Interacting with ubiquitous media

A research matrix

Abstract

How do you structure your HCI research programme when you are heading your own group? This article presents the Media Computing Group at RWTH Aachen University and its goal to explore the future of collaborative, ubiquitous interaction with audiovisual media. It explains how this initial research vision has led to work in a *research matrix*: contributions on different levels, from HCI theory and algorithms to toolkits, testbeds, and design patterns, intersect with three topical research directions – interaction with audiovisual, ubiquitous, and collaborative systems. It also introduces some of our external collaborations, in particular the Excellence Initiative, Germany's most fundamental change in government research funding to date, which supports RWTH and our group.

Introduction

Until recently, the German HCI research landscape had been relatively sparsely populated. Virtually no research groups contributed consistently to top international conferences such as CHI. This paper aims, on the one hand, to explain how our Media Computing Group, established in 2003, went from an initial broad research vision to creating successful contributions in many fields of HCI, from theories and algorithms, to toolkits and testbeds, to ways to capture lessons learned as design patterns. In this, we hope both to interest partners around the world to get in contact with us to exchange ideas, and to inspire new labs defining and implementing their own research programme.

On the other hand, it is productive external collaborations and well-aimed funding that really make research groups thrive. Therefore, we will also present our links to other partners, and discuss the Excellence Initiative, a recent move by the German government funding body DFG (the German National Science Foundation) to provide extra funding to Germany's top universities and research groups.

Initial research vision: New interactions with ubiquitous audiovisual media

We started out with the conviction that we were most interested in *qualitatively new* interactions with technology (such as conducting an electronic orchestra), rather than *incremental improvements* to existing paradigms and techniques (such as speeding up dropdown menu selection). This approach provides big benefits: it helps to create systems that are 'first of their kind', breaking new ground in how people can interact with technology. However, it also carries specific problems – evaluating a fundamentally new interaction technique quantitatively is hard when there are no competitors to compare it to.

Our research vision joined the domains of audiovisual media and ubiquitous computing. On the one hand, we wanted to provide new ways to interact with, for example, audio or video streams. On the other hand, we wanted to move beyond the existing desktop paradigm and into the world of post-desktop, mobile, ubiquitous, wearable interfaces and ambient computing environments and interaction techniques.

The *Aachen Media Space* initiative turned out to be a great melting pot for these two ingredients: This room prototype provides an environment for scenarios in which people collaborate. They work with text but also multimedia digital material. Post-desktop interaction techniques make sure that their flow of collaborative activity is not interrupted.

From visions to contributions

Many of our concrete research projects, including those leading to PhD theses, have a similar overall structure:

1. **Vision** An initial research vision suggests a fundamentally new way for people to interact with technology. Examples include conducting an orchestra, casting spells by gesturing with a cell phone, or navigating a video by dragging objects in it to get to the moment when they are at that location. Sometimes these research visions are suggested by project opportunities with external partners; in other cases they are the result of internal brainstorming and discussions.
2. **Theories & Algorithms** From that HCI-oriented vision of a particular user experience, technological problems quickly begin to emerge that require hard thinking in computer science. For example, conducting an actual orchestra recording requires time-stretching audio and video in real time at high quality; cellphone spell-casting requires robust gesture recognition on a device with little processing power; and dragging objects in a video to navigate requires computing forward and backward trajectories of every pixel in every frame. Also, after one or two research projects with a similar goal, we have often uncovered more general theories to describe our approach and model the problem or solution design spaces.
3. **Toolkits** To make our results more accessible and reusable, we also found ourselves bundling our technology in software toolkits and frameworks to help ourselves and other similar research projects.
4. **Testbeds** Solving the above problems leads to contributions on the technical side of HCI, and often to presentable prototypes of the idea at work. However, it has become our tradition to try to 'go the extra mile' and apply these techniques to create an actual real-world system – often an interactive exhibit for a public space such as a tech museum or similar exhibition centre. That way, there is only one customer to satisfy, but still a large number of people get to experience the new interaction over time as the system is on exhibition

or in use. An initial conducting system created by the author for the House of Music in Vienna, for example, has made its way into the top recommendations in *Lonely Planet* and virtually all other current Vienna travel guidebooks. This nicely grounds our work and provides closure to our initial vision of a new interaction.

5 **Patterns** We sometimes try to capture what we learned in our projects not just in scientific papers and testbeds, but as concrete design guidance. HCI Design Patterns [1] have been a suitable format to do this in a standard, accessible, and reusable way.

The remainder of this paper looks at the three major research directions introduced earlier: interaction with audiovisual media, ubiquitous interaction, and collaborative environments. For each of these three directions, we explain the above steps in more detail, and what projects and contributions emerged at each level.

Interaction with audiovisual media

In this research direction, we created several contributions on the level of ***algorithms and theories***, to better understand our research problems and ways to solve them. One good example is PhaVoRIT [2], an algorithm that time-stretches audio recordings without pitch shifts or major audio artifacts such as reverberation or distortion, with real-time response to user input with a latency of a few milliseconds, and without requiring any lengthy precomputing. Over several iterations, it greatly changed and improved the user experience of our interactive conducting exhibits. We also studied conducting behaviour of professionals and laypeople, and found clear differences in how these two groups thought about and interacted with a musical system while trying to conduct a recording [4]. At the same time, we noticed that writing software that time-stretched multimedia data was unnecessarily difficult because of a missing abstraction: we introduced *Semantic Time*, a concept that lets programmers think in beats of a music piece rather than in sample numbers or milliseconds of a recording. Another example for an algorithmic contribution in this research direction is the DRAGON [5] algorithm that computes optical flow fields from a given video recording, to enable the aforementioned dragging of objects to navigate a video timeline.

On the ***toolkit*** level, we continued to create the *Semantic Time Framework*, a new multimedia API that used the above concept of semantic time to create time-stretching applications much more easily.

We created ***testbeds*** for most of our research initiatives in this field, not least because interaction with audio and video lends itself perfectly to presentation in demonstrators. Our work in conducting systems was flanked by a series of exhibits; after the initial Vienna installation (Fig. 1), the Boston Children's Museum (2003) and the Betty Brinn Children's Museum Milwaukee (2005) received newer versions of interactive conducting systems. Other testbeds included REXband, an interactive music exhibit exploring computer-aided musical improvisation, and the open-source audio scrubbing tool DiMaß.

Figure 1 Personal Orchestra, an interactive conducting exhibit in the House of Music Vienna

Our testbed designs were guided by the author's set of HCI Design ***Patterns*** for interactive exhibits [1]. This initial collection proved quite valuable for our projects, as it helped to quickly communicate the most important design guidelines and design values behind our work. Sharing these effectively both with new members of our team and potential clients was a big bonus. At the same time, the guidelines evolved, as experience from subsequent projects strengthened some patterns in the language and weakened others.

Ubiquitous interaction

In this second research direction, our group had decided to look at interfaces beyond the desktop, for mobile and also wearable computing. Many of our initial experiments in this area, however, were hindered by the fact that, in our evaluations, we were trying to compare new, prototype-level interaction techniques (such as using phonecams as tracking devices to move cursors on large displays) with established, industry-strength interface devices such as mice or joysticks. One of the ***theories*** we gradually developed to address this was the idea of *selexels* to effectively measure the potential expressiveness of prototype mobile interaction techniques. In short, selexels are pixels in 'selection space', the movement space in which users make their input to a graphical display; this model extended Card's basic model of display and motor space for input devices, and allowed us to predict more accurately the future performance of our prototype input devices. The same work also led to a redefinition of Card's classic *design space of input devices*, which we extended to devices with multiple modalities. Finally, triggered by a research vision of a wearable device to help people learn snowboarding, we are developing a haptic language for full-body haptic feedback to communicate complex body stance corrections to someone practising sports or doing other physical activities (such as packing crates), or to those requiring therapeutic help with their body stance.

On the ***toolkit*** level, we discovered that when asked to create a new, post-desktop user interface, students would spend far too much time working on the low-level details of writing serial device drivers, soldering interfaces, or running cables through our lab. We therefore created *iStuff mobile* [3], a user interface framework that lets UI designers and researchers get from an idea for a new, post-desktop interaction (e.g., tilting a mobile phone to scroll through its address book) to a working prototype very easily. iStuff mobile was built on top of our earlier *iStuff* toolkit that enabled arbitrary input and output devices to talk to a ubicomp environment in a simple, standardised, and seamless way. iStuff mobile added off-the-shelf mobile phones as extremely versatile and 'hackable' input and output devices. It also provided a graphical signal-flow-diagram style editor to connect mobile phones to inputs or effects in the ubicomp environment.

One ***testbed*** for this research is *REXplorer* (Fig. 2), one of the first continuous, publicly accessible pervasive games. REXplorer let tourists explore the historical German city of Regensburg with a wand-like device; performing spell-casting gestures with the wand in specific locations let the tourist listen, through the device, to historical figures, who would provide historical information and little quests to continue exploring. Struggling with the balance of high-tech location sensing and gesture recognition versus the need for simplicity in an edutainment setting with casual visitors was an extremely educational, and sometimes sobering, experience. Another example is our *Wearable Snowboard Instructor* prototype to test our sensor network, body model, and haptic feedback languages.

Figure 2 Playing the REXplorer pervasive city game

Collaborative environments

This third research direction, finally, brings the first two together: spaces in which several people can work together, using post-desktop interfaces to access multimedia data streams, to achieve a common goal such as editing a newscast.

On the level of ***theories and algorithms***, we started out with an early *AudioSpace* project that provided audio output (access to the eight speakers around the Aachen Media Space) as a wireless service to any computer in the room. More recently, we have looked into multi-touch tabletop interaction, and developed initial theories on *Social Proxemics* in collaborative game settings, as well as a good foundation of technical knowledge on how to create multi-touch surfaces.

Figure 3 The Aachen Media Space, with its flexible furniture, AudioSpace speakers and Media Boards

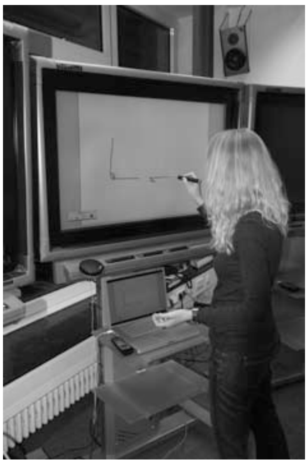

Figure 4 A Media Board, a mobile interactive group display in the Aachen Media Space

Figure 5 VICON tracking system in the Aachen Media Space to study precise interactions with mobile devices

The *iStuff mobile* toolkit [3], because it links mobile phones and sensors to an augmented environment, is one of our *toolkits* for this research direction.

The prime ***testbed*** for this research direction is clearly our *Aachen Media Space* (Fig. 3). Its highly flexible architecture, from furniture, to *Media Board* displays (Fig. 4), to collaborative middleware, has helped us use the room not only to evaluate (Fig. 5) and present our research prototypes, but also for our everyday meetings. As other testbeds, we have created exhibits that have allowed us to study social proxemics and the use of multi-touch technology, for display at the *Industrion* tech museum in the Netherlands. Our initial *Table Lemmings* game let multiple users play the classic *Lemmings* game, but by using their own hands as blocks, bridges, or obstacles – an interesting combination of the advantages of classic board games (many can act at the same time) with computer games (dynamically changing environment and actors).

On the ***patterns*** level for this third area, the author is currently completing a book chapter on Media Space design patterns that will hopefully serve to disseminate what we learned through our various projects.

Collaborations and funding

Our group was established by the B-IT foundation, and has worked closely with chairs in architecture and other disciplines in teaching and research, both at RWTH and internationally. RUFAE, for example, is an international network of research groups interested in augmented environments; partners include Electricité de Paris, Stanford, Darmstadt, CMU, Moscow, KTH, and RWTH.

The German Excellence Initiative has turned around the German government's research funding strategy in 2005, directing strong financial support to nine *Elite Universities* selected in a multiyear process based on both current standing and potential for innovation, to let them become more visible in today's global research landscape. Of these nine elite universities, RWTH Aachen University was awarded the highest amount of funding. In particular, our group is part of the UMIC excellence cluster at RWTH on high-speed mobile networks, contributing HCI expertise to the cluster, and of HumTec, a new interdisciplinary research centre at RWTH between the social sciences and engineering and an incubator for international young researchers.

Summary and outlook

Table 1 lists our main projects along our research directions and contribution levels. Our vision-based approach, with a blessing of great students, has worked out well so far: between 2003 and 2008, the group has been the most successful in Germany in terms of overall archival CHI publications [6], and our first PhD graduates have continued on to exciting careers at Nokia's Palo Alto Research Center and Apple's core development team in Cupertino. We are now pursuing some new directions with *Organic Interfaces*, and hope to continue our track record in HCI@Aachen. For more information about our group and projects, simply visit http://hci.rwth-aachen.de.

Vision: New Interactions with Ubiquitous Media			
Audiovisual Aspects	Ubiquitous Aspects	Collaborative Aspects	Requires work on:
PhaVoRIT Semantic Time DRAGON	Selexels Design Space of Mobile Input Haptic Languages	Social Proxemics Audio Space	Theories & Algorithms
Semantic Time Framework	iStuff Mobile	iStuff	Toolkits
You're The Conductor Maestro REXband DiMaß	REXplorer Associative PDA Snowboard Assistant	Aachen Media Space HumTec Spaces Table Lemmings The Logistics Table	Testbeds
Design Patterns for Interactive Exhibits		Media Space Design Patterns	Patterns

Table 1 Summary matrix of our 3 research directions and 4 contribution levels, with project names in each cell. Please see our home page URL at the end of this document for more details and other projects.

Acknowledgements

The work presented in this overview was funded in part by the German B-IT Foundation, and by the German Government through its UMIC Excellence Cluster for Ultra-High Speed Mobile Information and Communication at RWTH Aachen University. An earlier version of this article was presented as a non-archival research landscape overview at CHI 2008.

References

1 Borchers, J. (2001). *A Pattern Approach to Interaction Design*. Wiley & Sons.
2 Karrer, T., E. Lee and J. Borchers. (2006). PhaVoRIT: A Phase Vocoder for Real-Time Interactive Time-Stretching. *Proc. ICMC*, 708–715, ICMA.
3 Ballagas, R., F. Memon, R. Reiners and J. Borchers (2007). iStuff Mobile: Rapidly Prototyping New Mobile Phone Interfaces for Ubiquitous Computing. *Proc. CHI 2007*, 1107–1116. ACM.
4 Lee, E., M. Wolf and J. Borchers (2005). Improving orchestral conducting systems in public spaces: examining the temporal characteristics and conceptual models of conducting gestures. *Proc. CHI 2005*, 731–740. ACM.
5 Karrer, T., M. Weiss, E. Lee and J. Borchers (2008). DRAGON: A Direct Manipulation Interface for Frame-Accurate In-Scene Video Navigation. *Proc. CHI 2008*. ACM.
6 http://hci.rwth-aachen.de/chi-ranking