Symbiotic Attention Management in the Context of Internet of Things

Shahram Jalaliniya
Malmö University
IoTaP Research Center
shahram.jalaliniya@mah.se

Diako Mardanbegi
Lancaster University
d.mardanbegi@lancaster.ac.uk

Thomas Pederson
Malmö University
IoTaP Research Center
thomas.pederson@mah.se

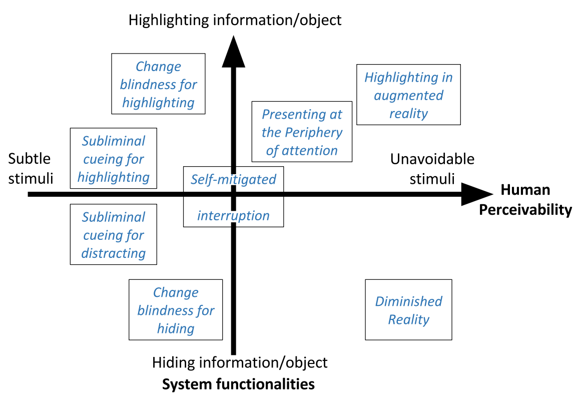

Figure 1: Symbiotic attention management quadrant.

UbiComp/ISWC'17 Adjunct, September 11-15, 2017, Maui, HI, USA

ACM ISBN 978-1-4503-5190-4/17/09.
https://doi.org/10.1145/3123024.3124559

Abstract
In this position paper we stress the need for considering the nature of human attention when designing future potentially interruptive IoT and propose to let IoT devices share attention-related data and collaborate on the task of drawing human attention in order to achieve higher quality attention management with less overall system resources. Finally, we categorize some existing strategies for drawing people's attention according to a simple symbiotic (human-machine) attention management framework.

Author Keywords
Eye tracking; Attention-aware systems; Internet of Things; smart environment

ACM Classification Keywords
H.5.m [Information interfaces and presentation (e.g., HCI)]: Miscellaneous

Introduction
The mechanism in our brains that controls our attention is optimized for our ancestors' most common living conditions 10.000 years ago or more. It is this attention mechanism which still today decides what we will consciously reflect on, and what to let slide. At the time of "design" there were obviously no Internet-connected interactive devices or other potential sources of sudden interruption except for perhaps

an aggressive lion attacking from the left, a warrior from a competing tribe from the right, or a rock falling down from a mountain hitting us in the head. While our attention mechanism still today is somewhat uselessly on its toes for such "Savannah events" and gradually prepares our bodies for action (e.g increasing adrenalin flow in case of danger), it fails miserably in coping with notifications that appear from perceptibly nowhere such as mobile phone call signals.

In this position paper we explore system design strategies for helping our biological attention mechanism deal with modern day interruptions. We adopt the widespread assumption amongst context-aware system designers that by better identifying the right time, place, and modality, smart IoT environments could potentially transmit more information (i.e. notifications about the status of otherwise imperceivable processes) to human agents without significantly disturbing ongoing tasks. Some of the inspiration for our approach comes from recent attempts to design for peripheral interaction [2].

Need for syncing and sharing attention information

We argue that to actually produce "graceful interruptions" and even more if we want to go further into actually *directing* human attention, we need to make our smart environments aware of the current focus and attention level of the target human agent(s). We see eye tracking as a starting point for future more advanced sensing and modeling approaches of human attention. In this position paper we discuss how IoT components could potentially share information about human agents' attention in order to gain qualitative advantages collectively (in this way, we don't need to completely fill our environments with expensive and carefully calibrated eye tracking devices), and we also discuss the need for an artificial attention management more in sync with how our biological attention mechanism seems to work. In the end of the paper we present a quadrant intended to inspire discussion on the workshop with regards to possible system design strategies for offering attention management in smart environments.

Human Visual Attention

The human brain processes an immense amount of information originating from body-external phenomena that is perceived by our senses (vision, auditory, olfactory, etc.) as well as from higher-level cognitive processes (our wishes, emotions, intentions) each second of our life. It has been estimated that around 95 percent of this brain processing is completely unconscious [8]. Furthermore, it is widely accepted that there is a regulatory mechanism, an "attention" mechanism, which selects among these many unconscious processes and lifts one or a few of them to our conscious awareness at a given point in time. The exact nature of this mechanism, and how it combines higher-level intentions with lower-level external stimuli to direct our conscious focused attention is still debated among psychologists. It is clear however, that human attention mechanism is affected by, and affect, the visual perception system: Eye movements in everyday life both form an important source of information to the attention mechanism for deciding where to focus our attention (the bottom-up information flow), as well as an indicator of what our current intentions are (the top-down information flow).

By tracking eye fixations in the environment that surrounds a given human agent, a computer system could, in theory, get an indication of what entities in that environment that matters to that human agent, in that moment.

Tracking Attention using Eye-Contact Sensors

Detecting eye movements in a smart environment is associated with several challenges such as need for many pre-calibrated eye tracking units in the environment or some computationally intensive computer vision algorithms when using head-mounted gaze tracking. We believe that the idea of using eye contact sensors similar to the one implemented in [13, 14] could be a better and more affordable solution for obtaining knowledge about user's attention in a smart environment. Compared to eye tracking solutions that estimate the user's gaze in her/his field-of-view, an eye contact sensor is meant to only provide a binary output indicating whether the user is looking at a specific area in their field-of-view or not. Distributed and connected eye-contact sensors could provide a pervasive interaction solution without any need for continuously monitoring the user's exact gaze point or what she/he is actually looking at in the environment. This could considerably optimize the computation which is more friendly for IoT-based infrastructures.

Distributing eye-contact sensors in the scene could also address some of the privacy issues that are associated with obtaining gaze data and eye information. Moreover, using eye contact sensors instead of remote or mobile gaze trackers enables us to move the smart objects in the environment without tagging and tracking the objects. In such smart environments, objects can communicate eye contact data with each other (as part of a bigger sensor network) or send it to the cloud in real-time. Subtly monitoring human agents' visual attention and their intention via distributed (visually) attention-aware objects could be a promising approach for improving future context models in the field of Context Aware Systems.

Although gaze information can be successfully used for supporting explicit control of devices around us ([10, 15, 1, 11]), many of our everyday interactions could also be facilitated by a system that more subtly monitors our natural eye movements (attentive user interfaces [9]). Different passive interaction strategies could be imagined in an attention-aware smart environment either when eye contact sensing objects are viewed as standalone and self contained objects or when they work in a co-operative scenario as part of a bigger infrastructure.

Changing state upon eye contact or eye leave

Smith et al. [14] proposed a number of interaction principles for using eye contact sensors for direct interaction with individual objects. Two main examples are when eye contact sensing objects could change their state upon direct eye contact (e.g. light goes ON after looking), or of particular interest, when the eye contact sensing object reacts upon eye leave (e.g TV pauses the movie when user looks away from screen).

Providing context

Eye contact can open up a communication channel through which the user can interact with the objects using other modalities such as speech, or other control devices. Here, user's attention provides context for an action by making the smart object stand by and ready for receiving commands (e.g. a voice command) from other modalities [14].

Activity Assistance

Eye-contact-aware objects could also work in a co-operative scenario where they share eye contact data with each other. On a higher level, with current development in IoT and ubiquitous computing, we envision a scenario where smart objects equipped with eye contact sensors allow us to monitor how visual attention is distributed in a smart environment. Information about human attention could be used as an indicator of intention and thus used by an intelligent system to prepare nearby computing devices for probable future requests from the side of the human agent, potentially reducing the need for the user to explicitly configure the devices. Such an automatic preemptive mechanism is

An example scenario of subtle highlighting:
A person is reading a book at home. According to the schedule of the person, she needs to call her dentist to book a time in 10 minutes. For avoiding an unwanted interruption, the system should be able to direct her attention towards the phone in a subtle way. The person and the phone are located in two different rooms; therefore to direct her attention from the book to the phone, several smart objects need to cooperate. First, the light located in her field of view flashes quickly (below the supraliminal threshold) but the user misses the first flash. Since the light is equipped with an eye contact sensor, it repeats flashing until the user looks at the light. The light communicates the eye contact with the TV. Then the TV tries to draw user's attention to the phone by displaying a visual cue. When she walks towards the TV, the phone is in her field of view and the phone reminds her to call the dentist.

very challenging to design and most likely to fail unless devices collaborate in both sensing visual attention, attracting visual attention (discussed in the next section of this paper), and collaborative coordination as to when and by what device this is done. By analyzing duration of users visual attention to a particular objects, the system could potentially identify whether the user needs supportive information in relation to a given device or not. In daily routine activities e.g. cooking, we usually follow a relatively fixed procedure. For some people with mental illnesses such as dementia, it is often not easy to remember even such routine procedures. In such cases, attention analysis could facilitate automatic recognition of the intended activity, detection of potentially missing steps, and providing supportive information.

Controlling Human Attention

The human brain is sometimes actually regarded to have several attention mechanisms but we will regard them as one in this discussion and focus on attention in relation to what is often called working memory. The working memory is highly volatile and contains basically everything we are consciously aware of at a given point in time: a handful selected long term memory items plus a handful selected current perceptions [7]. Any call for human attention (such as a notification from an interactive device) will alter the current composition of items in the working memory. It is obvious, then, that the better our artificial interruptive systems are at guessing the content of working memory of a given human agent, the better they will be in performing graceful interruption. But the fact is that we don't need to stop at just *monitoring* attention (working memory) - we could just as well try to more precisely *influence* its contents *in symbiosis* with the existing biological attention mechanism in the brain. This is our vision and what follows is a first attempt define some strategies for doing exactly this.

Symbiotic attention management strategies
In a smart environment with distributed interactive devices, each device can potentially initiate an interactive dialogue with a human agent. If the interaction request is not relevant to the task at hand it can be considered an unwanted interruption. One of the main goals of an artificial attention management mechanism would be to mitigate unwanted interruptions and also to assure that relevant notifications are perceived at the right moment. In short, such a mechanism should *suppress or hide* irrelevant information (e.g. [17]) and *highlight* relevant notifications (e.g. [3]). For example, if a human agent is not looking at a relevant smart device, the artificial attention management mechanism that tracks the human agent's eyes could highlight the smart device or somehow direct the visual attention towards the relevant smart object by e.g. flashing lights in the environment or proving visual cues on a head-mounted display if the person is wearing an augmented-reality head-mounted display.

Similar approaches to subtle notification can be taken to distract visual attention away from irrelevant pieces of information. This *highlighting* and *hiding* can be made extremely *subtle* or completely *unavoidable* (impossible to not notice) or something inbetween. By placing different types of attention control methods along the perceivability and highlight/hide dimensions, we get a quadrant (Fig. 1) which illustrates design alternatives for attention management in IoT-enabled smart environments. The vertical axis of the quadrant represents different strategies for artificial attention management available to system designers (system functionalities). The horizontal axis (perceivability) represent characteristics of human perception which we as system designers cannot affect.

We believe this quadrant could raise interesting discussions

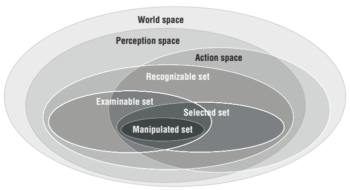

Figure 2: Precise tracking of available nearby stimuli generators: Systems intended to generate stimuli that lie in the spectrum between unconsciously noticeable and consciously noticeable would probably be facilitated by having precise spatial understanding of available actuators (e.g. visual displays and loudspeakers) in the visual and aural fields of view of the human agent. A human body-centric model similar to this figure suggested by [12] could help us categorize nearby objects with regards to human perception and action capabilities.

in the workshop about how to design different artificial attention management mechanisms in relation to humans' biological attention system. To illustrate the potential benefits of the proposed quadrant for this purpose, we briefly review some previous work in the light of it:

1. Subtle highlighting: perhaps the most subtle way of directing attention is subliminal cueing [5]. Aranyi et. al. [5] have shown the significant effect of subliminal stimuli on selection behavior of users in a virtual environments. Another slightly less subtle way of presenting a notification or highlighting a physical object in the human agent's field of view is to make use of the change blindness phenomenon [16]. In this method, the notification is displayed during the eyes' fast movements (saccades) resulting in a delayed perception due to the temporary blindness during saccades [16]. A third approach is to present the notification or highlight an object at the periphery of the user's attention, as investigated in *peripheral interaction* [6, 2]. As Figure 1 shows the perceivability of the message in each method determines how close to the extreme left side of the quadrant that method is located. We think all of these techniques could benefit from a network of attention-aware smart objects as discussed earlier in this paper.

2. Unavoidable highlighting: unavoidable notification (e.g. sound notifications) are the most dominant way of notifying users in smartphones, smart watches, and desktop computers. An example of unavoidable highlighting physical objects is the Attention Funnel system [3] that uses augmented reality user interface to direct the visual attention of a user towards a particular object in the environment.

3. Subtle hiding: attention control is not always about notifying users. Sometimes it is beneficial to hide irrelevant information. We notice any significant change in our field of view. One subtle way of hiding objects/information is to use the change blindness phenomenon [16] by e.g. removing irrelevant information from a near-eye display during saccades. Another approach to hide an irrelevant notification is to adjust the intensity of the stimuli based on the importance of the message tuned to how cognitively busy the recipient is. This approach (*self-mitigated interruption* [4]) probably constitutes the best example of a symbiotic combination of artificial biological attention management. It would also be possible to use subliminal cues similar to [5] for distracting users' attention from irrelevant information/object(s).

4. Unavoidable hiding: *diminished reality* (e.g. [17]) refers to hiding objects by superimposing the background texture on the object's image and is one of the classical approaches to visually remove physical object from real world.

Conclusions

While IoT is considered to be an ecology of interruptive computing devices, embedding eye contact sensor into the smart things seems to be an efficient way of designing attention-aware systems on the IoT platform. In this paper, we proposed a quadrant (Figure 1) including two important dimensions for integrating artificial attention control systems with our biological attention mechanism: one dimension explains perceivability of the stimuli from human point of view while the other dimension ranges system functionalities from hiding to highlighting information/objects. We used our proposed quadrant as a conceptual framework to review some of the previous attempts to control human attention.

REFERENCES

1. Kevin Purdy Alastair G. Gale. 2007. *The ergonomics of attention responsive technology*. Full research report.
2. Saskia Bakker. 2013. Design for Peripheral Interaction, PhD Thesis. (2013).

3. Frank Biocca, Arthur Tang, Charles Owen, and Fan Xiao. 2006. Attention Funnel: Omnidirectional 3D Cursor for Mobile Augmented Reality Platforms. In *Proc. of CHI '06*. ACM, New York, NY, USA, 1115–1122. DOI: http://dx.doi.org/10.1145/1124772.1124939

4. Frank Bolton, Shahram Jalaliniya, and Thomas Pederson. 2015. A Wrist-Worn Thermohaptic Device for Graceful Interruption. *Interaction Design & Architecture Journal* 26 (2015), 39–54.

5. Gabor et. al. 2014. Subliminal Cueing of Selection Behavior in a Virtual Environment. *Presence: Teleoperators and Virtual Environments* 23, 1 (2014), 33–50. DOI: http://dx.doi.org/10.1162/PRES_a_00167

6. Doris Hausen. 2014. Peripheral Interaction - Exploring the Design Space, PhD Thesis. (2014).

7. Jeff Johnson. 2014. *Designing with the Mind in Mind, Second Edition: Simple Guide to Understanding User Interface Design Guidelines* (2nd ed.). Morgan Kaufmann Publishers Inc., San Francisco, CA, USA.

8. G. Lakoff and M. Johnson. 1999. *Philosophy in the Flesh: The Embodied Mind and Its Challenge to Western Thought*. Basic Books. https://books.google.nl/books?id=KbqxnX3_uc0C

9. Päivi Majaranta and Andreas Bulling. 2014. Eye tracking and eye-based human–computer interaction. In *Advances in Physiological Computing*. Springer, 39–65.

10. Diako Mardanbegi, Dan Witzner Hansen, and Thomas Pederson. 2012. Eye-based head gestures. In *Proceedings of the symposium on eye tracking research and applications*. ACM, 139–146.

11. Petr Novak, O. Stepankova, M. Uller, L. Novakova, and P. Moc. 2009. Home and Environment Control. *COGAIN 2009 Proceedings. Lyngby: Technical University of Denmark* (2009), 35âĂŞ38.

12. Thomas Pederson, Lars-Erik Janlert, and Dipak Surie. 2011. A Situative Space Model for Mobile Mixed-Reality Computing. *IEEE pervasive computing* 10, 4 (2011), 73–83.

13. Jeffrey S Shell, Roel Vertegaal, and Alexander W Skaburskis. 2003. EyePliances: attention-seeking devices that respond to visual attention. In *CHI'03 extended abstracts on Human factors in computing systems*. ACM, 770–771.

14. John D Smith, Roel Vertegaal, and Changuk Sohn. 2005. ViewPointer: lightweight calibration-free eye tracking for ubiquitous handsfree deixis. In *Proceedings of the 18th annual ACM symposium on User interface software and technology*. ACM, 53–61.

15. Eduardo Velloso, Markus Wirth, Christian Weichel, Augusto Esteves, and Hans Gellersen. 2016. AmbiGaze: Direct Control of Ambient Devices by Gaze. In *Proceedings of the 2016 ACM Conference on DIS '16*. ACM, 812–817.

16. Mélodie Vidal, David H. Nguyen, and Kent Lyons. 2014. Looking at or Through?: Using Eye Tracking to Infer Attention Location for Wearable Transparent Displays. In *Proc. of ISWC '14*. ACM, New York, NY, USA, 87–90. DOI: http://dx.doi.org/10.1145/2634317.2634344

17. Siavash Zokai, Julien Esteve, Yakup Genc, and Nassir Navab. 2003. Multiview paraperspective projection model for diminished reality. In *Mixed and Augmented Reality, 2003. Proceedings. The Second IEEE and ACM International Symposium on*. IEEE, 217–226.