Chapter 13
Gaze–Aware Systems and Attentive Applications

Howell Istance
De Montfort University, UK

Aulikki Hyrskykari
University of Tampere, Finland

ABSTRACT

In this chapter, we examine systems that use the current focus of a person’s visual attention to make the system easier to use, less effortful and, hopefully, more efficient. If the system can work out which object the person is interested in, or is likely to interact with next, then the need for the person to deliberately point at, or otherwise identify that object to the system can be removed. This approach can be applied to interaction with real-world objects and people as well as to objects presented on a display close to the system user. We examine just what we can infer about a person’s focus of visual attention, and their intention to do something from studying their eye movements, and what, if anything, the system should do about it. A detailed example of an attentive system is presented where the system estimates the difficulty a reader has understanding individual words when reading in a foreign language, and displays a translation automatically if it thinks it is needed.

INTRODUCTION

In this chapter, we will examine how knowledge of what a person is currently interested in or attending to can be used to make computer-based systems easier to deal with, less effortful, and – it is hoped – more efficient to use. As we don’t want a person to keep telling the system what the object of current interest is, we will try to infer this by monitoring the users as they use the system. There are many clues we can use, such as contextual information on what they are doing, what they have just done, or what they usually do at this time of day. However, as we saw in the first chapter in this section (Chapter 11 by Mulvey and Heubner), the most reliable way of getting
information about the current focus of someone’s attention is by monitoring gaze behaviour. This can be supplemented by additional information, with brain activity – the topic of the second chapter of this section (Chapter 12 by Vidaurre, Kübler, Tangermann, Müller, and Millán) – being one of the future possibilities.

We will use the term ‘attentive system’ to describe a computer system that changes state on the basis of inferences about what the user of the system is currently attending to and, in some cases, intending to do as a consequence. An attentive system uses real-time measurement of gaze position and possibly of other physiological indicators to do this. This doesn’t mean that the system second-guesses what the user will do and takes some action on the user’s behalf automatically, although in certain circumstances this might be appropriate. It may simply mean that the action the user is most likely to want to perform next is offered as a default command.

We should consider other similar terms used to describe computer systems to clarify where the similarities and the differences lie. The term ‘adaptive systems’ refers to systems that make use of knowledge of the tasks a person is performing in order to modify or adapt the current state of the system. To do this, these often make reference to a model of individual user preferences, and of similar tasks undertaken previously. Knowledge of the tasks is obtained from a log of commands the user has entered into the systems at different times. Affective systems typically try to take account of a user’s current emotional (or affective) state. This state can be inferred from a variety of input modalities or sources, including facial expression, voice, heart rate, and pressure on a keypad during its use. This could be used to make feedback from the system appear to be sensitive or responsive to the current state of the person using it.

The term ‘ubiquitous computing’ refers to computation embedded in basic objects, environments, and activities of our everyday lives in such a way that no-one will notice its presence (Weiser, 1999). This includes the idea of tangible user interfaces, where input and output devices are integrated with real-world objects (rather than the familiar dedicated devices) that someone can pick up, squeeze, shake, move around, and so on. These objects are also candidates to be made gaze-aware and attentive.

The term ‘context-aware systems’ is a broad term that generally incorporates all of the above. It appears in different fields of computer systems research and development, not least in mobile technologies, which we will further examine below.

Importantly, in all of the above-mentioned types of system, there are two distinct phases: 1) determining what the current state of the person using the system is and 2) deciding what, if anything, to do with that information. The boundaries between the systems described by the terms used above are blurred, and these often overlap.

The focus of this chapter will be on systems that make active use of a person’s gaze position and eye movements.

**Context-Awareness and the Importance of Context**

Mobile devices (including phones) are equipped with the capability of continually sending information back to central computers. Importantly, this includes information about where the person is in relation to places and other people. This information allows different types of context to be defined and built for a person’s current activity. Mobile devices are often equipped with accelerometers, which can detect motion. Integrating this information with, say, a person’s online calendar can enable inferences to be drawn about the person’s current activity. If, for example, the position of the device is changing rapidly, as indicated by the GPS sensor in the device, but the device is not bouncing up and down, it is likely that, if the time is before 9am and it is a work day, the person is driving or being driven, say, to work. Mobile