Eye Contact Sensing Glasses for Attention-Sensitive Wearable Video Blogging

Connor Dickie, Roel Vertegaal, Jeffrey S. Shell, Changuk Sohn, Daniel Cheng and Omar Aoudeh
Human Media Lab, Queen's University
Kingston, ON K7L 3N6, Canada
{connor, roel, shell, changuk, cheng, aoudeh} @ cs.queensu.ca

ABSTRACT
We present ECSGlasses: eye contact sensing glasses that report when people look at their wearer. When eye contact is detected, the glasses stream this information to appliances to inform these about the wearer's engagement. We present one example of such an appliance, eyeBlog, a conversational video blogging system. The system uses eye contact information to decide when to record video from the glasses' camera.

Author Keywords
Attentive User Interfaces, Ubiquitous Computing, Context-Aware Computing, Eye Tracking.

ACM Classification Keywords
H.5.2. Information interfaces and presentation (e.g., HCI): User Interfaces: Input devices and strategies.

INTRODUCTION
With the introduction of ubiquitous and wearable computing, human-computer interaction has fundamentally changed from a one-to-one to a one-to-many relationship. We are now surrounded by *many* computers, including cell phones, cameras and PDAs. The design of mixed-initiative dialogue with these appliances is still based on the notion that each appliance is the user's only computer. When notifying the user of incoming communications or information, each device acts in isolation, without regard for other requests for user attention. As a consequence, users are continually interrupted in their activities. Rather than tool-based interaction, the Attentive User Interface paradigm [4] uses human group communication as a guiding metaphor for designing mixed-initiative dialogue. In both human group communication and human-device communication, human attention is a limited and competitive resource. In both situations, there is a simultaneous interest in taking the floor to deliver a message. The primary difference is that humans know whether and when it is appropriate to interrupt others, and can sense which mechanism for interruption is least disruptive. While humans observe eye contact patterns to detect when to take or release the floor [1], digital devices have no such capabilities.

In this paper, we present ECSGlasses, a pair of eyeglasses capable of detecting when others look at their wearer. Information about eye contact provides digital appliances with knowledge about the engagement of their user. This information can be used to contextualize mixed-initiative dialogue, for example, in estimating when to mute cell phone rings [6] or in automatically updating instant messaging status. In our demonstration, we will show how wearable eye tracking glasses may be used to reduce demands on user attention. We will focus our demonstration on the scenario of recording and posting first-person video blogs using attention-guided editing.

PREVIOUS WORK
The design of the glasses was based on the attentive cell phone prototype [6]. Unlike Eye-R glasses [2], which used infrared transceivers to detect when two pairs of glasses align, the attentive cell phone registered actual eye gaze of others. It used an eye contact sensor mounted on a hat to inform potential callers when the recipient was involved in a face-to-face conversation. However, due to the parallax between eyes and camera, eye contact detection with this system was not precise [1]. Our design was also inspired by EyeTap [5] and LAFCam [3]. In EyeTap, a visible parallax was avoided by mounting the camera on a pair of glasses. It allowed video to be shot from a first-person perspective and streamed to the Internet. LAFCam [3] allowed video footage to be edited without human intervention. Edit points were selected by monitoring nonverbal utterances made by the videographer during filming.

ECSGlasses
Eye Contact Sensing (ECS) Glasses are a novel type of wearable input device for detecting looking behavior of other people without *any* observable parallax [1]. ECSGlasses consist of a standard pair of eyeglasses, augmented with an eye contact sensing camera [6]. We have developed both clear and dark glass versions, the latter shown in Figure 1. Unlike Eye-R [2], ECSGlasses do not require others to wear any machine-recognizable tag. Instead, the glasses deploy computer vision to search for human eye gaze directed toward the wearer's eyes. The current design employs an infrared sensitive analog camera mounted near the bridge of the wearer's nose. Surrounding the lens, and thus placed on-axis with the camera, is a set of Infrared Light Emitting Diodes (LEDs) (see Fig. 1). Two similar sets of LEDs are placed off-axis near the arms

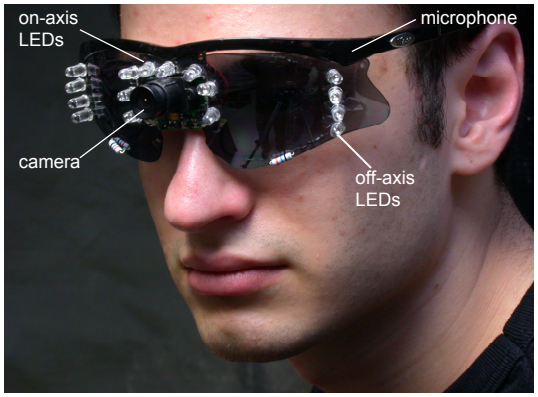

Figure 1. ECSGlasses: camera with LED illumination units.

of the glasses. The LEDs illuminate the scene in front of the camera. A human retina acts as a retro-reflector, sending light back in the direction from which it came. When the on-axis LEDs are illuminated, this produces a red eye effect in the pupils of onlookers. When the off-axis LEDs are illuminated the camera records a similar image, but with dark pupils instead. By syncing the LEDs with the camera clock, a bright and dark pupil effect is produced in alternate fields of each video frame. A simple algorithm finds any eyes in front of a user by subtracting the even and odd fields of each video frame. The LEDs also produce a reflection from the cornea of the eyes. These glints appear near the center of the detected pupils when the onlooker is looking at the eyes of the user, allowing the sensor to detect eye contact without calibration (see Fig. 2). Eye contact sensing range depends on focal length and camera resolution. With a 12 *mm* lens and a 320*x*120 effective resolution, range is 1 *m* with a 50 *cm* field of view. A digital version currently in testing provides sensing up to 4 *m*. The analog version of ECSGlasses uses an RJ45 cable that connects to a compact box worn at waist level. The box supplies power using a 12 volt DC battery, and contains the LED timer circuit board. The analog camera image is processed on a wearable computing system through a Firewire digitizer. This computer performs image analysis to isolate the pupils and glints and determine whether they are oriented at the wearer. After processing image data, the computer streams information about the coordinates and orientation of pupils found to connected appliances over a secure wireless network. Apart from pupils, ECSGlasses can also detect glyph patterns. This allows for identification of tagged objects or people.

Attentive Conversational Video Blogging

eyeBlog is an attentive video blogging application designed to automatically record conversations. It addresses one of the major problems in first-person video narratives: that of immediate editing of recordings. Digital camera appliances are typically fairly demanding of the attention of the user. This sometimes forfeits the experience of the event captured, particularly problematic when filming documentaries of one's life. By mounting a camera on glasses, systems such as EyeTap [5] facilitated streamed recordings that literally show the world through the eyes of

Fig 2. Processed Image: circles indicate eye contact detection.

the videographer. First-person perspective filming eases the creation of online video diaries, or video blogs, by allowing the user to monitor the event. To limit the amount of footage, however, archival systems such as video blogs require that the user either engage recording, or edit footage before posting. eyeBlog facilitates this process by using the signal from the eye contact sensor in the ECSGlasses to decide when to record and stream video from the ECSGlasses' camera unit. Running on the wearable computer unit, eyeBlog statistically processes eye contact data, recording images from the glasses whenever the percentage of eye contact over time reaches a preset threshold value. For additional robustness, eyeBlog also processes the audio energy from a microphone embedded in the glasses, determining when the user is speaking. When real-time editing is not required eyeBlog can be activated manually, recording engagement data with the video signal to inform subsequent editing. Applications are not limited to video blogging. Similar in function to opening one's office door, first-person perspective video can be streamed to allow others to monitor one's activities. ECSGlasses are also capable of communicating information about the user's face-to-face engagement to other online users through automated instant messaging status updates. This is useful when privacy is a concern. We are in the process of evaluating the latter technique.

CONCLUSION

We presented ECSGlasses: wearable eye contact sensing glasses that detect and report when people look at their wearer. One of the applications of these glasses is to allow automated editing of conversational video blogs.

REFERENCES

1. Argyle, M. and M. Cook. Gaze and Mutual Gaze. London: Cambridge University Press, 1976.
2. Selker, T, Lockerd, A. and J. Martinez. Eye-R, a Glasses-mounted Eye Motion Detection Interface. In Extended Abstracts of CHI'01. Seattle: ACM Press, 2001, pp. 179-180.
3. Lockerd, A. and F. Mueller. LAFCam: Leveraging Affective Feedback Camcorder. In Extended Abstracts of CHI'02. Minneapolis: ACM Press, 2002, pp. 574-575.
4. Shell, J. S., Selker, T., and R. Vertegaal. Interacting with Groups of Computers. *Communications of the ACM* Vol. 46 No. 3. New York: ACM Press, 2003, pp. 40-46.
5. Mann, S. EyeTap Technology for Wireless Electronic News Gathering. In Mobile Computing and Communications Review Vol 3. No. 4. ACM Press, 1999, pp. 19-25.
6. Vertegaal, R., Dickie, C., Sohn, C. and M. Flickner. Designing Attentive Cell Phones Using Wearable Eyecontact Sensors. In Extended Abstracts of CHI'02. ACM Press, 2002, pp. 646-647.