Künstl Intell
DOI 10.1007/s13218-015-0395-7

TECHNICAL CONTRIBUTION

# How *Companion*-Technology can Enhance a Multi-Screen Television Experience: A Test Bed for Adaptive Multimodal Interaction in Domestic Environments

**Jan Gugenheimer[1] · Frank Honold[1] · Dennis Wolf[1] · Felix Schüssel[1] · Julian Seifert[1] · Michael Weber[1] · Enrico Rukzio[1]**

Received: 5 June 2015 / Accepted: 7 September 2015


**Abstract** This article deals with a novel multi-screen interactive TV setup (smarTVision) and its enhancement through *Companion*-Technology. Due to their flexibility and the variety of interaction options, such multi-screen scenarios are hardly intuitive for the user. While research known so far focuses on technology and features, the user itself is often not considered adequately. *Companion*-Technology has the potential of making such interfaces really user-friendly. Building upon smarTVision, it's extension via concepts of *Companion*-Technology is envisioned. This combination represents a versatile test bed that not only can be used for evaluating usefulness of *Companion*-Technology in a TV scenario, but can also serve to evaluate *Companion*-Systems in general.

## 1 Introduction

Ever since television became the main source for entertainment and media in domestic environments, the static setup based on one fixed screen required all other things (including interiors and people) to be arranged around it [24]. More recently, this traditional setup is increasingly often supplemented by users taking advantage of *second screens* [4]. For instance, smartphones or tablet devices allow users to perform secondary tasks while the shared content on the main screen remains available for all users. For users, these second screens yield a number of advantages including social connectivity and sharing the experience with remote friends as well as quick access to additional background information supplementing the primary screen content [12].

These user needs were used to design and implement smarTVision, a *continuous projected display space* system. It enables users to create any number of second screens and place them in their environment in addition to their existing devices (Fig. 1). SmarTVision provides a flexible input and output space that enables diverse forms of interactions. Furthermore, we show how smarTVision can be enhanced using *Companion*-Technology [36] to create a highly adaptive system. We present smarTVision as a test bed to research *Companion*-Properties such as adaptability, individuality and availability.

This paper offers two main contributions: first, the design and implementation of smarTVision together with three example applications that illustrate how the flexible and novel design options can be utilized in different applications. Second, we present concepts for enhancing smarTVision to account for *Companion*-Properties such as adaptability, individuality and availability. We present how smarTVision can be used as a test bed to evaluate these and other characteristics of *Companion*-Technology.

## 2 Related Work

This work is influenced by a large body of previous research on second screen applications, augmented televisions, everywhere displays, and multimodal interaction in general and specific to *Companion*-Technology. After providing necessary definitions, this section gives insights in the state of the art of research-related topics.

✉ Jan Gugenheimer
jan.gugenheimer@uni-ulm.de

1 Institute of Media Informatics, Ulm University, James-Franck-Ring, 89081 Ulm, Germany

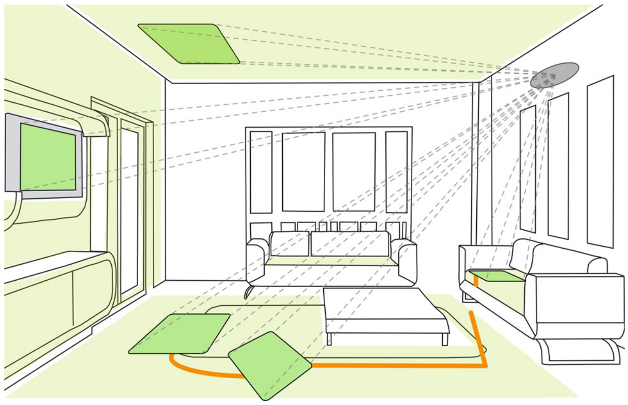

Fig. 1 Projected screens can be placed anywhere in the user's environment

## 2.1 Definitions

A *Companion*-System is a cognitive technical system that adjusts its functionality in a completely individual way to a particular user using so-called *Companion*-Technology. It realizes so-called *Companion*-Properties, such as individuality, adaptability, availability, cooperativeness and trustworthiness [36].

We use the term *modality* as equivalent to the term *interaction technique* as defined in [26]. The latter is defined as the combination of "a physical device *d* with an interaction language *L*: $\langle d, L\rangle$", where an interaction language defines a set of well-formed expressions (i. e., a conventional assembly of symbols) that convey meaning.

The term *Context of Use* (CoU) is based on the definition in [9], as "any information that can be used to characterize the situation of an entity. An entity is a person, place, or object that is considered relevant to the interaction between a user and an application, including the user and applications themselves."

## 2.2 Second Screen Applications

Second screen setups allow users to perform tasks parallel to other activities or other users without interfering with other users. Early work by Myers et al. presents first applications for second screen setups that facilitate interaction with distant displays and collaborative activities [25]. Further work investigated how personal second screens support collaborative planning tasks on a primary interactive surface [34]. Also in the context of television second screens have been used to provide additional information that supplements television content [1] and supports communication with the users' social network [2]. For instance, Robertson et al. used second screens to control media content displayed on a shared display [28]. More recently, a considerable amount of work investigated in large field studies how multiple devices are used for media play back [6] and what usage patterns of utilizing multiple devices simultaneously emerge through second screens [5].

## 2.3 Augmenting Television

The surrounding and immediate environment of television setups have been investigated as an option for extending the television experience, as well. For instance, Harboe et al. added an ambient display (i. e., ambient orb) to provide remote user presence information along with light weight messaging options [13]. With IllumiRoom, Jones et al. present projections to augment the immediate surroundings of a television to create a highly immersive experience [19]. In a follow-up project the TV was removed from the setup and the whole room became an interaction display [18]. Projections have also been used to render additional user interfaces next to a television set [35], a laptop computer (Bonfire) [21], and a mobile phone [37].

## 2.4 Everywhere Displays

Several approaches actuated the projected displays using motorized projectors that allow to freely position displays in the environment using a motorized projector [3, 11, 27]. Further, the whole floor of a room can be used as a display which enables versatile foot-based interaction options [30].

## 2.5 Multimodal Output Configuration

To configure the use of multiple modalities for input and outputs, we can distinguish between three different concepts. First, the multimodal combination is pre-defined by system developers before runtime. These hard-coded approaches leave little leeway for user-individual and CoU-specific adaptive behavior at runtime. Second, the user is able to configure the multimodal user interface at runtime (cf. [7, 29]). These approaches often make use of a so-called *meta UI*, which allows the user to re-arrange and configure the use of applied modalities for a specific setting. Since it is the user who is in charge to specify each adaptation at runtime, this concept is rather cumbersome to perform ongoing adaptations to the CoU. At last, the third class comprises approaches, where the system autonomously adapts itself to the sensed changes in the CoU. Such systems rely on an adequate sensory system to aggregate knowledge about the CoU, as described in [10, 15]. A specialized fission module is in charge to reason about the most adequate UI configuration with each change in the CoU [15, 16].

### 2.6 Multimodal Input Fusion

Gathering user inputs from multiple modalities and combining them where applicable has been a field of extensive research for several years now. Basically, there are four different strategies to tackle this problem, as there are procedural, frame-based, unification-based, and statistical/hybrid approaches. See [22, 23] for details and a comprehensive list of examples. Within the field of *Companion*-Systems, the input fusion has to deal with uncertain inputs from various sensors on an abstract level. Therefore, we will apply the approach from [31], as it allows complete handling of uncertainties within the tasks of combination, reinforcement, disambiguation, and conflict detection of user inputs.

Building upon and integrating the related work, we propose a complete interactive environment for the instantiation and evaluation of *Companion*-Systems. The overall scenario is that of a TV-like environment that is extended by optional and flexible second screens for the system output and flexible multimodal possibilities for user inputs as detailed in the next section.

## 3 User Interface and Interaction Concepts

We designed the concept for smarTVision to overcome limitations of current TV setups. Essentially, the concept comprises output options (for visualizing content) and input options (supporting different possibilities for interaction) via different devices and modalities [33].

### 3.1 Visualization Options

As illustrated in Fig. 1, smarTVision's conceptual display space spans across the ceiling, the wall, and the floor. We focus on GUI-dominated scenarios with one primary screen and multiple optional secondary screens. Within this space, the secondary screens can be freely placed to provide informative displays in any prominent places. This way, for instance, secondary screens placed on the ceiling may be suitable for content that is only of limited interest to users. Each screen can be subdivided to generate smaller screen patches (see the four screen patches on the floor in Fig. 2b). The output is based on logical content containers which can be rendered as full-screen widgets on each particular screen.

In addition to the projected screens, *visual links* can illustrate the coherence of multiple distributed screens. Such links support users to easily understand which semantic concepts from what screens belong together. As an example, a social media feed can be connected to the primary television content.

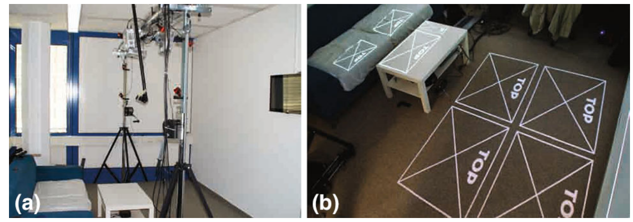


**Fig. 2** The prototype hardware setup: a traverse mounted on two tripods spans across the room, holding a depth camera and projectors (**a**). The projected display space of the prototype allows to create several surfaces (**b**)

A third visualization option of our implemented system are *visual indicators* that render highlights on physical objects in the user's environment. For instance, in a quiz game played by several users, results of each individual user's guess could be visualized by projecting such indicators onto each specific player.

The use of personal *mobile devices* forms the forth concept for visualization. With a dedicated app, a user's smartphone or tablet can also be used to represent information using the aforementioned widget concept.

### 3.2 Interaction Options

With smarTVision allowing various screen configurations at any position in their environment, several spatial constellations arise between a user and the interface, the user wishes to interact with. Due to this flexibility, interaction options need to support interaction with interfaces and displays across different distances. Therefore, the system offers interactions via touch, gestures, and speech, as well as multimodal interactions.

*Touch* In case the screen is in the user's immediate vicinity (e. g., on the couch table) or is on a mobile device, direct touch-based interaction is an option provided in the smarTVision concept.

*Gestures* In case a screen is placed at a remote position relative to the user, hand gestures in mid-air as well as pointing gestures can be conducted.

*Speech* Independent of screen positions, the user can always perform speech interactions, e. g. by uttering commands.

Using these modalities in a purely unimodal way, i. e. one at a time and with self-sufficient meaning, would severely restrict the interaction options. To overcome this, the concept should also allow *multimodal interactions* as defined by the CARE-properties of Coutaz et al. [8]. This way, especially complementary inputs combining multiple modalities like deictic references from speech and pointing gestures allow a much wider range of interactions and increase the naturalness for the user.

## 4 Implementation

In order to investigate the smarTVision concept more in depth, a prototype system was designed and implemented.

The hardware of the prototype setup comprises a stage lighting rig that is mounted on two tripods (see Fig. 2a). This rig spans across a couch and a couch table (which are typical pieces of furniture in most living rooms). In order to render projected second screens (Fig. 2b), three BenQ W1080ST full HD projectors are mounted to the rig. Two of them are facing the floor and one projector is responsible for projecting on the wall, which is facing the user sitting on the couch. A fourth projector is placed in front of the couch table facing the ceiling in order to provide the ceiling display.

These projectors yield a display space which allows to render any visual content (i. e., TV content or interfaces) that comprises the couch, the couch table, the floor around and in front of the table, the wall, and the ceiling (Fig. 2b).

In addition to the projectors, a Microsoft Kinect depth camera is attached to a pole that is mounted on the rig (Fig. 2a). The depth camera is facing the floor and in particular the area of the couch and couch table in order to support touch-based interaction on these. Finally, a Leap Motion sensor is attached to the border of the couch table, which is used to support mid-air hand-gestures.

The software architecture draws on the *UbiDisplays* framework by Hardy [14] and includes custom modifications to support the distribution of multiple second screen applications. In order to manage complex applications, the smarTVision implementation includes a central server for coordinating the internal application logic and corresponding states (in particular important if several surfaces access a shared data model or timing critical content). The server is written in Node.js [20].

## 5 Applications

In order to explore and to illustrate one possible scenario that is supported by smarTVision, a television scenario was implemented including several demo applications. Three selected examples illustrate how users can benefit from the smarTVision concepts.

### 5.1 Meta User Interface (Meta UI)

Therefore, the meta UI uses a straightforward concept to serve as a mediator for users to place, move, or delete widgets (Fig. 3c). This interface provides a schematic representation of the environment and offers predefined locations where widgets can be placed (e. g., a social media message feed next to the user on the couch). Instead of using predefined locations, the meta UI could be implemented to allow free positioning by using structured light for automatic surface recognition.

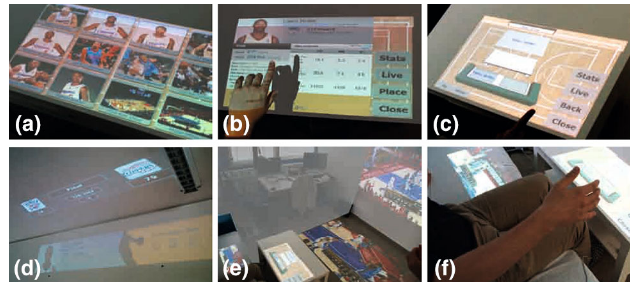


**Fig. 3** After selecting a player from the overview (**a**) the user can view detailed information (**b**) and place it via the screen manager (**c**). This way, screens can be placed within the whole environment of the user (**d**, **e**). Remote interfaces can be operated by mid-air gestures (**f**)

### 5.2 Sports Play Application

The first content-specific example application supports following a basketball game broadcast. This application aims for providing most different perspectives and views (on different players), as well as different content types (e. g., game statistics, social media etc.) in order to allow users to follow all kinds of aspects that matter during such complex game play.

A central menu serves as a player overview that is displayed on the couch table (Fig. 3a). By selecting a player via touch, a detailed player view is opened (see Fig. 3b). Here, users can select to open and place (Fig. 3c) the player-specific camera view (which constantly follows this particular player) using the metaUI. This allows users to arrange any number of different views in their environment such as camera views of specific players, an overview camera (Fig. 3e), and game statistics for instance at the ceiling (Fig. 3d). Users can easily browse through different statistics by using hand gestures to *swipe* to the next page (Fig. 3f).

### 5.3 Quiz Application

Another example application supports users to play along while watching a quiz show. This application can be played either by one or two players in the current implementation. Users are provided with a widget that contains the answer options (Fig. 4a). Next to the user on the couch, a small selection widget is projected, which allows users to select the answer option they think is correct. By the time, the answer is revealed in the quiz show, corresponding feedback is provided through a visual indicator, which illuminates the user with a red (wrong answer) or a green light (correct answer) (Fig. 4b).

While the presented demo applications offer a wide range of interaction possibilities, they were not specifically

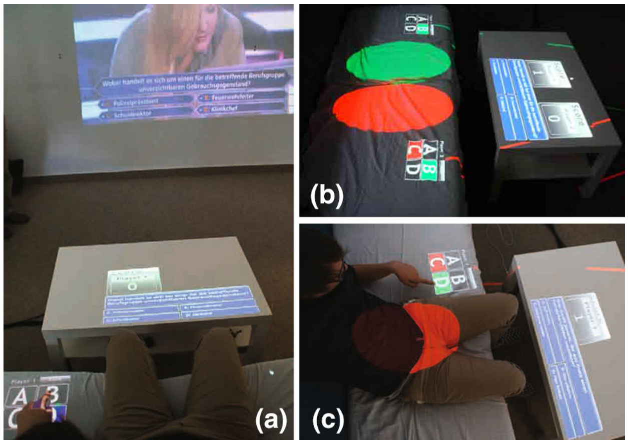


**Fig. 4** Answer options can be selected via a small interface next to the user (**a**). Depending on the selection, the user is illuminated in a *red* (wrong) or *green* (correct) light (**b**, **c**).

designed with *Companion*-Technology in mind. Therefore, the following section presents their enhancement with certain aspects of *Companion*-Technology in order to explicitly support *Companion*-Properties.

## 6 Interaction, Adaptation and Exploration

Currently, smarTVision only has some implicit *Companion*-Properties such as "high flexibility", which is achieved through multi-modality. This section describes how *Companion*-Technology can be used to enhance the interaction with the multi-screen television experience on the one hand, and how such a user-controlled TV setting can serve as a test bed on the other hand.

### 6.1 Automatic Output Configuration

The *Companion*-System's fission component can be used to reason about the applied devices and modalities in order to support properties of individuality, adaptability and availability. In that way, the system is able to automatically react to changes in the CoU as described in [16]. If the user moves around, e. g. to fetch some snacks, the primary display's TV content could accompany him e. g. with the use of the motorized projectors. The fission component also manages to distribute and locate the desired secondary screen contents depending on their importance flags (cf. ceiling projection as motivated in Sect. 3.1). During game play, e. g. in quiz-mode, the fission could decide about the most adequate location for each player's private widget for answer options (see Fig. 4). In multi-user settings, such widgets could be rendered on the sofa next to the individual user or on the user's mobile device. The coffee table would then be out of question to ensure each player's privacy.

Although the automatic fission process replaced the former manual configuration task via the system's meta UI controller, the meta UI concept can still be used to communicate a user's suggestions for improvement. In that way the meta UI can be used to offer nominations for particular modalities as motivated in [17].

The fission component of smarTVision decides not only about the output, it also reasons about the applied sensors for possible explicit user inputs. This setting for input and output, the so-called *abstract interaction model* (AIM), is provided as configuration for the fusion component, as it describes all possible and valid user interactions (cf. [17]).

### 6.2 Adaptive Input Fusion

In order to allow rich multimodal interactions as described in Sect. 3 a multimodal input fusion component is needed. As the possible inputs and their semantic meaning may greatly vary, depending on the application at hand and the variable output configuration, this input fusion can not be predefined, but must be configured at runtime. Using the approach described in [17], the *Companion*-System's input fusion component can be configured automatically at runtime using the aforementioned AIM that provides it with an abstract description of the currently possible interactions inferred from the current output configuration. Properties of availability and adaptability benefit most from using such an flexible concept of input fusion.

### 6.3 User-Centered Testing

The well-equipped TV setting offers diverse possibilities to test novel approaches for fission, fusion and *Companion*-Technology in general.

The use of motorized projectors [11] increases the number of possible areas for output. Depth cameras or structured light scans can be used to identify suitable areas for possible projections. This paradigm of steerable everywhere displays lead to new research questions, which can be further elaborated in the presented test bed. We are interested in how users prefer to interact with moving displays, and how such a dynamic aspect influences present reasoning approaches for modality arbitration. With the use of the presented Meta UI, the presented test bed allows to share control over this reasoning process, and further allows to gain insights into a user's decision process and his individual UI-specific desires and dislikes.

The *Companion*-Properties of availability and trustworthiness could be supported by increasing the robustness of input understanding. This could be achieved by implementing the concept of adapting to individual interaction history as described in [32], where sensor errors are detected and avoided on the basis of an individual user's known temporal input behavior. The demo applications allow for a wide range of multimodal inputs and therefore represent an ideal testbed. Furthermore, general insights in a user's individual preferences when it comes to the choice and usage of multiple modalities in such a scenario can be gathered as well.

Possible future enhancements include the extension of the system with additional components, like planning and dialog management (cf. [36]), that would allow the realization and evaluation of a holistic *Companion*-System.

## 7 Conclusion and Lessons Learned

We presented smarTVision, a continuous projected display space system that enables users to create any number of second screens and place them in their environment. To illustrate the utilization of this concept, we implemented three example applications that each draw on different design options. Furthermore, the presented system works as a test bed for *Companion*-Technology and especially adaptive multimodal interaction in the context of a TV-like environment. It allows a maximum of flexibility when it comes to interaction options. The output is not restricted to a fixed set of display devices, but can make use of any surface in the surroundings of the user. Likewise the possibly available inputs for the user cover touch, gesture and speech inputs, as well as multimodal interactions.

The presented and implemented scenarios showcase the flexibility and technical feasibility of the overall concept. Using modular and reusable components like the described fission and fusion, the test bed can not only be used to evaluate the interaction with a *Companion*-System, but also allows to integrate additional concepts, like dialog management and planning components. This way, it can serve as a complete test environment for *Companion*-Technology itself, the way users interact with it and the experiences they make. Future desirable extensions are the integration of sensory equipment and software components to infer the emotional state of users and the extension to multiple simultaneous users.

**Acknowledgments** This work was supported by the Transregional Collaborative Research Center SFB/TRR 62 "Companion-Technology for Cognitive Technical Systems", which is funded by the German Research Foundation (DFG).

## References

1. Basapur S, Harboe G, Mandalia H, Novak A, Vuong V, Metcalf C (2011) Field Trial of a Dual Device User Experience for iTV. In: Proceedings of the EuroITV '11. doi:10.1145/2000119.2000145
2. Basapur S, Mandalia H, Chaysinh S, Lee Y, Venkitaraman N, Metcalf C (20102) FANFEEDS: Evaluation of Socially Generated Information Feed on Second Screen As a TV Show Companion. In: Proceedings of the EuroiTV '12. doi:10.1145/2325616.2325636
3. Cauchard JR, Fraser M, Han T, Subramanian S (2012) Steerable projection: exploring alignment in interactive mobile displays. Pers Ubiquitous Comput. doi:10.1007/s00779-011-0375-3
4. Cesar P, Bulterman DCA, Jansen AJ (2008) Usages of the secondary screen in an interactive television environment: control, enrich, share, and transfer television content. In: Changing television environments. LNCS. Springer, Heidelberg . DOI 10.1007/978-3-540-69478-6_22
5. Courtois C, D'heer E (2012) Second screen applications and tablet users: constellation, awareness, experience, and interest. In: Proceedings of the EuroiTV '12. doi:10.1145/2325616.2325646
6. Courtois C, Schuurman D, De Marez L (2011) Triple screen viewing practices: diversification or compartmentalization? In: Proceedings of the EuroITV '11. ACM, New York, pp 75–78. doi:10.1145/2000119.2000132
7. Coutaz J (2007) Meta-user interfaces for ambient spaces. In: Coninx K, Luyten K, Schneider K (eds) Task models and diagrams for UI design. LNCS, vol 4385. Springer, Heidelberg, pp 1–15. doi:10.1007/978-3-540-70816-2_1
8. Coutaz J, Nigay L, Salber D, Blandford A, May J, Young RM (1995) Four easy pieces for assessing the usability of multimodal interaction: the CARE properties. In: Proceedings of INTERACT95, pp 115–120
9. Dey AK, Abowd GD (1999) Towards a better understanding of context and context-awareness. In: HUC '99: Proceedings of the 1st int. symposium on Handheld and Ubiquitous Computing. Springer, Heidelberg, pp 304–307
10. Glodek M, Honold F, Geier T, Krell G, Nothdurft F, Reuter S, Schüssel F, Hörnle T, Dietmayer K, Minker W, Biundo S, Weber M, Palm G, Schwenker F (2015) Fusion paradigms in cognitive technical systems for human-computer interaction. Neurocomputing 161:17–37. doi:10.1016/j.neucom.2015.01.076
11. Gugenheimer J, Knierim P, Seifert J, Rukzio E (2014) Ubibeam: An interactive projector-camera system for domestic deployment. In: Proceedings of the ITS '14. ACM, New York, pp 305–310. doi:10.1145/2669485.2669537
12. Häkkilä JR, Posti M, Schneegass S, Alt F, Gultekin K, Schmidt A (2014) Let me catch this! Experiencing interactive 3D cinema through collecting content with a mobile phone. In: Proceedings of the CHI '14). doi:10.1145/2556288.2557187
13. Harboe G, Metcalf CJ, Bentley F, Tullio J, Massey N, Romano G (2008) Ambient social Tv: drawing people into a shared experience. In: Proceedings of the CHI '08. doi:10.1145/1357054.1357056
14. Hardy J, Alexander J (2012) Toolkit support for interactive projected displays. In: Proceedings of the MUM '12. doi:10.1145/2406367.2406419
15. Honold F, Bercher P, Richter F, Nothdurft F, Geier T, Barth R, Hörnle T, Schüssel F, Reuter S, Rau M, Bertrand G, Seegebarth B, Kurzok P, Schattenberg B, Minker W, Weber M, Biundo S (2014) Companion-technology: towards user- and situation-adaptive functionality of technical systems. In: IE'14: Proceedings of the 10th international conference on intelligent environments, pp 378–381. IEEE (2014). doi:10.1109/IE.2014.60. Video: http://companion.informatik.uni-ulm.de/ie2014/companion-system.mp4
16. Honold F, Schüssel F, Weber M (2012) Adaptive probabilistic fission for multimodal systems. In: Proceedings of the 24th Australian computer-human interaction conference, OzCHI '12. ACM, New York, pp 222–231. doi:10.1145/2414536.2414575

17. Honold F, Schüssel F, Weber M (2014) The automated interplay of multimodal fission and fusion in adaptive HCI. In: IE'14: Proceedings of the 10th international conf.erenceon intelligent environments, pp 170–177. doi:10.1109/IE.2014.32
18. Jones B, Sodhi R, Murdock M, Mehra R, Benko H, Wilson A, Ofek E, MacIntyre B, Raghuvanshi N, Shapira L (2014) Roomalive: magical experiences enabled by scalable, adaptive projector-camera units. In: Proceedings of the UIST '14. doi:10.1145/2642918.2647383
19. Jones BR, Benko H, Ofek E, Wilson AD (2013) IllumiRoom: peripheral projected illusions for interactive experiences. In: ACM SIGGRAPH 2013 emerging technologies. doi:10.1145/2503368.2503375
20. Joyent Inc (2014) Node js. http://nodejs.org/. Access:ed 29 May 2015
21. Kane SK, Avrahami D, Wobbrock JO, Harrison B, Rea AD, Philipose M, LaMarca A (2009) Bonfire: a nomadic system for hybrid laptop-tabletop interaction. In: Proceedings of the UIST '09. doi:10.1145/1622176.1622202
22. Lalanne D, Nigay L, Palanque P, Robinson P, Vanderdonckt J, Ladry JF (2009) Fusion engines for multimodal input: a survey. In: Proceedings of the 2009 international conference on multimodal interfaces, ICMI-MLMI '09. ACM, New York, pp.153–160. doi:10.1145/1647314.1647343
23. aViola JJ Jr, Buchanan S, Pittman C (2014) Multimodal input for perceptual user interfaces, chap. 9. Wiley, New York, pp 285–312. doi:10.1002/9781118706237.ch9
24. Miller LJ (1995) Family togetherness and the suburban ideal. Sociol Forum 10(3):393–418
25. Myers BA (2001) Using handhelds and pcs together. Commun ACM. doi:10.1145/384150.384159
26. Nigay L, Coutaz J (1995) A generic platform for addressing the multimodal challenge. In: CHI '95: Proceedings of the SIGCHI conference on human factors in computing systems. ACM Press/Addison-Wesley Publishing Co., New York, pp 98–105. doi:10.1145/223904.223917
27. Pinhanez C (2001) The everywhere displays projector: a device to create ubiquitous graphical interfaces. In: Ubicomp 2001: ubiquitous computing. LNCS. Springer, New York
28. Robertson S, Wharton C, Ashworth C, Franzke M (1996) Dual Device User Interface Design: PDAs and Interactive Television. In: Proceedings of the CHI '96. doi:10.1145/238386.238408
29. Roscher D, Blumendorf M, Albayrak S (2009) Using meta user interfaces to control multimodal interaction in smart environments. In: Meixner G, Görlich D, Breiner K, Hußmann H, Pleuß A, Sauer S, den Bergh JV (eds) Proceedings of the IUI'09 workshop on model driven development of advanced user interfaces, CEUR Workshop Proceedings ISSN 1613-0073, vol 439
30. Schmidt D, Ramakers R, Pedersen EW, Jasper J, Köhler S, Pohl A., Rantzsch H, Rau A, Schmidt P, Sterz C, Yurchenko Y, Baudisch P (2014) Kickables: tangibles for feet. In: Proceedings of the CHI '14. doi:10.1145/2556288.2557016
31. Schüssel F, Honold F, Weber M (2013) Using the transferable belief model for multimodal input fusion in companion systems. In: Schwenker F, Scherer S, Morency LP (eds) Multimodal pattern recognition of social signals in HCI. LNCS, vol 7742. Springer, New York, pp 100–115. doi:10.1007/978-3-642-37081-6_12
32. Schüssel F, Honold F, Weber M, Schmidt M, Bubalo N, Huckauf A (2014) Multimodal interaction history and its use in error detection and recovery. In: Proceedings of the 16th ACM international conference on multimodal interaction, ICMI '14. ACM, New York, pp 164–171. doi:10.1145/2663204.2663255
33. Seifert J (2015) Mobile mediated interaction with pervasive displays. Ph.D. thesis, Ulm University,Ulm
34. Sugimoto M, Hosoi K, Hashizume H (2004) *Caretta*: a system for supporting face-to-face collaboration by integrating personal and shared spaces. In: Proceedings of the CHI '04. doi10.1145/985692.985698
35. Vatavu RD (2013) Designing Gestural Interfaces for the Interactive TV. In: Proceedings of the EuroITV '13. doi:10.1145/2465958.2465981
36. Wendemuth A, Biundo S (2012) A companion technology for cognitive technical systems. In: Esposito A, Esposito AM, Vinciarelli A, Hoffmann R, Müller VC (eds) Cognitive behavioural systems. LNCS, vol 7403. Springer, New York, pp 89–103. doi:10.1007/978-3-642-34584-5_7
37. Winkler C, Löchtefeld M, Dobbelstein D, Krüger A, Rukzio E (2014) SurfacePhone: a mobile projection device for single- and multiuser everywhere tabletop interaction. In: Proceedings of the CHI '14. doi:10.1145/2556288.2557075

**Jan Gugenheimer** holds a M.Sc. in Media Informatics (2014) from Ulm University. He joined the HCI group at the University of Ulm in March 2014. Before, he was a research associate at the ETH Zurich in the Auto-ID Labs. During his studies he worked at Daimler AG, IBM Research and Development Germany and Mercedes Benz Research and Development North America.

**Frank Honold** graduated from the Institute of Media Informatics, Ulm University. In 2009 he joined the institute as Research Associate and member of the scientific staff. He is doing his Ph.D. in the SFB/TRR 62 "Companion-Technology for Cognitive Technical Systems", which is funded by the German Research Foundation (DFG). His thematic focus is on the topics of multimodal fission and modality arbitration in the domain of multimodal systems

**Dennis Wolf** holds a B.Sc. in Computer Science from Ulm University. Since 2015, he is a Research Associate at the Institute of Media Informatics, Ulm University. His research interests include human-computer interaction and camera-projector systems.

**Felix Schüssel** studied Media Informatics at Ulm University. After graduation, he worked for the Fraunhofer Institute for Experimental Software Engineering (IESE) at Daimler AG, Ulm. Since 2009 he is a Ph.D. student and Research Associate at the Institute for Media Informatics at the University of Ulm in the SFB/TRR 62 "Companion-Technology for Cognitive Technical Systems" which is funded by the German Research Foundation (DFG). The thesis focuses on the fusion of user inputs with uncertainty and the abstract modeling of interactions.

**Julian Seifert** holds a Ph.D. in Computer Science from Ulm University. Between 2012 and 2014 he was a Research Associate at the Institute of Media Informatics, at Ulm University. His research interests include human-computer interaction and tabletop interaction.

**Michael Weber** holds a Ph.D. in computer science from University of Kaiserslautern. After some years in industry he joined Ulm University as a professor for computer science and was appointed director of the Institute of Media Informatics in 2000. His current research interests include mobile and ubiquitous computing systems and human-computer interaction.

**Enrico Rukzio** heads the Human-Computer-Interaction group at Ulm University. Prior to coming to Ulm he was an assistant professor at the University of Duisburg-Essen (2010 till 2012). He worked as a lecturer at Lancaster University from 2006 to 2013. He received his Ph.D. from University of Munich in 2007.