IDEA GROUP PUBLISHING
701 E. Chocolate Avenue, Suite 200, Hershey PA 17033-1240, USA
Tel: 717/533-8845; Fax 717/533-8661; URL-http://www.idea-group.com
ITB12676

This paper appears in the book, *Emerging Trends and Challenges in Information Technology Management, Volume 1 and Volume 2* edited by Mehdi Khosrow-Pour © 2006, Idea Group Inc.

Advanced Multi-Modal User Interfaces for Mobile Devices: Integration of Visualization, Speech Interaction and Task Modeling

Norman Biehl & Antje Düsterhöft
Hochschule Wismar, University of Business, Technology & Design, Dept of EE & Computer Science, Multimedia Systems Research Group, Phillipp-Müller-Strasse, 23952 Wismar, Germany, {n.biehl, duest}@et.hs-wismar.de

Peter Forbrig, Georg Fuchs, Daniel Reichart (Software Engineering Research Group) & Heidrun Schumann
Computergraphics Research Group, University of Rostock, Dept of Computer Science, Albert-Einstein-Str. 21, 18051 Rostock, Germany, {pforbig, georg.fuchs, dreichart, schumann}@informatik.uni-rostock.de

ABSTRACT

The use of techniques from the fields of visualization, natural language and task modeling provides a new complementary style of human computer interaction, where the computer becomes an intelligent, active and personalized collaborator.

In this paper we present an adaptive, platform independent integration strategy of visualization, speech and task modeling techniques for technically oriented applications with the special focus on advance human-computer interfaces for mobile devices. The implemented system is illustrated on a maintenance support case study.

INTRODUCTION

With the ongoing pervasion of mobile devices and wireless networks, the use of digital manuals is about to become feasible with commonplace devices. However, creating a digital manual is not simply a matter of digitizing printed material. The capabilities of mobile devices vary strongly from smartphones to laptops. The same is true for the goals and proficiency of the users, e.g. the chief engineer compared to a subcontractor's employee.

This entails the need for advanced user interfaces in such information systems. They should incorporate state-of-the-art visualization, speech and interaction techniques tailored for specific tasks in order to utilize the varying capabilities of the respective target platform to the largest extend possible.

In technology-mediated spaces communication tends to be overly explicit, and work is often interrupted by the need for 'interaction management'. This is exacerbated by the limited interaction capabilities of many mobile devices compared to PCs. This leads to breakdowns or cumulative misunderstandings.

This paper presents a strategy of developing interactive systems, which is a combination of using results from visualization, natural language processing and model-based design. It shows that extended task models can be used to specify information for visualization and natural language processing. Additionally, the transformation from analysis models to navigation and presentation models is demonstrated.

The above aspects have been addressed in a scenario that covers maintenance tasks for air conditioning units. A software prototype for the scenario was presented on CeBIT 2005 fair. This contribution reports on conceptual and implementation issues pertaining to this prototype.

The paper is organized as follows. In section 2 we briefly address the main issues and challenges of the three components of our interactive system. Then we illustrate the integration strategy of the interaction techniques on the basis of our example application. We conclude with a discussion of synergy effects and appropriateness of the used techniques.

VISUALIZATION, SPEECH INTERACTION AND TASK MODELLING

Visualization

Due to the limited screen space of mobile handhelds, we can only present small parts of an information space at once. So the question is which information has to be shown with what granularity to guarantee effective and adequate visualizations. Effectiveness in this context means that the visual interface has to allow for an exact interpretation with minimal effort for the user. Adequacy paraphrases the fact that visual representations should neither omit necessary, nor contain superfluous, information for a particular task; also its creation must not require excessive system resources.

Achieving adequacy is supported by the task model as it allows identification of all relevant information. Effectiveness on small screens can be boosted by advanced concepts from the field of information visualization.

Efficient use of screen space by specific presentation techniques like Focus & Context [Kea98] or lens techniques [SFB94]: These techniques combine a focus view, which shows a part of the layout at a high degree of detail, and a context view, which presents the whole information in lower detail to provide an overview. The distinction between the two is that Focus & Context is applied to a visual representation as a whole, whereas lenses only modify a small region (the lens area), usually by blending in additional details. [FS06] discusses the use of Focus & Context in multimodal mobile interfaces in more detail.

On mobile handhelds however, a good tradeoff between complexity and response time is important, since computing arbitrary distortion functions can be very expensive, usually resulting in strictly functional interfaces without "eye-candy", see e.g. [Buy00].

Efficient navigation. Often, technical data is inherently structured, e.g. expressed by assembly-parts relationships or circuit connectivity. Navigation can then be realized by browsing specific representation of the structure, allowing to reach arbitrary information starting from one point of interest. Examples for the efficient generation and navigation of structures on mobile handhelds can be found e.g. in [KKS03].

Figure 1. Architecture of the NLI Framework

Natural Language Interaction

As stated above, navigation on mobile devices is often difficult and interaction can be oppressive. A language interface would enable easier designation and navigation and hence help users to complete their tasks faster.

The recent years have witnessed the emphasis on NLI with speech input and output as speech synthesis and recognition have made substantial progress and increasingly intersected with NLI [DüsT05].

Speech. Due to the technological advance, designers now have choices in high quality recorded speech and the synthesis of text to speech (TTS). Although the synthesis misses the characteristics of normal human speech, especially in mobile context TTS is an efficient way of realizing common presentation tasks.

For the interaction with the computer, speech input can be divided up into command-and-control and dictation input. Those inputs can be handled in two ways. Depending on the resources of the target device, speech input is analyzed on the device itself, or the recorded files are interpreted by a more powerful speech server.

Natural language processing. The challenge when developing NLIs is the amount of work required for modeling the application specific grammar and vocabulary and to combine the different language processing tools.

To overcome this drawback we focus on an adaptive integration system of language tools (cf. figure 1). Describing natural language grammars and semantic definitions based on several common specifications (SALT, SGRS, SML), we are able to integrate different language tools and languages.

However, a NLI on mobile devices should minimize its required resources on that device itself. This will be handled by a dictionary database system including several common and application-dependant dictionaries, as well as ontologies for language applications. For a new application, textual information is scanned for new application terms. A generation system allows the designer to create the specific part of the language for the prospective NLI.

Task Modeling

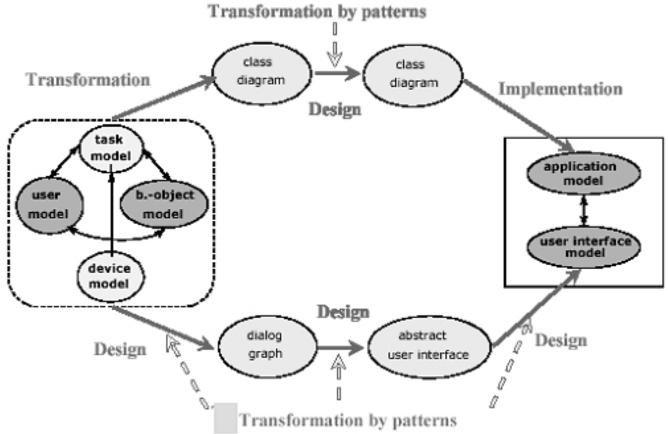
To meet the demands outlined above for both modes, visualization and natural speech, we propose a model-based approach that combines task, object and dialog models to specify platform-independent user interfaces. The introduction of new concepts like instance iteration for tasks or relating tasks to domain objects and dialog views allows us to generate abstract canonical prototypes.

Task modeling. In order to develop software based on user tasks and objects, several frameworks have been introduced. Our approach is characterized by figure 2. It demonstrates that the dialogue model and the application model have to be based on the same analysis specification, which consists of mutual related models of tasks, users, business objects and devices [ReiF05]. Software development is considered as a sequence of transformations mainly controlled by patterns. We already developed a tool for transformations controlled by design patterns.

Designing dialogues. There are different strategies to design the dialogue model. The Janus project [Cam05] uses information mainly from the object model, but most approaches are based on tasks. The project Teresa [Bal96] follows an idea of grouping tasks based on preconditions, which allows an automatic generation of dialogue models. Using our method of explicitly designing a dialogue graph an alternative strategy by designing a very abstract user interface can be employed (cf. figure 3).

We developed a dialogue graph editor that is able to read and write a file specifying task, user and object models as XML specification. A representation of dialogue graphs was developed as well. This editor

Figure 2. General view on a transformational model-based development process



allows authors to develop different dialogue graphs for the same task model, integrate visualization and speech functions, and transform them into a XML-GUI-description.

MAINTENANCE SUPPORT: INTEGRATION STRATEGY AND SYNERGY EFFECTS

Maintaining and repairing complex technical facilities (such as waterworks) requires comprehensive knowledge on a broad range of subsystems as well as operational and safety procedures on the part of the performing technician. Hence, there exists the need to have access to instruction sheets on site. This may not be practicable with printed manuals. So, there is a common need of a digital support of the maintenance work.

The mobile maintenance application is characterized as defined technical tasks that have to be executed goal-oriented dialogues the need of hand-free instructions while executing repairs.

The development of the multi-modal user interfaces is defined by general steps. First, a task model will be constructed on the basis of the printed manual. Secondly, the interaction types and parameters will be specified. After that, the description of the task model, the visualization functions and the natural language (cf. figure 4) will be transferred to a mobile device. The corresponding data will be stored into a database, which can reside on either server or client.

The integration of visualization, speech and task modeling techniques with the special focus on interfaces for mobile devices provides interaction components that can be fitted to work together. The task modeling component

- selects the interaction techniques depending on task, role and device
- supports users according to their level of experience
- realizes consistent updates for all interaction components and devices.

The visualization component derives appropriate visualization techniques according to specific tasks from the task model, while the speech component offers appropriate speech interaction techniques according to these tasks, mitigating the limited display size by ‘sonification’ of graphic content.

In compound interfaces, these two interface components are selected for output depending on their suitability. Visualization is used to help locate and identify component parts and to communicate interrelationships. Speech provides augmenting explanations, information that is requested repeatedly, and provides hands-free input while working on an assembly.

Figure 3 Parts of the task model of the application 'maintenance of air conditioning units'

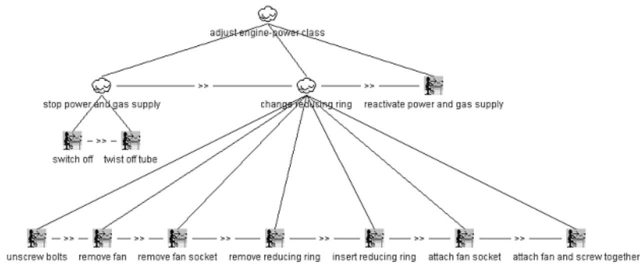


Figure 5. Screenshot of the GUI fragment corresponding to Figure 4

CONCLUSION AND FUTURE WORK

Above, we have outlined our strategy of integrating visualization techniques and speech interaction on the basis of a task model for realizing an advanced multimodal user interface. The strategy is implemented in a software design system that creates user interfaces in a consistent and platform independent way. In an example application, 'maintenance of air conditioning units', the effectiveness of the design system is shown. Actually, usability studies are performed that proof our assumptions about the in-work-handling of the implemented interface.

Ongoing research focuses on the development of a user model which is used in all interaction components of the system, the construction of a user-friendly dialogue structure for a mobile HCI on the basis of the presenting information coming from printed information (e.g. manuals), and the evaluation of special interaction techniques in visualization, speech interaction and task modeling that overcome the gaps between the different techniques.

REFERENCES

[Bal96] Balzert, H.: From OOA to GUIs: The JANUS System. In: Journal of Object-Oriented Programming, pp. 43-47, Feb. 1996.

[Buy00] Buyukokten, O., Molina, H.G., Paepcke, A., and Winograd, T.: Power browser: Efficient Web browsing for PDAs. Technical report. Stanford University, 2000.

[Cam05] Cameleon project, http://giove.cnuce.cnr.it/cameleon/SIGatCHI.html

[DüsT05] Düsterhöft A., Thalheim: DKE – Special Issue: Natural Language and Database and Information Systems, Volume 55(1) October 2005.

Figure 4. Example fragment of the GUI description

```
<Dialogdef:Name="Dialog" Text="Schritt 1" Load="step1_speak.Start">
<salt:prompt def:Name="step1_speak" Text="Strom- und Gasversorgung des
        Heizkessels unterbrechen"/>
<Controls>
<wf:Label def:Name="step1_text" Text="{step1_speak.Text}" Height="40"
        Dock="Top"/>
<wf:Button Text="^" Left="20" Top="40" Width="200" Height="20"
        Click="viscontrol.ScrollUp"/>
<wf:Button Text="&lt;" Left="0" Top="60" Width="20" Height="160"
        Click="viscontrol.ScrollLeft"/>
<wf:Button Text="&gt;" Left="220" Top="60" Width="20" Height="160"
        Click="viscontrol.ScrollRight"/>
<wf:Button Text="V" Left="20" Top="220" Width="200" Height="20"
        Click="viscontrol.ScrollDown"/>
<wf:Button Text="+" Left="95" Top="247" Width="20" Height="20"
        Click="viscontrol.ZoomIn"/>
<wf:Button Text="-" Left="122" Top="247" Width="20" Height="20"
        Click="viscontrol.ZoomOut"/>
<vis:VisualizationControl def:Name="viscontrol"
        BaseImageSrc="Abb07.png" Src="Abb07_bg.png"
        MaskImageSrc="Abb07_masks.png" MaskId="0" Left="20" Top="60"
        Width="200" Height="160"/>
<wf:Button Text="Weiter" Dock="Bottom" Click="Dialog.Weiter"/>
</Controls>
</Dialog>
```

[FS06] Fuchs G., Schumann H.: Visualization in Multimodal User Interfaces of Mobile Applications. Proceedings IRMA 2006, Washington D.C., USA, 2006.

[KKS03] Karstens, B.; Kreuseler, M.; Schumann,H.:Visualization of Complex Structures on mobile handhelds. Proceedings IMC'2003, International Workshop on Mobile Computing, Rostock, Germany, June 17-18, 2003.

[Kea98] Keahey, T. A. The generalized detail-in-context problem. In Proc. of the IEEE Symposium on Information Visualization, IEEE Visualization, October 1998.

[ReiF05] Reichart, D., Forbrig, P. Multiple-User Interfaces based on Task, User and Object Models. Workshop "Mobile Computing and Ambient Intelligence: The Challenge of Multimedia", Dagstuhl, May 2005.

[SFB94] Stone, M. C., Fishkin, K., Bier, E.A. The movable filter as a user interface tool. Proceedings of the ACM Conference on Computer Human Interaction, 1994.

Related Content