Subtle Gaze Direction

REYNOLD BAILEY
Rochester Institute of Technology
ANN MCNAMARA
Texas A&M University
NISHA SUDARSANAM
Mindjet Corporation
and
CINDY GRIMM
Washington University in St. Louis

This paper presents a novel technique that combines eye-tracking with subtle image-space modulation to direct a viewer's gaze about a digital image. We call this paradigm *subtle gaze direction*. Subtle gaze direction exploits the fact that our peripheral vision has very poor acuity compared to our foveal vision. By presenting brief, subtle modulations to the peripheral regions of the field of view, the technique presented here draws the viewer's foveal vision to the modulated region. Additionally, by monitoring saccadic velocity and exploiting the visual phenomenon of *saccadic masking*, modulation is automatically terminated before the viewer's foveal vision enters the modulated region. Hence, the viewer is never actually allowed to scrutinize the stimuli that attracted her gaze. This new subtle gaze directing technique has potential application in many areas including large scale display systems, perceptually adaptive rendering, and complex visual search tasks.

Categories and Subject Descriptors: I.3.3 [**Computer Graphics**]: Picture/Image Generation—*display algorithms*

General Terms: Eye-tracking, modulation

Additional Key Words and Phrases: Luminance, warm-cool, visual acuity, image-based

---

1. INTRODUCTION

When viewing traditional, static images, the pattern a viewer's gaze makes may be guided by a variety of influences. For example, the pattern of eye movements may depend on the viewer's intent or task [Yarbus 1967] [Henderson and Hollingworth 1998]. Image content also plays a role. For example, it is natural for humans to be drawn immediately to faces or other informative regions of an image [Mackworth

---

Author's addresses: R. Bailey, Rochester Institute of Technology, One Lomb Memorial Drive, Rochester, NY 14623; A. McNamara, Texas A&M University, 3137 TAMU, College Station, TX 77843; N. Sudarsanam, Mindjet Corporation, 1160 Battery Street, San Francisco, CA 94111; C. Grimm, Washington University in St. Louis, One Brookings Drive, St. Louis, MO 63130.
Permission to make digital/hard copy of all or part of this material without fee for personal or classroom use provided that the copies are not made or distributed for profit or commercial advantage, the ACM copyright/server notice, the title of the publication, and its date appear, and notice is given that copying is by permission of the ACM, Inc. To copy otherwise, to republish, to post on servers, or to redistribute to lists requires prior specific permission and/or a fee.
© 2001 ACM 0730-0301/2001/0100-0001 $5.00

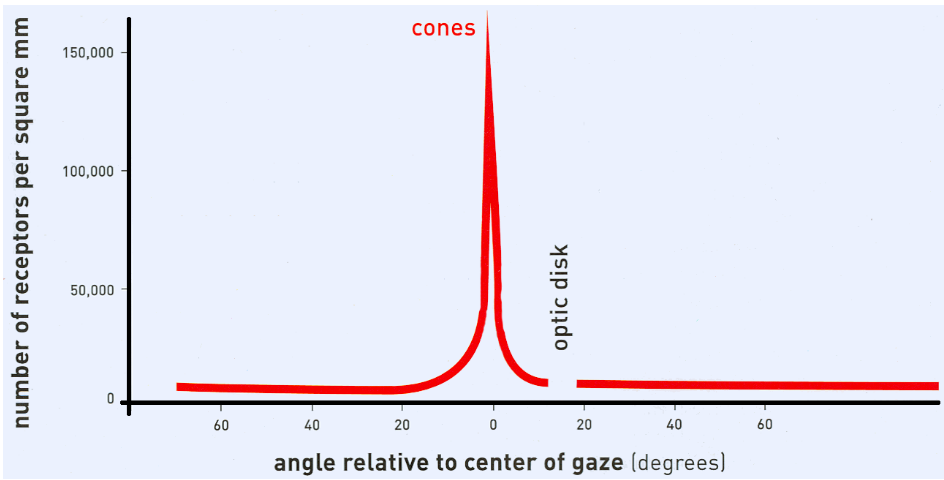


Fig. 1. Distribution of cones in the retina. Adapted from [Livingstone 2002]. Cones are densely packed in the center of gaze (fovea) and the density of cones falls off rapidly as angle from the center of gaze increases. The distribution of cones directly affects visual acuity. Visual acuity is highest in the center of gaze and falls off rapidly as angle from the center of gaze increases.

and Morandi 1967]. Additionally, research has shown that our gaze is drawn to regions of high local contrast or high edge density [Mannan et al. 1996] [Parkhurst and Niebur 2003]. Although traditional images are limited to these passive modes of influencing gaze patterns, digital media offers the opportunity for active control of the gaze pattern.

This paper demonstrates successful use of subtle image modulation to influence gaze direction. The subjects reported no conscious observation of the modulations, but the perceived quality of the modulated images changed, indicating that the altered gaze patterns did have an effect on the viewing experience. The technique works by subtly modifying specific locations of the image to direct the viewer's gaze to those locations. This new technique, which combines eye-tracking with subtle image-space modulation, exploits differences in visual acuity and stimuli response time (detection time) between the peripheral vision and the foveal vision of the human visual system (HVS).

In humans, foveal vision has very high acuity when compared to peripheral vision. The falloff in visual acuity as distance from the fovea increases is directly related to the distribution of the cones in the retina [Osterberg 1935]. Fig. 1 shows the distribution of cones as a function of angle (relative to the center of gaze). The density of cones, and hence the visual acuity, is very high in the fovea (0 degrees) and falls off rapidly as the angle increases[1].

Research has also shown that the peripheral vision responds (detects stimuli) faster than the foveal vision. This is due to the fact that the optic fibers that carry signals from the peripheral regions of the retina to the primary visual cortex for

[1]We ignore the behavior of the rods because our experiment was conducted in a well illuminated (photopic) environment where the response of the rods is completely saturated.

processing are fast-conducting while the optic fibers that carry signals from the fovea are slower [Ogden and Miller 1966].

When viewing a scene for the first time, the low acuity peripheral vision of the HVS locates areas of interest. The slower, high acuity foveal vision is then directed to fixate on these regions. The technique presented in this paper operates by modulating regions of the scene that appear only to the peripheral vision. This causes the eyes to move (saccade) to focus the foveal vision on the modulated region in an attempt to identify the stimuli detected. Both luminance modulation and warm-cool modulation were studied. These modulations were chosen because the HVS is very sensitive to luminance changes [Spillmann 1990], and research has shown that the responses of the photoreceptors in the retina are combined into color-opponent channels that differentiate warm and cool colors [Hurvich and Jameson 1957].

To achieve subtlety in the gaze directing technique, a small study was conducted to establish thresholds for each type of modulation so that they were just intense enough to be detected by the peripheral vision. This attention capture is the result of the motion cue triggered by the onset of the modulation. However, the viewer's foveal vision is never allowed to fixate on the modulated region. This is achieved by monitoring the direction component of the saccade velocity vector, to determine if the foveal vision is about to enter the modulated region. If this is the case, the modulation is immediately terminated. Taking advantage of the visual phenomenon known as *saccadic masking*[2] allows sufficient time to terminate the modulation before it can be scrutinized by the foveal vision. Saccadic masking, first described by Dodge [1900], is the temporary suppression of visual processing during eye movements (i.e. between fixations (see Fig. 2 inset)).

The results of a larger psychophysical experiment reveal that both luminance modulation and warm-cool modulation are effective at directing gaze. Subjective evaluations of the quality of the modulated test images tended to be slightly lower than the corresponding static images. One possible reason for this is that the subtle gaze direction technique forces viewers to violate their natural gaze pattern for a given image.

The remainder of this paper is organized as follows: a brief survey of previous work is given in section 2. In section 3 the subtle gaze direction technique is described in greater detail. The design of an experiment to test the effectiveness of this technique and to study its impact on perceived image quality is presented in section 4. Analysis and discussion of the experimental results are presented in section 5. In section 6, the paper concludes with a summary of the implications of the new technique along with potential avenues of future research.

2. PREVIOUS WORK

Eye tracking systems first emerged in the early 1900s [Dodge and Cline 1901] [Huey 1968] (see [Jacob and Karn 2003] for a review of the history of eye-tracking). Until the 1980s, eye trackers were primarily used to collect eye movement data during

[2]Saccadic masking is also called saccadic suppression. Saccadic masking prevents perception of the blur of the retinal image caused by the ballistic movement of the eye. Complex neurological processes ensure that the resulting gaps in the visual signal are not perceived.

psychophysical experiments. This data was typically analyzed after the completion of the experiments. During the 1980s, the benefits of real-time analysis of eye movement data were realized as eye-trackers evolved as a channel for human-computer interaction [Levine 1981; Hutchinson et al. 1989]. More recently, real-time eye tracking has been used in interactive graphics applications [Luebke et al. 2002; Levoy and Whitaker 1990; Duchowski 2002; O'Sullivan et al. 2003] and large scale display systems [Baudisch et al. 2003] to improve computational efficiency and perceived quality. These systems use real-time eye-tracking to *follow* the viewer's gaze. The technique described in this paper combines real-time analysis of eye movement data with subtle image space modulation to *direct* the viewer's gaze about a scene.

Up until now, computer graphics approaches for directing a viewer's gaze have typically relied on the fact that the foveal vision is naturally drawn to regions of sharp focus or high detail. The most commonly used approach to direct gaze in 2D images is to simulate the depth-of-field effect from traditional photography. This effect can be achieved in commercially available image editing packages by applying a sharpening filter to specific regions of an image and a blurring filter to others [Mitchell 2004]. This has the effect of bringing different areas of an image in or out of focus. The depth-of-field concept has also been applied to 3D scenes [Kosara et al. 2001]. DeCarlo and Santella [2002] used a different approach to direct viewer gaze. They recorded the gaze pattern of a single observer over an image and used it to direct the gaze of others. This was achieved by creating a stylized rendering of the image where only the areas attended to by the first observer are shown in high detail. Cole et al. [2006] applied similar stylized rendering techniques to direct gaze in 3D scenes. Unlike these previous approaches, which draw the viewer's gaze by manipulating focus or level of detail, the subtle gaze directing technique uses brief luminance or warm-cool image-space modulations presented to the peripheral regions of the field of view. Additionally, unlike the previous approaches, this technique does not affect the overall appearance of the image.

3. SUBTLE GAZE DIRECTING TECHNIQUE

Consider the hypothetical image shown in Fig. 2. Suppose that the goal is to direct the viewer's gaze to some predetermined area of interest A. Let F be the position of the last recorded fixation, let $\vec{v}$ be the velocity of the current saccade, let $\vec{w}$ be the vector from F to A, and let θ be the angle between $\vec{v}$ and $\vec{w}$. Either luminance modulation or warm-cool modulation is performed over the pixels in region A. Once the modulation commences, saccadic velocity is monitored using feedback from a real-time eye-tracking device and the angle θ is continually updated[3] using the geometric interpretation of the dot product:

$$\theta = \arccos\left(\frac{\vec{v}\cdot\vec{w}}{|\vec{v}|\,|\vec{w}|}\right) \tag{1}$$

[3]Our software polls the eye-tracker at a rate of 20Hz. This relatively low sampling rate works well in practice because research has shown that only 3-4 fixations occur per second [Gajewski et al. 2005]. A velocity threshold of 100 deg/sec was used to discriminate between saccades and fixations [Salvucci and Goldberg 2000].

A small value of θ ($\leq 10^\circ$) indicates that the center of gaze is moving toward the modulated region. In such cases, modulation is terminated immediately. It is important to note that the modulation is terminated *during* the saccade to take advantage of the gap in our perception caused by saccadic masking (see Fig. 2 inset). This contributes to the overall subtlety of the technique. By repeating this process for other predetermined areas of interest, the viewer's gaze is directed about the scene.

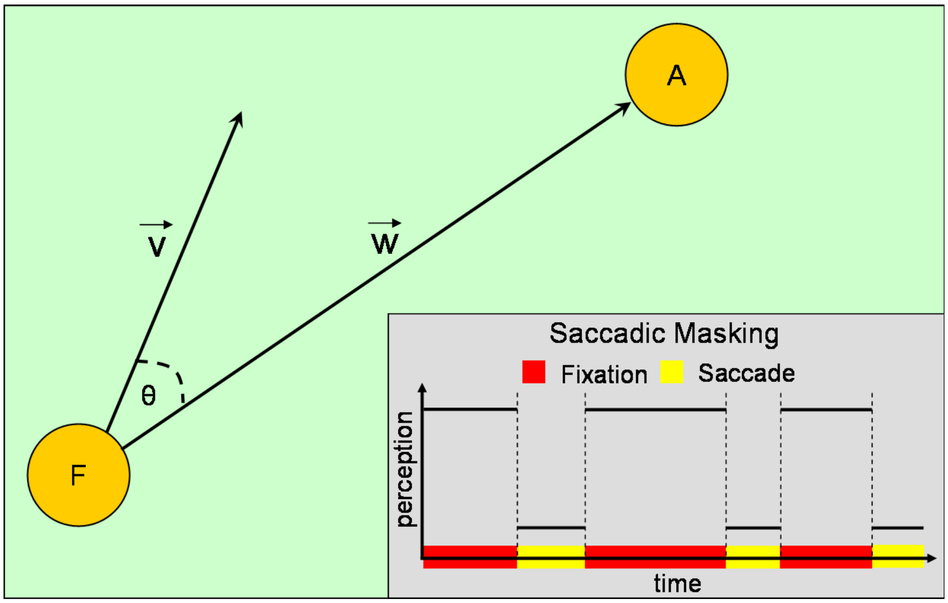


Fig. 2. Hypothetical image with current fixation region F and predetermined region of interest A. Inset illustrates saccadic masking.

The modulations are simply alternating interpolations of the pixels in A with black and white, in the case of luminance modulation, or with a warm and a cool color, in the case of warm-cool modulation (see Fig. 3). The alternating interpolations occur at a rate of 10 Hz. Consider a pixel $p = (x, y)$ in A with color $col(p)$. For warm-cool modulation, the color of the resulting pixel $col'(p)$ for the warm interpolation is given by:

$$col'(p) = ((w * i) + col(p) * (1 - i)) * f(p) + col(p) * (1 - f(p)) \qquad (2)$$

where w is a warm color, $f(p)$ is a Gaussian falloff function [Young 1987], and i is some scalar value in the range [0,1] which controls the intensity of the modulations. The cool interpolation of the warm-cool modulation cycle, and the black and white interpolations of the luminance modulation cycle, are analogous. For our experiments we used red and blue as our warm and cool color, respectively. We set the radius of the function $f(p)$ to be 32 pixels, which corresponds to approximately a $1cm$ diameter circular region on screen (0.76 degrees of the visual field).

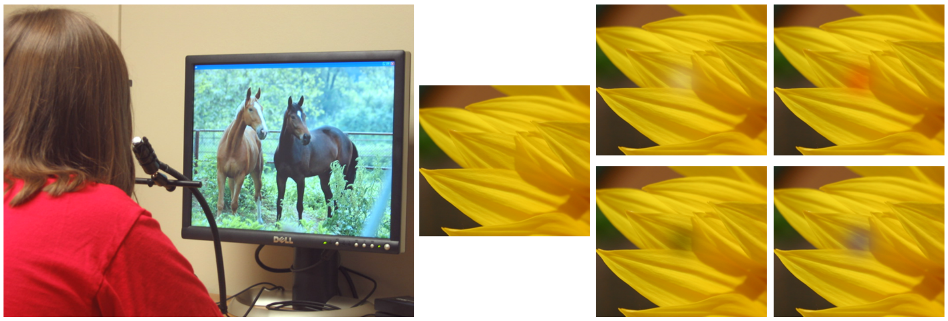


Fig. 3. Photograph of experiment setup (left). Small patch from a test image (center). Example of luminance modulation (third column). Example of warm-cool modulation (right column).

A small pilot study was conducted to determine the value of i for which the modulations are just intense enough to be detected by the peripheral vision. Three participants were involved in this pilot study. They were each presented with five randomly selected images from the complete test set (see Fig. 4). They were instructed to fixate on a cross in the center of the image while modulations were presented in random peripheral regions. The modulations were not noticeable to begin with. Using the keyboard (+/-), they adjusted the value i in step sizes of 0.005 until the modulations were just noticeable. The final value for i was obtained by averaging the results of the three participants. For luminance modulation between black and white with a Gaussian falloff, $i = 0.095$ and for warm-cool modulation between red and blue with a Gaussian falloff, $i = 0.105$. It should be noted that i needs to be recomputed if any changes are made to the falloff function or to the colors used for the interpolations.

4. EXPERIMENTAL DESIGN

This section describes an experiment that was conducted to test the effectiveness of the gaze-directing technique and to study its impact on perceived image quality.

4.1 Stimuli

Stimuli were presented on a 20 inch monitor, operating at 75Hz with a resolution of 1280 x 1024. The stimuli used in this experiment consisted of forty 1280 x 1024 images compiled from various sources. These images were chosen with no particular criteria and are shown in Fig. 4.

4.2 Participants

Ten participants (5 females, 5 males), between the ages of 18 and 45 volunteered to participate in this study. All participants reported normal or corrected-to-normal vision with no color vision abnormalities. Participants were randomly assigned to one of two groups:

—**Static group:** Participants (five) were presented with a randomized sequence of the 40 images with no modulation. This group served as the control group for the experiment.

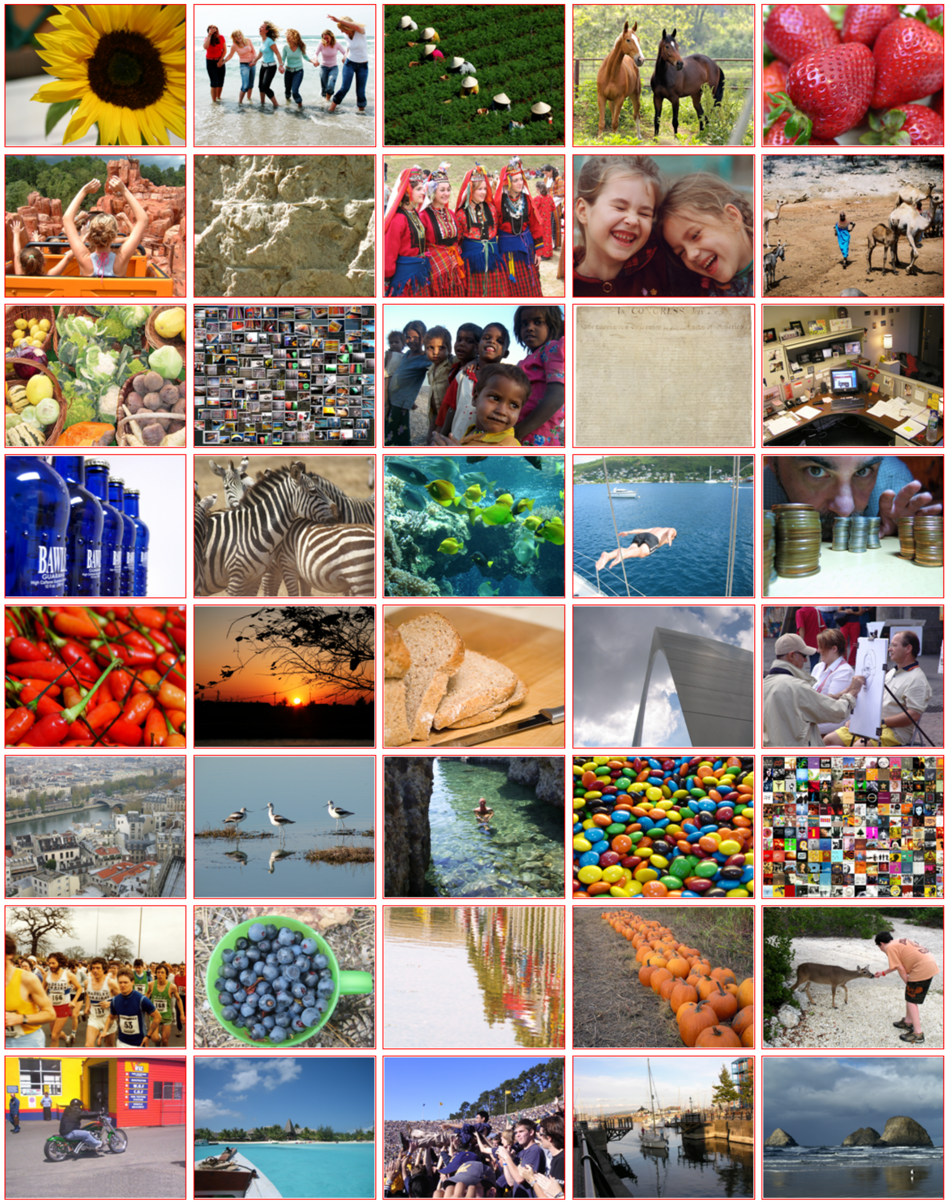


Fig. 4. The complete set of images used in this study. Images were chosen with no particular criteria and gathered from various sources on the web [Yahoo! Inc.] [Google Inc.].

—**Modulated group:** Participants (five) were presented with a randomized sequence of the 40 images with modulations at pre-selected image locations. These locations, manually selected by the researchers, included areas that are not visually significant - low contrast, low detail, low color saturation, and uninteresting objects. These are areas that an observer would not ordinarily attend to. The type of modulation for a given image was randomly selected to be either luminance or warm-cool.

To ensure that the onset of a modulation does not occur in the immediate vicinity of the current fixation or in the path of the current saccade, the eye-tracker is polled just before a new location is modulated. If the new location is within the foveal view the modulation is skipped. To further reduce the likelihood of this occurring, the researchers selected target locations for modulation that were well-spaced across the image.

For both groups, the images were presented for a duration of 8 seconds. Between each image, a black screen with a small white cross displayed at the center was presented for 2 seconds. This allows participants to rest briefly between images.

4.3 Procedure

Participants were seated in front of the computer screen in a well lit room with their chin comfortably resting on a chin-rest to reduce head movement. Using an infrared camera-based eye-tracking system[4], dominant eye position was recorded for each participant[5] (see Fig. 3). Participants were instructed to remain as still as possible while the eye-tracker was calibrated and the experiment was conducted. The chin-rest was positioned $75cm$ from the screen. At this distance, the actual perceptual span (area of high acuity) of the observer subtended an angle of 3.8 degrees of the visual field [Rayner 1975]. This corresponds to a circular region of diameter $5cm$ on the screen. To further promote subtlety, modulations were presented in a smaller $1cm$ diameter circular region equivalent to 0.76 degrees of the visual field.

The participants in each group were asked to assess the quality of each image on a scale from 1 (low) to 10 (high) and to report it verbally during the 2 second period between images. These scores were recorded by the researchers. The term *quality* was not defined by the researchers. Instead, it was left up to the participants to formulate their own notion of image quality. The complete set of instructions read verbatim to each participant are given in Appendix A. For participants in the modulated group, real-time analysis of the eye-tracking data was performed (as described in section 3) to determine when to terminate the modulations.

Data files for all participants were compiled during the experiment. The data files contain the following information: time elapsed, filename of image being displayed, position of modulated region, type of modulation, location of current fixation, and corresponding delta change from last fixation. Entries to the data files occurred at

[4]ViewPoint EyeTracker® by Arrington Research, Inc.

[5]The dominant eye was established by asking participants to stare at an object a few feet away from them and to point at that object with their index finger. With their eyes focused on the object their index finger would appear blurred in the line of sight. They were asked to close one eye and then the other. The dominant eye is the eye with which the index finger appears to be pointing directly at the object.

a rate of 20Hz. Testing for each participant (including calibration) lasted approximately 15 minutes.

5. RESULTS

To measure the effectiveness of our technique we look at activation times (how quickly the viewer responded), how close the viewer got to the target, overall change of viewing pattern, and perceived quality. In summary, the modulations drew the viewer's gaze, hence altering their viewing pattern, but their gaze did not always reach the modulated region, typically falling within one to two perceptual spans. This is not surprising, since the modulations were turned off before the viewer could fixate on them, and other studies have shown that saccades commonly undershoot their targets [Collewign et al. 1988], particularly when the target position is new (not visited before). Although none of the participants reported seeing anything which would cause confusion or distraction, they consistently ranked the modulated images as having lower quality. We hypothesize that this is because we disrupted their natural viewing pattern by drawing their gaze to uninteresting portions of the image (recall that the modulations were deliberately placed in these regions).

5.1 Activation Times

We define *activation time* as the time elapsed between the start of the modulation and the detection of movement in the direction of the modulation (Eq. 1). The main graph in Fig. 5 shows the activation times recorded during the experiment for all participants in the modulated group, sorted by time. The criteria for terminating the modulation was met within 0.5 seconds for approximately 75 percent of the target regions and within 1 second for approximately 90 percent of the target regions, indicating that the participants responded to the majority of the modulations. We also examined how the activation response times varied over time and saw no significant effect. The inset of Fig. 5 shows the activation times for one participant. The plots for the other participants were similar.

5.2 Target Acquisition

To quantify how close the viewer got to the modulation, we found the closest fixation point in the time period following the onset of the modulation and before the start of the next one (Fig. 6). Pixel distances are grouped into bins of 32 pixels (roughly 1cm on screen). Approximately 69% of the time the viewer's gaze came within one perceptual span (160 pixels or 5cm) of the target region for our viewing configuration. An additional 24% of the time (for a total of 93%) the viewer's gaze got within two perceptual spans (10cm).

Ideally, all of the fixations would be within one perceptual span of the target, since this corresponds to the area of high visual acuity. There are two possible causes for this inaccuracy. First, the simple chin-rest that we use does not completely eliminate head movement. Second, since the subtle modulations are terminated as soon as the eye starts moving toward them, it may be that the presentation time of the stimuli is too brief for the brain to accurately determine its location.

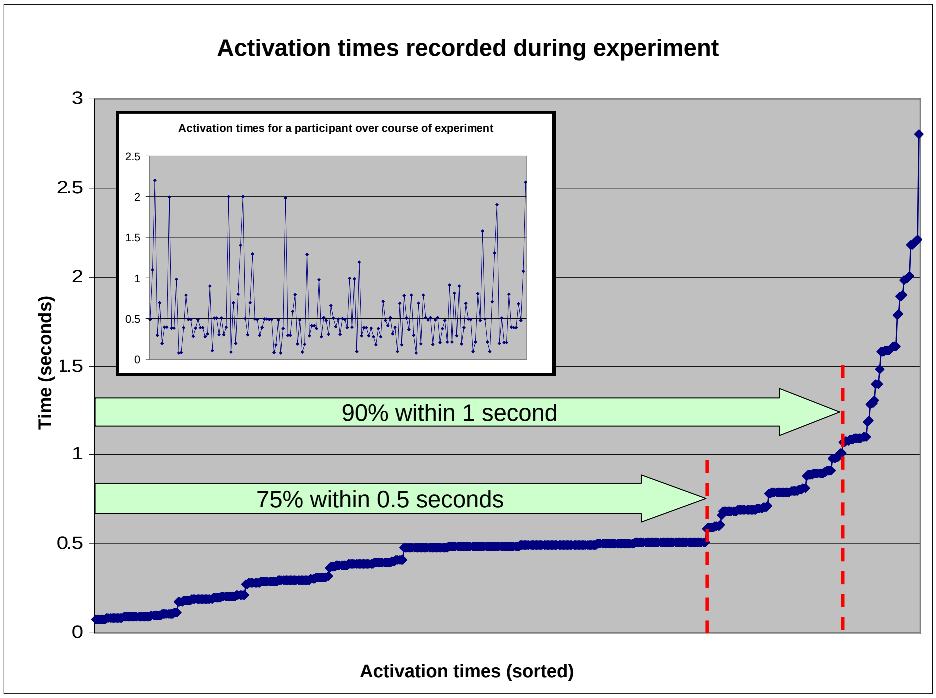


Fig. 5. Activation times recorded during the experiment for all participants in the modulated group, sorted by increasing time. Approximately 75% of the target regions were found within 0.5 seconds and approximately 90% within 1 second. The inset shows the activation times recorded over time for a single participant. This plot reveals no clear trend in the activation response times over time.

5.3 Gaze Pattern Changes

Since only areas of low visual significance were selected as regions of interest, it would be expected that an effective gaze directing technique would result in a gaze pattern that is significantly different than one obtained by natural viewing. We compare gaze patterns by partitioning the image into a 5 x 5 regular grid and counting the number of fixations in each region. Similar gaze patterns will produce similar counts in the cells.

To create the ground truth pattern for each static image we averaged the static image fixations across five participants (see Fig. 7). To verify this ground truth image, a new participant was asked to complete the experiment. As expected, we observed a high degree of correlation between her gaze distribution and that of the averaged distribution. This is especially true for the images with strong salient features where we observed a Pearson coefficient of correlation as high as $r = 0.96$. For images with no dominant salient features, the minimum Pearson coefficient of correlation was still a respectable value of $r = 0.67$.

The participants in the modulated group viewed a total of 200 images (5 participants, 40 images). Of these, 107 were luminance modulated, and 93 were warm-cool modulated (the type of modulation was chosen randomly). Fig. 8 plots all of the

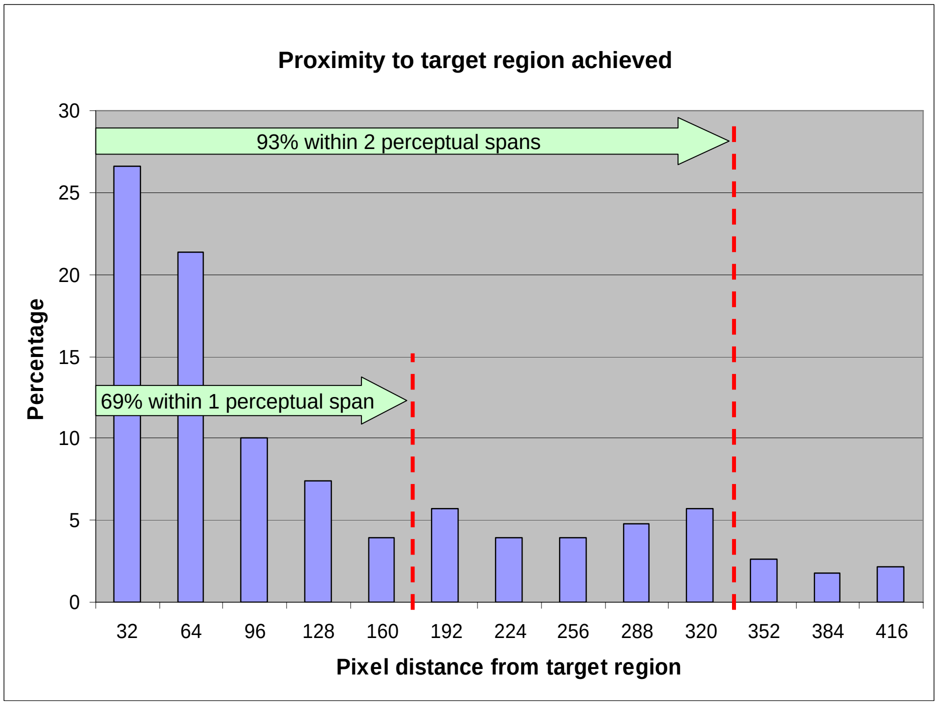


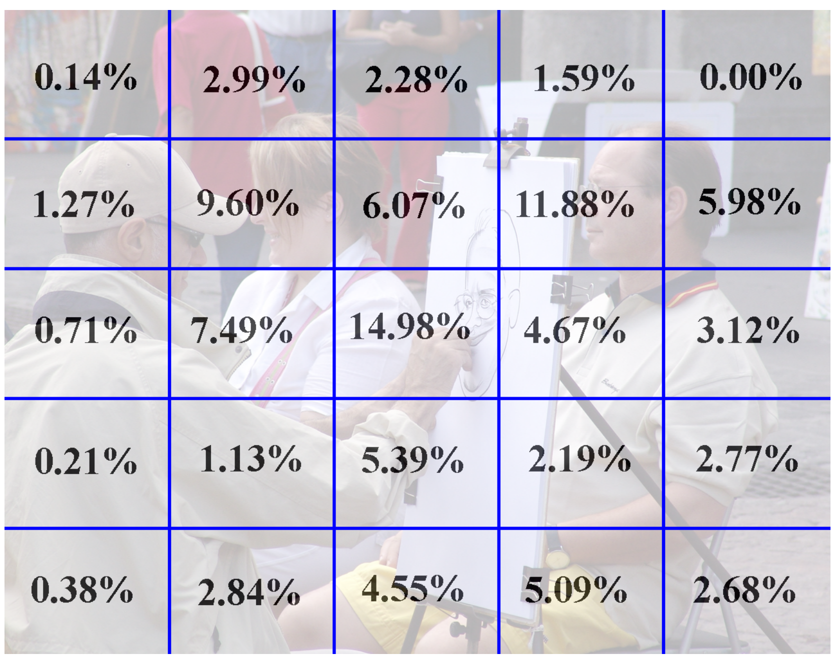
Fig. 6. Histogram showing how close the viewer's gaze got to the modulated regions in terms of pixel distances. Each bin corresponds to 32 pixels (roughly 1cm on screen). 5 consecutive bins (160 pixels) is roughly the size of the perceptual span on screen for our viewing configuration. Approximately 69% of the time the viewer's gaze comes within 1 perceptual span of the target. The viewer's gaze comes within 2 perceptual spans of the target 93% of the time.

correlation values obtained by comparing the gaze distribution of the modulated images with the corresponding ground truth gaze distribution. The results indicate a poor correlation between the gaze distribution of the two groups which implies that the image modulations were successful in shifting the natural viewing path by drawing the viewer's gaze into the modulated regions.

Notice also in Fig. 8 that the coefficients of correlation for the luminance modulated images occupy a narrower range ($-0.34 \leq r \leq 0.66$) than the coefficients of correlation for the warm-cool modulated images ($-0.46 \leq r \leq 0.75$). This suggests that the luminance modulation is more effective at directing gaze than the warm-cool modulation. One possible explanation for this is due to the fact that the luminance modulations are visible against a wider variety of image backgrounds whereas the warm-cool modulations are sometimes difficult to detect, especially against dark or greenish colored backgrounds.

While the small coefficients of correlation suggest that the overall gaze patterns are different for both groups, we are particularly interested to learn how the gaze pattern was affected close to the modulated regions. Recall that modulations were deliberately placed in areas of an image deemed uninteresting under normal viewing conditions. We would expect then, that the percentage of time spent in these

(a)



(b)

Fig. 7. (a) Source image. (b) Average fixation time for each grid cell for participants viewing image without modulation. Notice that regions containing visually significant features such as faces receive greater fixation time.

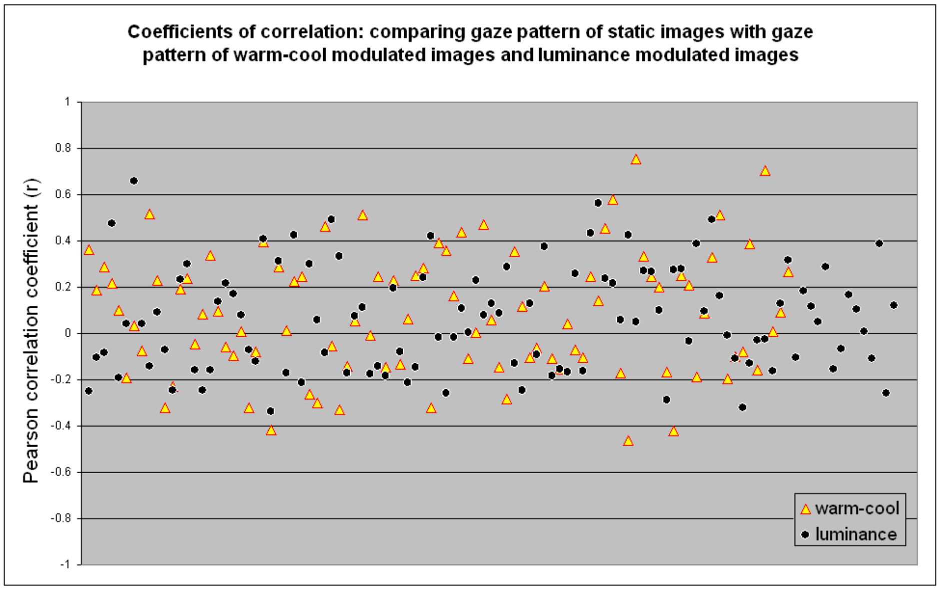


Fig. 8. Plot of all Pearson coefficient of correlation values obtained by comparing the gaze distribution of modulated images with the corresponding natural gaze distribution. Note that most of the coefficients fall in a very narrow range close to zero ($-0.2 \leq r \leq 0.4$). These small values show that there is poor correlation between the gaze patterns of the static group and modulated group.

	Luminance modulation	Warm-cool modulation
1 perceptual span	$F(1, 284) = 40.319; p <= 0.001$	$F(1, 268) = 37.630; p <= 0.001$
2 perceptual spans	$F(1, 284) = 41.078; p <= 0.001$	$F(1, 268) = 31.090; p <= 0.001$

Table I. Results from ANOVAs performed to determine if there was a significant change in percentage viewing time spent within 1 and 2 perceptual spans of the target regions. Separate ANOVAs were conducted for luminance and warm-cool modulations. The results show that there is a significant difference between the percentage viewing time spent within 1 and 2 perceptual spans of the target regions for both groups.

regions in static images would be significantly less than the time in the modulated cases. For each image and for each subject, we calculated the percentage of time that the viewer's gaze was within one perceptual span of any modulation point. An ANOVA was performed to determine if there was a significant difference in percentage times between the two groups. This is a mixed-model ANOVA with two random factors (images and subjects), and one fixed factor, change (modulated or static). Subjects are nested within change and images are crossed with change. We repeated the ANOVA using time spent within two perceptual spans as well, and saw similar results. The results, summarized in Table I, show that the effects of the modulation are indeed significant. Note that the luminance and warm-cool modulations were tested separately.

We can visually compare gaze patterns using a heat map (see Figs. 9 and 10). Although our numerical target analysis showed that the majority of the time the viewer's gaze came within one perceptual span of the target, it is clear from the heat map that their gaze did not always *stay* there. We hypothesize that, once the viewer's gaze was drawn towards the target and the modulation turned off that their gaze drifted to the nearest salient features.

In a more recent study [McNamara et al. 2008] we used both subtle and non-subtle modulations to help the participant perform a search task. In this case, the targets were located on features that the participant was searching for, and in this case the heat map aligns much better with the targets.

5.4 Evaluating the Overall Experience

On completing the experiment the participants were asked the following question: "Did you notice anything unusual about the images?" The term *unusual* was not defined by the researchers. Surprisingly, only one participant reported seeing anything atypical. However, even on noticing the presence of irregularities the participant reported ignoring them and completing the experiment regardless. This requires further studies to verify that this is indeed the case and that task performance (for example) would not be impeded as a result of gaze direction.

We use the perceived image quality to determine the effect of gaze modulation on the viewing experience. Fig. 11(a) summarizes the mean quality scores for both the static and the modulated images. Quality ratings for the static images average 7.07, while the modulated images receive a mean quality rating of 6.17. These mean scores were obtained by averaging the individual scores of the 5 participants over the forty images in each group. The difference in scores suggests that introducing modulations, even subtle ones presented to the peripheral regions of the field of view, results in a reduction in the perceived quality of the image being viewed.

An independent-samples t-test revealed that the effect of the modulation was significant and not due to chance:

$$t(398) = 4.884; p < 0.001$$

Recall that the modulated regions were specifically chosen because they had no features of interest. We believe that by drawing the viewer's gaze to these regions they were left with less time to view parts of the image that they would normally attend to. This altered pattern can be clearly seen in Figs. 9 and 10.

We also categorize the quality ratings of each static image with respect to the corresponding modulated image. A static image can fall into one of three categories:

—higher perceived quality

—equal perceived quality

—lower perceived quality

Fig. 11(b) shows the percentage of images that fall into each of these categories. More (80%) of the static images received higher quality ratings than their corresponding modulated images. The reverse was true for only a few (17.5%) images.

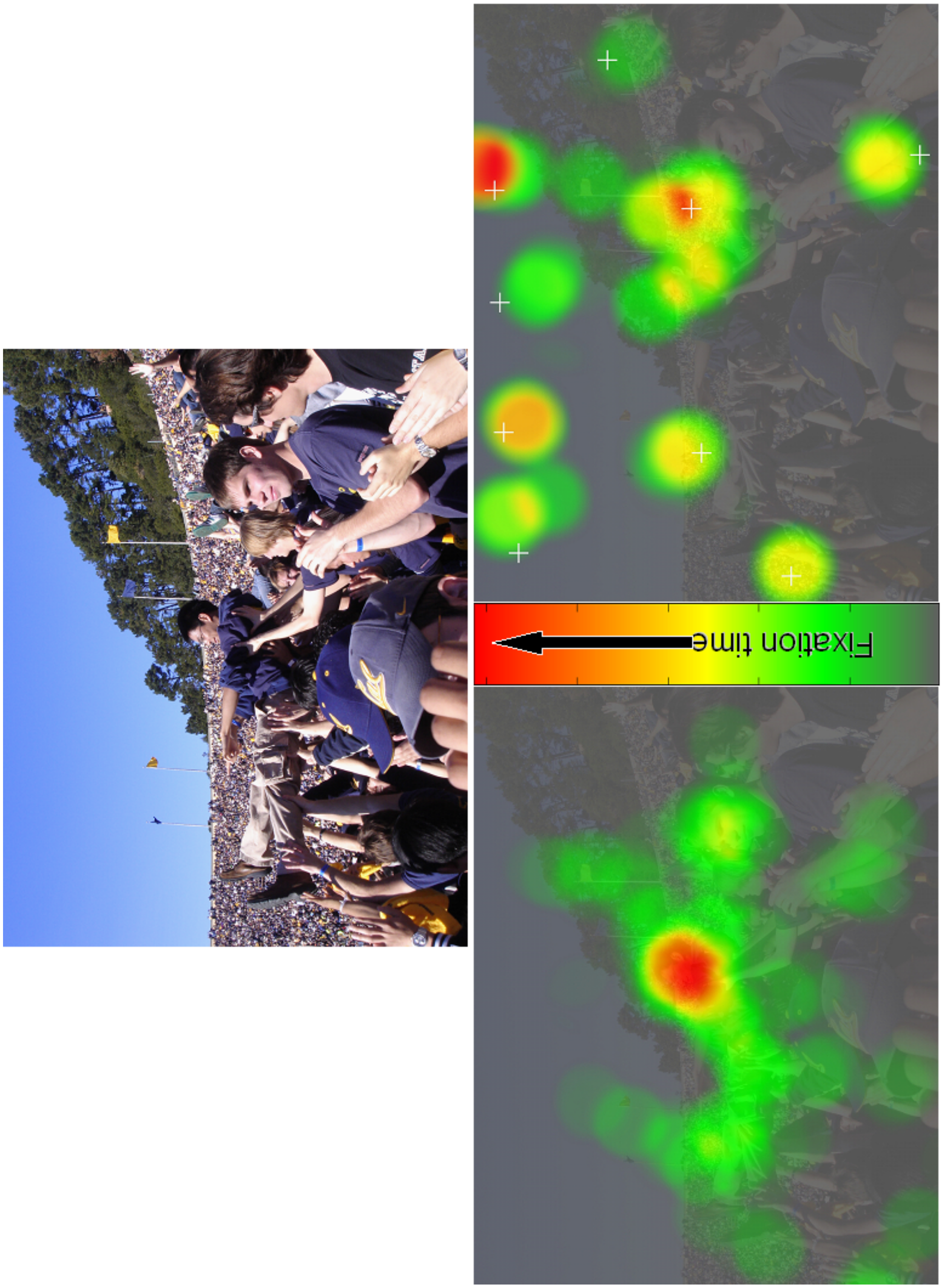


Fig. 9. Gaze distributions for an image under static and modulated conditions. Input image (top). Gaze distribution for static image (bottom left). Gaze distribution for modulated image (bottom right). White crosses indicate locations preselected by researchers for modulation.

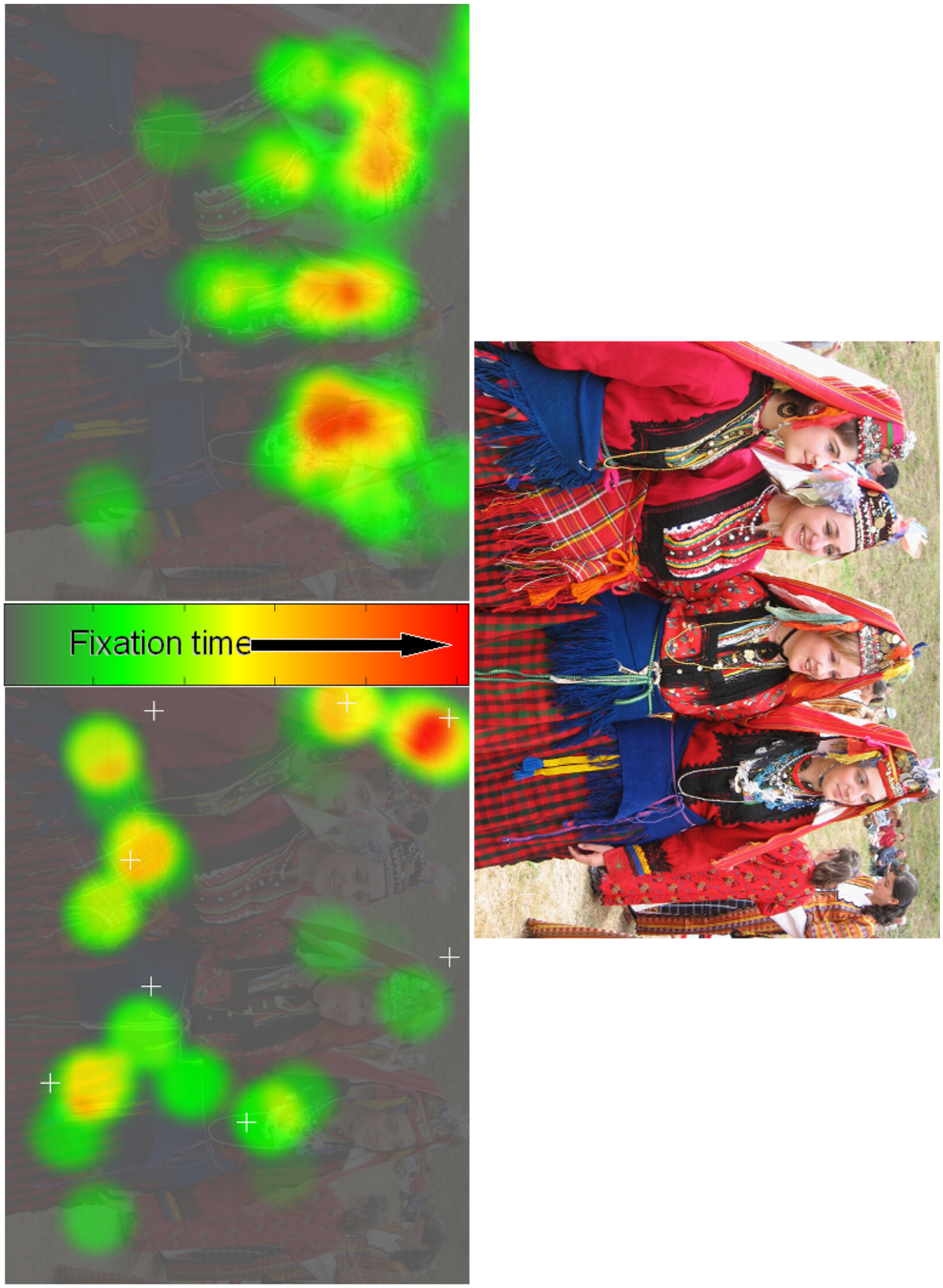
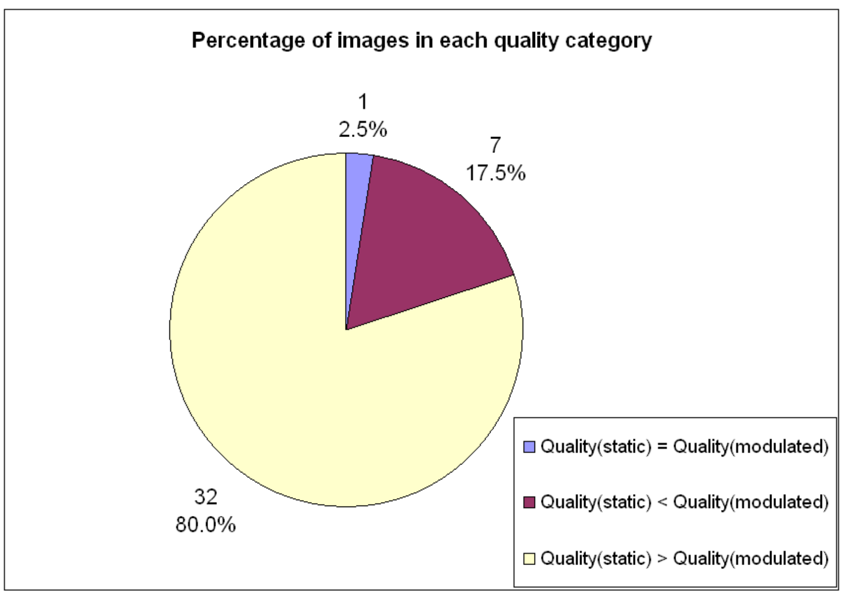


Fig. 10. Gaze distributions for an image under static and modulated conditions. Input image (top). Gaze distribution for static image (bottom left). Gaze distribution for modulated image (bottom right). White crosses indicate locations preselected by researchers for modulation.

(a)



(b)

Fig. 11. (a) Average quality scores for static and modulated images. (b) Percentage of images that fall into each quality category.

6. CONCLUSION

We have presented a method capable of directing an observer's gaze to chosen regions of a display. Using custom built software, we introduce brief subtle luminance or warm-cool modulations in still images. By monitoring the saccadic velocity of the dominant eye of the observer, we are able to terminate these modulations before the observer has an opportunity to scrutinize them. The results of an experiment show that this approach to gaze directing is highly effective. Subjective evaluations of image quality, however, tended to be lower for the modulated images compared to the corresponding static images.

Our technique has potential applications in a number of research and commercial areas[6]:

—**Perceptually Adaptive Rendering:** By dictating where in a scene a viewer looks, we can reap benefits in the rendering domain. If a particular region of a scene requires more rendering time (due to the complexity of the model or the need for higher resolution), our technique can be used to direct users to look at other regions of the scene that required less time to render, thus distracting their view from the more complex areas that are progressively updating. To fully make use of this idea, we would have to gather additional information on how long a user is distracted by a modulation and how strong the modulation must be in order to distract from changes elsewhere in the image. By coordinating modulations with image updates it might be possible to mask the updates with saccades.

—**Flight/Driving Simulation and Training:** In flight and driving simulators, the goal is to instill good navigation habits, such as checking certain information on the cockpit equipment, or routinely checking rear-view and side mirrors. Training simulators could be equipped with subtle gaze direction techniques to encourage users to frequently look at selected regions, such as mirrors, in the hope that the habit would transfer to the real world situation. Because the imagery in this case is dynamic, additional studies are necessary to determine if, and what kind of, modulations are sufficiently strong to attract the viewer's gaze without decreasing their performance.

—**On-line Training and Distance Education:** Educators employing on-line technology in distance learning courses could use subtle gaze direction to encourage student viewing of relevant sections of the on-line course display. For example, when displaying slides with a voice over, subtle gaze direction could be used to guide viewers to relevant text or imagery on the screen. This would involve carefully synchronizing the modulations to reflect the content in the audio.

—**Pervasive Advertising:** On large single-screen displays, advertisers could use subtle gaze directing to quickly guide the viewer to the important product in-

[6]The trade-off between cost and accuracy/sensitivity of eye-tracking equipment should be taken into account for each of the applications listed. In general, our technique uses eye-tracking primarily as a means to ensure subtlety. Hence, the higher the degree of subtlety required for a gaze-direction application, the more accurate/sensitive the method of eye-tracking should be.

formation. This could potentially be a cost saving approach especially for high viewer volume segments such as Super Bowl commercials which currently cost about $2.4M for 30 seconds of air time.

Related to all these applications is the question of image understanding and information recall. We have demonstrated the ability to guide a viewer's gaze to specific image locations. This however does not necessarily mean that the viewer fully processed the visual details of those regions or remembered them. Future studies would probe just how much information is retained. These studies would need to be content and/or application specific. Furthermore, since the experiment described in this paper attempted to guide the viewer's gaze to relatively unimportant regions of the image, additional experiments are necessary to determine if similar gains can be found for already prominent areas or for areas that have already been viewed.

There are also several other potential avenues of future research. We believe that the effectiveness of the gaze directing technique would be improved by making it more adaptive to the observer's viewing configuration. For example, by adjusting the intensity of the modulation and/or the size of the modulated area based on the distance between the current fixation point and the desired fixation point, the modulations can be made more salient for larger distances or more subtle for smaller distances. For larger distances, adapting the modulations in this manner will lead to quicker detection by the peripheral vision and also a greater degree of accuracy when the eye saccades to fixate on the modulated region. We also plan to experiment with other types of modulation such as sharpness, contrast, texture, and the introduction of noise.

An emerging area of research that explores the use of subliminal cues to trigger unconscious neural processing may offer a solution to the problem of reduced perceived quality that was observed for the modulated images. The approach that is currently being used presents subtle stimuli to the peripheral regions of one eye while a continuously flashing mask is presented to the other eye. The presence of the flashing mask completely suppresses the conscious perception of subtle stimuli. However, the subtle stimuli still results in electrical signals that trigger neural processing [Bahrami et al. 2007]. While this particular approach is impractical for our work, because we would like to facilitate natural binocular viewing of the display, it does encourage the exploration of subliminal approaches for gaze-directing.

Finally, a natural extension of the research presented in this paper is the application of the gaze-directing technique to video. We believe that the presence of motion in the video will further suppress the subtle modulations. Hence, the modulated videos may not suffer from the perceived quality degradation that was noted in the modulated images. There are however, several challenges such as maintaining frame-to-frame coherency of the modulations and changing the position of the modulations to follow moving regions of interest, that will need to be overcome for this to be feasible.

As display technology advances and screen sizes continue to increase, there is a danger that viewers will be overwhelmed by the amount of visual information available at any one time. If their gaze is subtly directed as they navigate through the vast amount of information, the effects will be positive for both viewer and presenter.

APPENDIX

A. INSTRUCTIONS TO PARTICIPANTS (READ VERBATIM)

In addition to the following instructions which were read to the participants at the start of the study, participants were also provided with documentation showing that the study was reviewed and approved by the Institutional Review Board (IRB) at the institution where this study was conducted.

The purpose of this study is to investigate the role of eye movement in quality judgments of computer graphics imagery. Your participation in this study will involve looking at a sequence of forty images on a computer screen, and giving each a value from 1 to 10 depending on how good you judge the quality of the image to be. During this process your eye movements will be recorded by a small non-invasive camera.

There is a short calibration phase before the experiment begins. Once calibration begins, we ask that you sit as still as possible and minimize your head-movement as this can adversely affect the quality of the results

The risks to you as a participant are minimal. These include loss of time and boredom. The experiment will be run as efficiently as possible in order to minimize boredom and loss of time. The experiment should take no longer than 15 minutes to complete.

The results of this study may be published in scientific research journals or presented at professional conferences. However, your name and identity will not be revealed and your record will remain anonymous. Your name will not be used in any data collection, so it will be impossible to tell your answers from other peoples answers.

Participation in this study will not benefit you directly. Your participation may benefit others by increasing the realism and efficiency of computer graphics.

Participation is entirely voluntary. Additionally, you may choose to withdraw from this study at any time. If you decide not to participate or to withdraw from this study, there will not be a penalty to you.

REFERENCES

Bahrami, B., Lavie, N., and Rees, G. 2007. Attentional load modulates responses of human primary visual cortex to invisible stimuli. *Current Biology 17,* 6 (March), 509–513.

Baudisch, P., DeCarlo, D., Duchowski, A., and Geisler, W. 2003. Focusing on the essential: considering attention in display design. *Commun. ACM 46,* 3, 60–66.

Cole, F., DeCarlo, D., Finkelstein, A., Kin, K., Morley, K., and Santella, A. 2006. Directing gaze in 3D models with stylized focus. *Eurographics Symposium on Rendering*, 377–387.

Collewign, H., Erkelens, C. J., and Steinman, R. M. 1988. Binocular co-ordination of human horizontal saccadic eye movements. *Journal of Physiology 404*, 157–182.

DeCarlo, D. and Santella, A. 2002. Stylization and abstraction of photographs. In *SIGGRAPH '02: Proceedings of the 29th annual conference on Computer graphics and interactive techniques.* ACM Press, New York, NY, USA, 769–776.

Dodge, R. 1900. Visual perception during eye movement. *Psychological Review 7*, 454–465.

Dodge, R. and Cline, T. S. 1901. The angle velocity of eye movements. *Psychological Review 8*, 145–157.

DUCHOWSKI, A. T. 2002. A breadth-first survey of eye-tracking applications. *Behav Res Methods Instrum Comput 34,* 4 (November), 455–470.

GAJEWSKI, D. A., PEARSON, A., MACK, M. L., BARTLETT III, F. N., AND HENDERSON, J. M. 2005. *Lecture Notes in Computer Science.* Vol. 3368/2005. Springer Berlin - Heidelberg, Chapter Human Gaze Control in Real World Search, 83–99.

GOOGLE INC. Google^{TM} Image Search. http://images.google.com/.

HENDERSON, J. M. AND HOLLINGWORTH, A. 1998. Eye movements during scene viewing: An overview. In *Eye Guidance in Reading and Scene Perception*, G. Underwood, Ed. Oxford: Elsevier., 269–293.

HUEY, E. B. 1968. *The Psychology and Pedagogy of Reading.* Cambridge, MA: MIT Press. (Originally published 1908).

HURVICH, L. M. AND JAMESON, D. 1957. An opponent-process theory of color vision. *Psychological Review 64*, 384–404.

HUTCHINSON, T. E., WHITE, K. P., MARTIN, W. N., REICHERT, K. C., AND FREY, L. A. 1989. Human-computer interaction using eye-gaze input. *Systems, Man and Cybernetics, IEEE Transactions on 19,* 6, 1527–1534.

JACOB, R. J. K. AND KARN, K. S. 2003. *The Mind's Eye: Cognitive and Applied Aspects of Eye Movement Research.* Elsevier Science, Amsterdam, Chapter Eye Tracking in Human-Computer Interaction and Usability Research: Ready to Deliver the Promises (Section Commentary), 573–605.

KOSARA, R., MIKSCH, S., AND HAUSER, H. 2001. Semantic depth of field. In *INFOVIS '01: Proceedings of the IEEE Symposium on Information Visualization 2001 (INFOVIS'01).* IEEE Computer Society, Washington, DC, USA, 97.

LEVINE, J. L. 1981. An eye-controlled computer. Research Report RC-8857, IBM Thomas J. Watson Research Center, Yorktown Heights, N.Y.

LEVOY, M. AND WHITAKER, R. 1990. Gaze-directed volume rendering. *SIGGRAPH Comput. Graph. 24,* 2, 217–223.

LIVINGSTONE, M. 2002. *Vision and Art: The Biology of Seeing.* Harry N. Abrams, Inc.

LUEBKE, D., WATSON, B., COHEN, J. D., REDDY, M., AND VARSHNEY, A. 2002. *Level of Detail for 3D Graphics.* Elsevier Science Inc., New York, NY, USA.

MACKWORTH, N. H. AND MORANDI, A. J. 1967. The gaze selects informative details within pictures. *Perception and Psychophysics 2*, 547–552.

MANNAN, S. K., RUDDOCK, K. H., AND WOODING, D. S. 1996. The relationship between the locations of spatial features and those of fixations made during visual examination of briefly presented images. *Spatial Vision 10*, 165–188.

MCNAMARA, A., BAILEY, R., AND GRIMM, C. 2008. Improving search task performance using subtle gaze direction. In *APGV '08: Proceedings of the 5th symposium on Applied perception in graphics and visualization.* ACM, New York, NY, USA, 51–56.

MITCHELL, G. E. 2004. *Taking Control Over Depth of Field: Using the Lens Blur Filter in Adobe Photoshop CS.* The Light's Right Studio. http://www.outbackphoto.com/workflow/wf_51/essay.html.

OGDEN, T. E. AND MILLER, R. F. 1966. Studies of the optic nerve of the rhesus monkey: Nerve fiber spectrum and physiological properties. *Vision Research 6*, 485–506.

OSTERBERG, G. 1935. Topography of the layer of rods and cones in the human retina. Acta Ophthal. suppl. 6, 11-97.

O'SULLIVAN, C., DINGLIANA, J., AND HOWLETT, S. 2003. Eye-movements and interactive graphics. *The Mind's Eyes: Cognitive and Applied Aspects of Eye Movement Research*, 555–571. J. Hyona, R. Radach, and H. Deubel (Eds.).

PARKHURST, D. AND NIEBUR, E. 2003. Scene content selected by active vision. *Spatial Vision 16*, 125–154.

RAYNER, K. 1975. The perceptual span and peripheral cues in reading. *Cognitive Psychology 7*, 65–81.

SALVUCCI, D. D. AND GOLDBERG, J. H. 2000. Identifying fixations and saccades in eye-tracking protocols. In *ETRA '00: Proceedings of the 2000 symposium on Eye tracking research & applications*. ACM, New York, NY, USA, 71–78.

SPILLMANN, L. 1990. *Visual Perception: The Neurophysiological Foundations*. Academic Press, San Diego.

YAHOO! INC. flickr™. http://www.flickr.com/.

YARBUS, A. L. 1967. *Eye Movements and Vision*. Plenum Press. Translated from Russian by Basil Haigh. Original Russian edition published in Moscow in 1965.

YOUNG, R. A. 1987. The Gaussian derivative model for spatial vision: I. Retinal mechanisms. *Spatial Vision 2,* 4, 273–293.