Interacting with Computers 23 (2011) 484–498

Contents lists available at ScienceDirect

Interacting with Computers

journal homepage: www.elsevier.com/locate/intcom

# DOGeye: Controlling your home with eye interaction

Dario Bonino, Emiliano Castellina, Fulvio Corno, Luigi De Russis *

*Politecnico di Torino, Dipartimento di Automatica ed Informatica, Corso Duca degli Abruzzi 24, 10129 Torino, Italy*

ARTICLE INFO

*Article history:*
Received 29 October 2010
Received in revised form 15 April 2011
Accepted 15 June 2011
Available online 23 June 2011

*Keywords:*
Human–home interaction
Smart homes
Domotics
Usability
User interface
User study

ABSTRACT

Nowadays home automation, with its increased availability, reliability and with its ever reducing costs is gaining momentum and is starting to become a viable solution for enabling people with disabilities to autonomously interact with their homes and to better communicate with other people. However, especially for people with severe mobility impairments, there is still a lack of tools and interfaces for effective control and interaction with home automation systems, and general–purpose solutions are seldom applicable due to the complexity, asynchronicity, time dependent behavior, and safety concerns typical of the home environment. This paper focuses on user–environment interfaces based on the eye tracking technology, which often is the only viable interaction modality for users as such. We propose an eye-based interface tackling the specific requirements of smart environments, already outlined in a public Recommendation issued by the COGAIN European Network of Excellence. The proposed interface has been implemented as a software prototype based on the ETU universal driver, thus being potentially able to run on a variety of eye trackers, and it is compatible with a wide set of smart home technologies, handled by the Domotic OSGi Gateway. A first interface evaluation, with user testing sessions, has been carried and results show that the interface is quite effective and usable without discomfort by people with almost regular eye movement control.


## 1. Introduction

In the last 5 years, (smart) home automation gained a new momentum, thanks to an increased availability of commercial solutions (e.g., X10 or Z-Wave) and to steadly reducing costs. The evergreen appeal of automated, intelligent homes together with a raising technology maturity has fostered new research challenges and opportunities in the field of "intelligent" or "smart" environments. According to the Mark Weiser definition, a Smart Home system, that in this paper we decline as domotic or environmental control system,[1] is "a physical world that is richly and invisibly interwoven with sensors, actuators, displays and computational elements, embedded seamlessly in the everyday object of our lives, and connected through a continuous network" (Weiser, 1999), providing ways for controlling, interacting and monitoring the house. The idea behind this vision is that homes of tomorrow would be smart enough to control themselves, understand contexts in which they operate and perform suitable actions under inhabitants' supervision (Bierhoff et al., 2007). Although smart and autonomous homes might raise controversial opinions on how smart are they or should they be, currently available commercial solutions can start playing a relevant role as enabling technology for improving the care of the elderly (Berlo, 2002; Zhang et al., 2009) and of people with disabilities (Chikhaoui and Pigot, 2010; Chan et al., 2009), reducing their daily workload in the house, and enabling them to live more autonomously and with a better quality of life. Even if such systems are far from cutting-edge research solutions, they are still really complex to master since they handle and coordinate several devices and appliances with different functionalities and with different control granularities.

In particular, among other disabilities, people who have severely impaired motor abilities can take great advantages from eye tracking systems to control their homes, since they generally retain normal control of their eyes, that become therefore their preferential stream of interaction (Hornof and Cavender, 2005). Eye tracking can transforms such a limited ability into both a communication channel and an interaction medium, opening possibilities for computer-based communication and control solutions (Donegan et al., 2005). Even if eye tracking is often used for registering eye movements in usability studies, it can be successfully exploited as alternative input modality to control user interfaces. Home automation can then bridge the gap between software and tangible objects, enabling people with motor disabilities to effectively and physically engage with their surroundings (Andrich et al., 2006). Several house control interfaces have been proposed in the literature, i.e., applications to allows users to control differ-

* Corresponding author. Tel.: +39 011 5647191; fax: +39 011 5647099.
*E-mail addresses:* dario.bonino@polito.it (D. Bonino), emiliano.castellina@polito.it (E. Castellina), fulvio.corno@polito.it (F. Corno), luigi.derussis@polito.it (L. De Russis).

[1] We only consider systems currently available on the market such as X10, Konnex, MyHome, Z-Wave and ZigBee HA.

0953-5438/$ - see front matter © 2011 British Informatics Society Limited. Published by Elsevier B.V. All rights reserved.
doi:10.1016/j.intcom.2011.06.002

ent types of devices in their homes, to handle triggered alarms, etc. Such interfaces, either based on conventional unimodal (Koskela and Väänänen-Vainio-Mattila, 2004) or multimodal interactions (Weingarten et al., 2010) (e.g., mouse, remote controller, etc.), are too often uncomfortable and/or useless for people with severe impaired motor abilities, and only few of them have been specifically designed and developed to be controlled with eye movements.

In 2004, applications based on gaze interaction have been analyzed by a European Network of Excellence, named COGAIN (*Communication by Gaze Interaction*)[2], to evaluate the state-of-the-art and to identify potential weaknesses and future developments. According to the report "D2.4 A survey of Existing 'de facto' Standards and Systems of Environmental Control" (Bates et al., 2006), the COGAIN Network identified different problems in eye-based house control applications, such as the lack of advanced functionalities for controlling some appliances of the house, the absence of interoperability between different smart house systems or the difficulty to use an eye tracker for realizing some actions. In a subsequent report (Corno et al., 2007), COGAIN members proposed solutions to overcome the discovered problems. In particular, they proposed 21 guidelines to promote safety and accessibility in eye tracking based environmental control applications.

This paper describes the design and development of DOGeye, one of the first home control applications designed for gaze-based interaction and by explicitly accounting for the COGAIN guidelines. DOGeye is a multimodal eye-based application for home management and control, based on state-of-the-art technologies in both tracking and home control. It enables people to control their domotic homes through different input devices, possibly combined, so that it does not limit itself to eye tracking only. The presence of various input modalities allows application use by other people present in the house and offers different alternatives to the persons affected by possibly evolving impairments such as the ALS (Amyothrophic Lateral Sclerosis).

The remainder of the paper is organized as follows: Section 2 presents the basic features of eye tracking technology and the characteristics of eye-based user interfaces while Section 3 presents the work accomplished by the members of the COGAIN Network and describes COGAIN guidelines for eye tracking based environmental control applications. Section 4 reports relevant related works and findings. DOGeye design and architecture is described in Section 5, while Sections 6 and 7 report the setup and results of a user test involving people in a controlled environment, thus building the basis for further considerations and research. Section 8 concludes the paper and outlines future works.

## 2. Eye tracking basics

To better understand the principles and implementation of eye controlled interface, this section defines some terms and features pertaining to eye movements and eye tracking.

The eye does not generally move smoothly over the visual field; instead, it makes a series of quick jumps, called *saccades*, along with other specialized movements (Haber and Hershenson, 1973). A saccade lasts 30–120 ms, and typically covers 15–20 degrees of visual angle (Jacob, 1995). Between saccades, the *gaze-point*, i.e., the point in a scene where a person is looking, stays at the same location (with a slighty tremor) for a *fixation* that lasts from 100 to 400 ms; a longer fixation is called *dwell* (Hornof and Cavender, 2005).

Eye positions and their movement relative to the head can be measured by using different methods, e.g., computer vision techniques. One of these techniques is the so-called *Corneal Reflection* technique that consists in sending a small infrared beam toward the center of the pupil and estimating the changes in its reflexion (*eye tracking*). Eye tracking has several distinguishing features (Jacob, 1995):

- it is *faster* than other input media, as Ware and Mikaelian (Ware and Mikaelian, 1987) observed; before the user operates any mechanical pointing device, she usually looks at the destination to which she wishes to move;
- it is *easy* to operate, since no training or particular coordination is required to look at an object;
- it shows where the *focus of attention* of the user is located; an eye tracker input could be interpreted as an indication of what the user points at, but it can also be interpreted as an indication of what the user is currently paying attention to, without any explicit input action on her part;
- it suffers from *Midas Touch* problem: the user expects to be able to look at an item without having the look cause an action to occur. This problem is overcome by using techniques such as dwell time or blink selection;
- it is *always on*; in fact, there is no natural way to indicate when to engage the input device, as there is with grasping or releasing the mouse;
- it is *noninvasive*, since the observed point is found without physical contact;
- it reduces *fatigue*; if the user uses an eye tracker input instead of other manual pointing devices, movements of arms and hands will be reduced and will cause less fatigue;
- it is *less accurate* than other pointing devices, such as a mouse.

Because of these features, an eye tracking based interface has some specific peculiarities: for example, graphical widgets and objects are bigger than in traditional user interfaces, due to eye tracking lower accuracy; and the pointer is often absent, since its presence could divert users' attention (Donegan et al., 2006), replaced by other forms of visual feedback.

To overcome the "Midas Touch" problem, many interfaces use the *dwell time* technique. By using such a technique, the user can select a widget of a user interface only if she continues to look at it for a sufficiently long time. The amount of time is, generally, customizable by the user itself.

Moreover, interaction with eye-based interfaces can be improved by exploiting the *Selection–Action strategy* (SA), already used in the iAble application[3] and whose basic principle was proposed by Razzak et al. (2009). This strategy permits to separate the selection of an object from the activation of its associated actions. The *selection* is the process of choosing an object and displaying its related options, while *action* permits to perform some task on the selected object. The selection–action strategy is generally implemented by showing two separate areas to interact with: one is used only for selection, with a really short dwell time; the other is used for actions, with a longer dwell time, controllable by users. Two interaction patterns lie at the basis of SA: the *non-command based interaction*, used for selection, and the *command based interaction*, used for actions. In the *non-command based interaction* pattern, the computer observes and interprets user actions instead of waiting for explicit commands. By using this pattern, interactions become more natural and easier to use, as indicated by the work of Tanriverdi and Jacob (2000). In *command based interactions*, instead, the user explicitly directs the computer to perform some operations.

[2] http://www.cogain.org.

[3] A SRLabs commercial software – http://www.srlabs.it/en/iable.html.