Chapter III
Exploring Multimodal Input Fusion Strategies

Arianna D’Ulizia
Istituto di Ricerche sulla Popolazione e le Politiche Sociali
Consiglio Nazionale delle Ricerche, Italy

ABSTRACT

Human-computer interaction is a discipline that aims at enabling the interaction between humans and computational machines. In the last years, several efforts have been made to make this interaction more intuitive and natural. In this direction, multimodal interaction has emerged as the future paradigm of human-computer interaction. In order to enable a natural dialogue between users and computer systems, in multimodal systems the two main challenges to face are: to combine and integrate information from different input modalities (fusion process) and to generate appropriate output information (fission process). Our specific concern in this chapter is with the fusion of multiple input modalities.

INTRODUCTION

In recent years, people are increasingly surrounded by objects in the everyday environment that are equipped with embedded software and wireless communication facilities and which they need to interact with. For instance, mobile phones, PDAs and portable PCs are used by an increasing amount of people to develop everyday activities. This phenomenon produces the need to simplify the access to these technological devices by rendering human-computer interaction more similar to human-human communication. As a consequence, the “universal accessibility” concept is acquiring an important role in the research area of human-computer interaction (HCI). This discipline deals with the design, development and implementation of computational systems devoted to enable the interaction with human beings. Three of the main emerging research directions of the HCI, in line
with the universal accessibility concept, are: (i) to make this interaction more intuitive, natural and efficient by integrating multiple input-output modalities, (ii) to enable a broader spectrum of users, with different ages and skill levels as well as users with disabilities, to access technological devices, and (iii) to increase the level of freedom offered to users. In particular, multimodal interaction, which refers to the simultaneous or alternative use of several modalities, has emerged as the future paradigm of human-computer interaction that advances the implementation of universal accessibility.

This enhancement is also demonstrated by the studies of Oviatt et al. (Oviatt et al., 1997; Oviatt and Cohen, 2000; Oviatt et al., 2000) that emphasize the benefits of multimodal systems in terms of usability, accessibility, flexibility and efficiency compared to unimodal ones. In particular, a multimodal interaction improves usability, as it provides users with the means to choose among different available modalities. Secondly, multimodality improves accessibility to the device by encompassing a broader spectrum of users, enabling those of different ages and skill levels as well as users with disabilities. Finally, it offers improved flexibility and interaction efficiency.

Several aspects characterize multimodal interaction compared to usual interaction through graphical user interfaces (GUIs). Firstly, a GUI requires atomic and unambiguous inputs (such as the selection of an element by mouse or the insertion of a character by keyboard), whereas a multimodal interaction involves several simultaneous inputs that have to be recognized and opportunely combined by managing the uncertainty of inputs through probabilistic techniques. The process of integrating information from various input modalities and combining them into a complete command is called multimodal fusion. Secondly, in a multimodal interaction temporal constraints of inputs have to be taken into account and consequently it requires a time-sensitive architecture and the recording of time intervals of each modality. Finally, in a GUI the output messages are conveyed to the user through a single medium (the graphical display), whereas in a multimodal system a way of disaggregating outputs through the various channels has to be found in order to provide the user with consistent feedback. This process is called multimodal fission, in contrast with multimodal fusion.

Consequently, in the design and development of a multimodal system the two main challenges to face concern the multimodal fusion and fission processes. Our specific concern in this chapter is with the fusion of multiple input modalities. In particular, we intend to give a comprehensive analysis of current fusion approaches, considering two main classifications, namely according to the data fusion level (e.g. the fusion process takes places in the dialogue management system, as well as at grammar level) and the mathematical method (e.g. based on statistical or artificial intelligence techniques).

The remainder of the chapter is organized as follows. After a brief introduction about basic terms and concepts of human-computer interaction theory, we investigate current multimodal fusion mechanisms, classifying them according to the data fusion level and the applied mathematical method. Final conclusions are given at the end of the chapter.

**BACKGROUND**

In this section we review some basic terms and concepts of human-computer interaction theory.

The success of the human-computer communication depends on the reaching of a common ground by exchanging information through the communication modalities. Such a communication modality refers to the medium or channel of communication that conveys information (Coutaz and Caelen, 1991). Multimodality refers to the quality of a system to allow more than one communication modality to be used during human-computer interaction.
Related Content

Predicting Psychological Characteristics by Smartphone Usage Behaviors
www.igi-global.com/chapter/predicting-psychological-characteristics-by-smartphone-usage-behaviors/130161?camid=4v1a

Technology Cultures
www.igi-global.com/chapter/technology-cultures/45039?camid=4v1a

Social Responses to Conversational TV VUI: Apology and Voice
www.igi-global.com/article/social-responses-to-conversational-tv-vui/121635?camid=4v1a

Model Analysis and Development Strategy on Building an Industrial Research and Development (R&D) Centre: Shanghai’s Practice and Inspiration
www.igi-global.com/chapter/model-analysis-development-strategy-building/58710?camid=4v1a