

EyePliances and EyeReason: Using Attention to Drive Interactions with Ubiquitous Appliances

Jeffrey S. Shell, Roel Vertegaal, Aadil Mamuji,
Thanh Pham, Changuk Sohn, and Alexander W. Skaburskis

Human Media Lab

Queen's University

Kingston, ON, Canada K7L 3N6

{ shell, roel, mamuji, pham, csohn, skaburs } @cs.queensu.ca

ABSTRACT

We present three prototype appliances that detect and respond to human attention. The Attentive Television pauses its feed when nobody is watching, and resumes playback when it receives eye contact; AuraLamp 'listens' and reacts to voice commands only if a person is looking at it; EyeProxy notifies its user of an incoming call by making eye contact, instead of producing a disruptive ring. These processes are managed by the EyeReason system, which acts as an interaction gatekeeper, limiting the distracting intrusions caused by attention seeking digital appliances. By recognizing and attenuating distractions, the EyeReason network preserves and fosters continuity in mixed interactions involving humans and devices.

Keywords Attentive User Interfaces, EyePliances, Context-Aware, Ubiquitous Computing, Notification Systems.

INTRODUCTION

Our attention is increasingly being fragmented by constant interruptions from digital devices. These devices relay volumes of email, instant messages, phone calls and appointment notifications without any regard for the user's activities. This produces an intricate web of annoying 'attention grabbers' within which a user can easily become entangled. By coordinating communications on the basis of human attention, devices may ultimately engage in more polite and respectful interactions with users [7]. To accomplish this, devices are augmented with eye contact sensors [9] that determine simply whether someone is looking at them. In multi-user, multi-device scenarios, the EyeReason system reduces the burden of interaction overload by negotiating the priorities of competing requests for attention from both remote contacts and digital devices. EyeReason also affords a cost efficient, extensible approach to enhance existing devices by adding gaze and

speech input. We will briefly describe the justification for using gaze as the primary source of attentional input, followed by illustrative descriptions of the prototypes.

Visual Attention and EyePliances

Of the eight nonverbal cues used to regulate turn taking in human group communication, only eye gaze cross-culturally indicates to *whom* the speaker wishes to yield the floor [8]. However, because it is difficult and often unnatural to overload the visual channel using an eye actuated selection mechanism, eye gaze is, in general, a poor means for control [4]. Nevertheless, it does convey interest in the target, and has been shown in wizard of oz experiments to implicitly select the target of spoken commands [1,5]. EyePliances, or appliances that detect and respond to human visual attention, use visual interest to disambiguate deictic targets in speech-enabled environments [6]. By doing so, EyePliances allow a turn taking process between humans and devices that is analogous to the method of speaker exchange found in human group communication [7]. We will discuss three prototypes that illustrate different aspects of EyePliances.

Attentive Television

The Attentive Television (fig 1b) uses an eye contact sensor to determine whether someone is watching it. If nobody is watching, the Attentive Television pauses its feed. When a viewer returns, the program resumes. This example demonstrates a scenario where the device implies the context of activity. Since a television serves only one functional purpose – to be watched, it simply requires eye gaze as input in order to be attentive to user interest. This concept can generalize to other devices that are fitted to use visual cues of attention to perform meaningful actions, such as automatically putting an unattended computer system to sleep, reducing power consumption.

AuraLamp

AuraLamp (fig 1a) illustrates gaze and speech enabled EyePliances. It is a lava lamp augmented with an eye contact sensor and speech recognition capability. By looking at the lamp, a person indicates attention to the device, thereby activating its speech vocabulary. The lamp

responds to the two actions it is capable of – turning on and turning off. In this case, the lexicon is implicit in the device, and the vocabulary is both small and intuitive. AuraLamp is a model for how we may use eye gaze with speech to interact with traditional household appliances.

EyeProxy

EyeProxy [1], seen in figure 1c, is a surrogate that represents a remote contact person. EyeProxy shows how a device can demonstrate its attention to a user by means of the anthropomorphic cue of eye gaze. It consists of an eye contact sensor mounted between two actuated eyeballs. When a remote contact wishes to engage in a phone conversation with the user, EyeProxy physically searches for the user's eyes. This movement in the periphery of the user's attention space triggers a visual saccade. The saccade allows the user to see which EyeProxy is moving, thus identifying the caller in a less disruptive manner than a ringing telephone. If the user fixates on the EyeProxy, in effect sharing eye contact, then a voice connection is established. This less intrusive method of communication is particularly important in environments where immersive, creative work requiring the continuity of thought is likely. An audible ring can derail a person's thought process, preventing completion of an initiated task, which can contribute to increased stress, further inhibiting a return to a productive mindset. EyeProxy allows requests for attention that minimally infringe upon the user's attention space.

EyeReason

The EyeReason system coordinates communications among many EyePliances and remote interlocutors by keeping track of the user's activities. It operates as a centralized server that EyePliance clients connect to. The EyeReason system monitors the devices in focus, and assigns priorities to each task on the basis of activity levels and duration. These constitute the user's level of engagement, which is considered in conjunction with the implicit importance of the task and the user's history with the task. This information is combined using Bayesian reasoning to attribute a value to the current interactions. This value is compared with competing requests for attention, such as an email from a contact. The email would be assessed a statistic using the mean response time and frequency of responses to the sender, in the tradition of Horvitz's Priorities system [1]. If the email is of higher importance, it is relayed to the device receiving the most amount of attention from the user that is capable of carrying such a message. In the case that the user is not using such a device, collocation is the default forwarding criteria. EyeReason communicates the users attentive status to her buddy list of personal contacts through both her Attentive Messaging System (AMS), and her Attentive Cell Phone [9] depending on the relationship between her buddy and the set of current tasks. For example, in a work scenario, an employee's boss would have elevated interruptive priority over social contacts. In the case of emergency situations these levels can be explicitly overridden.



Fig. 1. (a) AuraLamp, (b) Attentive TV, (c) EyeProxy
Videos can be found at http://hml.queensu.ca/hml_videos.html.

The purpose of EyeReason is to limit unnecessary interruptions, give remote interlocutors a sense of what activities they are intruding upon, and provide a facility to coordinate communications among EyePliances using a generalized model of user attention.

The EyeReason architecture simplifies the process of augmenting a standard appliance with gaze and speech capability. By embedding an eye contact sensor in an appliance and specifying an appropriate XML speech grammar, a device instantly becomes an EyePliance. If the appliance receives eye contact, a wireless headset processes speech commands using the XML lexicon specified in EyeReason to either perform tasks which can either be processed wirelessly through an X10 device, or directly interfaced into the appliance, depending on its construction. If neither is possible, then EyeReason still recognizes that a user is engaged with the device, holding particular attentive attributes. Because speech commands are processed through the centralized server, new forms of attentive interactivity are permitted without increasing the complexity of each appliance.

REFERENCES

1. Horvitz, E., Jacobs, A., and Hovel, D. Attention-sensitive alerting. In *Proceedings of UAI'99*. Stockholm: Morgan Kaufmann, 1999. pp. 305-313
2. Jabarin, B et al. "Establishing Remote Conversations Through Eye Contact With Physical Awareness Proxies." In *Extended Abstracts of CHI 2003* Ft.Lauderdale: ACM Press, 2003. pp 948-949.
3. Maglio, P. et al. Gaze and Speech in Attentive User Interfaces. In *Proceedings of the Third International Conference on Multimodal Interfaces (2000)*. Beijing, China.
4. Nielsen, J. Noncommand User Interfaces. *Commun ACM* 36(4), April 2003. pp. 83-99.
5. Oh, A. et al. Evaluating Look-to-Talk. In *Extended Abstracts of CHI 2002*, Minneapolis: ACM Press, 2002 pp.650-651
6. Shell, J. S. et al. EyePliances: Attention-Seeking Devices that Respond to Visual Attention. In *Extended Abstracts CHI 2003*. Ft.Lauderdale: ACM Press, 2003, pp 770-771.
7. Shell, J. S. et al. Interacting with Groups of Computers. *Commun ACM* 46(3), (March 2003), pp. 40-46.
8. Short, J., Williams, E., and Christie, B. The social psychology of telecommunications. London: Wiley, 1976.
9. Vertegaal, R. et al. Designing Attentive Cell Phones Using Wearable Eyecontact Sensors. In *Extended Abstracts of CHI 2002*. Minneapolis: ACM Press, pp. 646-647