

Engineering Gestures for Multimodal User Interfaces

Florian Echter

Chair for Media Informatics
University of Regensburg
florian.echter@ur.de

Dietrich Kammer

Chair of Media Design
TU Dresden
dietrich.kammer@tu-
dresden.de

Davy Vanacken

Expertise Centre for Digital Media
Hasselt University - tUL - iMinds
davy.vanacken@uhasselt.be

Lode Hoste, Beat Signer

Web & Information Systems
Engineering Lab, Vrije
Universiteit Brussel
{lhoste,bsigner}
@vub.ac.be

ABSTRACT

Despite increased presence of gestural and multimodal user interfaces in research as well as daily life, development of such systems still mostly relies on programming concepts which have emerged from classic WIMP user interfaces. This workshop proposes to explore the gap between attempts to formalize and structure development for multimodal interfaces in the research community on the one hand and the lack of adoption of these formal languages and frameworks by practitioners and other researchers on the other hand.

Author Keywords

gestures; formal languages; APIs; multimodal user interfaces.

ACM Classification Keywords

D.2.2. Software Engineering: Design Tools and Techniques: User Interfaces; H.5.m. Information Interfaces and Presentation (e.g. HCI): Miscellaneous

WORKSHOP TOPIC

During the last five years, interest in multimodal user interfaces (MMUIs) has increased significantly in research as well as in a commercial context. Popular examples are smartphones which usually provide a multitouch screen, speech input and motion sensors, or the Microsoft Kinect, which enables full-body interaction with a game console. However, the design of development tools and application programming interfaces (APIs) has not kept pace with this new trend. Most widely used APIs such as the Android SDK or the Microsoft Surface SDK still follow the decades-old paradigm of triggering event-based callbacks and support a few hard-wired gestures at best. While numerous research projects have attempted to address these issues, they have so far failed to gain widespread adoption, with the possible exception of the low-level TUIO protocol [4]. Reasons for this low rate of adoption may include complex programming paradigms, lack of support for diverse input devices, inflexible GUI libraries or limited availability for popular operating systems.

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the Owner/Author.

Copyright is held by the owner/author(s).

EICS'14, June 17–20, 2014, Rome, Italy.
ACM 978-1-4503-2725-1/14/06.

<http://dx.doi.org/10.1145/2607023.2610288>

We wish to stimulate a discussion on these issues by inviting position papers between 2 and 6 pages in length on any of the following topics:

- Definition of and relationship between gestural and multimodal user interfaces
- Novel programming paradigms for MMUIs, e.g. reactive programming or visual programming
- Architectural concepts for multimodal UI APIs
- Studies on performance of different programming paradigms
- Analysis of limitations of existing multimodal API concepts
- Improved event structures for MMUIs
- Formal languages and concepts for describing multimodal interaction
- Arguments for/against formalization of multimodal user interfaces
- Reports on real-world deployment scenarios using existing APIs
- Standardization efforts regarding multimodal UIs by ISO, W3C, ...

WORKSHOP AIMS & GOALS

We hope to foster a lively discussion on the issues of multimodal interface development with the ultimate goal of developing a set of design guidelines that multimodal UI APIs should follow. We aim to include framework developers who *provide* such APIs as well as application developers who *use* these APIs to avoid an "ivory tower" discussion which might lead to the most elegant, formally rigid or most universal API design without taking the real-world requirements of application developers into account. The combined knowledge and experience of the workshop participants should lead to a modular and extensible concept for programming gestures for multimodal user interfaces.

WORKSHOP FORMAT

This workshop follows in the footsteps of the "Engineering Patterns for Multitouch Interaction" workshops at EICS 2010 and 2011. We envision a half-day workshop with approximately 8 - 12 participants. The first part (about 2 h) will consist of short presentations in Pecha Kucha style and questions

on the submitted papers, while the second part (also about 2 h, after a coffee break) will consist of 3 - 4 small discussion groups who each focus on a different sub-topic (selected from submitted position papers). After a second coffee break, a final wrap-up session (about 1 h) will reconcile the outcome of the discussion groups into a set of design recommendations for multimodal APIs. Total duration including breaks is estimated to be about 6 h.

Possible contributors to the workshop are researchers and practitioners in the area of multimodal framework development. A number of researchers have published individual approaches before, who will be explicitly invited to participate in the workshop. Among these approaches are Proton [7] by Kin et al., GDL [6] by Khandkar and Maurer, Midas [9] by Scholliers et al., GISpL [2] by Echtler and Butz, GeForMT [5] by Kammer et al., [10] MARIA by Spano et al., ICO [8] by Navarre et al., NiMMiT [1] by De Boeck et al. and GestureAgents [3] by Julià et al.

ORGANIZER CV

Florian Echtler is an assistant professor at the Chair for Media Informatics at University of Regensburg. His research interests focus on formal languages for gesture descriptions and design of multimodal development frameworks. Additional topics include computer vision for HCI applications, sensor technology and rapid prototyping.

Dietrich Kammer is a postdoctoral researcher at Technische Universität Dresden, affiliated with the Chair of Media Design. His research is focused on the formalization of gestural input, especially with regards to multitouch technology. Further areas of research are semiotics in HCI, computer graphics, and information visualization.

Davy Vanacken is a postdoctoral researcher at Hasselt University, affiliated with the HCI group of the Expertise Centre for Digital Media (EDM). His research is primarily focused on multi-touch and mid-air gestural interfaces for public and shared interaction spaces.

Lode Hoste is a PhD student at the Vrije Universiteit Brussels within the Web & Information Systems Engineering laboratory (WISE) and the Software Languages Lab (SOFT). His research focuses on software engineering abstractions for multimodal interaction, ranging from programming language design to modular and reusable framework architectures.

Beat Signer is Professor of Computer Science at the Vrije Universiteit Brussel (VUB) in Belgium, where he is co-director of the Web & Information Systems Engineering laboratory (WISE). His research interests include pen-based gesture recognition, multi-touch and multimodal gesture frameworks as well as cross-media information architectures.

WORKSHOP OUTCOME

The primary outcome is a set of design recommendations for multimodal user interface APIs. These recommendations as well as selected notes on their development process will be published on the workshop website along with the submitted position papers. An extended version of the design recommendations will also be submitted to an appropriate computer

science journal or scientific conference for further dissemination of the results.

ACKNOWLEDGEMENTS

On behalf of Dietrich Kammer, this work was supported by the European Union and the Free State Saxony through the European Regional Development Fund (ERDF).

The work of Lode Hoste is funded by an IWT doctoral scholarship.

REFERENCES

1. J. De Boeck, D. Vanacken, C. Raymaekers, and K. Coninx. High-level modeling of multimodal interaction techniques using NiMMiT. *JVRB - Journal of Virtual Reality and Broadcasting*, 4(2007)(2), Aug. 2007.
2. F. Echtler and A. Butz. GISpL: gestures made easy. In Proc. TEI '12, page 233–240, New York, NY, USA, 2012. ACM.
3. C. F. Julià, N. Earnshaw, and S. Jordà. GestureAgents: an agent-based framework for concurrent multi-task multi-user interaction. In Proc. TEI '13, page 207–214, New York, NY, USA, 2013. ACM.
4. M. Kaltenbrunner, T. Bovermann, R. Bencina, and E. Costanza. TUIO: a protocol for table-top tangible user interfaces. In Proc. GW '05, 2005.
5. D. Kammer, J. Wojdziak, M. Keck, R. Groh, and S. Taranko. Towards a formalization of multi-touch gestures. In Proc. ITS '10, page 49–58, New York, NY, USA, 2010. ACM.
6. S. H. Khandkar, S. M. Sohan, J. Sillito, and F. Maurer. Tool support for testing complex multi-touch gestures. In Proc. ITS '10, page 59–68, New York, NY, USA, 2010. ACM.
7. K. Kin, B. Hartmann, T. DeRose, and M. Agrawala. Proton++: a customizable declarative multitouch framework. In Proc. UIST '12, page 477–486, New York, NY, USA, 2012. ACM.
8. D. Navarre, P. Palanque, J.-F. Ladry, and E. Barboni. ICOs: a model-based user interface description technique dedicated to interactive systems addressing usability, reliability and scalability. *ACM Trans. Comput.-Hum. Interact.*, 16(4):18:1–18:56, Nov. 2009.
9. C. Scholliers, L. Hoste, B. Signer, and W. De Meuter. Midas: a declarative multi-touch interaction framework. In Proc. TEI '11, page 49–56, New York, NY, USA, 2011. ACM.
10. L. D. Spano, A. Cisternino, and F. Paternò. A compositional model for gesture definition. In Proc. HCSE'12, page 34–52, Berlin, Heidelberg, 2012. Springer-Verlag.