Conveying Multimodal Interaction Possibilities through the use of Appearances

Florian Weingarten, Marco Blumendorf and Sahin Albayrak

DAI-Labor, TU-Berlin

Ernst-Reuter-Platz 7, D-10587 Berlin, Germany

{Florian.Weingarten, Marco.Blumendorf, Sahin.Albayrak}@DAI-Labor.de

ABSTRACT

With the increasing importance of computers in all areas of life, new and innovative interaction concepts gain importance as current windows, icons, menus, and pointing concepts are rendered unusable. Well known graphical user interface currently move towards enhanced and multimodal interaction capabilities. In this paper we describe our approach to support this transition by extending graphical interfaces with multimodal interaction capabilities. Major aspects we focus on are the conveyance of the usable modalities as well as the fluent transitions between different modality combinations when the interaction context changes.

KEYWORDS

HCI, GUI, Multimodality, enhancing GUIs, Adaptation, Touch, Voice, Gesture, Smart Home, Home Control

1.INTRODUCTION

As the computer moves away from a production tool towards a ubiquitous life support facility, interaction with computers is undergoing a radical change. Instead of requiring the user to adapt to the user interface (UI), a trend to build UIs that adapt to the user is emerging. Additionally, the utilization of computers for new tasks and within new environments calls for the utilization of new interaction techniques. Mouse and keyboard are fast and efficient, but not very well suited for ubiquitous computing scenarios.

In our work we explore the utilization of multimodal interaction techniques within smart home environments. Such environments encompass challenging scenarios for human-computer interaction as they move the computer from a tool for efficient problem solving to a ubiquitous companion in everyday life. There is a lack of attention and focus while using such systems, combined with changing situations during the engagement in activities and conversations (e.g. an incoming phone call). This raises the need for new interaction techniques, which we explore in a smart home testbed, set up at the

Fig.1 Smart Home set up at CeBIT

Technical University of Berlin, and that we showed at this years CeBIT (Figure 1). Recently a focus has been set to multimodal interaction and the integration of new modalities with well known interaction concepts and paradigms.

In this paper we describe the utilization of ‘appearances’ to convey the possibilities of multimodal interaction to the user. In contrast to other (pure multimodal) approaches [1, 2], we keep the graphical user interface (GUI) in the center of the approach, as it provides a well known basis for interactions. However, we enhance the current presentation of GUIs with a new type of multimodal widget, conveying its interaction possibilities to the user. This allows a UI design with interaction elements that can be simultaneously controlled by touch, voice and gesture commands while providing extensive feedback mechanisms in and for various modalities. The problem we address with this approach in the first place, is the lack of transparency of interaction possibilities of ubiquitous user interfaces, by making them comprehensible and easy to conceive for the user. Due to the dynamic of situations and interaction, an enhancement of UI elements with a fluent transition between presentations for changing modality combinations is crucial.

2.APPEARANCES

In order to represent the availability of alternative modalities for control and to provide assistance of possible interactions, we introduce ‘appearances’ for any interaction element of the graphical user interface, such as buttons, lists, menus and so forth.

This basically means interaction elements are capable of changing their appearance, according to available modalities and therefore visually underline possible ways to interact with them. To assure a consistent overall look and integrity of the graphical user interface, it is important to define a basic shape of each element that can also be found in any of the elements appearances.

As an example we show the visual change of a graphical user interface element in figure 2.

From top to bottom: The basic appearance shows an inactive action element, meaning no interaction is possible, but already showing a basic shape and description that is referred to in any appearance. Adding the

Fig.2 Appearances of a GUI element

possibility of pressing it with a mouse click or through a touch device, changes the appearance into a more three-dimensional and tangible representation. Adding speech as an input modality to an inactive element results in changing the appearance into a well known symbol for direct speech, while keeping the basic shape. Also adding quotation marks to the descriptive word, supports the

user by precisely showing a possible speech input command. Adding gesture input, includes a small symbol to the appearance, which shows the gesture that needs to be performed to trigger the action. The last part of the figure shows the appearance of the element combining representations of the availability of touch, speech- and gesture input. It is an appearance of a multimodal interaction element.

3.EXAMPLE

Positioned in the overall home scenario, we show the adaptation of a UI depending on the changing availability of different modalities by utilizing appearances. As a part of a home control center that includes independent control of interconnected devices and home appliances, we show the visual changes of a user

Fig.3 Touch Control Interface

Fig.4 Speech Control Interface

Fig.5 Multimodal Control Interface

interface to control an oven. The software can manage basic features like turning it on and off, adjusting the desired temperature or setting a mode or a timer.

The UI is divided into three columns, one for basic data and information and the other two for interaction elements, that are subject to change their appearance. In the first figure, the look and feel of the interaction elements -in this case buttons- is well-known. By exchanging the touch or click input modality by speech, the appearance changes into speech bubbles with quotation marks, while keeping the overall shape (figure 2). With all mentioned modalities available (touch, speech and gesture), the appearances change into 'touchable speech bubbles', with an extra area indicating a possible gesture command. By providing this information with the help of familiar imagery and traceable change of shapes, the user is supported while dealing with flexible and dynamic input possibilities and can finally benefit from advantages that might emerge by using one specific- or a combination of modalities, that best suits the desired activity.

4.DISCUSSION AND OUTLOOK

The main advantage of the presented approach is that it provides the possibility to easily enrich graphical user interfaces to convey multimodal input capabilities. This is an important first step towards the utilization of truly multimodal interaction, but already provides huge advantages for users. By supporting human ways of communication and providing assistance for input possibilities, the approach supports the transition from the Windows, Icons, Menus, and Pointing (WIMP) paradigm to user interfaces that are more humane.

However, an important next step is the integration of more complex changes of the graphical user interfaces, as the integration of additional modalities might also influence the workflow of the applications and the way tasks are handled at all. A shift away from graphics as main modality opens a whole new area of interaction techniques with all its usability and development problems.

This also raises an important aspect for our future work, the application of 'appearances' to other modalities. There seems to be a need to give clues about interaction possibilities in other modalities as well (e.g. explain a set of possible gestures via voice). Appearances in this sense could also be a word or sentences, a tactile feedback or light signal of any source. Especially interesting in this sense is also the application of appearances to mobile devices and reduced screen spaces.

Additional steps are the incorporation of the learning curve of users, meaning that experienced users might need less clues and information than novice users and thus appearances can be learned by users and might change over time. A smooth transition towards embedded on demand help systems might be a suitable approach.

REFERENCES

- [1] N. Reithinger, J. Alexandersson, T. Becker, A. Blocher, R. Engel, M. Löckelt, J. Müller, N. Pflieger, P. Poller, M. Streit, V. Tschernomas, SmartKom - Adaptive and Flexible Multimodal Access to Multiple Applications, ICMI 2003
- [2] R. A. Bolt, "Put-That-There": Voice and Gesture at the Graphics Interface, SIGGRAPH 1980