Augmenting Human Interaction Capabilities with Proximity, Natural Gestures, and Eye Gaze

Mihai Bâce
Department of Computer Science
ETH Zurich, Switzerland
mihai.bace@inf.ethz.ch

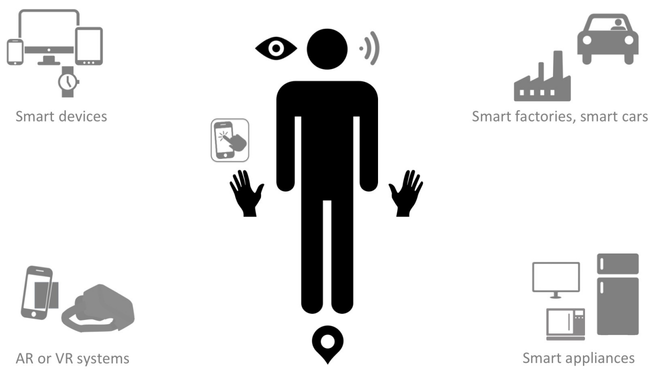

Figure 1: Existing interaction methods based on touch or voice control are not necessarily suitable for complex applications. We propose novel user interfaces that incorporate additional contextual cues like proximity, gestures, and eye gaze.

MobileHCI '17, September 04-07, 2017, Vienna, Austria
ACM 978-1-4503-5075-4/17/09.
https://doi.org/10.1145/3098279.3119924

Abstract
Nowadays, humans are surrounded by many complex computer systems. When people interact among each other, they use multiple modalities including voice, body posture, hand gestures, facial expressions, or eye gaze. Currently, computers can only understand a small subset of these modalities, but such cues can be captured by an increasing number of wearable devices. This research aims to improve traditional human-human and human-machine interaction by augmenting humans with wearable technology and developing novel user interfaces.

More specifically, (i) we investigate and develop systems that enable a group of people in close proximity to interact using in-air hand gestures and facilitate effortless information sharing. Additionally, we focus on (ii) eye gaze which can further enrich the interaction between humans and cyber-physical systems.

Author Keywords
Wearable technology; Augmented human; Eye gaze; Deep learning; Proximity detection; Collaboration; HCI; CPS.

ACM Classification Keywords
H.5.2 [Information interfaces and presentation (e.g., HCI)]: User Interfaces

Introduction

While getting ever smaller, cheaper, and more interconnected, computing technology comes even closer to humans, currently witnessed by wearable devices of various forms. What makes these devices particularly interesting is the insight to an egocentric perspective: smartwatches feel how we move, smartglasses see what we see, and smartphones know where we are. Besides having a personal point of view, these devices can enhance our senses much more: they can hear things we cannot hear, see things we cannot see, or have access to other sensors, and to practically endless knowledge via the Internet. We have witnessed a trend of smartphones becoming universal interaction devices where smart devices can ”outsource” their user interfaces, however, this form of interaction is limited in multiple respects. It typically requires a handheld device, it is mainly bound to spoken commands, and the physical context is rarely considered.

Research questions

1. How can we facilitate effortless information sharing between people who are physically close to one another?

2. How can we enrich our interaction capabilities with the gaze direction?

3. How can we predict the gaze direction on mobile devices using only commodity cameras?

Research Goals and Questions

We intend to broaden and facilitate the interface between complex systems and people by leveraging wearable devices, substituting today's use of smartphones by less explicit and less obtrusive wearable technology. We argue that advanced interactive technologies and systems based on cooperating wearables should also be able to understand additional contextual cues like physical proximity, interaction with our hands, and eye-gaze interaction (Figure 1). Given our research questions (Sidebar), we identify three work packages.

In the first work package (WP1), we focus on multi-user scenarios and collaboration when interacting with a system. While it is easy for humans to get contextual cues about who is working with whom, for machines this is nearly impossible at the moment. We leverage wearable devices like smartwatches to understand what users are doing with their hands and take into account physical proximity to enable effortless interaction between multiple people. Similar frameworks or toolkits have been proposed before, but they require additional infrastructure [7] or focus on smartphones and explicit interaction [6]. One way to further improve interaction and better understand users' intention is through eye gaze.

Eye-tracking and the gaze direction have a long history in research as a tool for understanding human behavior. Recent technological advances have made eye trackers more affordable which, in turn, has created interest to develop attention-aware systems [3]. In the second work package (WP2), we investigate how wearable eye trackers like the Pupil [4] can be used for interaction. Most wearable eye trackers lack any processing capabilities and need to be connected to a powerful computer, making such a system less portable. Connecting such an eye-tracker to a mobile device will greatly increase portability. Additionally, we would like to identify suitable applications which can benefit from wearable eye tracking (e.g., simplifying object detection when trying to recognize objects).

Wearable eye-trackers have come a long way, but requiring additional hardware hinders the wide deployment of gaze-aware systems. Advances in machine learning, more specifically deep learning, together with large public datasets have created a few methods that eliminate the need for dedicated eye tracking hardware [5]. In the third work package (WP3), we would like to explore how to further develop robust, real-time gaze prediction methods which rely on only the front facing camera found on standard mobile devices. Given such a method, we also intend to investigate what kind of applications can benefit from this technology and build prototypes.

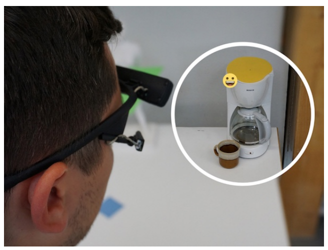

Figure 2: ubiGaze: an augmented reality system that allows embedding virtual messages into any object through gaze gestures [1]. Gaze gestures are captured using the Pupil wearable eye tracker.

Research Conducted and Expected Results

As part of WP1, we have developed a system that enables multiple users who are physically close to one another to effortlessly exchange information. HandshakAR [2], our first prototype, shows how two users who share the same greeting gesture and are close to each other can exchange contact information. We have further extended our system to allow multiple users, evaluated our system more rigorously, and developed two additional application scenarios. This work has not been published yet and is currently being prepared for submission.

Wearable eye-tracking (WP2) has the potential to enable new mobile interactive applications. We developed an augmented reality system which allows embedding virtual messages into any object (Figure 2) [1]. We are currently working on building an open-source platform which allows connecting the eye-tracker to a smartphone and doing all the computations locally. This coupling opens new challenges: implementing efficient pupil detection algorithms, exploring and choosing an appropriate calibration method. The expected outcome of this work is an open-source platform which others could use to develop mobile gaze-aware systems.

In WP3, we have started exploring models based on deep neural networks to predict the gaze direction. Most systems require as input an image of the face and the eyes. Our initial tests show that face and facial landmark detection are a bottleneck to real time systems. Given the GazeCapture dataset [5] we have developed an initial prototype which does not need an explicit face detection step. The outcome of this work will be a new algorithm that can predict the gaze direction in real-time on commodity mobile devices using only the front-facing camera.

REFERENCES

1. Mihai Bâce, Teemu Leppänen, David Gil de Gomez, and Argenis Ramirez Gomez. ubiGaze: Ubiquitous Augmented Reality Messaging Using Gaze Gestures. In *Proc. SIGGRAPH ASIA '16 MGIA*. 11:1–11:5.
2. Mihai Bâce, Gábor Sörös, Sander Staal, and Giorgio Corbellini. HandshakAR: Wearable Augmented Reality System for Effortless Information Sharing. In *Proc. Augmented Human '17*. 34:1–34:5.
3. Andreas Bulling and Kai Kunze. Eyewear Computers for Human-computer Interaction. In *Interactions '16*. 70–73.
4. Moritz Kassner, William Patera, and Andreas Bulling. Pupil: An Open Source Platform for Pervasive Eye Tracking and Mobile Gaze-based Interaction. In *Proc. UbiComp '14 Adjunct*. 1151–1160.
5. Kyle Krafka, Aditya Khosla, Petr Kellnhofer, Harini Kannan, Suchendra Bhandarkar, Wojciech Matusik, and Antonio Torralba. Eye Tracking for Everyone. In *Proc. CVPR '16*.
6. Andrés Lucero, Jussi Holopainen, and Tero Jokela. Pass-them-around: Collaborative Use of Mobile Phones for Photo Sharing. In *Proc CHI '11*. 1787–1796.
7. Nicolai Marquardt, Robert Diaz-Marino, Sebastian Boring, and Saul Greenberg. The Proximity Toolkit: Prototyping Proxemic Interactions in Ubiquitous Computing Ecologies. In *Proc. UIST '11*. 315–326.