Adaptive Multimodal HCI with Uncertain Data by Collaborative Fission and Fusion

Felix Schüssel, Frank Honold, and Michael Weber

Institute of Media Informatics, Ulm University, Ulm, Germany
{felix.schuessel,frank.honold,michael.weber}@uni-ulm.de

Abstract. Multimodal systems that adapt their interface to the surroundings and the user shall be able to handle uncertain data provided by ubiquitous sensors. This has to be realized for the complete communication cycle and therefore demands collaboration between the system's output and input processes. In this article we demonstrate how dedicated fission and fusion components that handle uncertainty can be connected via an additional interaction content management component. We present the overall architecture of the resulting adaptive system and discuss the status quo of our implementation.

Keywords: Adaptive HCI, Multimodal Interaction, Fission, Fusion, Collaboration of Output and Input Processing.

1 Introduction

While interacting, human beings continuously adapt their way of communication to their surroundings and their communication partner. Although present context-aware ubiquitous systems gather a lot of information to maximize their functionality, they predominantly use rather static ways to communicate. In order to fulfill the user's communication needs and demands, ubiquitous sensors' diverse information can be used to dynamically adapt the user interface. Until now, reasoning algorithms used for that purpose are bound to clearly defined information, often ignoring the uncertain and sometimes ambiguous nature of information provided by diverse sensors. This means that the way of how different user input information fragments are fused should consider this uncertainty. In the same way the method of how a system represents its communicable information has to respect the same contextual real-world decision knowledge that is fraught with uncertainty.

In this paper we demonstrate, how multimodal fusion based on evidential reasoning and probabilistic fission with adaptive reasoning can act together to form a highly adaptive and model driven interactive system. A special component, called interaction content management facilitates this. The resulting system can handle uncertain or ambiguous data throughout the complete interaction cycle with a user.

2 Importance of Uncertainty in HCI

In order to realize flexible, adaptive interactive systems, a model-based procedure can be seen as the de facto standard approach in HCI. Starting from abstract models of tasks,

C. Stephanidis (Ed.): Posters, Part I, HCII 2013, CCIS 373, pp. 372–375, 2013.

users, dialog, information, and context knowledge, refinement steps automatically lead to a final user interface that is realized on the end use devices. The incorporation of uncertainty in this process is important, because only this allows for informed decisions in any situation, where non-perfect sensors gather user and context data. This not only holds true for the observed user actions containing possible inputs, but also for the information output, that shall be adapted to uncertain contextual knowledge. Although many approaches for each direction of communication exist, integrated systems, which offer the best of both worlds, are virtually non-existent. To fill this gap, and to form a highly dynamic and adaptive user interface, we introduce a so-called interaction content management, as explained in the next section.

3 Own Approach

The simplified overall architecture of our adaptive multimodal system is depicted in Figure 1. In order to realize a complete interaction loop, an abstract description of a dialog output is handed over from the application/dialog management to the multimodal fission component. Using an extensible set of decision functions, the best way to output the given information in the current situation and the currently available devices is computed and realized [5]. Our model-based implementation extends the often theoretical approaches of the state of the art [1-4] and introduces probabilistic reasoning to integrate uncertain and ambiguous knowledge [5]. It is designed to easily integrate new reasoning knowledge at runtime. With respect to the reasoning time, the approach is qualified to be used without any noticeable time lag while interacting.

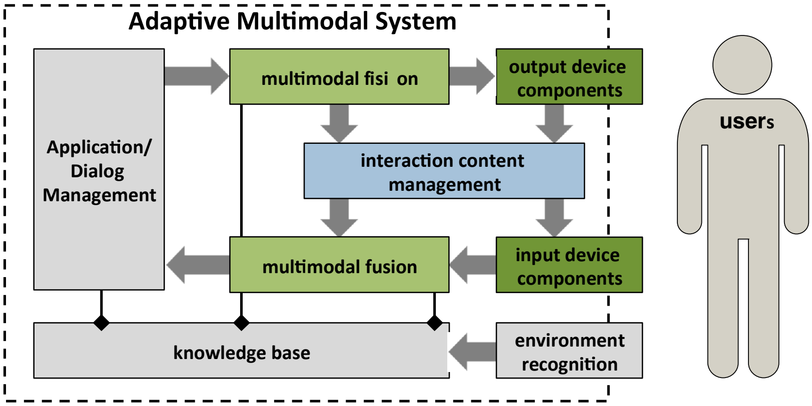

Fig. 1. Architectural overview of our adaptive multimodal system where the interaction content management mediates between the fission and fusion components

Once the concrete final user interface is realized on the assigned device components, the interaction content management (ICM) provides an interaction description of the realized output to the multimodal fusion component. This description is generated automatically from the fission's interaction output and the linked output device components' models. It contains information about all possible multimodal inputs the system may encounter. In addition, the ICM's information can be used as configuration data for the input device components. The fusion component then uses the approach described in [6] to receive uncertain observations from the available input device components, fuses them and decides on the most likely input. At the same time the robustness of input recognition is increased by the combination, disambiguation, reinforcement, and conflict detection capabilities of the adapted Transferable Belief Model (TBM) when compared to existing approaches [7-9]. This input is then handed back to the application/dialog management to finally proceed in the application logic. The aforementioned approaches, which only focus on one direction of interaction (input or output), are extended by the ICM, facilitating the full collaboration of both directions. Figure 2 shows an example for one complete interaction cycle of an assistant that helps users wiring up a home cinema system.

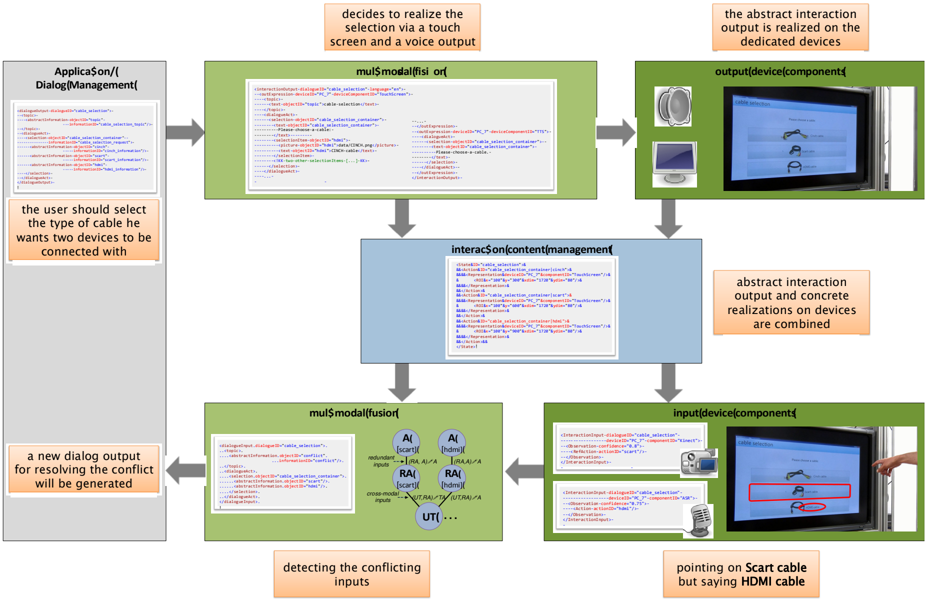

Fig. 2. Example of an interaction cycle from an assistant that helps the user to wire a home cinema system. The content management provides the necessary information, so that the multimodal fusion component can correctly detect a conflicting user input. As the graph shows, the action for scart "A [scart]" and the reference to hdmi "RA [hdmi]" are not directly connected via an edge and therefore yield a conflict if observed at the same time. All components communicate via XML messages, from which the most relevant are shown.

4 Discussion

The presented architecture and the embedded components represent a complete implementation of Norman's action cycle [10] on the side of the multimodal system. Both directions of interaction (fission and fusion) are realized via new methods yielding unprecedented capabilities compared to other approaches [5, 6] and are connected via the addition of the dedicated ICM component. Besides, our system is fully functional as presented in [11]. At this stage, the implementation allows for a user and context adaptive system of dialog navigation and selection tasks. Currently we are working on the full utilization of the fission and fusion capabilities in collaboration to allow a broader range of application domains. For the future, we plan to conduct a case study emphasizing the architecture's functionality and the ICM's benefit.

References

1. Costa, D., Duarte, C.: Adapting multimodal fission to user's abilities. In: Stephanidis, C. (ed.) Universal Access in HCI, Part I, HCII 2011. LNCS, vol. 6765, pp. 347–356. Springer, Heidelberg (2011)
2. Foster, M.E.: State of the art review: Multimodal fission. University of Edinburgh, Edinburgh (2002)
3. Rousseau, C., Bellik, Y., Vernier, F., Bazalgette, D.: A framework for the intelligent multimodal presentation of information. Signal Processing 86(12), 3696–3713 (2006)
4. Hina, M.D., Tadj, C., Ramdane-Cherif, A., Levy, N.: A Multi-Agent based Multimodal System Adaptive to the User's Interaction Context. In: Alkhateeb, F. (ed.) Multi-Agent Systems – Modeling, Interactions, Simulations and Case Studies, pp. 29–56. InTech (2011)
5. Honold, F., Schüssel, F., Weber, M.: Adaptive Probabilistic Fission for Multimodal Systems. In: Proceedings of the OzCHI 2012, pp. 222–231. ACM, New York (2012)
6. Schüssel, F., Honold, F., Weber, M.: Using the Transferable Belief Model for Multimodal Input Fusion in Companion Systems. In: Schwenker, F., Scherer, S., Morency, L.-P. (eds.) MPRSS 2012. LNCS, vol. 7742, pp. 100–115. Springer, Heidelberg (2013)
7. Holzapfel, H., Nickel, K., Stiefelhagen, R.: Implementation and evaluation of a constraint-based multimodal fusion system for speech and 3d pointing gestures. In: Proceedings of the 6th International Conference on Multimodal Interfaces, ICMI 2004, pp. 175–182. ACM, New York (2004)
8. Pfleger, N.: Context based multimodal fusion. In: Proceedings of the 6th International Conference on Multimodal Interfaces, ICMI 2004, pp. 265–272. ACM, New York (2004)
9. Dumas, B., Ingold, R., Lalanne, D.: Benchmarking fusion engines of multimodal interactive systems. In: Proceedings of the 2009 International Conference on Multimodal Interfaces, ICMI-MLMI 2009, pp. 169–176. ACM, New York (2009)
10. Norman, D.A.: Psychology of Everyday Things. Basic Books, New York (1988)
11. Honold, F., Schüssel, F., Nothdurft, F., Kurzok, P.: Companion Technology for Multimodal Interaction. In: Proceedings of the 14th ACM International Conference on Multimodal Interaction, ICMI 2012, pp. 67–68. ACM, New York (2012)