



UNIVERSITY OF GOTHENBURG

Pushing Towards Embodied Interactions

Ludwig Weschke
Gustav Börman

Bachelor of Software Engineering and Management Thesis

Report No. 2010:031
ISSN: 1651-4769

Pushing Towards Embodied Interactions

Ludwig Weschke
l.weschke@gmail.com

Gustav Börman
gustav.borman@gmail.com

IT University of Gothenburg, Software Engineering and Management. Gothenburg, Sweden

CHALMERS



UNIVERSITY OF GOTHENBURG

Abstract— New interaction techniques are coming more than ever before, but are they meaningful? Is there a purpose behind these novel technologies or is it the technology itself that is emphasized? In this paper we have taken the approach of embodied interaction and evaluated current interaction modalities in order to design a next generation input device focusing on the user instead of the technology behind it. We have established the relevance of embodied interaction and our result brings interesting insights into design considerations when designing novel interaction mediums. We have also discussed the technical considerations around our proposed interaction technique.

Keywords: Embodied Interaction, HCI, Interaction Devices, Ubiquitous Computing

1 Introduction

Today the current paradigm of human computer interaction (HCI) consists mostly of standard computer interaction techniques such as keyboards and mouse devices that has been around for the last 20 years (Dourish, 2004). However, in recent research new novel approaches to bring the interaction between humans and computers closer together through more natural and embodied interaction solutions implies a paradigm shift in the way we interact with the digital environment around us (Jaimes and Sebe, 2007; Kurkovsky, 2007). Interaction with a system is today no longer seen as interaction with standard computer systems alone, e.g. desktop environments, but also seen as the interaction with everyday objects such as interactive- whiteboards (Klemmer, 2005), -furniture, -spoons, -cups (Loke, 2006) or even digital jewellery (McCarthy et al., 2006).

Due to advances in hardware technologies during the last decades, the fields like ubiquitous computing, pervasive computing and context-awareness have recently started to flourish. Computers, sensors and actuators are constantly shrinking and dropping in price making actualization of these areas feasible in practice (Dourish, 2004; Satyanarayanan et al., 2001). Interaction methods such as haptic, voice, gaze, wearable devices and gestures are due to the progress in the hardware market today possible to implement using more or less standard computer peripherals. This plays an important role in reducing the borders between humans and

computers in order to make interactive systems embodied in the real-world environment. The question is, what is needed to design new interaction mediums to make it more than a technology hype—to make it part of peoples lives as something meaningful they can interact with?

The Wiimote used in the Nintendo Wii™¹ gaming console has proved to reduced the borders between participants and the games they are playing making the system more enjoyable and understandable for a broader audience (Bowman et al., 2008), however, it is generally not preferred by the hard core gamers (LaViola Jr, 2008). Alternate interaction techniques for personal computers have been proposed, however these solutions are more or less intrusive on the users experience as they e.g. have to be equipped with colored markers placed on their hands (Fredriksson et al., 2008). Wearable computers used for augmented reality are often expensive and intrusive, its focus lie mostly on technological aspects and its scope of use is often fuzzy (Kolsch et al., 2004; Plouznikoff et al., 2007).

We have chosen to take the approach of embodied interaction (Dourish, 2001) to design and start the development of a 3D input device aimed at computer desktop interaction, namely the AiRL3D Project. We argue for taking this approach into consideration when designing the next generation input devices as it puts the user in focus of the design instead of the technology behind it. It combines both tangible- and social computing thus focus on peoples physical skills as they are as well as their social understandings.

“ By embodiment, I do not mean simply physical reality, although that is often one way in which it appears. Embodiment, instead, denotes a form of participative status. Embodiment is about the fact that things are embedded in the world, and the ways in which their reality depends on being embedded. ” - Paul Dourish 2001 p.18

This paper and prototype will make use of a set of design principles (Dourish, 2001) together with related HCI techniques in order to incorporate embodied interaction into computer systems. Previous research will be evaluated and we will identify what is good, what is bad and what is missing in this area and compare the proposed interaction solution

¹<http://www.nintendo.com/wii>

towards the previous findings and discuss what needs to be considered when designing novel interaction mediums in general. The contributions of this study thus to provide insights into design considerations, technical aspects of the prototype and state-of-research within embodied interaction.

The aim is not to build state-of-the-art 3D-interaction application but rather show that it is possible to realize embodied interaction with the use of general design principles to develop a non-intrusive prototype out of low-cost hardware. It is not meant to be seen as a replacement to current technology, since both the keyboard and the mouse have their specific purpose, merely as an extension to the available interaction space.

The rest of this paper is structured as follows. Section 2 will go through research related to embodied interaction and its related areas, interaction modalities and technical aspects. Section 3 will present the research design and methods used in this paper followed by section 4 which presents the data collected in the literature review and from the interviews. Section 5 will analyze these findings followed by technical details of the prototype in section 6. We conclude this thesis with our results and propose further research in section 7.

2 Related Research

Embodied Interaction

The area which is of particular interest in this paper is embodied interaction which is a richer form of interaction with computer systems. Embodied interaction is focused on how humans create and communicate meaning with a system in the realms of both physical and social reality i.e. in contrast to the traditional interaction using mouse devices and keyboards controlling graphical user interfaces consisting of windows, icons and menus (WIMP) (Dourish, 2001). Embodied interaction is a common ground for social- and tangible computing (Dourish, 2001) which in itself is part of the broader area of context-awareness which is characterized by Dey et al. (2001) as the interaction between humans, application and the surrounding environment. Dey also acknowledge the challenge that the notion of context is ill defined within the area of HCI. The research area of context-awareness have evolved and is now consisting of several different fields such as pervasive-, ubiquitous-, and mobile computing. However, embodied interaction spans over several of these areas combining the foundations to create the common notion *embodiment*. This notion promotes the creation of more natural interfaces by incorporating social and/or physical skills in order to create a more meaningful interaction space, allowing participants to engage in systems in a more natural way (Dourish, 2001) and this is why we choose the forum of embodied interaction rather than covering the aspect of context-awareness. Social computing incorporates the area of sociology into interactive computer systems design whereas tangible computing takes use of real world objects into interactive systems enabling humans to interact with several senses (Ishii and Ullmer, 1997).

Dourish (2001) presents six design principles to be considered when designing interactive systems:

1. *Computation is a medium*
 2. *Meaning arises on multiple levels*
 3. *Users, not designers, create and communicate meaning*
 4. *Users, not designers, manage coupling*
 5. *Embodied technologies participate in the world they represent*
 6. *Embodied interaction turns action into meaning*
- Paul Dourish 2001 p.163

Dourish means that these principles should not be used as rules but rather describe the features of embodied interaction and design considerations to keep in mind for a designer implementing interactive systems both at a theoretical- and design level. We intend to follow these principles as a tool to guide our attention during the design.

Embodied interaction, or embodiment, is a commonly occurring concept used in previously presented papers. An important note to take is that embodiment is about how we interact with technology and not a technology itself (Dourish, 2001) thus as we go on presenting related design considerations and approaches, technologies such as tangible user interfaces and context-aware applications are technologies related to embodiment as they withhold more or less embodiment.

SignalPlay is a prototype system designed in order to get an understanding for how users explore and understand the space that are occupied with ubiquitous computing technologies. It builds on the concept of embodied interaction and focus on how people act together in space rather than focus on interaction (Williams et al., 2005). They reach conclusions that peoples understanding of interacting with ubiquitous computing applications was in terms of individual actions and components rather than an understanding of the whole system. Designers of systems should emphasize individual features, not as elements of the system but rather as objects who have their own meaning and interpretation. Second they concluded that current design approaches considers mostly the sequential organization of interaction but lacks to understand the temporal aspect such as pace and rhythm when designing systems. Design models should address space as something explicitly constructed and managed as people come to understand infrastructures through ubiquitous technologies. The concept of *space* can mean multiple things, not only the digital interaction space but also the physical space. It can mean the space a two- or three dimensional computer desktop occupy, the space in which sounds travel through and how we interpret it as well as something people share and interact within. Space comes hand in hand with *place* whereas space is concerned with the physical environment and place is where social understandings unfold in that space (Dourish, 2001; Williams et al., 2005).

Empirical work has been done on comparing traditional graphical user interfaces to tangible user interfaces where the effects on user's spatial cognition have been studied when modeling 3D-blocks. It was shown that tangible user interfaces gave a richer sensory experience thus off-loading their cognition. In addition, the correlation between perceptual and immersive gestures eased the creativeness of people. It was concluded that off-loading cognition and immersion should be considered when designing novel user interfaces and that it should be either through multimodal user in-

terfaces or spatial systems with augmented reality (Kim and Maher, 2008).

Plouznikoff et al. (2007) presents a novel interaction approach for meaningful information manipulation through virtually embodied software processes in a users environment with wearable computing. Here the design and implementation, to bring embodiment and meaningful interactions to a system, is described in a more direct and technical manner as opposed to the design considerations previously mentioned. The result from the research showed that this interaction techniques where feasible at least in a controlled environment.

Interaction Modalities

There are several areas where gestures, embodied interaction and spatially convenient applications has found its way to. LaViola Jr (2008) presents the evolution of gaming and draws conclusions that gaming consoles such as the Wii™ gaming console might not provide the precision and expressive powers as others might but that its more appealing to the casual gamer. Its hard to remember complicated button sequences if your not a hardcore gamer, thus its more appealing for a casual gamer to play more naturally using gestures. He also explains that 3D interaction is no longer a gimmick since hardware have become cheaper and faster making it more feasible to implement. Wingrave et al. (2009) discusses the Wiimote's capabilities and shows that spatially convenient hardware indeed is becoming mainstream. However as applications are designed with different input hardware by industry, academia etc. it is going to be hard to truly introduce it on a large scale since no common input hardware platform exists.

Spatially convenient interaction techniques have also found its way into the automotive industry where the current interaction technologies with infotainment systems can distract the drivers attention. Previous research and surveys have looked at interaction techniques such as gestures and speech in in-vehicle systems (Alpern and Minardo, 2003; Bach et al., 2008; Pickering et al., 2007) where they found that using gestures and similar interaction techniques for controlling in-vehicle secondary task is viable, reduces interaction complexity and that it has much potential in the future accordingly. An overview of potential problems regarding 3D user interfaces and interaction in cars is given by Tönnis et al. (2006) where they find that keeping it simple and provide multimodal interaction techniques have high potential when interacting with in-vehicle systems. Althoff et al. (2005) does experimental results on a head- and hand gesture video-based system implemented on a BMW limousine and shows that in-vehicle gestures are both effective and intuitive since a driver can focus on its primary driving-tasks.

Technical

At this point we have established the relevance of an embodied interaction perspective to related areas, and a number of existing interaction modalities related to the prototype we are developing and assessing in this paper. To accomplish this we also need to understand the technical aspects in developing this prototype and will move on to this now.

The main technique needed by our prototype is stereo vision which has a variety of different approaches such as correlation-based stereo vision (Hirschmüller, 2001; Hirschmüller et al., 2002), global, semi-global (Hirschmüller, 2006), local and local-global (Zhao et al., 2006). The different techniques listed here have more or less been successfully implemented in real-time or near real-time systems which is one of the prerequisites for the design of our implementation since the interaction medium needs to respond instantly to make the users feel like they are directly interacting with the system to keep their attention focused. Techniques for detecting fingertips and gestures in real-time have been proposed by several authors such as (Malik, 2003; Manresa et al., 2000).

3 Research Design

The research conducted in this paper have been addressed with an extensive literature review, lead user methodology (Hippel, 1988) together with the repertory grid interview technique (Tan and Hunter, 2002) while using a interpretive research approach (Walsham, 2006) to analyze the data.

Literature Review

When the topic was identified, an extensive literature review was conducted for the purpose of evaluating the current state of research within embodied interaction, related areas and interaction techniques. Most of the previous research presented here was gathered from the top journals *Human-Computer Interaction (HCI)*², *Communications of the ACM (CACM)*³, *IEEE Transactions on Software Engineering*⁴, *IEEE Computer Graphics and Applications*⁵ and *IEEE Pervasive Computing*⁶ which are seen as some of the leading journals within HCI and the areas of context awareness, pervasive- and ubiquitous computing among others. These sources have also been complemented with articles found through keyword search from other scientific databases such as *IEEE Xplorer*⁷ and *Science Direct*⁸.

Lead User Methodology

As we wanted to investigate if this novel interaction technology was feasible, we decided to make use of lead user methodology. The reason for this, as described by Hippel (1986, 1988), is that lead users can foresee the future of products in the market. When new products reach the markets e.g. a new phone, its generally not so much different than its predecessor and "normal" users can participate to provide input to those sort of products as they are familiar with them. However, regarding novel products, "normal" users might feel alienated by them, hence a lead user, or expert, was a natural choice to provide input to our prototype.

²<http://hci-journal.com/>

³<http://cacm.acm.org/>

⁴<http://www.computer.org/portal/web/tse/>

⁵<http://www.computer.org/portal/web/cga/home>

⁶<http://www.computer.org/portal/web/pervasive/home>

⁷<http://ieeexplore.ieee.org>

⁸<http://www.sciencedirect.com/>

Hippel defines lead users in two aspects: “1. *Lead users face needs that will be general in a marketplace, but they face them months or years before the bulk of that marketplace encounters them, and 2. Lead users are positioned to benefit significantly by obtaining a solution to those needs*” (Hippel, 1988).

We adapted a four-step process described by Hippel (1986, 1988) to best utilize the lead users. The process consists of the following steps: 1) *Identifying an important trend*, 2) *Identifying Lead users*, 3) *Analyzing Lead User insights* and 4) *Testing Product Concept Perceptions and Preferences*. The fourth step was ignored in this article as it was out of the scope. In the first step, we identified the current state of research within alternate interaction techniques and interaction trends through the literature review. In the second step we identified lead users with higher technical knowledge and sufficient insight in to the domain of our prototype. The lead users were then interviewed with the repertory grid technique.

The Repertory Grid Technique

The repertory grid technique is a cognitive mapping technique (Tan and Hunter, 2002). The technique evaluates peoples views on objects and events in terms of bipolar personal constructs (Kelly, 1955), e.g. a cup of coffee may be seen as hot while a glass of soda is seen as cold, the construct would then be ‘hot-cold’. The repertory grid is composed of elements, constructs and links. The elements is the objects that the interviewees will evaluate within a specific domain (coffee, soda), the constructs are the bipolar personal constructs which they label the objects with (hot-cold) and the links relates and shows the interpretations an interviewee have between a construct and an element (Tan and Hunter, 2002).

There are several different ways on how to design and execute the repertory grid. The general steps it consists of are 1) element selection, 2) construct elicitation, 3) linking elements to constructs (Tan and Hunter, 2002). Below is a description of the setup used in this paper deriving from Tan and Hunter (2002).

1. *Element selection*, in this step we selected a number of elements related to the proposed prototype in this paper.
2. *Construct elicitation*, here we chose to use the triadic sort method where three elements were selected from the whole set of elements. The interviewee were then asked to point out one of the three elements that was different in some aspect from the other two to then put bipolar personal construct labels on them. This step was repeated with different triads until a required amount of constructs was obtained from an interviewee.
3. *Linking the elements to constructs*, In the final step to link the elements with the constructs we let the interviewee rate the elements on a scale from one to five on each of the constructs obtained.

When the steps above were performed with all interviewees, the data was analyzed by transforming it using the FOCUS cluster analysis algorithm (Shaw and Thomas, 1978).

In order to do this we used the tool Web Grid 5⁹ in which the data was inserted into and which automatically performed the cluster analysis.

Interpretative Research

The data that have been analyzed consists of data found in the literature review, the repertory grid results, the proposed prototype and related interaction devices. From this, themes have been identified and interpreted around the design principles proposed by Dourish (2001), this have been done on a continuous basis. The interaction techniques or proposed solutions presented in the current research are also evaluated based on attributes such as degrees-of-freedom (DOF), intrusiveness, naturalness, accurateness, mobility and feedback.

4 Reflections of Data Collected

On Interaction Devices

Several different approaches to extend the interaction space from pure mouse and keyboard or touchscreen interfaces have been invented such as voice recognition, gaze detection, facial recognition, motion detection and tangible interfaces. Of these modalities the most commonly known technique is interaction via hand gestures or facial recognition which is done by processing the images captured by various input devices such as web- or a infrared cameras. Other techniques such as haptic input devices like force feedback game controls, -gloves and -pens have also been proposed to decrease the boarder between the user and the system by giving feedback from the virtual environment back to the user. Some less known devices have been also proposed such as Blowable User Interfaces (Patel and Abowd, 2007) or a device for improving human-chicken interaction through the internet (Lee et al., 2006)

Frohlich et al. (2006) presents an overview of their solutions for interaction in three dimensional space in order to extend the two dimensional available standard input devices such as the mouse. The GlobeFish is a 6-DOF input device consisting of an elastically suspended trackball allowing rotation around all three axes as well as spatial movement by applying force to the trackball. The same authors proposed a hand-held device for spatial 3D interaction in virtual reality environments called Two-4-Six consisting of a touchpad for rotating an elastic ring for viewpoint motion and an analog rocker sensor controlled by the users index finger for controlling depth movements (Bowman et al., 2008; Frohlich et al., 2006).

Holman et al. (2005) introduces a system called Paper-Windows for bridging the gap between the digital environment and the real world while working with documents based on paper enhanced with IR markers where the document is projected on the paper and easily changed by rubbing the

⁹<http://gigi.cpsc.ucalgary.ca:2000/>

paper against a LCD display. The interaction with the interactive paper is done by gestures captured by a motion detection system based on a marker placed on the users index finger. The technique allows for documents to be handled in a more natural way combining the properties of both worlds.

Plouznikoff et al. (2007) introduces a prototype where meaningful real-world actions in a mobile setting can take place with help of a wearable computer consisting of a head mounted display, a webcam and a vest containing the computational hardware. Their idea is that e.g. when you are walking on the street and an e-mail arrives or the weather report is given, that an avatar will be displayed in a head mounted display in front of you. These events can then be handled by picking up the avatar (accepting the event) or brushing it away (rejecting the event), other examples where also given such as stomping on the avatar (rejecting the event) and kicking it away (delaying the event). Several other researchers have created prototypes for AR by the use of wearable computation devices such as vests as in the previous example or by a backpack containing the computational platform as proposed in (Avery et al., 2005; Piekarski et al., 2004; Smith et al., 2005) and others.

Games like Guitar Hero¹⁰, Rock Band¹¹, flight simulators and racing games allows the user to be engaged in the games in a more realistic way by offering custom made controllers shaped like guitars, drums, basses, flight controls and steering wheels. Even though these controllers are a success they are quite expensive, especially if you look at the area they are applicable to, these kind of controllers tends to be tailor made for one specific game. These controllers takes gaming to a more realistic level, however, they are still only adaptations to the original controllers still interacting in two dimensions mapped to a three dimensional world causing the controllers to score low on the DOF scale.

In 2006, Nintendo released their gaming console Wii™ introducing a brand new way of interacting with the gaming environment. Instead of using a static input modality such as a standard hand-controller, mouse or keyboard, they extended the controller (Wiimote) to take advantage of the users hand movement taking the first step towards incorporating the users themselves into the actions of the game. Research that have been done on the new modality have shown to create a better understanding and involvement than previous ways of interacting with games making the console a success to a wider audience. (Bowman et al., 2008; LaViola Jr, 2008). LaViola Jr (2008) comes to the conclusion that even if the Wiimote allows for spatial coarse 3D interaction, few of the available games takes advantage of this due to developers lack of knowledge on how to exploit the opportunity and that the extra freedom it offers is usually seen as an afterthought. LaViola Jr (2008) also points out that the Wiimote initially is not able to function as a input device complying with the requirements for the 6-DOF, however, the coarse precession may be improved by the additional MotionPlus¹² accessory to the control.

Sony's upcoming motion controller Move¹³ resembles the Wii™ interaction method combining advanced motion sen-

sors and visual input via Playstation's Eye™¹⁴ camera promising to "Take core gaming to a new level or bring your whole family in to the adventure.."¹⁵. Microsoft's upcoming Project Natal™¹⁶ to the Xbox360™¹⁷ console offers similar approach to embodiment, tracking the users body movement by replacing the physical controller with the users own body movements. At the time of writing the information on Microsoft's Project Natal™ and Sony's Move™ is sparse and most speculative, however one can see a trend of growing emphasis on embodiment within this area.

Image Processing Techniques

Since our idea for the prototype was to create an input device based on natural hand interaction, extensive research was done in the area of computer vision and gesture based interaction via webcams and/or infrared sensors. Manresa et al. (2005) determines the location of the users hands by the use of skin segmentation, others such as Shan et al. (2007) make use of particle based tracking to determine areas of interest. Other widely used techniques for hand recognition include contour tracking as used by Bowden et al. (2004), background subtraction used by e.g. Coogan et al. (2006); von Hardenberg and Bérard (2001) or by colored markers to identify fingertips or other features based on histogram analysis (Fredriksson et al., 2008; Joslin et al., 2005; Keskin et al., 2003). Ogawara et al. (2003) presents a device for tracking hands by using infrared cameras, however, this approach was dismissed due to the need of specialized equipment or the need for hardware modification of a standard webcam to be able to capture the spectrum of infrared light. Agarwal et al. (2007) presents a solution to make any surface interactable like a tabletop by using two cameras mounted above the surface and the use of stereo vision to determine the location of the hand as well as the distance from the surface that the image is projected on. Schlattmann et al. (2007) are using similar solutions for detecting posture and/or hand gestures using stereo vision however most of the techniques mentioned above tracks gestures and/or 2D movement of only one point and all of them have their own benefits and drawbacks.

There are several approaches to detect hand properties such as fingertips, palm and wrist once the hand have been segmented from the rest of the scene. Principal component analysis (PCA) is one of the common techniques for comparing the eigenvalue of the input segment to a predefined set of reference images. Hidden Markov Model (HMM) or adaptations of the method is another common approach for detecting hand properties and gestures.

Repertory Grid Results

As described in the Research Design section, lead user interviews with the repertory grid technique were conducted. Four lead users were identified who had the required technical- and domain knowledge. The area from which

¹⁰<http://guitarhero.com/>

¹¹<http://www.rockband.com/>

¹²<http://www.nintendo.com/wii/console/accessories/wiimotionplus>

¹³<http://us.playstation.com/ps3/playstation-move/>

¹⁴<http://us.playstation.com/ps3/accessories/scph-98047.html>

¹⁵<http://us.playstation.com/ps3/playstation-move/>

¹⁶<http://www.xbox.com/en-us/live/projectnatal/>

¹⁷<http://www.xbox.com/en-US/>

the elements was selected from was *interaction technologies* ranging from traditional devices to unreleased devices including our own prototype. The elements that were chosen are listed in Table 1. The participants did have good knowledge about most of these elements beforehand and information was given before the interviews started in the cases they lacked sufficient knowledge. From these elements, five triad combinations were selected and shown to the participants as a basis for the construct elicitation. The participants were then given the task to rate their own constructs on a scale from one to five.

Elements	Description
BMW Gestures	In-vehicle system for controlling in-fo-tainment systems with predefined hand- and head gestures (Althoff et al., 2005).
Project Natal™ ¹⁸	Controller-Free gaming for the Xbox 360™ using your own body to control the games.
Nintendo Wii™ ¹⁹	Gaming console where players control the games with physical gestures in combination with button presses.
AiRL3D ²⁰	The proposed prototype in this paper.
iPad ²¹	A tablet computer with multi-touch touchscreen.
Mouse Devices	Standard mouse devices.
HandVu ²²	Wearable computer hand-tracking for mouse-based interaction or augmented reality interaction.

Table 1: Elements used in the repertory grid.

After the data was collected from all lead users, everything was inserted into Web Grid 5²³ where the FOCUS cluster algorithm generated an initial, unmodified view of the repertory grid data which is illustrated in Figure 1. What can be seen in the figure is that some of the constructs resembles each other and in some cases they are almost identical. The majority of the constructs address to what extent you are dependent to use something physical to interact with e.g. 'Tied - Independent', 'Natural - Mechanical' and 'Restricted - Free'. Followed by constructs that indicate how flexibly you can interact with these devices or interaction technologies e.g. 'All-round - Predefined', 'Freestyle - Predefined' and 'Invariable - Versatile'.

The next step was to look at the FOCUS view more closely in order to narrow it down. First constructs that were related less than 88% where sorted out followed by grouping of the constructs that resembled each other. This was done in order to get the most outstanding differences and to get a comprehensible set with unique constructs to interpret. This generated another FOCUS view shown in Figure 2.

Notes were also taken during the interviews, as the participants discussed why they wanted specific constructs, in

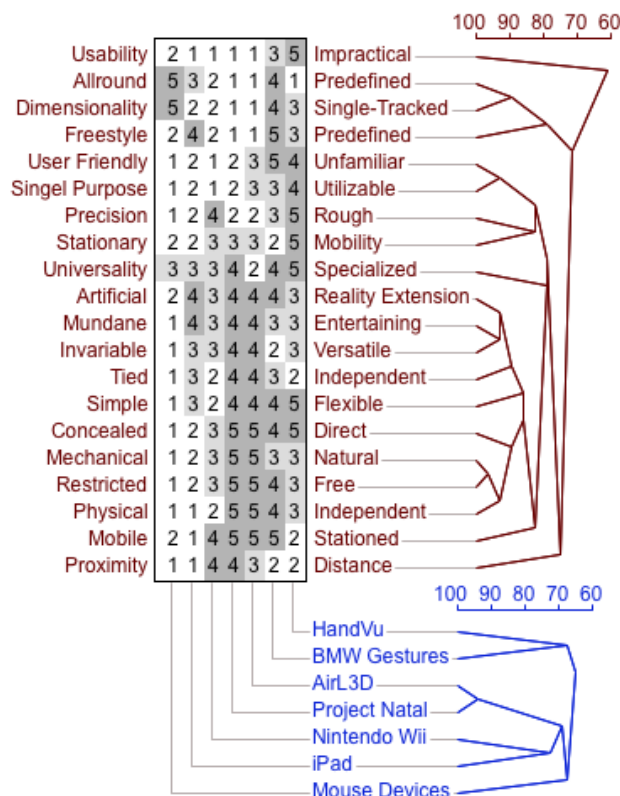


Figure 1: FOCUS, Domain: Interaction Technologies 20 Constructs, 7 Elements

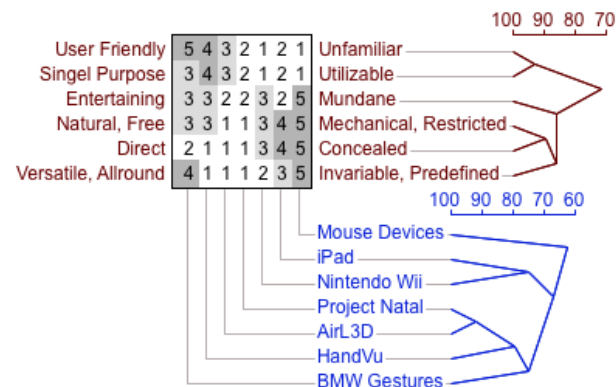


Figure 2: FOCUS, Domain: Interaction Technologies 6 Constructs, 7 Elements

order to clarify what these constructs means as they can be perceived differently depending on the context. Their reasonings are briefly described below and will be looked at more closely in the analysis section.

User Friendly - Unfamiliar Novel interaction techniques and less explored areas of use were predicted to be seen as alienating and unfamiliar as they can be perceived as complicated or overambitious. Interaction devices like the mouse were seen as more user friendly as they are and have been broadly established interaction techniques.

¹⁸<http://www.xbox.com/en-us/live/projectnatal/>

¹⁹<http://www.nintendo.com/wii/console>

²⁰<http://sourceforge.net/projects/airl3d/>

²¹<http://www.apple.com/ipad/>

²²<http://www.movesinstitute.org/kolsch/HandVu/HandVu.html>

²³<http://gigi.cpsc.ualgary.ca:2000/>

Single Purpose - Utilizable This construct describes to what extent these interaction techniques could be utilized in different areas and also within the current destined area of use.

Entertaining - Mundane Entertaining was here seen as fun and engaging whereas mundane was seen as standard or traditional interaction.

Mechanical, Restricted - Natural, Free This mainly means how much you need to be, or are, aware of the physical presence of the hardware you are interacting with. However, it is also related to the construct '*Versatile, Allround - Predefined, Invariable*' as devices using predefined gestures that have to be learnt was seen as mechanical behaviour.

Direct - Concealed This construct was selected based on how "hands-on" your experience would be. While e.g. a mouse device was seen as *concealed* as it acts as a tool between you and your computer, other interaction techniques were seen as having more *direct* control of your interactions.

Versatile, Allround - Predefined, Invariable This construct tells us about how flexibly you can interact e.g. *allround* and *versatile* indicates that there are several plausible ways in which artifacts can be acted upon as they unfold in real-time. The interaction technique was seen as *predefined* or *invariable* if there were limited ways or if you have to configure they ways in which you interact.

5 Analysis

In this section we will discuss design considerations regarding the proposed prototype to other interaction techniques together with the repertory grid constructs using Dourish (2001) six design principles.

Computation Is A Medium

When designing interactive systems, Dourish means that one should not focus on the technologies capabilities but rather how they are embedded in a set of practices. "*Meaning is conveyed not simply through digital encodings, but through the way that computation enlivens those encodings with semantic and effective power*" (Dourish, 2001). He means that computer systems should be seen as augmentations and amplifications of our own activities.

The design of the prototype strived for a more realistic interaction experience, but also realistic in the sense that it should not be considered as a replacement to current interaction devices as they still serve their purpose. So how do you put focus on practice instead of technology? And what does practice really mean? As an example, in the paper on the wearable computer presented by Plouznikoff et al. (2007), they do realize that the natural actions proposed could be socially unaccepted and that a more applicable area of use could be in-vehicle interaction to minimize distracting events, but should it then be necessary to be equipped with a wearable computer inside a car? Could not an augmented

reality windshield display (Kim and Dey, 2009) better serve that purpose since a head mounted display could be considered to be intrusive and superfluous to a user. Even if it would be highly applicable in an area, is it really understood by a broader audience that kicking away an avatar delays e.g. a notification message? What we want to illustrate here is that it is not enough to have a technologically advanced device which people can naturally interact with, when designing an interactive system one also needs to consider the area and context it is supposed to be used within.

This can be related to the construct '*User Friendly - Unfamiliar*' which seems to be rooted in what expectations a user have about an interaction device. The clearer the purpose and meaning is, the more user friendly and welcoming the interaction device or technique is. To enable people to see this, it should be important to steer the focus of the design early on by thinking of about what practice really means, it is not enough to look at practice or natural interaction as only e.g. a hand movement by itself. The AirL3D project was rated right in the middle of this construct, this was not all surprising considering that it is a novel interaction technique. Nevertheless, it was seen as more user friendly than the wearable computer (HandVu) and in-vehicle gestures (Althoff et al., 2005) which is seen as promising.

Meaning Arises On Multiple Levels

With this principle Dourish means that representations of artifacts in interactive systems also can be acted upon and as well as understood by the user that it is possible. As an example he describes the traditional WIMP desktop approach where icons represents files but that the embodiment of the interaction comes into play when the icon can be treated as an artifact as well. The new meanings that arises with the AiRL3D project can be illustrated in a comparison with a normal mouse device, which in this case are the artifacts. With a normal mouse, users can move the cursor, navigate through a system, move things around and resize windows etc. it has been around for a long time and therefore feels natural for most people, as recognized by Dourish, but it is still just a cursor that we move around. Guided by the principles of Dourish and the analysis of different input devices for 3D space we came to the conclusion that the design of our prototype should be non-intrusive, spatial, free and allow for a high DOF. The design should be able to simultaneously track multiple fingertips and posture of the hand without the use of intrusive markers. The idea with the prototype is to not restrict the user to predefined gestures but instead freely interact with the environment much like a multi-touch system and at its simplest acting as a single mouse cursor while at the same time allow for extending the interaction to multiple fingertips rendering it to a more iconic interaction than the graphical cursor.

This meaning on multiple levels leads to the construct '*Versatile, Allround - Predefined, Invariable*' where the proposed prototype together with HandVu and Project Natal was seen as the most *allround* and *versatile* devices. As seen by the ratings on all devices, this seem to indicate what degree of spatial freedom you have, however, this was not the case, BMW gestures was seen as *predefined* and *invariable* although it is more or less spatially free. What is also seen

in Figure 2, is that the ratings in this construct is similar to the ratings in the construct 'Mechanical, Restricted - Natural, Free' although they address slightly different aspects. This observation is important as it shows that even if you do not have to be aware of the physical hardware its more or less perceived as *predefined* or *invariable*.

Interestingly to also be seen in the ratings is that HandVu was seen as *versatile* and *allround* while it was seen as something in between the construct 'Mechanical, Restricted - Natural, Free' unlike the other elements that got fairly similar ratings on these constructs. Since the HandVu wearable computer did not have clear area of use, the participants reasonings behind seeing it as *allround* and *versatile* was that it had the potential of being used in many areas because of its mobility.

Users, not designers, create and communicate meaning; Users, not designers manage coupling

Traditionally a designers responsibilities concerns how an artifact looks like and functions together with other artifacts as well as how users will use them. However, artifacts are not always used as intended by the designer "*Meaning, and its coupling to the features and representations the system offers, emerge from actions of users, not designers*" (Dourish, 2001). This is the reason we do not want to implement e.g. pure gestures but instead allow for several plausible actions to be made upon artifacts in a desktop environment. As identified by Suchman (1987), that plans can not fully determine nor reconstruct an action in a specific situation and means that we need to consider how people make use of what their given in specific situations. It is not only the gestures per se we argue against, but rather that you should not consider and design for every possible use case scenario that might arise while using a system in order to support embodied interaction. This is not seen this as a flaw, rather as a feature of the system enabling users to potentially create the meaning they are striving to achieve.

Looking at the current state of HCI, one can see a clear trend of using statically defined hand gestures as an alternate interaction technique where standard input devices might not be applicable, but these might limit the user to explore the medium and create their own meaning as proposed by Dourish. Even though statically defined gestures might be applicable in certain areas such as within the automotive industry where the visual feedback needs to be kept to a minimum (Althoff et al., 2005), other areas where visual feedback is high we argue for a more flexible approach to allow for the exploration and creation of meaning within the system.

Embodied Technologies participate in the world they represent

The embodied perspective means that mind and body, representation and object are not separate entities but that they exist in the same world, "*..embodiment does not denote physical reality, but participative status.*" (Dourish, 2001). As an example, a number of years ago before the internet and the advanced search engines where available to the broader community and you wanted to find information about a cer-

tain topic, it felt natural to look for the answer in e.g. an encyclopaedia. However, today it is just as, or even more, natural to go to a computer to find the information you are want, it has become participative in the world we live in. The construct 'Direct - Concealed' points out that both our proposed prototype as well as Project Natal allows the user to become more participative and engaged in the virtual environment due to the direct interaction via your own actions instead of being conveyed through a combination of e.g. mechanical button presses. The intention with our prototype is that it should become just as participative as a mouse or a keyboard in the meaning of being seen as something established, but with a higher sense of freedom, allowing the user to adapt the interaction to his/her meaning instead of adapting themselves around the interaction medium.

As mentioned earlier, we strived to design for a realistic life-like interaction, but one should also be careful to overdo it as embodiment implies participative status not physical reality. The envisioned technique in PaperWindows Holman et al. (2005) is a good example that illustrates that there are benefits both in the physical and digital world and that there are advantages in combining them both. They take the physical properties of paper which can be archived and queried as efficient as in the documents you have on your computer. Even though their envisioned technique is not quite there yet due to the cost of for instance OLED displays, their vision to implement their proposed interaction technique on high-resolution paper-thin displays seems promising and the results from their experiments have got positive feedback from the users. The reason why we see their prototype as promising is that once the required hardware is out on a consumer wide market it can provide the freedom and flexibility to support the guidelines for embodied interaction where the user can create their own meaning of the device.

Even if the user should be the center of attention when designing interaction mediums, the technological properties should not be totally neglected as they do have advantageous properties the physical world does not have.

Embodied Interaction Turns Action Into Meaning

It all boils down to this final principle which could be argued as the main principle for supporting embodied interaction— "*Embodied interaction turns action into meaning as part of a larger system. Meaning, after all, does not reside in the system itself, but in the ways in which it is used.*" (Dourish, 2001). The question is, does the AiRL3D project support embodied interaction?

To this point we have looked at several different interaction devices and lessons have been learnt from all of them. The Nintendo Wii™ gaming console offers a higher degree of spatial interaction by tracking the users control via IR sensors aided by motion sensors within the controller itself. As mentioned earlier, the system has been a huge success in a wide audience without emphasis on the graphical experience in comparison to other systems, since focus have been put on an engaging interaction technique instead. However, the precision is very coarse and does not really allow for real interaction in three dimensional space due to the low DOF.

Input devices such as the GlobeFish and Two-4-Six ranks high on the DOF scale and allows for interaction in three dimensional space suitable for virtual reality environments and CAD applications, but they still suffer from the same drawbacks as the previous controllers. They all require the user to use physical controllers and they only allows control of a single interaction point at a time restricting the user to freely explore the interaction space.

HandVu is also available for computer desktop interaction without the intrusive wearable computer equipment and can roughly track the hand motion of the user for controlling a mouse cursor, however the precision is rather low and the interaction is in a plain two dimensional space restricting the user for real three dimensional interaction. We ignored most of the techniques used in the area of augmented reality, even though some of the techniques aids the merge of virtual and real world entities and also share the same vision of embodiment, however, at the current state most of them are not mature enough to not be intrusive on the user. Even though embodiment is clearly existing within the research field of augmented reality the focus tends to lean more towards the technology behind the sensors and actuators rather than the aspect of embodiment.

Guided by the principles proposed by Dourish together with our findings we made, we designed an interaction medium that should be capable of tracking fingers separately functioning much as a multi-touch display with an additional dimension without being intrusive on the user. The reason why we chose not to include static gesture recognition in our proposed design is to allow the user to interact freely with the system as if it was extended right to the top of their fingertips without having to learn multiple predefined patterns. It is realized that interaction by using only your hands can cause fatigue since the user does not have a place to off-load the weight of the arms compared to using a keyboard or mouse but as previously mentioned our proposed prototype is not meant to be a replacement merely an extension to the interaction space.

The results from the repertory grid technique were useful as they highlighted important properties of interaction devices. The constructs '*Single Purpose - Utilizable*', '*Versatile, Allround - Predefined, Invariable*', '*Mechanical, Restricted - Natural, Free*' and '*Direct - Concealed*' are all closely related to each other as they more or less address flexibility but on slightly different levels. In broad terms, the '*Single Purpose - Utilizable*' looks at to what purpose and scope of use the interaction technology can be utilized for followed by '*Versatile, Allround - Predefined, Invariable*' which address the flexibility within the scope of use. The construct '*Mechanical, Restricted - Natural, Free*' tells us to what extent you are restricted or liberated from the presence of the hardware while '*Direct - Concealed*' address the sense of how much your physical body is in control of the interaction. This illustrates how much a designer needs to keep in mind when designing interactive devices, there is a fine line between these constructs. The proposed prototype was on average seen as the most flexible within these constructs followed by Project Natal and HandVu which is seen as promising results since it was qualities we aspired for. It should be understood that these results does not indicate that it is better, this depends on what you want to achieve with the interaction device.

The prototype was also seen as the most entertaining in the construct '*Entertaining - Mundane*' together with Project Natal and iPad. This construct seemed to address entertaining in the sense of the interaction itself, not for the area which it was applied to. This can be observed in the Nintendo Wii™ case where it was ranked lower than the iPad and AiRL3D as they are meant to be utilized in other aspects than purely gaming. Therefore a construct such as '*Entertaining - Serious*' or '*Playful - Dull*' or similar could also have been useful to take into consideration.

Even though all of the devices mentioned earlier have their specific purpose and are great input devices few of them clearly shows the support for embodied interaction and they are designed around the current interaction paradigm which restricts the user of the freedom of exploring the medium. That is why we do not only propose a device derived from Dourish principles of embodied interaction together with a high DOF but also the need for a shift in the environment the user interacts within, the pure WIMP paradigm simply does not allow the kind of embodiment we are striving for.

Based on our findings and the feedback received from the lead users we can see indications that the proposed prototype points towards supporting embodied interaction.

6 Technical Discussion & The Prototype

Technical Discussion

As mentioned earlier, several different approaches can be taken in order to detect the points of interest needed in vision based interaction and they all have their benefits and their drawbacks.

Skin segmentation is quite computational cheap, however detecting the right segments are not the easiest task since the approach is based on filtering pixels within the range of a skin classifier which could vary widely due to ethnicity and illumination conditions. The use of background subtraction is dependent on a reference image either statically taken before the user enters the frame which are sensitive to fluctuations in the background scene or by adaptive classification which tends to be quite computational expensive (Stauffer and Grimson, 1999; Zivkovic and van der Heijden, 2006). The use of color histograms based on markers placed on the users hands is relative easy to implement and computational cheap, however, it has an obvious drawback due to the fact that intrusive specialized markers needs to be placed on the hand.

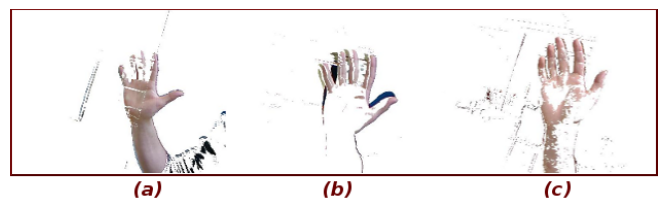


Figure 3: Samples of tested segmentation methods.

Using haptic devices for interaction such as gloves usually provides good precision at a high update rate and are often



Figure 4: *Samples of obtained disparity maps.*

able to detect the posture of the hand. But as previously mentioned, this technique suffers from the same intrusiveness as the previous approach since it requires the user to be equipped with a specialized glove. The stereo vision approach is computational expensive and the images retrieved from the two cameras must be perfectly aligned before the computation of the depth map based on epipolar geometry, also known as disparity map, can be performed. Another drawback of the stereo vision technique is that it relies on careful camera calibration often done by capturing images of a specialized checkerboard. However, the main benefit of stereo vision is the disparity map crucial for obtaining the Z-coordinates in a 3D environment as well as the possibility of extending the technique from purely hand detection to a system that are capable of detecting other objects the user could interact with as well.

Since the prototype needs to be robust in the sense that it should function outside a controlled lab environment and to be non intrusive on the user while preserving the freedom and naturalness of the interaction medium, the stereo vision approach came to our main attention.

The Prototype

Figure 3 shows a few of the segmentation techniques tested during the design phase of AiRL3D. Figure 3A shows the segmentation done by filtering based on the sum of absolute difference, comparing the difference between a pixel in current- and reference frame to a threshold value. Figure 3B is the output of motion detection by comparing the difference of the current- and previous frame. Figure 3C shows one of the skin segmentation algorithms tested, however this was as mentioned before, sensitive to illumination and ethnicity conditions. These methods for segmentation were dismissed due to the unreliable results caused by fluctuations in the surrounding environment. Figure 4 shows a few samples of the segmentation obtained by the stereo vision approach which proved more robust as well as the possibility to extract the three dimensional coordinates.

At the time of writing AiRL3D is in a early phase of the implementation, however, acquisition of synchronized frames with a resolution of 640x480 and generation of the disparity map was successfully implemented and running with a update frequency of ~ 26 frames per second without any optimized code nor any graphic hardware acceleration on an Intel Core 2 Duo T7300 2.0 GHz with 2 GB RAM.

Our prototype has been implemented in ANSI C currently only in a Linux distribution, however, to be able to port the prototype to other platforms the only modification needed is to the webcam interfaces and to the output host such as a multi-touch interface.

Figure 5 shows an overview of the AiRL3D project and a brief explanation to each step is given in Table 2. a

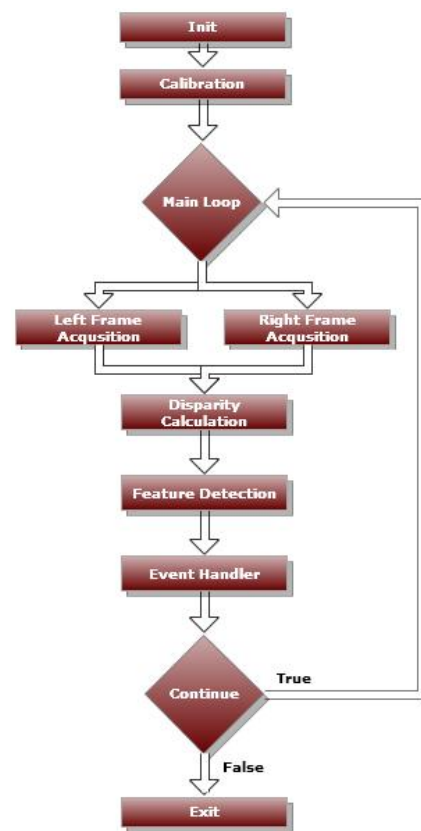


Figure 5: *AiRL3D*

Step	Description
Init	The main initiation procedures for the CCD sensors & RGB to YUV Look-Up-Tables (LUT).
Calibration	Captures one frame from each input stream and applies image stitching to calculate the X, Y and angular offsets, sets up reference LUT to avoid recalculating the rectification for each acquired frame.
Main Loop	Spawns two threads for image acquisition if the host system is a multi-core system.
Left / Right Frame Acquisition	Synchronizes threads to capture frames simultaneously.
Disparity Calculation	Waits for frame threads to complete and handles the disparity algorithm proposed by Shafik and Mertsching (2009). If the host system is a multicore system the disparity calculations are divided between the multiple cores to improve performance of the algorithm.
Feature Detection	Handles identification of fingertips, palm and wrist.
Event Handler	Takes the feature points X-,Y-, and Z-coordinates and sends it to the host system as mouse events.

Table 2: *AiRL3D*

7 Conclusion

In this paper we have, with the help of design principles proposed by Dourish and the findings we have made when investigating the current interaction techniques, defined the design for our prototype.

To allow the interaction medium to become participative one should not restrict the user to predefined interaction rules by designing specific actions, one should leave the interaction space open to allow the user to adapt the medium around their needs instead of adapting their needs around the medium. The interaction should also be natural, in the meaning of not being intrusive on the user and allow the user to freely explore the medium forming their own meaning. However, the interaction technique is not enough, the design should also reflect on the context in which it should be used.

Solutions such as Nintendo Wii have shown to create a more natural and graspable interaction that is understandable by a wider audience. Microsoft's upcoming Project Natal, and Playstation's Move are examples of this as well, showing that at least that the gaming industry is striving towards an approach to support a higher degree of embodied interaction, however, the same techniques can be applied in other fields as well.

We used the guidelines proposed by Dourish and started the implementation of the AiRL3D project and performed validation of our proposed medium towards lead users both showing that the creation of a 3D embodied input devices

is feasible both technically as well as likely to support the perception of a meaningful medium for the user. However to enable this, our prototype is not enough, one needs to push further towards changing the current static two dimensional WIMP paradigm, towards a more flexible environment.

We can not yet know for sure if this will be meaningful for the user or not, after all its the user that decides if something is meaningful when they take action upon an artifact. Nevertheless, what we have tried to show here, is that we as designers have to start thinking outside of the box of the old WIMP paradigm.

Further Research

The interfaces existing today are unfortunately not adapted to this type of interaction, not only does the standard interfaces today mainly support interaction in two dimensions but most of them are also only intended for use of with a single interaction point at a time. To bridge the gap further we propose further research in how to extend the current WIMP paradigm to move from a mainly iconic representation towards a more entity based platform to allow the user to a more embodied state.

References

- Agarwal, A., Izadi, S., Chandraker, M., and Blake, A. (2007). High Precision Multi-touch Sensing on Surfaces using Overhead Cameras. 5
- Alpern, M. and Minardo, K. (2003). Developing a car gesture interface for use as a secondary task. In *CHI '03: CHI '03 extended abstracts on Human factors in computing systems*, pages 932–933, New York, NY, USA. ACM. 3
- Althoff, F., Lindl, R., Walchshausl, L., and Hoch, S. (2005). Robust multimodal hand- and head gesture recognition for controlling automotive infotainment systems. *VDI BERICHTE*, 1919:187. 3, 6, 7, 8
- Avery, B., Thomas, B., Velikovskiy, J., and Piekarski, W. (2005). Outdoor augmented reality gaming on five dollars a day. In *Proceedings of the Sixth Australasian conference on User interface-Volume 40*, page 88. Australian Computer Society, Inc. 5
- Bach, K., Jæger, M., Skov, M., and Thomassen, N. (2008). You Can Touch, but You Can't Look: Interacting with In-Vehicle Systems. *Proceedings of the Human Factors in Computing Systems (CHI'08)*, Florence, Italy, ACM Press, pages 1139–1148. 3
- Bowden, R., Windridge, D., Kadir, T., Zisserman, A., and Brady, M. (2004). A linguistic feature vector for the visual interpretation of sign language. *Computer Vision-ECCV 2004*, pages 390–401. 5
- Bowman, D., Coquillart, S., Froehlich, B., Hirose, M., Kitamura, Y., Kiyokawa, K., and Stuerzlinger, W. (2008). 3D User Interfaces: New Directions and Perspectives. *IEEE Computer Graphics and Applications*, 28(6):20–36. 1, 4, 5

- Coogan, T., Awad, G., Han, J., and Sutherland, A. (2006). Real time hand gesture recognition including hand segmentation and tracking. *Lecture Notes in Computer Science*, 4291:495. 5
- Dey, A. K., Abowd, G. D., and Salber, D. (2001). A conceptual framework and a toolkit for supporting the rapid prototyping of context-aware applications. *Hum.-Comput. Interact.*, 16(2):97–166. 2
- Dourish, P. (2001). *Where the action is: the foundations of embodied interaction*. The MIT Press, Cambridge, MA. 1, 2, 4, 7, 8
- Dourish, P. (2004). What we talk about when we talk about context. *Personal and ubiquitous computing*, 8(1):19–30. 1
- Fredriksson, J., Ryen, S., and Fjeld, M. (2008). Real-time 3D hand-computer interaction: optimization and complexity reduction. In *Proceedings of the 5th Nordic conference on Human-computer interaction: building bridges*, pages 133–141. ACM New York, NY, USA. 1, 5
- Frohlich, B., Hochstrate, J., Kulik, A., and Huckauf, A. (2006). On 3D input devices. *IEEE computer graphics and applications*, 26(2):15–19. 4
- Hippel, E. V. (1986). Lead users: a source of novel product concepts. *Manage. Sci.*, 32(7):791–805. 3, 4
- Hippel, E. V. (1988). *The sources of innovation*. Oxford University Press, Inc., 200 Madison Avenue, New York, New York 10016. 3, 4
- Hirschmüller, H. (2001). Improvements in real-time correlation-based stereo vision. In *Proceedings of IEEE Workshop on Stereo and Multi-Baseline Vision*, pages 141–148. Citeseer. 3
- Hirschmüller, H. (2006). Stereo Vision in Structured Environments by Consistent Semi-Global Matching. In *Proceedings of the 2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition-Volume 2*, page 2393. IEEE Computer Society. 3
- Hirschmüller, H., Innocent, P., and Garibaldi, J. (2002). Real-time correlation-based stereo vision with reduced border errors. *International Journal of Computer Vision*, 47(1):229–246. 3
- Holman, D., Vertegaal, R., Altosaar, M., Troje, N., and Johns, D. (2005). Paper windows: interaction techniques for digital paper. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 591–599. ACM. 4, 8
- Ishii, H. and Ullmer, B. (1997). Tangible bits: towards seamless interfaces between people, bits and atoms. In *CHI '97: Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 234–241, New York, NY, USA. ACM. 2
- Jaimes, A. and Sebe, N. (2007). Multimodal human-computer interaction: A survey. *Computer Vision and Image Understanding*, 108(1-2):116–134. 1
- Joslin, C., El-Sawah, A., Chen, C., and Georganas, N. (2005). Dynamic gesture recognition. In *IEEE Instrumentation and Measurement Technology Conference Proceedings*, volume 22, pages 1706–1711. IEEE; 1999. 5
- Kelly, G. (1955). *The psychology of personal constructs (Vols. 1 & 2)*. 4
- Keskin, C., Erkan, A., and Akarun, L. (2003). Real time hand tracking and 3D gesture recognition for interactive interfaces using HMM. *ICANN/ICONIPP*, pages 26–29. 5
- Kim, M. and Maher, M. (2008). The Impact of Tangible User Interfaces on Designers' Spatial Cognition. *Human-Computer Interaction*, 23(2):101–137. 3
- Kim, S. and Dey, A. (2009). Simulated augmented reality windshield display as a cognitive mapping aid for elder driver navigation. In *Proceedings of the 27th international conference on Human factors in computing systems*, pages 133–142. ACM. 7
- Klemmer, S. (2005). Integrating physical and digital interactions. *Computer*, pages 111–113. 1
- Kolsch, M., Turk, M., Hollerer, T., and Chainey, J. (2004). Vision-based interfaces for mobility. In *Proc. of Intl. Conference on Mobile and Ubiquitous Systems*. Citeseer. 1
- Kurkovsky, S. (2007). Pervasive computing: Past, present and future. In *Proceedings of ITI 5th International Conference on Information and Communications Technology (ICICT 2007)*, pages 16–18. 1
- LaViola Jr, J. (2008). Bringing VR and spatial 3D interaction to the masses through video games. *IEEE Computer Graphics and Applications*, 28(5):10–15. 1, 3, 5
- Lee, S., Cheok, A., James, T., Debra, G., Jie, C., Chuang, W., and Farbiz, F. (2006). A mobile pet wearable computer and mixed reality system for human-poultry interaction through the internet. *Personal and Ubiquitous Computing*, 10(5):301–317. 4
- Loke, S. (2006). Context-aware artifacts: two development approaches. *IEEE Pervasive Computing*, 5(2):48–53. 1
- Malik, S. (2003). Real-time Hand Tracking and Finger Tracking for Interaction. *CSC2503F Project Report*. 3
- Manresa, C., Varona, J., Mas, R., and Perales, F. (2000). Real-Time Hand Tracking and Gesture Recognition for Human-Computer Interaction. *Electronic Letters on Computer Vision and Image Analysis*. 3
- Manresa, C., Varona, J., Mas, R., and Perales, F. (2005). Hand tracking and gesture recognition for human-computer interaction. *Electronic letters on computer vision and image analysis*, 5(3):96–104. 5
- McCarthy, J., Wright, P., Wallace, J., and Dearden, A. (2006). The experience of enchantment in human-computer interaction. *Personal and Ubiquitous Computing*, 10(6):369–378. 1

- Ogawara, K., Hashimoto, K., Takamatsu, J., and Ikeuchi, K. (2003). Grasp recognition using a 3D articulated model and infrared images. In *IEEE/RSJ Proceedings of Conference on Intelligent Robots and Systems*, volume 2, pages 27–31. Citeseer. 5
- Patel, S. N. and Abowd, G. D. (2007). Blui: low-cost localized blowable user interfaces. In *UIST '07: Proceedings of the 20th annual ACM symposium on User interface software and technology*, pages 217–220, New York, NY, USA. ACM. 4
- Pickering, C., Burnham, K., and Richardson, M. (2007). A Research Study of Hand Gesture Recognition Technologies and Applications for Human Vehicle Interaction. In *Automotive Electronics, 2007 3rd Institution of Engineering and Technology Conference on*, pages 1–15. 3
- Piekarski, W., Smith, R., and Thomas, B. (2004). Designing backpacks for high fidelity mobile outdoor augmented reality. In *Proceedings of the 3rd IEEE/ACM International Symposium on Mixed and Augmented Reality*, page 281. IEEE Computer Society. 5
- Plouznikoff, N., Plouznikoff, A., Desmarais, M., and Robert, J. (2007). Gesture-based interactions with virtually embodied wearable computer software processes competing for user attention. In *IEEE International Conference on Systems, Man and Cybernetics, 2007. ISIC*, pages 2533–2538. 1, 3, 5, 7
- Satyanarayanan, M. et al. (2001). Pervasive computing: Vision and challenges. *IEEE Personal communications*, 8(4):10–17. 1
- Schlattmann, M., Kahlesz, F., Sarlette, R., and Klein, R. (2007). Markerless 4 gestures 6 dof real-time visual tracking of the human hand with automatic initialization. In *Computer Graphics Forum*, volume 26, pages 467–476. Blackwell Science Ltd, Osney Mead, Oxford, OX 2 0 EL, UK,. 5
- Shafik, M. and Mertsching, B. (2009). Real-Time Scan-Line Segment Based Stereo Vision for the Estimation of Biologically Motivated Classifier Cells. *KI 2009: Advances in Artificial Intelligence*, pages 89–96. 11
- Shan, C., Tan, T., and Wei, Y. (2007). Real-time hand tracking using a mean shift embedded particle filter. *Pattern Recognition*, 40(7):1958–1970. 5
- Shaw, M. and Thomas, L. (1978). FOCUS on education—an interactive computer system for the development and analysis of repertory grids. *International Journal of Man-Machine Studies*, 10(2):139–173. 4
- Smith, R., Piekarski, W., and Wigley, G. (2005). Hand tracking for low powered mobile AR user interfaces. In *Proceedings of the Sixth Australasian conference on User interface-Volume 40*, page 16. Australian Computer Society, Inc. 5
- Stauffer, C. and Grimson, W. (1999). Adaptive background mixture models for real-time tracking. In *Computer Vision and Pattern Recognition, 1999. IEEE Computer Society Conference on.*, volume 2. 9
- Suchman, L. (1987). *Plans and situated actions: The problem of human-machine communication*. Cambridge Univ Pr. 8
- Tan, F. and Hunter, M. (2002). The repertory grid technique: A method for the study of cognition in information systems. *MIS Quarterly*, 26(1):39–57. 3, 4
- Tönnis, M., Broy, V., and Klinker, G. (2006). A Survey of Challenges Related to the Design of 3D User Interfaces for Car Drivers. In *Proceedings of the 1st IEEE Symposium on 3D User Interfaces (3D UI)*, pages 127–134. 3
- von Hardenberg, C. and Bérard, F. (2001). Bare-hand human-computer interaction. In *Proceedings of the 2001 workshop on Perceptive user interfaces*, pages 1–8. ACM. 5
- Walsham, G. (2006). Doing interpretive research. *European Journal of Information Systems*, 15(3):320–330. 3
- Williams, A., Kabisch, E., and Dourish, P. (2005). From interaction to participation: Configuring space through embodied interaction. *Lecture Notes in Computer Science*, 3660:287–304. 2
- Wingrave, C., Williamson, B., Varcholik, P., Rose, J., Miller, A., Charbonneau, E., Bott, J., and LaViola, J. (2009). Wii Remote and Beyond: Using Spatially Convenient Devices for 3DUIs. *Computer Graphics and Applications*. 3
- Zhao, J., Yu, S., and Cai, H. (2006). Local-global stereo matching algorithm. *Aircraft Engineering and Aerospace Technology: An International Journal*, 78(4):289–292. 3
- Zivkovic, Z. and van der Heijden, F. (2006). Efficient adaptive density estimation per image pixel for the task of background subtraction. *Pattern recognition letters*, 27(7):773–780. 9