Chapter 3

Improving Interaction with TV-Based Applications through Adaptive Multimodal Fission

David Costa
University of Lisbon, Portugal

Carlos Duarte
University of Lisbon, Portugal

ABSTRACT

In this chapter, the authors describe the design and implementation of an adaptive multimodal fission component integrated in the multimodal GUIDE framework. This component is able to adapt any HTML-based application’s UI to a specific user’s characteristics, making it possible for elderly and impaired users to interact by offering several output modalities that try to overcome possible interaction difficulties. They also present an evaluation of the proposed solution, conducted with more than 50 participants, that shows the efficiency of multimodal adaptation in increasing task perception and task execution.

INTRODUCTION

Nowadays, interaction with TV is no longer a passive viewing experience, where all the viewer does is choose what channel to watch or change volume settings. Today, with services like Google TV, Samsung smart TV or even Apple TV, viewers have active roles consuming and producing content like browsing the web, uploading media such as images or videos, using different TV applications available or simply recording TV programs.

Since everyone can and should be able to access these new technologies regardless of their age, knowledge and physical, sensorial or cognitive abilities it makes sense to employ adaptation techniques and offer a wide range of possible interaction devices for both input and output feedback.

Given that it cannot be expected that software developers quickly become experts in designing and developing multimodal adaptive systems, the way to bring the benefits of adaptation and multimodality to end-users is to provide an easy way for application developers to include multimodal adaptation in their applications, or to integrate them into an existing adaptive multimodal framework.

Keeping in mind the e-inclusivity of all possible users of smart TVs and the feasibility for developers to apply those features, a framework was
envisioned with particular focus on elderly people and their specific impairments and preferences.

This chapter focuses on our approach to handle output adaptations based on users’ abilities and characteristics. To do so, we demonstrate our implementation of a fission engine and its applicability on a specific project, a multimodal adaptive framework named GUIDE. This framework is a layer between the end-user and the application execution platform, which is capable of endowing applications with multimodal and adaptive features. The fission component uses an abstract description of the user interface (UI) to perceive what elements are displayed in the screen and adapts the UI to best suit the user and interaction context.

Although the approach designed can be applied to different execution environments, the GUIDE project focuses on TV as a channel to increase accessibility to Web application for older adults. As such, in this chapter, we will focus on the current implementation, that targets Web applications running on TVs. Consequently, the examples provided in the chapter will reflect this context.

In the following sections we describe what adaptive multimodal systems are and present earlier research work and projects on this subject. Next, we describe our results from earlier user studies to design and implement this user centered framework. In the main section we do an in depth description of the adaptive multimodal fission engine, by explaining the engine’s process flow, the approaches and techniques used and its main sub-components. Later on, we present the conclusions of study made to evaluate the usability and adaptive capacities of the framework with more than 50 users from 3 different countries.

**ADAPTIVE MULTIMODAL SYSTEMS**

Dumas, Lalanne, and Oviatt (2009) define multimodal systems as “computer systems endowed with multimodal capabilities for human-computer interaction and able to interpret information from various sensory and communication channels”. These systems offer users a set modalities to allow them to interact with machines and “are expected to be easier to learn and use, and are preferred by users for many applications” (Oviatt, 2003).

Adaptive multimodal systems enable a more effective interaction by adapting to different situations and different users according to their skills, physical or cognitive abilities.

The flexibility of these systems allows them to adapt not only to users but also to the environment (context awareness). For example, the system can use speech to warn or present information in an eyes-busy situation. Combining adaptation capabilities with multimodal interaction can result in an interface even more flexible and natural to the user. In the past, several works have tried to explore this combination as a way to improve interaction. Most of these have targeted adaptation within specific modalities, instead of adaptation across modalities (Duarte & Carriço, 2006).

Past attempts at designing adaptive systems have shown their potential and usefulness though they fail or are hard to be translatable to different types of applications or application execution environments. Examples of earlier projects include ELM-ART II (Weber & Specht, 1997) and I-Mailer (Kabassi & Virvou, 2003).

Other authors tried to address this issue and developed frameworks that are indeed able to perform adaptations, independently of the application. Examples of these systems are SmartKom (Reithinger et al., 2003) and SUPPLE (Gajos, Weld, & Wobbrock, 2008). The SUPPLE project considers the generation of personalized user interfaces as an optimization issue. It automatically generates a personalized user interface based on four types of input information: 1) Functional specification of the interface; 2) Device-specific constrains; 3) Usage trace and 4) Cost function. SUPPLE then generates the user interface that is the most efficient while satisfying all the constraints.