221-228). Savannah, GA: ACM Press.

Jacob, R. (1991). The use of eye movements in human-computer interaction techniques: What you look at is what you get. *ACM Transactions on Information Systems, 9*(3), 152-169.

Jacob, R. (1993). Eye movement-based human-computer interaction techniques: Toward non-command interfaces. *Advances in Human-Computer Interaction, 4*, 151-190.

Jacob, R. J. K. (1995). *Eye tracking in advanced interface design: Virtual environments and advanced interface design,* New York, NY: Oxford University Press.

Jacob, R. J. K., & Karn, K. S. (2003). Eye tracking in human-computer interaction and usability research: Ready to deliver the promises. In Hyona, Radach, & Deubel (Eds.), *The Mind's Eye: Cognitive and Applied Aspects of Eye Movement Research* (pp. 573-605). Oxford, England: Elsevier Science BV.

Jarodzka, H., Scheiter, K., Gerjets, P., & van Gog, T. (2010). In the eyes of the beholder: How experts and novices interpret dynamic stimuli. *Learning and Instruction, 20*(2), 146–154.

Jasper, H., & Walker, R. (1931). The Iowa eye-movement camera. *Science, 74*(1916), 291-294.

Judd, C. H., McAllister, C. N., & Steel, W. M. (1905). General introduction to a series of studies of eye movements by means of kinetoscopic photographs. In J. M. Baldwin, H. C. Warren, & C. H. Judd (Eds.), *Psychological Review, Monograph Supplements, 7*, 1-16. Baltimore, MD: The Review Publishing Company.

Karsten, G., Goldberg, B., Rood, R., & Sulzer, R. (1975). *Oculometer measurement of air traffic controller visual attention*. (FAA Report No. FAA-NA-74-61). Atlantic City, NJ: National Aviation Facilities Experimental Center.

Kasarskis, P., Stehwien, J., Hickox, J., Aretz, A., & Wickens, C. (2001). Comparison of expert and novice scan behaviors during VFR flight. In *Proc. 11th International Symposium on Aviation Psychology*, Columbus, OH.

Kratchounova, D. (2016). From Inter*face* to Inter*space*: Conceptual Framework for Multimodal Flight Deck Controls. In *Springer Lecture notes in computer science Vol. 9740, Proceedings of the 8th International Conference, Virtual, Augmented and Mixed Reality 2016, HCI International 2016*, Toronto, Canada, July 17-22, 2016 pp 406-415.

Lambert, R. H., Monty, R. A., & Hall, R. J. (1974). Methods & designs high-speed data processing and unobtrusive monitoring of eye movements. *Behavior Research Methods & Instrumentation, 6*(6), 525-530.

Law, B., Atkins, M. S., Kirkpatrick, A. E., & Lomax, A. J. (2004). Eye gaze patterns differentiate novice and experts in a virtual laparoscopic surgery training environment. In A. T. Duchowski and R. Vertegaal, Roel (Eds.), *ETRA 2004 - Proceedings of the Eye Tracking Research and Application Symposium*, San Antonio, Texas (pp. 41-48). New York, NY: ACM Press.

Leger, A. (1998). Synthesis and expected benefits analysis. *RTO Lecture Series 215 Alternative Control Technologies: Human Factors Issues, 9-1*, 9-10.

Mackworth, J. F., & Mackworth, N. A. (1958). Eye fixations recorded on changing visual scenes by the television eye-marker. *Journal of the Optical Society of America, 48*(7), 439-445.

Mackworth, N. H., & Thomas, E. L. (1962). Head-mounted eye-marker camera. *Journal of the Optical Society of America, 52*(6), 713-716.

Majaranta, P. (2012) Communication and text entry by gaze. In Majaranta et al. (Eds.), *Gaze interaction and applications of eye tracking: Advances in assistive technologies* (pp. 63-77), Hershey, PA: IGI Global.

Majaranta, P., Ahola, U. K., & Špakov, O. (2009). Fast gaze typing with an adjustable dwell time. Department of Computer Sciences, Finland: University of Tampere, Finland. In *Proceedings of the 27th international conference on human factors in computing systems* (pp. 357-360), New York, NY: ACM Press.

Majaranta P., Bates R., & Donegan M. (2009) Eye tracking. In C. Stephanidis (Ed.), *The universal access handbook* (pp. 1-17). Boca Raton, FL: CRC Press.

Majaranta, P., & Bulling, A. (2014). Eye tracking and eye-based human-computer interaction. In S. H. Fairclough & K. Gilleade (Eds.), *Advances in physiological computing* (pp. 39-65). London: Springer-Verlag.

McMillan, G. R., Eggleston, R. G., & Anderson, T. R. (1997). Nonconventional controls. In G. Salvendy (Ed.), *Handbook of Human Factors and Ergonomics* (2nd ed.), pp. 729-771. Hoboken, NJ: Wiley.

Merchant, S., & Schnell, T. (2000). Applying eye tracking as an alternative approach for activation of controls and functions in aircraft. In *Proceedings of the 19th Digital Avionics Systems Conference* (pp. 5A5/1 - 5A5/9), New York, NY: IEEE Computer Society.

Miles, W. R., & Shen, E. (1925). Photographic recording of eye movements in the reading of Chinese in vertical and horizontal axes method and preliminary results. *Journal of Experimental Psychology, 8*(5), 344-362.

Nielsen, J. (1994). Enhancing the explanatory power of usability heuristics. In *Proceedings of the SIGCHI conference on Human Factors in Computing Systems* (pp. 152-158). ACM.

O'Brien, J. (1922). *Silent reading*. New York, NY: The MacMillan Company.

Ohno, T. (1998). Features of eye gaze interface for selection tasks. In J. Bruer (Ed.), *The Third Asia Pacific Computer Human Interaction, APCHI '98* (pp. 176-181), New York, NY: IEEE Computer Society.

Ottati, W. L., Hickox, J. C., & Richter, J. (1999). Eye scan patterns of experienced and novice pilots during visual flight rules (VFR) navigation. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting, 43*(1), 66-70.

Oviatt, S. & Cohen, P. (2000, March). Multimodal interfaces that process what comes naturally. In *Communications of the ACM, 43*(3), 45-53. New York, NY: ACM Press. doi:10.1145/330534.330538

Pomplun, M., Ivanovic, N., Reingold, E. M., & Shen, J. (2001). Empirical evaluation of a novel gaze-controlled zooming interface. *Proceedings of HCI International 2001, 1*, Mahwah, NJ: Lawrence Erlbaum.

Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin, 124*(3), 372-422.

Reiser, L. (2015, May 6). *Mesa eye tracking company aims to curb distracted driving* [Video file]. Retrieved from http://www.azfamily.com/story/28999246/mesa-eye-tracking-company-aims-to-curb-distracted-driving?autostart=true

Rood, G. M. (1998). Human factors issues for the integration of alternative control technologies. *RTO Lecture Series 215 Alternative Control Technologies: Human Factors Issues, 8-1*, 1-7.

Salvucci, D. D., & Anderson, J. R. (2000). Intelligent gaze-added interfaces. In *Proceedings of CHI 2000 Human Factors in Computing Systems* (pp. 273-280). New York, NY: ACM Press.

Shaviv, B. D. (1993). *The design and improvement of an eye controlled interface.* Unpublished Report, Center for Visual Computing, Stony Brook University, Stoney Brook, NY. Retrieved from http://www.cs.sunysb.edu/~vislab/projects/eye/Reports/report/report.pdf

Sheena, D., & Flagg, B. N. (1978). Semiautomatic eye movement data analysis techniques for experiments with varying scenes. In J. W. Senders, D. F. Fisher & R. A. Monty (Eds.), *Eye Movements and the Higher Psychological Functions* (pp. 65-75). Hillsdale, NJ: Lawrence Erlbaum.

Sibert, L. E., & Jacob, R. J. K. (2000). Evaluation of eye gaze interaction. In *Proceedings of the SIGCHI conference on Human Factors in Computing Systems* (pp. 281-288). New York, NY: ACM Press.

Skovsgaard, H., Mateo, J. C., & Hansen, J. P. (2011). Evaluating gaze-based interface tools to facilitate point-and-select tasks with small targets. *Behaviour & Information Technology, 30*(6), 821-831.

Song, J., Cho, S., Baek, S., Lee, K., & Bang, H. (2014). GaFinC: Gaze and finger control interface for 3D model manipulation in CAD application. *Computer-Aided Design. 46*, 239-245.

Tanriverdi, V., & Jacob, R. J. K. (2000). Interacting with eye movements in virtual environments. In *Proceedings of the SIGCHI conference on Human Factors in Computing Systems* (pp. 265-272). New York, NY: ACM Press.

Thomas, P., Biswas, P., and Langdon, P. (2015). State-of-the-art and future concepts for interaction in aircraft cockpits. In M. Antona & C. Stephanidis (Eds.), *Universal Access in Human-Computer Interaction: Access to Interaction* (pp. 538-549). New York, NY: Springer.

Tinker, M. A. (1936). Eye movements in reading. *Journal of Educational Research, 30*, 241-277.

Tinker, M. A. (1963). *Legibility of print*. Ames, IA: Iowa State University Press.

Tobiipro. (2015). *How do eye trackers work*. Retrieved from http://www.tobiipro.com/learn-and-support/learn/how-do-tobii-eye-trackers-work/

Tong, H. M., & Fisher, R. A. (1984). Progress Report on an Eye-Slaved Area-of-Interest Visual Display. (Report No. AFHRL-TR-84-36) *Proceedings of IMAGE III Conference*. Brooks Air Force Base, TX: Air Force Human Resources Laboratory.

TopGunMilitary. (2013, May). *Army AH064 attack helicopter* [Video file]. Retrieved from https://www.youtube.com/watch?v=UruW0N2mPsY

Van der Lans, R., Wedel, M., & Pieters, R. (2011). Defining eye-fixation sequences across individuals and tasks: The Binocular-Individual Threshold (BIT) algorithm. *Behavior Research Methods, 43*(1), 239–257. doi:10.3758/s13428-010-0031-2

Ware, C., & Mikaelian, H. H. (1987). An Evaluation of an Eye Tracker as a Device for Computer Input. In *CHI '87 Proceedings of the SIGCHI/GI Conference on Human Factors in Computing Systems and Graphics Interface* (pp. 183-189). Toronto, Ontario, Canada.

Wendt, P. R. (1952). Development of an eye camera for use with motion pictures, *Psychological Monographs: General and Applied, 7*(66), 1-18.

Wilcox, T., Evans, M., Pearce, C., Pollard, N., & Sundstedt, V. (2008). Gaze and voice based game interaction: the revenge of the killer penguins. Poster presented at *ACM SIGGRAPH 2008* (p. 1). New York, NY: ACM Press.

Yarbus, A. (1967). *Eye movements and vision*. New York, NY: Plenum Press.

Yamato, M., Monden, A., Matsumoto, K. Inoue, K., & Torii, K. (2000). Quick button selection with eye gazing for general GUI environment. *International Conference on Software: Theory and Practice (ICS2000)* (pp. 712-719).

Zhai, S. Morimoto, C., & Ihde, S. (1999). Manual and gaze input cascaded (MAGIC) pointing. *Proceedings of the Conference on Human Factors in Computing Systems (CHI'99)* (pp. 246-253). New York, NY: ACM Press.

APPENDIX A

History of Eye Tracking Prior to Gaze Control

There are three identified eras of eye movement research (Rayner, 1998):

- Era 1: 1879 to 1920;
- Era 2: 1920 to the mid-1970s; and
- Era 3: The mid-1970s to 1998.

For the purposes of this literature review, the ending of the third era was identified at the beginning of the 1980s, when gaze control technology began to emerge.

Era 1: 1879 to 1920.

The first era of eye movement studies could be characterized as a period of basic eye movements (i.e., saccades, fixation, saccadic suppression,[7] and saccade latency) research (Rayner, 1998). Huey (1918) and Rayner (1998) both credited the earliest eye movement research to Louis Émile Javal in 1879. Javal's research involved an eye tracker and examined discontinuous eye movements during reading.[8] In 1891, Ahrens attempted to record eye movements on a smoked drum by attaching a light ivory cup to the reader's cornea and fastened a bristle pointer to the cup. While his attempt in recording eye movements was unsuccessful, he was credited with creating the first head-mounted eye-tracking device (Huey, 1918; O'Brien, 1922). Delabarre (1898) conducted a similar experiment with a plaster of Paris cup. The results for recording slow movements were reported to be accurate; however, faster, natural movements led to results difficult to interpret.

Dodge and Cline's (1901) contribution to eye tracking was the successfully and accurately photographed beam of light that was reflected from the cornea at varying angles, thus recording of eye movements without physically placing a cup on the participant's eye. Referred to as "corneal reflection," this technique was further refined during the next fifty years of eye-tracking research and serves as the basis for one of the eye-tracking methods used in gaze control technology today (Huey, 1918; Mackworth & Mackworth, 1958; O'Brien, 1922).

Judd, McAllister, and Steele (1905) reported an experiment in which they recorded the temporal aspects of eye movements in two dimensions by applying kinetoscopic photography to a small particle placed on the cornea. They noted that while Dodge and Cline (1901) measured "the bright spot" in the eye, several bright spots could be seen on the front of the cornea, which did not move in parallel with normal eye movements. Instead, they inserted a small white speck of material into their participants' eyes and recorded the movement of the speck (Jacob & Karn, 2003; Judd, McAllister, & Steele, 1905)

By the end of the first era, basic understanding of saccades, fixation, saccadic suppression, and saccade latency had significantly advanced. Initial development of eye-tracking technology and the resulting rudimentary data collection methods were also established at this time. The next fifty years of eye-tracking research built on this foundation.

Era 2: 1920 to mid-1970s.

Research from this period was more applied than the research from the first era. However, little progress was made in better understanding the cognitive processes involved in eye movement (Rayner, 1998). Rather, the advances that occurred in this period occurred in data collection methods, head movement recording, new types of eye attachments, and the inclusion of closed circuit video cameras in data collection.

As data collection methods began improving in the 1920s, Miles and Shen (1925) recorded both horizontal and vertical movements by using the same piece of film for horizontal and vertical data

---

[7] Rayner (1998) defines saccadic suppression as the fact that we do not perceive information during an eye movement. Rayner (1998) defines saccade latency as the time it takes to initiate an eye movement. Saccades are the quick movements where the eye moves. Fixations are the periods of time that an eye remains relatively focused on an area.

[8] The discontinuous eye movements consist of saccades and fixations.

collections through the same camera. The result was four different data sets, depending on the direction of the reading and the motion of the recording film. Jasper and Walker (1931) built upon this research by measuring vertical and horizontal movement on two separate pieces of moving film using corneal reflections (Figure 2 and Figure 3). However, unlike Miles and Shen, Jasper and Walker collected both categories of data simultaneously.

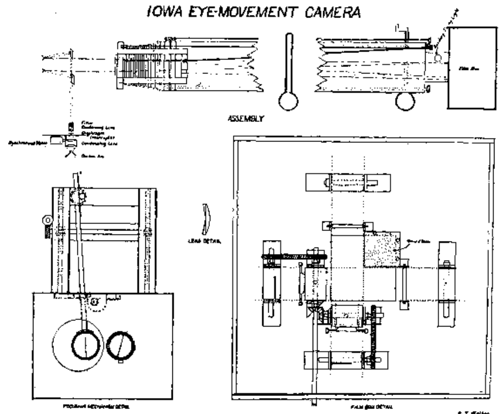

Figure 2. Diagram of the Iowa Eye Movement Camera (Jasper & Walker, 1931).

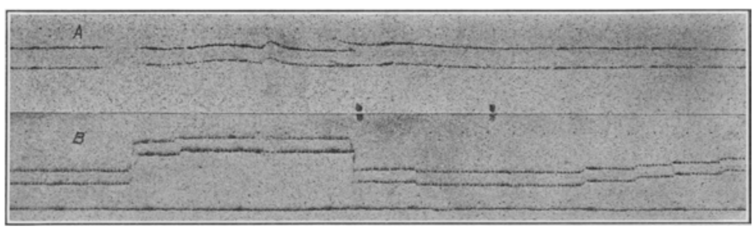

Figure 3. Image of a sample of film that collected the vertical movements of both eyes (Jasper & Walker, 1931).

Buswell (1935) recorded horizontal and vertical eye movement simultaneously, but instead used two film strips oriented 90° from each other (Figure 4 and Figure 5). While prior research mainly focused on eye movements in reading, which relies primarily on horizontal eye motion, Buswell integrated vertical eye movement analysis in addition to horizontal eye movement analysis.

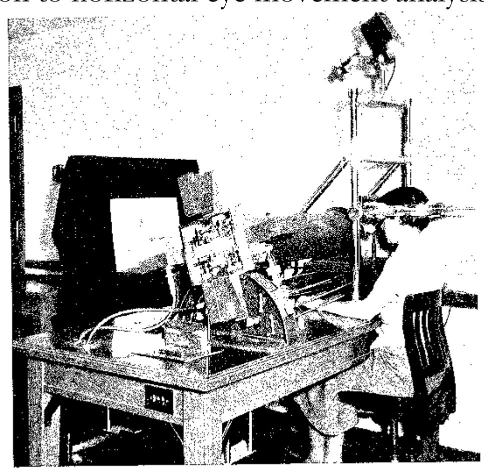

Figure 4. Image of Buswell's eye-tracking apparatus (Buswell, 1935).

Buswell also recorded head movements as part of his data collection. Figure 5 depicts a sample of the film Buswell used. Each film recorded either horizontal or vertical movement of the eyes and head.

Horizontal movement was recorded on the upper sample, and the vertical movement on the lower sample. The data identified with an "H" show head movements, while the data labeled with an "E" show eye movements.

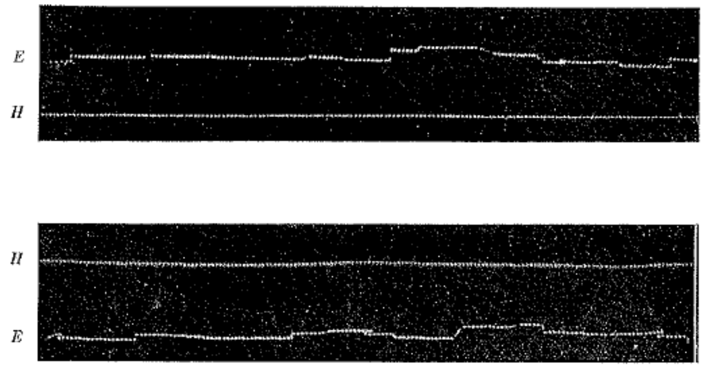

Figure 5. Image of a sample of film used to collect eye and head movement (Buswell, 1935).

Brandt (1937) used a 35-mm camera to collect horizontal and vertical eye movement data, and the results were printed on one strip of film. Most of the data collected by Brandt used dynamic imagery as opposed to the static imagery used by his predecessors (Buswell, 1935; Brandt, 1937; Mackworth & Mackworth, 1958).

Hartridge and Thomson (1948) developed a head-mounted eye-tracking system that consisted of a low-powered microscope attached to a plaster of Paris hat (Figure 6).

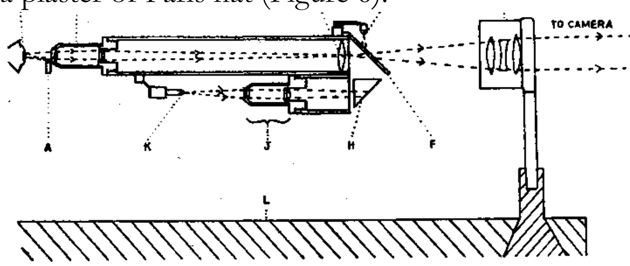

Figure 6. Diagram of Hartridge and Thomson's original head mounted eye-tracking system.

They eventually improved upon this design by enabling the head-mounted tracker to hang from a counterpoise frame, adding a mouth plate for head stabilization, and adding a screw to allow vertical height adjustment (Figure 7).

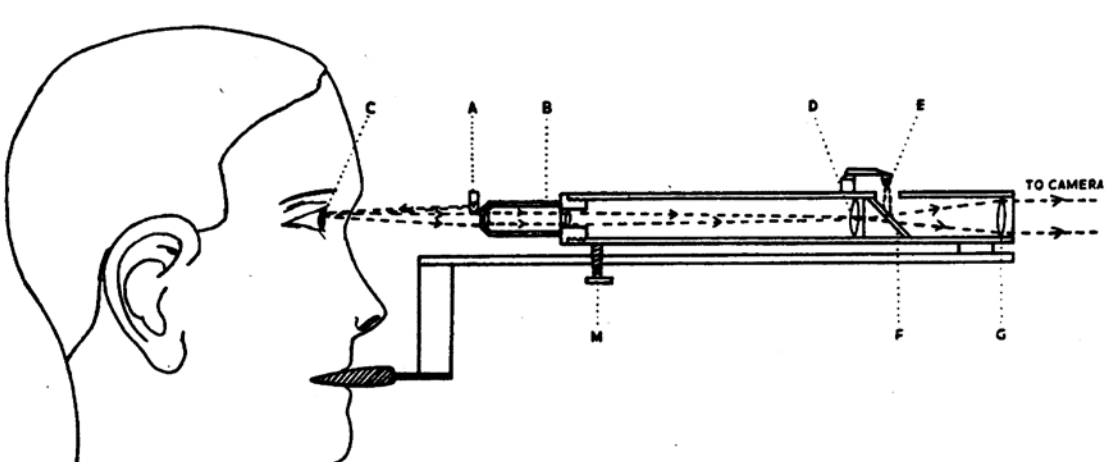

Figure 7. Diagram of Hartridge and Thomson's improved head-mounted eye-tracking system.

The updated equipment provided the following benefits:

- enabled the measurement of eye movements separate from head movements,
- eliminated the need to restrain head movement,
- eliminated problems arising from immobilizing the participant's head,
- allowed identification of movements resulting from head movement instead of eye movement, and
- eliminated any impact that fluctuations in vascularity might have on the position of the equipment in relation to the eye.

Wendt (1952) modified the Brandt camera by increasing the candle power of the corneal light sources, changing the false eye type, adapting the viewing angle, using an adjustable chair, and installing a bite stick (Figure 8).

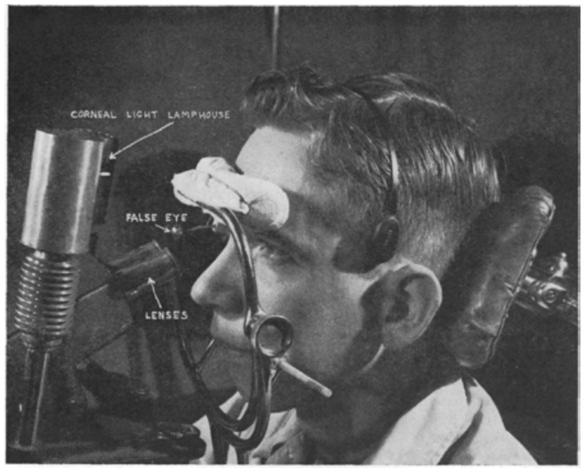

Figure 8. Photograph of Wendt's Eye Movement Recording Device (Wendt, 1952).

In 1956, Yarbus created a system of eye-tracking devices—referred to as "caps"—that attached directly to the human eye. The cap was applied to the eye by suction and could be modified to suit the needs of the individual experiment. Two examples are shown in Figure 9 and Figure 10.

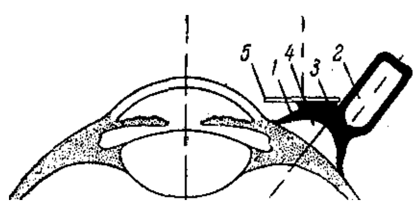

Figure 9. Diagram of the P1 suction cap (Yarbus, 1967).

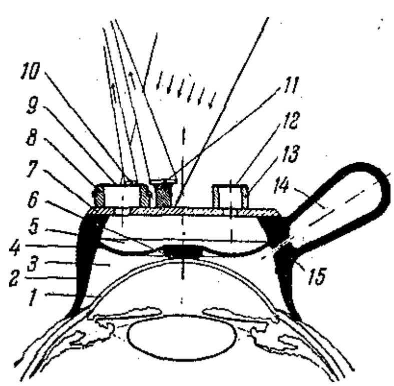

Figure 10. Diagram of the P5 Cap (Yarbus, 1967).

Mackworth and Mackworth (1958) developed a method to track eye movement using two closed-circuit video cameras to produce a composite picture. The first camera, shown in Figure 11, recorded the area the participant was observing. The second camera used the bright patch from the participant's corneal reflection as an eye-marker and superimposed it over the first picture. The resulting image identified the area at which the participant was currently looking. Mackworth and Thomas (1962) improved on this design through the integration of a head-mounted 8 mm movie camera.

Figure 11. Mackworth and Thomas' (1962) head mounted 8mm eye marker camera.

Tinker (1936, 1963) classified the corneal reflection method and the electrical method as the two primary methods of eye tracking. The corneal reflection method was a system where light was reflected off the cornea of the eye into an enlarging camera and recorded on a 35 mm film. The electrical recording method measured the electric potential changes between the front of the eyes (cornea) and the back of the eyes (retina) by placing electrodes on the temples near the eye. The electrodes were then connected to an amplifying device and either an ink writing oscillograph or photographic device.

In summary, the second era of eye tracking brought new methods that allowed for the collection of larger, higher-quality data sets. However, during the 1950s through the mid-1970s little eye movement research was conducted. The lull in activity into the 1970s was likely due to the time and effort required to collect and analyze vast quantities of data collected during experimentation (Jacob & Karn, 2003).

Era 3: Mid-1970s to 1980s.

Research during this era resulted in further improvements of eye-tracking technology and the creation of new methods of eye tracking. These improvements granted increased accuracy and ease of obtaining data (Jacob & Karn, 2003; Rayner, 1998).

Cornsweet and Crane (1973) developed a method of eye tracking that used the first and fourth Purkinje images[9] and the spatial separation between the two images to measure eye rotation eye translation. This comparison allowed the differentiation between eye rotations and head movement because the two Purkinje images moved together during head movement and differently during eye rotation.

Between 1974 and 1980, the U.S. Army Human Engineering Laboratory sponsored three symposia on eye movements. The proceedings were gathered into three texts: *Eye Movements and Psychological Processes* (1976)*, Eye Movements and the Higher Psychological Functions* (1978), and *Eye Movements: Cognition and Visual Perception* (1981). Eye-tracking technology and methods were emphasized only during the first symposium. Methods discussed included double Purkinje, corneal-reflection-pupil-center measurement technique, a Digital Eye Camera, and the development of the Programmed Eye-track Recording System and Eye-coupled Ubiquitous Scene (PERSEUS) generator. At the second conference, Sheena and Flagg (1978)

[9] Purkinje images are images formed when light is reflected off different parts of the eye. A first Purkinje image is formed by light reflected off the front of the cornea. The second Purkinje forms when light reflects off the rear surface of the cornea. A third Purkinje image forms when light is reflected off the front surface of the lens. The fourth Purkinje image is formed when light reflects off the rear surface of the lens.

presented a study on eye-tracking data analysis techniques that included the development of a semi-automatic digitizer system that manually tracked changes in the x-y coordinates of items of interest in changing scenes.

In 1974, Lambert, Monty, and Hall reported the creation of an eye-tracking system that did not impose unnatural constraints or mechanical attachments on the participant, and included high-speed on-line data processing. The system, called the Edgerton, Germeshausen, & Grier, Inc./United States Army Human Engineering Laboratory (EG&G/HEL) oculometer, or the Lambert oculometer, occupied an entire laboratory split into three areas: (a) the test room/studio, (b) the camera room, and, and (c) the control room/data reduction facility. One of the key components of the control room was the electronic processor, which computed the participant's gaze by determining the distance between the center of the pupil and the location of a highlight. The experimenter in the control room manually entered an initial targeting of the participant's pupil, after which an automatic tracker maintained the connection until the pupil was obscured. Eye position and behavior were fed into the PDP-11/20 computer, which performed two vital functions: (a) it processed all the data in real time, allowing the experimenter to change the visual material if desired, and (b) it relayed the processed raw data to a tape deck for storage.

Bolt (1982) described two eye-tracking systems, one commercially available and one prospective. The commercially available system was an eyeglass-mounted system that used the corneal reflection of an infra-red LED and a space sensor used to measure head position. The prospective system was not worn by the user and determined point-of-regard by using a corneal reflection technique. The system was very expensive but also reportedly highly accurate; however, no additional details were given to support this statement.

In summary, the third era of eye-tracking resulted in several advancements that are key parts of gaze control technology today. These included a more robust set of eye-tracking methods, enabled tracking of fixation point on an x-y coordinate scale, and development of the first head free eye trackers. The integration of computers into eye-tracking technology, however, was the advancement that allowed gaze control technology to begin development.

APPENDIX B

State of the Art (Market Survey)

Gaze control and multimodal user interfaces are more prevalent in domains other than aviation. Such a trajectory of aviation technology development is common. That is, development and implementation in other domains allows time for the technology to mature with often associated improvements in reliability and predictability. The current state of the art for gaze control technology is in a state of transition, a place where decades of development have laid the foundation for the change from laboratory concept research and development to productionisation and operational reality, particularly in the automotive and computer gaming industries. This recent growth in gaze control hardware and software in conjunction with reports of eye-tracking companies working with undisclosed businesses (Automotive HMI Research Project, 2015) may signal the arrival of readily available and affordable eye-tracking and gaze control systems to the consumer electronics market. This section will review some of the publicly available information about the use of gaze as an input control device for non-aviation applications including assistive technology for disabled users, consumer electronics (e.g., cell phones), medical use, advanced gaming applications, automobile industries, and virtual reality.

Assistive Technologies

One of the early motivations to develop gaze control technology began with the desire to help disabled users access computer technology in order to improve quality of life, including systems that will speak for the user when they input words and sentences into a computer while using unimodal gaze control. Currently, the Communication by Gaze Interaction (COGAIN) Association is one organization that works to improve quality of life of individuals with motor-control-disorders through use of gaze control communications and entertainment technology. Research has been conducted to minimize the limitations of typing or writing through unimodal gaze control. Traditionally, the user's gaze focuses on a letter display screen with the user's dwell providing the icon selection or confirmation. However, a current product limitation is the difference in communicating by speech (150-250 words per minute [wpm] and gaze control [10-15 wpm]; Majaranta, 2012).

Consumer electronics (Samsung Galaxy S®4, and later models)

The Samsung Galaxy S®4 is able to detect the user's face and eyes and react to eye movement using a built-in eye tracker. The system features include (a) no time-out if the user maintains eye contact with the screen, (b) automatic updates to match the angle at which the user is viewing the screen, (c) video pausing when the user looks away from it, and (d) screen scrolling.

Medical Use

Fujii et al. (2013) proposed a hands-free system that enabled the surgeon to be part of the robot control feedback loop, allowing user-friendly camera panning and zooming and avoiding the limitations of using dwell-time camera control methods. Additionally, the system featured a completely hands-free setup without the need of obtrusive sensors mounted on the surgeon or the use of a foot pedal. The quantitative and qualitative system and usability performance measures employed in the evaluation of the proposed system included (a) recognition accuracy of the probability-based Hidden Markov Modeling (HMM) algorithm for predicting gaze gestures, (b) task completion times, (c) workload measures for each of the three control modes, and (d) difficulty of system skill acquisition ratings. The results indicated that the proposed system was ergonomically sound—the cognitive workload was comparable to using a camera assistant or a foot pedal activated camera system. These results were corroborated by the difficulty of system skill acquisition ratings. In addition, the system enabled the surgeon to achieve comparable performance without the need of additional foot pedals.

Computer Games

The use of gaze control in computer games has generated considerable growth over the last decade. A few examples of games that have been used to test gaze control technology include “Halflife,” “World of Warcraft,” “Quake 2,” and “Neverwinter Nights.” In addition to these games, “Assassin's Creed Rogue” has recently demonstrated the integration of gaze control technology with keyboard and mouse through deliberately integrating a feature where the player can use the keyboard and mouse normally, while at the same time using gaze technology to control the direction of the computer screen in a 3-D real-time combat environment.

Automotive Applications

The Hyundai Genesis HCD-14 Concept Car integrates multimodal gaze control technology (gaze and 3-D gesture) to allow the driver to control information and entertainment systems without manual controls. This multimodal system was developed by Tobii™ and uses gaze control to select functions and hand-gesture control to confirm selections (Gehm, 2013). Another gaze control application to automobiles is using eye-tracking technology to monitor drivers and alert them when they become distracted from the task of driving (Reiser, 2015). A third application was developed by EyeTech© (Automotive HMI Research Project, 2015) for an unidentified automotive company in 2014. The system prototyped a system that used eye tracking, gestures, and touch as a form of driver input and had the ability for eye-tracking and gesture devices to be used at the same time.

Military Aviation

The Apache AH-64 Attack Helicopter contains the Modernized Target Acquisition Designation Sight/Pilot Night Vision Sensor (M-TADS/PNVS) system, also referred to as Arrowhead. The gaze control system on the M-TADS allows the pilot to aim their weapon systems by merely looking at the desired target (TopGunMilitary, 2013).

Virtual Reality (VR)

Three companies, FOVE-inc., Starbreeze, and Oculus VR, are developing virtual reality systems that include eye tracking and gaze control. FOVE’s system is projected to include integrated gaze control and head tracking technology and eye tracking with 0.2 degrees of accuracy, as well as a high resolution display with a field of view of over 100 degrees. A concept called “foveated rendering” is used. Foveated rendering identifies where in 3D space the user is looking and then adjusts the focus on that area by reallocating rendering resources. Starbreeze is working with Tobii to integrate eye-tracking technology to their virtual reality StarVR headset (http://uploadvr.com/hands-on-with-starvrs-premium-headset-and-the-computers-that-power-it/). Tobii is also offering improvements on VR that include foveated rendering, natural targeting, and accurate 3D stereoscopic rendering enhancements to VR graphics rendering. Tobii’s natural targeting function is used in computer games to allow the user to pick up, aim, or throw objects in an interactive gaming environment using only their eyes. Accurate 3D stereoscopic rendering can be achieved only when the system knows where the user is looking (http://www.tobii.com/tech/products/vr/).

No details are publicly available about Oculus VR ongoing research into integrating eye tracking into their products, however SensoMotoric Instruments offers an upgrade kit to the Oculus Rift DK2 headset that provides eye-tracking functions (http://www.smivision.com/en/gaze-and-eye-tracking-systems/products/eye-tracking-hmd-upgrade.html).

Summary

Gaze control input technologies and multimodal interfaces that include gaze control capabilities are currently under research, development, and production in many domains including assistive technology, consumer electronics, medical use, gaming applications, automobile industries, and virtual reality applications. Recent growth and the projected expansion, especially of gaming applications, automobile industries, and virtual reality domains, indicate continued growth and technology readiness for transition from laboratory settings to more presence in operational environments. However, further development and testing are still being conducted. Therefore, the current *state* of the art across domains is in a *state* of transition, and should continually be monitored during this period.