Chapter 4
Design and Evaluation of Vision–Based Head and Face Tracking Interfaces for Assistive Input

Chamin Morikawa
Motion Portrait Inc., Japan

Michael J. Lyons
Ritsumeikan University, Japan

ABSTRACT

Interaction methods based on computer-vision hold the potential to become the next powerful technology to support breakthroughs in the field of human-computer interaction. Non-invasive vision-based techniques permit unconventional interaction methods to be considered, including use of movements of the face and head for intentional gestural control of computer systems. Facial gesture interfaces open new possibilities for assistive input technologies. This chapter gives an overview of research aimed at developing vision-based head and face-tracking interfaces. This work has important implications for future assistive input devices. To illustrate this concretely the authors describe work from their own research in which they developed two vision-based facial feature tracking algorithms for human computer interaction and assistive input. Evaluation forms a critical component of this research and the authors provide examples of new quantitative evaluation tasks as well as the use of model real-world applications for the qualitative evaluation of new interaction styles.

INTRODUCTION

The past decade has witnessed dramatic change in the field of Human Computer Interaction (HCI). Interface technologies that were previously found only in the laboratory have started to show up in consumer devices. Tablet computers and smart phones allow smooth and nearly faultless tactile interaction that computer interface researchers of the past could only dream about. The use of digital devices is moving

beyond button, keyboard, and mouse, and towards more intuitive interaction styles better suited to the human brain and body.

Interaction methods based on computer-vision hold the potential to become the next powerful technology to support breakthroughs in HCI (Jaimes, 2007; Porta, 2002). Computer-vision algorithms can obtain real-time knowledge about the actions of a human user for use in HCI (Ahad 2008). Most research on vision-based HCI has naturally focused on movements of the arms and hands (Jaimes, 2007; Porta, 2002). However, it is also interesting to consider other parts of the body such as the head and features of the face, because these afford detailed and expressive movements (Lyons, 2004).

Action of the face plays an important role in many human behaviors including speech and facial expression. It is therefore not unreasonable to think that the actions of the face could play an important role in man-machine interactions. While there is a considerable body of prior research on automatic facial expression recognition and lip reading, there has been relatively little work examining the possibility of using these for intentional interactions with computers or other machines or, notably, for assistive input. This may be partly due to technological limitations: how can information about motor actions of the face be acquired in an unencumbering, non-invasive fashion? As we will show in this chapter, this is no longer a consideration: robust, real-time acquisition of facial movements makes only modest technological demands. The strangeness and novelty of the idea of using the face or features of the face for intentional interaction may be another factor in the relative dearth of precedent studies, however novelty should not be a deterrent to research. Furthermore we will discuss several of our recent applications which focus on using mouth movements for HCI, in which our studies with users show the concept to be quite natural and advantageous.

Adoption of new interface paradigms depends not only on the invention and development of novel technology, but on how the technology is used to create engaging interaction styles. Therefore we have concretely examined what kinds of interaction styles may be suitable for vision-based interfaces with systematic evaluation tasks and also in the context of potential applications. In the current chapter we apply standard human factors evaluation methodology that has evolved in the HCI field, to characterize a facial gesture UI developed in our group, which allows the user to provide input to a computer using movements of the head and mouth. To demonstrate the value of the standard evaluation methodology we show how it may be used to accurately predict performance with real-world applications.

An outline of the chapter is as follows: Section 2 reviews the state of the art in the relevant areas of human-computer interaction, computer vision, and more specifically in the specialized area of vision-based human-computer interaction; Section 3 concretely illustrates the concept of vision-based facial gesture interaction by describing in some detail algorithms we have studied in our own research projects. Specifically these systems (a) initialize the face tracking by blink detection and nostril detection (b) track the tip of the nose for pointing (c) detect mouth opening to allow the user to input click events and continuous data using the size and shape of the mouth cavity; Section 4 concretely illustrates how new vision-based interaction techniques can be evaluated with standard and newly developed methods from the field of human-computer interaction. After briefly introducing Fitts' law and standard methodology for measuring the information throughput of a pointing device, we then describe a measurement of the throughput of a user-interface, comparing it to that of a standard computer mouse; in some cases newly designed tasks are needed: this is concretely illustrated with a measurement of the accuracy and precision of input derived from the shape of the mouth; Section 5 illustrates the application of qualitative measures for usability assessment using user questionnaires; Section 6 contains a brief overview of examples of applications for vision-based facial gesture interfaces; after general observations, in Section 7, Section

Related Content

Aggregate Searchable Encryption With Result Privacy
Dhruti P. Sharmaand Devesh C. Jinwala (2021). *Research Anthology on Artificial Intelligence Applications in Security (pp. 922-945).*
www.irma-international.org/chapter/aggregate-searchable-encryption-with-result-privacy/270633

Eliciting User Preferences in Multi-Agent Meeting Scheduling Problem
Mohammad Amin Rigiand Farid Khoshalhan (2011). *International Journal of Intelligent Information Technologies (pp. 45-62).*
www.irma-international.org/article/eliciting-user-preferences-multi-agent/54066

Automating the Generation of User Activity Timelines on Microsoft Vista and Windows 7 Operating Systems
Stephen O'Shaughnessyand Anthony Keane (2012). *International Journal of Ambient Computing and Intelligence (pp. 35-47).*
www.irma-international.org/article/automating-generation-user-activity-timelines/66858

Synthetic Neuron Implementations
Snorre Aunet (2009). *Encyclopedia of Artificial Intelligence (pp. 1555-1561).*
www.irma-international.org/chapter/synthetic-neuron-implementations/10445

On the Similarity Search With Hamming Space Sketches
Vladimir Micand Pavel Zezula (2021). *Intelligent Analytics With Advanced Multi-Industry Applications (pp. 97-127).*
www.irma-international.org/chapter/on-the-similarity-search-with-hamming-space-sketches/272781