Attentive Display: Paintings as Attentive User Interfaces

David Holman, Roel Vertegaal, Changuk Sohn, and Daniel Cheng
Human Media Lab
Queen's University
Kingston, ON, Canada, K7L 3N6
{holman,roel,csohn,cheng}@cs.queensu.ca

ABSTRACT
In this paper we present ECS Display, a large plasma screen that tracks the user's point of gaze from a distance, without any calibration. We discuss how we applied ECS Display in the design of Attentive Art. Artworks displayed on the ECS Display respond directly to user interest by visually highlighting areas of the artwork that receive attention, and by darkening areas that receive little interest. This results in an increasingly abstract artwork that provides guidance to subsequent viewers. We believe such attentive information visualization may be applied more generally to large screen display interactions. The filtering of information on the basis of user interest allows cognitive load associated with large display visualizations to be managed dynamically.

Categories & Subject Descriptors: H.5.2. [Information interfaces and presentation]: User Interfaces.

General Terms: Human Factors.

Keywords: Paintings, Eye Tracking, Attentive User Interfaces.

INTRODUCTION

Recent increases in the size of computer displays represent a challenge for the management of on-screen visual information. Firstly, large screens may contain more information, and thus more visual clutter. Secondly, critical information on the display may be missed when the display area is larger than the user's field of view [1]. In this paper, we discuss how large displays may employ knowledge of the user's looking behavior in addressing the first of these two problems. When viewing an image, users tend to distribute their eye gaze in a manner that reveals the relevance of the presented information. Because there are no reliable techniques for tracking user eye movements at a distance, there are currently no large screens that utilize user viewing behavior. Knowledge about the user's looking behavior may be applied in the following ways:

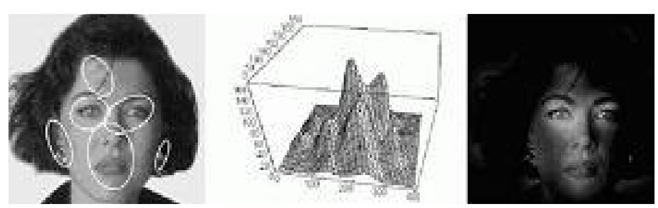
CHI 2004, April 24–29, 2004, Vienna, Austria.
ACM 1-58113-703-6/04/0004.

1. *Information Filtering*. Information on a display can be filtered on the basis of user fixations at an area, with unattended areas being abstracted or removed [3]. Such displays are known as Gaze-contingent displays, and have been deployed to optimize graphics rendering capacity [4].
2. *Dynamic Interaction*. Areas that attract user attention can be used to trigger specific sound, motion, or other meaningful responses [7]. Displays can ensure information is visible by presenting it within field of view.

In this paper, we present ECS Display, a 50" plasma screen capable of tracking user eye movements from a distance, and without calibration. We discuss how we applied ECS Display in an Attentive Art exhibit, with interactive artworks that explore the *Information Filtering* approach. Artworks displayed on ECS Display respond directly to user interest by visually highlighting areas of the artwork that receive eye fixations, and by attenuating areas that receive little interest. This results in an increasingly abstract artwork that provides visual interpretation and guidance to subsequent viewers.

BACKGROUND

Painters have been long known to use lighting and detail to guide the eyes of observers to significant areas of the artwork. By doing so, the artist aims to reduce the scene to its most essential elements, thus lowering cognitive demand while aiding comprehension of the image [14]. In Attentive Art, this process is made interactive. By responding to user attention, attentive artworks become a form of Attentive User Interface (AUI).



(a) (b) (c)

Figure 1. Salient features in complex imagery [8].

AUIs are user interfaces that sense and process user attention with the explicit purpose of tuning information presentation to the cognitive resources of the user. To achieve this, AUIs typically sense non-verbal cues such as presence, proximity, body orientation and eye gaze of the user [9]. In particular, eye fixations have been shown to correlate well with the locus of attention [10]. Yarbus found early on that when observing an image, fixation paths depend not just on the content of the image, but also on what questions participants had to answer about the image [13]. Pomplin, Ritter, and Velichovksy [8] used eye fixation data to investigate how participants interpreted ambiguous imagery. Figure 1 shows how they represented mean attention within their images using 2D Gaussian filters. For each fixation point in 1a, a Gaussian of unit height was added at its corresponding location. The Gaussian distributions were summed to illustrate overall attention (figure 1b). These distributions were normalized and applied to control the luminance in the figure (fig. 3c).

Public Eye Tracking Exhibits

Several researchers have also experimented with public eye tracking exhibits in order to explore and understand perception of complex imagery. Buquet et al. developed Eye-Follower [2], a museum exhibit that mapped eye fixations by visitors on a painting. They concluded eye gaze recordings can be applied to a broad range of public installations, with minimal supervision. With the "Telling Time" project, Wooding [11,12] carried out a study of eye movements of gallery visitors observing paintings during an exhibit. Eye gaze fixation data was used to generate images similar to those in Figure 1c. These studies inspired us to create an exhibit similar in nature, but with interactive content and real-time visualization maps. In Wooding's study, users were required to sit down behind a computer screen, go through a calibration process, and keep their heads still. We tried to eliminate this process by removing calibration, as well as the need to sit down altogether.

ATTENTIVE ART SCENARIO

Figure 2 shows an image of a painting displayed on a large eye tracking display positioned similarly to a painting found in a gallery. Our Attentive Art software measures user fixations on the display. The software continuously processes the image such that areas that do not receive attention fade away over a user-specified interval. The software emphasizes areas of the display that do receive eye fixations by increasing their luminance level. Over time, the result is a visual map of areas in the image that are statistically most likely to be observed. This process alters the image, guiding and influencing subsequent users' perception. It allows passive observers to become subtle participants in the artistic process, by emphasizing through a vote of visual interest what aspects of the artwork are most salient. For our exhibit, we commissioned artists to deliver ambiguous imagery in their artworks, as portrayed by the rendition of "Infant" by artist Mirjam Netten (see Figure 2a). The artist's intent here was to create a visually challenging scenario that allows viewers to initially

Figure 2. Mirjam Netten's *Infant* as an Attentive Art piece, Human Media Lab, Kingston. Initial image (*a*) and image after multiple viewings (*b*).

appreciate the visual composition without recognizing what is represented. By gradual filtering of the content over time, the outline of the represented objects is enhanced. This facilitates recognition of the semantics of the artwork and recognition of the objects portrayed, in this case a pet nursing a sibling (see Figure 2b) [8].

IMPLEMENTATION

One of the central components of the Attentive Art system is a novel calibration-free technique for tracking user eye fixations. First, we will discuss how our ECS Display uses this technique to track eye fixations on a large display surface. We then present our visual design rationale, discussing how we used user eye fixation data in the filtering of information presented on this display.

Eye Contact Sensing (ECS) Display

The tracking of eye fixations by visitors of an art gallery represented two significant challenges. Firstly, simulating the arrangement of an artwork in a gallery required a large wall-mounted display with a distance of at least one meter to the observer. None of the currently available commercial remote (i.e., not head-mounted) eye trackers are capable of tracking users at distances beyond 2 feet. Current systems are also not capable of tracking fixations on displays larger than 17 inches. Secondly, current eye tracking systems require an initial calibration in order to correctly map user eye fixations to the coordinate system of the display. During calibration, users are forced to follow a calibration pattern displayed as dots on the screen. We felt such calibration requirement would affect the behavior of gallery visitors and reduce the usefulness of attentive art. The Eye Contact Sensing (ECS) Display addresses both of the above problems (see Fig. 4). The ECS Display builds upon prior work on Eye Contact Sensors (ECS) at our lab [9]. Eye Contact Sensors are eye trackers capable of detecting whether users are looking at a device, without calibration.

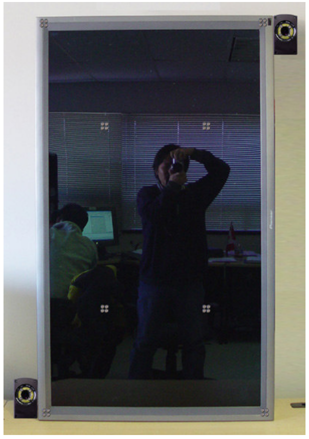


Figure 4. ECS Display with diagonal camera arrangement and LED markers in visible light.

Figure 5. ECS Display with diagonal camera arrangement and LED markers in infrared light. Note the illumination of the LED markers on the screen.

An ECS is composed of an infrared video camera with an on-axis ring of infrared Light Emitting Diodes (LED) surrounding the lens. Another set of LEDs is placed off-axis from the lens. The on-axis LEDs produce a bright pupil image in the eyes of an onlooker, while the off-axis LEDs produce a dark pupil image [5]. The on-axis and off-axis LEDs are flashed alternately with each camera frame. Rolling subtraction of the camera images allows for easy detection of the onlooker's pupil. The LEDs also produce a reflection or glint on the onlooker's cornea. When this glint appears near the center of the pupils, the ECS determines that the onlooker is looking straight at the camera.

Our ECS Display applies the same principle to the calibration-free tracking of large surfaces. Here, multiple off-axis LEDs are placed across the display surface (see Fig. 5). Two 2MPixel cameras with on-axis illuminators are placed at diagonal corners of the display. Each camera has a field of view of approximately 1.8 meters at 2 meters distance. The off-axis LED markers on the display cause multiple corneal glints to appear in the images captured by the cameras. However, as with the original ECS, when a user is looking at an LED on the display, it appears centered in his pupil. The location of the LED on the display is determined through the geometrical arrangement of the LED marker reflections in the user's eye. Interpolation between the known locations of these markers allows our computer vision algorithm to determine where the user is looking. At a viewing distance of one meter, the algorithm yields point-of-gaze measurements with a resolution better than 2 inches on the display.

VISUAL DESIGN

We experimented with several visual parameters such as hue, saturation, and luminance of an area to emphasize or attenuate visual information in the artworks. The most obvious visual parameter, and the one most frequently used by artists, is luminance. Varying the luminance value of an area on the painting according to a Gaussian distribution creates an effect similar to that of illumination by a spotlight. This behavior forms a natural analog to the representation of points of interest [8,12]. Other parameters such as hue and saturation were discarded as they overly distorted the colors of the artwork.

Design Rationale

We based our visual design on the approach taken by Pomplun et al. [8]. In order to represent user attention, or the lack of it, we needed two filters:

1. *Illumination Filter.* This filter projects a two dimensional Gaussian distribution over the luminance values of the original image at the location of a user fixation. If the luminance value in the distribution is greater than or equal to the current level, its value is updated. This method is used to create a smooth spotlighting effect (see figure 6b). Here, the center area appears brightest and regions further out are diminished according to a bell curve.

2. *Darkening Filter.* To represent lack of attention, we chose to darken the entire image using a timed interval filter. This filter subtracts the luminance value of each pixel by one for each user-defined time interval, until the relative pixel luminance reaches a user-defined lower threshold. Figure 6 demonstrates the effect. The initial image is seen in Fig. 6a. Fig. 6b shows the image after applying the darkening filter with an interval of 1 second for 1 minute. After three minutes, parts of the image reach the lower threshold, with areas that did not receive eye gaze appearing faded (Fig. 6c).

There are two techniques to reset the luminance value of the faded image represented in Figure 6c. In the first technique, luminance value is blended back to the original immediately, allowing new viewers to immediately observe the underlying image. In the second technique, the relative luminance value for each part of the picture increases with overall user fixation time for that area. We chose to the first technique because it allows users to more readily view contents of the faded artwork.

(a) 0 seconds (b) 1 minute (c) 3 minutes

Figure 6. Illuminating and Darkening. Mirjam Netten's *The Frog*, Human Media Lab. Queen's University.

Implementation

Illuminating and darkening the image in real time is a computationally demanding task. All image calculations were therefore implemented in C++ using the Open Computer Vision (OpenCV) Library [6], allowing the image to be filtered without any perceivable lag. We deployed one thread to darken the image at a predetermined time interval, with another thread gathering eye fixation data and calculating the Gaussian illumination filters for those areas. The width of the Gaussian spotlight is typically set to approximate the resolution of the macula [8]. With approximately 10 cm at 1 m distance or 5 degrees, this corresponds well with the resolution of the ECS Display.

Applications to Large Display User Interfaces

While the examples discussed in this paper are specific to artistic representation, we believe the underlying principles may be applied more generally to large screen display interactions. The cognitive load of visual information representation on large displays may be actively managed by dynamically filtering information on the basis of user interest.

CONCLUSION

We presented ECS Display, a large screen that tracks the user's point of gaze from a distance without any calibration. We discussed how we applied ECS Display in the design of Attentive Art. Artworks displayed on the ECS Display respond directly to user interest by visually highlighting areas of the artwork that receive attention, and by darkening areas that receive little interest. This results in an increasingly abstract artwork that provides guidance to subsequent viewers. We believe the dynamical filtering of information on the basis of user interest allows cognitive load associated with large display visualizations to be managed more effectively.

REFERENCES

1. Baudisch P. et al. Keeping things in context: a comparative evaluation of focus plus context screens, overviews, and zooming. *Proc. CHI 2002*, ACM Press (2002), 492-493.
2. Buquet, C., Charlier, J.R. and Paris, V. Museum Application of an Eye Tracker. *Medical Biomedical Engineering Computing,* 26 (1998), 277-281.
3. Geisler, W. S. and Perry J. S. Real-time simulation of arbitrary visual fields. *Proc. ETRA 2002*, ACM Press (2002), 83-87.
4. Loschky C.L. and McConkie, G.W. User performance with gaze contingent multi-resolutional displays. In *Proc. ETRA 2000*, ACM Press (2002), 97-103.
5. Morimoto, C. H., Koon., D., Amir, A. and Flickner, M. Pupil detection and tracking using multiple light sources. *Image and Vision Computing* 18, 4, 2000.
6. OpenCV. http://sourceforge.net/projects/opencvlibrary/.
7. Paradiso J. A. Tracking contact and free gesture across large interactive surfaces. *Special Issue on A Game Experience in Every Application.* Communications of the ACM 46, 7, 2003.
8. Pomplun, M., Ritter. H. and Velichkovsky B. Disambiguating Complex Visual Information: Towards Communication of Personal Views of a Scene. *Perception*, 25, 1995, pp. 931-948.
9. Shell, J., Selker, T., and Vertegaal, R. Interacting with Groups of Computers. *Special Issue on Attentive User Interfaces*, Communications of the ACM 46, 3, 2003.
10. Vertegaal, R. and Ding, Y. Explaining Effects of Eye Gaze on Mediated Group Conversations: Amount or Synchronization? In *Proc. of CSCW 2002.* New Orleans: ACM Press, 2002, pp. 41-48.
11. Wooding, D. Eye Movement of Large Populations I: Implementation and performance of an autonomous public eye tracker. *Behavior Research Methods, Instruments, & Computers 34,* 2, 2002, pp. 509-517.
12. Wooding, D. Eye Movement of Large Populations: II. Deriving regions of interest, coverage, and similarity using fixation maps. *Behavior Research Methods, Instruments, & Computers 34,* 2, 2002, pp. 518-528.
13. Yarbus, A.L. *Eye Movements and Vision.* Plenum Press, New York, USA, 1967.
14. Zeki, S. *Inner Vision: An Exploration of Art and the Brain*. Oxford University Press, Oxford, UK, 1999.