PRESENCE: Teleoperators and Virtual Environments, Volume 15, Number 1, February 2006, pages 16–32.

COVER PAGE

TITLE:

The Effects of Text Drawing Styles, Background Textures, and Natural Lighting on Text Legibility in Outdoor Augmented Reality

Joseph L. Gabbard
jgabbard@vt.edu
Systems Research Center
Virginia Tech
Blacksburg, VA USA

J. Edward Swan II
swan@cse.msstate.edu
Computer Science and Engineering Department
Mississippi State University
Starkville, MS USA

Deborah Hix
hix@vt.edu
Systems Research Center
Virginia Tech
Blacksburg, VA USA

Unusual fonts – need "Lucida Console" font (or similar monospaced font) for Equations

PRESENCE: Teleoperators and Virtual Environments, Volume 15, Number 1, February 2006, pages 16–32.

Abstract

A challenge in presenting augmenting information in outdoor augmented reality (AR) settings lies in the broad range of uncontrollable environmental conditions that may be present, specifically large-scale fluctuations in natural lighting and wide variations in likely backgrounds or objects in the scene. In this paper, we motivate the need for research on the effects of text drawing styles, outdoor background textures, and natural lighting on user performance in outdoor AR. We present a pilot study and a follow-on user-based study that examined the effects on user performance of outdoor background textures, changing outdoor illuminance values, and text drawing styles in a text identification task using an optical, see-through AR system. We report significant effects for all these variables, and discuss user interface design guidelines and ideas for future work.

1 Introduction

By providing direct, heads-up access to information correlated with a user's view of the real world, augmented reality (AR) has the potential to redefine the way information is presented and accessed. A challenge in presenting augmenting information, particularly in outdoor, optical see-through AR applications, lies in the broad range of uncontrollable environmental conditions that may be present, specifically large-scale fluctuations in natural lighting and wide variations in likely backgrounds or objects in the scene. In many cases, a carefully designed AR user interface may be easily legible under some lighting and background conditions, and moments later become totally illegible. Since lighting and background conditions may vary from moment to moment in dynamic AR usage contexts, there is a need for basic research to understand the relationship among natural lighting, outdoor background textures, real-world objects, and associated AR graphics.

Table 1. Representative lighting levels of a typical outdoor scene (Halsted, 1993).

Outdoor Condition	**Illuminance (lux)**
Bright sun	50,000 – 100,000
Hazy day	25,000 – 50,000
Cloudy bright	10,000 – 25,000
Cloudy dull	2,000 – 10,000
Very dull	100 – 2,000
Sunset	1 – 100
Full moon	0.01 - 0.1
Starlight	0.001 - 0.001

One obvious strategy is to make the AR graphics *active*, so they sense and adapt in real time to changing environmental conditions. Natural lighting conditions can vary by orders of magnitude in very short time periods. For example, daytime outdoor illuminance can vary from as little as 1 lux to 100,000 lux. Halstead (1993) presents a detailed discussion of typical lighting conditions and identifies ranges of light measurements typically found in outdoor environments (see Table 1). These ranges vary from a poorly lit dark street, to indoor lighting, to the brightest sunny day.

In the past decade, AR research efforts have successfully tackled several challenging hardware integration problems, so that current AR systems are beginning to function robustly. As a result, the AR field is now to the point where meaningful, systematic human-computer interaction (HCI) research can be conducted and applied. Despite the fact that this technology can fundamentally change the way we visualize, use, and interact with information, very little HCI work has been done specifically in AR. We conducted a survey (Swan & Gabbard, 2005) of four primary publishing venues for AR research, which comprised a corpus of 1104 papers. This survey revealed only 21 user-based studies of AR systems over a period of 12 years.

The work described in this paper has been motivated by the empirical results of our application of usability engineering activities to the design and development of the Battlefield Augmented Reality System (BARS) (Gabbard et al., 2002; Livingston et al., 2002). Usability engineering activities can be used to determine *what* information should be presented to users, or, for example, *where* information should be presented to users. Of particular interest in our work, however, is *how* augmenting information should be visually presented for optimal usability. The focus of the work reported here is studying the effects of environmental conditions on AR text legibility, with a motivation of designing text drawing styles that are optimal for dynamic environmental conditions.

Since our work relies on dynamic outdoor viewing conditions, we created a research testbed that employs an optical "see-through" display (as opposed to a video-see-through display) so that users see through the display directly to the real world (as opposed to viewing LCD screens with real-time camera video feeds). This see-through approach to AR maximizes the extent of presence and degree of "realspace imaging" (Naimark, 1991) by maximizing the visual fidelity of outdoor background textures and by allowing a large range of natural light to the user's eye.

Section 2 contains a broad discussion of related work, and Section 3 describes a pilot study conducted in preparation for a more comprehensive user-based experiment. Section 4 describes this user-based experiment, including a discussion of the experimental task and stimuli. Section 5 then presents hypotheses, Section 6 describes results, Section 7 discusses related user interface design guidelines, and Section 8 offers ideas for future work.

This paper repeats and extends results we published at the IEEE Virtual Reality 2005 conference (Gabbard et al., 2005). Relative to this prior publication, the current paper makes two additional contributions. First, it greatly expands the description of the problem domain, experiment, analysis, and results. Second, it details a pilot study along with an analysis of how it informed the experiment's design. In addition to presenting the pilot study and experiment, this paper also constitutes a case study of experimental design in AR; readers contemplating experiments in augmented or virtual reality may find this meta-information helpful.

2 Related Work

Although it is well-known (at least anecdotally) that color and texture of the background environment have a direct bearing on the visibility and legibility of text in AR systems, very little research has investigated and quantified these effects. To our knowledge, the experiment described in this paper is the first attempt to identify the range of outdoor conditions over which information from an optical see-through AR display can be effectively observed.

Several researchers have empirically examined the effects of natural lighting on user performance in AR. (Feiner, MacIntyre, Höllerer & Webster, 1997) reported that AR displays are very difficult to read in bright sunlight. Others have suggested that this commonly noted problem can be mitigated by using high contrast between the outdoor background texture and the display contents (e.g., text and graphics), which helps a user to view the display in sunlight (Gleue & Dähne, 2001; Azuma, Bruce, Hoff, Neely & Sarfaty, 1998).

Other attempted solutions to the outdoor lighting problem have ranged from dimming the real-world light that reaches the eye by using a sunglass effect to enhance visibility of the AR display (Pasman, 1997; Azuma, Hoff, Neely & Sarfaty, 1999), to creating novel displays such as the virtual retinal display (VRD), which creates images that can be easily seen in ambient room light and in ambient daylight (Azuma, 1999) by writing directly on the user's retina with a low-powered laser.

Many AR systems, such as the online maintenance system described by Lipson, Shpitalni, Kimura & Goncharenko (1998), for example, depict labels as white objects with solid black backgrounds. Although such backgrounds are possible for video AR displays, they cannot be used for optical see-through AR displays because see-through displays are additive, which means the color black is effectively transparent and thus does not obscure the background view. The AR Quake system, developed by Piekarski and colleagues (Piekarski & Thomas, 2002; Thomas et al., 2002), modified the textures of monsters in the AR Quake game to make them visible against the real world, and also provided recommendations on text color given ambient outdoor lighting conditions.

Numerous HCI studies have examined the use of color in traditional 2D user interfaces. As long as three decades ago, researchers studied the contribution of color coding in comparison to other graphical encoding techniques (e.g., Christ, 1975; Cleveland & McGill, 1985), while others attempted to determine the optimum number of colors to use in displays, as well as which colors should be used (e.g., Cahill & Carter, 1976; Carter, 1982; Smallman & Boynton, 1990).

Harrison and Vicente (1996) describe an "anti-interference" font, designed to produce transparent 2D menus superimposed over different GUI background content, such as applications overlaid on the desktop. The work also includes an empirical evaluation of the effect of varying transparency levels, visual interference produced by different types of background content, and performance of anti-interference fonts on text menu selection tasks. While we did not use the anti-interference font for this study, we are planning on using a version of it for a follow-on study. Similar efforts have been addressed in the television and entertainment business, since there is often a need to overlay text-based information onto a real-world video scene. However, in most cases, the overlaid text is done post-production (i.e., not in real-time) with *a priori* knowledge of the exact scene content and layout. In the rare cases where real-time augmentation is performed, there are known camera angles, with (generally) known scene content, so that a set of working assumptions can be used to facilitate real-time design and placement of overlaid information. The most common examples are televised sporting events, such as football and basketball.

3 Pilot Study: Controlling Lighting and Backgrounds using the CAVE™

One of our first concepts for exploring the effects of outdoor background texture and illuminance on the legibility of augmented text involved placing an AR user in an immersive VR environment such as a CAVE. In this *AR-within-VR* concept, the AR user would wear an optical, see-through display and observe AR graphics registered with VR graphics – representing the real world – projected onto the CAVE walls. We anticipated several benefits from such a system. First, we hypothesized that using photographic images projected onto the CAVE walls would result in a strong sense of presence. Second, we could *control* (as oppose to just measure) the amount of ambient and background lighting. Third, we could use the same head tracking for both the AR and VR systems, which would theoretically allow us to perfectly register the AR graphics[1].

[1] Because of this potential for perfect AR registration, the AR-within-VR concept may also form a useful testbed for studying the effects of AR tracking errors (e.g., MacIntyre and Coelho, 2000). In such a testbed, a controlled amount of registration error could be added to the zero baseline error.

We developed a prototype of this concept and conducted a pilot study to assess its effectiveness. We used the Virginia Tech CAVE, and developed an application-based context involving navigation and wayfinding in an urban environment. We photographed and modeled five local intersections, ranging from a visually complex, paved, downtown scene to a simple, rural gravel road. We created our virtual real world with DIVERSE (Kelso, Arsenault, Satterfield & Kriz, 2003), an API created at Virginia Tech for developing interactive VR applications. In most VR scenes, the rendering software uses a geometric model of the virtual world to draw scene objects. We presented our virtual real-world scenes to each user in the CAVE using a large simple geometric cylinder around which we wrapped a panoramic photograph such as that shown in Figure 1. That is, we did not create models of buildings, roads, trees, etc., and composite them in the virtual worlds, but instead created a simple (albeit compelling) immersive cylinder. This approach was *much* faster than modeling an entire street scene (although it did not allow us to experiment with registered AR graphics), and potentially provided a compelling experience of presence in the CAVE as a substitute for the real world normally seen through an AR display.

FIGURE 1 HERE

We then conducted a pilot study. This study employed three users performing a basic visual identification task that consisted of locating an augmented arrow registered with one of the three street intersections. The qualitative results indicated that the CAVE gave a compelling experience of presence, and the realism of the cylindrical scene was adequate to examine the effect of background and scene lighting on the legibility of various augmenting drawing styles. However, we also identified several factors that would significantly limit our research. The most profound was that the CAVE projectors (and even newer high-intensity DLP projectors) do not provide enough light to simulate an outdoor environment. In fact, almost any augmenting text presented on the AR display was legible *independent of the virtual background*. Using a light meter, we determined that the maximum amount of light that could be created using our current CAVE configuration is 90 lux. Specifically, we flooded each wall and the floor with a completely white scene and measured 20 lux against the wall, 40 lux in the CAVE center, and a max of 90 lux midway between the CAVE center and the CAVE wall – which approximates that of an outdoor sunset, or an indoor corridor and living room (see Table 1).

We attempted to flood the CAVE room with additional lighting to approach realistic outdoor lighting levels, but doing so washed out the rear-projected CAVE walls to the point that none of the virtual real-world scene was visible. Our last attempt was to reduce the illumination on the display so that the *relative* illumination between the augmenting text and the virtual background approximated the *relative* illumination between a fully illuminated AR display and a typical outdoor scene. However, we felt that this approach would not produce results that could be generalizable to real-world AR usage settings, since we would be approximating or simulating more real-world factors (i.e., first simulating the real world, and then simulating relative illuminance). We did determine however, that the CAVE would be an innovative and well-suited approach to support a controlled environment for studying AR usage at night, dawn, dusk, or indoor AR.

Based on this pilot study, we decided that instead of attempting to control the ambient scene lighting (which was very difficult since we were aiming to create an artificial outdoor scene), we would use “real” outdoor lighting and *measure* the amount of light present at any given time.

Under this approach, our user-based studies are restricted to lighting conditions that fall within a given range of light measurements. However, despite our wishes and efforts to control the lighting, we felt that measuring the light and restricting our studies to a fixed set of lighting levels was a scientifically valid choice, especially since it will produce more generalizable results for use in real outdoor AR usage contexts. We chose to perform studies with illuminance levels between 2,000 lux and 25,000 lux, or lighting no darker than that of a "cloudy dull" day and no brighter than that of a "cloudy bright" day (as defined in Table 1).

As a result of piloting we identified a set of key tradeoffs for our research agenda: controlling the amount of ambient light versus measuring it, controlling distances to objects versus measuring, and controlling the complexity of the scene versus measuring.

Controlling the amount of ambient light while retaining realistic outdoor illuminance levels was not possible, and as such we decided that all subsequent studies would have to be performed outdoors. Unfortunately, from prior experience we have learned that it is very difficult to incorporate accurate head tracking in outdoor studies, and moreover, conducting a study under varying outdoor weather conditions is hard on equipment, experimenters, and users alike!

Controlling the distance to objects in an outdoor study is difficult as well, and in most cases requires the user's position to be moved, as opposed to moving an object. Moving users around (and thus moving equipment around) in the middle of a study is very difficult, if not intractable, for our studies. While our stated alternative was to measure the distance to various objects, we felt that it was more important to proceed along a research path that let us control the distances between users and outdoor background textures. Our experience also indicated that the complexity or visual makeup of an outdoor scene is also very difficult to control, as there inevitably are passing cars, pedestrians and general curiosity seekers.

To address the lighting tradeoffs and environmental challenges, we opted to conduct the study in a well-lit greenhouse on the Virginia Tech campus – ensuring ample lighting at both the user's position as well as at the outdoor background texture's position. We located an underutilized greenhouse that allowed us to control the complexity of the scene and eliminate the opportunity for random persons or automobiles to alter the scene. Lastly, the greenhouse allowed the experimenter, equipment, and users to be protected from some climatic elements.

4 The Empirical User-Based Study

Through previous user studies (e.g., Livingston et al., 2003; Hix et al., 2004), we have observed that *reading* text in outdoor AR environments is not only difficult, but more importantly, a necessary precursor to many other typical user tasks. In addition, as discussed above, outdoor background textures, distance from the user to these backgrounds, and ambient illuminance have a noticeable effect on text legibility. Moreover, we have also observed that many AR displays are built with fixed focal lengths, meaning that the user's eyes must focus at a fixed distance (e.g., two meters for the Sony Glasstron) independently of the *virtual* distance to the augmenting text or the actual distance to the outdoor background texture. Based on these observations, we conducted a study that examined the effects on user performance of outdoor background textures, changing outdoor illuminance values, text drawing styles, and distance from the user to the background, in text identification task (Christ, 1975). We captured user response time, user error, and measured, and controlled for, variance in natural illumination levels. Table 2 summarizes the variables we systematically examined.

Table 2. Summary of variables studied in experiment.

Independent Variables

subject	18	*random variable*
distance	3	near (*1 meter*), medium (*2 meters*), far (*4 meters*)
outdoor background texture (Figure 2)	6	pavement, granite, red brick, sidewalk, foliage, sky
text drawing style (Figure 2)	6	*static*: billboard, red, green *active*: complement, maximum HSV complement, maximum brightness contrast
repetition	4	1, 2, 3, 4

Dependent Variables

response time	in *milliseconds*
error	0 (*correct*), 1 (*incorrect*)
ambient illuminance	in *lux*

4.1 User Task and Experimental Setup

We decided to examine user performance of a text identification task. We chose to make the identification task text-based (as opposed to icons-, lines-, or bitmap-based), since it can be argued that text is one of the most fundamental graphical elements in any user interface. Further, since very little work has been done to research how best to display any graphical elements in AR imposed on top of a complex outdoor background texture, we suggest that researching text specifically will provide insight into how to design and construct more complex AR user interface elements.

Specifically, we designed a task that abstracted the kind of short reading tasks, such as reading text labels, which are prevalent in many AR applications. For this study, we purposefully designed the experimental task to be a low-level perceptual identification task, as opposed to a higher-level cognitive task. Since we chose not to address the notion of semantics (e.g., cognitively understanding the contents/meaning of the text), we were not concerned with users' semantic interpretation of the data, but simply whether or not users could quickly and accurately *read* information (i.e., text legibility). Our basic motivation is that if the user cannot see and read text, then certainly the user cannot understand its meaning.

The user's task was to identify (find and read) a single numeral presented in a text string of randomized, distracting letters. Specifically, each text string contained one, and only one, numeral – either 4, 5, or 6. For each trial, users entered the numeral, using the numeric keypad of a standard extended keyboard, by pressing a key in the middle row of numbers (i.e., 4 through 6). Users entered a 0 if either they could not find the text string at all (i.e., the string was effectively invisible), or if they could not see a numeral in the text string.

All text strings were presented in the middle of the user's field of view, and users were instructed to minimize their head movement. Since outdoor background textures were placed directly in front of each user (using large posters, as described below in Section 5.2.1), all text strings were effectively presented in the middle of each outdoor background texture. This allowed users to concentrate on reading, and not searching for, the target text strings.

We measured the amount of ambient lighting (illuminance) at the user's position, both to quantify the effect of varying ambient illumination on user task performance, and to ensure that ambient illuminance fell into a pre-determined acceptable range. We used a Cooke Corporation Cal-Light 400 light meter to measure illuminance at the user's position. We only ran subjects during the middle of the day, when the sky was clear to at most partly cloudy. During the study, the range of measured illuminance values was never outside our acceptable range of 2000 to 25,000 lux.

Our image generator was a Pentium M 1.6 GHz computer (comparable to a Pentium IV 2.4 GHz) with 772 megabytes of RAM and an NVidia GeForce4 4200 Go graphics card generating monoscopic images. The computer ran under the Linux Mandrake operating system. We used the same computer to collect user data. For the display device, we used a Sony Glasstron PLM A55 biocular, optical see-through display, at NTSC resolution. Because our user task did not require world-centered graphics, we did not use a tracking device.

4.2 Independent Variables

4.2.1 Outdoor Background Texture

We chose six outdoor background textures to be representative of commonly-found objects in an urban setting, which is the intended location for BARS use (Gabbard et al., 2002): 'pavement', 'granite', 'red brick', 'sidewalk', 'foliage', and 'sky' (Figure 2). To display these textures in the greenhouse, as well as to easily control the distance between each background and the user, we created large (40" x 60") posters of each background texture. We captured the textures by taking high-resolution digital photographs, except for 'sky', which we generated using an Adobe Photoshop cloud-rendering algorithm. We made large, matte-finished prints of each texture, which we mounted onto foamcore posterboard. We scaled the prints so that texture features

FIGURE 2 HERE

were life-sized, e.g., bricks on the poster were the same size as the actual bricks on the building we photographed, leaves were the same size as the actual leaves on the tree we photographed, and so forth. Posters were set up on tripods at specific distances from the user during an experimental session.

4.2.2 Text Drawing Style

We created six text drawing styles (Figure 2) based on previous research in typography, color theory, and HCI text design (Gabbard, 2003). Three of the text styles (‘billboard’, ‘red’, and ‘green’) were *static*, meaning that the text color did not change, and three of the text styles were *active,* meaning that the text color changed depending upon the outdoor background texture presented on the current poster. We wanted to examine both static and active text styles to determine whether or not active styles would result in better user performance (e.g., faster text identification) than static styles. For the active text drawing styles we used the average pixel color value (Figure 2) of each outdoor background texture (calculated from the digital images) as an input to the algorithms described below. We used a sans serif font (Helvetica), and presented the text at a size that appeared approximately two inches tall at a distance of two meters. Text size did not vary during the experiment.

Billboard (static): We designed the 'billboard' style using a saturated blue text ([0, 0, 255] in RGB space) on a white ([255, 255, 255] in RGB space) rectangular background. The solid white background for 'billboard' nearly completely occluded the background texture, resulting in easily readable text regardless of the background. We chose a saturated white background both to fully occlude the background texture and because white is a commonly used background color for GUIs and print media. We chose blue text instead of black text (as most GUI-based reading studies use, or as often used in newspaper and other print media) because black is transparent in optical see-through displays.

Red (static), Green (static): Our choice of 'red' and 'green' as conditions was based on the physiological fact that cones in the human eye are most sensitive to long wavelength (red) and medium wavelength (green) (Hecht, 1987; Williamson & Cummins, 1983). Both colors were fully saturated in RGB space; 'red' was [255, 0, 0] and 'green' was [0, 255, 0]. Of course, this means that the spectrum of red and green generated by our display device was unlikely to be the spectrum to which our subjects' medium- and long-wavelength cones were maximally sensitive; however, we were also motivated by replicating the sort of easy, fully-saturated color choices that are commonly used for GUIs.

*Complement (active)***:** We chose the 'complement' text drawing style based on color theory and graphic arts guidelines asserting that strong color contrast supports efficient text reading (Legge, Parish, Leubker & Wurm, 1990; Knoblauch & Arditi, 1994). The 'complement' style (also called 'inverse' in computer graphics) is defined in RGB space as shown in the equation below, where (`R`, `G`, `B`) is the average pixel background color, and (`R′`, `G′`, `B′`) is the resulting text drawing style color.

```
R′ = ABS (R - 255)
G′ = ABS (G - 255)
B′ = ABS (B - 255)
```

A potential problem of using the complement in this setting is that some of our backgrounds' average pixel color fell into the midrange of the RGB scale; that is, their respective `R`, `G`, and `B` values were closer to 128 than 0 or 255. As a result, the complement of these colors does not provide strong color contrast as compared to the source background, especially for 'pavement' and 'granite' backgrounds. Even so, we wanted to study the complement, because it is prevalent in graphic arts and 2D GUIs.

*Maximum HSV Complement (active)***:** Our observations of problems with the 'complement' style motivated us to design this text drawing style. To achieve further contrast, we first designed a 'saturated complement' style to fully saturate the complementary color. However, pilot testing suggested that saturating the complement of our chosen backgrounds resulted in mostly dark (approaching black) text, which becomes increasingly difficult to read using optical see-through AR displays. We then refined the style with the following goals: retain the notion of employing color complements, account for the fact that optical see-through AR displays cannot present the color black, and use the HSV color model (Foley, van Dam, Feiner, Hughes & Phillips, 1993) so we could easily and independently modify saturation. The result is

```
H′ = (H + 180) - (((H + 180) DIV 360) * 360)
S′ = |100% - S|
V′ = 100%,
```

where (`H`, `S`, `V`) is the average pixel background color in HSV space, and (`H′`, `S′`, `V′`) is the resulting text drawing style color. This algorithm rotates `H` by 180 degrees, calculates the complement of `S` (defined as |100% – `S`|), and maximizes `V`, the value (brightness) of the color.

Maximum Brightness Contrast (active): We wanted to create a text drawing style that maximized the perceived brightness contrast between the augmented text and the outdoor background texture. This style is based on MacIntyre's maximum luminance contrast technique (MacIntyre, 1991; MacIntyre & Cowan, 1992). Like MacIntyre, our algorithm calculates within the *Commission Internationale de l' Éclairage* (CIE) XYZ color model (Foley et al., 1993), because the `Y` basis function models human luminance sensitivity. To calculate this style, we had to convert our known average pixel colors from RGB space to CIE XYZ space. This process requires colorimeter equipment to physically measure the XYZ basis functions, which was impractical for our experimental setup. Algorithmically converting between RGB colors and XYZ colors requires assuming values for two parameters: object size (measured in degrees of subtended field of view on the retina), and white point. The XYZ basis functions are only defined for 2° objects and 10° objects, and since 2° is much closer to the size of our text strings than 10°, we used the 2° standard. Because we ran our experiments outdoors with natural lighting, we used CIE standard illuminant D65 as our white point, since D65 represents sunlight. More details on these parameters and values are available in (Wyszecki & Stiles, 1982 and Foley et al., 1993).

Our algorithm for calculating the 'maximum brightness contrast' is described by the equation below. It only manipulates the CIE `Y` value. Let (`X`, `Y`, `Z`) be the CIE values of the background's average pixel color, and (`X′`, `Y′`, `Z′`) be the resulting text drawing style color. The algorithm

PRESENCE: Teleoperators and Virtual Environments, Volume 15, Number 1, February 2006, pages 16–32.

maximizes `Y′` if `Y` is less than or equal to 0.5, otherwise it minimizes `Y′`. This maximization (minimization) is the value closest to 1.0 (0.0) subject to the constraint that each component of the resulting (`R′`, `G′`, `B′`) tuple is in the valid range of 0 to 255.

$$
\begin{aligned}
X' &= X \\
Y' &= \begin{cases}
\max\limits_{i \in [0..1]} i : \begin{cases} 0 \le R(X',Y',Z') \le 255 \\ 0 \le G(X',Y',Z') \le 255 \\ 0 \le B(X',Y',Z') \le 255 \end{cases} & \text{if } Y \le 0.5 \\[2ex]
\min\limits_{i \in [0..1]} i : \begin{cases} 0 \le R(X',Y',Z') \le 255 \\ 0 \le G(X',Y',Z') \le 255 \\ 0 \le B(X',Y',Z') \le 255 \end{cases} & \text{if } Y > 0.5
\end{cases} \\
Z' &= Z
\end{aligned}
$$

4.2.3 Other Independent Variables

As summarized in Table 2, we varied the distance from the user to the outdoor background texture (poster) between three different levels: ‘near’ (one meter), ‘medium’ (two meters), and ‘far’ (four meters). We chose these distances based on the fixed accommodative focal distance of our Glasstron PLM A55 display (specifically, two meters), so that one condition would present augmenting text *beyond* the background (display focal distance of two meters, poster at one meter from user), the second condition would present augmenting text *on* the background (display focal distance of two meters, poster at two meters from user), and the third condition would present augmenting text *in front of* the background (display focal distance of two meters, poster at four meters from user). We call these conditions ‘near’, ‘medium’, and ‘far’, respectively.

In addition, users saw four repetitions of each combination of independent variables.

FIGURE 3 HERE

4.3 Dependent Variables

Also as summarized in Table 2, we collected values for three dependent variables: response time, errors, and ambient illuminance at the user's position. For each trial, our custom software recorded both the user's four-alternative forced choice (0, 4, 5, or 6), and the user's response time. Whenever the distance or outdoor background texture changed, we measured and recorded the illuminance at the user's position.

4.4 Experimental Design and Users

Figure 3 describes how we counterbalanced presentation of stimuli to users. We used a factorial nesting of independent variables for our experimental design, which varied in the order they are listed in Table 2, from slowest (subject) to fastest (repetition). When the distance variable changed, experimenters had to move the background posters to a different set of tripods, and when the outdoor background texture changed, experimenters had to replace the background posters on all the tripods. Because these activities were cumbersome and time-consuming, we wanted to minimize the number of times they occurred, and therefore distance has the slowest variation rate, and outdoor background texture has the next slowest rate. We collected a total of 7776 response times and errors (18 subjects * 3 distances * 6 outdoor background textures * 6 text drawing styles * 4 repetitions), and 324 illuminance measurements (18 subjects * 3 distances * 6 outdoor background textures). We counterbalanced presentation of the independent variables using a combination of Latin squares (Box, Hunter & Hunter, 1978) and random permutations (Gabbard, 2003). Each subject saw all levels of each independent variable, so all variables were within-subject.

Eighteen subjects participated, twelve males and six females, ranging in age from 20 to 31. All volunteered and received no compensation. We screened the subjects, via self-reporting, for color blindness and visual acuity. Over half of our users (11 out of 18) wore glasses or contact lenses, and one had a slight green-red color deficiency. Twelve of our 18 users reported being moderate to heavy computer users (between three and eight hours of use per day), and about half were familiar with virtual reality systems either from class or through direct experiences. Subjects did not appear to have any difficulty learning the task or completing the experiment.

5 Hypotheses

Prior to conducting the study, we made the following hypotheses:

(1) Because the ‘billboard’ style obscures the background (and therefore some visual interference with the stimulus string), it will result in the fastest and most accurate task performance.

(2) The ‘green’ and ‘red’ styles will result in fast and accurate performance, because a human’s eyes are most sensitive to the two colors.

(3) The ‘maximum HSV complement’ style will result in faster and more accurate task performance than the ‘complement’ style, since it takes into account how optical see-through displays present the color black.

(4) The active styles will result in faster and more accurate task performance than the ‘green’ and ‘red’ (static) styles, since the active styles take background color into account.

(5) The more visually complex outdoor background textures (‘red brick’ and ‘foliage’) will result in slower and less accurate task performance, since their complexity will interfere with the task.

(6) When the distance is ‘medium’ (and therefore matches the display’s accommodative focus), user performance will be faster and more accurate than with the distance is ‘near’ or ‘far’ (and does not match the display’s accommodative focus).

6 Results

We analyzed our results with analysis of variance (ANOVA) and regression analysis. With ANOVA we modeled our experiment as a repeated-measures design that considers *subject* a random variable and all other independent variables as fixed (Table 2). When deciding which results to report, in addition to considering the p value, the standard measure of *effect significance*, we considered two different measures of *effect size*: eta-squared (η^2), and partial eta-squared ($p\eta^2$). Both are measures of how much variance is accounted for by an effect; η^2 is normalized across all the experiment’s independent variables, while $p\eta^2$ is normalized by each variable individually (Cohen, 1973).

6.1 Error Analysis

Out of 7776 total error measurements, there were 7628 correct responses, 137 incorrect responses, and 11 "target string was not visible" responses. Because the error rate was so small (~1.9%), we did not further analyze errors.

6.2 Response Time and Illuminance

FIGURE 4 HERE

FIGURE 5 HERE

Figure 4 shows, as expected, an effect of outdoor background texture on response time ($F(5,85) = 6.16$, $p < .001$, $\eta^2 = 1.36\%$, $p\eta^2 = 26.6\%$). Subjects performed fastest with the 'red brick' background, and they had comparable times for 'pavement', 'granite', 'foliage', and 'sky'. Subjects performed slowest with the 'sidewalk' background. Subjects' superior performance using the 'red brick' background may be explained by our observation that subjects adopted a strategy of moving their head slightly to center the text string within a single brick, and this framing coupled with the visual homogeneity of each brick may have enhanced legibility of the target string. These results refute hypothesis 5; the visually complex background textures performed very well ('red brick') and intermediately well ('foliage'). Clearly, factors other than visual complexity dominated background performance.

Figure 5 shows a main effect of text drawing style on response time ($F(5,85) = 19.86$, $p < .001$, $\eta^2 = 2.83\%$, $p\eta^2 = 53.9\%$). The 'billboard' drawing style supported the fastest performance, followed by 'green'. These results support hypothesis 1, and, to a lesser degree, hypothesis 2. Surprisingly, the 'red' drawing style gave the worst performance. For the active styles, 'maximum HSV complement' was significantly faster than 'complement', which supports hypothesis 3. We did not hypothesize about our 'maximum brightness contrast' style, and its performance overlapped the other active styles. These results do not support hypothesis 4: the 'green' style did at least as well as our best active style.

FIGURE 6 HERE

FIGURE 7 HERE

Figure 6 shows a response time interaction between background and text drawing style ($F(25,425) = 5.47$, $p < .001$, $\eta^2 = 2.09\%$, $p\eta^2 = 24.4\%$). Interestingly, the static styles gave much lower interaction effects than the active styles; performance varied widely depending on the combination of background and active style. Considering only the active styles, for some backgrounds ('red brick', 'granite', 'foliage') performance was relatively constant, while for the remaining backgrounds ('sky', 'pavement', 'sidewalk'), it varied considerably according to text drawing style. These results indicate that 'billboard' and 'green' were the only globally effective text drawing styles.

There was no main effect of distance ($F(2,34) < 1$), contradicting hypothesis 6. We believe this result is explained by two aspects of our task: (1) it may not require a sharp accommodative focus to read a number in a distracting text string, and (2) because our task did not require subjects to attend to the background, they may have ignored the background and just focused on the text string. But as Figure 7 shows, there was a response time interaction between background and distance ($F(10,170) = 3.59$, $p < .001$, $\eta^2 = 1.05\%$, $p\eta^2 = 17.42\%$). At the 'near' distance, background had a much greater effect on response time, both positive ('red brick') and negative ('sidewalk').

FIGURE 8 HERE

The effect of background was mitigated as the distance increased to 'medium' and then to 'far'. However, the lack of a main distance effect, combined with the reasoning above, leads us to suspect that this interaction is caused by the closer distances making the texture features more salient. At least it seems clear that the interaction is not caused by accommodative match ('medium') or mismatch ('near', 'far').

We found a main effect of repetition on response time ($F(3,51) = 25.11$, $p < .001$, $\eta^2 = .76\%$, $p\eta^2 = 59.6\%$). Subjects showed a standard learning effect: their response times dropped by 117.6 msec between the first and second repetition, and thereafter remained constant (varying by less than 3 msec). We also found a three-way response time interaction, which we do not consider because it has both low η^2 and $p\eta^2$, and does not mask any lower-order interactions. Finally, for the same reasons, we found but do not consider a nearly-significant two-way response time interaction.

Figure 8 shows that distance had a main effect on illuminance ($F(2,34) = 5.71$, $p = .007$, $\eta^2 = 1.79\%$, $p\eta^2 = 25.2\%$). The closer the subject was to the background posters, the less illuminance reached the subject's position. This can be explained by (1) the spatial layout of the experimental setup, time of day, and hence sun angle when we conducted the experiment, and (2) the slightly opaque white coating on the greenhouse roof, which diffused and softened entering sunlight. The combination of geometry and diffuse sunlight resulted in the background posters casting subtle, diffused shadows in the direction of the subject.

FIGURE 10 HERE

Figure 9 shows that background had a main effect on illuminance ($F(5,85) = 2.16$, $p = 0.066$, $\eta^2 = .93\%$, $p\eta^2 = 11.3\%$). The different backgrounds reflected different amounts of light to the subject, and the brightest backgrounds ('sidewalk', 'sky') reflected the most light. Because the posters had matte surfaces, and entering sunlight was soft and diffused, the reflected light was also soft and diffuse.

Figure 10 shows that the amount of ambient illuminance at the subject's position had an effect on response time that depended on the text drawing style. Because we sampled the illuminance 324 times and the response time 7776 times, we calculated the regression in Figure 10 with a 1944-line data set, where response times are averaged over 4 repetitions, and illuminance readings are 0-order interpolated over 6 drawing styles. We considered higher order interpolations, but because both distance (Figure 8) and background (Figure 9) affected illuminance, we decided it was more accurate to model illuminance with first-order discontinuities at boundaries where distance and / or background change.

Figure 10 also shows that subjects performed faster under the condition of less illuminance. This result can be explained by the fact that brighter illuminance tends to wash out the AR display, reducing the contrast between augmenting text and background texture. However, the strength of this effect depends on the text drawing style. In order of decreasing slope, 'red', 'maximum brightness contrast', and 'complement' showed significant ($p < .005$) regressions, while 'green' was significant at a weaker ($p < .05$) level. While 'maximum HSV complement' and 'billboard' have slight positive slopes, the *t*-test does not indicate that the regressions are different from 0. The effect of Figure 10 is why the amount of illumination reflected to the subject (Figures 8 and 9) is important.

7 Discussion

Our most important finding, which is not surprising, is clear empirical evidence that user performance on a task, which we believe is representative of a wide variety of imagined and realized AR applications, is significantly affected by background texture (Figure 4), text drawing style (Figure 5), and their interaction (Figure 6). Furthermore, the background affected the amount of ambient illuminance at the subject's position (Figure 9), and the combination of this illuminance and text drawing style also affected user performance (Figure 10).

In terms of design guidelines, the current study suggests using the 'billboard' and 'green' text drawing styles, and avoiding the 'red' style (Figures 6 and 10). However, the 'billboard' style is likely effective because the solid white background of the text string obscures the outdoor background texture, an effect that would be detrimental to many AR applications. Therefore, the main design guideline findings are evidence for the global effectiveness of fully-saturated green labels, and the global ineffectiveness of fully-saturated red labels. Interestingly, several currently manufactured monochromatic AR displays use red.

We are somewhat surprised that our active text drawing styles did not perform better relative to the static styles, refuting hypothesis 4. Nevertheless, 'maximum HSV complement', our active style that took the nature of optical see-through AR displays into account, did perform better than the 'complement' style (hypothesis 3; Figures 5 and 10). We still believe that the right active styles will result in better performance than static text drawing styles. In the current study, we only actively manipulated color, and only according to the averaged pixel color of the entire background texture. Figure 10 indicates that ambient illuminance can also affect text string reading performance. As discussed below, in the future we plan to design active styles that react to both ambient illuminance and to more precise measurements of background color.

Like most controlled user-based studies, this one had many limitations that restrict the generality of our findings. All these limitations suggest future text drawing style implementations and subsequent user-based studies.

- For reasons discussed above, we printed our outdoor background textures onto matte posterboards. The reflective properties of the poster surface are, of course, different than a real surface; e.g. imagine an actual brick wall as opposed to a photograph (albeit a very realistic photograph) of one.
- Furthermore, our background textures were two-dimensional; many textures, such as foliage, have large depth variation, which would likely affect the results.

- Although we tested six qualitatively very different textures, we still managed to test only a small sample of the hundreds (thousands?) of possible urban textures. Furthermore, the appearance of even our tested textures varies widely with differing illumination. It may be possible to use different texture analysis techniques to systematically select or generate a more comprehensive texture set.
- Like all optical see-through AR user-based studies to date, perhaps our greatest limitation is the capabilities of the display hardware itself. Among the serious limitations of our Glasstron display are that it does not support true occlusion and cannot display dark colors, its shifting head-fit makes precise alignment between augmentations and real-world objects difficult, and, like all common computer displays, its dynamic range does not come close to eight orders of magnitude of outdoor illuminance variance (Halsted, 1993).
- Although our experiment examined outdoor illuminance values, we only sampled a fraction of the available outdoor dynamic range, which varies from a starlit landscape to direct noon sunshine (Halsted, 1993).
- Finally, as discussed above, our task did not require the subject to integrate augmented information with real-world objects, but many potential AR tasks (such as product maintenance (Lipson et al., 1998)) would require this type of integration.

In summary, more important than the particular effects we found is the empirical confirmation that user performance for text legibility, closely related to the fundamental AR task of reading text, was strongly affected by text drawing style, background texture, and their interaction, which strongly motivates and informs further research.

8 Concurrent and Future Work

This work is the beginning of a series of research efforts designed to increase legibility in outdoor AR user interfaces. At the current time, we have conducted, but have not yet reported on, a study which involves a text reading task that more tightly integrates augmented and real-world information, further studies the effects of accommodative demand, and utilizes a Microvision laser-scanning optical see-through AR display (Gupta, 2004). This laser-scanning technology could potentially match the full dynamic range of outdoor illuminance.

In addition, we are concurrently running two empirical studies, and have several additional user-based studies planned. First, we are studying the effects of text/graphics drawing styles, environmental lighting, and outdoor background textures on user task performance in outdoor AR tasks, using physical real-world objects, rather than posters, and more sophisticated text drawing styles. Second, we are further exploring design styles, such as altering the brightness of augmentations, without changing their fundamental color and thus preserving color encoding, and dynamically altering the opaqueness of a background rectangle.

In the future, we intend to examine other potential active text drawing styles, such as a more accurate maximum luminance contrast style as was originally designed (MacIntyre, 2003), the anti-interference font (Harrison & Vincente, 1996), as well as other text drawing styles taken from graphics arts and the television and movie industries (e.g., drop shadows, halos, and so forth). This work will help identify static text rendering styles that not only preserve color-coding, but are flexible and robust enough to use in varying outdoor conditions. This includes developing methods to quantify the luminance of real-world objects and augmentations within the AR display, and using these methods to better measure the actual luminance contrast between augmentations and real-world objects.

Acknowledgements

BARS is supported by the Office of Naval Research (ONR) and the Naval Research Lab, and our usability engineering research is also supported by ONR under Program Managers Dr. Larry Rosenblum and Dr. Behzad Kamgar-Parsi. We would like to thank everyone who helped design, run, and analyze these experiments, especially Dr. Robert S. Schulman, John Lucas, and Divya Gupta. Lastly, we would like to thank Dr. Ron Kriz for his continued commitment to support the Virginia Tech CAVE and all its related research activities, as well as Dr. Lloyd Hipkins for use of the Virginia Tech greenhouse.

References

Azuma, R. T., Bruce R., Hoff, B., Neely III, H.E., & Sarfaty, R. (1998). "Making augmented reality work outdoors requires hybrid tracking". Proceedings of the First International Workshop on Augmented Reality, pp. 1-6, San Francisco, CA.

Azuma R. T. (1999). "The challenge of making augmented reality work outdoors". Mixed reality: Merging real and virtual worlds, pp. 379-39, Springer-Verlag.

Azuma, R. T., Hoff, B., Neely III, H., & Sarfaty, R. (1999). "A motion-stabilized outdoor augmented reality system". Proceedings of the IEEE Conference on Virtual Reality, pp. 252-259, Houston, TX.

Box, G. E. P., Hunter, W. G. & Hunter, J. S. (1978). Statistics for experimenters: an introduction to design, data analysis, and model building, New York. Wiley.

Cahill, M.C. & Carter, R.C. (1976). "Color Code Size for Searching Displays of Different Density", Human Factors, Volume 18, Issue 3, pp. 273–280.

Carter, R.C. (1982). “Visual Search with Color”, Journal of Experimental Psychology: Human Perception and Performance, Volume 8, Issue 1, pp. 127–136.

Christ, R.E. (1975). “Review and Analysis of Color Coding Research for Visual Displays”, Human Factors, Volume 17, Issue 6, pp. 542–570.

Cleveland, W.S. & McGill, R. (1985). “Graphical Perception and Graphical Methods for Analyzing Scientific Data”, Science, Issue 229, pp. 828–833.

Cohen, J. (1973). “Eta-Squared and Partial Eta-Squared in Fixed Factor ANOVA Designs”, Educational and Psychological Measurement, Volume 33, pp. 107–112.

MacIntyre, B. and Coelho, E. (2000) “Adapting to Dynamic Registration Errors Using Level of Error (LOE) Filtering” In International Symposium on Augmented Reality (ISAR 2000), pp. 85–88, Oct 5-6, 2000, Munich, Germany.

Devlin, K. (2002). “A review of tone reproduction techniques,” Department of Computer Science, University of Bristol. Online at http://www.cs.bris.ac.uk/Tools/Reports/Ps/2002-devlin-1.pdf

Feiner, S., MacIntyre, B., Höllerer, T., & Webster, A. (1997). A touring machine: prototyping 3D mobile augmented reality for exploring the urban environment. IEEE First International Symposium on Wearable Computers, 74-81.

Foley, J.D., van Dam, A., Feiner, S.K., Hughes, J.F., & Phillips, R.L. (1993). Introduction to Computer Graphics (2nd edition), Reading, MA, Addison-Wesley.

Gabbard, J.L., Swan II, J.E., Hix, D., Lanzagorta, M., Livingston, M.A., Brown, D., & Julier, S. (2002). “Usability Engineering: Domain Analysis Activities for Augmented Reality Systems”, The Engineering Reality of Virtual Reality 2002, A Woods, J Merritt, S Benton, M

Bolas, Editors, Proceedings SPIE Volume 4660, Stereoscopic Displays and Virtual Reality Systems IX, pp. 445–457.

Gabbard, J.L. (2003). Real-world Scenes, Objects, and Textures as a GUI Background: Designing Effective Text Styles for Outdoor Augmented Reality, PhD Dissertation Preliminary Report, Virginia Tech.

Gabbard, J.L., Swan II, J.E., Hix, D., Schulman, R.S., Lucas, J., & Gupta, D. (2005). "An Empirical User-Based Study of Text Drawing Styles and Outdoor Background Textures for Augmented Reality", In Proceedings of IEEE Virtual Reality, pp. 11-18.

Gleue, T., & Dähne, P. (2001). "Design and implementation of a mobile device for outdoor augmented reality in the ARCHEOGUIDE project". Virtual Reality, Archaeology, and Cultural Heritage International Symposium (VAST01), Glyfada, Nr Athens, Greece, 28-30.

Gonzalez, R.C., & Woods, R.E. (1992). Digital Image Processing. Addison-Wesley, Reading, MA.

Gupta, D. (2004). An Empirical Study of the Effects of Context-Switch, Object Distance, and Focus Depth on Human Performance in Augmented Reality, Masters Thesis in Industrial and Systems Engineering, Virginia Tech.

Halsted, C.P. (1993). "Brightness, Luminance and Confusion", Information display, March.

Harrison B.L. & Vicente K.J. (1996). "An Experimental Evaluation of Transparent Menu Usage", In Proceedings CHI '96, pp. 391–398.

Hecht, E. (1987). Optics (2nd edition), Addison Wesley.

Hix, D., Gabbard, J.L., Swan II, J.E., Livingston, M.A., Höllerer, T.H., Julier, S.J., Baillot, Y. & Brown, D. (2004). A Cost-Effective Usability Evaluation Progression for Novel Interactive Systems, Hawaii International Conference on System Sciences (HICSS-37).

Jacobson, R. E. (1995). “Image Quality Metrics”, The Journal of Photographic Science, 43(2), pp. 42–43.

Kelso, J., Arsenault, L.E., Satterfield, S.G., Kriz, R.D. (2003) *DIVERSE*: A Framework for Building Extensible and Reconfigurable Device Independent Virtual Environments and Distributed Asynchronous Simulations. *Presence: Teleoperators and Virtual Environments,* 12(1), pp. 19-36.

Knoblauch, K. & Arditi, A. (1994). “Choosing Color Contrasts in Low Vision: Practical Recommendations”, In AC Kooijman, PL Looijestijn, JA Welling, GJ van der Wildt (Eds.), Low vision: Research and new developments in rehabilitation. Amsterdam, The Netherlands, IOS Press, pp. 199–203.

Legge, G., Parish, D., Leubker, A., & Wurm, L. (1990). “Psychophysics of Reading XI: Comparing color contrast and luminance contrast”, Journal of the Optical Society of America, Volume 7, pp. 2002–2010.

Lipson, H., Shpitalni, M., Kimura, F., Goncharenko, I. (1998). “On-line Product Maintenance by Web-Based Augmented Reality”, In New Tools and Workflows for Product Development, Berlin, pp. 131–143.

Livingston, M.A., Rosenblum, L., Julier, S.J., Brown, D., Baillot, Y., Swan II, J.E., Gabbard, J.L., & Hix, D. (2002). “An Augmented Reality System for Military Operations in Urban

Terrain", In Proceedings of the Interservice / Industry Training, Simulation, & Education Conference (I/ITSEC '02), Orlando, FL, December 2–5.

Livingston, M.A., Swan II, J.E., Gabbard, J.L., Höllerer, T.H., Hix, D., Julier, S.J., Baillot, Y. & Brown, D. (2003). Resolving Multiple Occluded Layers in Augmented Reality, International Symposium on Mixed and Augmented Reality (ISMAR '03), October 7-10, Tokyo, Japan, pp. 56-65.

MacIntyre, B. (1991). A Constraint-Based Approach To Dynamic Colour Management For Windowing Interfaces, Master's thesis, University of Waterloo, Available as Department of Computer Science Research Report CS-91-55.

MacIntyre, B. & Cowan, W. (1992). "A Practical Approach to Calculating Luminance Contrast on a CRT", ACM Transactions on Graphics, Volume 11, Issue 4, pp. 336–347.

MacIntyre, B. (2003). Personnel Communication.

Naimark, M. (1991). "Elements of realspace imaging: A proposed taxonomy", Proceedings SPIE Vol. 1457, Stereoscopic Displays and Applications II.

Pasman, W. (1997). Perceptual requirements and proposals for the Ubicom augmented reality displays. Ubicom Technical Report.

Piekarski, W. & Thomas, B. (2002). "ARQuake: The Outdoor Augmented Reality Gaming System", Communications of the ACM, Volume 45, Issue 1, pp. 36–38.

Smallman, H.S. & Boynton, R.M. (1990). "Segregation of Basic Colors in an Information Display", Journal of the Optical Society of America, Volume 7, Issue 102, pp. 1985–1994.

Swan II, J.E. & Gabbard, J.L. (2005). Survey of User-Based Experimentation in Augmented Reality, to appear in 1st International Conference on Virtual Reality, Las Vegas, NV.

PRESENCE: Teleoperators and Virtual Environments, Volume 15, Number 1, February 2006, pages 16–32.

Thomas, B., Close, B., Donoghue, J., Squires, J., De Bondi, P., Piekarski, W. (2002). "First Person Indoor/Outdoor Augmented Reality Application: ARQuake", Personal and Ubiquitous Computing, Volume 6, pp. 139–146.

Watson, A. B., Hu, J. & McGowan III, J. F. (2001). "A Metric for Visual Quality of Digital Video," NASA Tech Brief ARC14236, Ames Research Center, Moffett Field, California, Online at http://www.nasatech.com/Briefs/Apr01/ARC14236.html.

Williamson, S.J. & Cummins, H.Z. (1983). Light and Color in Nature and Art, Wiley and Sons, NY.

Wyszecki, G. & Stiles, W.S. (1982). Color Science – Concepts and Methods, Quantitative Data and Formula, John Wiley, NY.

Figure Captions

Figure 1. Panoramic image of a 3-way road intersection used to simulate the real world in a CAVE™-based study. (Note that the panoramic image is split and presented in two halves in this figure).

Figure 2. The six outdoor background textures, respective average pixel colors (shown in small box in each background), and six text drawing styles. When the text strings were perceived in an AR display in front of actual background posters, the perceptually experienced contrast was different than it appears in this printed image.

Figure 3. Our experimental design for the greenhouse study. We varied the distance from user to outdoor background texture the least, followed by outdoor background texture, text drawing style, and repetition.

Figure 4. Effect of outdoor background texture on mean response time.

Figure 5. Effect of text drawing style on mean response time.

Figure 6. Effect of outdoor background texture by text drawing style interaction on mean response time.

Figure 7. Effect of outdoor background texture by distance interaction on mean response time.

Figure 8. Effect of distance on mean illuminance.

Figure 9. Effect of outdoor background texture on mean illuminance.

Figure 10. Regression between response time and illuminance, grouped by text drawing style. '*' and '**' indicate significant regressions.

Figure Sizes

Figure 1.	1.31 x 6.51	
Figure 2.	4.63 x 6.51	we will likely pay for this in color
Figure 3.	3.47 x 6.5	we will likely pay for this in color
Figure 4.	2.25 x 3.37	
Figure 5.	2.25 x 3.37	
Figure 6.	2.25 x 3.37	
Figure 7.	2.25 x 3.37	
Figure 8.	2.25 x 3.37	
Figure 9.	2.25 x 3.37	
Figure 10.	5 x 3.37	

Table Captions

Table 1. Representative lighting levels of a typical outdoor scene (Halsted, 1993).

Table 2. Summary of variables studied in experiment.

PRESENCE: Teleoperators and Virtual Environments, Volume 15, Number 1, February 2006, pages 16–32.

Figure 1

Figure 2

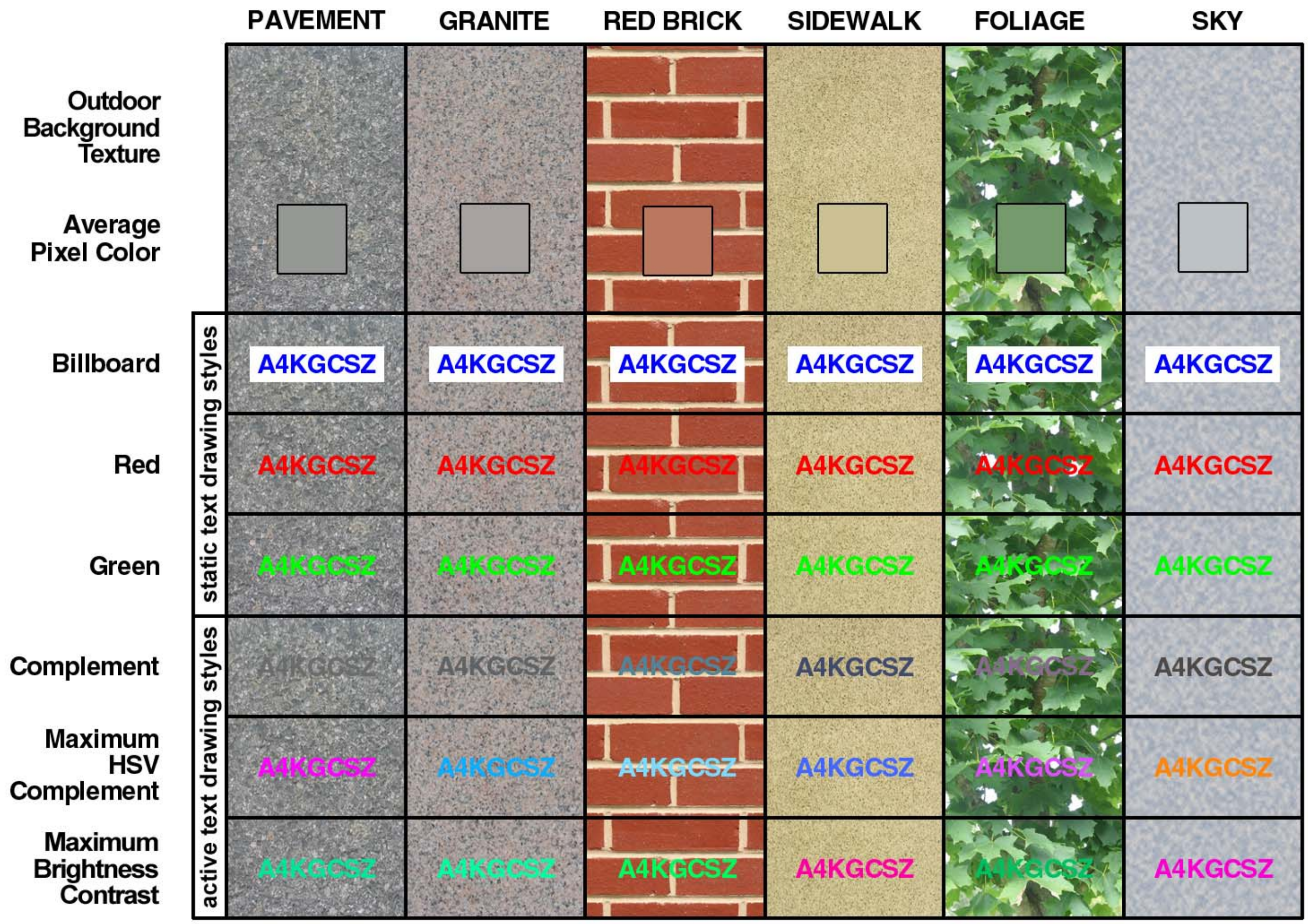

PAVEMENT
GRANITE
RED BRICK
SIDEWALK
FOLIAGE
SKY
Outdoor Background Texture
Average Pixel Color
Billboard
Red
Green
Complement
Maximum HSV Complement
Maximum Brightness Contrast
static text drawing styles
active text drawing styles
A4KGCSZ

Figure 3

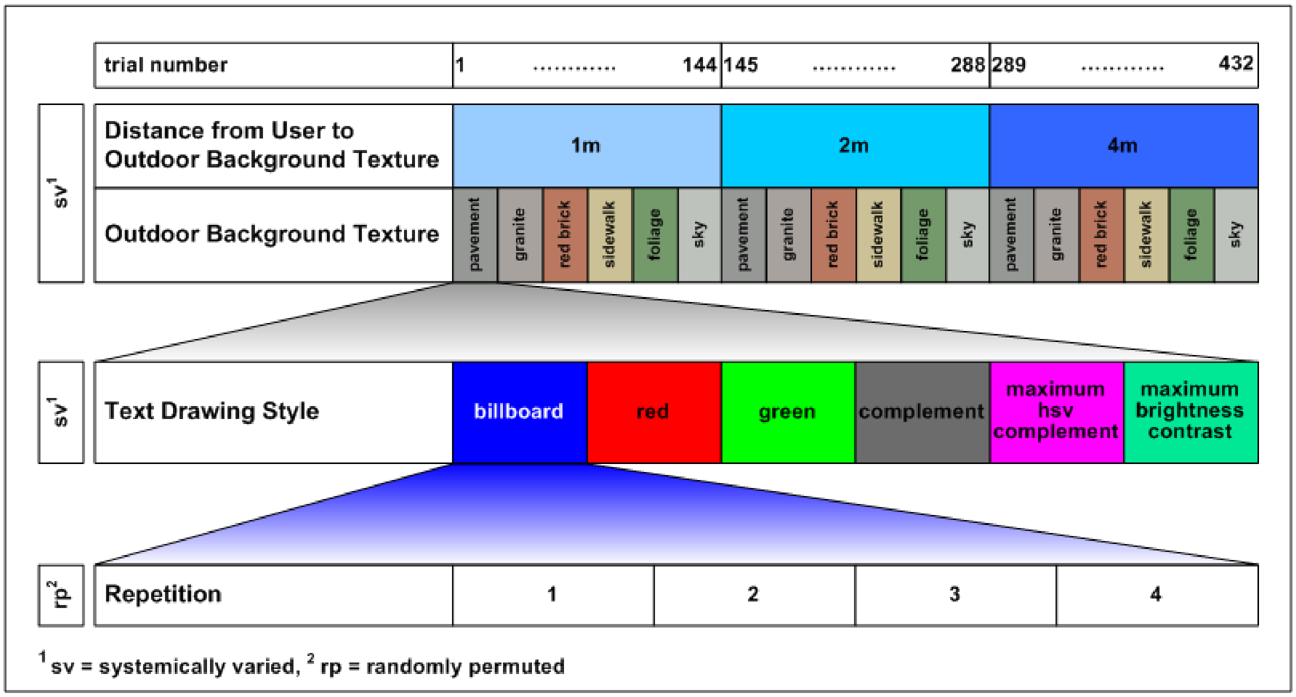

trial number
1 144
145 288
289 432
sv[1]
Distance from User to Outdoor Background Texture
1m
2m
4m
Outdoor Background Texture
pavement
granite
red brick
sidewalk
foliage
sky
sv[1]
Text Drawing Style
billboard
red
green
complement
maximum hsv complement
maximum brightness contrast
rp[2]
Repetition
1
2
3
4
[1] sv = systemically varied, [2] rp = randomly permuted

Figure 4

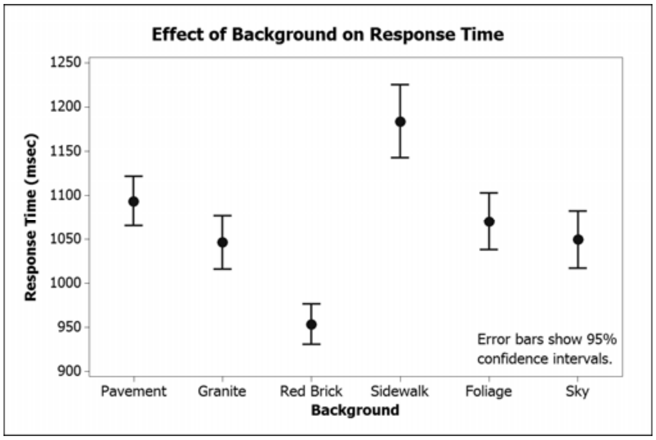

Effect of Background on Response Time
Response Time (msec)
1250
1200
1150
1100
1050
1000
950
900
Pavement
Granite
Red Brick
Sidewalk
Foliage
Sky
Background
Error bars show 95%
confidence intervals.

Figure 5

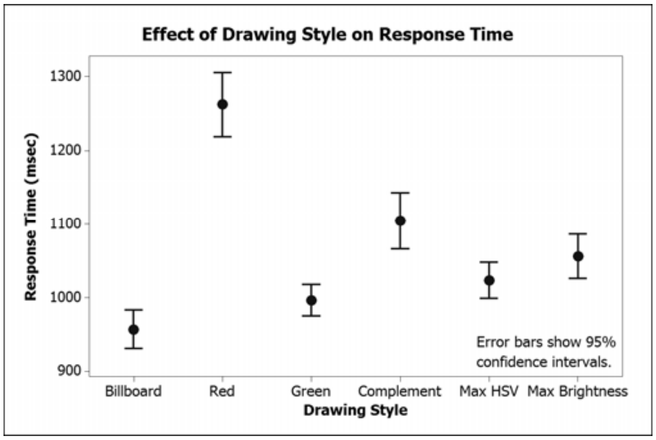
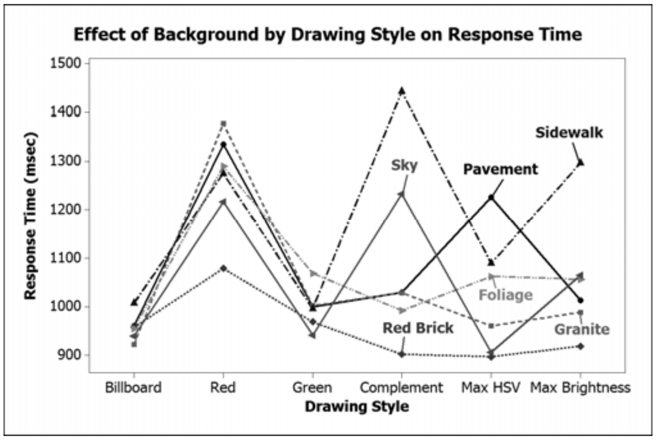

Effect of Drawing Style on Response Time
Response Time (msec)
1300
1200
1100
1000
900
Billboard
Red
Green
Complement
Max HSV
Max Brightness
Drawing Style
Error bars show 95% confidence intervals.

Figure 6



Effect of Background by Drawing Style on Response Time
Response Time (msec)
1500
1400
1300
1200
1100
1000
900
Sidewalk
Sky
Pavement
Foliage
Red Brick
Granite
Billboard
Red
Green
Complement
Max HSV
Max Brightness
Drawing Style

Figure 7

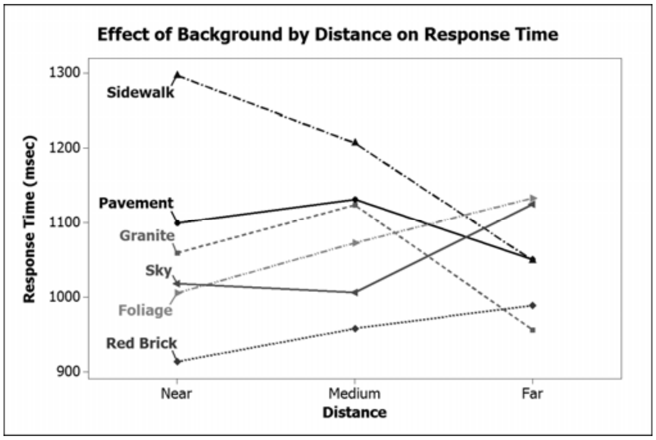

Effect of Background by Distance on Response Time
Response Time (msec)
1300
1200
1100
1000
900
Sidewalk
Pavement
Granite
Sky
Foliage
Red Brick
Near
Medium
Far
Distance

Figure 8

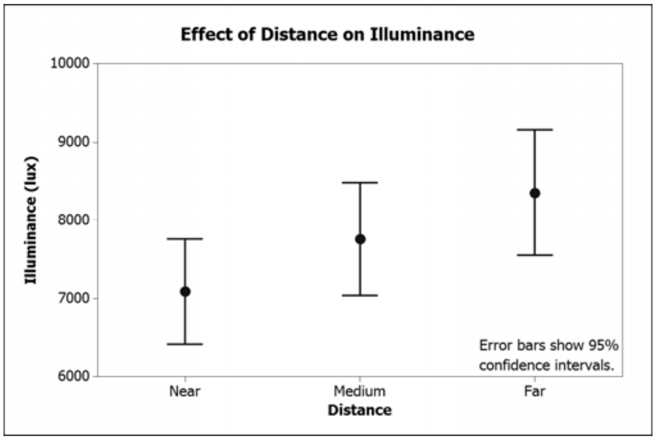

Effect of Distance on Illuminance
10000
9000
8000
7000
6000
Illuminance (lux)
Near
Medium
Far
Distance
Error bars show 95%
confidence intervals.

Figure 9

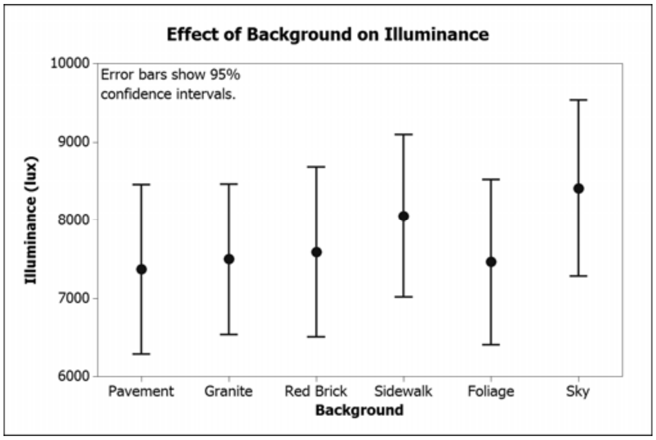
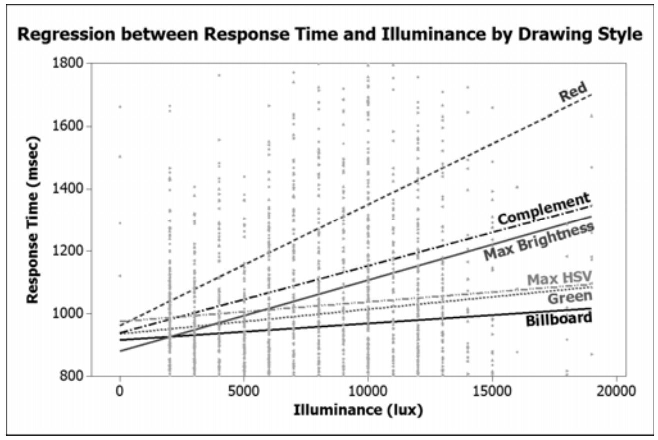

Effect of Background on Illuminance
Error bars show 95% confidence intervals.
Illuminance (lux)
10000
9000
8000
7000
6000
Pavement
Granite
Red Brick
Sidewalk
Foliage
Sky
Background

Figure 10



Billboard	r^2 = 0.49%	$t(322)$ = 1.25	p = .211
Red	r^2 = 7.31%	$t(322)$ = 5.04	p < .000 **
Green	r^2 = 1.35%	$t(322)$ = 2.10	p = .0364 *
Complement	r^2 = 2.70%	$t(322)$ = 2.99	p = .003 **
Max HSV	r^2 = 0.63%	$t(322)$ = 1.42	p = .156
Max Brightness	r^2 = 4.95%	$t(322)$ = 4.09	p < .000 **