Running head: GAZE-CONTINGENT MULTI-RESOLUTIONAL DISPLAYS

Gaze-Contingent Multi-resolutional Displays:

An Integrative Review

Eyal M. Reingold[+], Lester C. Loschky*, George W. McConkie* and Dave M. Stampe[+]

[+]Department of Psychology, University of Toronto

100 St. George Street, Toronto, Ontario, Canada M5S 3G3

*Beckman Institute for Advanced Science and Technology,

405 North Mathews, Urbana, IL 61801

Address correspondence to:

Dr. Eyal M. Reingold
Department of Psychology
University of Toronto
100 St. George Street
Toronto, Ontario, Canada, M5S 3G3
Phone: (416) 978-3990
E-mail: reingold@psych.utoronto.ca

Abstract

Gaze-contingent multi-resolutional displays (GCMRDs) place high-resolution information only in the area to which the user's gaze is directed. This portion of the display is referred to as the *area of interest* (AOI). Image resolution and details outside the AOI are reduced, lowering the requirements for processing resources and transmission bandwidth in demanding display and imaging applications. This review provides an integrative survey of the current literature on GCMRDs across a wide range of applications. It also provides a general framework within which such research can be integrated, evaluated, and guided. Within this framework, a GCMRD (or "moving window") is analyzed in terms of (1) its multi-resolutional images (also called "variable-resolution", "space-variant", or "level of detail"), and (2) its gaze-contingent (or "foveated" or "eye-slaved"), movement of the AOI. We also synthesize the known human factors research on GCMRDs, and point out important questions for future research and development. Actual or potential applications of this research include flight, medical and driving simulators, virtual reality, remote piloting and teleoperation, infrared and indirect vision, image transmission and image retrieval, telemedicine, video teleconferencing, robotics and artificial vision systems.

Introduction

Technology users often need or want large, high-resolution displays that exceed possible or practical limits on bandwidth and/or computation resources. In reality, however, much of the information that is generated and transmitted in such displays is wasted since it cannot be resolved by the human visual system, which resolves high-resolution information in only a small region. One way to reduce computation and bandwidth requirements is to reduce the amount of unresolvable information in the display by presenting lower-resolution in the visual periphery. Over the last two decades, a great amount of work has been put into developing and implementing gaze-contingent multi-resolutional displays (or GCMRDs). A GCMRD is a display showing an image with high resolution in one area, and lower resolution elsewhere, and the high-resolution area is centered on the viewer's fovea by means of a gaze tracker or other mechanism. Work on such displays is found in a variety of research areas, often using different terms for the same essential concepts. Thus, the gaze-contingent aspect of such displays has also been referred to as "foveated" or "eye-slaved" and the multi-resolutional aspect is often referred to as "variable-resolution", "space-variant", "area of interest", or "level of detail". When considered together, gaze-contingent multi-resolutional displays have been referred to with various combinations of the above terms, or simply as a "moving window." Figure 1 shows examples of a short sequence of a viewer's gaze locations in an image, and two types of multi-resolutional images that might appear during a particular eye fixation.

[Insert Figure 1 about here]

Note that the gaze-contingent display methodology has also had a tremendous influence in basic research on perception and cognition in areas such as reading and visual search (for a review, see Rayner, 1998), however the present review exclusively focuses on the use of such displays in applied contexts.

Why Use Gaze-Contingent Multi-resolutional Displays?

Saving bandwidth and/or processing resources and the GCMRD solution. The most demanding display and imaging applications have very high resource requirements for resolution, field of view, and frame rates. The total resource requirement is proportional to the product of these factors, and usually not all can be met simultaneously. An excellent example of such an application is seen in military flight simulators that require a wraparound field of view, image resolution approaching the maximum resolution of the visual system (which is at least 60 cycles/degree or 120 pixels/degree; e.g., Thibos, Still, & Bradley, 1996, figure 7), and fast display updates with minimum delay. Because it is not feasible to create image generators, cameras, or display systems to cover the entire field of view with the resolution of the foveal region, the GCMRD solution is to monitor where the observer's attention is concentrated and to supply higher resolution and greater image transfer or generation resources to this area, with reduced resolution elsewhere. The stimulus location to which the gaze is directed is generally called the *point of gaze*. We will refer to the local stimulus region surrounding the point of gaze, which is assumed to be the center of attention, as the *attended area of interest* (A-AOI), and the area of high resolution in the image as the *displayed area of interest* (D-AOI)[1]. GCMRDs integrate a system for tracking viewer gaze position (by combined eye and head tracking) with a display that can be modified in real time to center the D-AOI at the point of gaze. If a high-resolution D-AOI appears on a lower-resolution background, one can simultaneously supply fine detail in central vision and a wide field of view with reasonable display, data channel, and image source requirements.

In general, there are two sources of savings from GCMRDs. First, there is a reduction in the bandwidth required for transmitting images, since information encoding outside the D-AOI is greatly reduced. Second, in circumstances where images are being computer-generated, there are reduced rendering requirements, because it is simpler to render low-resolution than high-resolution image regions, and therefore there is a reduction in computer processing resources (see Table 1 for examples).

[Insert Table 1 about here]

Unfortunately, GCMRDs can also produce perceptual artifacts, such as perceptible image blur or image motion, which have the potential to distract the user (Loschky, 2002; Loschky & McConkie, 2000; Loschky & McConkie, 2002; McConkie & Loschky, in press; Parkhurst, Culurciello, & Neibur, 2000; Reingold & Loschky, in press; Shioiri & Ikeda, 1989; van Diepen & Wampers, 1998; Watson, Walker, Hodges, & Worden, 1997). Ideally, one would like a GCMRD that maximizes the benefits of processing and bandwidth savings while minimizing perception and performance costs. However, depending on the needs of the users of a particular application, greater weight may either be given to perceptual quality or to processing and bandwidth savings. For example, in the case of a GCMRD in a flight simulator, maximizing perceptual quality of the display may be more important than minimizing the monetary expenses associated with increased processing (i.e., in terms of buying larger capacity, faster processing hardware). However, in the case of mouse-contingent multi-resolutional internet image downloads for casual users, minimizing perceptible peripheral image degradation may be less important than maximizing bandwidth savings in terms of download speed. In addition, it is worth pointing out that perceptual and performance costs are not always the same. For example, a GCMRD may have moderately perceptible peripheral image filtering, yet not reliably disrupt visual task performance (Loschky & McConkie, 2000). Thus, when measuring perception and performance costs of a particular GCMRD configuration, it is important to decide how low or high one's cost threshold should be set.

Are GCMRDs really necessary? A question that is often asked about GCMRDs is whether they will become unnecessary when bandwidth and processing capacities are greatly expanded in the future. As noted by Geisler (2001), in general, one will always want bandwidth and processing savings whenever they are possible, which is the reason nobody questions the general value of image compression. Furthermore, as one needs larger, higher resolution images, and faster update rates, the benefits of GCMRDs, in terms of compression ratios and processing savings, become greater. This is because larger images have proportionally more peripheral image information, which can be coded with increasingly less detail and resolution, resulting in proportionally greater savings. These bandwidth and processing savings can then be traded for larger images, with higher resolution in the area of interest, and faster update rates. Even if the bandwidth problem was eliminated in the context of certain applications in the future, and thus GCMRDs might not be needed for them, the bandwidth problem will still be present in other applications into the foreseeable future (e.g., Virtual Reality (VR), simulators, teleconferencing, teleoperation, remote vision, remote piloting, telemedicine, etc.). Finally, even if expanded bandwidth and processing capacity makes it possible to use a full resolution display of a given size for a given application, there may be good reasons to reduce the computational requirements where possible. Reducing computational requirements saves energy, and energy savings are clearly an increasingly important issue. This is particularly true for portable, wireless applications, which tend to be battery-powered, where added energy capacity requires greater size and weight. Thus, for all of the above reasons, it seems reasonable to argue that GCRMDs will be useful for the foreseeable future (see Geisler, 2001 for similar arguments).

Why Should GCMRDs Work?

The concept of the GCMRD is based on two characteristics of the human visual system. First, the resolving power of the human retina is multi-resolutional. Second, the region of the visual world from which highest resolution is gathered is changed from moment-to-moment by moving the eyes and head.

The multi-resolutional retina. The multi-resolutional nature of the retina is nicely explained by the sampling theory of resolution (e.g., (Thibos, 1998), which argues that variations in visual resolution across the visual field are due to differences in information sampling. In the fovea, it is the density of cone photoreceptors that best explains the drop-off in resolution. However, in the visual periphery, it is the cone-to-ganglion cell ratio that seems to explain the resolution drop-off (Thibos, 1998). Using such knowledge, it is possible to model the visual sampling of the retina, and estimate, for a given viewing distance and retinal eccentricity, how much display information is actually needed in order to support normal visual perception (Kuyel, Geisler, & Ghosh, 1999), though such estimates require empirical testing.

The most fundamental description of visual acuity is in terms of spatial frequencies and contrast, as described by Fourier analysis (Campbell & Robson, 1968), and the human visual system seems to respond to spatial frequency bandwidths (De Valois & De Valois, 1988). An important finding for the creation of multi-resolutional displays is that the human visual system shows a well-defined contrast sensitivity by retinal eccentricity relationship. As shown in Figure 2, Panel A, contrast sensitivity to higher spatial frequencies drops off as a function of retinal eccentricity[2] (e.g., Peli, Yang, & Goldstein, 1991; Pointer & Hess, 1989; Thibos et al., 1996). Therefore, in order to save bandwidth, a multi-resolutional image can exclude high-resolution information that is below contrast threshold at each eccentricity. However, if above-threshold spatial frequencies are excluded from the image, this will potentially degrade perception and/or distract the user, a point discussed in greater detail below.

[Insert Figure 2 about here]

Gaze movements. The concept of a gaze-contingent display is based on the fact that the human visual system compensates for its lack of high-resolution outside of the fovea by making eye and head movements. During normal vision, one simply points the fovea at whatever is of interest (i.e., the A-AOI) in order to obtain high-resolution information whenever needed. For small movements (e.g., under 20°) only the eyes tend to move, but as movements become larger, the head moves as well (Guitton & Volle, 1987; Robinson, 1979). This suggests that, in most GCMRD applications, eye tracking methods that are independent from, or compensate for, head movements are necessary to align the D-AOI of a multi-resolutional display with the point of gaze. Furthermore, just prior to, during, and following a saccade, perceptual thresholds are raised (for a recent review, see Ross, Morrone, Goldberg, & Burr, 2001). This saccadic suppression can help mask the stimulus motion that accompanies the updating of the D-AOI in response to a saccadic eye movement.

In sum, the variable resolution of the human visual system provides a rationale for producing multi-resolutional displays that reduce image resolution, generally describable in terms of a loss of higher spatial frequencies, with increasing retinal eccentricity. Likewise, the mechanisms involved in eye and head movements provide a rationale for producing dynamic displays that move the high-resolution D-AOI in response to the changing location of the point of gaze. Based on these ideas, a large amount of work has been has been carried out in a number of different areas including engineering design work on the development of GCMRDs, multi-resolutional image processing, and multi-resolutional sensors, and human factors research on multi-resolutional displays, gaze-contingent displays, and human-computer interaction. Unfortunately, it appears that many of the researchers in these widely divergent research areas are unaware of the related work done in the other areas. Thus, this review provides a useful function in bringing information from these different research areas to the attention of workers in these related fields. Moreover, the current review provides a general framework within which research across these areas can be integrated, evaluated, and guided. Accordingly, the remainder of this article begins by discussing the wide range of applications in which GCMRDs save bandwidth and/or processing resources at present or in which they are expected to do so in the future. The article then goes on to discuss research and development issues related to GCMRDs, which necessarily involves a synthesis of engineering and human factors considerations.

Finally, the current review points out key unanswered questions for the development of GCMRDs and suggests promising human factors research directions.

Applications of GCMRDs

Simulators

Simulation, particularly flight simulation, is the application area in which GCMRDs have been used the longest, and is still the GCMRD application area that has been most researched, due to the large amount of funding available (for examples of different types of flight simulators with GCMRDs, see Barrette, 1986; Dalton & Deering, 1989; Haswell, 1986; Thomas & Geltmacher, 1993; Tong & Fisher, 1984; Warner, Serfoss, & Hubbard, 1993). Flight simulators have been shown to save lives by eliminating the risk of injury during the training of dangerous maneuvers and situations (Hughes, Brooks, Graham, Sheen, & Dickens, 1982) and save money by reducing the number of in-flight hours of training needed (Lee & Lidderdale, 1983) as well as reducing airport congestion, noise, and pollution because of fewer training flights.

GCMRDs are useful in high-performance flight simulators because of the wide field of view and high-resolution needed. Simulators for commercial aircraft do not require an extensive field of view, as external visibility from the cockpit is limited to ahead and 45° to the sides. However, military aircraft missions require a large instantaneous field of view, with visibility above, to the sides and more limited visibility to the rear (Quick, 1990). Requirements vary between different flight maneuvers, but some demand extremely large fields of view, such as the barrel roll, which needs a 299° (horizontal) x 142° (vertical) field of view (Leavy & Fortin, 1983). Likewise, situational awareness has been shown to diminish with a field of view less than 100° (Szoboszlay, Haworth, Reynolds, Lee, & Halmos, 1995). Added to this are the demands for fast display updates with minimum delay and the stiff resolution requirements for identifying aircraft from various real-world distances. For example, aircraft identification at 5 nautical miles requires a resolution of 42 pixels/deg (21 cycles/degree) and recognition of a land vehicle at 2 nautical miles requires resolution of about 35 pixels/deg (17.5 cycles/degree)(Turner, 1984). Other types of simulators, e.g., automotive, have shown benefits from using GCMRDs as well (Kappe, Erp, & Korteling, 1999)(see also the Medical simulations and displays section below).

Virtual Reality

Other than simulators, VR is one of the areas in which GCMRDs will be most commonly used. In immersive VR environments, as a general rule, the bigger the field of view the greater the sense of ‘presence’ and the better the performance on spatial tasks, such as navigating through a virtual space (Arthur, 2000; Wickens & Hollands, 2000). Furthermore, update rates should be as fast as possible, because of a possible link with VR motion sickness (Frank, Casali, & Wierwille, 1988; Regan & Price, 1994; but see Draper, Viirre, Furness, & Gawron, 2001). For this reason, while having high-resolution is desirable, in general, speed of updating is given greater importance than the resolution of the display (Reddy, 1995). In order to create the correct view of the environment, some pointing device is needed to indicate the viewer’s vantage point, and head tracking is one of the most commonly used devices. Thus, in order to save scene-rendering time, which can otherwise be quite extensive, multi-resolutional VR displays are commonly used (for a recent review, see Luebke et al., 2002), and these are most often head-contingent (e.g., Ohshima, Yamamoto, & Tamura, 1996; Reddy, 1997; Watson et al., 1997). Reddy (1997, p. 181) has, in fact, argued that head tracking is often all that is needed to provide substantial savings in multi-resolutional VR displays, and he showed that taking account of retinal eccentricity created very little savings in at least two different VR applications (Reddy, 1997; Reddy, 1998). However, the applications he used had rather low maximum resolutions (e.g., 4.8-12.5 cycles/degree, or 9.6-25 pixels/degree). Obviously, if one wants a much higher resolution VR display, having greater precision in locating the point of gaze can lead to much greater savings than is possible with head tracking alone (see Research and Development Issues

Related to D-AOI Updating). In fact, several *gaze*-contingent multi-resolutional VR display systems have been developed (e.g., Levoy & Whitaker, 1990; Luebke, Hallen, Newfield, & Watson, 2000; Murphy & Duchowski, 2001, September). Each used different methods of producing and rendering gaze-contingent multi-resolutional 3D models, but all have resulted in a savings, with estimates of rendering time savings of roughly 80% over a standard constant-resolution alternative (Levoy & Whitaker, 1990; Murphy & Duchowski, 2001, September).

Infrared and Indirect Vision

Infrared and indirect vision systems are useful in situations where direct vision is poor or impossible. These include vision in low-visibility conditions (e.g., night operations and search and rescue missions) and in future aircraft designs with windowless cockpits. The requirements for such displays are similar to those in flight simulation: pilots need high-resolution for target detection and identification, and wide fields of view for orientation, maneuvering, combat, and tactical formations with other aircraft. However, these wide field of view requirements are in even greater conflict with resolution requirements because of the extreme limitations of Infrared Focal Plane Array and indirect vision cameras (Chevrette & Fortin, 1995; Grunwald & Kohn, 1994; Rolwes, 1990).

Remote Piloting and Teleoperation

Remote piloting and teleoperation applications are extremely useful in hostile environments such as deep sea, outer space, or combat, where it is not possible or safe for a pilot or operator to go. These applications require real-time information with a premium placed on fast updating so as not to degrade hand-eye coordination (e.g., Rosenberg, 1994).

Remote piloting of aircraft or motor vehicles. These applications have a critical transmission bottleneck because low bandwidth radio is the only viable option (DePiero, Noell, & Gee, 1992; Weiman, 1994). This is because line of sight microwave is often occluded by terrain and it exposes the vehicle to danger in combat situations, while fiberoptic cable is only possible for short distances and breaks easily. Remote driving requires both a wide field of view and enough resolution to be able to discern textures and identify objects. Studies have shown that operators are not comfortable operating an automobile (e.g., jeep) with a 40° field of view system, especially turning corners, but feel more confident with a 120° field of view (Kappe et al., 1999; McGovern, 1993; van Erp & Kappe, 1997). In addition, high-resolution is needed to identify various obstacles, and color can help distinguish such things as asphalt versus dirt roads (McGovern, 1993). Finally, frame rates of at least 10 frames per second (fps) are necessary for optic flow perception, which is critical in piloting (DePiero et al., 1992; Weiman, 1994).

Teleoperation. Teleoperation allows performance of dexterous manipulation tasks in hazardous or inaccessible environments. Examples include firefighting, bomb defusing, underwater or space maintenance or nuclear reactor inspection. In contrast to remote piloting, a narrower field of view is often acceptable in many teleoperation applications (Weiman, 1994). Furthermore, context is generally stable and understood, thus reducing the need for color. However, high-resolution for proper object identification is generally extremely important, and update speed is critical for hand-eye coordination. Multi-resolutional systems have been developed, including those that are head-contingent (Pretlove & Asbery, 1995; Tharp et al., 1990; Viljoen, 1998) and gaze-contingent (Viljoen, 1998), with both producing better target acquisition results than a joy-stick based system (ibid).

Image Transmission

Images are often transmitted through a limited-bandwidth channel, due to distance or data access constraints (decompression and network, disk, or tape data bandwidth limitations). This is illustrated below by considering two examples of applications involving image transmission through a limited-bandwidth channel (image retrieval and video teleconferencing).

Image retrieval. Image filing systems store and index terabytes of data. Compression is required to reduce the size of image files to a manageable level, for both storage and transmission.

Sorting through images, especially from remote locations over bandwidth-limited communication channels, is most efficiently achieved via progressive transmission systems, so that the user can quickly recognize unwanted images and terminate transmission early (Frajka, Sherwood, & Zeger, 1997; To, Lau, & Green, 2001; Tsumura, Endo, Haneishi, & Miyake, 1996; Wang & Bovik, 2001). If the point of gaze is known, then the highest resolution information can be acquired for that location first, with lower resolution being sent elsewhere (Bolt, 1984; To et al., 2001).

Video teleconferencing. Video teleconferencing is the audio and video communication of two or more people in different locations; typically there is only one user at a time at each node. It frequently involves sending video images over a standard low-bandwidth ISDN communication link (64 or 128 Kb/sec) or other low bandwidth medium. Transmission delays can greatly disrupt communication, and with current systems, frame rates of only 5 fps at a resolution of 320 x 240 pixels are common. In order to achieve better frame rates, massive compression is necessary. The video sent in teleconferencing is highly structured (Maeder, Diederich, & Niebur, 1996) in that the transmitted image usually consists of a face or head-and-shoulders, and the moving parts of the image are the eyes and mouth, which along with the nose, comprise the most looked-at areas of faces (Spoehr & Lehmkule, 1982). Thus, it makes sense to target faces for transmission in higher resolution than the rest of the image (Basu & Wiebe, 1998).

Development of GCMRDs for video teleconferencing has already begun. Kortum and Geisler (1996a) first implemented a GCMRD system for still images of faces, and followed this up with a video-based system (Geisler & Perry, 1998). Sandini, and colleagues (Sandini et al., 1996; Sandini, Questa, Scheffer, Dierickx, & Mannucci, 2000) have implemented a stationary retina-like multi-resolutional camera for visual communication by deaf people by videophone with sufficient bandwidth savings that a standard phone line can be used for transmission.

Medicine

Medical imagery is highly demanding of display fidelity and resolution. Fast image updating is also important in many such applications in order to maintain hand-eye coordination.

Telemedicine. This includes both teleconsultation with fellow medical professionals to get a second opinion, and telediagnosis and telesurgery by remote doctors and surgeons. Telediagnosis involves inspection of a patient either by live video or other medical imagery such as X-rays, and should benefit from the time savings provided by multi-resolutional image compression (Honniball, 1999). Telesurgery involves actual remote manipulation of surgical instruments. An example would be laproscopy, in which a doctor operates on a patient through small incisions, and cannot directly see or manipulate the surgical instrument inside the patient, therefore relying on video feedback. This is essentially telesurgery, whether the surgeon is in the same room or on another continent (intercontinental surgery was first performed in 1993; Rovetta et al., 1993). Teleconsultation may tolerate some loss of image fidelity, whereas in the case of telediagnosis or telesurgery, the acceptable level of compression across the entire image is more limited (Cabral & Kim, 1996; Hiatt, Shabot, Phillips, Haines, & Grant, 1996). Furthermore, telesurgery requires fast transmission rates to provide usable video and tactile feedback, since non-trivial delays can degrade surgeons' hand-eye coordination (Thompson, Ottensmeyer, & Sheridan, 1999). Thus, real-time foveated display techniques, such as progressive transmission, could potentially be used to reduce bandwidth to useful levels (Bolt, 1984).

Medical simulations and displays. As with flight and driving simulators, medical simulations can save many lives. Surgical residents can practice a surgical procedure hundreds of times before they see their first patient. Simple laparoscopic surgery simulators have already been developed for training. As medical simulations develop and become more sophisticated, their graphical needs will increase to the point that GCMRDs provide important bandwidth savings. Levoy and Whitaker (1990) have already shown the utility of gaze-contingent volume rendering of medical data sets.

Gaze-tracking could also be useful in controlling composite displays consisting of many different digital images, such as the patient's computerized tomography (CT) or magnetic resonance imaging (MRI) scans with real time video images, effectively giving the surgeon "x-ray vision." Yoshida, et al. (Yoshida, Rolland, & Reif, 1995a; Yoshida, Rolland, & Reif, 1995b), suggested that one method of accomplishing such fusion is to present CT, MRI, or ultrasound scans inside gaze-contingent insets, with the "real" image in the background.

Robotics and Automation

Having both a wide field of view, and an area of high-resolution at the 'focus of attention' is extremely useful in the development of artificial vision systems. Likewise, reducing the visual processing load by decreasing resolution in the periphery is of obvious value in artificial vision. High-resolution information in the center of vision is useful for object recognition, and lower-resolution information in the periphery is still very useful for detecting motion. Certain types of multi-resolutional displays, e.g., those involving log-polar mapping, make it easier to determine heading, motion, and time to impact than displays using Cartesian coordinates (Dias, Araujo, Paredes, & Batista, 1997; Kim, Shin, & Inoguchi, 1995; Panerai, Metta, & Sandini, 2000; Shin & Inokuchi, 1994).

Research and Development Issues Related to GCMRDs

Although ideally GCMRDs should be implemented in a manner undetectable to the observer (see Loschky, 2002 for an existence proof for such a display), in practice such a display may not be feasible, or indeed, needed for most purposes. The two main sources of detectable artifacts in GCMRDs are image degradation produced by the characteristics of multi-resolutional images, and perceptible image motion resulting from image updating. Accordingly, we summarize the available empirical evidence for each of these topics and provide guidelines and recommendations for developers of GCMRDs to the extent possible. However, many key issues remain unresolved or even unexplored. Thus, an important function of the present review is to highlight key questions for future human factors research on issues related to GCMRDs, as summarized in Table 2.

[Insert Table 2 about here]

Research and Development Issues with Multi-resolutional Images

Methods of producing multi-resolutional images. Table 3 summarizes a large body of work focused on developing methods for producing multi-resolutional images. Our review of the literature suggests that the majority of research and development efforts related to GCMRDs have focused on this issue. The methods that have been developed include (1) computer-generated images (e.g., rendering 2D or 3D models) with space-variant levels of detail, (2) algorithms for space-variant filtering of constant high-resolution images, (3) projection of different levels of resolution to different viewable monitors (e.g., in a wraparound array of monitors) or the projection of different resolution channels and/or display areas to each eye in a head-mounted display, and (4) space-variant multi-resolutional sensors and cameras. All of these approaches have the potential of making great savings in either processing or bandwidth, though some of the methods are also computationally complex.

[Insert Table 3 about here]

Using models of vision to produce multi-resolutional images. In most cases, the methods of multi-resolutional image production in Table 3 have been based on neurophysiological or psychophysical studies of peripheral vision, under the assumption that these research results will scale up to the more complex and natural viewing conditions of GCMRDs. This assumption has been explicitly tested in only a few studies that investigated the human factors characteristics of multi-resolutional displays (Duchowski & McCormick, 1998; Geri & Zeevi, 1995; Kortum & Geisler, 1996b; Loschky, 2002; Luebke et al., 2000; Peli & Geri, 2001; Sere, Marendaz, & Herault, 2000; Yang, Coia, & Miller, 2001), but the results have been generally supportive. For example, Loschky (2002) tested the psychophysically derived Yang and Miller resolution drop-off function shown in

Figure 2, panel A, by creating multi-resolutional images based on it and on functions with steeper and shallower drop-offs (as in Figure 2, panel D). Consistent with predictions, a resolution drop-off shallower than that in Figure 2, panel A was imperceptibly blurred, but steeper drop-offs were all perceptibly degraded compared to a constant high-resolution control condition. Furthermore, these results were consistent across multiple dependent measures that were both objective (e.g., blur detection and fixation durations) and subjective (e.g., image quality ratings).

However, there are certain interesting caveats. Several recent studies (Loschky, 2002; Peli & Geri, 2001; Yang et al., 2001) have noted that sensitivity to peripheral blur in complex images is somewhat lower than predicted by contrast sensitivity functions (CSFs) derived from studies using isolated grating patches. They have argued that this lower sensitivity during complex picture viewing may be due to lateral masking from nearby picture areas. In contrast, Geri and Zevi (1995) used drop-off functions based on psychophysical studies using Vernier acuity tasks, and found that sensitivity to peripheral blur in complex images was *greater* than predicted. They attributed this to the more global resolution discrimination task facing their subjects in comparison to the positional discrimination task in Vernier acuity. Thus, it appears that the appropriate resolution drop-off functions for GCMRDs should be slightly *steeper* than suggested by CSFs, but *shallower* than suggested by Vernier acuity functions. Consequently, to create undetectable GCMRDs, it is still advisable to fine-tune previously derived psychophysical drop-off functions based on human factors testing. Similarly, working out a more complete description of the behavioral effects of different detectable drop-off rates in different tasks is an important goal for future human factors research.

Discrete versus continuous resolution drop-off GCMRDs. A fundamental distinction exists between methods in which image resolution reduction is produced by having discrete levels of resolution (*discrete drop-off methods*: e.g., Loschky & McConkie, 2000; Loschky & McConkie, 2002; Parkhurst et al., 2000; Reingold & Loschky, in press; Shioiri & Ikeda, 1989; Watson et al., 1997), and methods in which resolution drops off gradually with distance from a point or region of highest resolution (*continuous drop-off methods*: e.g., Duchowski & McCormick, 1998; Geri & Zeevi, 1995; Kortum & Geisler, 1996b; Loschky, 2002; Luebke et al., 2000; Peli & Geri, 2001; Sere et al., 2000; Yang et al., 2001). Of course, using a sufficient number of discrete regions of successively reduced resolution approximates a continuous drop-off method. Figure 1 illustrates these two approaches. Panel C has a high-resolution area around the point of gaze with lower resolution elsewhere, whereas in Panel D the resolution drops off continuously with distance from the point of gaze.

These two approaches are further illustrated in Figure 2. As shown in panel A of Figure 2, we assume that there is an ideal useful resolution function that is highest at the fovea and drops off at more peripheral locations. Such functions are well established for acuity and contrast sensitivity (e.g., Peli et al., 1991; Pointer & Hess, 1989; Thibos et al., 1996). Nevertheless, the possibility is left open that the “useful resolution” function may be different from these in cases of complex, dynamic displays, perhaps on the basis of attentional allocation factors (e.g., Yeshurun & Carrasco, 1999). In panels B-E of Figure 2, we superimpose step functions representing the discrete drop-off methods, and smooth functions representing the continuous drop-off method.

With the discrete drop-off method there is a high-resolution D-AOI centered at the point of gaze. An example in which a bi-resolutional display would be expected to be just barely undetectably blurred is shown in panel B of Figure 2. If such thresholds can be established, or estimated from existing psychophysical data, for a sufficiently large number of levels of resolution, they can be used to plot the resolution drop-off function, as shown in panel C of Figure 2. Ideally, such a discrete resolution drop-off GCMRD research program would (1) test predictions of a model of human visual sensitivity that could be used to interpolate and extrapolate from the data, (2) parametrically and orthogonally vary the size of the D-AOI and level of resolution outside it, and (3) use a universally applicable resolution metric (e.g., cycles/degree). In fact, several human factors studies have used discrete resolution drop-off GCMRDs (Loschky & McConkie, 2000; Loschky & McConkie, 2002; Parkhurst et al., 2000; Shioiri & Ikeda, 1989; Watson et al., 1997), and each identified one or more combinations of D-AOI size and peripheral resolution that did not differ appreciably from a full high-resolution control condition. However, none of the above studies meet all three of the above-stated criteria, and thus are of limited use for plotting a widely generalizable resolution drop-off function for use in GCMRDs.

A disadvantage of the discrete resolution drop-off method, as compared to the continuous drop-off method, is that it introduces one or more relatively sharp resolution transitions, or edges, into the visual field, which may produce perceptual problems. Thus, a second question concerns whether such problems occur, and if so, would more gradual blending between different resolution regions eliminate them? Anecdotal evidence suggests that blending is useful, as suggested by a simulator study in which it was reported that having non-existent or small blending regions was very distracting, whereas a display with a larger blending ring was less bothersome (Baldwin, 1981). However, another simulator study found no difference between two different blending ring widths in a visual search task (Browder, 1989), and more recent studies have found no differences between blended versus sharp-edged bi-resolutional displays in terms of detecting peripheral image degradation (Loschky & McConkie, 2000, experiment 3) or initial saccadic latencies to peripheral targets (Reingold & Loschky, in press). Thus, further research on the issue of boundary-related artifacts using varying levels of blending and multiple dependent measures is needed to settle this question.

A clear advantage of the continuous resolution drop-off method is that, to the extent that it matches the visual resolution drop-off of the retina, it should provide the greatest potential image resolution savings. However, it also has a disadvantage relative to the discrete drop-off approach. As shown in panel E of Figure 2, if the loss of image resolution at some retinal eccentricity causes a perceptual problem, it is difficult to locate the eccentricity where this occurs, since image resolution is reduced across the entire picture. With the discrete drop-off method it is possible to probe more specifically to identify the source of such a retinal/image resolution mismatch. Furthermore, the discrete drop-off method can also be a very efficient method of producing multi-resolutional images under certain conditions. When images are represented using multi-level coding methods such as wavelet decomposition (Moulin, 2000), producing discrete drop-off multi-resolutional images is simply a matter of selecting which levels of coefficients are to be included in reconstructing the different regions of the image (e.g., (Frajka et al., 1997).

In deciding whether to produce continuous or discrete drop-off multi-resolutional images, it is also important to note that discrete levels of resolution may cause more problems with animated images than still images (Stampe & Reingold, 1995). This may involve both texture and motion perception, and therefore studies on ‘texture-defined motion’ (e.g., Werkhoven, Sperling, & Chubb, 1993) may be informative for developers of live video or animated GCMRDs (Luebke et al., 2002). Carefully controlled human factors research on this issue in the context of GCMRDs is clearly needed.

Color resolution drop-off. Importantly, the visual system also shows a loss of color resolution with retinal eccentricity. Though numerous studies have investigated this function and found important parallels to monochromatic contrast sensitivity functions (e.g., Rovamo & Iivanainen, 1991), to our knowledge this property of the visual system has been largely ignored rather than exploited by developers and investigators of GCMRDs (but see Watson et al., 1997, Exp. 2). We would encourage developers of multi-resolutional image processing algorithms to exploit this color resolution drop-off in order to produce even greater bandwidth and processing savings.

Research and Development Issues Related to D-AOI Updating

We now shift our focus to issues related to updating the D-AOI. Whether in a continuous or a discrete drop-off display, every time the viewer's gaze moves, the center of high-resolution must be quickly and accurately updated to match the viewer's current point of gaze. Of critical importance is the fact that there are several options as to how and when this updating occurs that can affect human performance. Unfortunately, much less research has been conducted on these issues than on those related to the multi-resolutional characteristics of the images. Accordingly, our discussion below primarily focuses on issues that should be explored by future research. Nevertheless, we attempt to provide developers with a preliminary analysis of the available options.

Overview of D-AOI movement methods. Having made the image multi-resolutional, the next step is to update the D-AOI position dynamically so that it corresponds to the point of gaze. As indicated by the title of this article, we are most interested in the use of gaze tracking information to position the D-AOI, but other researchers have proposed and implemented systems that use other means of providing position information as well. Thus far, the most commonly proposed means of providing positional information for the D-AOI include:

(a) true GCMRD—typically combined eye and head tracking are used to specify the point of gaze as the basis for image updating. Gaze position is determined by both the eye position in head coordinates and head position in space coordinates (Guitton & Volle, 1987),

(b) methods using pointer device input that approximates gaze tracking with lower spatial and temporal resolution and accuracy (e.g., head- or hand-contingent D-AOI movement), and

(c) methods that try to predict where gaze will move without requiring input from the user.

Gaze-contingent D-AOI movement. Gaze control is generally considered to be the most natural method of D-AOI movement because it does not require any act beyond making normal eye movements. No training is involved. And if the goal is to remove from the display any information that the retina cannot resolve, making the updating process contingent on the point of gaze allows maximum information reduction. The most serious obstacle for developing systems employing GCMRDs is the current state of gaze tracking technology. To illustrate, consider the following specifications of a gaze tracking system that would likely meet the requirements of the most demanding GCMRD applications: (a) plug-and-play, (b) unobtrusive (e.g., a remote system with no physical attachment to the observer), (c) accurate (e.g., < + 0.5° error), (d) high temporal resolution (e.g., 500 Hz sampling rate) to minimize updating delays, (e) high spatial resolution and low noise to minimize unnecessary image updating, (f) ability to determine gaze position in a wraparound 360° field of view, and (g) affordable. In contrast, current gaze tracking technologies tend to have trade-offs between factors such as ease of operation, comfort, accuracy, spatial and temporal resolution, field of view, and cost (Istance & Howarth, 1994; Jacob, 1995; Young & Sheena, 1975). Thus, we are faced with a situation in which the most natural and perceptually least problematic implementation of a GCMRD may be complex and uncomfortable to use, and/or relatively expensive and thus impractical for some applications.

Nevertheless, current high-end eye trackers are approaching practical usefulness, if not yet meeting ideal specifications, and are more than adequate for investigating many of the relevant human factors variables crucial for developing better GCMRDs. In addition, some deficiencies in present gaze tracking technology may be overcome by modifications to the designs of GCMRDs (e.g., enlarging the high-resolution area to compensate for problems caused by lack of spatial or temporal accuracy in specifying the point of gaze). Furthermore, recent developments in gaze tracking technology (e.g., Matsumoto & Zelinsky, 2000; Stiefelhagen, Yang, & Waibel, 1997), suggest that user-friendly systems (e.g., remote systems requiring no physical contact with the user), are becoming faster and more accurate. In addition, approaches that include prediction of the next gaze location based on the immediately prior one (Tannenbaum, 2000) may be combined with prediction based on salient areas in the image (Parkhurst, Law, & Niebur, 2002) to improve speed and accuracy.

Moreover, as more applications come to use gaze tracking within multi-modal human-computer interaction systems (e.g., Sharma, Pavlovic, & Huang, 1998), gaze tracking devices should begin to enjoy the economy of scale and become more affordable. Yet, even at current prices, levels of comfort, and levels of spatial and temporal resolution/accuracy, certain applications depend on the use of GCMRDs and work quite well (e.g., flight simulators).

Head-contingent D-AOI movement. At the present time, Head-contingent D-AOI movement seems generally better than gaze-contingent D-AOI in terms of comfort, relative ease of operation and calibration, and lower price. However, it is clearly worse in terms of resolution, accuracy and speed of the D-AOI placement. This is because for gaze movements to targets closer than 20°, head movements often do not occur (Guitton & Volle, 1987; Robinson, 1979). Thus, with a head-contingent D-AOI, if the gaze is moved to a target within 20° eccentricity, the eyes will move but the head may not, nor, consequently, will the D-AOI. This would result in lower spatial and temporal resolution and accuracy in moving the D-AOI to the point of gaze, and could cause perceptual and performance decrements (e.g., increased detection of peripheral image degradation, and longer fixation durations and search times).

Hand-contingent D-AOI movement. Likewise, hand-contingent D-AOI movement, although easy and inexpensive to implement (e.g., with mouse input), may suffer from slow D-AOI movement. This is because hand movements tend to rely on visual input for targeting. In pointing movements, the eyes are generally first sent to the target, and the hand follows after a lag of about 70 ms (e.g., Helsen, Elliot, Starkes, & Ricker, 1998), with visual input also being used to guide the hand towards the end of the movement (e.g., Heath, Hodges, Chua, & Elliott, 1998). Similar results have been shown for cursor movement on CRTs through manipulation of a mouse, touchpad, or pointing stick (Smith, Ho, Ark, & Zhai, 2000). The latter study also found another pattern of eye-hand coordination in which the eyes only slightly led the cursor, continually monitoring its progress. All of this suggests that perceptual problems may occur and task performance may be slowed because the eyes must be sent into the low-resolution area ahead of the hand, or the eyes (and hand) must make shorter than normal excursions in order to avoid going into the low-resolution area, or the eyes must follow the D-AOI at a lower than normal velocity.

Predictive D-AOI movement. A very different approach is to move the D-AOI predictively. This can be done either based on empirical eye movement samples (Duchowski & McCormick, 1998; Stelmach & Tam, 1994; Stelmach, Tam, & Hearty, 1991), or based on saliency-predicting computer algorithms (Milanese, Wechsler, Gill, Bost, & Pun, 1994; Parkhurst et al., 2002; Tanaka, Plante, & Inoue, 1998). The latter option seems much more practical for producing D-AOIs for an infinite variety of images. However, a fundamental problem with the entire predictive approach to D-AOI movement is that it may often fail to accurately predict the exact location that a viewer wants to fixate at a given moment in time (Stelmach & Tam, 1994). Nevertheless, the predictive D-AOI approach may be most useful when the context and potential areas of interest are extremely well defined, such as in video teleconferencing (Duchowski & McCormick, 1998; Maeder et al., 1996). In this application, the A-AOI can generally be assumed to be in the speaker’s face, particularly, as noted earlier, the eyes, nose, and mouth (Spoehr & Lehmkule, 1982). An even simpler approach in video teleconferencing is simply to have a D-AOI that is always at the center of the image frame (Woelders, Frowein, Nielsen, Questa, & Sandini, 1997), based on the implied assumption that people spend most of their time looking there, which is generally true (e.g., Mannan, Ruddock, & Wooding, 1997).

Causes of D-AOI update delays. Depending on the method of D-AOI movement one chooses, the delays in updating the D-AOI position will vary. As mentioned earlier, such delays constitute another major issue facing designers of GCMRDs. Ideally, image updating would place highest resolution at the point of gaze instantaneously. However, such a goal is virtually impossible to achieve, even with the fastest GCMRD implementation. The time required to update the image in response to a change in gaze position depends on a number of different processes including the method used to update the location of the D-AOI (e.g., gaze-contingent, head-contingent, hand-contingent), multi-resolutional image production delays, transmission delays, and delays associated with the display method.

In most GCMRD applications, the most important update rate bottleneck is the time to produce a new multi-resolutional image. If it is necessary to generate and render a 3D multi-resolutional image, or filter a constant high-resolution image, the image processing time can take anywhere from 25-50 ms (Geisler & Perry, 1999; Ohshima et al., 1996) to 130-150 ms (Thomas & Geltmacher, 1993) or longer, depending on the complexity of the algorithm being used. Thus, increasing the speed of multi-resolutional image processing should be an important goal for designers working on producing effective GCMRDs. In general, image-processing times can be greatly reduced by implementing them in hardware rather than software. The multi-resolutional camera approach, which can produce an image in as little as 10 ms (Sandini, 2001), is a good illustration of such a hardware implementation. In this case, however, there is an initial delay due to rotating the multi-resolutional camera to its new position. This can be done using mechanical servos, the speed of which depends on the weight of the camera, or by leaving the camera stationary and rotating a mirror with a galvanometer, which can move much more quickly.

Problems caused by D-AOI updating delays. There are at least two ways in which delays in updating the D-AOI position can cause perceptual difficulties. First, if the D-AOI is not updated quickly following a saccade, the point of gaze may initially be on a degraded region. Luckily, due to saccadic suppression, the viewer’s visual sensitivity is lower at the beginning of a fixation (e.g., Ross et al., 2001), and thus brief delays in D-AOI updating may not be perceived. But stimulus processing rapidly improves over the period of 20-80 ms after the start of a fixation, and thus longer delays may allow perception of the degraded image (McConkie & Loschky, in press). Second, when updates occur well into a fixation (e.g., 70 ms or later), the update may produce the perception of motion, and this affects perception and task performance (e.g., Reingold & Stampe, 2002; van Diepen & Wampers, 1998).

Simulator studies have shown that delays between gaze movements and the image update result in impaired perception and task performance and, in some cases, can cause simulator sickness (e.g., Frank et al., 1988; but see Draper et al., 2001). Turner (1984) compared delays ranging from 130-280 ms, and found progressive decrements in both path following and target identification tasks with increasing levels of throughput delay. In addition, two more recent studies demonstrated that fixation durations increased with an increase in image updating delays (Hodgson, Murray, & Plummer, 1993; Loschky & McConkie, 2000, Exp. 6).

Questions for Future Research

Below we outline several important issues for future human factors evaluation of GCMRDs. The first set of issues concerns the useful resolution function, the second set concerns issues that arise when producing multi-resolutional images, and the third set of issues concerns D-AOI updating (see Table 2).

Although the resolution drop-off functions shown in Figure 2, panel A, are a good starting point, an important goal for future human factors research should be to further explore such functions and variables that may affect them. These may include image and task variables such as lateral masking (Chung, Levi, & Legge, 2001), attentional cuing (Yeshurun & Carrasco, 1999), and task difficulty (Bertera & Rayner, 2000; Loschky & McConkie, 2000, Exp. 5; Pomplun, Reingold, & Shen, 2001), and subject variables such as user age (e.g., Ball, Beard, Roenker, Miller, & Griggs, 1988) and expertise (Reingold, Charness, Pomplun, & Stampe, 2001). In addition, human factors research should extend the concept of multi-resolutional images to the color domain. For example, can we construct a GCMRD using a hue resolution drop-off function that is just imperceptibly different from a full-color image and has a substantial information reduction? Furthermore, if the drop-off is perceptible, what aspects of task performance, if any, are negatively impacted? Finally, further research should quantify the perception and performance costs associated with removing above-threshold peripheral resolution (i.e., detectably degraded GCMRDs) (for related studies and discussion see Kortum & Geisler, 1996b; Loschky, 2002; Loschky & McConkie, 2000; Loschky & McConkie, 2002; Parkhurst et al., 2000; Shioiri & Ikeda, 1989; Watson et al., 1997).

Human factors research should assist GCMRD developers by exploring the perception and performance consequences of important implementation options. One of the most fundamental choices is whether to use a continuous or a discrete resolution drop-off function. These two methods should be compared with both still and animated images. There are numerous additional design choices that should also be explored empirically. For example, it is known that the shape of the visual field is asymmetrical (e.g., Pointer & Hess, 1989). This raises the question of whether the shape of the D-AOI (ellipse vs. circle vs. rectangle) in a bi-resolutional display has any effects on users' perception and performance. Likewise, any specific method of multi-resolutional image production may require targeted human factors research. For example in the case of rendering 2D or 3D models with space-variant levels of detail (e.g., in VR), it has been anecdotally noted that object details (e.g., doors and windows in a house) appear to pop in and out as a function of their distance from the point of gaze (Berbaum, 1984; Spooner, 1982). It is important to explore the perception and performance costs associated with such 'popping' phenomena. Similar issues can be identified with any of the other methods of multi-resolutional image production (see Table 3).

Human factors research into issues related to D-AOI updating is almost non-existent (but see Frank et al., 1988; Grunwald & Kohn, 1994; Hodgson et al., 1993; Loschky & McConkie, 2000, Exp. 6; McConkie & Loschky, in press; Turner, 1984). Two key issues for future research concern the D-AOI control method, and the D-AOI update delay. Given that a number of different methods of moving the D-AOI have been suggested and implemented (i.e., gaze-, head-, and hand-contingent, and predictive movement), an important goal for future research is to contrast these methods in terms of their perception and performance consequences. The second key question concerns the effects of a systematic increase in update delay on different perception and performance measures in order to determine when and how updating delays cause problems. Clearly, the chosen D-AOI control method will influence the update delay and resultant problems. Consequently, in order to compensate for a D-AOI control method having poor spatial or temporal accuracy/resolution, the size of the area of high-resolution may have to be enlarged (e.g., Loschky & McConkie, 2000, Exp. 6).

Conclusions

The present review is primarily aimed at two audiences: 1) Designers and engineers working on the development of applications and technologies related to GCMRDs and 2) Researchers investigating relevant human factors variables. Of course, given that empirical validation is an integral part of the development of GCMRDs, these two groups partially overlap and collaborations between academia and industry in this field are becoming more prevalent. Indeed, we hope that the present review may help facilitate such interdisciplinary links. Consistent with this goal, we recommend that studies of GCMRDs should, whenever appropriate, report information both on their effects on human perception and performance and on bandwidth and processing savings. To date, such dual reporting is rare (but see Luebke et al., 2000; Murphy & Duchowski, 2001, September; Parkhurst et al., 2000).

As is evident from the above review, research into issues related to GCMRDs is truly in its infancy with many unexplored and unresolved questions and with few firm conclusions. Nevertheless, the preliminary findings we reviewed clearly demonstrate the potential utility and feasibility of GCMRDs. The ultimate goal for GCMRDs is to produce savings by substantially reducing peripheral image resolution and/or detail yet be undetectably different from a normal image to the user. This has recently been shown in a few studies using briefly flashed (Geri & Zeevi, 1995; Peli & Geri, 2001; Sere et al., 2000; Yang et al., 2001) and gaze-contingent (Loschky, 2002) presentation conditions. Other studies (see Table 1) have shown that using GCMRDs can result in substantial savings in processing and/or bandwidth. Thus, the GCMRD concept is now beginning to be validated. Furthermore, general perceptual disruptions and performance decrements have been shown to be caused by (a) peripheral degradation removing useful visual information or inserting distracting information (Geri & Zeevi, 1995; Kortum & Geisler, 1996b; Loschky, 2002; Loschky & McConkie, 2000; Loschky & McConkie, 2002; Parkhurst et al., 2000; Peli & Geri, 2001; Reingold & Loschky, in press; Shioiri & Ikeda, 1989; Watson et al., 1997; Yang et al., 2001) and by (b) D-AOI update delays (Frank et al., 1988; Grunwald & Kohn, 1994; Hodgson et al., 1993; Loschky & McConkie, 2000, Exp. 1 & 6; McConkie & Loschky, in press; Turner, 1984; van Diepen & Wampers, 1998). Such studies illustrate the manner in which some of the performance costs associated with detectably degraded GCMRDs can be assessed.

Any application of GCMRDs must involve the analysis of tradeoffs between computation and bandwidth savings and the degree and type of perception and performance decrements that would result. Ideally, for most tasks where a GCMRD is appropriate, a set of conditions can be identified that will provide substantial computation/bandwidth reduction while still maintaining adequate, and perhaps even normal, task performance. Simply because an implementation results in a detectably degraded GCMRD does *not* mean that performance will deteriorate (Loschky & McConkie, 2000) and consequently, performance costs must be assessed directly. Developers must set a clear performance-cost threshold as part of such an assessment. A prerequisite for this step in the design process is a clear definition of tasks that are critical and typical of the application (i.e., a task analysis). In addition, a consideration of the characteristics of potential users of the application is important.

The specific target application provides important constraints (e.g. budgetary) that are vital for determining the available development options. For example, whereas gaze-contingent D-AOI update is a feasible and arguably the optimal choice in the context of flight simulators, given the cost of gaze trackers such a method may not be an option for other applications such as video teleconferencing and Internet image retrieval. Instead, hand-contingent and/or predictive D-AOI updating are likely to be the methods of choice for the latter applications.

Finally, as clearly demonstrated in the present article, human factors evaluation of relevant variables is vital for the development of the next generation of GCMRDs. The current review outlines a framework within which such research can be motivated, integrated and evaluated. The human factors questions listed in the above sections require investigating the perception and performance consequences of manipulated variables using both objective measures (e.g., accuracy, reaction time, saccade lengths, fixation durations) and subjective report measures (e.g., display quality ratings). Such investigations should be aimed at exploring the performance costs involved in detectably degraded GCMRDs and the conditions for achieving undetectably degraded GCMRDs. Although the issues and variables related to producing multi-resolutional images and to moving the D-AOI were discussed separately, potential interactions and trade-offs between these variables should also be explored. As our review indicates, the vast majority of these issues related to the human factors of GCMRDs are yet to be investigated and therefore represent a fertile field for research.

References

Arthur, K. W. (2000). Effects of field of view on performance with head-mounted displays. , University of North Carolina at Chapel Hill.

Baldwin, D. (1981). Area of interest: Instantaneous field of view vision model. Paper presented at the Image Generation/Display Conference II.

Ball, K. K., Beard, B. L., Roenker, D. L., Miller, R. L., & Griggs, D. S. (1988). Age and visual search: Expanding the useful field of view. Journal of the Optical Society of America, 5(12), 2210-2219.

Barrette, R. E. (1986). Flight simulator visual systems--An overview, Proceedings of the Fifth Society of Automotive Engineers Conference on Aerospace Behavioral Engineering Technology: Human Integration Technology: The Cornerstone for Enhancing Human Performance (pp. 193-198).

Basu, A., & Wiebe, K. J. (1998). Enhancing Videoconferencing Using Spatially Varying Sensing. IEEE Transactions on Systems, Man, & Cybernetics Part A: Systems & Humans, 28(2), 137-148.

Berbaum, K. S. (1984). Design criteria for reducing "popping" in area-of-interest displays: Preliminary experiments (NAVTRAEQUIPCEN 81-C-0105-8). Orlando, FL: Naval Training Equipment Center.

Bertera, J. H., & Rayner, K. (2000). Eye movements and the span of the effective stimulus in visual search. Perception & Psychophysics, 62(3), 576-585.

Bolt, R. A. (1984). The human interface: Where people and computers meet. London: Lifetime Learning Publications.

Browder, G. B. (1989). Evaluation of a helmet-mounted laser projector display. Proceedings of the SPIE: The International Society for Optical Engineering, 1116, 85-89.

Cabral, J. E., Jr., & Kim, Y. (1996). Multimedia systems for telemedicine and their communications requirements. IEEE Communications Magazine, 34, 20-27.

Campbell, F. W., & Robson, J. G. (1968). Application of Fourier analysis to the visibility of gratings. Journal of Physiology, 197, 551-566.

Chevrette, J. F., & Fortin, J. (1995). Wide angle coverage infrared surveillance system. Proceedings of the SPIE: The International Society for Optical Engineering.

Chung, S. T. L., Levi, D. M., & Legge, G. E. (2001). Spatial-frequency and contrast properties of crowding. Vision Research, 41(14), 1833-1850.

Dalton, N. M., & Deering, C. S. (1989). Photo based image generator (for helmet laser projector). Proceedings of the SPIE: The International Society for Optical Engineering, 1116, 61-75.

De Valois, R. L., & De Valois, K. K. (1988). Spatial Vision. New York, NY, US: Oxford University Press.

DePiero, F. W., Noell, T. E., & Gee, T. F. (1992). Remote driving with reduced bandwidth communication, Proceedings of the Sixth Annual Space Operation, Application, and Research Symposium (SOAR '92) (pp. 163-171). Houston, Texas: NASA Johnson Space Center.

Dias, J., Araujo, H., Paredes, C., & Batista, J. (1997). Optical normal flow estimation on log-polar images: A solution for real-time binocular vision. Real-Time Imaging, 3(3), 213-228.

Draper, M. H., Viirre, E. S., Furness, T. A., & Gawron, V. J. (2001). Effects of image scale and system time delay on simulator sickness within head-coupled virtual environments. Human Factors, 43(1), 129-146.

Duchowski, A. T. (2000). Acuity-matching resolution degradation through wavelet coefficient scaling. IEEE Transactions of Image Processing, 9(8), 1437-1440.

Duchowski, A. T., & McCormick, B. H. (1998). Gaze-contingent video resolution degradation. Paper presented at the Human Vision and Electronic Imaging II, Bellingham, WA.

Fernie, A. (1995). Helmet-mounted display with dual resolution. Journal of the Society for Information Display, 3(4), 151-153.

Fernie, A. (1996). Improvements in area of interest helmet mounted displays, Proceedings of the Conference on Training - Lowering the cost, maintaining the fidelity; (pp. 21.1-21.4). London, UK: Royal Aeronautical Society.

Frajka, T., Sherwood, P. G., & Zeger, K. (1997). Progressive image coding with spatially variable resolution, Proceedings of the International Conference on Image Processing (pp. 53-56).

Frank, L. H., Casali, J. G., & Wierwille, W. W. (1988). Effects of visual display and motion system delays on operator performance and uneasiness in a driving simulator. Human Factors, 30(2), 201-217.

Geisler, W. S. (2001). Space Variant Imaging: Frequently Asked Questions. Center for Vision and Image Sciences, University of Texas. Retrieved March 23, 2002, from http://fi.cvis.psy.utexas.edu/foveated_questions.htm

Geisler, W. S., & Perry, J. S. (1998). A real-time foveated multi-resolution system for low-bandwidth video communication. Proceedings of the SPIE: The International Society for Optical Engineering, 3299, 294-305.

Geisler, W. S., & Perry, J. S. (1999). Variable-Resolution Displays for Visual Communication and Simulation, The Society for Information Display (Vol. 30, pp. 420-423).

Geri, G. A., & Zeevi, Y. Y. (1995). Visual Assessment of Variable-Resolution Imagery. Journal of the Optical Society of America A-Optics & Image Science, 12(10), 2367-2375.

Grunwald, A. J., & Kohn, S. (1994). Visual field information in low-altitude visual flight by line-of-sight slaved helmet-mounted displays. IEEE Transactions on Systems, Man and Cybernetics, 24(1), 120-134.

Guitton, D., & Volle, M. (1987). Gaze control in humans: Eye-head coordination during orienting movements to targets within and beyond the oculomotor range. Journal of Neurophysiology, 58(3), 427-459.

Haswell, M. R. (1986). Visual systems developments, Advances in flight simulation--Visual and motion systems (pp. 264-271). London, UK: Royal Aeronautical Society.

Heath, M., Hodges, N. J., Chua, R., & Elliott, D. (1998). On-line control of rapid aiming movements: Unexpected target perturbations and movement kinematics. Canadian Journal of Experimental Psychology, 52(4), 163-173.

Helsen, W. F., Elliot, D., Starkes, J. L., & Ricker, K. L. (1998). Temporal and spatial coupling of point of gaze and hand movements in aiming. Journal of Motor Behavior, 30(3), 249-259.

Hiatt, J. R., Shabot, M. M., Phillips, E. H., Haines, R. F., & Grant, T. L. (1996). Telesurgery - Acceptability of compressed video for remote surgical proctoring. Archives of Surgery, 131(4), 396-400.

Hodgson, T. L., Murray, P. M., & Plummer, A. R. (1993). Eye movements during "area of interest" viewing. In G. d'Ydewalle & J. V. Rensbergen (Eds.), Perception and cognition: Advances in eye movement research (pp. 115-123). New York: Elsevier Science.

Honniball, J. R. T., P.J. (1999). Medical image databases: the ICOS project, Proceedings of the IEE Colloquium on Multimedia Databases and MPEG-7 (pp. 3/1-3/5).

Hughes, R., Brooks, R., Graham, D., Sheen, R., & Dickens, T. (1982). Tactical ground attack: On the transfer of training from flight simulator to operational red flag range exercise, Proceedings of the Human Factors Society - 26th Annual Meeting (pp. 596-600).

Istance, H. O., & Howarth, P. A. (1994). Keeping an eye on your interface: The potential for eye-based control of graphical user interfaces (GUI's). Paper presented at the HCI'94.

Jacob, R. J. K. (1995). Eye Tracking in Advanced Interface Design. In W. Barfield & T. A. Furness (Eds.), Virtual Environments and Advanced Interface Design (pp. 258-288). New York, NY, USA: Oxford University Press.

Kappe, B., Erp, J. v., & Korteling, J. E. (1999). Effects of head-slaved and peripheral displays on lane-keeping performance and spatial orientation. Human Factors, 41(3), 453-466.

Kim, K. I., Shin, C. W., & Inoguchi, S. (1995). Collision avoidance using artificial retina sensor in ALV, Proceedings of the Intelligent Vehicles '95 Symposium (pp. 183-187).

Kooi, F. L. (1993). Binocular configurations of a night-flight head-mounted display. Displays, 14, 11-20.

Kortum, P. T., & Geisler, W. S. (1996a). Implementation of a foveated image-coding system for bandwidth reduction of video images. In B. Rogowitz & J. Allebach (Eds.), Proceedings of the SPIE: The International Society for Optical Engineering (Vol. 2657, pp. 350-360).

Kortum, P. T., & Geisler, W. S. (1996b). Search performance in natural scenes: The role of peripheral vision. Paper presented at the ARVO.

Kuyel, T., Geisler, W., & Ghosh, J. (1999). Retinally reconstructed images: Digital images having a resolution match with the human eye. IEEE Transactions on Systems, Man, & Cybernetics Part A: Systems & Humans, 29(2), 235-243.

Leavy, W. P., & Fortin, M. (1983). Closing the gap between aircraft and simulator training with limited field-of-view visual systems, Proceedings of the 5th Interservice/Industry Training Equipment Conference (pp. 10-18).

Lee, A. T., & Lidderdale, I. G. (1983). Visual scene simulation requirements for C-5A/C-141B aerial refueling part task trainer (AFHRL-TP-82-34). Williams Air Force Base, AZ: Air Force Human Resources Laboratory.

Levoy, M., & Whitaker, R. (1990). Gaze-directed volume rendering. Computer Graphics, 24, 217-223.

Loschky, L. C. (2002). Investigating perception and eye movement control in natural scenes using gaze-contingent multi-resolutional displays. Doctoral dissertation, University of Illinois, Urbana-Champaign.

Loschky, L. C., & McConkie, G. W. (2000). User performance with gaze contingent multiresolutional displays. In A. T. Duchowski (Ed.), Proceedings of the Eye Tracking Research & Applications Symposium 2000 (pp. 97-103). Palm Beach, FL.: ACM.

Loschky, L. C., & McConkie, G. W. (2002). Investigating spatial vision and dynamic attentional selection using a gaze-contingent multi-resolutional display. Journal of Experimental Psychology: Applied, 8(2), 99-117.

Luebke, D., Hallen, B., Newfield, D., & Watson, B. (2000). Perceptually driven simplification using gaze-directed rendering (Technical Report CS-2000-04): University of Virginia.

Luebke, D., Reddy, M., Cohen, J., Varshney, A., Watson, B., & Huebner, R. (2002). Level of detail for 3D graphics. San Francisco: Morgan-Kaufmann Publishers.

Maeder, A., Diederich, J., & Niebur, E. (1996). Limiting human perception for image sequences. Proceedings of the SPIE: The International Society for Optical Engineering, 2657, 330-337.

Mannan, S. K., Ruddock, K. H., & Wooding, D. S. (1997). Fixation patterns made during brief examination of two-dimensional images. Perception, 26(8), 1059-1072.

Matsumoto, Y., & Zelinsky, A. (2000). An Algorithm for Real-time Stereo Vision Implementation of Head Pose and Gaze Direction Measurement, Proceedings of IEEE Fourth International Conference on Face and Gesture Recognition (FG'2000) (pp. 499-505). Grenoble, France.

McConkie, G. W., & Loschky, L. C. (in press). Perception onset time during fixations in free viewing. Behavioral Research Methods, Instruments, and Computers.

McGovern, D. E. (1993). Experience and results in teleoperation of land vehicles. In S. R. Ellis (Ed.), Pictorial communication in virtual and real environments (pp. 182-195). London: Taylor & Francis.

Milanese, R., Wechsler, H., Gill, S., Bost, J.-M., & Pun, T. (1994). Integration of bottom-up and top-down cues for visual attention using non-linear relaxation, Proceedings of Computer Vision and Pattern Recognition (pp. 781-785).

Moulin, P. (2000). Multiscale image decompositions and wavelets. In A. C. Bovik (Ed.), Handbook of Image and Video Processing (pp. 289-300). New York, USA: Academic Press.

Murphy, H., & Duchowski, A. T. (2001, September). Gaze-contingent level of detail rendering. Paper presented at the EuroGraphics 2001, Manchester, UK.

Ohshima, T., Yamamoto, H., & Tamura, H. (1996). Gaze-directed adaptive rendering for interacting with virtual space, Virtual Reality Annual International Symposium (Vol. 267, pp. 103-110): IEEE.

Panerai, F., Metta, G., & Sandini, G. (2000). Visuo-inertial stabilization in space-variant binocular systems. Robotics & Autonomous Systems, 30(1-2), 195-214.

Parkhurst, D., Culurciello, E., & Neibur, E. (2000). Evaluating variable resolution displays with visual search: Task performance and eye movements. In A. T. Duchowski (Ed.), Proceedings of the Eye Tracking Research & Applications Symposium 2000 (pp. 105-109). Palm Beach, FL.: ACM.

Parkhurst, D., Law, K., & Niebur, E. (2002). Modeling the role of salience in the allocation of overt visual attention. Vision Research, 42(1), 107-123.

Peli, E., & Geri, G. A. (2001). Discrimination of wide-field images as a test of a peripheral-vision model. Journal of the Optical Society of America A-Optics & Image Science, 18(2), 294-301.

Peli, E., Yang, J., & Goldstein, R. B. (1991). Image invariance with changes in size: The role of peripheral contrast thresholds. Journal of the Optical Society of America, 8(11), 1762-1774.

Pointer, J. S., & Hess, R. F. (1989). The contrast sensitivity gradient across the human visual field: With emphasis on the low spatial frequency range. Vision Research, 29(9), 1133-1151.

Pomplun, M., Reingold, E. M., & Shen, J. (2001). Investigating the visual span in comparative search: The effects of task difficulty and divided attention. Cognition, 81, B57-B67.

Pretlove, J., & Asbery, R. (1995). The design of a high performance telepresence system incorporating an active vision system for enhanced visual perception of remote environments. Proceedings of the SPIE: The International Society for Optical Engineering, 2590, 95-106.

Quick, J. R. (1990). System Requirements for a high gain dome display surface. Proceedings of the SPIE: The International Society for Optical Engineering, 1289, 183-191.

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

Reddy, M. (1995). A perceptual framework for optimizing visual detail in virtual environments, Framework for Interactive Virtual Environments (FIVE) Conference (pp. 179-188). London.

Reddy, M. (1997). Perceptually Modulated Level of Detail for Virtual Environments. Unpublished Doctoral thesis, University of Edinburgh, Edinburgh, U.K.

Reddy, M. (1998). Specification and evaluation of level of detail selection criteria. Virtual Reality: Research, Development and Application, 3(2), 132-143.

Regan, E. C., & Price, K. R. (1994). The frequency of occurrence and severity of side-effects of immersion virtual reality. Aviation Space & Environmental Medicine, 65(6), 527-530.

Reingold, E. M., Charness, N., Pomplun, M., & Stampe, D. M. (2001). Visual span in expert chess players: Evidence from eye movements. Psychological Science, 12(1), 48-55.

Reingold, E. M., & Loschky, L. C. (in press). Saliency of peripheral targets in gaze-contingent multi-resolutional displays. Behavioral Research Methods, Instruments, and Computers.

Reingold, E. M., & Stampe, D. M. (2002). Saccadic inhibition in voluntary and reflexive saccades. Journal of Cognitive Neuroscience, 14(3), 371-388.

Robinson, G. H. (1979). Dynamics of the eye and head during movement between displays: A qualitative and quantitative guide for designers. Human Factors, 21(3), 343-352.

Rojer, A. S., & Schwartz, E. L. (1990). Design considerations for a space-variant visual sensor with complex-logarithmic geometry, Proceedings of the 10th International Conference on Pattern Recognition (Vol. 2, pp. 278-285).

Rolwes, M. S. (1990). Design and flight testing of an electronic visibility system, Proceedings of the SPIE: The International Society for Optical Engineering (Vol. 1290, pp. 108-119).

Rosenberg, L. B. (1994). The use of virtual fixtures to enhance operator performance in time delayed teleoperation. USAF AMRL Technical Report. AL/CF TR, 0139, 49.

Ross, J., Morrone, M. C., Goldberg, M. E., & Burr, D. C. (2001). Changes in visual perception at the time of saccades. Trends in Neurosciences, 24(2), 113-121.

Rovamo, J., & Iivanainen, A. (1991). Detection of chromatic deviations from white across the human visual field. Vision Research, 31(12), 2227-2234.

Rovetta, A., Sala, R., Cosmi, F., Wen, X., Sabbadini, D., Milanesi, S., Togno, A., Angelini, L., & Bejczy, A. (1993). Telerobotics surgery in a transatlantic experiment: Application in laparoscopy. Proceedings of the SPIE: The International Society for Optical Engineering, 2057, 337-344.

Sandini, G. (2001). Giotto: Retina-like Camera. Retrieved January 7, 2002, from http://www.lira.dist.unige.it/expsetup/svavisca2.html
"specs of GIOTTO the space-variant camera (pdf format)"

Sandini, G., Argyros, A., Auffret, E., Dario, P., Dierickx, B., Ferrari, F., Frowein, H., Guerin, C., Hermans, L., Manganas, A., Mannucci, A., Nielsen, J., Questa, P., Sassi, A., Scheffer, D., & Woelder, W. (1996). Image-based personal communication using an innovative space-variant CMOS sensor, 5th IEEE International Workshop on Robot and Human Communication (pp. 158 -163): IEEE.

Sandini, G., Questa, P., Scheffer, D., Dierickx, B., & Mannucci, A. (2000). A retina-like CMOS sensor and its applications, 2000 IEEE Sensor Array and Multichannel Signal Processing Workshop (pp. 514 -519): IEEE.

Sekuler, A. B., Bennett, P. J., & Mamelak, M. (2000). Effects of aging on the useful field of view. Experimental Aging Research, 26(2), 103-120.

Sere, B., Marendaz, C., & Herault, J. (2000). Nonhomogeneous resolution of images of natural scenes. Perception, 29(12), 1403-1412.

Sharma, R., Pavlovic, V. I., & Huang, T. S. (1998). Toward Multimodal Human-Computer Interface. Proceedings of the IEEE, 86(5), 853-869.

Shin, C. W., & Inokuchi, S. (1994). A new anthropomorphic retina-like visual sensor, The 12th IAPR International Conference on Pattern Recognition, Conference C: Signal Processing (Vol. 3, pp. 345 -348).

Shioiri, S., & Ikeda, M. (1989). Useful resolution for picture perception as a function of eccentricity. Perception, 18, 347-361.

Smith, B. A., Ho, J., Ark, W., & Zhai, S. (2000). Hand eye coordination patterns in target selection. In A. T. Duchowski (Ed.), Proceedings of the Eye tracking research & applications symposium (pp. 117-122). Palm Beach, FL.: ACM.

Spoehr, K. T., & Lehmkule, S. W. (1982). Visual information processing. San Francisco, CA: W.H. Freeman.

Spooner, A. M. (1982). The trend towards area of interest in visual simulation technology, Proceedings of the 4th Interservice/Industry Training Equipment Conference (pp. 205-215).

Stampe, D. M., & Reingold, E. M. (1995). Real-time gaze contingent displays for moving image compression (DCIEM Contract Report W7711-3-7204/01-XSE).

Stelmach, L. B., & Tam, W. J. (1994). Processing image sequences based on eye movements. Proceedings of the SPIE: The International Society for Optical Engineering, 2179, 90-98.

Stelmach, L. B., Tam, W. J., & Hearty, P. J. (1991). Static and dynamic spatial resolution in image coding: An investigation of eye movements. Proceedings of the SPIE: The International Society for Optical Engineering, 1453, 147-151.

Stiefelhagen, R., Yang, J., & Waibel, A. (1997). A model-based gaze-tracking system. International Journal of Artificial Intelligence Tools, 6(2), 193-209.

Szoboszlay, Z., Haworth, L., Reynolds, T., Lee, A., & Halmos, Z. (1995). Effect of field-of-view restriction on rotorcraft pilot workload and performance: Preliminary results. Proceedings of the SPIE: The International Society for Optical Engineering, 2465, 142-153.

Tanaka, S., Plante, A., & Inoue, S. (1998). Foreground-background segmentation based on attractiveness, IASTED International Conference on Computer Graphics and Imaging (pp. 191-194).

Tannenbaum, A. (2000). On the eye tracking problem: a challenge for robust control. International Journal of Robust & Nonlinear Control, 10(11-12), 875-888.

Tharp, G., Liu, A., Yamashita, H., Stark, L., Wong, B., & Dee, J. (1990). Helmet mounted display to adapt the telerobotic environment to human vision, Third Annual Workshop on Space Operations, Automation, and Robotics (SOAR 1989) (pp. 477-481). Washington, DC: NASA.

Thibos, L. N. (1998). Acuity perimetry and the sampling theory of visual resolution. Optometry & Vision Science, 75(6), 399-406.

Thibos, L. N., Still, D. L., & Bradley, A. (1996). Characterization of spatial aliasing and contrast sensitivity in peripheral vision. Vision Research, 36(2), 249-258.

Thomas, M., & Geltmacher, H. (1993). Combat simulator display development. Information Display, 9, 23-26.

Thompson, J. M., Ottensmeyer, M. P., & Sheridan, T. B. (1999). Human factors in telesurgery: Effects of time delay and asynchrony in video and control feedback with local manipulative assistance. Telemedicine Journal, 5(2), 129-137.

To, D., Lau, R. W. H., & Green, M. (2001). An adaptive multiresolution method for progressive model transmission. Presence-Teleoperators & Virtual Environments, 10(1), 62-74.

Tong, H. M., & Fisher, R. A. (1984). Progress Report on an Eye-Slaved Area-of-Interest Visual Display, Report No. AFHRL-TR-84-36. In E. G. Monroe (Ed.), Proceedings of 1984 Image Conference III (pp. 279-294). Phoenix, AR: Air Force Human Resources Laboratory.

Tsumura, N., Endo, C., Haneishi, H., & Miyake, Y. (1996). Image compression and decompression based on gazing area. Proceedings of the SPIE: The International Society for Optical Engineering, 2657, 361-367.

Turner, J. A. (1984). Evaluation of an eye-slaved area-of-interest display for tactical combat simulation, The 6th Interservice/Industry Training Equipment Conference and Exhibition (pp. 75-86).

van Diepen, P. M. J., & Wampers, M. (1998). Scene exploration with Fourier-filtered peripheral information. Perception, 27(10), 1141-1151.

van Erp, J. B. F., & Kappe, B. (1997). Head-slaved images for low-cost driving simulators and unmanned ground vehicles [final report] (TD97-0216). Soesterberg, Netherlands: TNO Human Factors Research Institute.

Viljoen, G. T. (1998). Comparative study of target acquisition performance between an eye-slaved helmet display and unaided human vision. Proceedings of the SPIE: The International Society for Optical Engineering, 3362, 54-65.

Wang, Z., & Bovik, A. C. (2001). Embedded foveation image coding. IEEE Transactions on Image Processing, 10(10), 1397-1410.

Warner, H. D., Serfoss, G. L., & Hubbard, D. C. (1993). Effects of area-of-interest display characteristics of visual search performance and head movements in simulated low-level flight. Final technical report (Sep 89-Jul 92) (AL-TR-1993-0023). Williams Air Force Base, AZ: Armstrong Laboratory.

Watson, B. A., Walker, N., Hodges, L. F., & Worden, A. (1997). Managing level of detail through peripheral degradation: Effects on search performance with a head-mounted display. ACM Transactions on Computer-Human Interaction, 4(4), 323-346.

Weiman, C. F. R. (1990). Video compression via log polar mapping. Proceedings of the SPIE: The International Society for Optical Engineering, 1295(266-277).

Weiman, C. F. R. (1994). Video compression by matching human perceptual channels. Proceedings of the SPIE: The International Society for Optical Engineering, 2239, 178-189.

Werkhoven, P., Sperling, G., & Chubb, C. (1993). The dimensionality of texture-defined motion: A single channel theory. Vision Research, 33(4), 463-485.

Wickens, C. D., & Hollands, J. G. (2000). Engineering psychology and human performance. (3rd ed.). Upper Saddle River, NJ: Prentice Hall.

Wodnicki, R., Roberts, G. W., & Levine, M. D. (1995). A foveated image sensor in standard CMOS technology, Proceedings of the IEEE Custom Integrated Circuits Conference (pp. 357-360).

Wodnicki, R., Roberts, G. W., & Levine, M. D. (1997). A Log-Polar Image Sensor Fabricated in a Standard 1.2-Mu-M Asic Cmos Process. IEEE Journal of Solid-State Circuits, 32(8), 1274-1277.

Woelders, W. W., Frowein, H. W., Nielsen, J., Questa, P., & Sandini, G. (1997). New developments in low-bit rate videotelephony for people who are deaf. Journal of Speech Language & Hearing Research, 40(6), 1425-1433.

Yang, J., Coia, T., & Miller, M. (2001). Subjective evaluation of retinal-dependent image degradations, Proceedings of PICS 2001: Image Processing, Image Quality, Image Capture Systems Conference (pp. 142-147). Springfield, VA: The Society for Imaging Science and Technology.

Yeshurun, Y., & Carrasco, M. (1999). Spatial attention improves performance in spatial resolution tasks. Vision Research, 39(2), 293-306.

Yoshida, A., Rolland, J. P., & Reif, J. H. (1995a). Design and applications of a high-resolution insert head-mounted-display, Proceedings of Virtual Reality Annual International Symposium '95 (Cat. No.95CH35761) (pp. 84-93). Los Alamitos, CA, USA: IEEE Comput. Soc. Press.

Yoshida, A., Rolland, J. P., & Reif, J. H. (1995b). Optical design and analysis of a head-mounted display with a high-resolution insert. Proceedings of the SPIE: The International Society for Optical Engineering(2537), 71-73.

Young, L. R., & Sheena, D. (1975). Survey of eye movement recording methods. Behavior Research Methods & Instrumentation, 7(5), 397-429.

Author Note

This research was supported by grants to Eyal Reingold from the Natural Science and Engineering Research Council of Canada (NSERC) and the Defence and Civil Institute for Environmental Medicine (DCIEM) and to George McConkie from the US Army Federated Laboratory under Cooperative Agreement DAAL01-96-2-0003. We thank William Howell, Justin Hollands and an anonymous reviewer for their very helpful comments on earlier versions of this manuscript.

Correspondence should be sent to Eyal M. Reingold (email: reingold@psych.utoronto.ca), Department of Psychology, University of Toronto, 100 St. George Street, Toronto, Ontario, Canada M5S 3G3.

Footnotes

1. It is common in the multi resolutional display literature to refer to a high-resolution area placed at the point of gaze as an *area of interest* (AOI). However, from a psychological point of view, the term *area of interest* is more often used to indicate the area that is currently being attended. We have attempted to distinguish between these two uses through our terminology.
2. Retinal eccentricity refers to the distance of a retinal location from the center of the fovea. This distance is usually measured in degrees of visual angle. In general, the more eccentric a retinal location, the lower the visual resolution (i.e., acuity) at that location.

Table 1

Examples of Processing and Bandwidth Savings Due to Use of Multi-resolutional Images

Measure	**Savings due to use of multi-resolutional image versus constant resolution image**
3D image rendering time	4-5 times faster (Levoy & Whitaker, 1990; Murphy & Duchowski, 2001, September; Ohshima et al., 1996, p. 108)
Reduced polygons in 3D model	2-6 times fewer polygons, with greater savings at greater eccentricities, and no difference in perceived resolution (Luebke et al., 2000)
Video compression ratio	3 times greater compression ratio in the multi-resolutional image (Geisler & Perry, 1999, p. 422), with greater savings for larger field of view images and same maximum resolution
Number of coefficients used in encoding a wavelet reconstructed image	2-20 times fewer coefficients needed in the multi-resolutional image, depending on the size of the D-AOI and the level of peripheral resolution (Loschky & McConkie, 2000, p. 99)
Reduction of pixels needed in multi-resolutional image	35 times fewer pixels needed in the multi-resolutional image as compared to constant high-resolution image (Sandini et al., 2000, p. 517)

Table 2

Key Questions for Human Factors Research Related to GCMRDs.

Question	References
Can we construct just undetectable GCMRDs that maximize savings in processing and bandwidth while eliminating perception and performance costs?	(Geri & Zeevi, 1995; Loschky, 2002; Luebke et al., 2000; Peli & Geri, 2001; Sere et al., 2000; Yang et al., 2001)
What are the perception and performance costs associated with removing above-threshold peripheral resolution in detectably degraded GCMRDs?	(Geri & Zeevi, 1995; Kortum & Geisler, 1996b; Loschky, 2002; Loschky & McConkie, 2000; Loschky & McConkie, 2002; Parkhurst et al., 2000; Peli & Geri, 2001; Reingold & Loschky, in press; Shioiri & Ikeda, 1989; Watson et al., 1997; Yang et al., 2001)
What is the optimal resolution drop-off function that should be used in guiding the construction of GCMRDs?	(Geri & Zeevi, 1995; Loschky, 2002; Luebke et al., 2000; Peli & Geri, 2001; Sere et al., 2000; Yang et al., 2001)
What are the perception and performance costs and benefits associated with employing continuous versus discrete resolution drop-off functions in still versus full-motion displays?	(Baldwin, 1981; Browder, 1989; Loschky, 2002; Loschky & McConkie, 2000, experiment 3; Reingold & Loschky, in press; Stampe & Reingold, 1995)
What are the perception and performance costs and benefits related to the shape of the D-AOI (ellipse vs. circle vs. rectangle) in discrete resolution drop-off GCMRDs?	(No empirical comparisons to date)
What is the effect, if any, of lateral masking on detecting peripheral resolution drop-off in GCMRDs?	(Loschky, 2002; Peli & Geri, 2001; Yang et al., 2001)
What is the effect, if any, of attentional cuing on detecting peripheral resolution drop-off in GCMRDs?	(Yeshurun & Carrasco, 1999)
What is the effect, if any, of task difficulty on detecting peripheral resolution drop-off in GCMRDs?	(Bertera & Rayner, 2000; Loschky & McConkie, 2000, Exp. 5; Pomplun et al., 2001)
Do older users of GCMRDs have higher resolution drop-off thresholds than younger users?	(Ball et al., 1988; Sekuler, Bennett, & Mamelak, 2000)
Do experts have lower resolution drop-off thresholds than novices when viewing multi-resolutional images relevant to their skill domain?	(Reingold et al., 2001)
Can we utilize a hue resolution drop-off algorithm that is just imperceptible in the construction of GCMRDs?	(Watson et al., 1997, Exp. 2)
What are the perception and performance costs and benefits associated with employing the different methods of producing multi-resolutional images?	(See Table 3)
How do the different methods of moving the D-AOI (i.e., gaze-, head-, and hand-contingent, and predictive movement) compare in terms of their perception and performance consequences?	(No empirical comparisons to date)
What are the effects of a systematic increase in update delay on different perception and performance measures?	(Draper et al., 2001; Frank et al., 1988; Grunwald & Kohn, 1994; Hodgson et al., 1993; Loschky & McConkie, 2000, Exp. 1 & 6; McConkie & Loschky, in press; Reingold & Stampe, 2002; Turner, 1984; van Diepen & Wampers, 1998)
Is it possible to compensate for poor spatial and temporal accuracy/resolution of D-AOI update by decreasing the magnitude and scope of peripheral resolution drop-off?	(Loschky & McConkie, 2000, Exp. 6)

Table 3

Methods of Combining Multiple Resolutions in a Single Display

Method of making images multi-resolutional	**Suggested application area(s)**	**Basis for resolution drop-off**	**References**
Rendering 2D or 3D models w/ multiple levels of detail and/or polygon simplification	Flight simulators, VR; medical imagery; image transmission	Retinal acuity or CSF x eccentricity &/or velocity &/or binocular fusion &/or size	(Levoy & Whitaker, 1990; Luebke et al., 2000; Murphy & Duchowski, 2001, September: Luebke et al., 2002; Ohshima et al., 1996; Reddy, 1998; Spooner, 1982; To et al., 2001)
Projecting image to viewable monitors	Flight simulator, driving simulator	(No vision behind the head)	(Kappe et al., 1999; Thomas & Geltmacher, 1993; Warner et al., 1993)
Projecting 1 visual field to each eye	Flight simulator (head-mounted display)	Unspecified	(Fernie, 1995; Fernie, 1996)
Projecting D-AOI to 1 eye, periphery to other eye	Indirect vision (head-mounted display)	Unspecified (emphasis on binocular vision issues)	(Kooi, 1993)
Filtering by retina-like sampling	Image transmission	Retinal ganglion cell density and output characteristics	(Kuyel et al., 1999)
Filtering by "super pixel" sampling and averaging	Image transmission, video teleconferencing, remote piloting, telemedicine	Cortical magnification factor or eccentricity-dependent CSF	(Kortum & Geisler, 1996a; Kortum & Geisler, 1996b; Yang et al., 2001)
Filtering by low-pass pyramid w/ contrast threshold map	Image transmission, video teleconferencing, remote piloting, telemedicine VR, simulators	Eccentricity-dependent CSF	(Geisler & Perry, 1998; Geisler & Perry, 1999; Loschky, 2002)
Filtering by Gausian sampling w/ varying kernel size w/ eccentricity	Image transmission	Human Vernier acuity drop-off function (point spread function)	(Geri & Zeevi, 1995)
Filtering by Wavelet transform with scaled coefficients w/ eccentricity or discrete bands	Image transmission, video teleconferencing, VR	Human minimum angle of resolution x eccentricity function or empirical trial and error	(Duchowski, 2000; Duchowski & McCormick, 1998; Frajka et al., 1997; Loschky & McConkie, 2002; Wang & Bovik, 2001)
Filtering by log-polar or complex log-polar mapping algorithm	Image transmission, video teleconferencing, robotics	Human retinal receptor topology or Macaque retino-cortical mapping function	(Basu & Wiebe, 1998; Rojer & Schwartz, 1990; Weiman, 1990; Weiman, 1994; Woelders et al., 1997)
Multi-resolutional sensor (log-polar or partial log polar)	Image transmission, video teleconferencing, robotics	Human retinal receptor topology & physical limits of sensor	(Sandini, 2001; Sandini et al., 1996; Sandini et al., 2000; Wodnicki, Roberts, & Levine, 1995; Wodnicki, Roberts, & Levine, 1997)

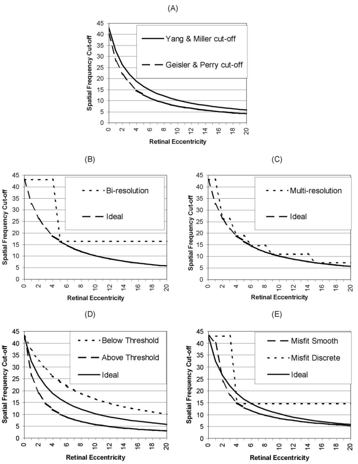
Figure Caption

Figure 1. Gaze-contingent multi-resolutional imagery. (A) A constant high-resolution image. (B) Several consecutive gaze locations of a viewer who looked at this image; with the last in the series indicated by the cross mark. (C) A discrete drop-off, bi-resolutional image, having two levels of resolution, high and low. The high-resolution area is centered on the viewer's last gaze position in (A). (D) A continuous drop-off multi-resolutional image, with the center of high-resolution at the above viewer's last gaze position.

Figure 2. Visual resolution drop-off as a function of retinal eccentricity and spatial frequency. (A) Two different contrast sensitivity cut-off functions from Yang and Miller (in (Loschky, 2002) and Geisler and Perry (1998). The functions assume a constant Michaelson contrast ratio of 1.0 (maximum) and show the contrast threshold as a function spatial frequency for each retinal eccentricity in degrees visual angle. Viewers should be unable to discriminate spatial frequencies *above* the line for any given eccentricity in a given function (i.e., those frequencies are *below* perceptual threshold). Note the overall similarity of the two functions, each of which is based on data from several different psychophysical studies using grating stimuli. (The small differences between the plots can be characterized as representing a band-pass vs. low-pass foveal CSF, but could be reduced by changing some parameter values). For simplicity, the Yang and Miller model is designated the "ideal" in the following figures. (B) The spatial frequency cut-off profile of a discrete drop-off, bi-resolutional display matching an ideal sensitivity cut-off function. Although much spatial frequency information is dropped out of the bi-resolutional image, it should be imperceptible because the spatial frequency information removed is always below threshold. (C) The profile of a multi-resolution display with many discrete bands of resolution. It is similar to the bi-resolution display except that the number of resolution bands is greater, thus increasing the potential processing and/or

bandwidth savings. (D) A comparison of two continuous drop-off multi-resolutional displays with the ideal. One drop-off function is below threshold, i.e., imperceptible, but contains more information than necessary. The other function is above threshold and will likely cause perceptual difficulties. The two functions differ from the ideal on only a single parameter, thus making it relatively easy to determine the ideal fit. (E) Two multi-resolutional drop-off schemes that do not match the ideal. One is a continuous drop-off function. Note that the function is below threshold at low eccentricities, but above threshold at higher eccentricities. It may be difficult to determine the particular parameter(s) that must be varied in order to correct the lack of fit, or whether a function of a different form is needed. The second is a discrete drop-off (bi-resolutional) step function. Similar to the continuous drop-off function, it is above and below threshold at different eccentricities. However, it should be relatively easy to fix the lack of fit by varying a single parameter, either the eccentricity at which the drop-off occurs, or the level of drop-off at a given eccentricity.


(A)
(B)
(C)
(D)


(A)
Spatial Frequency Cut-off
Yang & Miller cut-off
Geisler & Perry cut-off
Retinal Eccentricity
0
5
10
15
20
25
30
35
40
45
0
2
4
6
8
10
12
14
16
18
20
(B)
Spatial Frequency Cut-off
Bi-resolution
Ideal
Retinal Eccentricity
(C)
Spatial Frequency Cut-off
Multi-resolution
Ideal
Retinal Eccentricity
(D)
Spatial Frequency Cut-off
Below Threshold
Above Threshold
Ideal
Retinal Eccentricity
(E)
Spatial Frequency Cut-off
Misfit Smooth
Misfit Discrete
Ideal
Retinal Eccentricity