Attention Meter: A Vision-based Input Toolkit for Interaction Designers

Chia-Hsun Jackie Lee
MIT Media Laboratory
20 Ames ST. E15-324
Cambridge, MA 02139 USA
jackylee@media.mit.edu

Jon Wetzel
MIT Media Laboratory
20 Ames ST. E15-322
Cambridge, MA 02139 USA
jwwetzel@mit.edu

Chiun-Yi Ian Jang
Graduate Institute of Architecture
National Chao Tung Univ.
Hsin-Chu, Taiwan
yvainj@yahoo.com

Yang-Ting Bowbow Shen
Information Architecture Lab.
National Cheng Kung Univ.
Tainan, 700 Taiwan
bowbowshen@yahoo.com.tw

Ting-Han Daniel Chen
Graduate Institute of Architecture
National Chao Tung Univ.
Hsin-Chu, Taiwan
daniel@arch.nctu.edu.tw

Ted Selker
MIT Media Laboratory
20 Ames ST. E15-322
Cambridge, MA 02139 USA
selker@media.mit.edu

CHI 2006, April 22–27, 2006, Montréal, Québec, Canada.
ACM 1-59593-298-4/06/0004.

Abstract
This paper shows how a software toolkit can allow graphic designers to make camera-based interactive environments in a short period of time without experience in user interface design or machine vision. The Attention Meter, a vision-based input toolkit, gives users an analysis of faces found in a given image stream, including facial expression, body motion, and attentive activities. This data is fed to a text file that can be easily understood by humans and programs alike. A four day workshop demonstrated that some Flash-savvy architecture students could construct interactive spaces (e.g. TaiKer-KTV and ScreamMarket) based on body and head motions.

Keywords
Multimodal interaction, visual attention, interactivity

ACM Classification Keywords
C.3 Special-Purpose and Application-based System.

Introduction
Visual works of art are often sitting quietly inside galleries or museums. It is important to understand how people react to the artwork and provide feedback at the

right time. Krueger presented an artificial reality approach for digital art installations in which cameras interact with the viewers [1]. His system took racks of equipment and was tuned to a particular interaction. Can current technology make this easier? Traditionally, providing a dynamic interactivity environment has been difficult. Designing novel visual experiences usually involves understanding how people are paying attention. Since attention is a limited resource [2] and all exhibits in galleries or museums have their own stories to tell, the visual interface must be designed carefully. Visual attention can be tracked and measured by understanding the patterns of eye gestures [3]. Attention-based augmentations can be deployed to create a digitally-switchable domestic environment [4]. ScanEval [5] is a toolkit that processed eye movement and provided a real-time attention assessment and data summary that could be used for a wide variety of purposes, including user interface design. The expense and effort to set up systems such as Seeing Machines' FaceLAB [6] has limited applications so far. This paper demonstrates a technique which takes available face tracking software from the Intel Open Source Computer Vision (OpenCV) libraries [7], augments it to perform facial gesture and expressions recognition, and then creates a simple interface for Adobe Flash [8], which non-programmers can use as a simple develop environment for augmented reality and spatial interfaces.

Attention Meter

Attention Meter is a Visual C++ program that measures attention using camera-based input. In Figure 1, different levels of attentive engagement (i.e. passing by, glancing, standing and watching, reading carefully, and engaging) can be observed by monitoring people's behavior patterns using a camera. The input is a video stream from a camera mounted near or within the target of attention. The camera is positioned such that subjects attending to the target are looking almost directly at it, allowing the attention meter to analyze their attention based on cues from their faces.

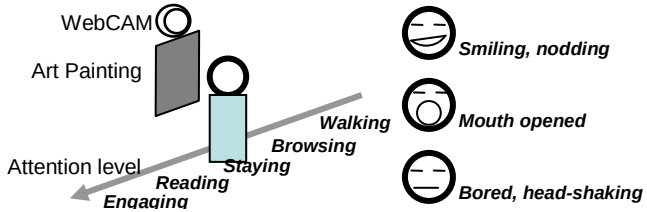

Figure 1: The system measures the attention level of people by using computer vision techniques to monitor their head movement, eye blinking frequency, and proximity.

Face Tracking

Attention Meter monitors and analyzes faces found in the camera view. Each frame taken from the video stream is run through a face detection algorithm from the Intel Open Computer Vision (OpenCV) library, as shown in Figure 2. This algorithm gives us the location and sizes of all faces in the image that are turned towards the display.

Head Movements: Large Motion, Nodding, and Shaking

By keeping track of the position of individual faces from frame to frame, the Attention Meter can detect when faces are moving laterally with respect to the target of attention. Using a finite state machine to analyze sequences of small movements, the Attention Meter also recognizes the smaller gestures of nodding and shaking. A further improvement could be to incorporate size,

allowing the detection of movement towards or away from the target as well.

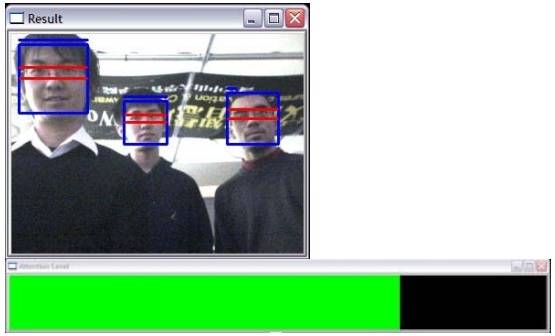

Figure 2. The Attention Meter calculates Attention Scores as a result of reasoning human behaviors from head and eyes movements. Individual Attention Scores are displayed above the face in the image window, while the group Attention Score is displayed as a large green bar.

Facial Expressions: Blinking and Mouth Position
By using basic knowledge of the structure of the face and looking for the distinctive brightness gradients of the eye, the Attention Meter can quickly find and detect eyes in faces, and over several frames measure the face's blink rate. One feature in development is detecting expressions of the mouth. In a manner similar to the eyes, the position of the mouth is determined.

Attention Scores
Every face being tracked is given an attention score which varies over time. The score starts at 0 and increases up to some predefined maximum as the face exhibits more attention. The individual scores are then summed together to form a group attention score. Remaining still allows the score to increase, while lateral motion halts it. Nodding and/or shaking, moving closer to the target (becoming larger), and blinking less often (eyes visible more often) will also increase the attention score. Various constants affecting the attention score calculation can be set by the designer using the Attention Meter's GUI at runtime.

Text Interface to Adobe Flash
The Attention Meter also outputs a summary of its collected data into a plain text file, which can be read by many other programs, including Adobe Flash. This data includes the group attention score, total number of faces, and for each individual face: coordinates attention score, size (proximity) and position, blink rate, and whether the face is moving laterally, nodding, or shaking. A single function call in Flash will read these variable/value pairs into the local environment, allowing developers to access the input data.

User Experience in Design Workshop

Workshops and educational forums are often created to bring new kinds of techniques and technologies to other communities. Asian Reality design workshop 2005 was used to test whether flash developers without a background in computer science could make cutting edge interactive demonstrations in a few days. The following examples show how these architecture students transcended their lack of technical background to create a real-time physical interaction with digital arts. In our case, one goal was to see if there is a new approach to design that can work for a group that has very little experience or background in creating computer interactive systems.

The Attention Meter system was demonstrated as a toolkit for a four-day design workshop. 23 students who

have design-related background (i.e. architecture, design, and art) without any formal training of computer science were divided into 8 groups for quick prototyping ideas. Students were given one three-hour lecture with tutorial for understanding immersive and interactive spaces and how to use the tool to integrate visual attention and multimodal interaction. Their assignment was to explore and implement interactive installations in the context of a night marketplace in Taiwanese culture. Three groups of students quickly integrated the Attention Meter system into their proposals (i.e. ScreamMarket and TaiKer-KTV). Students believe that they could build interactive that took human figure, shape, number of people into account to build interactive. In the course of this project three interactive working prototypes in big physical spaces with cameras and projectors were demonstrated after four days.

The value of these projects is that these people never had been involved in building prototypes to demonstrate technology and new ways of interacting with computers. They had been involved with classic flash kinds of interactions in which a cartoon or a button interacts. These research-worthy projects done in four days with three or four people are striking.

First Example: ScreamMarket

ScreamMarket is an interactive night-market show that interacts with audience's attention and feedbacks. This system demonstrated how audience engaged with the performance by monitoring their visual attention and audio feedbacks. The interactive show is implemented in Adobe Flash with Attention Meter as an attention-based triggering mechanism.

ScreamMarket transformed the Taiwanese traditional night market experience into a virtual and simulated space. In the beginning, an image of stage in night marketplace is blurred, but it gets clearer when the audiences pay attention to it. If more people are gathering in front of the stage, the dancers will show up, as show in Figure 3. The audience can yell to respond to the stage and get visual feedback. By using a microphone, as the volume of the audience increases, the performers become more active and entertaining.

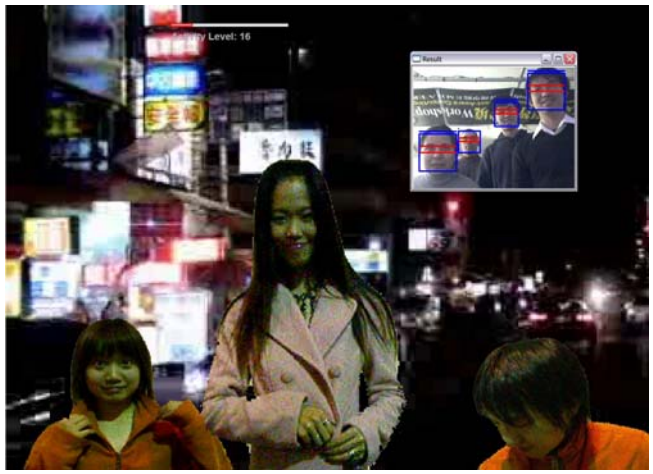

Figure 3. The Scream Market presents an animation of two Taiwanese girls if the audiences are paying attention to the stage. While the crowd is interested in the show and screams, the virtual girls will dance and entertain the audience.

The process of interaction is similar to the behavior that we watch the interactive show or bargain with the hawker in the night market. According to the method of interaction, the users are not simply viewers, but also performers in their own right. 30 people took turns in an

exhibition interacting with ScreamMarket. People were able to figure out and use it within a minute. It constrains output based on the regular environment noise, so that people may need to scream very loudly to interact with the Flash movie. The atmosphere of the interaction creates a realistic simulation of the night market.

Second Example: TaiKer-KTV

TaiKer-KTV enhances the interactivity of the performer and the environment for a more responsive and joyful karaoke space. TaiKer-KTV demonstrates how karaoke players engaged with the song can interact with the whole physical space based on their physical reaction and body movement. As shown in Figure 4, 'head-shaking dance' enriched the projection as a way to support group performance. The purpose is to amplify the group activity phenomenon in KTV and to create an interactive way to enhance the joyful and relaxing atmosphere.

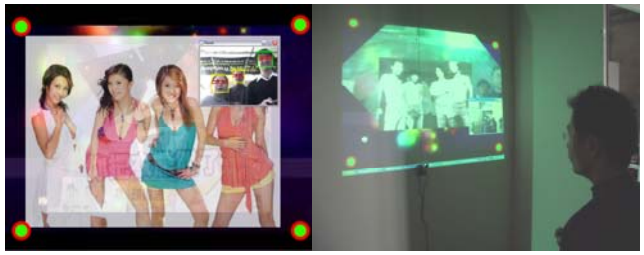

Figure 4. TaiKer characteristics were implemented into this music Flash MTV for people to actively interact with the karaoke environment.

In the implementation, a webcam is used to observe participants, a video projector outputs the media for singer-machine-audience interaction, and an interactive music video, created in Adobe Flash, was projected on a wall. The lyrics display includes a nodding head indicator to lead singing and dancing. The video is kept still and dull if it doesn't get enough attention. The more people are engaged, the more special visual effects are applied. If one moves his/her body, the image gets clearer. If one nods his/her head, image switches faster. If one shakes his head, the light flashes more dramatically. If there are multiple people participating, the media elements (i.e. symbols, visual effects, texts and recorded screaming voices) have an additive effect, creating a vivid mix of sound and imagery.

This system was completely compelling. Around 100 people came up to it and immediately began making strong movements to make the people on the screen dance. Participants spontaneously tried to get others to join. The ease and success at creating a feeling of inhibition in the user was striking.

Discussion

The process of creating interactive art can be intuitive and accessible**.** Interactive techniques for computer graphics should not only belong to computer scientists. We present the Attention Meter system which allows novice graphic designers to quickly make interactive spaces, not only using analysis of human facial behavior, but also through a calculated measure of attention, the Attention Score.

This experience demonstrates the ability for modern tools to allow visual designers to make innovative arts based on new interface technologies in a very short time. Visual artists and designers usually have limited tools to develop art installations that interact with the audience. ScreamMarket and TaiKer-KTV were built upon the

Attention Meter system and were all done in a four-day workshop. These two examples demonstrated the value and opportunity of giving visual designers good understanding and tools, so that even with limited technological background, they can still succeed to make interactive art installations.

The primary users of Attention Meter were graduate and undergraduate students with architectural background. Most of them have Flash experience that they can make visual arts quickly. Given instructions on using Attention Meter with Flash, they were able to pick it up quickly.

The Attention Score made it possible to monitor the visual attention of the audience. In the ScreamMarket example, designers used the scores to give different feedback depending on how many people are facing to the show. The performers are added on the stage if more people present. In the Taiker-KTV example, the overall group attention scores are used for determining the visual interaction between people and the screen. The visual effects get fancier and crazier when more people focus on the screen. Playing with the Taiker-KTV, people tend to exaggerate their actions when the actions become means of control, but the vision recognition system was implemented for normal actions like nodding or shaking head naturally. Thus over-exaggerated actions were usually not as effective as the normal ones. This obstacle was overcome through tuning of the user-side parameters. All groups were able to experiment and tune them so that they could find conditions that made the interactive experience consistent and successful.

The Attention Meter system shows the capability for a single camera to interpret attentive actions and transform the head, face, and eyes movements into computational models. It also demonstrates extensibility in interfacing with other software systems. The Attention Meter system is being extended with modularized sensors as a complete toolkit for designers to quickly prototype ideas. This toolkit demonstrates that research grade user interface tools can be designed in such a way that they allow novices to use them in innovative ways.

Acknowledgements

We thank Francis Lam, Kristy Liao, Scottie Huang, Yu-Dang Chen, Sheunn-Ren Liou, Mao-Lin Chiu, Sheng-Fen Chien in the Asian Reality workshop 2005 in Taiwan.

References

[1] Krueger, M., Gionfriddo, T., Hinrichsen, K., Videoplace- An Artificial Reality, in proceedings of CHI '85, pp.35- pp.40

[2] Pashler, H., The Psychology of Attention. Bradford Books. Reprint edition, 1999

[3] Selker, T., Visual Attentive Interfaces. In BT Technology Journal, Vol 22 No 4, (2004), 146-150.

[4] Bonanni, L., Lee, C.H., Selker, T., Attention-Based Design of Augmented Reality Interfaces, Ext. Abstracts CHI 2005, ACM Press (2005).

[5] Weiland, W., Stokes, J., Ryder, J., ScanEval - A toolkit for eye-tracking research and attention-driven applications, Human Factors and Ergonomics Society 42nd Annual Meeting, Chicago, IL 5-9 Oct. 1998.

[6] Seeing Machines-faceLAB, seeingmachines.com

[7] OpenCV- Intel Open Source Computer Vision Library, http://www.intel.com/

[8] Adobe Flash, http://www.adobe.com/