The Contribution of Multimodal Adaptation Techniques to the GUIDE Interface

José Coelho and Carlos Duarte

LaSIGE, Faculty of Sciences, University of Lisbon, Lisbon, Portugal,
jcoelho@lasige.di.fc.ul.pt, cad@di.fc.ul.pt

Abstract. This paper describes the European Union funded project GUIDE, focusing on the issues of multimodal and context adaptation techniques, as well as in the importance of having a multimodal system architecture based in user models and integrated with fusion and fission mechanisms in order to give elderly and impaired users several input and output modalities in their interaction with TV and set-top box based technology. The possible future role of GUIDE in the development of accessible applications is also focused.

Keywords: Multimodal adaptation, context adaptation, user modeling, GUIDE.

1 Introduction

The old days are gone, and TV interaction is now more complicated with complex TV remote controllers and too many buttons or complex interaction just to perform regular actions, like changing channels, record a TV show, or simply turn off the TV. More, different users have different ways to interact, as each one has a set of unique capabilities and preferences. However, disabled and elderly users have a much reduced set of capacities when compared with typical users, being limited to the ones they understand or can perform. Motor, physical or mental impairments make simple interactions more complicated to these users, so the platform they use to interact with a TV must be prepared to understand their unique range of interaction modalities, and potentiate their preferences in order to give them the most comfortable and effective interaction experience.

This paper presents the GUIDE[1] project, a new TV and set-top box based system which employs adaptation techniques, in order to provide accessibility and well-suited interaction to users with different impairments and skills, as well as being capable to reflect their preferences and the different contexts of use.

2 The GUIDE Project

GUIDE (Gentle User Interfaces for Elderly People), is an EU funded project focused on accessibility, that offers to elderly and impaired users a TV and set-top box platform, giving them a more intuitive way to interact with this type of technology, making use of gestures, voice commands, or just pointing to the screen, in order to

C. Stephanidis (Ed.): Universal Access in HCI, Part I, HCII 2011, LNCS 6765, pp. 337–346, 2011.

perform various kinds of actions without ever having to touch or understand the complexity of the traditional controllers. With GUIDE, interactions can be seen as a set of intuitive combination of actions, mixing all types of input modalities a user can and likes to perform.

GUIDE interaction is based on simple and intuitive actions that require wispy learning from the users (it's the only way it will be adopted by them), and a development process based in an extensive user requirement phase, focused on fully understanding the different impairments and what is needed for each type of disabled user to easily interact with the system. To accomplish his objectives, any user interacting with GUIDE can use a range of devices: TV, remote control, speech synthesis and recognition, Microsoft Kinnect, Nintendo Wii remote and Tablet PC. In addition to this technology, and to raise user's empathy, Avatar personas (face, body and voice) are also used to impersonate the system. With all of this input and output possibilities, GUIDE interfaces can communicate with the user's trough visual, audio and haptic feedback.

In what is related to scenarios of use, originally GUIDE can be used to perform regular TV operations, as well as performing video-conference, home automation, accessing media and social interaction and performing telelearning.

As it was already referred, GUIDE also has adaptation mechanisms capable of "translating" an interface into several types of presentation, each one of them focusing the different types of users the system aims (for example, for a visual impaired user a interface with a menu would have its content presented using only speech and sound, or having bigger buttons and larger fonts, or a combination of both). Additionally, thanks to these mechanisms, and following other of its objectives, GUIDE will offer to developers, a toolbox for accessible development of applications which performs a kind of automatic "translation" of "generic" interfaces to interfaces tailored for specifically impaired users.

3 GUIDE User Model: First Step to Adaptation

In general, humans differ in background, sex, education, personality, cognitive skills, preferences, motivation, goals and mood [2]. In addition, we know from literature [3,4] that elderly and handicapped users differ from the so called "regular" users, in what concerns aims, interests, experience and abilities (for example: elderly users prefer historic and cultural information and content; For vision-impaired users the screen display should be enlarged or modality presented must be changed to tactile or audio output; etc). Considering this many differences between users and the necessity of having their many representations in the system so GUIDE can adapt its interfaces to the complexity of each user, the first step of adaptation is having a User Model ("a representation of the knowledge and preferences of users that the system believes the user posses" [6]) capable of clustering groups of users with different characteristics.

For making this possible, GUIDE User Model [8] is based on the Model Human Processor [5] model for reasons of simplicity, and it is characterized by having 1) a user profile section which stores application-relevant characteristics about the user:

- Physical and Cognitive Abilities
- Domain knowledge
- Environmental data (software, hardware, user location, etc)
- Competence in handling with GUIDE system

2) A section responsible for increasing overall simulation accuracy by incorporating the following detailed action models:

- Perception Model: Simulates the phenomenon of visual perception as is capable of simulating the effects of different visual impairments.
- Cognitive Model: Simulates performance from novice to expert.
- Motor Behavior Model: Developed from cursor analysis of cursor trace from motor impaired users and can be used to predict paths and pointing time.

3) An inference machine section called simulator [7], which uses user information, takes a description of an interaction with the interface (Interface Model) as well as a task definition (Application Model) and predicts possible interaction and task completion times for each type of user in that given interface.

So, GUIDE User model is based on the empirical data described above, and collected from a wide range of users with different abilities (and trough a process of questionnaires, focus groups and user trials with multimodal prototypes). The User model will cluster these users to different user profiles, and the simulator will predict interaction patterns for each of those profiles, making static and dynamic adaptation possible (described in the next section of this paper). Additionally, since this User Model does not store data about user preferences, GUIDE also has a complementary Preference Model which stores this type of information and makes personalization on interfaces also possible within every cluster of users.

4 GUIDE Architecture: Adaptive/Adaptable System Behavior

4.1 GUIDE Architecture

The User model is a crucial component of any adaptive multimodal system, but it is only a small part of it (by its own, it can't do anything). GUIDE, as one of this type of systems, has a complex architecture based on recognizers and sensors to perceive the user and the environment, as well as an integration module which is capable of receiving all the information from the input modalities, make sense of it and select an adapted response making use of the output modalities and always considering the user characteristics and preferences as well as the different interaction contexts:

Figure 1 shows how input modalities are first perceived though recognizers, outputting their results to the fusion engine [19], which is in charge of making a common interpretation of the inputs and communicating it to the dialog manager [20]. Here, the dialog state is identified, as well as the transition to perform, the action to

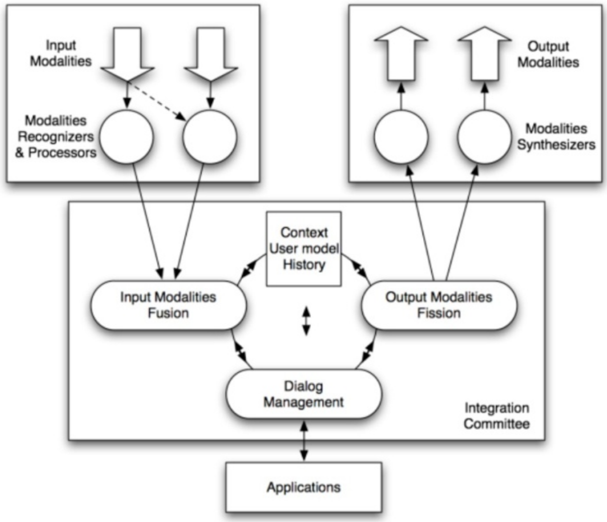

Fig. 1. The architecture of an adaptive multimodal system [18]

communicate to a given application, and the message to return to the user through the fission component. Then, depending on the user profile and context of use, the fission engine [21] is responsible for sending back a message to the user trough the most adequate modality or combination of output modalities. Finally, the context manager communicates any environment, content or user profile changes to fission, fusion and dialog mechanisms, so they can adapt their interpretations and perform context adaptation.

4.2 Machine-User Adaptation

Benyon said "Humans are adaptive systems. Computer systems can be made adaptive" [10]. When people have characteristics which they cannot easily change, or would prefer not to change, GUIDE automatically changes the way it appears in order to better suit those user characteristics. It adapts the user interface (font-size, background colors, number and size of buttons, etc) as well as the content shown (language, news, etc) to every single user interacting, whatever his context of interaction. So, we can say that GUIDE is a self-modifying adaptive system, because it is capable of automatically changing its components to reflect the ad-hoc needs of every user [2, 22].

Adaptation in GUIDE is a process, based on User models (profiles) and sets of rules, that needs data for producing output that makes sense. This data can be obtained initially by an initialization application (Initial Adaptation) and after that, inferred from the extended user interaction with the system (Run-time Adaptation).

Initial Adaptation. GUIDE adaptation begins trough an "Initialization application" that allows for the acquisition of primary assumptions about the user and is therefore a valuable source for initially assigning the user to a certain user model (or user profile). Only by doing this, is possible to offer each user the most appropriate interface and to pre-define a (multimodal) way of interaction. Knowing that each user model contains assumptions about interesting characteristics of user subgroups, after "going trough" the initialization application a user is assigned to a user model as certain preconditions are met (adaptability [23]). From that moment on, every time the user interacts with a GUIDE application, the system is "initially" adapted to him.

It's relevant to say that GUIDE initialization application is presented to the user as a simple step-by-step configuration of a "general" interface. In each step, different types of contents and different contexts of interaction are presented, so the user can test different components and parameters, and the system learns the user characteristics, from his impairments to his preferences.

Run-time Adaptation. After the initial adaptation, any user can already start using the system as it is already calibrated to his capabilities. However, each interface instance continues to be enhanced at run-time (adaptivity [23]), in order to provide a continuous and high-quality adaptation as the user evolves and gets more experienced with the system. So, at run-time the mechanisms of adaptivity need to infer likely events, based on their representations of users, contexts and applications, and on the experience of previous interactions. With the ability to alter the interface, functions are programmed into the system in a rule-based manner, and are selected according to specific well-defined criteria, refining the interaction with the system [2, 3]. For better understanding, we give some examples of GUIDE adaptation rules:

```
IF user_orientation = Low AND input_mod = Speech THEN volume = High
…
IF input_mod = Pointing AND menu_opt_number = 20 AND selection_time > 500
THEN menu_opt_number = 10 AND menu_opt_size = 20
…
IF user_experience = Veteran  THEN application_help_explanation = Off
```

In the first example, if the user is interacting with speech and GUIDE perceives he is disoriented, the volume is raised; In the second example, if the user is taking too much time to do a selection the system reduces the quantity and duplicates the size of options shown in a menu; In the third rule, if the user is experienced with the system, the help explanation is turned off (because he already knows how to interact).

User-Machine Adaptation. GUIDE is not only an adaptable system because it automatically adapts interactions and interfaces to the user, but also because it allows manual configuration of the system aspects by the user [22]. This means, the user also adapts to the system and not only the opposite, particularly by adjusting interface or interaction components to reflect his preferences at any time in any interaction. So, for example, if the user is reading information on the screen and he decides that he wants a bigger font size he just speaks to the system saying "Bigger text" and the interface automatically updates to fit that preference. In this way, the system is always "listening" to the user, who by using any type of input modality has the capacity of changing any part of the system (this also ensures that the user is fact in control of the system and not the opposite).

All this preferences are saved in a Preference Model that adds more efficiency and personalization to the user model to which the user was allocated when he begun using the system.

5 Interaction Context Adaptation

Context-awareness allows user interfaces to be adapted to new situations into an environment and interaction that evolves continuously. Considering applications for disabled or impaired users, and the necessity of adapting the user graphical interface to almost every single user, the coupling of information coming from context awareness and multimodality is indispensable [10, 16]. In GUIDE interaction, adaptation to several contexts is also important, because many users require applications that can be run in different environments, or with different devices. In context adaptation every situation is defined by a set of entities defining the system (applications, tasks and computational objects), the environment (place, physical objects and atmosphere conditions) and the users; by the relations these entities have with each other (ex: every task "has a" set computational objects capable of describing it, and every place "has a" certain environment and a set of physical objects); and a set of roles assigned to the entities, which typically describes every scenario of use of the application. A small change in the number of entities, or in the assignment of a role or even in a type of relation between two entities, forces a new situation and a new context, asking for context adaptation so the system can continue to perform well [9, 10, 11].

This means that for context-awareness and adaptation to be possible in GUIDE, we have to consider there are physical and social contexts related to the environment where every action is performed, meaning that is important to know how the user understands the physical place he is in, and also how he relates to all the social aspects related with the activity he is involved in. Also, there are several aspects concerning the context of the application that is being used, related to how every user should perform each action to achieve certain goals given a set of technical characteristics. Of course all of this only makes sense if we also take into account the most important contexts in the use of GUIDE, the ones related with physical, sensorial and cognitive capacities that each user has, defining which modalities one can use for interacting with the system.

Next, we present a list of context dimensions and variables that can interfere during or between interactions in GUIDE as in any typical adaptive system [14, 15, and 17]:

- Environment:
  - Physical: heating, lighting, noise level, acceleration and pressure (atmospheric conditions), user location, user position, object in the way (placed on the user or placed between the user and a sensor)
  - Social: background culture, user culture, people other than the user, time of day

- Application:
  - Software: application, task, modality of interaction, provision of help and training, time for interaction, communication
  - Technical: device of interaction, bandwidth (latency), device malfunction, resources available
- User:
  - Physical: motor impairments, tremors, reflexes
  - Sensorial: blindness, vision impairments, deafness, hearing impairments
  - Cognitive: reflexes, goals, memory, interest/boredom levels, affective state, focus of attention, interests

After defining all these dimensions, we present (Table 1) some use cases of context awareness and adaptation in GUIDE, showing in what way the system can adapt to every change in order to provide elderly and disabled users an intuitive interaction at all times.

Adaptation to contexts in GUIDE is also based in a set of rules. The activation of this rules, is caused by a specific change in environment, application or user contexts, which results in system (interface) adaptations as a mean to provide the user adequate interaction. So, for every single adaptation described in the table, will exist in the system an adaptation rule (or a set of them) responsible for the changes in GUIDE interface and interaction

6 The Role in Development of Accessible Applications

The development of accessible applications is one of the most important issues in the world of today's technology, as it is currently one of the main focuses considered by developers when designing new applications and interfaces. Despite this, if we ask any developer if he would be able to make different interfaces for each application he designs, so that every type of user could use it, his response would be a resounding "no".

Considering the adaptation mechanisms present in GUIDE, and all the efforts done to develop a system focused on accessibility, and capable of addressing all the differences between several types of users in one single platform, makes all sense that we extend its use to external developers. GUIDE can, in this way, be presented as a toolbox of accessibility and adaptation, a platform capable of almost automatically "translate" standard applications into several others suitable for other clusters of users. However, for this to be possible, developers would be "obligated" to use a specific programming language like HTML (with JavaScript and CSS) to make a "generic" interface of the application, and the GUIDE platform, making use of an impairment simulator (to infer about which content could be used by which cluster of users, and which could be not) and a set of translation rules, would automatically generate a set of different versions of the original interface. This way, every time a different user interacts with the application a "tuned" and "adapted" interface would be presented to him, making the application accessible to everyone.

Table 1. Use cases of context awareness and adaptation in GUIDE

Context of interaction	GUIDE solution
User takes off glasses during interaction.	Using the image from the motion sensors, GUIDE recognizes this situation and increases the size of content rendered on the screen (bigger buttons, font, etc.)
During user interaction, dog, cat or another person gets in the way, between the TV and the couch.	Using motion sensors, GUIDE captures a sudden change in the image captured, stops interaction (locking the screen), and sends a confirmation message to the user (using the output modality the user prefers or that is more appropriate). When the user is ready, confirms to GUIDE (by selecting a button, or issuing a voice command) and the system continues the interaction.
During user interaction, dog, cat or another person are captured by the motion sensor, side by side with the user, but closer to the screen.	Using motion sensors, GUIDE captures an additional and sudden presence, but ignores it focusing still on the first presence. The system sends a notification message to the user asking if there is a change of user and if nobody responds in a configurable number of seconds it ignores the second presence.
During user interaction, dog barks or person, other than the user, talks loud.	Using voice recognition the system… A) Can't identify the dog barks as GUIDE commands and ignores them. B) Recognize the other person tone of voice as not being from the original user, and ignores all the commands issued.
Change of lighting conditions caused by time of the day, or by turning on/off a light.	GUIDE interface adapts to the lighting conditions by changing the interface background, text and button color. For weak lighting, background will be darker, font and buttons will be lighter, and vice-versa.
Noise level increases in the living room or in the room where the user is interacting with the system.	System perceives the noise conditions making use of microphones, and… A) suggests the user changes from speech modality to another modality if the noise is too loud (typically to the second preferred modality of interaction) B) Asks the user for confirmation after each speech command. Or suggests a change between speech interaction and the use of at least two modalities (example: speech and pointing) for redundancy in each selection.
User is choosing a channel to watch on the TV and he receives a call from his son.	A notification is sent to the user asking him if he wants to receive the call. If user responds affirmatively GUIDE changes to the video-conference application and to the respective preferred modality of interaction.
User change positions while interacting with GUIDE.	Checks if there is still an image of a person captured by the motion sensor. If it can see the user it continues the same interaction automatically calibrating the motion sensor mechanism. If the user "disappeared", it sends a notification to the user and changes to the next modality of interaction preferred.
User change locations while interacting with GUIDE using pointing.	
The device in use runs out of battery.	Sends a notification to user and changes to next modality of interaction preferred.

7 Conclusion and Future Work

Multimodality is the central question in all GUIDE interaction. In this paper we prove that, multimodal adaptation techniques can be applied to the GUIDE interfaces to bring more consistency to all types of actions performed by any end user: First, the system can automatically adapt to the user by learning in each case of interaction, which modality he feels more comfortable with and with which he can deliver more actions with less effort. Second, the system can adapt to the user by inferring his intentions from the extended use of various modalities and the way he interacts, as well as from the interpretation of his User model and Preference model. Third, the system can recognize different users every time they make use of the TV. Forth, at any point in time, in any interaction, the user can modify every aspect of the interface to suit his preferences or needs, retaining the control of the system. Fifth, the system automatically adapts its interfaces to the context of every interaction, so that every time a change occurs the user can continue using the system in an effective manner.

In the future, to prove the statements, assumptions and concepts described above, we will continue implementing GUIDE platform and related applications, with special focus on easing accessible application development trough the GUIDE toolbox, and testing the system with users of several European countries.

References

1. European Project GUIDE – Gentle User Interfaces for Elderly People, `http://www.guide-project.eu/`
2. Benyon, D.: Accommodating Individual Differences trough and Adaptive User Interface.: Results and Prospects. Elsevier, North-Holland (1993)
3. Fink, J., Kobsa, A., Nill, A.: Adaptable and Adaptive Information Access for All Users, Including the Disabled and the Elderly. In: Proc. HYPERTEXT 1998, Pittsburgh (1998)
4. Kennel, A., Perrochon, L., Darvishi, A.: WAB: World-Wide Web access for blind and visually impaired computer users. New technologies in the education of the visually handicapped. ACM SIGCAPH Bulletin 55, 10–15 (1996)
5. Card, S., Moran, T., Newell, A.: The Psychology of Human-Computer Interaction. Lawrence Erlbaum Associates, Hilldale (1983)
6. Benyon, D., Murray, D.: Applying User Modeling to Human Computer Interaction Design. Artificial Intelligence Review 7(3), 199–225 (1993)
7. Biswas, P., Robinson, P.: Evaluating the design of inclusive interfaces by simulation. In: Proc. IUI 2010, pp. 277–280. ACM Press, New York (2010)
8. Biswas, P., Langdon, P.: Towards an inclusive world – a simulation tool to design interactive electronic systems for elderly and disabled users. In: Proc. SRII (2011)
9. Shackel, B.: Human Factors and Usability. In: Preece, J., Keller, L. (eds.) Human-Computer Interaction. Prentice Hall, Hemel Hempstead (1990)
10. Benyon, D.: Adaptive Systems: A Solution to Usability Problems. User modeling and User-adapted Interaction 3(1), 65–87 (1993)
11. Brusilovsky: Methods and techniques of adaptive hypermedia. User Modeling and User-Adapted Interaction 6(2-3), 87–129 (1996)

12. Anisetti, M., Bellandi, V., Ceravolo, P., Damiani, E.: Multimodality in Pervasive Environment. In: Semantic Changes in Sensor Networks. In: Dagstuhl Seminar Proceedings (2010)
13. Coutaz, J., Crowley, J.L., Dobson, S., Garlan, D.: Context is Key. Communications of the ACM 48(3) (March 2005)
14. Feki, M.A., Renouard, S., Abdulrazak, B., Chollet, G., Mokhtari, M.: Coupling context awareness and multimodality in smart homes concept. In: Miesenberger, K., Klaus, J., Zagler, W.L., Burger, D. (eds.) ICCHP 2004. LNCS, vol. 3118, pp. 906–913. Springer, Heidelberg (2004)
15. Schmidt, A., Beigl, M., Gellersen, H.W.: There is more to Context than Location. In: Proc. IMC 1998, Rostock, Germany (November 1998)
16. Dumas, B., Lalanne, D., Oviatt, S.: Multimodal interfaces: A survey of principles, models and frameworks. In: Lalanne, D., Kohlas, J. (eds.) Human Machine Interaction. LNCS, vol. 5440, pp. 3–26. Springer, Heidelberg (2009)
17. Costa, D., Duarte, C.: Adapting Multimodal Fission to User's Abilities. In: Stephanidis, C. (ed.) 14th International Conference on Human-Computer Interaction (HCII). LNCS, vol. 6765, pp. 357–366. Springer, Heidelberg (2011)
18. Costa, D., Duarte, C.: Self-adapting TV based applications. In: Stephanidis, C. (ed.) 14th International Conference on Human-Computer Interaction (HCII). LNCS, vol. 6765, pp. 367–374. Springer, Heidelberg (2011)
19. Feiteira, P., Duarte, C.: Adaptive Multimodal Fusion. In: Stephanidis, C. (ed.) 14th International Conference on Human-Computer Interaction (HCII). LNCS, vol. 6765, pp. 384–391. Springer, Heidelberg (2011)
20. Tran, T., Cimiano, P., Ankolekar, A.: A Rule-based Adaption Model for Ontology-based Personalization. In: Advances in Semantic Media Adaptatation and Personalization. SCI, vol. 93 (2008)
21. Stephanidis, C., Paramythis, A., Sfyrakis, M., Stergiou, A., Maou, N., Leventis, A., Paparoulis, G., Karagiannidis, C.: Adaptable and adaptive user interfaces for disabled users in the AVANTI project. In: Campolargo, M., Mullery, A. (eds.) IS&N 1998. LNCS, vol. 1430, pp. 153–166. Springer, Heidelberg (1998)