A Component-Based Approach: ICARE
Workshop: The Future of User Interface Design Tools

Jullien Bouchet, Benoit Mansoux, Laurence Nigay
CLIPS-IMAG, University of Grenoble 1
BP 53 38000 Grenoble, FRANCE
{jullien.bouchet, benoit.mansoux, laurence.nigay}@imag.fr

INTRODUCTION
Parallel to the development of the Graphical User Interface technology, natural language processing, computer vision, 3-D sound, and gesture recognition have made significant progress. In addition recent interaction paradigms such as perceptual User Interface (UI) [13], tangible UI [8], embodied UI [7] and Augmented Reality (AR) open a vast world of possibilities for interaction modalities including modalities based on the manipulation of physical objects and modalities based on the manipulation of a PDA and so on. Since the seminal "Put that there" demonstrator [2] that combines speech, gesture and eye tracking, significant achievements have been made in terms of both modalities and real multimodal systems. We distinguish two types of modalities: the active and passive modalities. For inputs, active modalities are used by the user to issue a command to the computer (e.g., a voice command). Passive modalities are used to capture relevant information for enhancing the realization of the task, information that is not explicitly expressed by the user to the computer such as eye tracking in the "Put that there" demonstrator [2] or location tracking for a mobile user.

In addition to many interaction modalities that are more and more robust, conceptual and empirical work on the usage of multiple modalities (CARE properties [10], etc.) are now available for guiding the design of efficient and usable multimodal interfaces.

Due to this conceptual and predictive progress and the availability of numerous modalities, real multimodal systems are now built in various domains including medical and military ones. One of our application domains is medical: Computer-Assisted Surgery (CAS) systems. CAS systems aim to help a surgeon in defining and executing an optimal surgical strategy based on a variety of multimodal data inputs. CAS systems often rely on AR interaction techniques [6], which are based on the fusion of the digital world (e.g. MRI, scan images, computed trajectory) with the real world (e.g. the patient's body, a needle). The variety of surgical needs and specialties as well as the multitude of interaction devices have led to many specific AR interaction techniques and therefore CAS systems that bring real clinical improvements. Nevertheless the design and development of such interaction techniques is still ad hoc and generally driven by technologies. Another of our application domains is military. We are working on multimodal augmented reality commands in the cockpit of French military planes. For example while flying, the pilot can mark a point on the ground by issuing the voice command "mark" and looking at a particular point [4]. Moreover multimodal interfaces are now playing a crucial role for mobile systems since multimodality offers the required flexibility for variable usage contexts, as shown in our empirical study of multimodality on PDA [15].

EXISTING SOFTWARE TOOLS
As pointed out in the description of the workshop, although several systems have been built, their development still remains a difficult task. The existing frameworks dedicated to post-WIMP interaction are currently few and limited in scope. Either they address a specific technical problem including the fusion mechanism [9], the composition of several devices [5] and mutual disambiguation [11], or they are dedicated to specific modalities such as tangible interaction, vision-based interaction [1], gesture recognition [14], speech recognition or the combined usage of speech and gesture .

ICARE APPROACH AND PLATFORM
We adopt a more global approach and we propose a component-based approach, called ICARE (Interaction CARE -Complementarity Assignment, Redundancy and Equivalence-), which allows the easy and rapid development of multimodal interfaces. ICARE framework is based on a conceptual component model [3][4] that describes the manipulated software components and we are currently finishing a tool, a graphical platform for specifying a multimodal interface by direct manipulation. From this specification (ICARE schema), the code of the multimodal interaction is automatically generated. While the ICARE framework provides the services of a toolkit, the ICARE platform must be compared to a UIMS built on top of the toolkit.

We identify two kinds of ICARE components: (1) elementary components that enable the designer to define "pure interaction modality", and (2) generic composition components (CARE composition components, Complementarity, Assignment, Equivalence and Redundancy) that enable the designer to specify combined usage of modalities. As opposed to elementary components, composition components are generic in the sense that they are not dependent on a particular modality. Elementary components are dedicated to interaction modalities. In [9] we define an *interaction modality* as the coupling of a physical device d with an interaction language L: <d, L>. A *physical device* is an artifact of the system that acquires (input device) information. Examples of devices include the mouse, microphone, GPS and magnetometer. An *interaction language* defines a set of well-formed expressions (i.e., a conventional assembly of symbols) that convey meaning. The generation of a symbol, or a set of symbols, results from actions on physical devices. Based on this definition of an interaction modality, we identify two types of elementary ICARE components, namely Device and Interaction Language components.

We concentrate on input (i.e., from the user to the system) although our model holds for output as well. Nevertheless we did not test our approach for output so far. For input we have developed several interfaces using our ICARE framework [3][4]: a Multimodal IDentification system (MID) and a mobile system MEMO. MEMO allows users to annotate physical locations with digital notes which have a physical location [12] and are then read/removed by other mobile users: two versions are currently running, one using a PDA and another one using a Head-Mounted Display. Figure 1 shows the MEMO ICARE specification for input multimodal interaction as an example of ICARE schema. Three tasks are possible using the modalities. They define what the rest of the system receives from the ICARE components: (1) orientation and localization of the user (T1) so that the system is able to display in the HMD the visible notes according to the current position and orientation of the mobile user (2) manipulation of a note (create, pick and remove a note) (T2) and (3) exit the system (T3). We also developed a larger system, FACET, a simulator of Rafale (a French military plane) [4]. And we are currently applying our ICARE approach for augmented surgery systems.

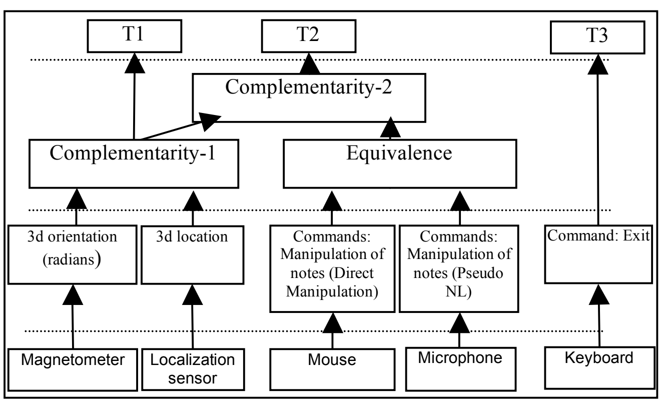

Figure 1. ICARE schema of MEMO input interaction [3].

SUMMARY OF CONTRIBUTION

In summary, we can contribute to the workshop in three ways: (1) use our ICARE approach as a starting point for further discussions on component-based approaches for developing this next generation of interfaces; (2) provide participants of the workshop with an interaction modality/multimodality perspective on post-WIMP interfaces; (3) discuss how software tools can support the development of various types of post-WIMP interfaces.

ACKNOWLEDGMENTS: This work is partly funded by DGA (French Army Research Dept.) under contract #00.70.624.00.470.75.96 and by the SIMILAR European FP6 network of excellence (http://www.similar.cc).

REFERENCES

1. Bérard, F. The Magic Table: Computer-Vision Based Augmentation of a Whiteboard for Creative Meetings. *IEEE workshop on Projector-Camera Systems, PROCAM*, Nice, France, 2003.
2. Bolt, R. Put that there: Voice and gesture at the graphics interface. *Computer Graphics* (1980), 262-270.
3. Bouchet, J., Nigay, L. ICARE: A Component-Based Approach for the Design and Development of Multimodal Interfaces. *Extended Abstracts CHI'04* (2004), 1325-1328.
4. Bouchet, J., Nigay, L., Ganille, T. ICARE software components for rapidly developing multimodal interfaces. *Proc. of ICMI'04* (2004), 251-258.
5. Dragevic, P., Fekete, J.-D. ICON: Input Device Selection and Interaction Configuration. *Demonstration, UIST 2002 Companion* (2002), 47-48.
6. Dubois, E, Nigay, L, Troccaz, J. Classification Space for Augmented Surgery, an Augmented Reality Case Study. *Proc. of Interact'99*, (1999), 353-359.
7. Harrison, B. et al., R. Squeeze me, Hold me, Tilt Me! An exploration of Manipulative User Interface. *Proc. of CHI'98* (1998), 17-24.
8. Ishii, H., Ullmer, B. Tangible Bits: Towards Seamless Interfaces between People, Bits and Atoms. *Proc. of CHI'97* (1997), 234-241.
9. Nigay, L., Coutaz, J. A Generic Platform for Addressing the Multimodal Challenge. *Proc. of CHI'95* (1995), 98-105.
10. Nigay, L., Coutaz, J. Multifeature Systems: The CARE Properties and Their Impact on Software Design. *Intelligence and Multimodality in Multimedia Interfaces*, AAAI Press (1997).
11. Oviatt, S. Taming recognition errors with a multimodal interface. *Comm. of the ACM*, 43, 9 (2000), 45-51.
12. Persson, P., Espinoza, F., Cacciatore, E. GeoNote: Social Enhancement of Physical Space. *Proc. of CHI2001 Ext. Abstracts* (2001), 43-45.
13. Turk, M., Robertson, G. Eds, Perceptual user Interfaces. *Comm. of the ACM*, 43, 3 (2000), 32-70.
14. Westeyn, T. et al. Georgia tech gesture toolkit: supporting experiments in gesture recognition. *Proc. of ICMI03* (2003), 85–92.
15. Zouinar, M. et al. Multimodal Interaction on Mobile Artifacts. *Communicating with smart objects-developing technology for usable pervasive computing systems*, Kogan Page Science (2003).