Technology in Society 43 (2015) 33–49

Contents lists available at ScienceDirect

Technology in Society

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

Understanding the timing of economic feasibility: The case of input interfaces for human-computer interaction

Karthik Nandakumar [b], Jeffrey L. Funk [a, *]

[a] Division of Engineering and Technology Management, National University of Singapore, Faculty of Engineering, 9 Engineering Drive 1, Singapore 117576
[b] Institute for Infocomm Research, A*STAR, Fusionopolis, Singapore 138632

ARTICLE INFO

Article history:
Received 5 March 2013
Received in revised form
5 October 2015
Accepted 9 October 2015
Available online 13 November 2015

Keywords:
Economic feasibility
Technological feasibility
Systems
Components
Technological trajectories
Technological trends
Human–computer interaction
Speech interfaces
Gesture interfaces
Direct neural interfaces

ABSTRACT

The objective of this paper is to understand when new types of input interfaces for Human–Computer Interaction (HCI) such as Natural User Interfaces (NUI) (e.g., speech and gesture) and Direct Neural Interfaces (DNI), or combinations of them, might become technologically and economically feasible. This problem is addressed by analyzing the performance trajectories of key components in these HCI systems. In the case of speech interfaces, we observe that microphones and automated speech recognition systems are no longer experiencing rapid improvements along key dimensions of performance, which inhibits their technical and economic feasibility. On the other hand, 2D image sensors and depth sensors, which constitute the core components of gesture interfaces, are continuing to improve at a significant rate in terms of characteristics like spatial resolution, pixel sensitivity, and depth resolution. When coupled with the exponential improvements in the memory and processing power of computing systems, the above improvements in image sensors are enabling gesture-based natural user interfaces to reach acceptable levels of technical performance and economic feasibility. Similarly, simultaneous improvement in the spatial and temporal resolution of non-invasive brain scanning technologies is likely to accelerate the development of direct neural interfaces (DNI). However, a number of challenging obstacles such as lack of robust magnetic shielding systems, high cost, and poor usability continue to hinder the economic feasibility of DNI systems.

1. Introduction

A new technology can be considered as technologically feasible and economically viable when it attains sufficient levels of performance and cost so as to enable widespread diffusion. Understanding when a new technology might become economically viable and begin to diffuse remains an elusive goal. The technology forecasting literature offers a number of general techniques (see Ref. [1–2] for a detailed review) such as scanning and technology monitoring [3], technology roadmapping [4–5], trend analysis [6–7], simulation [8–9], and expert opinion [10]. With the exception of expert opinion, which has its own set of problems [10], the major challenge in applying these techniques is to find good data and little data is available for a new technology before it has begun to diffuse.

The economics literature focuses on cumulative production as a key driver of trends in cost and performance. According to this theory, the cost of a new technology (also known as a “technology discontinuity” [11–13]) falls as cumulative production increases, following a so-called learning or experience curve [14–16]. However, if cost reductions primarily come from cumulative production as suggested by the learning curve, by definition, cost reductions cannot occur before production occurs. This makes it very difficult to use a learning curve to analyze when a new technology might become economically viable and hence, begin to diffuse.

One approach is to use data pre-commercialization rates of improvement, sometimes called trajectories [17–20], to assess when new technologies might become economically feasible. One can compare the performance and cost of a new technology with the needs in the marketplace and with the performance and cost of existing technologies, and then use the rates of improvement to estimate when a new technology might become economically feasible and thus begin to diffuse. For example, one study [21] analyzed the rates of improvements for a number of new technologies [22–23] in order to determine the types of design changes

---

* Corresponding author.
E-mail addresses: karthiknandakumar@gmail.com (K. Nandakumar), jeffreyleefunk@gmail.com (J.L. Funk).

http://dx.doi.org/10.1016/j.techsoc.2015.10.001
0160-791X/

that led to the improvements. The study found that new materials and reductions in scale enabled many of the improvements in cost and performance [21]. Other studies [24–26] have shown that the efficiency of some mathematical/computing algorithms has been experiencing significant improvements.

A different type of approach focuses on systems [27–29] and the improvements in systems that come from rapid improvements in components. This is because the performance and cost of some systems are driven more by improvements in components than by changes in system design [30–31]. For example, some argue that integrated circuits (ICs) have had a much larger impact on the performance and cost of computers than have changes in the design of computers [32]. Some also argue that these improvements in ICs have enabled the emergence of smaller computers such as personal computers, laptops, and tablet computers [30–33]. One can take this a step further and define ICs as a "general purpose technology" [34–35] whose progress has had a large impact on the performance and cost of many higher-order electronic systems and not just those of computers [36]. Thus, one can use Moore's Law to analyze when new electronic products might become economically viable. Similar arguments have been made for improvements in the recording density of magnetic tape and the emergence of new types of magnetic-tape based systems [37].

This paper uses these two approach, and in particular the second approach, to analyze new input interfaces for Human–Computer Interaction (HCI) and discuss when they might become technologically and economically feasible. It focuses on Natural User Interfaces (NUI) (e.g., speech and gesture) and Direct Neural Interfaces (DNI), or combinations of them for two reasons. Firstly, NUI and DNI are new input interface technologies that are at various stages of technological development and these technologies are likely to have a significant impact on how computers will evolve. For example, one can argue that performance improvement in touch interfaces played a key role in making smartphones and tablet computers economically viable. Secondly, one can identify a number of technological trajectories within NUI and DNI and the rates of progress of these trajectories depend on a wide variety of factors such as ICs, materials, and software algorithms.

The rest of this paper is organized as follows. Section 2 introduces the field of HCI and the key components in these systems, specifically in relation to input interfaces for HCI. Section 3 outlines our research methodology. The subsequent sections present a detailed analysis of three different input interface technologies, namely, speech (Section 4), gesture (Section 5), and direct neural interfaces (Section 6). This is followed by a discussion in Section 7 about the common patterns underlying these three analyses.

2. Key concepts

2.1. Human computer input interfaces

Human–computer interaction (HCI) refers to the technology that connects humans and machines. In other words, it is the system of hardware and software components that allows a human to make inputs and receive results from the computer [38]. The working of a typical HCI system is shown in Fig. 1. Generally, when a person wishes to interact with a computer, he must translate his thoughts into specific actions (e.g., type or touch a key, move the mouse, speak a command, etc.). These actions are delivered to the machine through an input device (keyboard, mouse, camera, touch sensor, etc.) [39]. The computer recognizes the action (or set of actions) performed by the user, decodes these actions into a task, and executes the task. The results of the executed task are rendered in an appropriate format (e.g., text, images/videos, sound, etc.) and conveyed to the user via the output device (e.g., display, speaker, etc.). The user is able to perceive the results through his sensory organs (e.g., eyes, ears, etc.). In this work, we focus only the *input interface*, which includes both the hardware device used to acquire the input from the user and the software component that performs action recognition.

2.2. New technologies for HCI

In the case of input interfaces for HCI, five major technologies and a number of trajectories can be identified. The fundamental difference between the input interface technologies is the nature of actions required from the user to interact with the computer. Historically, batch interfaces, command line interfaces (CLI), and graphical user interfaces (GUI)[1] have been used for human–computer interaction. Some of the devices used in these traditional input interface technologies are shown in Fig. 2. More recently, new technologies such as natural user interfaces (NUI) and direct neural interfaces (DNI) have been introduced. Natural user interfaces are based on natural interactions that people use to communicate among themselves. Examples of such natural interactions are speech, touch, and gestures. Consequently, NUI can be further categorized into speech, touch, and gesture interfaces. Finally, direct neural interfaces take the concept of human–computer interaction to the extreme and attempt to directly decode human thoughts, without requiring the user to perform any explicit physical action. Fig. 3 presents some illustrations of natural and direct neural interfaces.

In addition to the five major input interface technologies discussed above, other input interfaces based on eye gaze, facial expressions, head and body movements, and hand pressure are also possible. However, these additional input interfaces are more likely to be used only in specific niche applications (e.g., to enhance visual displays or robotic control). Furthermore, many applications demand multimodal interfaces, where more than one basic input interface technologies will be used. Examples of such applications include smartphones and gaming consoles, where a combination of different NUI technologies (e.g., touch and speech) is typically used. In the multimodal scenario, the success of the input interface will depend on how the different technologies are seamlessly fused in to form a well-designed input interface system that can quickly switch between different technologies depending on the context.

The following common criteria can be used to assess the different input interface technologies. An ideal input interface should perform well along all these dimensions.

- *Accuracy* refers to the precision in recognizing the actions made by the user.
- *Throughput* is the amount of information that can be input to the computer per unit time.
- *Affordability* is inversely proportional to the cost of the input interface.
- *Sociability* refers to ability of the input interface to allow multiple persons to interact simultaneously with the computer.
- *Mobility* or portability depends on the size and mass of the system and factors like power consumption.

---

[1] Note that graphical user interfaces (GUI) involve both input and output components. The typical input devices for GUI are keyboard, mouse, joystick, etc. The primary output module of a GUI is a 2-dimensional display that contains graphic elements (e.g., window, icon, menu, pointer, etc.). Since the focus of this work is only on the input interface, the term GUI refers only to the input devices like keyboard and mouse. Thus, modern tablet computers such as iPad use touch interfaces (a form of NUI) for input and not GUI.