Fabian Hennecke, Alina Hang
Hendrik Richter, Andreas Butz
Heinrich Hussmann
(Editors)

Over Distance

Hauptseminar Medieninformatik WS 2011/2012

Technical Report
LMU-MI-2012-1, Feb. 2012
ISSN 1862-5207

University of Munich
Department of Computer Science
Media Informatics Group

Fabian Hennecke, Alina Hang, Hendrik Richter
Andreas Butz, Heinrich Hussmann (Editors)

Over Distance

An overview of current research dealing with various problems over distance

Preface

This report provides an overview of current research trends dealing with various problems ‘over distance’. There are various applications domains ranging from haptic interfaces, interactive surfaces, collaboration to telepresence robots.

During the winter term 2011, students from the Computer Science Department at the Ludwig-Maximilians-University in Munich did research on specific topics related to ‘over distance’ problems and analyzed various publications. This report comprises a selection of papers that resulted from the seminar.

Each chapter presents a survey of current trends, developments, and research with regard to a specific topic. Although the students’ background is computer science, their work includes interdisciplinary viewpoints such as theories, methods, and findings from interaction design, ergonomics, hardware design and many more. Therefore, the report is targeted at anyone who is interested in the various facets of ‘over distance’ interaction and representation.

Munich, February 2012

The Editors

Fabian Hennecke, Alina Hang
Hendrik Richter, Andreas Butz
Heinrich Hussmann

Contents

Application Scenarios for Virtual Environment Interfaces with Haptic Feedback

Julia Polleti

Abstract— Virtual environments (VE) create the illusion of being in another world. In our daily lives, many of our decisions depend on haptic impulses, like temperature, texture, motion or force. In order to make the interaction with virtual worlds more realistic and entertaining, haptic feedback can be provided by special interfaces and actuators. Both research and industry use the combination of VE and haptic feedback devices. This paper presents an overview of scenarios with virtual environment interfaces with haptic feedback in the fields of medicine, molecular research, gaming, industrial design and others. Augmented reality (AR) is considered to be a variation of VE. Therefore, a basic understanding of its technology is presented in the end of this paper. To conclude, pros and cons of both VE and AR are discussed. Additionally, an estimation about the future evolution of virtual environments, virtual environment interfaces and haptic feedback is given.

Index Terms—Human Machine Interaction, Virtual Reality, Virtual Environment Interfaces, Haptic Feedback, Simulation

1 Introduction

For the purpose of research, warfare or just entertainment, the crew of Star Trek's Deep Space Nine can enter the HoloDeck. This is a special room on the spaceship that can simulate another world. In one episode, the character Data enters 19th century London as Sherlock Data Holmes. Architecture and people of this place and time appear real to him, even though they are far away in both distance and time. He can touch and interact with creatures and objects in the same way as in the real world. When moving, the world around Data adapts to his position and creates therefore a perfect virtual illusion of another world.

Apart from some limitations, the progress in modern technology in the field of virtual reality (VR) makes it more and more possible to create virtual worlds like in Star Trek's HoloDeck. One of the first projects that used VR technology was developed by Krueger et al. between 1970 and 1985 [23]. In his project "Videoplace", the user is tracked by a camera and the image is projected on a video screen. Virtual objects are added to the image. The user can interact with those elements, for instance by moving her hands. Then, the representation of the user on the screen "touches" the virtual objects [23].

Virtual reality has several advantages. It can help people to learn specific tasks, for example in the field of molecular research or laparoscopic surgery, that are hardly possible to learn in the real world [43]. It is a good way to simulate environments that are too dangerous, dirty or small to visit in person [43, 21]. And last but not least, it can add a high degree of entertainment to applications like games [43, 1].

In order to make the experience and interaction with the virtual world more realistic, haptic feedback can be added via special interfaces. Actuators and commercially available haptic devices are outlined in 3.1. In 3.2, several scenarios of the fields medicine, molecular research, gaming, industrial design and others, that all exploit the advantages of both VR technology and haptic feedback, are presented. Then, the related field of augmented reality is presented. Pros and cons of VR and AR are discussed. An estimation of the future of VE, VEI and haptic feedback is given in 5.

2 Definition

This paper presents different application scenarios of virtual environments with haptic feedback. Therefore, all discussed scenarios contain two main technologies. On the one hand, there is the virtual environment. It is an interactive display that makes the user feel to be in another, virtual world [15, 36]. On the other hand, user interaction happens with interfaces that provide haptic feedback. This can be either tactile feedback or kinesthetic feedback [45]. In order to give a better understanding of both technologies, definitions of "virtual environment" and "haptic feedback" are presented in 2.1 and 2.2.

2.1 Virtual Reality

The term of VE or VR has been defined in various ways, so far. In his foreword to Wexelblat's "Virtual Reality: Applications and Explorations", Cadigan calls *the book* the oldest form of virtual reality [43]. When looking at the field of Human Machine Interaction, one can find two main approaches of defining virtual environments. Early definitions often focus on the technological equipment that is needed and used for creating VE. The terms "goggles and gloves", "head mounted display" and "three-dimensional" can often be found in those definitions. For Krueger, virtual environments are "three-dimensional realities implemented with stereo viewing goggles and reality gloves" [22]. Ellis describes it as "interactive, computer-graphics based, head-referenced displays that create the illusion that their users are in a place other than where they actually are" [15].

Another way of defining virtual environment was presented by Steuer. In his definition, he focuses on the terms of "presence" and "telepresence" rather than on specific hardware. In Steuer's opinion, presence is the "natural perception of an environment" [36] in the real world. In contrast, telepresence or virtual reality, is the "mediated perception of an environment" [36].

Other approaches to define virtual reality discuss its purpose of usage. For Cadigan, it is a "computer-generated graphic environment for purposes of education, work, and/or recreation" [43].

All definitions have one thing in common: Distance is an important fact in virtual environments. A virtual reality display makes it possible for the user to communicate with people over distance, to act in a "temporally or spatially distant 'real' environment" [36] or to virtually control distant tools.

2.2 Haptic Feedback

In most cases, human machine interaction happens via the visual, auditory or haptic senses. In contrast to audiovisual sensing, haptic is the only modality that enables both input (acting) and output (sensing) of information [45, 40, 6]. Therefore, a bidirectional way of communication between a machine and a user is possible. The user can get haptic feedback by the computer while controlling the system with his or her body *(see figure 1)*.

The human haptic sense is divided in the tactile and the kinesthetic sense.

- *Julia Polleti is studying Media Informatics at the University of Munich, Germany, E-mail: julia.polleti@googlemail.com*
- *This research paper was written for the Media Informatics Advanced Seminar 'Over Distance', 2011/2012*

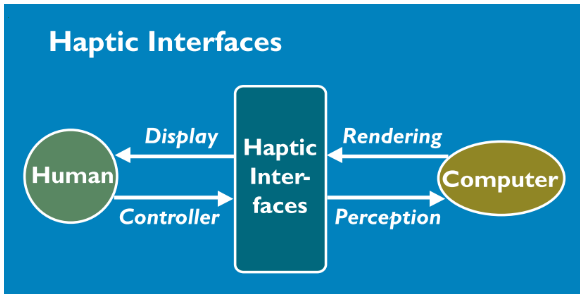

Fig. 1. Definition of Haptics by Tan [40].

Tactile Feedback

The tactile, or cutaneous, sense refers to temperature, texture, roughness, softness, shape, slipping, surface compliance, elasticity, viscosity, and electrical conductivity [5, 6, 45]. This kind of information is gathered by receptors in the human skin [5, 40].

Kinesthetic Feedback

Kinesthetic feedback, or prioreceptive feedback, refers to the feeling of motion, force and weight [5, 45]. The human body gathers this kind of information by the motion of limb segments, muscle tension and joint position [5, 6, 45, 40].

3 Systems and Scenarios of VE with Haptic Feedback

Why is haptic feedback so important in virtual environment scenarios? In order to find a response to that question, one has to take a look at the "real world". How we act and decide in our daily lives is influenced by many different kinds of natural haptic feedback that we get from our environment. We feel a train coming on by vibrations on the subway station floor; we try if the ice of a lake is thick enough by feeling the hardness of its surface; we check if a cooktop is already hot enough for putting food in the pan; we test how slippery wet floors are to estimate how slow and cautious we have to walk on them. Haptic feedback is important for us to take the right decisions in daily live. Being an essential part of the real world, haptic feedback can make virtual environment more realistic and improve their performance [6, 24]. Studies showed that the error rate for grasping virtual objects can be reduced by over 50% by force feedback [18]. In some scenarios, the existence of haptic feedback is indispensable.

3.1 Systems with Haptic Feedback

Haptic interfaces can be classified in active and passive devices. The mouse, the keyboard and the trackball are well-known passive haptic interfaces. The user can feel their existence via tactile and kinesthetic sensing. Although, the forces that they apply on the user's hand are not controlled by a program. Interfaces that provide systematic and directed haptic feedback are called active [6]. There are several possibilities to provide interaction interfaces with active haptic feedback. They range from simple actuator technology to fully developed and commercially available haptic interfaces that are used by researchers and companies throughout different scenarios of VE.

Actuators

Actuators can be used to provide different kinds of force or tactile feedback. In virtual environment interfaces, pressure can be simulated by pneumatic systems, vibration by voice coils, temperature by heat pump systems, shape by solenoids, piezoelectric crystal, electromagnetic vibration needles/heads or shape-memory alloy technologies [45, 5].

PHANToM

Originally developed at the Massachusetts Institute of Technology (MIT), the *PHANToM* device is often used to provide haptic feedback in VE scenarios. Today, it is marketed by Sensable, which is a company that develops 3D touch-enabled technology []. The *PHANToM* devices are desktop-based interfaces that enable the user to feel and manipulate virtual objects via a stylus in her hand. Some *PHANToM* products provide a thimble, wherein the user can put her finger. With this, tactile feedback is directly sensible on the user's fingertip. The *PHANToM* interfaces provide force feedback between three and six degrees of freedom.

Argonne Remote Manipulator

The Argonne Remote Manipulator (*ARM*) was developed by the Argonne National Laboratory for teleoperation applications. It is a so called master-slave manipulator which was designed to remotely handle reactor components in nuclear industry [25]. The *ARM* provides force feedback via servo-motors and potentiometers [45].

Impulse Engine

Immersion Corporation is a company that produces technology interfaces with haptic feedback [13]. With its product family *Impulse Engine*, Immersion Corporation created different haptic devices that provide force feedback by servo-motor actuators. They can be connected with a PC, Mac or hardware from Silicon Graphics. As the Impulse Engine products are mostly designed for virtual reality simulations of surgery procedures, devices like the *Laparoscopic Impulse Engine* or *Needle Insertion Simulator* can be controlled via surgical instruments and tools [45].

3.2 Scenarios of VE with Haptic Feedback

The early work on tele-operation and tele-existence, which later created the field of virtual reality, focuses on scenarios of robot simulation in hostile environments. Hostile environments are places, where it is too dangerous or impossible for the human being to live and work. Therefore, early work included scenarios of nuclear research, aerospace industry or environments with extreme temperature [39].
Since the development of virtual environments, research has been done on adding haptic feedback in order to make the VE more realistic. In some fields, acting in a virtual environment and getting haptic feedback from the system is particularly important.
One can find several scenarios, products and research projects for topics like medicine, molecular research, gaming and industrial design. This is due to the fact that in those fields, the necessity of virtual environments that provide haptic feedback is very high. In medicine and molecular research, haptic feedback is very important. However, specific tasks can not be performed on real objects. In gaming, existing virtual environments get much more interesting for gamers with the addition of haptic feedback. In the field of industrial design, forming objects by hand is still an important fact and can not be done only by using computer aided design (CAD) software.

3.2.1 Medicine

In the field of medicine, VE with haptic feedback are used for rehabilitation processes. Also, simulations can help students to learn how to perform laparoscopic surgery. Practicing this activity in a non-virtual environment is hardly possible. However, haptic feedback is very useful for this purpose [3, 44, 9].

Telerehabilitation System, 2000

In 2000, Popescu et al. presented a virtual-reality-based telerehabilitation system that provides force feedback [29]. The system was designed to help people practicing orthopedic rehabilitation from home. With this system, patients can exercise whole hand or single finger movement. Data about the patient's progress is sent to the clinic's server. The telerehabilitation system includes a library of different virtual exercises that can be adopted to the user's needs. Those tasks include interacting with virtual objects by grasping (either with the whole hand or with two fingers), pointing or releasing them. The user can see the objects and a virtual hand representing the movements of her real hand on the PC screen. In order to feel the virtual objects, haptic feedback is provided by a *RM-II* haptic glove. In one exercise, the user has to squeeze a virtual rubber ball. In contact with the virtual hand, the ball deforms and the user gets force feedback by the glove, simulating the pressure of a real rubber ball [29].

Laparoscopic Surgery Simulation, 2001

Much work has been done on simulating laparoscopic surgery with virtual environment interfaces and haptic feedback. Laparoscopy is a technique which is used for medical surgery like operations of gallon stones. In this process, a small video camera and instruments are inserted into the human body. With this, larger openings of the human body can be avoided [4, 3].
In 2001, Basdogan et al. developed a tool for the training of laparoscopic surgery [4]. With this system, people can exercise laparoscopic catheter insertion - a task which is performed in gallon stone detection. The tool presented by Basdogan et al. consists of a haptic interface and a PC screen which displays the operating scene and the virtual instruments *(see figure 2)*. The user can control the virtual instruments on

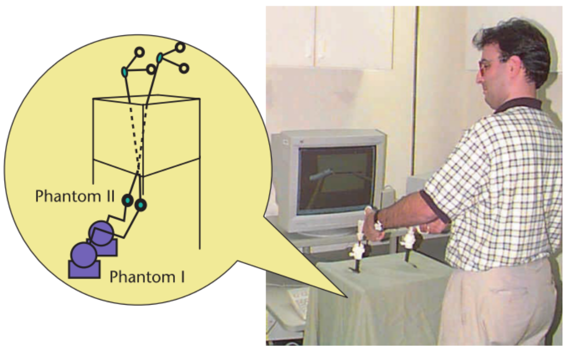

Fig. 2. The training tool for laparoscopic surgery by Basdogan et al. [3].

the screen by using real laparoscopic instruments connected to two *PHANToM* stations []. They add haptic feedback to the instruments. With this, the user has a tactual sense of the virtual organs and gets force feedback during the exercise [4].

LAP Mentor, 1997-today (2012)

A commercial product for training laparoscopic surgery situations with haptic feedback is the *LAP Mentor* [34]. The system was developed by Simbionix, a company producing medial training simulators since 1997. The *LAP Mentor* displays a virtual surgery situation on a screen in front of the user. He can manipulate the virtual organs with two real forceps which are connected to the system. A virtual representation of the instruments is shown on the screen. Contact with the virtual objects is rewarded as force feedback to the interfaced instruments. Therefore, the system is used to provide exercising of bariatric (“The branch of medicine that deals with the causes, prevention, and treatment of obesity” [17]), colorectal (“Relating to or involving both the colon and the rectum” [17]), urologic or gynecologic surgery processes [34]. Yiasemidou et al. evaluated the *LAP Mentor* in 2011, comparing it to a similar simulation tool without haptic feedback. The study showed that the existence of haptic feedback reduced the task completion time for laparoscopic procedures significantly [44].

Endonasal Sinus Surgery Simulation, 2009

In the field of surgery simulation, a tool for teaching students endonasal sinus surgery was presented by Tolsdorff et al. in 2009 [42]. In this process, endoscopes have to be inserted in the patient’s nose. So far, training is mostly done on cadavers. In order to avoid disadvantages from cadaver dissections, such as unrealistic color and firmness, the research on virtual simulations is important on this field. In Tolsdorff’s system, the user can virtually exercise endonasal surgery with both realistic visualization and haptic feedback. The organs and virtual instruments are displayed to the user through shutter glasses in a three-dimensional real-time visualization *(see figure 3)*. He can interact with

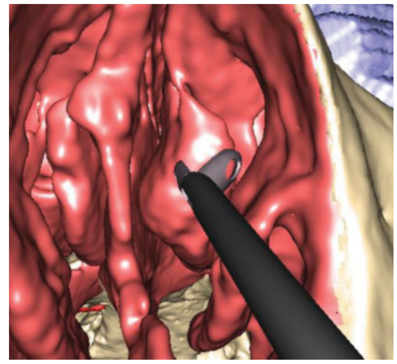

Fig. 3. The user’s view on the virtual surgery scenery in Tolsdorff’s surgery simulator [42].

the virtual instruments with a *PHANToM Omni* device []. The *PHANToM* station provides force feedback when the virtual instrument collides with the virtual organs or tissue. Tolsdorff’s system provides a “high congruency between the visual and haptic feedback”[42] which is indispensable in this kind of surgery simulation.

3.2.2 Molecular Research

In the field of molecular research, VE with haptic feedback are used to train the manual docking or modeling of molecules. Molecular docking and modeling is needed to find out, for example, which compound is suited as a medicine for special diseases. A virtual simulation with haptic feedback can help untrained users to learn those processes and to reduce task completion time [7, 10].

GROPE 1970-1990

In 1990, Brooks et al. presented their work on the *GROPE* project, a haptic display for interacting with protein molecules [10]. Their latest simulation tool, the *GROPE-III* has been developed for chemistry research and education. Virtual molecules are displayed on a large screen offering stereo vision. The user can perform molecular docking tasks via an Argonne Remote Manipulator (*ARM*) device *(see figure 4)*. Individual dials that simulate the twistable bonds in the drug are

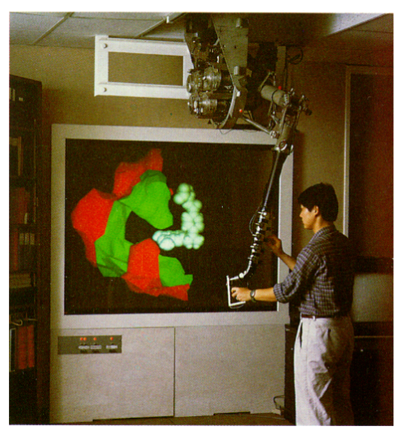

Fig. 4. The *GROPE-III* haptic display system by Brooks et al. [10].

added to the *ARM* [10]. As the system provides prioreceptive feedback, the user knows of the position of the virtual instrument. The *ARM*’s handgrip applies pressure feedback to the user’s hand. A study showed that some parts of molecular docking could be performed significantly faster with the haptic feedback of the *GROPE-III* than without [10].

Molecular Docking Simulation, 2008

Another tool for molecular docking problems was created by Subasi et al. in 2008 [38]. Their device is especially designed for the virtual docking of rigid-body molecules. The user can see visualized proteins and ligands on a PC screen. First, he has to find possible docking places on the protein. In the second step, he can test if the ligand connects well with the detected binding sites. During this process, he gets haptic feedback via a desktop instrument which is not specified by Subasi et al[1]. The haptic feedback gives information about the texture of the protein's surface [38].

Molecular Docking Simulation, 2009

In 2009, Daunay and Régnier presented a tool for getting haptic feedback in molecular docking simulations [14]. The device can be used for "any molecular simulator based on a force field minimization process"[14]. The docking scene appears on a large display in front of the user. A *Virtuose* haptic device from Haption company provides haptic feedback in six degrees of freedom during the docking process. The user can choose two kinds of haptic feedback. In order to get an impression of the approximate interaction energy trend, she can choose smooth haptic feedback. Second, she can opt for the realistic energy profile of the molecules. All in all, the build-up of Daunay and Régnier enables the user to interpret micro-forces of the molecules via visual and haptic feedback [14].

Molecular Docking Simulation, 2010

Bolopion et al. developed a tool for learning nanoscale molecular docking phenomena [7]. The molecular docking simulation software *SAMSON* is connected to a *PHANToM* haptic interface. The user has two possibilities to manipulate the virtual molecules with the haptic device. First, he can change the representation or deform the molecule's structure by setting the position of the *PHANToM*. Second, manipulating the molecules is also possible via applying force on the virtual objects that is on the haptic device. Several applications can be used with this build-up. For instance, the user can also get haptic feedback to feel the interactions of inhibitors with HIV molecules [14].

3.2.3 Gaming

Virtual environments with haptic feedback are often used in the field of gaming. Especially so called first-person shooters have the intention to let the user feel like actually being in the virtual game environment. Haptic feedback can make this feeling more realistic and add interesting experience to the game [1].

Game Consoles, 1986-today (2012)

In 1986, the arcade game *Out Run* by SEGA was the first example, where haptic feedback was displayed to the user in a first-person virtual environment game *(see figure 5)*. The player sits in a car seat in front of a monitor. He can control the car with a gear selector and a steering wheel. The car seat is connected to a hydraulic motor and simulates the motion of a real car. Also, the player can feel it on the seat when he hits another car or an obstacle. Since *Out Run*, several products were developed by the gaming industry to increase a game's entertainment factor and realism with haptic feedback. Sony's *DualShock* controller for PlayStation consoles provides vibration feedback for games like *Burnout Paradise* [35]. The controller vibrates when the virtual character's car collides with an obstacle. Nintendo users can add a so called *Rumble Pak* to controllers or consoles, like the Nintendo DS or controllers of the Nintendo 64 [27]. In games like the first-person shooter *Quake 64* or car racing games like *Ferrari Challenge*, it adds haptic vibration effects to the game, for instance, when the user's virtual character hits an obstacle or gets injured. In virtual environment games like *EA Sports Grand Slam Tennis* for Nintendo's Wii, the user can feel the vibration when the tennis ball hits his virtual racket in the game [27]. Those effects can make games more realistic and fun for users [1].

Fig. 5. The arcade game *Out Run* by SEGA [31].

[1]By pictures in Subasi's paper, I assume that the device is a *PHANToM* station.

Hapticast, 2006

Hapticast is a haptic 3D game that was presented by Andrews et al. in 2006 [1]. The game is a typical first-person adventure game where the user has to fight enemies in a virtual fantasy world. The game environment can be displayed on any screen connected to a Windows2000/XP system. Haptic feedback is provided by a *PHANToM Omni* or Desktop device offering three degrees-of-freedom feedback to the user *(see figure 6)*. Playing the role of a wizard, the gamer can use four different wands. When spelling a cast with one of the wands, she gets haptic feedback via the *PHANToM* device. Each wand displays a specific haptic effect, depending on its function and utilization. Using the "Lift and Swing wand", the gamer can touch objects in the virtual environment. When firing on enemies with the "Bolt and Blast wand", she can feel the recoil effect on the haptic device. With the "Lob wand", the user can throw grenades at the enemy. She can feel the force that the grenade will be thrown from the wand on the *PHANToM* station [1].

Fig. 6. The *Hapticast* game with a *PHANToM* haptic device by Andrews et al. [1].

Battlefield 3 Simulator, 2011

In 2011, the british TV show *The Gadget Show* built and presented the *Battlefield 3 Simulator* [26]. The virtual game environment of the first-person shooter Battlefield 3 is displayed by five high-definition (HD) projectors in a 360-degree projection dome with a height of four meters. At the beginning of the game, the user is standing in the middle of the dome. When she wants to move her virtual character through the game setting, she can walk in any speed or direction on treadmills which are placed in a circle around the middle of the dome. In order to shoot enemies in the game, the user can use a haptic controller, formed like a real gun. When pulling the gun's trigger, the virtual gun shoots. The *Battlefield 3 Simulator* provides additional haptic feedback by 12 paintball guns placed around the user. Every time the virtual character gets injured in the game, also the real player gets shot. With this technique, the simulator provides three kind of haptic interaction. First, the haptic gun device which the user is using to shoot enemies. Second, the treadmills, on which the user can walk and feel the floor of the virtual anvironment. And third, the pain which is caused by hits of the paintball guns [26].

3.2.4 Industrial Design

In the field of industrial design, haptic has always played a major role. Dinnerware out of ceramics or clay is traditionally formed by hand. Today, industrial design is often done with CAD software. However, the ability to form special elements by hand is still needed. If so, combining the CAD software with haptic feedback for the user can be an important fact [12, 16].

VADE, 1995

VADE, the Virtual Assembly Design Environment was presented by Jayaram et al. in 1995 [20]. The tool helps users to plan and evaluate assemblies. In contrast to the real world, it is possible to un- or redo assemblies in virtual reality. *VADE* can be connected to any parametric CAD software system and provides a virtual environment in which the user can directly design and plan an assembly. The user can see the assembly scene and his virtual hand(s) on a head-mounted display. In order to grab and manipulate a virtual object, he has to wear one or two Cybergloves. With this, the system is able to track information about the user's hand movement [20].

Jewelery Design, 2001

In order to evaluate the benefits of haptic feedback in the industrial design process, Cheshire et al. run a user study in 2001 [12]. Participants should design and manually form a piece of jewelery that provides basic functions of a mobile phone. For this process, *Freeform* software by Sensable Technologies was used. It displays the created objects in a virtual three-dimensional environment on a PC screen. Additionally, a *PHANToM* haptic device provided haptic feedback and was the input method for the manipulation of virtual objects. Physical interaction and hands-on model building is considered as an important part of the industrial design process [12]. Therefore, Cheshire et al. assumed that haptic feedback would make the usage of CAD software more realistic and easy. Their study showed that the modelling of some forms and shapes would not have been possible without haptic feedback. Nevertheless, the connection of the *PHANToM* station and CAD software required a "high level of hand-eye coordination"[12] and has limitations. With the system build-up that was used in 2001 for this study, designers had problems in obtaining the shapes and forms that they wanted to create with the haptic input method. Instead, they often switched back to the "normal" usage of the CAD software which means interacting with the mouse and the keyboard. Aside from the technical limitations of the evaluated system, Cheshire et al. still think that tactile modeling itself can benefit professional design processes [12].

Perfume Container Design, 2002

Sener et al. also evaluated the combination of *Freeform* software and a *PHANToM* haptic device [32]. In 2002, they run a user study with both design professionals and industrial design students. In the study scenario, participants had to design a perfume container in two hours. The *PHANToM* device simulates the surface and haptic feedback that would be provided by clay and foam in a real-world hands-on creation process. The study had similar results like the study run by Cheshire et al. in 2001. On the one hand, virtual environments with haptic feedback can create a realistic experience of modelling industrial design objects. On the other hand, the combination of *Freeform* software and *PHANToM* device was not elaborate enough for the designers at this time. They had problems in using the keyboard, mouse and *PHANToM* device at the same time. Some participants stated that their wrist got tired while using the *PHANToM* station. With six degrees of freedom, the *PHANToM* had its limitations in movement of the virtual perfume container. Surface and decoration modeling was considered to be not accurate enough with haptic input. All participants claimed that they still preferred using the CAD software via mouse and keyboard, even though they appreciated the haptic feedback [32].

Toaster Design, 2005

A third evaulation of *Freeform* with haptic feedback by a *PHANToM* station was done by Evans et al. in 2005 [16]. In this scenario, the participants should design a toaster. With the *PHANToM* device, the users could feel and manipulate the virtual toaster's surface. In his case study, Evans compared the design process with traditional styrofoam forming techniques with the virtual creation with the *Freeform*/*PHANToM* system. The haptic input method was especially appreciated by the participants during the creation of hammered effects on the toaster's surface, as shown in *figure 7*. Also the possibility of undoing steps in the virtual creation process was an advantage. Nevertheless, the *Freeform*/*PHANToM* system has its limitations. Participants claimed that they could not feel fine details - a problem which refers to the resolution of the system [16]. The studies

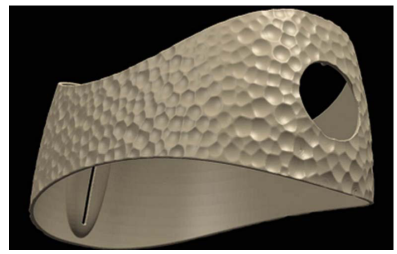

Fig. 7. The hammered surface of the virtual toaster created with haptic input [16].

of Cheshire, Sener and Evans show that the potential of virtual environments with haptic feedback is high for the industrial design process [12, 32, 16]. Some advantages are the possibility of working elsewhere than in a workshop, undoing creation steps and the haptic experience itself. Nevertheless, using other haptic feedback devices like gloves could solve some of the problems that occured in the studies [16].

3.2.5 Other Scenarios

Many projects of virtual environments with haptic feedback have been created so far in the field of medicine, molecular research, gaming and industrial design. However, this technology is also utilized in other topics, like land mine clearance, collaboration, art and music. Due to space limitations and less examples in the field of art and music, one scenario of each of above mentioned areas is presented in this section.

Land Mine Clearance, 2000

In 2000, SimTeam, a subsidiary of MUSE Technologies Inc., created a tool that helps training soldiers of the French Army to detect land mines [37]. The system consists of a head-mounted display, a *PHANToM* haptic device and a standard military probe which is usually used for land mine clearance. The user can see a virtual representation of the probe in a ground area on the display. She can feel the collision of the probe and land mines through the *PHANToM* device [37].

Collaboration, 2000

Sallnäs et al. run a user study to evaluate the impact of haptic feedback on collaborative desktop virtual environments [30]. The subjects had to manipulate the position of virtual objects in a collaborative way. Two participants had to work together, both sitting in two different rooms. They were able to communicate via audio. Each participant had two *PHANToM* stations as input- and haptic feedback devices in front of him. The virtual environment scenery was displayed to them on a PC screen. For instance, in order to lift virtual objects, the users had to touch it on two opposite sides. Results from Sallnäs' study showed that haptic force feedback in virtual collaborative environments can improve task performance and perceived virtual presence [30].

Art, 1996

Jeffrey Shaw is a new media artist, who focuses on interactive virtual reality installations [33]. One of his art works is *ConFIGURING the CAVE*, which was created in cooperation with Agnes Hegedues and Bernd Lintermann in 1996. This art piece is an interactive video installation created with CAVE technology stereographic virtual reality environment (*see figure 8*). It can be manipulated by a wooden puppet user interface. By changing the mannequin's limb positioning, the user can influence the virtual environment that is projected on a three-dimensional wall display around him. *ConFIGURING the CAVE* emphasizes the "relationship between corporeal and spatioal coordinates"[33] even though it does not provide active, that is computer-driven, haptic feedback.

Fig. 8. *ConFIGURING the CAVE* - a virtual reality video installation with a haptic user interface by Jeffrey Shaw et al. [33].

Music, 2010

Immersive Technologies created virtual turntables with haptic feedback in 2010 [19]. The user can scratch virtual turntables by using a *PHANToM* haptic interface with thimbles. In order to create a realistic scratching experience, force feedback is displayed to three of the user's fingers. The user can see the virtual turntables in stereo vision through special goggles [19].

4 Augmented Reality

Virtual reality is often compared to Augmented Reality (AR) due to some similar attributes, yet there are some key differences. VR replaces the real world with a virtual environment that makes the user feel to be in another world. In contrast, AR supplements the real world with three-dimensional virtual objects [2]. The technology of AR overlays computer-generated pictures over "a user's view of the physical world"[28]. For Azuma, it has three main characteristics [2]:

- AR is a combination of real and virtual objects.
- AR provides interactivity in real time.
- AR is displayed to the user in 3-D.

AR is a variation of VR and can help users to do real-world tasks. Amongst others, one can find scenarios of AR in fields like medicine and entertainment/gaming [2]. Those are topics, where both VR and AR is often used. Since several examples for the usage of VR in medicine and gaming were presented in 3.2.1 and 3.2.3, we will now take a look at scenarios in medicine and gaming from the AR point of view.

AR in Medicine

The *ProMIS* tool by CAE Healthcare is a laparoscopic simulator that uses the technology of augmented reality [11, 8]. It consists of a life-size puppet torso with an integrated instrument tracking system. The user can interact with any real laparoscopic instruments that are placed in the mannequin's torso *(see figure 9)*. The motion and position of the instruments is tracked by three cameras in the torso. The user can perform both virtual reality and augmented reality tasks. When doing virtual reality tasks, the position of the real instruments is tracked and the user can see a virtual surgery scenery with virtual representations of the instruments on a screen. When performing augmented reality tasks, the screen displays the scene that the camera is filming in the puppet torso. Augmented reality effects are added to this image.

Fig. 9. The *ProMIS* simulator for laparoscopic surgery by CAE Healthcare [11].

In 2007, Botden et al. run a study to compare the augmented reality technique of the *ProMIS* device with the virtual reality technique of the *LapSim*, a virtual reality laparoscopic simulator with haptic feedback [8]. The subjects had to perform tasks like suturing or translocation with both simulators. Results of this study showed that the realistic experience was far better with the augmented reality simulator than with the VR simulator *LapSim*. One important reason for that was that the haptic feedback was more realistic with the AR simulator. This is due to the nature of the system. The AR *ProMIS* device uses real tissue in order to provide haptic feedback [8].

The advantage of AR simulators is the better haptic feedback. In contrast, VR simulators do not use real tissue. As "the acquisition of human cadaver donors is getting more difficult from year to year" [42], the investigation on better haptic feedback for VR simulators may be an important step for future research.

AR in Gaming

In 2000, Piekarski and Thomas developed the game *ARQuake* [41]. The basic game software is taken from the popular first-person shooter *Quake*. Piekarski et al. changed the software and used special AR hardware in order to make an augmented reality version of *Quake*. The user has to wear a head-mounted display and a wearable computer system on a backpack. Furthermore, he has a plastic gun with recoil haptic feedback. Through the display, the user can see the real world, with overlaying virtual representations of objects and monsters of the game *(see figure 10)*. AR games like *ARQuake* make it possible

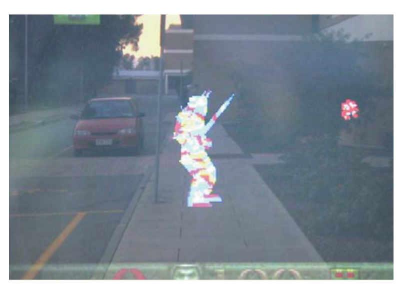

Fig. 10. *ARQuake*: The user's view on the real world with overlaying virtual objects [28].

to do physical outdoor activity while playing computer games. Therefore, they can combine both health and fun. As the technical hardware requirements are often cheaper, VR gaming technology is easier to use for domestic purpose than AR games. Additionally, AR games require a specific environment. The user's position has to be tracked via Global Positioning System (GPS) or with specific landmarks that a camera on his head can track [28].

5 Conclusion

In this paper, an overview about scenarios of virtual environments with haptic feedback was given. First, technical terms that are often used in this context, like virtual reality/environment and haptic feedback, were explained. Some devices were presented that can be used to add haptic feedback to VE. Both research and commercial projects show that the *PHANToM* station can provide satisfying haptic feedback for VE. However, its shape does not fit perfectly to every scenario of VE. Sometimes, one may prefer another form than the *PHANToM*-specific stick. Also, some scenarios may require interfaces that let the user feel haptic feedback not only on his hand or fingers but also on other parts of the body. Examples like the *Battlefield 3 Simulator* show that the amount of projects that focus on more specific or non-manual haptic feedback will probably increase in the next years.

VEI with haptic feedback are used in the fields of medicine, molecular research, gaming, industrial design, land mine clearance, collaboration, art and music. Often, this technology is used for learning environments. It also can make scenarios of gaming, art and music more entertaining. The future evolution of VE in these areas depends amongst others on the evolution of adequate interfaces. Even though studies in the field of industrial design found no advantages of using the *PHANToM* station with haptic feedback for the design process [12, 32, 16], I think that research will still go on in this field. In the field of medicine, VE with haptic feedback are frequently used as learning environments. The gaming industry always aims to make games more realistic and, thus, more fun. This makes it to an interesting field for the future development of VEI with haptic feedback, too.

In the last part of this paper, information about the related field of augmented reality was given. Both VR and AR have its pros and cons. AR may provide more realistic effects in some cases and it can fully utilize conditions provided by the real world, like natural haptic feedback. VR may be less realistic in some applications, but does not necessarily require complex hardware or a specific real world environment. AR environments provide real haptic feedback for real world elements. In contrast, VR environments with haptic feedback devices can provide tactile or force feedback for virtual elements. An interesting evolution could be the combination of VEI with haptic feedback and AR technology. Therein, haptic feedback could be provided for both real world elements and virtual, computer-generated elements. The scenarios that were presented in this paper show that VR may be a good choice for some examples. In my opinion, the suitability of VR or AR depends on the specific scenario it shall be used for. One should think about financial feasibility and technical requirements of a scenario. Sometimes, a combination of both VR and AR can be a good choice.

In the end of December 2011, Microsoft Research presented its latest innovations on 3D-displays with the goal to make the vision of Star Trek's HoloDeck real. With the combination of sophisticated virtual environments and elaborate interfaces that provide realistic haptic feedback, the realisation of the HoloDeck will probably be possible at some point in the near future.

References

[1] S. Andrews, J. Mora, J. Lang, and W. Lee. Hapticast: A physically-based 3d game with haptic feedback. *Proceedings of FuturePlay*, 2006.

[2] R. Azuma et al. A survey of augmented reality. *Presence-Teleoperators and Virtual Environments*, 6(4):355–385, 1997.

[3] C. Basdogan, S. De, J. Kim, M. Muniyandi, H. Kim, and M. Srinivasan. Haptics in minimally invasive surgical simulation and training. *Computer Graphics and Applications, IEEE*, 24(2):56–64, 2004.

[4] C. Basdogan, C. Ho, and M. Srinivasan. Virtual environments for medical training: Graphical and haptic simulation of laparoscopic common bile duct exploration. *Mechatronics, IEEE/ASME Transactions on*, 6(3):269–285, 2001.

[5] M. Benali-Khoudja, M. Hafez, J. Alexandre, and A. Kheddar. Tactile interfaces: a state-of-the-art survey. In *Int. Symposium on Robotics*. Citeseer, 2004.

[6] S. Biggs and M. Srinivasan. Haptic interfaces. *Handbook of virtual environments*, pages 93–116, 2002.

[7] A. Bolopion, B. Cagneau, S. Redon, and S. Régnier. Comparing position and force control for interactive molecular simulators with haptic feedback. *Journal of Molecular Graphics and Modelling*, 29(2):280–289, 2010.

[8] S. Botden, S. Buzink, M. Schijven, and J. Jakimowicz. Augmented versus virtual reality laparoscopic simulation: What is the difference? *World journal of surgery*, 31(4):764–772, 2007.

[9] S. Botden and J. Jakimowicz. What is going on in augmented reality simulation in laparoscopic surgery? *Surgical endoscopy*, 23(8):1693–1700, 2009.

[10] F. Brooks Jr, M. Ouh-Young, J. Batter, and P. Jerome Kilpatrick. Project grope - haptic displays for scientific visualization. In *ACM SIGGRAPH Computer Graphics*, volume 24, pages 177–185. ACM, 1990.

[11] CAE. Cae healthcare - promis simulator. `http://www.cae.com/en/healthcare/promis.simulator.asp`, 2011. visited 05.12.2011.

[12] D. Cheshire, M. Evans, C. Dean, et al. Haptic modeling an alternative industrial design methodology. In *Proceedings of eurohaptics*, volume 2001, pages 124–128. University of Birmingham, UK, 2001.

[13] I. Corporation. Impulse engine. `http://www.immersion.com/`, 2011. visited 27.11.2011.

[14] B. Daunay and S. Régnier. Stable six degrees of freedom haptic feedback for flexible ligand–protein docking. *Computer-Aided Design*, 41(12):886–895, 2009.

[15] S. Ellis. What are virtual environments? *Computer Graphics and Applications, IEEE*, 14(1):17–22, 1994.

[16] M. Evans, D. Wallace, D. Cheshire, and B. Sener. An evaluation of haptic feedback modelling during industrial design practice. *Design Studies*, 26(5):487–508, 2005.

[17] Farlex. The free dictionary - medical dictionary. `http://www.thefreedictionary.com`, 2011. visited 31.12.2011.

[18] D. Gomez, G. Burdea, and N. Langrana. Integration of the rutgers master ii in a virtual reality simulation. In *Virtual Reality Annual International Symposium, 1995. Proceedings.*, pages 198–202. IEEE, 1995.

[19] ImmersiveTechnology. Virtual turntables with haptic feedback. `http://www.youtube.com/watch?v=IniGM8ASUkk`, 2011. visited 04.12.2011.

[20] S. Jayaram, U. Jayaram, Y. Wang, H. Tirumali, K. Lyons, and P. Hart. Vade: A virtual assembly design environment. *Computer Graphics and Applications, IEEE*, 19(6):44–50, 1999.

[21] M. Krueger. An easy entry artificial reality. *This early Snowscape world was generated with GenWorld. Here, an invisible cameraman agent has chosen the camera position and angle. Copyright© by Brian Gardner.*, page 147, 1986.

[22] M. Krueger. *Artificial reality II*. Addison-Wesley Reading (Ma), 1991.

[23] M. Krueger, T. Gionfriddo, and K. Hinrichsen. Videoplacean artificial reality. In *ACM SIGCHI Bulletin*, volume 16, pages 35–40. ACM, 1985.

[24] K. Kuchenbecker, J. Fiene, and G. Niemeyer. Improving contact realism through event-based haptic feedback. *Visualization and Computer Graphics, IEEE Transactions on*, 12(2):219–230, 2006.

[25] A. N. Laboratory. Argonne national laboratory. `http://www.anl.gov/`, 2011. visited 16.11.2011.

[26] C. . B. Ltd. Battlefield 3 simulator. `http://www.youtube.com/watch?v=eg8Bh5iI2WY&feature=youtu.be&t=11m40s`, 2011. visited 10.11.2011.

[27] Nintendo. Nintendo products. `http://www.nintendo.de`, 2011. visited 10.11.2011.

[28] W. Piekarski and B. Thomas. Arquake: the outdoor augmented reality gaming system. *Communications of the ACM*, 45(1):36–38, 2002.

[29] V. Popescu, G. Burdea, M. Bouzit, and V. Hentz. A virtual-reality-based telerehabilitation system with force feedback. *Information Technology in Biomedicine, IEEE Transactions on*, 4(1):45–51, 2000.

[30] E. Sallnäs, K. Rassmus-Gröhn, and C. Sjöström. Supporting presence in collaborative environments by haptic force feedback. *ACM Transactions*

on Computer-Human Interaction (TOCHI), 7(4):461–476, 2000.

[31] SEGA. Outrun. http://blamethecontrolpad.com/outrun/ORarcade/outrun.jpg, 2011. visited 19.12.2011.

[32] B. Sener, P. Wormald, and I. Campbell. Evaluating a haptic modelling system with industrial designers. In *proceedings of the EuroHaptics International Conference*, pages 8–10. Citeseer, 2002.

[33] J. Shaw. The legible city. http://www.jeffrey-shaw.net, 1989-1991. visited 10.11.2011.

[34] Simbionix. Lap mentor. http://simbionix.com/wp-content/uploads/2011/01/LAPBrochure10-2011-Web.pdf, 2011. visited 10.11.2011.

[35] Sony. Playstation. http://de.playstation.com/, 2011. visited 30.11.2011.

[36] J. Steuer. Defining virtual reality: Dimensions determining telepresence. *Journal of communication*, 42(4):73–93, 1992.

[37] R. Stone. Haptic feedback: A potted history, from telepresence to virtual reality. In *The First International Workshop on Haptic Human-Computer Interaction*, pages 1–7. Citeseer, 2000.

[38] E. Subasi and C. Basdogan. A new haptic interaction and visualization approach for rigid molecular docking in virtual environments. *Presence: Teleoperators and Virtual Environments*, 17(1):73–90, 2008.

[39] S. Tachi, H. Arai, and T. Maeda. Development of an anthropomorphic tele-existence slave robot. In *Proceedings of the International Conference on Advanced Mechatronics (ICAM)*, volume 385, page 390, 1989.

[40] H. Tan. Perceptual user interfaces: haptic interfaces. *Communications of the ACM*, 43(3):40–41, 2000.

[41] B. Thomas, B. Close, J. Donoghue, J. Squires, P. De Bondi, M. Morris, and W. Piekarski. Arquake: An outdoor/indoor augmented reality first person application. In *Wearable Computers, 2000. The Fourth International Symposium on*, pages 139–146. IEEE, 2000.

[42] B. Tolsdorff, A. Pommert, K. Höhne, A. Petersik, B. Pflesser, U. Tiede, and R. Leuwer. Virtual reality: A new paranasal sinus surgery simulator. *The Laryngoscope*, 120(2):420–426, 2010.

[43] A. Wexelblat. *Virtual reality: applications and explorations*. Academic Press Professional, 1993.

[44] M. Yiasemidou, D. Glassman, P. Vasas, S. Badiani, and B. Patel. Faster simulated laparoscopic cholecystectomy with haptic feedback technology. *Open Access Surgery*, 4:39–44, 2011.

[45] C. Youngblut. Review of virtual environment interface technology. Technical report, DTIC Document, 1996.

Discovering User Similarity in Social Media

Sylvia Kempe

Abstract— Social media websites like Facebook, LastFM or Flickr have millions of users, who share their private information with their friends, if not with the whole Internet. Recommender systems try to manage this huge amount of data. Thus, they help discovering user similarity and provide different functions based on these shared information. Clustering online friends into groups or suggesting new friends are just some of the features which are based on discovering user similarities. This paper gives an insight into the history of social media. Moreover, information revelation is discussed, in order to discover which information such recommender systems rely on. In the main section, a number of different approaches of previous research dealing with "user similarity in social media" are analyzed. The summary with regard to this topic gives a critical discussion of privacy in social media and supplies some ideas for improvements concerning user acceptance.

Index Terms—Social Media, Recommender Systems, Shared Information

1 Introduction

Today we couldn't imagine life without social media. Checking our mails at Facebook[1], composing a tweet on Twitter[2] , sharing links to music sites or even looking up a recipe for dinner on food blogs, all these things have made social media an important part of our life. Since the capacity of information on social media sites continues to grow, it is difficult for the user to keep a clear overview in which sites the user is interested and in which he is not [12].

Moreover, social media supports people in keeping in touch with old friends, in finding new friends with similar interests and in sharing opinions and ideas with others. Another provided support is the clustering of online friends to simplify the e-mail correspondence. Nevertheless, using these data is not always just in the user's interest. Collected user information enables customized advertising et cetera. With the ever-increasing number of online features provided by social media sites, the user is tempted to spend more and more time on these sites. For this reason, the user must also be sensitized as to which data he's willing to share and which data must be handled in a confidential way.

All in all the number of social media applications and the inclusion of users is rapidly growing. They can be separated in many different types, like social networking sites (Facebook or Myspace[3]), blogs (Twitter), wikis (Wikipedia[4]) and media sharing sites (Youtube[5], Flickr[6]), just to name a few [14]. Social media sites are based on user generated content and allow the user to share information with other people. Just by looking at the registration on a social media site, the user is seduced to divulge personal information in order to create an online profile [10]. This information is used to identify the user but also to customize applications for the user. Furthermore, the user leaves a virtual trace by leaving comments, tagging, bookmarking or making new friends. This data can be used as additional information about customer behavior. This is the point where recommender systems start to work and make use of all these information. Based on these data sets, the similarity between users is determined with the help of different techniques (see chapter 4). Because of the steadily increasing number of different application areas, there are more and more different algorithms to assimilate this growing data. All the collected information is used by social media to facilitate the handling, to support the user in finding interesting things and to connect people with each other. Websites are becoming more and more personalized and customized. A good example could be Amazons personalized recommendations.

Basically the use of these algorithms can be subdivided into two areas. *Familiarity* evidences and *similarity* evidences [11]. To find out if two people might know each other, familiarity evidences can be consulted. One clue of familiarity is for example being friends on a social network site or frequent mail traffic between people. Many studies related to this topic focus on buddy list weighting, people recommendation or tie strength prediction to discover these familiarity evidences.

More research is done on the calculation of similarity behavior/activities. This category emphasizes matching behavior patterns like bookmarking the same websites or knowing the same people. These clues can lead to shared interests of individuals who don't know each other. There are two ways of collecting information that can be used for these purposes. On the one hand, explicit information for instance obtained by directly adding new friends. On the other hand, implicit information which is revealed by the user's interaction [23].

To give some background knowledge of social media, chapter 2 concentrates on the history of social media and explains a concept of social media applications. In order to understand the development of information needed for recommender algorithms, section 3 is concerned with information revealing in social media. Chapter 4 introduces different approaches of previous research in the field of social media. A few different algorithms of recommender systems in social media are presented as well in the following sections. Finally, the discussion provides a critical insight into the privacy handling in order to reach the needed information and provides some suggestions to improve the user acceptance of such recommender systems.

2 History and Evolution of Social Media

The roots of social media can be traced back to 1979. Tom Truscott and Jim Allis from Duke University developed a system where internet users could discuss topics by sending public messages. This system was called "Usenet", a simple prototype of today's extensive Internet [14].

But social media as we understand it now has its origin most likely in a social network site called "Open Diary" which was developed by Bruce and Susan Abelson in 1998. They created a social network site, where online diary writers where connected with each other. At this time, the term "weblog" appeared for the first time, which became "Blog" one year later, when a blogger changed the word "weblog" into "we blog" [14].

The foundation of social network sites like Myspace in the year 2003 or Facebook in 2004 made the concept of social networking popular within the masses. Latest services in social media are the so called

• *Sylvia Kempe is studying Media Informatics at the University of Munich, Germany, E-mail: sylvia.kempe@campus.lmu.de*
• *This research paper was written for the Media Informatics Advanced Seminar 'Over Distance', 2011/2012*

[1] https://www.facebook.com/
[2] http://www.twitter.com/
[3] http://de.myspace.com/
[4] http://de.wikipedia.org/
[5] http://www.youtube.com/
[6] http://www.flickr.com/

		Social presence/ Media richness		
		Low	Medium	High
Self-presentation/ Self-disclosure	High	Blogs	Social networking sites (e.g., Facebook)	Virtual social worlds (e.g., Second Life)
	Low	Collaborative projects (e.g., Wikipedia)	Content communities (e.g., YouTube)	Virtual game worlds (e.g., World of Warcraft)

Table 1. Classification of social media in social presence/mediarichness and self-presentation/self-disclosure [14]

virtual worlds. The most famous of them is probably "Second life" [14]- a virtual world where users can create characters and interact with the characters of others.

The differences between the various online services are far extensive. However, most sites have one thing in common: They are based on the presentation of your own profile. The intention is that other individuals become aware of your personal data to get, for example, new friends, a new job or any recommendations of things they like [10].

The success of social media speaks for itself. In January 2009 the social network application "Facebook" boasted 175 million registered, active users. In the same year, 10 hours of video content per minute, were uploaded to the video sharing platform "Youtube" [14]. Only two years later in 2011 Facebook comprises 800 Million users, which corresponds to more than one-tenth of the global population.

The rising usage of social media and the consequent rise of available online data is misleading services to collect social network profiles and other personal information and create recommender systems.

Because there is a large amount of different social media types, Kaplan and Haenlein [14] created a classification scheme, based on different research.

Their classification is based on two main elements of social media: Social presence/media richness and self-presentation/self-disclosure. Table 1 shows the classification of the different areas of social media according to their distribution of self presentation/ self-disclosure and social presence/media richness. In social presence/media richness, social networks for example are rated higher than blogs because there is more personal communication between people. The degree of self-presentation is the same yet. In addition, some of the most popular social media applications are classified into these schemes.

Social presence theory [13], which was founded by Short Williams and Christie in 1976, claims that a medium's social impact can be measured on the degree of social presence. Social presence here denotes the feeling that other people are a party to the communication. It can be measured in varying degrees. Asynchronous communication like e-mail transfer is rated lower than synchronous communications e.g. live chat.

"The higher the social presence, the larger the social influence that the communication partners have on each other's behavior." [14] Similar to the concept of social presence, there is the concept of media richness. The theory of media richness was stated by Daft and Lengel [4]. They claim that the goal of "any communication is the resolution of ambiguity and the reduction of uncertainty." Media has a different measure of value for the wealth of information given by the users. This leads to resolving ambiguity and uncertainty. Probably the main reason why people reveal their personal information in social media is self-representation. Users create a digital identity to represent their personal unique identity. Moreover people classify themselves into different interest groups to show their social identity. This process commonly occurs with self-disclosure or rather conscious or unconscious revelation of personal data.

3 Information Revelation

As already mentioned in chapter 2, a large part of social media is based on self-representation, mostly in form of online profiles and sharing additional personal data for example by classifying friends into groups, sharing links, or liking movies. The more information the user is willing to reveal about him, the more social media features he can use. This brings the user to divulge more and more personal information. Also the fact that users increasingly identify themselves with their online profile makes a contribution to this. This enables recommender systems to use all available information unnoticeable [14].

4 Existing Approches

To give an insight into the various approaches of previous research, different algorithms are explained roughly in the following sections. A lot of research has dealt with the topic of information sharing in social media so far and has offered different ways to process user data.

The considered research findings are divided into two areas, in order to clarify the different operating conditions of this subject: There are Algorithms that deal with the degree of familiarity and algorithms that compare the similarity of users.

4.1 Familiarity Evidence

Most persons use social media applications to stay in contact with people they already know from their real life, for example family members, old friends or colleagues. For users with a large network it may be difficult to find certain persons. Moreover, it can be very time consuming to find all old friends from school. Algorithms working on a basis of familiarity evidences give advice whether two people know each other or not. Recommender systems on social network sites also take advantage of these familiarity evidences to suggest new friends to the user, who he might know. Previous research on this topic dealt with several different approaches, some of which are now explained in more detail.

4.1.1 Implicit Social Graph

Many social media sites permit the user to organize online friends into various groups. However, this supply is not used in many cases, because the progress is relatively time consuming. Sometimes this kind of feature is not even noticed by the user. Due to the fact that people mostly communicate with the same group of contacts, automated algorithms suggesting e-mail recipients would be useful and timesaving. Nevertheless, even if the user does not cluster his friends into groups like "family", "co-workers" et cetera, implicit groups are formed only by the user's interaction. Thereby, appropriate contacts easily could be added to the CC field of an e-mail.

The *implicit social graph* helps with grouping online contacts without manually clustering. The implicit social graph is composed of interactions with specific contacts differing from the explicit social graph, which is caused by the distinct befriending of people [23].

As a result of an experiment with Google employees, more than 40% of the sent e-mails was dispatched to multiple recipients. Even 10% of the e-mails were mailed to 5 or more recipients. Likewise, it turned out that individuals habitually communicate with the same groups. Since many people do not take the time to cluster their friends into groups, sending e-mails to multiple persons could be wasteful. The use of an automated algorithm in this area seems to make sense. The now proposed algorithm is based on the analysis of the implicit social graph. This involves interactions between contacts and groups of contacts.

The graph contains weighted edges, which arise from the frequency, recency and direction of these interactions. It is a non-content-sensitive algorithm. The edges of the graph are generated by sending and receiving messages. When a message is sent from a single user to a group of contacts, it is realized by a single edge. This creates a directed hypergraph. A hyper edge is called implicit group, even if it consists of a single contact. All incoming or outgoing edges of the node of a single user are called the user's egocentric network.

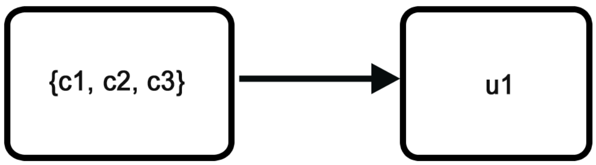

Fig. 1. Hyperedge of user u1. One incoming edge of the contacts c1, c2, c3 in the node of u1.

Each group represents a single implicit hyperedge, even if it consists only of a single user contact. Figure 1 shows the egocentric network of user u1, if contact c1 sends an e-mail to the user u1 and the contacts c2, and c3. It is represented as an incoming edge from the group c1, c2, c3 to u1. Continuative this could be complemented by weighting the hyperedge.

The edge weight is calculated from the following points:

Frequency Frequent contact between the user and a specific group is considered more important than rare contact with groups.

Recency Contact to a group that has taken place recently, are considered more important than older interactions. The result is a dynamic list of group importance.

Direction Sent mails are considered more important than user-perserved messages.

With the help of an interaction ranking and the three points mentioned above, the edge weight can be calculated. Interaction Rank ($\mathscr{I}\mathscr{R}$) can be estimated with a set of e-mail interactions I = { I_{out}, I_{in} }, considering the equation:

$$\mathscr{I}\mathscr{R} \leftarrow \omega_{out} \sum_{i \in I_{out}} \left(\frac{1}{2}\right)^{\frac{t_{now}-t(i)}{\lambda}} + \sum_{i \in I_{in}} \left(\frac{1}{2}\right)^{\frac{t_{now}-t(i)}{\lambda}}$$

The collected e-mail exchange between a user and a specific implicit group can weight the interaction by its recency. The interaction weight of an edge declines exponentially over time with the half-life, λ, as a tunable parameter. ω_{out} is also a parameter which can affect the importance of outgoing versus incoming e-mails (Direction).

I_{out} represents the set of outgoing edges between a user and a group of contacts and I_{in} indicates a set of incoming interactions, t_{now} is the current time compared to t(i) which stands for the timestamp of the specific interaction i $\in I$.

Interactions Rank weights the hyperedges according to their timestamps (Recency) and thereby regards the frequency by summing up these values.

The result is an indicator of the strength of a relationship between the user and a certain group. Combining the user's egocentric network one gets the global sociocentric network. Even though other research in this area works with friend-of-friend edges to cluster friends, Roth et al. confine theirselves to use only the egocentric network of the user. In this way they try to protect the user's privacy and try to prevent the user from extraneous information. In their study, Roth et al. use only the following metadata: timestamp, sender and recipient of the incoming and outgoing mails. The mail content was ignored.

4.1.2 Tie strength predicion

Most social media applications offer its users the concept of befriending. Because of this one could conclude that you either can be a friend or a stranger to another person. It is not taken into account that not all relationships are equal. For this purpose, relationships can be measured in tie strength.

Gilbert and Karahalios [6] try to map social media data to tie strength, because relationships in social media can differ. Using a data set of 2.000 Facebook friendships, evaluated according to tie strength and tagged with more than 70 numeric evidences, they try to distinguish between weak and strong ties. Casual acquaintanceships, which can help for example in finding a job, are referred to **weak ties** in this context. Close friends and family fall into the category **strong ties**. Wheareas the communication of weak ties mostly is based on a few ordinary media, the interaction with strong ties mostly takes place across multiple channels.

The concept of tie strength was presented in 1973 by Mark Granovetter in his paper "The strength of weak ties" [9]. Granovetter defines the strength of a tie as a combination of several factors. Here the lapse of time, familiarity, emotional intensity and the reciprocal services play an important role for the calculation of tie strength.

Granovetter's resulting factors to calculate tie strength :
Amount of time, intimacy, intensity, reciprocal services.

Subsequent work extended this list with the following points:
Structural factors (like network topology) [3], providing emotional support[26], social distance[20].

In theory, there are these seven criteria as a result. In practice, the following factors have become prevalent: communication reciprocity [5], possessing at least one mutual friend [25], recency of communication [18] and interaction frequency [7].

The study, conducted by Gilbert and Karahalios [6] with 35 participants and the help of the Mozilla plug-in Greasemonkey[7], tries to prove that tie strength can be predicted with the help of social media data. The Greasemonkey script randomly selected Facebook friends of the participants and integrated five questions according to tie strength into each of the friend's profile. The subjects had to work through as many prepared profiles as possible during a thirty minute session. Hence, a dataset with 2,184 rated Facebook friendships resulted from this study (on average 62.4 friends per participants). During this test, a script automatically gathered data on all Facebook interactions between the subject and the corresponding friends to compare this information with the results of the survey. The collected data was analyzed according to the previously mentioned seven criteria.

Just to gives a little insight into the 70 used variables, some of them are explained in the following: The Facebook wall stands for a public communication channel which is mostly only available for online friends. *Wall words exchanged* indicates the number of words interchanged between the user and her online friends. *Inbox messages exchanged* counts the occurrences of a friend in the participant's Facebook Inbox, a private communication channel. The category "Intimacy Variables" covers variables which needed content analysis [22]. Therefore, Linguistic Inquiry and Word Counts (LIWC) dictionary were used. *Wall intimacy words* compare the number of Wall words matching at least one of the eleven following LIWC categories: Family, Friends, Home, Sexual, Swears, Work, Leisure, Money, Body, Religion and Health. Equally, *Inbox Intimacy Words* counts the number of Inbox words, fitting at least one of these sections. *Days since last communication* observes the recency of written communication in Facebook channels like Wall, Inbox, photo comments.

[7]http://www.greasemonkey.net/

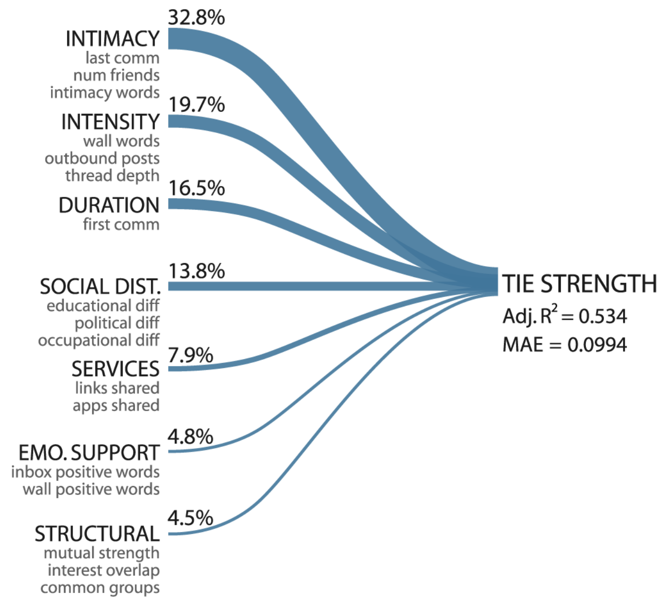

Fig. 2. The seven dimensions of predicing tie strength in the "How strong?" model [6].

The variable *Days since first communication* gives a clue regarding the length of the friendship. *Links exchanged by wall post* measures the number of URL exchanges between friends, which is a common Facebook practice. Other variables are *Groups in common*, *Number of mutual friends*, *Age difference* and many more.

According to the first tie strength question "How strong is your relationship with this person?" the model provides good results: Adj. R^2 = 0.534, p < 0.001 and a Mean Absolute Error of 0.0994% on a constant 0-1 scale ,which means an average model calculates tie strength within one-tenth of its true value.

Figure 2 illustrates the power of prediction according to the seven tie strength dimensions as part of the *How strong?* model. In addition, the figure shows the three best performing variables of each dimension. Summing up the coefficients of the associated variables, the weight of a dimension can be calculated. It is recognizable that no dimension has a absolute monopoly on tie strength, even if all dimensions are variably distributed. The result of this study shows that social media can predict tie strength with 85% accuracy [6].

4.1.3 CC Prediction with Graphical Models

Pal and McCallum [21] deal with the topic of how to propose additional e-mail recipients to users with the help of an automated system. For example, if an employee working on a certain project forgets to add an important team member to CC on an e-mail, an automated system suggesting appropriate recipients would be useful. Pal and McCallum use a graphical model, to investigate words in the body of an e-mail as well as the subject line and the recipients of the e-mail.

As a basis they used a multinomial naive Bayes model [19][15]. This is a procedure that can automatically rate texts by assigning labels from predefined label sets to the document. The multinomial naive Bayes model is a special version of the naive Bayes. It is based on the occurrence of a predefined set of words that represent the document. Here, the order of the occurrences is irrelevant; however, the total number of each occurrence is recorded.

For each mail in the user's sent folder, every receiver was used as a target label y. The variable n represents the number of words that are observed. The model was used to calculate the distribution over y and it provides a list that is sorted according to their probabilities. Graphical models are built upon these results and provide a more detailed structure of the e-mail.

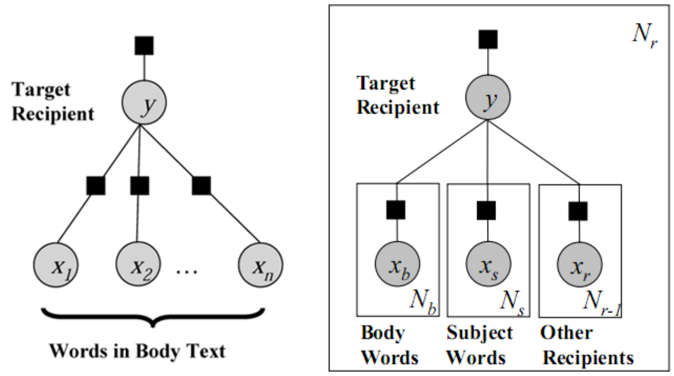

Fig. 3. (Left) Factor graph for naive Bayes model (Right) Plated factor graph for a naive model containing different alphabets for words in the body of an e-mail, words in the subject line and for recipients [21].

Figure 3 (Left) represents a classic naive Bayes document model with n remarks of the discrete random variables x, i= 1 . . . n standing for each word in the document using factor graph notation [16]. The number of words n can change for different e-mails.

Figure 3 (Right) shows the structure of the model using plated factor graph notation. A distinction is made between recipients, the words in the body of the mail and the subject words. N_b stands for the words in the body of the e-mail, whereas N_s represents the words in the subject line. N_{r-1} labels the residual recipients of the e-mail. The e-mail is duplicated for each of the N_r receivers. No difference was made between recipients of the TO and the CC field, as previous research has shown that the benefits of this strategy are negligible.

A combination of factor graphs and sheet or plate notation [8] is suggested for illustrating these types of models. In most cases plates help to explain reproduced variables in Bayesian networks. With the use of plated factor graphs, a mix of directed and undirected graphs can clearly be presented.

As a result of the study, Pal and McCallum [21] discovered the importance of adding co-recipient information to their calculations. Table 2 shows the effect of adding the co-recipients. An average daily increase from .364 to .448 confirms the positive effect of working with factor graphs.

Due to this information the performance of CC prediction could be increased. Furthermore, the graphical model presented by Pal and McCallum helps to predict e-mail recipients, CC and BCC recipients adding co-recipient information to their calculations. Due to this information the performance of CC prediction could be increased. Furthermore, the graphical model presented by Pal and McCallum helps to predict e-mail recipients, CC and BCC recipients.

Model	First Month	Last Month	Avg. Daily
Naive Bayes	.301	.326	.364
Factor Graph	.364	.395	**.448**
Thread Info	.357	.403	.448

Table 2. A comparison implemented by Pal et al.[21] to measure the differences between naive Bayes models and plated factor graph models.

4.2 Similarity Evidence

An additional feature of online services is the connection of people with similar interests. Internet users could share the same bookmarks, the same tags or the same circle of friends even without knowing each other. To extract these clues for similarity, previous research has presented different approaches.

4.2.1 Discovering Shared Interests using Graph Analysis

Schwartz and Wood [24] set their sights on locating people with similar interests or expertise. Usually, lists consisting of interest groups are generated from collected data. However, this entails the need of knowing which groups should be formed and who should be added to these lists. Schwartz and Wood present a technique that provides a more accurate and dynamically adaptive assumption of shared interests.

Analyzing subject lines and bodies of e-mails would cause an obvious privacy threat and would require difficult natural language recognition studies in addition. Therefore, Schwartz and Wood decided to address the problem by testing the structure of the graph, resulting from the e-mail fields ”From:To:”, using different heuristic algorithms to find common interests between people. They tested the algorithms on data collected from 15 websites around the world within two months. The resulting graph contained nearly 50.000 people in 3.700 different sites world-wide.

Collecting the history of electronic mail communication, shared interests between different persons could be ascertained. With the help of the participants of the study, networks were created that include individuals with similar interests. Based on a clustering relationship called Specialization Subgraph (SSG), Schwartz and Wood try to simulate these resulting networks of colleagues with shared interests. An SSG of a communication graph contains nodes (people sharing a common interest) and edges (communication). So a person can be part of many different SSGs due to her or his different interests and responsibilities. The goal of the algorithms presented by Schwartz and Woods is to infer SSGs regarding a special interest from a communication graph.

Starting with one person, the first algorithm isolates a list of other people some of which have shared interests with the target person. From this little group, whose individuals are certain to have shared interests, they create a list of individuals that are scored more highly. An Aggregate Specialization Graph (ASG), which contains several SSGs, associates topics of high interest with a special person. A second algorithm determines the people who share that certain interest. These algorithms allow them to discover shared interests successfully on collected personal data.

The results of Schwartz and Wood show that the SSG algorithm is a very powerful technique and could be useful for social media platforms allowing them to monitor traffic and customize their advertisement to special target groups Therefore the SSG approach should be used with care and only with the consent of the user involved.

4.2.2 Tag based social interest discovery

To discover similar interests between users and groups of users, Li et al. [17] work with a concept based on user-generated tags. This principle benefits from the fact that user-generated tags reflect the user’s interests and the web content he is attached to. In social media, people share a large amount of information like web blogs, bookmarks, photographs, music and many more. Discovering these shared interests is a challenging task considering the size of social communities. Li et al. make use of user generated tags to detect shared social interest. In their study, they used data from the online service del.icio.us[8] where users have the possibility to label the content they are interested in or want to share with other individuals. By analyzing these tags, a specific user vocabulary can be assembled, reflecting the user’s interests.

They have processed 4.3 million tagged bookmarks that where tagged by 0.2 million users on 1.4 million URLs, omitting stopwords in the user’s tags and the applied websites. After filtering the remaining keywords, they normalized them for the research.

To check whether the utilization of tags is corresponding with social interest, Li et al. had to analyze the user generated vocabulary. Subsequently they had to compare it with the keyword vocabulary of the website. If the most important words of a web document are covered by the user generated vocabulary this document is considered relevant.

The study of Li et. also reveals that user generated tags describe the content of a website even more concisely and detailed. Hence they have proven that tags can be used to describe the content of a websites and therefore to represent the user’s interest. Consequently sets of tags often used by many persons unite them to a kind of community of interest. In that case, these tags represent the user’s interest by general terms while the tagged URLs display the web content the user is focusing on.

Li et al. suggested the following architecture for Internet social interest discovery (ISID):

Find topics of interests Search all topics of interests based on a given set of bookmarks. Each topic of interests represents a set of tags plus the number of their occurences

Clustering For each topic, a user cluster and a URL cluster should be generated containing all involved users and all URLs.

Indexing The processed topic, user and URL clusters should now be imported into an indexing system for application queries.

On the one hand, Li et al. show that the user is interested in the content of a webpage if he repeatedly bookmarks pages with the same keywords. On the other hand they prove that in most cases user generated tags capture the content of a website.

5 Discussion

Since the use of social media is becoming more and more popular, the variety of research concerning social media is increasing as well. A rising number of algorithms are increasing the functional possibilities of social services such as social networking sites. Nevertheless, the opinions concerning user acceptance differ. Functions such as the suggestion of new friends or the presentation of subjects you might like, facilitate the handling of information overload of the internet.

On the other hand, it is more and more confusing what kind of personal data will be used by the operators and for what purposes [2]. Algorithms that are invisible to the user determine which information should be passed on to us. A balanced flow of information is no longer available, because almost every page tries to customize their information to our needs. Perhaps it would be useful to involve the user in this process. Algorithms would be a bit more transparent, if the user could decide what information should be used by recommender systems.

5.1 Privacy

The issue of visibility of information is one of the main problems accruing when talking about social media. It can’t be taken for granted that visible information is treated in a confidential way [1]. There might be a situation where you want to make special information visible for only a certain group of friends, for example, best friends (strong ties) and other information only for work colleagues (weak ties). Therefore it must be possible to distinguish several levels of privacy in social media. Information about the visibility of personal data is generally difficult to detect. In addition the variable handling of each different application makes it difficult for the user to keep track of his digital information.

Furthermore, not every user is aware of the visibility of his data. Without giving it serious consideration, they publish personal information on social media sites [10]. Here, the level of identifiability plays a certain role. The more the user identifies himself with his online profile, the more information he is willing to reveal. For example, previous business concerns or schools are listed online to stay in contact with old colleagues or friends. Most users are not aware that their data is not only revealed to other users, but also to the operator of the website. That website can now use the published information, for example, to make recommender systems more precise when proposing new friends. The same process appears for example at Amazon when a user orders a book. The recommender system proposes him other books he might like.

Table 3 shows the results of a study carried out by Gross and Acquisti [10]. It visualizes the percentage of Carnegie Mellon University (CMU) profiles that give away different kind of information. In general, CMU users publish a lot of private information in their online

[8] http://www.del.icio.us/

profiles. 90,8% of the users put an image online. Favorite music, books and movies are released each by over 60% of the study participants.

All provided information can be utilized by recommender systems in order to suggest better matches. However the boundary between privacy and social media is still a prevailing topic. To endure for example in the German market with its strong data protection laws, a lot of social media applications will have to improve their protection of entrusted information.

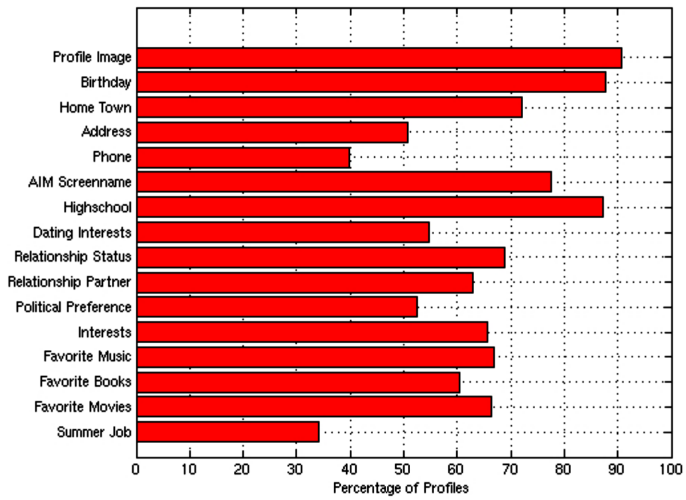

Table 3. Percentage of CMU profiles revealing various types of personal information [10]

5.2 User Acceptance

A large part of the research tries to improve the algorithms; however, the design of such recommender systems should also be improved. Just a few social media sites integrate the source information of these proposing functions. This could help users understanding these proposals and make it easier for them to decide whether to accept them or dismiss them. A lot of research deals with the investigation of new algorithms, but there is very little analysis of design and information protection. This could potentially improve the user acceptance.

The transparency of social networks should be increased, so users can understand what actually happens with their data and subsequently change their privacy options where applicable. Social media applications also have to face the problem of data security and improve the protection of entrusted information. At last, they have to take data protection laws of different countries into account, in order to still be able to interact at these markets.

REFERENCES

[1] The dark side of social media and privacy. http://www.markevanstech.com/2010/02/21/the-dark-side-of-social-media-and-privacy/. Last accessed on 2012-23-01.

[2] Should we care about privacy? http://www.markevanstech.com/2010/02/28/should-we-care-about-privacy/. Last accessed on 2012-23-01.

[3] R. S. Burt. *Structural holes: The social structure of competition.* Harvard University Press, 1992.

[4] R. L. Daft and R. H. Lengel. Organizational information requirements, media richness and structural design. *Manage. Sci.*, 32:554–571, May 1986.

[5] N. Friedkin. A test of structural features of Granovetter's strength of weak ties theory. *Social Networks*, 2(4):411–422, 1980.

[6] E. Gilbert and K. Karahalios. Predicting Tie Strength With Social Media. *Group*, 2009.

[7] E. Gilbert, K. Karahalios, and C. Sandvig. The network in the garden: an empirical analysis of social media in rural life. In *Proceeding of the twenty-sixth annual SIGCHI conference on Human factors in computing systems*, pages 1603–1612. ACM, 2008.

[8] W. Gilks and D. Spiegelhalter. *Markov chain Monte Carlo in practice*. Chapman & Hall/CRC, 1996.

[9] M. Granovetter. The Strength of Weak Ties. *The American Journal of Sociology*, 78(6):1360–1380, 1973.

[10] R. Gross, A. Acquisti, and H. J. H. Iii. Information Revelation and Privacy in Online Social Networks (The Facebook case). *Human Factors*, 2005.

[11] I. Guy, M. Jacovi, A. Perer, I. Ronen, E. Uziel, and M. Carmel. Same Places , Same Things , Same People ? Mining User Similarity on Social Media. *Media*, pages 41–50, 2010.

[12] I. Guy, N. Zwerdling, I. Ronen, D. Carmel, and E. Uziel. Social Media Recommendation based on People and Tags. *Evaluation*, (Lc):194–201, 2010.

[13] B. C. J.Short, E. Williams. *The social psychology of telecommunications*. 1976.

[14] A. M. Kaplan and M. Haenlein. Users of the world, unite! The challenges and opportunities of Social Media. *Business Horizons*, 53(1):59–68, Jan. 2010.

[15] A. M. Kibriya, E. Frank, B. Pfahringer, and G. Holmes. Multinomial naive bayes for text categorization revisited.

[16] F. Kschischang, B. Frey, and H.-a. Loeliger. Factor graphs and the sum-product algorithm. *Information Theory, IEEE Transactions on*, 47(2):498–519, 2001.

[17] X. Li, L. Guo, and Y. Zhao. Tag-based social interest discovery. 2008.

[18] N. Lin, P. Dayton, and P. Reenwald. Analyzing the Instrumental Use of Relations in the Context of Social Structure. *Sociological Methods and Research*, 7(2):149–166, 1978.

[19] A. McCallum and K. Nigam. A comparison of event models for naive bayes text classification, 1998.

[20] J. C. . V. Nan Lin , Walter M . Ensel. Social Resources and Strength of Ties : Structural Factors in Occupational Status Attainment. *History*, 46(4):393–405, 2010.

[21] C. Pal and A. Mccallum. CC Prediction with Graphical Models. *Text*, pages 4–6, 2004.

[22] J. Pennebaker and M. Francis. Linguistic inquiry and word count: LIWC 2001. *Mahway: Lawrence*, 2001.

[23] M. Roth, G. Flysher, Y. Matias, A. Leichtberg, and R. Merom. Suggesting Friends Using the Implicit Social Graph Categories and Subject Descriptors. *Human Factors*.

[24] M. F. Schwartz and D. C. M. Wood. Discovering Shared Interests Among People Using Graph Analysis of Global Electronic Mail Traffic 2 . Specialization Subgraph Structure. *System*, 1992.

[25] X. Shi, L. Adamic, and M. Strauss. Networks of strong ties. *Physica A: Statistical Mechanics and its Applications*, 378(1):33–47, 2007.

[26] B. Wellman and S. Wortley. Different strokes from different folks: Community ties and social support. *American Journal of Sociology*, 96(3):558–588, 1990.

Interaction Over Distance On Large Displays

Bertram Schmitt

Abstract— Nowadays, using large displays is a common way to provide people with various kinds of information. When people work collaboratively with large amounts of information, a typical way to view and organize them is using large, wall-sized displays. However, interacting with them is still an issue and a major subject of current research. Typical problems are that parts of the display might not be reachable for the user, that she cannot see all parts of the display as she stands right in front of it or that she has to interact with the display from a distance. Many of the existing interaction techniques do not provide proper solutions for these problems that occur while interacting with large displays. Typical situations for the usage of large displays are control or meeting rooms, school, university or public places like airports. This paper gives an overview of existing interaction techniques and the problems they are trying to solve. The techniques are divided into software- and hardware-based approaches and include a wide range of possible solutions to the problems mentioned before. These methods are then critically discussed to find out whether they can provide solutions to the problems they are trying to overcome. The findings of the discussion show that the analyzed techniques are either not able to solve all problems or that they can not achieve this without introducing new problems.

Index Terms—Large Displays, Wall Displays, Tabletops, Interaction Techniques, Over Distance

1 INTRODUCTION

As a result of displays constantly growing in size while decreasing in price, the availability of large displays has increased during the last years [20]. Applications for large displays include a wide range, reaching from the usage in control or meeting rooms to applications in school, university or public places like shopping malls or airports [2]. When people work collaboratively with large amounts of information, a typical way to view and organize them is using large, wall-sized displays [9]. However, interacting with these displays is often difficult [16]. Typical problems are that parts of the display might not be reachable for the user, that she cannot see all parts of the display as she stands right in front of it or that she has to interact with the display from a distance. Additionally, large displays often do not provide a way to interact with them either because they might be non-interactive or simply lacking a proper interaction technique. Therefore, providing intuitive and user-friendly interaction techniques with large displays is an important issue in today's research.

In this paper, large displays are defined as displays that, due to their size and the context they are used in, need different techniques to provide a proper way of interaction compared to conventional approaches like mouse or touch input. Examples for this kind of displays are wall-size displays, projectors, video walls and tabletops.

This paper gives an overview of existing interaction techniques for large displays. The techniques are divided into software- and hardware-based approaches and include a wide range of possible solutions to the problems mentioned before. The software-based solutions will be presented in chapter 2 and the hardware-based in chapter 3. After that, the advantages and disadvantages of these techniques will be discussed and the three major problems while interacting with large displays are extracted. The paper then evaluates whether the presented techniques could overcome these problems. The last chapter gives a summary of the results of this paper and states possible topics for future research.

2 SOFTWARE-BASED INTERACTION TECHNIQUES

In this chapter, interaction techniques that only use software-based approaches are presented. They use different methods to solve the problems that exist while interacting with large-scale displays. They range from adaptions of common interaction principles like drag-and-drop to the design and implementation of new interaction metaphors using widgets.

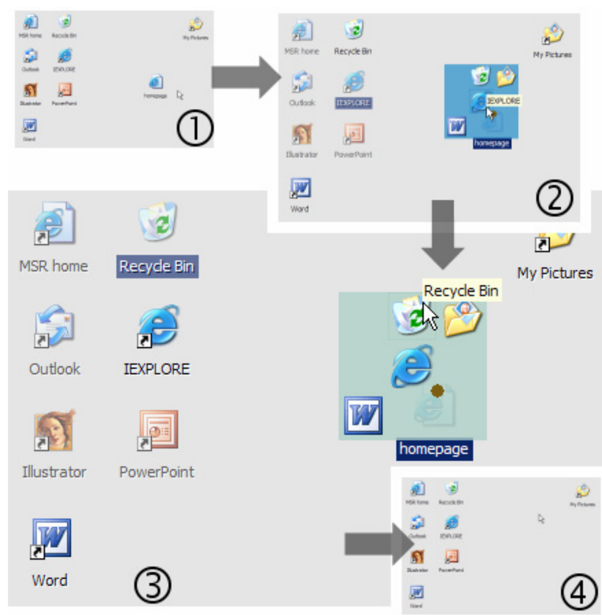

Fig. 1. The four steps of a *push-and-pop* [7] sequence: (1) The sequence is started by the user dragging a document. (2) The icons of compatible applications are displayed near the document. (3) The user drags the document over the recycle bin, which gets highlighted to show its selection. (4) The user releases the document and it is moved into the recycle bin.

2.1 Push-and-pop

Push-and-pop [7] is an interaction technique that addresses the problem, that users might have difficulties to reach all parts of the display on wall-size displays with touch or pen input. While interacting on a large display with regular drag-and-drop, the user might face the situation that she wants to drag content from or to a location that is inaccessible for her. *Push-and-pop* tries to solve this problem. It is a combination of the previously presented techniques *push-and-throw* [10], a method that allows the user to extend her reach by offering the possibility to "throw" information to other parts of the display, and *drag-and-pop* [3], which brings potential targets near to the user's cur-

- *Bertram Schmitt is studying Media Informatics at the University of Munich, Germany. E-mail: schmittbe@cip.ifi.lmu.de*
- *This research paper was written for the Media Informatics Advanced Seminar 'Over Distance', 2011/2012*

rent location [7].

Figure 1 shows an exemplary walkthrough of a *push-and-pop* sequence. In this example the user wants to drag the word document on the right side into the recycle bin. The sequence begins when the user starts dragging the word document (step 1). The system responds by surrounding the cursor with the *take-off area* – a miniature version of the wall-size display showing the icons of applications the currently selected document can be combined with (step 2). The user then drags the document over the recycle bin, which gets highlighted by a colored frame to show that it has been selected (step 3). Finally, the user releases the document and it disappears in the recycle bin (step 4) [7].

To evaluate their method, the authors conducted a user study and compared *push-and-pop* to five other adaptions of regular drag-and-drop (including *pick-and-drop*, *push-and-throw*, *accelerated push-and-throw* and *drag-and-pop*) regarding interaction with large displays. The task of the participants was similar to the one shown in figure 1. The study results showed that the lowest task completion time could be achieved using *push-and-pop*, with *drag-and-pop* being only slightly slower. Conventional drag-and-drop also worked well as long as the target was located near the user's position. However, with an increasing distance between an item and the target, the completion time increased dramatically. In terms of error rates, all of the tested techniques performed well. After the study, the users were asked to rank all six techniques according to how satisfied they were with the respective technique. Consistent with the fastest completion time, the participants ranked *push-and-pop* as their favorite method [7].

2.2 I-Grabber

Another approach to solve the problem of interacting with objects that are out of the user's reach is the *I-Grabber* [1]. The *I-Grabber* is a multi-touch "interactive grabber" that acts as a virtual hand extension and allows users to manipulate any object from their current location. As opposed to the *push-and-pop* method, the *I-Grabber* was designed for large tabletop interfaces and especially for situations in which multiple users work on the tabletop simultaneously.

The *I-Grabber* is based on the idea that in a multi-user environment the users typically divide the workspace into territories, like a personal and a common workspace. The personal workspace is limited by the user's physical reachability, making it impossible to reach content outside of her territory without physically moving, for example standing up or walking around the tabletop. By doing so, the user could reach the desired content, but might also bother other users by disturbing their physical space. The *I-Grabber* uses a grabbing metaphor and allows the user to reach out for both reachable and unreachable objects while maintaining her current position without disturbing the work of other users [1].

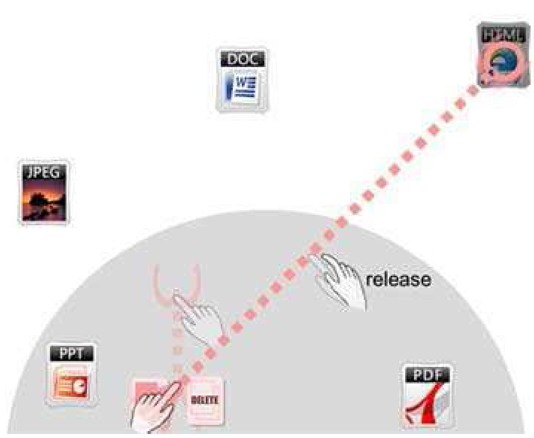

Fig. 2. The I-Grabber reaching out for a distant object [1]

Figure 2 shows how the *I-Grabber* can be used to access data that is out of the user's reach. Whenever the user places two fingers at a distance of 20 cm on the tabletop, two points (in the following referred to as P1 and P2) are created and the *I-Grabber* is initialized. It is composed of three parts, the holder, the stick and the hand. P1 defines the location of the holder and P2 defines the location of the hand as well as the length of the stick. By increasing the distance between these two points, the user can determine the length of the stick, that is the desired reach of the *I-Grabber*. When the hand is located over an object, it will be highlighted. When the user releases P2, the object will be selected and can now be moved to another location or opened in the user's personal territory. Furthermore, additional options like "copy" or "delete" are shown near P1 to provide fast access to often used commands.

Unfortunately, the authors did not conduct a user study to evaluate the performance of the *I-Grabber* compared to similar techniques. However, the *I-Grabber* was implemented in a multi-display environment, making it possible to manipulate data on remote screens. The authors also mentioned the possibility to use the *I-Grabber* on a handheld device to create a controller for distant large displays [1].

2.3 Vacuum

The *Vacuum* [4] is an interaction technique that provides access to items on large displays that are difficult to reach without the user having to move physically. The idea is similar to the approach of *drag-and-pop/pick* [3], bringing non-reachable items to the user. Figure 3 shows the structure of the *Vacuum*. It consists of a circular widget with a bull's-eye in the center and the so-called *arc of influence*. A proxy of every object within this arc is displayed near the center of the widget to allow easy interaction. The angle of this arc can be controlled by the user in order to select those objects of which a proxy should be created. Around the widget, a buffer zone is defined which contains objects that already lie within the user's reach. This prevents the creation of proxies for objects that are already near to the user's location. The proxies are displayed as shrunken versions of the original objects to allow more proxies to be represented near the vacuum without overlapping [4].

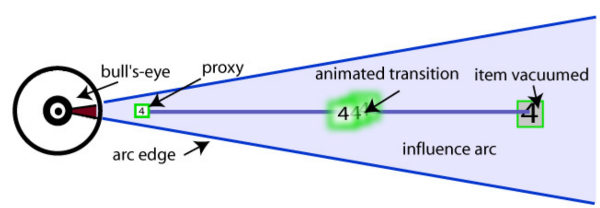

Fig. 3. The design of the *Vacuum* [4]

The authors used *drag-and-pop/pick* in early design explorations and found out that users were unaware of the area of the display that is influenced by this method. This led to the design of the arc and the principle of giving the user constant feedback by coloring the area of the arc's influence. Interaction with the proxies does not differ from the original objects. This means, for example, that dropping an object on a proxy has the same effect as dropping it on the original object [4].

Two user studies were conducted to compare the performance of the *Vacuum* to the *drag-and-pick* technique and direct picking without any assistance on a large display. For the first study, the participants' task was to select a discrete target from a collection of targets acting as distracter targets. The task of the second study was similar to the first one, except that in this case the three techniques were examined in the context of selecting multiple targets in sequence. The results indicated that the *Vacuum* performs similarly to *drag-and-pick* and direct picking in single target selection tasks, except for the case with many distracter targets over a large distance, where *drag-and-pick* performed better. However, in the case of selecting multiple targets, the *Vacuum* performed significantly better than the other two techniques [4].

2.4 Frisbee

A similar approach to the idea of using widgets to provide better usability of large displays is *Frisbee* [12]. The main goal of this approach is to provide an interaction method that uses direct manipulation techniques while minimizing the distance the user has to move in the physical space. The *Frisbee* technique consists of two circular GUI components, the *telescope* and the *target*. The idea is to provide a portal to another part of the display to allow the user to see and manipulate distant data. Furthermore, the target serves as a way to allow other viewers to follow the actions of the user manipulating the data. This is important as the user might be interacting with the display while not standing physically close to the data. The user can also use the telescope to move the target to the desired area of the display, as shown in figure 4. When the user is manipulating the distant data through the target, the input events are performed as if they occurred at the remote location. Furthermore, zooming operations on the target can be performed via the telescope [12].

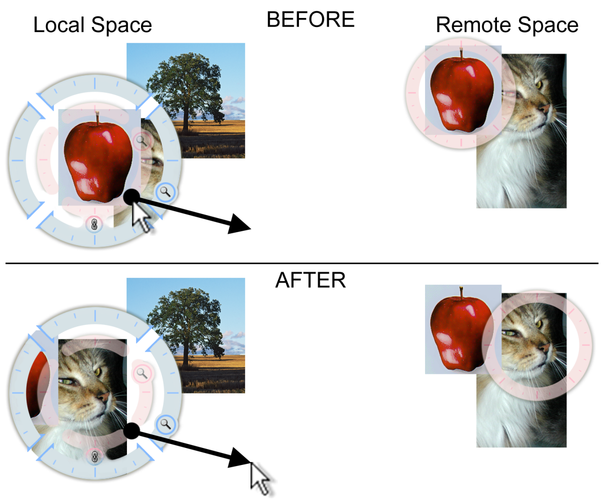

Fig. 4. *Frisbee*: the *target* (right) can be moved using the target position controls on the *telescope* (left) [12]

A user study was conducted to evaluate the efficiency of *Frisbee*. The participants were asked to perform a task in which they should move a set of blocks that were spread across two displays. They had to alternately perform a task on each of the screens, comparing the two conditions *frisbee* and *walking*. Whereas *frisbee* used the *Frisbee* technique to interact on the remote display, the *walking* condition needed the participants to physically move between the two displays to perform the next action. The results showed that the *Frisbee* method performed better than physically walking between the two displays.

3 Hardware-based Interaction Techniques

In this chapter, methods that need additional hardware to provide interaction are presented. The interaction techniques are further divided into approaches that use any kind of motion-tracking to recognize the user's movements, as well as techniques that use handheld devices like mobile phones to provide ways of interaction. The paper focuses on these two categories, as the spread of applications using motion-tracking or gesture-recognition and smartphones is currently increasing. Therefore, these techniques might become more important in the future.

3.1 Interaction Techniques Using Motion-Tracking

The following section presents techniques that use any kind of motion-tracking to provide interaction with large displays. These approaches have in common that they use different types of sensors and cameras to compute a 3D representation of the user which serves as a base to determine the user's actions.

3.1.1 Shadow Reaching

Shadow Reaching [18] is an interaction technique that uses a shadow representation of the user to allow interaction with a display over large distances. It tries to overcome two main problems that exist during interaction with large wall displays: the user may not be able to reach all areas of the display and the user's interactions with the display should be recognizable for collaborators. Figure 5 illustrates how *Shadow Reaching* takes advantage of a perspective-based transformation of the user's shadow. By varying the distance between herself and the light source, the user can determine which areas of the display are in her reach. Furthermore, both user and collaborators have no problems following the user's interactions due to the use of the shadow as a pointing technique [18].

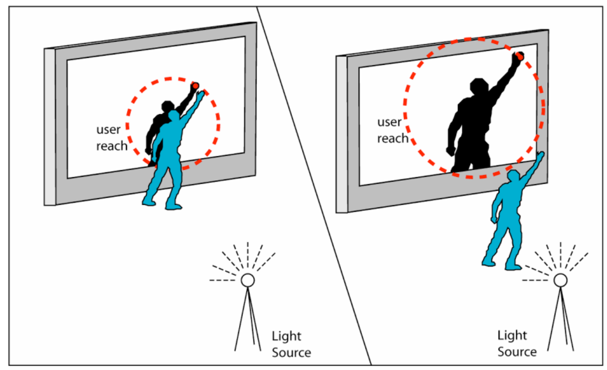

Fig. 5. *Shadow Reaching*: The user can determine which areas she wants to reach by increasing or decreasing the distance between herself and the light source [18].

The idea of *Shadow Reaching* is based on *VIDEOPLACE* [13], an artificial reality system which uses the projection of a user's silhouette on a large screen to interact with an artificial environment. Three different applications were implemented, each based on a different interaction metaphor. The first application serves as a replacement for a mouse cursor to allow pointing and interacting with the wall display, the second is based on the full-body interaction metaphor introduced by *VIDEOPLACE*, and the third extends the idea of *Magic Lenses* [5] to *Magic Shadows* [18]. *Magic Lenses* are movable see-through windows which are used to filter on-screen data, for example to remove the street names when moving the *Magic Lens* over a map.

The hardware that is used by *Shadow Reaching* differs between the three applications. To determine the position of the user's hand, the first implementation uses the known geometry of the light source and the display as well as the data from a handheld device the user has to hold while interacting with the display. This handheld device contains a position sensor to determine the hand's location and allows the user to "click" on the display by pushing a button on the handheld. The other two implementations use a light source behind the screen to accomplish shadow sensing. The light is captured with an infrared camera in front of the screen and is then processed by computer vision techniques to compute the user's location [18].

3.1.2 Combining Touch and Pointing

Another system that presents a solution to the problem that not every point is reachable when using touch input on large displays was presented by Schick et al. [17]. The idea is to extend the intuitivity of touch by allowing the user to point at non-reachable parts of the display to interact with them. The method is based on a 3D reconstruction of the user's body using standard RGB cameras which are placed freely around the display. This means that no modifications to the display are required and every vertical surface can be provided with touch and point interaction.

The main goal of Schick et al.'s approach is to implement a system that does not require the user to change the way she interacts with the system depending on whether she is touching or pointing at the display. To allow this kind of interaction, a method that does not distinguish between touching and pointing is necessary. Therefore, a 3D representation of the people standing in front of the display is created from which the hand and arm positions of the users can be extracted. In this 3D view, the only difference between touching and pointing at the display is the distance between the user's arm and the display. This goes hand in hand with the idea of not changing the way of interaction. The user can interact with the display by pointing at it while holding her arm still, which is interpreted as touching the display. By withdrawing her arm or pointing at another point of the display the "untouch" event is performed [17].

This system uses multiple RGB cameras and voxel carving [19], a technique to compute the visual hulls of objects, to create a 3D reconstruction of the user. In the next step, the arm needs to be detected to get the direction the user is pointing at. This is done by clustering the voxels depending on their distance to the display. When the user is interacting with the display, the arm is always the part of the body that is closest to it. This allows the system to extract the direction of the arm and to create a straight line in the 3D view representing the user's pointing direction. In this view, the desired point of interaction is the intersection of this line and the plane of the display. Knowing the dimensions of the display in 3D as well as in pixel coordinates, the intersection point can be computed and then converted from the 3D model to 2D pixel coordinates [17].

Schick et al. used two cameras in their experimental setup which were positioned near the top corners of the screen. They conducted a user study at which the participants were asked to point at specific spots on the display being indicated by crosses. The mean error was 105,7 Pixels (relates to 103.2 mm), which is good considering that the users got no feedback on their interaction with the display and the low resolution of the cameras of 640 x 480 pixels. By using more than two cameras or cameras with higher resolutions, the accuracy can be further improved, making this technique a solution with rather low hardware requirements [17].

3.1.3 Distant Freehand Pointing and Clicking

The techniques presented by Vogel and Balakrishnan [20] are based on the problem that in certain environments large displays are not reachable for the user and need to be manipulated from a distance. They focus on direct manipulation through pointing and clicking, as it is an intuitive interaction paradigm in conventional user interfaces. Unlike earlier works that needed additional handheld devices like flying mice [22] or laser pointer-style devices [15] to interact with the display, they present two clicking and three gestural pointing techniques that allow clicking and pointing using only the human hand. For the development and evaluation of their methods, they used the *Vicon* motion tracking system which was attached to the hand in a glove-like manner. However, they refer to methods of real time motion-tracking systems that should be able to achieve similar results as they did in the near future [20].

The two clicking techniques are shown in figure 6. The *AirTap* method is similar to a click on a conventional mouse or a tap on a touch screen and is performed by moving the index finger down and up. One problem the authors faced was that they could not define a definite start and stop position for the downward movement of the index finger without violating their freehand interaction paradigm. To solve this, they measured relative features like velocity and acceleration additionally to the absolute position and movement axis of the finger. The second clicking technique, called *ThumbTrigger*, uses a similar approach but uses the thumb to recognize clicks by moving it to the index finger side of the hand. The theoretical advantage of this method is that it provides an absolute down position as well as kinesthetic feedback when the thumb touches the side of the hand. However, early prototype studies showed that users thought it was too uncomfortable and tiring to "click" on their side of the hand, and therefore a recognition algorithm using relative features similar to *AirTap* was implemented [20].

In addition to these two clicking techniques the authors presented three gestural pointing methods. The first, *RayCasting*, takes a similar approach as the method presented by Schick et al. in section 3.1.2. *RayCasting* is based on the natural human pointing gesture at which the index finger is extended while the palm is facing downwards. The cursor on the screen is then placed at the intersection of the extended line indicated by the direction of the index finger and the display. The second technique uses the "safe hand" gesture (the palm is facing forward with all fingers loosely extended) to navigate the cursor on the screen. To define an absolute starting position for the hand movement tracking, a clutching gesture is used allowing the user to decide when the tracking should begin by clenching her fist. All other movements are then relative to this starting point. Therefore, they refer to this method as *Relative*. The last method, called *RayToRelative*, is an hybrid of the previous two techniques and uses *RayCasting* to set the pointer to an absolute starting point and the "safe hand" gesture to control the cursor [20].

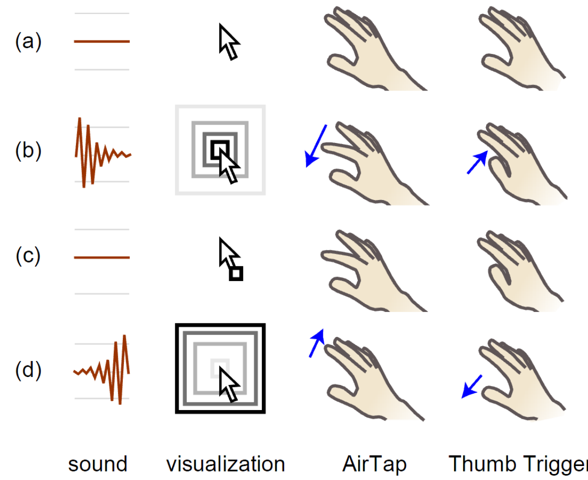

Fig. 6. The two clicking techniques *AirTap* and *Thumb Trigger* with their respective sounds and visual feedback. (a) shows the default state, (b) the click-down gesture, (c) the drag gesture and (d) the click-up gesture [20].

The authors evaluated their clicking and pointing techniques in a user study in which the participants should perform various tasks containing multiple point and click operations. They could not measure a difference in trial performance time or error rate between the two clicking gestures. However, they decided that *AirTap* might be the better alternative of the two because of the gesture's similarity to mouse and touch screen interaction. In terms of the pointing methods, they found that *RayCasting* lead to a significantly higher error rate than the other two techniques, achieving quite similar results in the user study. With these two techniques, they measured error rates of 9.6% and 15.4% when the participants should try to select small targets (width of 16mm) from a distance of 1.3m [20]. However, the need of a glove-like motion-tracking system is not applicable for everyday use and the prediction, that video-based motion-tracking systems could achieve similar levels of accuracy, still needs to be proven.

3.2 Interaction Techniques Using Handheld Devices

In this section, two interaction techniques using camera-equipped handheld devices to perform interaction on large displays are presented. As modern mobile phones are typically equipped with a camera and a big display, these techniques have a high amount of potential users due to the increasing spread of smartphones. Furthermore, the techniques would not need additional pointing devices as users could use their own mobile phones to interact with a display.

3.2.1 Direct Pointer

Direct Pointer [11] is an interaction technique that allows users to intuitively interact with large displays using cameras that are equipped in handheld devices, for example in mobile phones. It is an extension to the concepts of *Sweep* and *Point&Shoot* [2]. The primary advantage of *Direct Pointer* is that it needs no extra hardware to allow interaction on the display. The goal of this method is to allow direct manipulation of the cursor while giving the user continuous visual feedback, compared to the use of a laser pointer [11].

The basis of the system is a closed-loop feedback between the handheld device and the display. The camera captures the screen and sends the data to the server (the computer to which the display is connected) via a wireless connection. The server then updates the position of the cursor, placing the cursor in the center of the camera frame. To identify the cursor on the camera image, different characteristics can be used like color, shape or motion of the cursor or a combination of all three. The motion of the cursor can be estimated by identifying the motion of the background in the camera frame. By comparing consecutive frames, the affine transformation between these frames can be computed and the position of the cursor is detected as an area of difference between the frames [11].

The authors conducted a user study to compare the *Direct Pointer* against other input devices, like a trackball, a joystick and a laser pointer. The evaluation was based on ISO 9241-9, a standard which defines the ergonomic requirements for non-keyboard input devices and also includes suggestions for tests to prove a device's conformity to the standard. The prototype that was used in this study was a combination of a webcam to capture the video data and a wireless presenter to allow the participants to perform clicks using the presenter. The study revealed that the *Direct Pointer* performed better than the three other devices [11]. The fact that it does not need more than a camera-equipped handheld device makes the *Direct Pointer* an interaction technique that already has a large amount of potential users, considering the ubiquity of mobile phones and smartphones nowadays.

3.2.2 Touch Projector: Mobile Interaction through Video

An approach that follows a similar idea as the *Direct Pointer* is the *Touch Projector* [6]. It also provides an interaction method for distant displays through a live video image, similar to the *Direct Pointer*. However, *Touch Projector* does not use direct cursor manipulation, but gives the user the opportunity to interact with the distant display by using the touch input of the handheld device. When the user touches a point in the live video image, the input is projected to the target display and the input is processed as if it had occurred there. Furthermore, *Touch Projector* allows to not only manipulate the content on one display but offers the possibility to interact in an environment with multiple displays, for example by transferring information from one display to another. Figure 7 illustrates the idea of the *Touch Projector* [6].

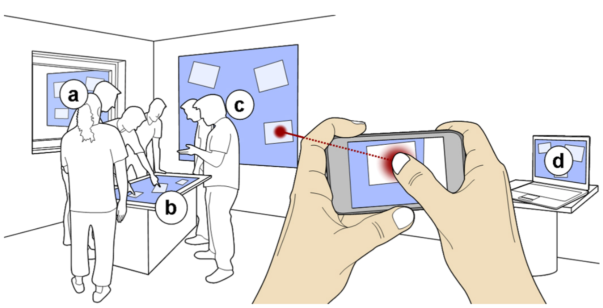

Fig. 7. *Touch Projector* allows users to interact with distant displays [6].

The system of *Touch Projector* consists of the *Touch Projector* device (an Apple iPhone 3GS was used in the prototype), a software that is installed on all systems in the environment and a server that controls all communication between the devices. The live video image of the handheld device is transferred to the server, which computes the current position and orientation of the *Touch Projector* by comparing the frames from the video stream with the content that is shown on the displays. This way, the server instantly knows whether the mobile device is currently pointed at a display or not. An advantage of the *Touch Projector* is that it adds multi-touch input to any kind of display, even to non-interactive displays, as the handheld device determines the possible input methods [6].

In addition to the described approach, the authors presented three extensions to improve the usability of *Touch Projector*. These extensions included the *Manual Zoom*, allowing users to apply a digital zoom for easier interaction, as well as the *Auto Zoom*. This method zoomed in automatically to keep the size of the area that is shown on the mobile device constant, independent from the distance between the *Touch Projector* and the distant screen. The third extension is the *Freeze*, which provides the possibility to freeze the image that is currently shown on the handheld device to allow easier interaction on a still image. In a user study, the four conditions were compared to each other by letting the participants perform various tasks under all conditions. The results showed that the zoom-enabled approaches performed better than the original method. The study also revealed that interaction with a still image of the display allows a more precise manipulation, leading to the conclusion that automatic zooming in combination with the possibility to freeze the current image leads to the best results [6].

4 Discussion

In this chapter, the advantages and disadvantages of the techniques presented before are stated. Furthermore, the three major problems that exist while interacting with large-scale displays are extracted. The techniques are then evaluated whether they can overcome these problems.

4.1 Software-based Interaction Techniques

The *Push-and-Pop* as well as the *Vacuum* technique both offer a functioning and easy-to-use way to access remote data. However, both techniques only work in one direction. This means, that the user can only apply manipulation to data she can actually reach. If she wants to interact with an object that is out of her reach, she cannot do that as she needs to physically reach and touch the object. Both techniques only work when the target of a manipulation is out of reach, and not the source. Therefore, one can say that both techniques did not entirely solve the problem to reach every point of the display.

The *I-Grabber* provides an interaction method that is based on an intuitive grabbing metaphor. The advantage compared to *push-and-pop* and *Vacuum* is that the user is given the opportunity to freely move objects on the screen, allowing interaction in both directions. Due to the fact that the authors did not conduct a user study, a comparison to other techniques is not possible.

Allowing distant manipulation through a portal metaphor, the *Frisbee* technique achieved its main goal by providing a technique that is based on the paradigm of minimizing the physical movement. Due to the portal metaphor, every part of a large display can be reached. Furthermore, *Frisbee* overcomes the problem of *Push-and-Pop* and *Vacuum*, which both have the problem that they cannot provide interactions when the source of a manipulation is out of reach.

4.2 Hardware-based Interaction Techniques

A disadvantage of *Shadow Reaching* is the problem that the room needs to be darkened to achieve "good" shadows with sharp edges, as bright rooms might lead to washed-out shadows. This might result in a decreased performance in recognizing the user's body [18]. The system also needs a display that has a light source behind the screen to perform motion-detection and to ensure that the content is always visible. Otherwise, using a projector as the display, the person standing in front of the projection surface might "block" the display when she is using her shadow to point at it.

The distant freehand pointing and clicking techniques presented good approaches. The use of a motion-tracking device that is attached to the hand contradicts to the paradigm that the technique is "freehand". The authors indeed proposed to use video image recognition instead of the motion-tracking-device, probably knowing that this might lead to a decrease in accuracy [8]. Comparing the results of the study to the combined touch and pointing method, the difference in accuracy becomes obvious.

The two techniques using mobile devices provided both solutions that required low hardware costs while offering a technique that performed well in their user studies. Due to the growing ubiquity of smartphones, these techniques already have a lot of potential users and might become more important in future research.

4.3 Problems While Interacting With Large Displays

Concerning the advantages and disadvantages of the techniques described in this paper, three major problems that exist while interacting with large-scale displays can be extracted. These problems are:

- Parts of the display are out of the user's reach
- The user can not see all parts of the display when she stands right in front of it
- The user has to interact with the display from a distance (for example, when the display is located high above the user)

All of the software-based techniques described in this paper present solutions to the problem that the user might not reach certain parts of the display. Especially the *Frisbee* approach provides interaction on a large display as if the manipulation would occur at the remote part of the display. However, in all of these approaches the user needs to touch the display, as no further hardware should be used to allow interaction. This necessarily means that the user always has to stand near to the display, which results in the problem that the user is not able to see all parts of the display equally or to get an overview of the display. Furthermore, software-based solutions cannot be used when the user has to interact with the display from a distance. In this case, the user always needs some kind of hardware to manipulate the display. Therefore, software-based techniques can not solve the major problems that exist while interacting with large displays.

Compared to the software-based approaches, the hardware-based techniques can overcome these problems. However, all of these techniques have the disadvantage of high costs compared to the software solutions. The techniques that are based on motion-tracking need multiple cameras and additional sensors to achieve good results, making these techniques harder to use and in most scenarios not applicable for everyday usage. Furthermore, *Shadow Reaching* has the problem that the user "blocks" parts of the display when a projector is used as a display. This leads to the problem that other people might not be able to follow the user's interactions on the display. The techniques using handheld devices do not have these problems. However, with the usage of a second display to navigate on another display, the problem of reorientation occurs [21]. As the displays differ in size, the smartphone can only show a certain section of the large display while still providing precise interaction. This means that the user has to reorientate when she switches between looking at the large-scale display and the display on the handheld device as she might lose orientation. In conclusion, one can say that the hardware-based techniques can solve the problems mentioned before, but they also introduce new problems.

5 Conclusion and Future Work

This paper gave an overview of software- and hardware-based interaction techniques for large displays. After presenting each technique, the advantages and disadvantages of these techniques were discussed. Based on these findings, the three major problems that occur while interacting on large displays were extracted and discussed. The analysis of the presented techniques leads to the conclusion, that these approaches are either not able to solve all of the problems or that they cannot achieve this without introducing new problems.

Comparing software-based and hardware-based techniques, it is hard to determine general situations in which only one of these techniques would make sense. There are, however, two basic situations in which only one of the techniques is useful. As software-based techniques usually require touch input, it makes no sense to use these techniques in situations where the display is positioned in a way that the user cannot reach it. According to this, hardware-based techniques should be used when the display is out of the user's reach.

Future work might involve, for example, the *Kinect* system by Microsoft to perform motion-tracking tasks, providing a cheap alternative compared to commercial motion-tracking systems. This would also overcome the need of multiple cameras or additional sensors attached to the users body. The official support of the *Kinect* system for Windows will begin in February 2012 [14], which might lead to *Kinect*-based solutions for interaction on large-scale displays in the future. Furthermore, the usage of smartphones or tablets as input devices might become more important as their distribution is constantly increasing. Using tablets instead of smartphones comes with the advantage that the user has a bigger display to interact with the large display. This could provide a more accurate way to manipulate the large display. Additionally, it might reduce the reorientation problem, as the tablet's display would be able to show a larger part of the display, helping the user to reorientate faster.

References

[1] M. Abednego, J.-H. Lee, W. Moon, and J.-H. Park. I-Grabber: Expanding Physical Reach in a Large-Display Tabletop Environment Through the Use of a Virtual Grabber. In *Proceedings of the ACM International Conference on Interactive Tabletops and Surfaces (ITS '09)*, pages 61–64. ACM, 2009.

[2] R. Ballagas, M. Rohs, and J. Sheridan. Sweep and Point & Shoot: Phonecam-Based Interactions for Large Public Displays. In *CHI '05 Extended Abstracts on Human Factors in Computing Systems (CHI EA '05)*, pages 1200–1203. ACM, 2005.

[3] P. Baudisch, E. Cutrell, D. Robbins, M. Czerwinski, P. Tandler, B. Bederson, and A. Zierlinger. Drag-and-Pop and Drag-and-Pick: techniques for accessing remote screen content on touch- and pen-operated systems. In *Proceedings of INTERACT*, pages 57–64, 2003.

[4] A. Bezerianos and R. Balakrishnan. The Vacuum: Facilitating the Manipulation of Distant Objects. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '05)*, pages 361–370. ACM, 2005.

[5] E. A. Bier, M. C. Stone, K. Pier, W. Buxton, and T. D. DeRose. Toolglass and Magic Lenses: The See-Through Interface. In *Proceedings of the 20th Annual Conference on Computer Graphics and Interactive Techniques (SIGGRAPH '93)*, pages 73–80. ACM, 1993.

[6] S. Boring, D. Baur, A. Butz, S. Gustafson, and P. Baudisch. Touch Projector: Mobile Interaction through Video. In *Proceedings of the 28th International Conference on Human Factors in Computing Systems (CHI '10)*, pages 2287–2296. ACM, 2010.

[7] M. Collomb, M. Hascoët, P. Baudisch, and B. Lee. Improving drag-and-drop on wall-size displays. In *Proceedings of Graphics Interface 2005 (GI '05)*, pages 25–32. Canadian Human-Computer Communications Society, 2005.

[8] T. Grossman, D. Wigdor, and R. Balakrishnan. Multi-Finger Gestural Interaction with 3D Volumetric Displays. In *Proceedings of the 17th Annual ACM Symposium on User Interface Software and Technology (UIST '04)*, volume 24, pages 61–70. ACM, 2004.

[9] F. Guimbretière, M. Stone, and T. Winograd. Fluid Interaction with High-resolution Wall-size Displays. In *Proceedings of the 14th Annual ACM Symposium on User Interface Software and Technology (UIST '01)*, pages 21–30. ACM, 2001.

[10] M. Hascoët. Throwing models for large displays. In *Proceedings of HCI 2003*, pages 73–77, 2003.

[11] H. Jiang, E. Ofek, N. Moreveji, and Y. Shi. Direct Pointer: Direct Manipulation for Large-Display Interaction using Handheld Cameras. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '06)*, pages 1107–1110. ACM, 2006.

[12] A. Khan, G. Fitzmaurice, D. Almeida, N. Burtnyk, and G. Kurtenbach. A Remote Control Interface for Large Displays. In *Proceedings of the 17th Annual ACM Symposium on User Interface Software and Technology (UIST '04)*, volume 6, pages 127–136. ACM, 2004.

[13] M. W. Krueger, T. Gionfriddo, and K. Hinrichsen. VIDEOPLACE - An Artificial Reality. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '85)*, volume 16, pages 35–40. ACM, 1985.

[14] Microsoft. Kinect SDK for Developers. http://www.microsoft.com/en-us/kinectforwindows/, last viewed on 2012-01-22.

[15] D. R. Olsen Jr. and T. Nielsen. Laser Pointer Interaction. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*

(CHI '01), pages 17–22. ACM, 2001.

[16] G. Robertson, M. Czerwinski, P. Baudisch, B. Meyers, D. Robbins, G. Smith, and D. Tan. The Large-Display User Experience. *Computer Graphics and Applications, IEEE*, 25(4):44–51, 2005.

[17] A. Schick, F. van de Camp, J. Ijsselmuiden, and R. Stiefelhagen. Extending Touch: Towards Interaction with Large-Scale Surfaces. In *Proceedings of the ACM iIternational Conference on Interactive Tabletops and Surfaces (ITS '10)*, pages 117–124. ACM, 2009.

[18] G. Shoemaker, A. Tang, and K. S. Booth. Shadow Reaching: A New Perspective on Interaction for Large Wall Displays. In *Proceedings of the 20th Annual ACM Symposium on User Interface Software and Technology (UIST '07)*, pages 53–57. ACM, 2007.

[19] R. Szellski. Rapid Octree Construction from Image Sequences. *CVGIP Image Understanding*, 58(1):23–32, 1993.

[20] D. Vogel and R. Balakrishnan. Distant Freehand Pointing and Clicking on Very Large, High Resolution Displays. In *Proceedings of the 18th Annual ACM Symposium on User Interface Software and Technology (UIST '05)*, pages 33–42. ACM, 2005.

[21] D. Wigdor, C. Shen, C. Forlines, and R. Balakrishnan. Effects of Display Position and Control Space Orientation on User Preference and Performance. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '06)*, pages 309–318. ACM, 2006.

[22] S. Zhai. User Performance in Relation to 3D Input Device Design. *ACM Siggraph Computer Graphics*, 32(4):50–54, 1998.

Tactile Tangibles: Areas of Application

Simon Gurn

Abstract—
This paper introduces tactile tangibles, physical devices to control digital data and give tactile feedback to the user. Classical Graphical User Interfaces (GUIs) usually present the user visual and mostly also acoustical information as feedback for interactions. Nowadays, researches try to adress a third sense: the sense of touch. In Tangible User Interfaces (TUIs) the physical input device also serves as an output device for more immediate representation of data. In tactile TUIs the user even is addressed through touch. Through the sense of touch, the interaction is enriched in a lot of different ways to be shown. This is not solely limited to interaction between human and machine but also includes interpersonal interaction through machines, for example mobile devices. Thus tactile tangibles show different specifications and implementations based on their later use. Divided into two groups, interaction between human and machine and interpersonal interaction, a selection of applications and their scenarios is presented.

Index Terms—
tactile, tangibles, haptic, interaction

1 Introduction

While using computers, no matter whether working on a personal computer or talking with a friend on the phone, one always has to overcome a certain distance. Whereas the physical distance in the second case is instantly evident, the logical distance while working on a computer may be not, as we unconsciously overcome it everyday as it is an inherent part of our lives. In both cases we have to bridge the gap to interact properly. The established form of interaction is done through interfaces with visual and acoustic output. The sense of touch is barely used, if ever then through vibration as most mobile phones are capable of, but a very important one that should not be ignored.

In human-machine interactions touch permits a better ease of use, as for example the reaction time is far shorter than that of vision or hearing [3]. Surfaces and their characteristics can be identified more immediate. And while visual feedback requires the undivided visual attention of the user and acoustic feedback can be hard to keep track of while in a noisy environment, tactile feedback can be perceived as long as there is physical contact [8, 12].

In interpersonal interaction touch provides more private means of communication. Expressing emotions is possible in more various ways, and as Park et al. state touch is considered to be the most fundamental and primitive form of non-verbal communication methods [2, 9].

Overall touch offers a more immediate way of interaction and improves the connection of the physical and digital world.

In the following, tactile tangibles are introduced, serving to improve the connection of both worlds through the sense of touch. Several applications and their scenarios will be shown to provide an insight into the different areas of application of tactile tangibles.

2 Classification

Tactile tangibles are physical objects used to interact with digital data and provide tactile feedback for the user. Tactile User Interfaces (TUIs) give physical form to digital information, making it directly graspable and manipulable [4]. Compare Urp (an Urban Planning Workbench, developed by the Tangible Media Group in 1999) as an example for a basic TUI, where scaled physical representations of buildings are arranged on a table and their shadows, light reflections and wind flows are simulated [4]. While all tangibles are haptic, not all are also tactile. Tactile tangbiles offer feedback that is felt on the surface of the skin or can be felt through the skin. The most important types of tactile feedback are: pressure, friction, temperature and malleability. Tactile TUIs let the user feel the digital information, rather than just grasp it. For example when pressing a digital button in such an interface the user could get tactile feedback similar to when pushing a real physical button, making the digital information more immediate. Haptic feedback also includes kinetic feedback, that is commonly also known as force feedback and includes forces acting on muscles or joints. This paper focuses on tangibles with tactile feedback [8, 12].

3 Applications and Scenarios

The following section shows a variety of different scenarios for tactile tangibles. As every single scenario focusses on its own unique task there are many different motivations for all those scenarios. They all will be presented, as well as the configuration of the prototypes for each application, their implementation and testing. Mentioned earlier there are two main settings for the use of touch. One to overcome logical and one to overcome physical distance. Or also: the use of touch in interaction between human and machine only or interpersonal interaction as well. Those two are the main areas of applications tactile tangibles are used in and thus the following application examples for tactile tangibles will be divided into these two main categories.

3.1 Logical Distance

First, applications focussing solely on human machine interaction and overcoming logical distance are presented. This category mostly contains tangibles like touchpads, pens and smartphones provided with tactile feedback and in rare cases other unique objects especially designed for specific tasks. Main goals of this tangibles are to improve performance and provide the user with a more effective, comfortable and enjoyable user experience through the sense of touch [12].

3.1.1 The Tactile Touchpad

One of the first touchpads implemented with tactile feedback, *The Tactile Touchpad* from I. MacKenzie and A. Onisczczak was introduced in 1997 [7]. At the time touchpads became the main pointing device for notebook computers, replacing the trackball and isometric joystick. Mice were now the main choice (as they are for desktop computers) if it were not for the usage of notebook computers mostly in constrained spaces.

But it is not only because of the space availability why mice are the favourite choice for desktop computers, but because they offer a better usability. While both touchpads and mice use relative positioning to navigate, there are two common implementations for selecting used by touchpads, that are both inferior to the ones of the mouse: the usage of either physical buttons or lift-and-tap. Physical buttons, either operated with the index finger or the thumb, come with two problems. First: if the index finger is used for operating the user has to frequently switch between the touchpad and the button which reduces

- *Simon Gurn is studying Media Informatics at the University of Munich, Germany, E-mail: simon.gurn@campus.lmu.de*
- *This research paper was written for the Media Informatics Advanced Seminar 'Over Distance', 2011/2012*

performance. Second: using the thumb for operating the touchpad is suboptimal because of the interference between the muscle and the limp groups. A problem not occurring while using a mouse where positioning occurs via the wrist and forearm and selecting through the fingers. Lift-and-tap comes with a similar performance problem. Certain actions performed require more input compared to when using a mouse. For example when performing the action of dragging the user first has to lift off his finger from the touchpad to be able to click on an item as with a mouse the user always can click and drag without interruption.
Thus a prototype for the *Tactile Touchpad* is designed to eliminate the disadvantages compared to the mouse. Instead of using the lift-and-tap concept the touchpad now reacts to pressure to perform tasks like dragging. So it is no longer necessary to lift off a finger. Different pressure thresholds for different tasks shall prevent unintended clicks. To provide a smooth usage the thresholds are based on the state the user is in. The three states are (1) not-tracking, (2) tracking and (3) dragging. In addition to that the *Tactile Touchpad* has a build in relay to provide the user with a "click" sensation when pushing a digital button. The user experiences a similar feeling as when pushing a mechanical button. The touchpad gives the impression to move slightly at the pressure thresholds. This tactile feedback of a "click" sensation combined with the sound of one improves the connection of the digital and physical world [7].

3.1.2 The Haptic Pen

The *Haptic Pen* by J. C. Lee et al., presented in 2004, is a stylus for touchscreens that provides tactile feedback [6]. Implementing the tactile feedback inside the stylus rather than the touchscreen negates two problems. While Personal Digital Assistants (PDAs) or other small devices can be effectively equipped with a physical actuator behind the touch surface, this method does not work well with larger screen sizes. Also this method cannot provide individualized feedback for multiple users simultaneously. By placing the actuator inside the stylus it is possible to support multiple users simultaneously and feedback can be given regardless of screen size. In addition to that, the *Haptic Pen* can detect tip pressure, is able to give constantly feedback even when not pressing on the surface and can utilize this hovering state to locate data without contact.
The *Haptic Pen* is capable of producing many different sensations similar to mechanical switches. Those are used to improve button simulation. Most behaviours of buttons can be mapped to different applications in a classical user interface, while for example dragging interactions are less uniformly appropriate. Eight behaviours for haptic buttons are implemented: (1) No Click, (2) Light Click, (3) Basic Click, (4) Hard Click, (5) Buzz, (6) Force Buzz, (7) Two-Click and (8) Buzz-Click. Similar to the *Tactile Touchpad* the reaction of the pen and hence the type of the tactile feedback given is based on the state the user is in. For example when performing a Two-Click the user receives a light-click sensation when pressed halfway followed by a stronger full click if pressed harder. This way a double click is simulated with technical one single pen-down action but a sensation of two clicks.
Besides these basic types the button behaviours of the *Haptic Pen* can be modified for all kinds of application usage. For example when using a settings dialogue each selection may only give light feedback while the confirmation button may be very stiff. This way the user must be confident when confirming his actions, but provided with a sense of closure and completeness. Or when checking emails the button to call new ones becomes stiffer depending on how many mails are available. When not physically touching the screen the stylus still can provide feedback. For example the Buzzing strength can be driven by proximity, region or direction to guide users toward a target area. Also when not interacting with the surface at all the stylus can function as a consistent feedback channel, especially valuable when visual or audio feedback is impossible.
Responses of informal usage experience interviews indicated a high degree of believability in the tactile simulations [6].

3.1.3 The Haptic Tabletop Puck

The *Haptic Tabletop Puck (HTP)* by N. Marquardt et al. (presented in 2009) is a device designed for tabletops that provides haptic feedback [8]. The interaction with real objects on tables is strongly connected with the sense of touch. Whether it is the texture of a paper that makes it immediate differentiable from the table or the relief of a ruler allowing it to be placed accurately, the tactile information of surfaces and objects help us with the execution of many different taks. But when interacting with digital objects nearly all those tactile cues are sacrificed. The *HTP*, used on digital tabletops, provides this tactile informations to make the touch interface more realistic and also extend it in new ways.
The *HTP* is a small wooden box equipped with a vertically installed rod to put a finger on while holding the device with the rest of the hand. The Rod has a sensor on top to measure pressure if applied. On the bottom a break pad is installed to provide friction. Following information can be communicated through tactile feedback: height and texture of elements of the table, malleability of different materials, friction while moving and location based as well as multi-user related feedback.
Various applications demonstrate the use of the *HTP*. A map application allows the exploration of geographical features. The relief and different kinds of terrain as well as the ocean temperature are mapped to different independent tactile responses. To improve orientation a digital arrow is added to the front of the puck representing the position of the user on the table, because otherwise the device itself would hide the active area. An application to explore paintings allows the user to feel and explore hidden areas of a painting, so called pentimenti, previous versions of the painting that have been painted over by the artist. A painting application improves working with multiple layers at the same time. Depending on how much pressure is applied the user can only paint on a new layer or the original one. Another example for the pressure sensor is used on a shopping site. The web page buttons have different pressure thresholds. Adding an item to a cart only requires low pressure while checking out and purchasing requires a strong click.

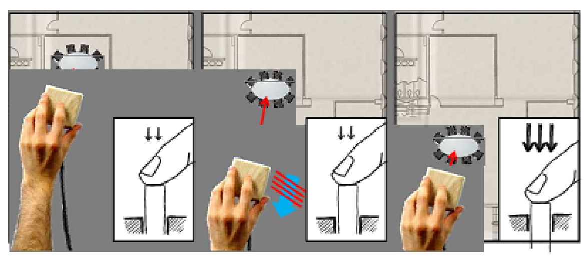

Fig. 1. HTP layout application: Moving an object inside the room with low pressure (left). Trying to move it outside its boundaries is not possible (middle) until enough pressure is applied (right) [8].

Final example for the use of the *HTP* is a office layout application to rearrange furniture (Figure 1). A combination of height and pressure is used. While objects can be moved inside a room with only light pressure, high pressure must be applied to move a object through a wall, otherwise the object stays in the room. While the pressure value is not reached the object does not pass the wall and the digital arrow stretches. Additionally the friction increases, making the *HTP* harder to move. This combination of feedback illustrate to the user that his actions do not have any effects. The office layout application also supports a multi-user environment. Changes made by a user let other pucks on the table oscillate. The feedback varies depending on the identity and distance of the initiator. This kind of feedback can easily be noticed without being distracted from an ongoing activity. Also this form of awareness information is highly useful when other is not available, such as peripheral vision or sound. This would be the case for example when working on very large tables.
Observing visitors of the lab using the devices started in varying ways

but reported that the feedback through the sense of touch was a very active and engaging experience [8].

3.1.4 TouchEngine

The *TouchEngine* (by I. Poupyrev et al. in 2002) is a miniature tactile actuator designed for mobile interfaces [11, 12, 13]. Using a mobile device like a PDA in public spaces mostly requires the users visual attention. The audio channel often cannot be used because of the interference of real-world noise with the mobile devices sounds. Adding an ambient tactile interface can help redirect information from the visual and audio channel to touch. Not only is the skin the largest human sensory organ with most of it unused, the sense of touch also is about five times faster than vision [3]. Thus an ambient tactile interface improves especially peripheral awareness where the user receives and processes information without having to switch from his current activity.
The *TouchEngine* placed in a PDA can provide direct tactile feedback, when pushed on the display, to simulate various tactile sensations, like the feeling of a mechanical click. Indirect tactile feedback is also possible, where the whole device functions as a tactile display. This is similar to the vibration of mobile phones, but faster and able to create more sharp and distinct force pulses. Three main scenarios for the use of ambient touch interfaces are: (1) notification through touch, (2) tactile monitoring and (3) tactile feedback for gestural mobile interfaces. While handheld devices already communicate tactile information, for example an incoming phone call, no further information is given, like who is calling or how important the phone call is. Thus the user has to interrupt his ongoing activity even so. This problem can be neglected through tactile notification with more complex vibration patterns that can process more information to the user.
For observation tasks, for example when keeping track of a progress bar, tactile monitoring is used. A tactile progress bar informs the user with two repeating tactical impulses. As the progress processes the time between the two impulses decreases. This way the user does not have to check the display for updating information, as he can feel the progress.
Gestural mobile interfaces are used in devices enhanced with tilt sensors that let the user interact with data by physically interacting with the device. Tilting interfaces, where the user tilts the device back and forth to scroll through data, can be greatly enriched through tactile feedback. A main problem is known as *overshoot* where the target destination is reached but the user is not able to stop scrolling in time by tilting the device back in its natural state. Also users often overshoot in the wrong direction, meaning while trying to stop scrolling, the neutral position is missed and the user starts scrolling in the opposite direction. Third downside is the need of constant visual attention, since a short distraction can result in a loss of control. Using tilting in a tactile enhanced interface improves all those weaknesses.

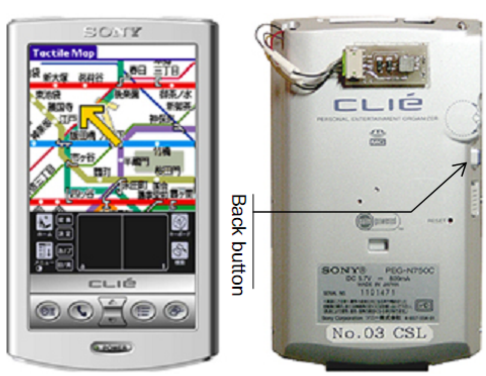

Fig. 2. TouchEngine: Example of a tilting interface. Browsing through a Tokyo subway map [12].

Figure 2 shows the first example application of tilting with the PDA: browsing through a Tokyo subway map. A button at the back activates the mode. Tilting the devices now results in scrolling over the map, while the yellow arrow points the direction and increases in length as the speed increases. Every time the image moves on the screen, the user can feel a scratching tactile pattern.
A second application is simple but often used: scrolling through lists of text. Every time the user scrolls one line further he feels a simple "tap". With increased scrolling speed the intervals of the tapping decreases, giving the user full control without the need of vision.
User studies showed a great acceptance of the with *TouchEngine* enhanced PDA. An experimental user study evaluating the use of the tactile tilting interface showed an improvement of a 22 percent faster task completion, decreasing the general task duration time as well as reducing the overshoot problem. All users preferred to have tactile feedback and reported an overall better user experience [11, 12, 13].

3.1.5 MudPad

A newer addition to the list of tactile tangibles is the in 2010 introduced *MudPad* by Y. Jansen et al. [5]. While touchpads are more and more common they are not flawless. When interacting for example with a digital keyboard temporary graphical overlays are used. Because the user has to touch the device, this touched area is consequently hidden and therefore those hidden, in this case, buttons are temporary displayed above the touched area. This results in occlusion of other parts of the interface. Touchpads with tactile feedback are single user only and are not able to give multiple tactile feedbacks at the same time.
The *MudPad* (Figure 3) is a multi-user touchpad and the first device of its kind capable of localized active haptic feedback. It is able to produce a wide range of tactile sensations at multiple positions at the same time. Each graphically displayed user interface element can also be associated with a different tactile sensation, that go beyond the simulation of a mechanical button. As the touch surface contains a liquid whose physical properties can be controlled, it is possible to create different variations of stiffness, textures and roughness and to change those states in a very short time. While hovering on touchpads is mostly reserved to interfaces supporting an additional input device like a stylus, the *MudPad* allows the user to explore the interface with the sense of touch. In this "hovering state" the user can get tactile feedback, but does not initiate any actions while touching the device. Not until the pressure reaches a certain threshold.

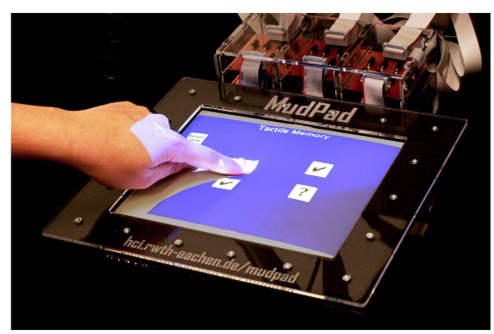

Fig. 3. MudPad: Picture of the MudPad prototype [5].

Being able of multi-touch input and output with rich and quick changing tactile feedback the *MudPad* offers new ways for interface design. Different kinds of fluid states and changes can be used to communicate information (Figure 4). For example background tasks can be felt in certain areas of the display, allowing the user to gather that information while hovering and without having to stop his ongoing activity.
One of the most intuitive applications is one of a virtual keyboard with tactile feedback, as virtual keyboards are generally harder to use because of the lack of physical response and their need of constant visual attention. Distinguishing keys from each other and being able to feel if input was successful is a huge improvement. Also imaginable would be a digital keyboard that changes the viscosity of the keys, depending on how likely it is for each key to be used. For example making unlikely keys stiffer so they do not get pressed unintentional.

Magnet Signal	Fluid State	UI Mapping (System View)	UI Mapping (User View)
	stiff	inactive areas, i.e., no user input possible	prevent interaction
press release	quick on/off transition	active UI elements, e.g., buttons	acknowledge user input
	(rapidly) changing vibrating	active areas, demanding user attention	communicate system processes, e.g., progress bar
	fluid	active areas, allow Interaction	neutral

Fig. 4. MudPad: Elementary building blocks from which feedback patterns can be constructed [5].

A more specialized application would be a music sequencer or digital instrument, enhanced with tactile feedback. While musicians generally benefit from a mobile digital recording environment, this could make them more viable. Similar to the keyboard problem, it is problematic to play an instrument on a digital UI. Being able to feel them would be great improvement. Each slider could play the signal it controls and the music or rhythm could be felt similar to the background-task scenario.
Currently user tests are being designed to evaluate the system and test new ideas, for example if it is possible to distinguish an *OK* from an *Cancel* button through touch [5].

3.2 Physical Distance

The following section presents tactile tangibles designed for the interaction between people that are not co-located. Everyone uses mobile phones to bridge distance and communicate with each other. While voices and acoustics are communicated this way for a long time now and video transmissions are more common than ever, those are the only two senses being used. Adding the sense of touch enriches the communication and enables new designs for interpersonal communication devices. A variety of those devices are presented, classical mobile devices like smartphones enriched with tactile feedback as well as specially designed devices for the solely purpose of communicating through the sense of touch over distance.

3.2.1 inTouch

InTouch by S. Brave and A. Dahley (from 1997) is a physical device designed solely to communication through the sense of touch over distance [1]. Touch is a particularly important communicator of affection and a fundamental aspect of interpersonal communication. Physical contact plays a huge role in close personal relationship, to achieve a sense of connection, indicate intention or express emotion. Goal of *inTouch* is to enable this kind of interpersonal interaction for two people that are not co-located.
The *inTouch* device consists of two objects, one for each user. Each objects consists of three cylindrical rollers mounted on a base. When one user rotates the rollers of his device, the rollers of the second device rotate in the same way. They can be rotated in either a clockwise or counter clockwise direction. The rollers can be rotated indefinitely in each direction, so that thrashing between bounds is not possible. This shall prevent an aggressively manipulation of the device, and accentuate the communication of subtle emotion states. Rollers were chosen because users could interact both active and passive, meaning they can either manipulate the device and rotate the rollers or they can just feel them and sense how the complement user is manipulating them. Three rollers are used so the whole device can be felt and activated with one hand only.
A mechanical prototype (Figure 5) was used to observe people interacting with it. While it did not support the desired distance it was enough for benchmarks and user testing. Users reported a ”playful” interaction, and while some complained the lack of the ability to communicate concrete information, others commended the abstract nature of it, as it was most fitting to intimate relationships [1].

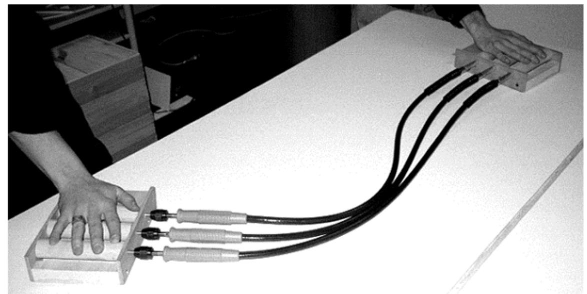

Fig. 5. First mechanical prototype of the inTouch [1].

3.2.2 ComTouch

The *ComTouch* device (by A. Chang et al. in 2002) augments remote voice communication with touch in form of a vibrotactile device sleeve that fits over the back of a mobile phone [2].
The *ComTouch* connects audio and tactile data and enriches conversations over the phone with the sense of touch and the private gains that come along. It also enables deaf people to communicate over distance with anyone who has a sense of touch. While using the device, one can simultaneously send and receive as its input and output channels are separated. Input occurs through pressure applied with the finger tip. The pressure is converted into vibration which can be felt at the own hand at the middle of the finger, so the user has more control about how much pressure he has to apply, and a clear picture about what the other will feel. The bottom of the finger receives the remote signal also in form of vibration. Compare Figure 6 for the touch-to-vibration mapping.

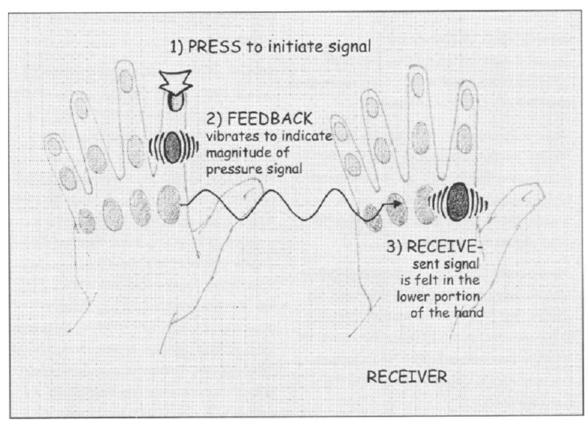

Fig. 6. ComTouch touch-to-vibration-mapping [2].

Extensive user testing was done to examine the use of the *ComTouch*. One scenario examined the use of both aural and tactile communication simultaneously and a second examined the use of tactile communication predominantly. Both scenarios showed that users would use three of the same tactile gestures: emphasis, turn-taking and mimicry. While communicating with both senses users tended to emphasize important points of the conversation by pressing more heavily. When they wanted to interrupt the conversion in order to say something, frequently buzzing was observed. In silent phases of the conversations as well as while talking, users tended to send individual tapping patterns which were mimed by their conversation partner. While some participants did not find the tactile channel that useful most of them used one of the tactile gestures at least unconsciously [2].

3.2.3 Connexus

The *Connexus* by E. Paulos (in 2003) is one of the more unconventional devices [10]. It is a communal interface, part of a new research

effort that studies non-verbal human cues such as their intent, motion, meaning, subtleties and importance in communication. A big part of interpersonal human communication takes place only through touch. People with established relationships show the need to maintain a communication channel most of the times when co-located and mainly is done in form of tactile interaction. This form of emotional ambient connection involves for example rather subtle touching of fingers, hands or backs than direct handholding.
The *Connexus* is a small, simple, wrist-worn personal object augmented with simple sensing, actuation, and ad hoc networking support [10]. Based on observed behaviour of interaction of co-located humans, its specification fulfils the following criteria: (1) continuous in- and output, (2) always on, (3) personal association to the communication device, (4) support for non-verbal communication and (5) providing an emotional interface. The device provides three sensors: a first sensor for pressure detection that detects both simple touching and rich signals, like for example swirling a finger along the surface. A second sensor that detects the pulse to provide a "life signal" for the other person. And a third sensor that senses ambient light and generates a signal when the user takes the *Connecus* on or off, or when he occludes it with his hand. Three actuators are used for output: a vibration motor implements both simple and complex vibration patterns that are felt on the users skin. An array of Light-Emitting Diodes (LEDs) allows various subtle glowing colours to be displayed. The third output changes the temperature to either hotter or colder than room temperature. The range of the device itself attains up to ten meters, but since it is able to communicate with a mobile phone and use the General Packet Radio Service (GPRS) to send and receive messages the range is accordingly higher.
When a pair of the *Connexus* is first exchanged between users face to face, according to internal studies, the exposed emotional connection between those two is higher. Also in this case the communication is only possible between those devices. Theoretically it is possible to receive and send messages between more than two devices, for example between a group of people. In this scenario questions about the possible downsides of personal privacy come to mind and must first be answered by conducting further user tests [10].

3.2.4 Haptic Instant Messaging

A framework for *Haptic Instant Messaging (HIM)* is presented in 2004 by A.F. Rovers and H.A. van Essen [14]. Communication today often takes place through internet-enabled technologies such as email or instant messaging. As most of the communication is done through textual messages, it is not as emotionally rich as interpersonal communication in the real word. A popular method to solve that problem, or at least a workaround, is the use of so called emoticons, specified icons to express emotional states. Also the exchange of instant messages is not only done to send messages but to stay in touch with friends and share intimate and emotional information. As touch is the most direct and intimate manner of interpersonal interaction the *HIM* framework enriches instant messages by adding haptic interaction.
A *Haptic Instant Messenger* is developed that can be used like an ordinary instant messenger (such as list people that are online and send messages) but with the addition of haptic input and output. The devices used for input and output can be selected by the user. To couple meaning and touch, hapticons are used: they are defined as small programmed force patterns that can be used to communicate basic notion in a similar manner as ordinary icons are used in graphical user interfaces. [14] A basic predefined set is implemented but they can be simply expanded by the user, as the devices used for input and output. The framework was from the start designed to be easily expandable. A haptic effect is triggered by the use of a hapticon. This is either the case when a special input device is used or the textural representation for it, such as an emoticon, is sent. The haptic effect is determined through frequency, amplitude and duration. Also the haptic effect the user receives depends on what output device is used. By default not only common devices like joysticks and mice are supported but also custom designed devices can be used.
A sample scenario describes the use of the *HIM*. For input and output a specially designed device is used. It consists of a touchpad on the desk for input and a vibration device for haptic feedback, worn in the pocket (Figure 7). To request a chat, instead of writing a message, the touchpad can be used by tickling on it. The person getting the request now gets feedback from his output device. When hapticons are triggered in form of emoticons, different tactile sensations are communicated through the vibration device. For example when using the big-smile-emotion (:D) the generated haptic effect would be a fast vibration with increasing amplitude burst that ends abruptly [14].

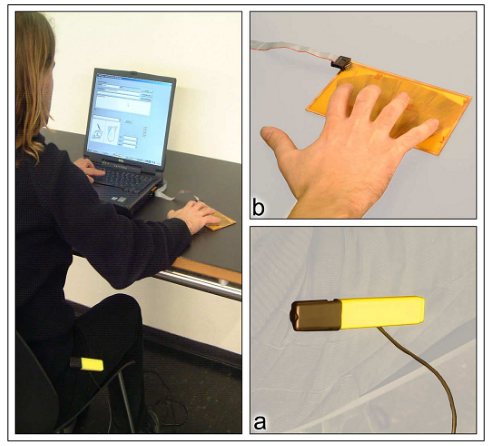

Fig. 7. HIM: Example of custom made devices: vibration device (a) and touch pad (b) [14].

3.2.5 CheekTouch

CheekTouch by Y.-W. Park et al. (in 2010) is an affective interaction technique while speaking on the mobile phone [9]. It underlies a similar concept as Chang et al. *ComTouch* but is implemented differently in almost every design choices. Also the design choices start from different point of views. While the *ComTouch* focuses on tactile communication the *CheekTouch* adds tactile information to voice communication to overcome missing non-verbal cues [14].
Instead of putting the mobile phone in a device sleeve that handles input and output, the *CheekTouchs* sensors and actuators are implemented inside the mobile phone. This way the user can hold it naturally while communicating tactile information. Tactile feedback is delivered through the cheek, which is profitable as the rich receptors on the cheek are able to detect various affective gestures. Actuators inside the device deliver the tactile information through vibration. The users input is done through the back of the mobile phone. Multiple touch patterns as input can be used to communicate different gestures. As not all of the most common touch gestures of human interaction can be expressed with multi-finger input, six types were selected that were both important for interpersonal interaction as well as realizable. The six gestures are: patting, slapping, pinching, stroking, kissing and tickling. Each gesture has a specific input that is done with the index and middle finger, since those fingers are free to use while holding a mobile phone naturally. Figure 8 displays a list of the mapping of all possible gestures. For example to communicate kissing, the two fingers are gathered slightly together.
A user study was held to evaluate both the usability of the input gestures as well as the vibrotactile feedback on the cheek and their mapping appropriateness according to the touch behaviours. The overall feedback was positive above average. Some gestures, like pinching, were less often identified correct then others, like kissing, but most participants complimented the ease of learning. Intention of future use was mostly reported by female participants while the overall majority reported the need for an implementation of expressing shapes freely and not be limited to six gestures [9].

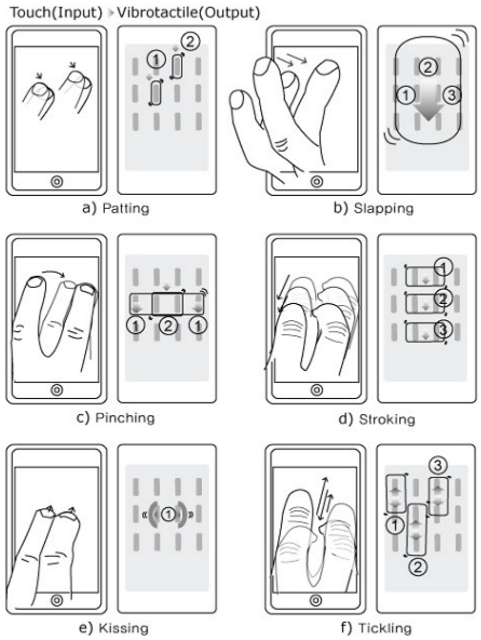

Fig. 8. CheekTouch: Mapping between touch input with fingers and vibrotactile feedback pattern on the cheek [9].

4 Conclusion and Future Prospects

Tactile tangibles help overcome the problems that occur while communicating over distance. Since communication and interaction over distance is done for different purposes depending on whether the gap between logical or physical distance needs to be bridged, the way in which touch is used variates highly.

When the communication over logical distance is enhanced through the sense of touch, then with the goal to improve the human and machine interaction and make it more effective, comfortable and enjoyable [12]. The *Tactile Touchpad* and *Haptic Pen* improve the usage of the classical devices and let them feel more like their mechanical counterparts, the *Haptic Tabletop Puck* communicates touch through applications designed for tabletops. The *TouchEngine* adds touch to mobile phones and improves user performance while the *MudPad* and its technology opens the door for entirely new applications.

Using the sense of touch to overcome physical distance, interpersonal interaction is about to be improved. The early *InTouch* lets two user communicate abstract tactile information, *Connexus* expands this approach to focus on ambient communication, optional between more than two. *HIM* converts emoticons in chat conversations into tactile feedback. While *ComTouch* enables tactile communication between people using a mobile phone, *CheekTouch* users can directly communicate specific touch behaviours through input patterns.

Mobile phones and tabletop computers are already part of our daily lives. While today only few of them have tactile feedback and mostly in form of vibration, a lot of research is done in this area and especially devices such as the *MudPad* and *Cheektouch* have promising features. The first domain implementing tactile feedback in devices for a large number of consumers will most likely be the smartphone market.

References

[1] S. Brave and A. Dahley. intouch: a medium for haptic interpersonal communication. In *CHI EA '97 CHI '97 extended abstracts on Human factors in computing systems: looking to the future*, pages 363–364, New York, NY, USA, 1997. ACM Press.

[2] A. Chang, S. O'Modhrain, R. Jacob, E. Gunther, and H. Ishii. Comtouch: design of a vibrotactile communication device. In *DIS '02 Proceedings of the 4th conference on Designing interactive systems: processes, practices, methods, and techniques*, pages 312–320, New York, NY, USA, 2002. ACM Press.

[3] M. Heller. *The Psychology of Touch*, page 351. Lawrence Erlbaum Associates, 1991.

[4] H. Ishii. The tangible user interface and its evolution. *Communications of the ACM - Organic user interfaces*, pages 32–36, 2008.

[5] Y. Jansen, T. Karrer, and J. Borchers. Mudpad: tactile feedback and haptic texture overlay for touch surfaces. In *ITS '10 ACM International Conference on Interactive Tabletops and Surfaces*, pages 291–294, New York, NY, USA, 2010. ACM Press.

[6] J. C. Lee, P. H. Dietz, D. L. und William S. Yerazunis, and S. E. Hudson. Haptic pen: a tactile feedback stylus for touch screens. In *UIST '04 Proceedings of the 17th annual ACM symposium on User interface software and technology*, pages 291–294, New York, NY, USA, 2004. ACM Press.

[7] I. S. MacKenzie and A. Oniszszak. The tactile touchpad. In *CHI EA '97 CHI '97 extended abstracts on Human factors in computing systems: looking to the future*, pages 309–310, New York, NY, USA, 1997. ACM Press.

[8] N. Marquardt, M. A. Nacenta, J. E. Young, S. Carpendale, S. Greenberg, and E. Sharlin. The haptic tabletop puck: Tactile feedback for interaction tabletops. In *ITS '09 Proceedings of the ACM International Conference on Interactive Tabletops and Surfaces*, pages 85–92, New York, NY, USA, 2009. ACM Press.

[9] Y.-W. Park, C.-Y. Lim, and T.-J. Nam. Cheektouch: an affective interaction technique while speaking on the mobile phone. In *CHI EA '10 Proceedings of the 28th of the international conference extended abstracts on Human factors in computing systems*, pages 3241–3246, New York, NY, USA, 2010. ACM Press.

[10] E. Paulos. Connexus: a communal interface. In *Proceedings of the 2003 conference on Designing for user experiences*, pages 1–4, New York, NY, USA, 2003. ACM Press.

[11] I. Poupyrev and S. Maruyama. Tactile interfaces for small touch screens. In *Proceedings of the 16th annual ACM symposium on User interface software and technology*, pages 217–220, New York, NY, USA, 2003. ACM Press.

[12] I. Poupyrev, S. Maruyama, and J. Rekimoto. Ambient touch: designing tactile interfaces for handheld devices. In *UIST '02 Proceedings of the 15th annual ACM symposium on User interface software and technology*, pages 51–60, New York, NY, USA, 2002. ACM Press.

[13] I. Poupyrev, J. Rekimoto, and S. Maruyama. TouchEngine: a tactile display for handheld devices. In *CHI'02 extended abstracts on Human factors in computing systems*, pages 644–645, New York, NY, USA, 2002. ACM Press.

[14] A. Rovers and H. van Essen. Him: a framework for haptic instant messaging. In *CHI EA '04 CHI '04 extended abstracts on Human factors in computing systems*, pages 1313–1316, New York, NY, USA, 2004. ACM Press.

Interaction Over Distance In Multi-Display Environments - Analyzing Interaction Techniques

Christine Wagner

Abstract— As multi-display environments have gained more and more popularity in the last few years, the need for suitable interaction techniques also increased. Complex display settings combine multiple heterogeneous display types into one single environment. Apart from traditional devices like desktop computers and external monitors, they also include displays capable of touch input like tabletop or wall-size displays. Common existing techniques like drag-and-drop are often not able to meet all the resulting requirements as they were originally designed for interaction with indirect input devices. As a result of this trend towards complex environments, new suitable techniques have to be developed. Therefore, various factors must be examined thoroughly. Some main factors to be considered are for example the physical arrangement of displays, the used input devices, the reachability of displays in a distributed environment or the way how displays are connected to each other. Although many different interaction methods already exist, only a few of them can overcome most of the difficulties. The examination of several interaction techniques in this paper shows particularly two main problems. One problem is the use of different input devices and the other is the bridging of long distances between displays, especially if they are located across displayless space or behind bezels.

Index Terms—Interaction, Distance, Multi-Display, Display Factors, Input Devices, Reachability, Software-based, Hardware-based

1 Introduction

The term multi-display environment describes a setting of displays composed of several different display types. They can be arranged in many different ways, reaching from simple environments like common multi-monitor systems to complex ones consisting of many heterogeneous displays (see figure 1). Underlying technologies of displays are getting higher developed and as well cheaper. Therefore, using complex and advanced display environments has also become more and more popular in every day life over the last few years. A large field of research is the usage of multi-display settings in group meetings. Several approaches have been proposed in order to support collaborative group tasks like information sharing, problem solving or brainstorming. Examples are *IMPROMPTU* [5], the brainstorming system of Hilliges et al. [9], Rekimoto's multi-display approach [17] or his hybrid workspace [18], *UbiTable* [19] or *i-LAND* [20].

Nowadays, the use of touch displays is widespread, containing not only tabletops or wall-size displays but also smartphones and tablets, which are increasingly used among the population. Therefore, touch capable devices are often combined with traditional devices like desktop computers and external monitors in multi-display environments. This raised complexity of display settings led to an increased demand for new interaction techniques. Moreover, many common interaction methods cannot be fully adapted to this modern systems because they were originally designed for interaction with indirect input devices [4].

This paper first discusses several existing difficulties concerning the interaction over distance in multi-display settings. Section 3 gives an overview of different software- and hardware-based interaction techniques. After that, the introduced techniques are discussed and compared to each other based on different factors. Moreover, the section analyzes the question to what extent the presented techniques could overcome existing problems regarding the interaction over distance. The paper finishes with a summary of the discussed techniques and a conclusion regarding the findings of section 4.

2 Difficulties in Multi-Display Settings

Multi-display settings often contain all kinds of different display types ranging from external monitors to tabletop displays or mobile devices.

- *Christine Wagner is studying Media Informatics at the University of Munich, Germany, E-mail: wagnerchr@cip.ifi.lmu.de*
- *This research paper was written for the Media Informatics Advanced Seminar 'Over Distance', 2011/2012*

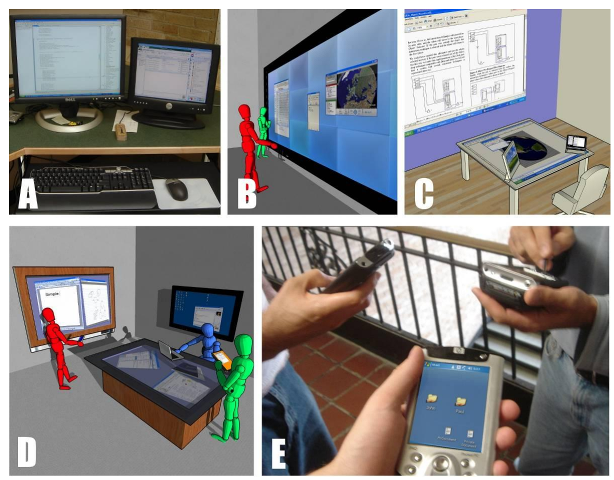

Fig. 1. Different types of multi-display settings: (a) multi-monitor system, (b) large composite display, (c) mixed display work environment, (d) meeting room and (e) multi-mobile environment [15]

As these display types were all developed separately from each other to serve various needs, they differ in many attributes, for example underlying technology, display orientation or used input devices. Moreover, the enhanced use of multi-display environments over the last years also created a need for interaction techniques between two or more displays. In many cases, existing methods cannot be applied to environments which use touch input as these methods were developed for indirect input devices like mice, track pads or joysticks [4]. Therefore, new interaction techniques have to be developed. However, finding a suitable technique for a specific setting can be difficult as various factors (like the characteristics mentioned above) have to be considered [15]. In the following, some of these factors will be further discussed.

2.1 Input Devices

An important aspect in multi-display settings are the used input devices. If several different display types are combined in one environment, it might occur that not all of them can use the same input device [8]. Some displays support only indirect devices like mice or track pads, others are only responsive to direct input as touch input using fingers or a digital pen [4]. As a result, the user has to switch between different devices when interacting with heterogeneous displays. When a user, for example, wants to interact with data on a tablet computer and on an attached external monitor, she has to permanently swap between the digital pen for the tablet and the mouse, as the external monitor is incapable of touch input [4]. Rekimoto already faced the problem of restricted input devices when presenting an early manipulation technique for multi computer environments in 1997 [16].

Moreover, especially interaction techniques using touch or multi-touch input can often only be used on one single display [2] [4]. This constraint is based on the underlying touch technologies as most methods work with sensors situated directly behind or inside of the display screen. The same situation applies for technologies using diodes or cameras which are coplanar to the surface. In order to achieve continuous interaction between several displays, the detection method to capture user input on the display has to be independent from a specific display. In general, indirect input devices like mice are independent from a certain display as the events are captured separately on each display. Still, in flexible or more complex multi-display settings with many displays, tracing the mouse movement between the displays becomes difficult because the single displays are not explicitly linked to each other.

2.2 Physical Arrangement

Another important issue is the physical arrangement of displays. Multi-display environments can consist of many different possible combinations, for instance a common multi-monitor system, a large composite display or a complex environment with various display types (see figure 1). Important factors regarding the arrangement include display size, angle and distance between the displays, number of displays and their alignment to each other. These aspects have to be considered as they can have great impacts on the interaction [23]. Su and Bailey investigated the effects of several display arrangements on user interaction [21] and the results of their experiment confirmed the importance of the physical configuration.

As a part of their research on cross-display movement in multi-display environments, Nacenta et al. defined the physical configuration of displays as their arrangement in logical space [15]. Furthermore, the input model of an interaction technique is considered as important in that context, too. The term input model describes the processing of user input into operations performed on the screen. The mapping between a model and a physical arrangement can be various as the implementation of an interaction technique is based on the logic of an input model. Therefore, interaction techniques can be classified into three types according to their underlying input model: *planar*, *perspective* and *literal*. An interaction method can be considered as *planar* if all used displays are aligned on the same plane and are nearly merged seamlessly into each other when moving the cursor across displays. *Perspective* techniques rely on the user's perspective on the display environment in order to obtain a natural-looking mapping between user input and the resulting output on the screen. Interaction methods which only depend on the physical connection of displays are called *literal*. A specific input model is not needed as the configuration can completely replace it. Furthermore, it is essential to choose the most suitable technique for a desired arrangement. Planar techniques work well in simple environments where all displays are located on the same plane. However, if the setting is more complex and the displays are, for example, located at different angles to each other or to the user [23], the so called *perspective problem* occurs [15]. This term describes the situation when a user is not directly in front of the display and therefore the user's perspective on a display is distorted due to the viewing angle [13] [14]. This can result in many limitations for the user concerning the interaction between displays. Besides reading problems, it can become difficult to move items or interact with them in other ways.

2.3 Reachability

An associated topic is the reachability of other displays in a multi-display setting. The interaction between multiple displays can raise several problems as it is necessary to cross spatial distance between the displays, for instance off-screen areas or bezels [4]. The common technique drag-and-drop uses only a single pen-down event for moving items. If a user, for example, wants to transfer an object into a folder on another display behind a bezel, the object first has to be moved to the display containing the folder. Therefore, it has to be dropped near the bezel in a first drag-and-drop interaction. After the item has been transferred across the bezel, it can be picked up and dragged to the target folder with a second drag-and-drop move. This can be inconvenient and time-consuming.

Furthermore, bridging *displayless space* often requires a reorientation as the user does not receive direct feedback about the current position of the cursor [23]. Targeting across *displayless space* is further examined by Nacenta et al. [12]. The mentioned reorientation problems occur especially when the distance between the displays is large or the displays are not placed in the same angle [23] [3]. Size and resolution of a display or vertical and horizontal offsets between screens can pose orientation problems as well. Figure 2 shows two examples for this problem in a multi-monitor setting. In the first scenario, the mouse cursor is moved from a display with high resolution (*start*) to one with lower resolution (*target*). In the user's perspective, the cursor does not show up at the targeted position as it is shifted by a certain vertical offset (see figure 2, scenario 1a). This results from the fact that the operating system ignores the different resolutions and therefore just relocates the cursor in a supposedly straight way (1b). The second scenario describes a mouse movement between two displays with the same resolution. When the cursor is moved diagonally, it again shows up on the second screen with some offset (2a). As the system is unaware of the passage between the displays, it always shifts the cursor on a straight line (2b).

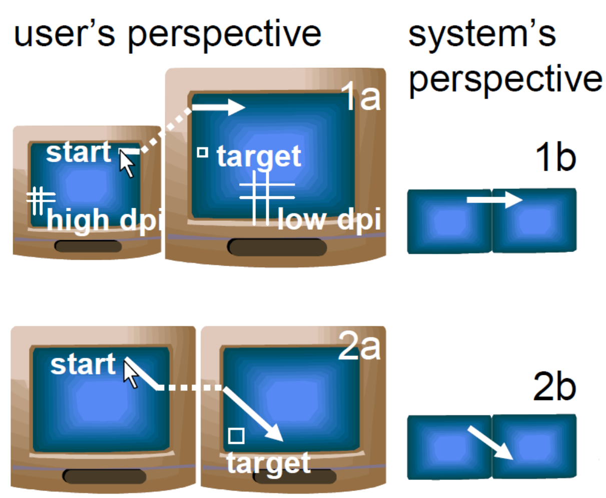

Fig. 2. Resolution (1) and offset (2) issues in a multi-monitor setup [3]

2.4 Connection

The way how displays are connected to each other is another factor that has to be considered when choosing an interaction technique for a multi-display environment. A rough classification can be made regarding the configuration of the environment's system. In a common multi-display setting, all displays are either connected to the same machine or distributed among several independent machines [7]. In the first configuration type, the communication between the displays is simple as they act as one single virtual space and all events are handled by only one common system. This makes a continuous movement of items possible. Though, the underlying system has a certain capacity limit for handling input and only some displays can ensure a wired connection with the system [15]. In the second case, the displays are managed by different, not connected systems, which makes the communication more difficult. One possibility to facilitate exchanging data is connecting the different systems and their attached

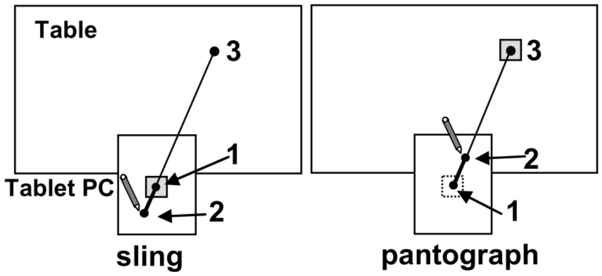

Fig. 3. *Slingshot* and *Pantograph*: (1) starting position of the object, (2) pen location for positioning the object and (3) target position of the object [11]

Fig. 4. *Drag-and-Pop* with rubber band [4]

displays through a network. At the same time, the specification and configuration of these networks often pose a problem due to the need of finding the requested systems among all existing devices in the network. Another possibility is the interaction between several displays via a wireless network [10]. As portable devices are getting more and more popular, the need to spontaneously interact with each other or share information has become essential. In order to enable a dynamic wireless communication, simple methods for quickly establishing a network connection are required. This brings up the *spontaneous device sharing problem*. This problem deals with the question how such networks can be set up easily although the interaction partners have to exchange their individual network addresses first.

3 Existing Interaction Techniques

Traditional methods like drag-and-drop often cannot meet the requirements of today's complex multi-display environments or settings containing a mix of different display types [4]. This situation results in an increased research for new techniques that are able to accomplish an interaction across multiple heterogeneous or distant displays. The existing techniques can be divided into software-based and hardware-based approaches. In the following, a selection of both approaches will be presented and later examined regarding to the difficulties mentioned in section 2.

3.1 Software-based Techniques

This section introduces several techniques that only make use of software for realizing an interaction between multiple displays. Some of these methods are based on the traditional technique drag-and-drop, but are modified in order to meet the new challenges brought up with the usage of heterogeneous or remote displays in multi-display settings. Another approach is the application of representations on the user's current display. These representations can either be certain data objects or the whole surrounding environment. A third group of interaction techniques relies on the concepts of throwing objects to their desired position.

3.1.1 Pick-and-Drop

Rekimoto proposed *Pick-and-Drop* [16] as one of the first techniques for interacting in multi-computer environments. The approach is derived from the commonly used drag-and-drop technique and allows users to treat computer stored data like a real physical object. In order to transfer data from one device to another, the user first has to pick it up by selecting it with the pen device. After lifting the pen, the data disappears and the pen can be moved to the target position without touching the surface. When the user taps the display again, the data appears and is copied to the touched location. As a visual feedback for the user, a shadow of the data is visible if the pen moves near enough to the display's surface. The *Pick-and-Drop* technique uses the concept of *Pen IDs* to transfer the data: Each pen device receives a unique ID and when an object is picked up, a network server called "pen manager" links the ID of the object with the pen ID. Once the user taps the target position, the pen manager initiates the transfer of the data over the network.

3.1.2 Radar View

Radar View [11] is a world-in-miniature technique for interaction with remote displays. A miniature version of the surrounding multi-display environment is displayed on the local screen. The small map only shows up if the user touches an object in order to move it from one display to another. The user can now drag it to its designated position on the appeared map. While the object is moved, a line between the current location of the object and the starting position is drawn. The user simply drops the item when she reaches the target position. Then, the interaction is finished and the object is transferred to the remote display.

Similar techniques using world-in-miniature representations were investigated by Aliakseyeu et al. [1] or Wigdor et al. [22].

3.1.3 Slingshot & Pantograph

Two other software-based interaction techniques are *Slingshot* and *Pantograph* [11], which are both similar to the techniques *Drag-and-Throw* and *Push-and-Throw*, first presented by Hascoët [8]. In the following, the terms *Slingshot* and *Pantograph* are used to describe the underlying concepts. Both techniques are based on the idea of throwing data objects from one point to another while still retaining a high accuracy and low error rates. With this throwing model, dragging over long distances and switching between multiple displays is rendered unnecessary.

Slingshot originates from the metaphor of a physical slingshot or the picture of archery. The user selects the data icon by clicking on it. In order to aim, she then drags it into the opposite direction of the desired throwing direction. For specifying the precise direction, the user can move the cursor to the right or to the left. The range of the shot is defined by the range of the backward movement. If the object is released, it is thrown towards the targeted location. In advance, the covered path of the object and its target position is shown to the user as visual feedback. This makes it easier to estimate the throwing move and to hit the correct position. The left part of figure 3 illustrates the technique.

Pantograph works similar to the *Slingshot* technique but uses a different aiming concept. Instead of defining direction and range of position by moving the selected item backwards, the item is moved forwards. Again, a visual feedback line is shown to support the user. An illustration of *Pantograph* is shown in the right part of figure 3.

3.1.4 Drag-and-Pop & Drag-and-Pick

Drag-and-Pop and *Drag-and-Pick* [4] are similar interaction techniques for touch- and pen-operated systems and are based on the drag-and-drop method. They were developed to support users in interacting with remote content that is hard or impossible to reach. This problem occurs particularly on large displays where the user has to cover a large distance or in multi-display settings when the user wants to interact across physically separate devices. Another problem with interaction can occur in multi-display settings where not all displays support

Fig. 5. *Drag-and-Pick*: Accessing content on an external monitor that does not support touch input [4]

the same type of input device, e.g. settings using external monitors which cannot handle touch events (see section 2.1). Baudisch et al. try to solve these problems with the two techniques *Drag-and-Pop* and *Drag-and-Pick*.

Drag-and-Pop uses representations of potential interaction candidates to overcome the distance to remote display content. These representations (*tip icons*) show up around the current position of the cursor if a user starts to drag a selected icon towards other icons (see figure 4). For reasons of clarity, only icons which are compatible to the dragged data and lie within a certain range of the dragging direction pop up near the cursor. A rubber band links the tip icon with its original icon by illustrating the trace of the tip icon from the original icon to the cursor area (as shown in figure 4). This visualization helps the user to keep the orientation and to find his intended target icon. In order to complete an interaction, the user has to drop the selected icon on the target icon. Moreover, the user has several options to abort the interaction: (1) dropping the icon on empty space or (2) dragging the icon away from the tip icon cluster in the opposite direction. In both cases, all tip icons disappear. However, the user retains the possibility to do a regular drag operation when performing option 2.

Drag-and-Pick is a modification of the *Drag-and-Pop* technique. Instead of only showing a subset of compatible icons, all icons within the angle of the motion direction cluster around the cursor. Furthermore, the interaction is not started by dragging an icon as it happens when using *Drag-and-Pop*. Instead, the user only has to start a drag operation on empty space and all tip icons pop up. Now the user can launch a program or open a new folder because the dropping of the cursor on an icon of any type replaces the double click on that icon. Figure 5 shows an interaction with *Drag-and-Pick* in a multi-display setting consisting of an external monitor and a tablet computer.

3.2 Hardware-based Techniques

In the following, some interaction techniques that are based on hardware-assisted systems will be presented. Nearly all of the methods use camera-based systems, sensors and visual markers in order to track the position and orientation of the user and the displays in the multi-display environment. Apart from this, one paper presents an approach for recognizing hand positions or hand gestures in a three-dimensional input space. Furthermore, another work uses a mobile device for interaction with remote displays.

3.2.1 Hyperdragging

Hyperdragging [18] is an interaction technique which is supported by a camera-based object recognition system. The work of Rekimoto and Saitoh mainly focuses on a smooth interaction between pre-installed (such as tabletop or wall-size displays) and portable computers in working environments. Their main idea was to provide the users with a virtually bigger desktop extending their portable computers in a multi-display setting. For the navigation between the different displays, the input devices of the portable computers are used for the whole environment in order to avoid an unnecessary change of input devices. The interaction technique *Hyperdragging* is used to transfer data between the screens. If a user joins the environment with his laptop and gets seated at the interactive table, a camera installed over the table identifies the laptop by its unique marker, which tells not only the ID of the computer but also its position and orientation in the room. This visual marker is a 2D matrix code attached to the computer. For moving data objects from the laptop to the table, the user can drag the item to the edge of the screen. As soon as it arrives at the edge, the cursor and the item show up on the next display and the user can continue his move. It is also possible to just navigate the cursor over several displays or to move data located on one of the pre-installed displays. Figure 6 shows an example of *Hyperdragging* in a multi-display setting. To determine the correct position of the cursor, all four edges of the laptop screen are marked as mouse-sensitive areas. The cursor position is remapped to the screen when the cursor reaches one of these areas and the computed offset of this remapping is used to show the position of a virtual cursor on the table. Meanwhile, the original cursor stays in the edge area. Additionally, the precise position and orientation of the laptop received from its visual marker is needed to define the position of the cursor accurately. Objects are transferred over the wireless network which is established between all devices in the environment.

3.2.2 Perspective Cursor

Another hardware-based technique for interaction between several displays is *Perspective Cursor* [14]. "Perspective is defined as the appearance to the eye of objects in respect to their relative distance and positions [14]." *Perspective Cursor* uses this principles of perspective to solve the problem that different displays are not spatially aligned to each other and therefore make traditional stitching of multiple displays difficult. Nacenta et al. developed a method which is based on a relative positioning input device and the user's point of view on the displays. The idea of the technique is to combine a three-dimensional model of the environment and all positions of its displays in it with the 3D location of the user's head in this environment. With this information, adjacent displays in the user's point of view can be determined and from this, the movement and position of the cursor can be computed and displayed continuously for the user. In the experimental setting, a Polhemus Liberty tracker with 6-DOF sensors was used to track the positions. The sensor for measuring the position of the user's head was fixed to a cap on the head of the user.

3.2.3 Lift-and-Drop

For an interaction across multiple displays, Bader et al. developed the video-based input device *Airlift* [2]. User input is captured with a sensing system that consists of a stereo camera and several infrared LEDs for a well illuminated environment. The sensors of the camera are

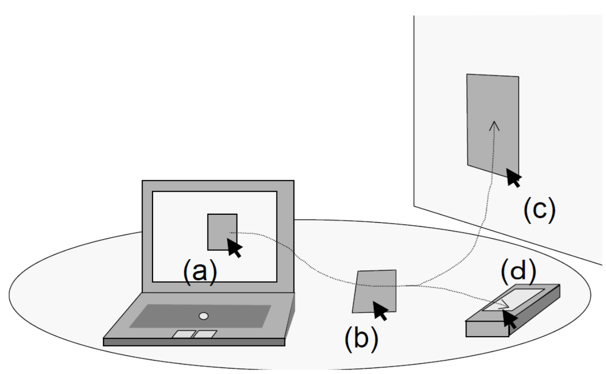

Fig. 6. *Hyperdragging*: Steps (a)-(d) show possible paths of interaction in a spatially continuous display setting [18]

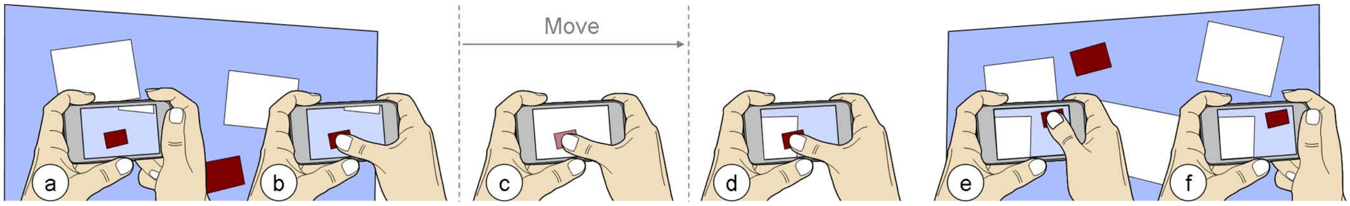

Fig. 7. Interaction with the *Touch Projector*: After aiming at the desired item (a), the user can select it by touching it (b). As the device leaves a display and moves off-screen, the item disappears and is displayed as a thumbnail (c). When the user reaches a second display, the dragged object reappears (d) and can be placed on its target position (e). It is dropped by lifting the finger (f). [6]

mounted above the display setting and can recognize the position of hands, fingertips and learned hand symbols in the three-dimensional input space. Therefore, the acquisition of input data remains independent from the displays and it is possible to interact with remote displays without being on or close to a display surface. The positions of fixed displays only have to be measured once. A global reference system tracks the positions of mobile displays continuously by using attached location markers. Additionally, the system determines the spatial relation between the displays and the user in the environment.

Furthermore, the paper introduces the new interaction technique *Lift-and-Drop*, which was especially designed for the use of a display independent input device like the presented *Airlift*. *Lift-and-Drop* resembles Rekimoto's *Pick-and-Drop* [16] (for details see section 3.1.1) as data can be moved by selecting it and dropping it on its target position. As opposed to *Pick-and-Drop*, the moved object is selected by finger and the user gets permanent feedback on the object's current position due to the continuous acquisition of the user's actions. The position on a display is an orthogonal projection of the finger position in the three-dimensional environment.

3.2.4 Touch Projector

Boring et al. presented the *Touch Projector* [6], a system which uses live video images on a mobile device to interact with content on remote displays. Besides bridging long distances between displays, the technique also addresses the problem of heterogeneous displays with different input devices (see section 2.1) as it is possible to interact with all kinds of displays, regardless of the used input device. Using the *Touch Projector* works as follows: To transfer digital data from one display to another, a mobile device is used to aim at the desired object. The display of the device shows a live video image of the content and now the user can select the object by touching it on the mobile screen. The selected object is dragged on its display as the user moves the device. When the user reaches an edge of the display and moves further into an off-screen area, the object disappears from its original screen and is only displayed as a thumbnail on the mobile device. As soon as the user points at another display with the mobile device, the dragged item shows up again and can be placed on its target position by moving the finger on the mobile's display. For dropping the object, the finger has to be lifted. Figure 7 illustrates the steps of this procedure. Besides a mobile device with a built-in camera, the system uses a software on all participating displays (which are connected through wireless LAN) and a server called "environment manager" for managing the interactions. All participating displays and their content are registered with the manager. The manager is responsible for the overall communication in the environment, especially the handling of touch events and display updates. The built-in camera of the mobile permanently forwards live video frames to the manager in order to identify the orientation and position of the device relative to the other displays in the environment. For tracking and display identification, the technique uses full image processing which basically means finding the best fitting match between the camera image and the other displays.

4 Discussion

All introduced techniques deal with difficulties of the interaction over distance in multi-display environments. As most of them were only developed for specific display settings or to address specific tasks, not every technique is designed to solve all of the existing problems regarding interaction. In the following, the main question will be analyzed whether the presented techniques could overcome some of the problems described in section 2.

4.1 Input Devices

Some displays in multi-display settings only support a special type of input devices. This makes an interaction between varying displays difficult and often requires a permanent switching between different devices. Being one of the first techniques of its kind, *Pick-and-Drop* could not yet solve this limitation, as the method only works among pen sensitive displays. Even though the author emphasizes the possibility of transferring the concept into settings using indirect input devices, the technique still cannot provide the opportunity of interacting between displays with different input technologies. The world-in-miniature technique *Radar View* solves the problem as the interaction only takes place on the display in front of the user. Although the user does not have to switch between displays or input devices within one single interaction, he might have to do it when starting a new interaction. *Drag-and-Pop* and *Drag-and-Pick* addressed the problem, but could not work it out totally. Switching input devices is avoided *within* a single interaction because the operation only has to be performed once on one of the participating displays. Nevertheless, *between* separate interactions, the device still has to be exchanged if the user wants to manipulate an object which is located on a different display compared to the one of the previous interaction. The throwing models *Slingshot* and *Pantograph*, however, support all kinds of input devices. An evaluation of the hardware-based interaction technique *Hyperdragging* shows that it is not necessary to switch input devices while interacting among the environment. The user always performs tasks with the input device assigned to the portable display as the environment serves as a single extended workspace. Moreover, the technique is more suitable for using indirect input devices as they are supported on every display in the setting as opposed to touch input. As *Perspective Cursor* is based on one common mouse input, the technique can be used on all kinds of displays and requires no changing of devices. *Lift-and-Drop* uses touch input for interaction and is consequently only suitable for touch supporting displays. However, the proposed input device *Airlift* can be trained to recognize pointing gestures and hand symbols and therefore can work with all kinds of displays. The system *Touch Projector* addresses the problem of limitations on input devices as it uses a mobile device in order to execute all interactions. Even displays which do not support touch input can be manipulated using the touch input of the mobile as the touch points are transformed onto the remote screen by using full image processing.

Another issue regarding input devices is the fact that many devices are not independent from a specific display. All presented interaction techniques solved this problem by using a server which manages the communication between all participating displays over a network. The server recognizes touch events on the displays and handles them correctly.

4.2 Physical Arrangement

All interaction methods described in this paper were tested in heterogeneous environments with displays varying in factors regarding their physical arrangement. Among several other factors, important ones were display size, angle and distance between the displays, number of displays and their alignment to each other. As all techniques use network managers to communicate within an environment, the physical arrangement did not have an immediate effect on the underlying communication. Difficulties which are related to distance between displays were addressed and also solved in nearly all of the techniques. *Drag-and-Pop* and *Drag-and-Pick* temporarily move representations of the designated objects to the current cursor position, whereas *Slingshot* and *Pantograph* bridge the distance by throwing objects towards their target location. *Radar View* solves the distance problem by showing a miniature representation of the environment on the display the user is currently interacting with. When using *Pick-and-Drop* or *Lift-and-Drop*, the user still has to walk to a distant display in order to drop the grabbed object. Therefore, these two techniques work better for displays within arm range. Both techniques *Perspective Cursor* and *Hyperdragging* use a mouse device on the virtually extended desktops for reaching remote data. For the application *Touch Projector*, a common mobile device can interact with distant screens by using its touch input on live video images. Distortions caused by the viewing angle from the user's current position can be fixed by applying the *Perspective Cursor* technique, which is based on the user's perspective of the room. Furthermore, the angle between displays or to the user and the alignment of the displays were examined for some of the techniques but did not show a main impact on the interaction of the users. The size of a display posed a problem especially when using the mobile device of the *Touch Projector*. The interaction with small objects on the mobile screen led to the fat finger problem. Targeting or dragging tasks were difficult concerning small objects. However, this could be resolved by implementing a special automatic zooming function. Moreover, the technique *Radar View* can also be difficult to use if the display on which the miniature map shows up is rather small. In that case, the shown map can possibly be to small for an easy interaction or the user could have problems to find the desired display on the small representation.

4.3 Reachability

Reachability problems can occur for displays that are located behind bezels or across displayless space. The solutions for this problems are the same as already explained in subsection 4.2 when discussing the difficulties of distance between displays. The need of reorientation poses a problem as display factors like angle or display resolution vary between two interacting displays or the displays are separated by a certain distance. *Slingshot* and *Pantograph* help the user in keeping the orientation in the display environment by giving additional visual feedback. With this feedback, the user can easier estimate the desired location of the object as the target position, the range and the direction of the shot are indicated by a visual line. When *Radar View* is used on a very small display, the user could have problems to orient herself in the even smaller miniature map of the environment. Moreover, a feedback line is drawn as the user is moving an object on the map. The technique *Lift-and-Drop* provides feedback to the user as the already covered distance and the current location of the object are continuously displayed. *Touch Projector* gives immediate visual feedback, too, by showing live video images of the aimed-at objects or displays. Therefore, reorientation is not necessary. The concept of *Perspective Cursor* addressed the reorientation issue by compensating the offsets concerning the start and destination positions on the displays. In this way, screens which are not aligned on the same plane appear to seamlessly blend into each other from the user's angle.

4.4 Connection

The last factor that was explained earlier is the connection between several displays in a distributed setting. Display events can either all be handled by the same system or managed by different machines independent from each other. All presented techniques of this paper were managed through one common server on the network. Neither this configuration nor the type of the network (wired or wireless) did result in any problems regarding the interaction methods.

5 Conclusion and Future Work

This paper presented a selection of four display factors to be considered when choosing an interaction technique for multi-display environments. Moreover, an overview of existing software- and hardware-based techniques was given. The findings of the discussion in section 4 showed that various mentioned factors can effect the interaction over distance in a multi-display setting. Especially the use of different input devices posed a major problem and was addressed in all introduced methods. Nearly all of them just tried to avoid the switching of devices instead of finding a solution to suit all input types. Regarding the physical arrangement of displays and their reachability in an environment, the central problem was reaching screen content on remote displays, especially if they are located across displayless space or behind a bezel. Nevertheless, all described methods could overcome this problem. In conclusion, it can be said that a thorough consideration of several display factors is important to find or develop a suitable technique for interacting in multi-display environments.

For future research, approaches similar to the technique presented by Boring et al. [6] might become more important as the spread of smartphones or tablets is currently increasing. With the wide availability, techniques using a mobile device to interact with multiple displays should be evaluated further because the amount of potential users is growing. Moreover, several problems mentioned above can be solved by using a mobile device, for example the need to switch locations between multiple displays or the reachability of remote displays. Another way to interact with multiple remote displays are free-hand gesture techniques using systems like *Airlift* [2]. Users could easily transfer content between displays in an intuitive way by using their hand as a pointing or interaction device. This could also solve the problem of reachability or the problem of having different input methods for every display. However, research concerning gesture-based approaches is still in its infancy and methods to identify free-hand gestures have to be improved to allow recognition with a low error rate. Moreover, the need of additional hardware makes it more difficult and expensive to use gestures in real-life scenarios. Finally, problems that were not mentioned in this paper like privacy, data security or costs of interaction techniques have to be analyzed as well. These are important aspects for designing methods that can easily be applied in every day display environments.

References

[1] D. Aliakseyeu, M. Nacenta, S. Subramanian, and C. Gutwin. Bubble Radar: Efficient Pen-Based Interaction. In *Proceedings of AVI '06*, pages 19–26. ACM, 2006.

[2] T. Bader, A. Heck, and J. Beyerer. Lift-and-Drop: Crossing Boundaries in a Multi-Display Environment by Airlift. In *Proceedings of AVI '10*, pages 139–146. ACM, 2010.

[3] P. Baudisch, E. Cutrell, K. Hinckley, and R. Gruen. Mouse Ether: Accelerating the Acquisition of Targets Across Multi-Monitor Displays. In *Proceedings of CHI '04*, pages 1379–1382. ACM, 2004.

[4] P. Baudisch, E. Cutrell, D. Robbins, M. Czerwinski, P. Tandler, B. Bederson, and A. Zierlinger. Drag-and-Pop and Drag-and-Pick: techniques for accessing remote screen content on touch- and pen-operated systems. In *Proceedings of INTERACT '03*, volume 3, pages 57–64, 2003.

[5] J. Biehl, W. Baker, B. Bailey, D. Tan, K. Inkpen, and M. Czerwinski. IMPROMPTU: A New Interaction Framework for Supporting Collaboration in Multiple Display Environments and Its Field Evaluation for Co-located Software Development. In *Proceedings of SIGCHI '08*, pages 939– 948, 2008.

[6] S. Boring, D. Baur, A. Butz, S. Gustafson, and P. Baudisch. Touch Projector: Mobile Interaction through Video. In *Proceedings of CHI '10*, pages 2287–2296. ACM, 2010.

[7] M. Collomb and M. Hascoët. Extending drag-and-drop to new interactive environments: a multi-display, multi-instrument and multi-user approach. *Interacting with Computers*, 20(6):562–573, 2008.

[8] M. Hascoët. Throwing Models for Large Displays. In *Proceedings of HCI '03*, volume 2, pages 73–77, 2003.

[9] O. Hilliges, L. Terrenghi, S. Boring, D. Kim, H. Richter, and A. Butz. Designing for Collaborative Creative Problem Solving. In *Proceedings of C&C '07*, 2007.

[10] K. Hinckley, G. Ramos, F. Guimbretiere, P. Baudisch, and M. Smith. Stitching: Pen Gestures that Span Multiple Displays. In *Proceedings of AVI '04*, pages 23–31. ACM, 2004.

[11] M. Nacenta, D. Aliakseyeu, S. Subramanian, and C. Gutwin. A Comparison of Techniques for Multi-Display Reaching. In *Proceedings of SIGCHI '05*, pages 371–380. ACM, 2005.

[12] M. Nacenta, R. Mandryk, and C. Gutwin. Targeting across Displayless Space. In *Proceedings of SIGCHI '08*, pages 777–786. ACM, 2008.

[13] M. Nacenta, S. Sakurai, T. Yamaguchi, Y. Miki, Y. Itoh, Y. Kitamura, S. Subramanian, and C. Gutwin. E-conic: a Perspective-Aware Interface for Multi-Display Environments. In *Proceedings of UIST '07*, pages 279–288. ACM, 2007.

[14] M. Nacenta, S. Sallam, B. Champoux, S. Subramanian, and C. Gutwin. Perspective Cursor: Perspective-Based Interaction for Multi-Display Environments. In *Proceedings of SIGCHI '06*, pages 289–298. ACM, 2006.

[15] M. A. Nacenta, C. Gutwin, D. Aliakseyeu, and S. Subramanian. There and Back Again: Cross-Display Object Movement in Multi-Display Environments. *Human-Computer Interaction*, 24(1-2):170–229, 2009.

[16] J. Rekimoto. Pick-and-Drop: A Direct Manipulation Technique for Multiple Computer Environments. In *Proceedings of UIST '97*, pages 31–39. ACM, 1997.

[17] J. Rekimoto. A Multiple Device Approach for Supporting Whiteboard-based Interactions. In *Proceedings of SIGCHI '98*, pages 344–351. ACM, 1998.

[18] J. Rekimoto and M. Saitoh. Augmented Surfaces: A Spatially ContinuousWork Space for Hybrid Computing Environments. In *Proceedings SIGCHI '99*, pages 378–385. ACM, 1999.

[19] C. Shen, K. Everitt, and K. Ryall. UbiTable: Impromptu Face-to-Face Collaboration on Horizontal Interactive Surfaces. In *Proceedings of UbiComp '03*, pages 281–288. Springer, 2003.

[20] N. A. Streitz, J. Geiß ler, T. Holmer, S. Konomi, C. Müller-Tomfelde, W. Reischl, P. Rexroth, P. Seitz, and R. Steinmetz. i-LAND: An interactive Landscape for Creativity and Innovation. In *Proceedings of CHI '99*, pages 120–127, 1999.

[21] R. Su. Put Them Where? Towards Guidelines for Positioning Large Displays in Interactive Workspaces. In *Proceedings of INTERACT '05*, pages 337 – 349, 2005.

[22] D. Wigdor, C. Shen, C. Forlines, and R. Balakrishnan. Table-Centric Interactive Spaces for Real-Time Collaboration. In *Proceedings of AVI '06*, pages 103–107, 2006.

[23] R. Xiao, M. Nacenta, R. Mandryk, A. Cockburn, and C. Gutwin. Ubiquitous Cursor: A Comparison of Direct and Indirect Pointing Feedback in Multi-Display Environments. In *Proceedings of GI '11*, pages 135–142. Canadian Human-Computer Communications Society, 2011.

Collaboration over Distance

Tim-Fabian Köck

Abstract— Collaboration over distance nowadays takes place as a result from globalization. Companies do global dispersal or even ”outsourcing” whereby company departments are managed by an external company. Making use of this strategy is based on intenting to achieve a reduction of costs.

This business strategy also involves disadvantages: due to distances so called face-to-face meetings are only practicable in a limited or hindered way due to costs and time expenses come along with it. To avoid those costs, technical tools are increasingly used to collaborate over distance, trying to replace expensive co-located meetings that way. For making use of technical tools, disadvantages appear as well, which are justified by a lower quality level in different channels of human perception compared to a face-to-face meeting.

Perceptional problems caused by using technical tools for communication over distance can lead to communication problems in general which in turn affects the productivity of working schedules in a negative way. For that reason the question is which techniques might play an important role in the future to successfully enhance the productivity for collaboration over distance.

This paper gives an overview of currently available technologies for communicating over distances. In addition advanced prototype technologies from research are being introduced. Finally a combination of prototype functionalities is theoretically sketched to act upon a bunge of problems which occur on collaboration over distance.

Index Terms—collaboration over distance, communication technologies, collaborative systems, communication problems, prototypes

1 Introduction and Motivation

Collaboration over distance takes place in many companies across many different markets. The reasons therefor mainly lead back to geographically distributed company structures and terms of “Outsourcing” as results from globalization. Referring to Kinkel and Lay [2], outsourcing was already in 2003 a commonly used instrument by german companies. For example 38% of the companies were outsourcing parts of their IT departments. The advantages of outsourcing are justified by the achievement of the reduction of costs and rating that outsourcing-partners provide demands more competent and efficient compared to the appropriate company.

Nevertheless negative impacts appear through collaboration over distance. James D. Herbslebs et al., in a study of distance, dependencies, and delay in a global collaboration state “that the frequency of communication drops off sharply with physical seperation among co-workers’ offices and that the sphere of frequent communication is surprisingly small.”[1]. Yankelovich et al. state “effective communication is an important factor in successful business outcomes” as well as “lack of opportunity to build relatonships through social interactions decreases trust, and trust among co-workers is equally important in fostering team effectiveness” [9]. A reduction of communication and consequential less effectiveness and productivity can be caused for different reasons.

First of all the geographical distance between co-workers leads to hindered conditions for co-located so-called face-to-face meetings. As a result less of those meetings take place due to travel costs and loss of time would result for it’s achievement [3]. For that reason communication needs to be carried out differently for example using telephone, electronic mail or further more modern opportunities of technical communication which will be named in the related work section of this paper. Not being co-located for communication can delimit the possibilities of underlining the own point of view, for instance using gestures or mimic. In case of telephone or electronic mail those two nonverbal ways of communicating are not provided.

Besides the loss of nonverbal communication, split working environments due to distance are a further reason, that might lead to a loss of understanding a colleagues’ issue. Not understand his current problem leads to not being able to help him directly, as it would be possible if people were working co-located. For instance, a remote colleague does not have the ability to touch or move things that reside inside the working environment of a remote colleague.

A further reason for reduced communication over distance can be organisational problems. In case co-workers don’t know which colleague to refer to in a remote-located office for solving a specific problem, the process of beginning a communication is hindered. Hence finding a solution takes more time compared to the situation that colleagues work close together in a co-located environment.

In case cooperation is distributed over diverse countries, further problems can result from different cultures, languages and time zones and lead to less intense communication. James D. Herbslebs et al. [1] evaluated in an empiric study that people prefer the telephone for non-face-to-face meetings, in case the communication was with a native speaker of the same language. In case it was with a native speaker of a different language, communication using electronic mail was preferred. Obviously people tend to avoid communcation the direct way in case that communication problems might occur due to native speakers of different languages are involved.

As described collaboration over distance contains a diversity of possible communication problems. Less communication leads to less speed of the working processes and finally to less productivity [1]. In a more and more globalized world speed is determined as one of the most important factors to succeed with new products to markets and especially in technically based markets to claim an innovative reputation [1]. This paper reviews existing technical concepts for remote based communication. Furthermore it describes visions on how to possibly make use of different techniques in combination with each other to finally lead to a better collaboration over distance with less communication problems.

2 Requirement Analysis

Referring to the described communication problems, which can result from cooperation over distance, the requirement analysis defines the needs in a theoretical way which technical systems should provide for an improvement of remote-based communication.

Remote-based colleagues should be able to find each other quickly in case of related problems that can be solved by teamworking together with people having commonly or more specific competence regarding the topic of the problem. To make it easier finding someone, a system should exist where all employees of a company are listed including

- *Tim-Fabian Köck is studying Media Informatics at the University of Munich, Germany, E-mail: tim.koeck@campus.lmu.de*
- *This research paper was written for the Media Informatics Advanced Seminar ’Over Distance’, 2011/2012*

their contact data, competences and current projects they are working on for instance.

To enhance nonverbal communication, techniques are required which also convey a video channel.

Furthermore advanced technology setups are necessary to make a virtual sharing of distant working environments possible and leading to a better understanding of a remote colleagues problem.

To reduce the problem of language barriers, caused through different native speakers, a system that provides a computer-based real time translation function would be visionary.

3 RELATED WORK

After introducing a chain of communication problems that can occur on communication over distance and after deriving requirements from those problems, this chapter will give an overview of available systems for collaborating over distance. In addition concerning topics from research are being introduced.

3.1 Available technologies

To ensure communication over distance, a wide spectrum of internet based services is available for satisfying different requirements.

3.1.1 JIRA

JIRA is a web-based application which is often used for managing and tracking projects, tasks, problems and errors. Primary it was developed for managing different steps in the process of software development, but is as well popular and often used for non technical task management in general. It is built on Java EE which makes it work platform independently on different operating systems and databases.

JIRA contains a management functionality for projects and issues. For each entry (ticket) inside one of the two categories, the system provides the functionality for users (holding appropriate rights) to comment on the project/issue and upload media to it. By that way content can be collected inside a ticket and the different process states are automatically being documented.

Tickets can be assigned to users by each other and different tickets belonging together can be linked in case a logical dependency exists between them. Furthermore filters can be setup, for example ”only show tasks which are currently assigned to me”.

In addition Jira provides detailed statistics on the progressing and resolving of tasks. Furthermore it contains a customizable dashboard functionality where activitiy streams and statistics can be displayed as an overview on the startpage after successfully logging into the system.

3.1.2 Skype

Skype is an application that is available for many different computer platforms, as well as for mobile devices. It provides an instant messaging service, phone and video conferencing as well as screen sharing. Besides contacts can be grouped and organized.

Skype offers two possibilities to make phone calls. While different people using skype are online at the same time, free phone calls can be made using the internet connection. In addition so-called ”Skype-Out” calls are possible to do which connect the caller with a persons' telephone connection of the conventional telephone network or of a mobile phone network by paying an amount per minute.

Moreover Skype offers the possibility of phone conferencing. Talking to more people on one call is for free in case everybody uses Skype. It also enables the possibility to connect people to a conference call using a connection to a mobile or conventional telephone network which is not free of charge.

Video calls are also provided. Establishing a video-connection between two people is for free, whereas a multi-point connection is not free of charge.

Screen sharing is a further feature offered by the Skype software. The business model is the same as for video calls: one-point screen sharing is free of charge whereas multi-point transmission is not.

3.1.3 Webex

Webex is an online web conferencing system. It runs inside the users browser and supports audio, video and screen sharing.

3.1.4 Wiki

A Wiki is a hypertext based system for websites. It is a simplified content management system which provides editing capability of contents and a set up and management of user accounts. Usually it is used for informative purposes.

It contains a simplified markup language. Using it is essential to format texts, embed pictures and create links while it does offer only limited functionality for the user to design the layout of a wiki page. The reason for it's limitation is the intention to achieve a consistent template structure and style for the information pages.

3.1.5 Shared calendars

Shared Calendars are used to see the appointments a person has planned and when free time slots are available to arrange meetings involving him.

3.1.6 WebDAV

WebDAV (Web-based Distributed Authoring and Versioning) is an open standard to allocate Data online. Users can access Data like using an online harddrive.

Technically WebDAV is an extended HTTP/1.1 protocol that overrides limitations of HTTP. In contrast to other protocols like the File Transfer Protocol (FTP), WebDav uses the HTTP typical Port number 80. For that reason administration complexity is kept low and the security level high.

Moreover user permissions can be managed by the WebDAV system. In normal case, accessing files remotely requires username and password. To establish a connection between a personal computer and a WebDAV system usually no software is needed (compared to FTP: FTP client software needed) due to many operating systems support access to a WebDAV system by using the file manager.

WebDAV is a solution for sharing files which offers a high degree of safety with minimal administrational complexity. In Addition it is a very comfortable way for sharing files by accessing data using the users' system-specific file manager he is used to.

3.2 Research

This chapter introduces advanced prototypes from different fields of research which could play an important role for future development of advanced technologies for collaboration over distance.

3.2.1 VideoDraw

Based on studies how people collaborate using shared drawing spaces, John C. Tang and Scott L. Minneman [7] developed VideoDraw. It is a prototype of a shared sketchbook among two or more remote collaborators which is based on video technology. Each collaborator draws on a screen. Above the screen a camera is placed that transmits the image of the current sketch to each one of the other collaborators screens (see Figure 1). Hence the system at all time provides a complete image to all participants, persisting of real and video marks. The setup of the system provides concurrent acces to the drawing space by all participants and in this way simulates sharing a piece of paper. As an implication of the cameras positioning even hand gestures are being transmitted. For that reason one nonverbal way of communication is made available. In the course of studying the use of VideoDraw by Tang and Minneman, it was observed as a very important feature to enhance the quality of the collaborative process. The technical features of VideoDraw contain two advantages compared to a co-located environment: hands can be closer together and can even overlap without collision occurance as it would happen in a non-virtual environment. Besides every participant has access to the full drawing canvas at all time without a need of standing up and moving around the table or walking up to the chalkboard or whiteboard.

Next to the drawing and writing functionality and transmission of hand gestures above the sketching canvas, VideoDraw provides further communication links. In front of each participant a video camera and a screen is positioned.

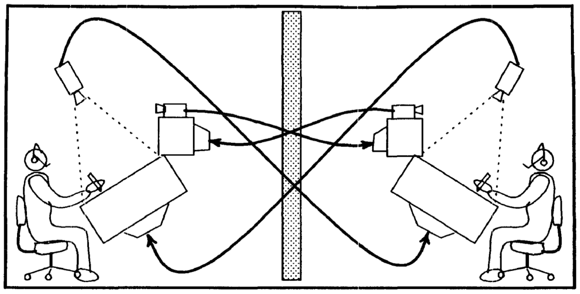

Fig. 1. Schematic diagram of 2 person videodraw [7]

The camera transmits real time video of the user while the monitor displays the frontal view of the other collaborator's faces and upper bodies captured by the cameras located at their working places. To transmit audio signals and provide talking, telephone headsets are used.

VideoDraw is a prototype of an advanced technology to enhance Collaboration over Distance. It provides a virtual shared working environment for real time concurrent sketching. Moreover it provides further perceptional channels like talking, looking at each other and gesturing activity while working together distributedly. All provided features by this system create a novel sense of spatial relationship among remote located collaborators and might lead to a better and faster communication.

3.2.2 VideoWhiteboard

On empirical studies of VideoDraw the development of a further sketching and drawing prototype is based which is called VideoWhiteboard [6]. The limitations observed from VideoDraw were the following ones: a relatively small video display which only provides a small canvas for texts and drawings, parallax problems, only being able to erase own selfmade marks and blocking the overhead camera with one's head. Instead of modifying VideoDraw this further prototype was developed. It simulates sharing of a whiteboard between distant colleagues. Technically each site is provided with a wall-mounted rear projection screen and a video camera plus a video projector which are both positioned behind the screen (see Figure 2).

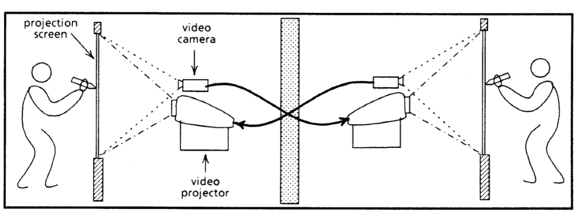

Fig. 2. Schematic of VideoWhiteboard system between two sites [6]

Each camera captures video of the user working on the board and sends it to the projector of the remote colleagues, which presents the image on their screen. The camera records the marks each user does on the board as well as each users silhouette due to it's shadow that shines through the board. In that way gestures and actions of colleagues that are physically remote from each other are transmitted aditionally to the shared drawings. The drawings in turn contain co-located and remote marks. Second ones were created by the remote collaborators and are added through the projection. Furthermore an audio channel is provided for talking with each other. The advantage of VideoWhiteboard compared to VideoDraw is the larger size of the screen. It offers a bigger canvas and even makes it possible for more users to work together around the screen on one site. Whereas VideoDraw transmits colored pictures of the face and and hands, VideoWhiteboard transmits the upper body gestures but only as a silhouette. For that reason the perception of gesture is enhanced but mimic is not possible to identify due to the fact that only the silhouette is being displayed.

As VideoWhiteboard was developed after the observation of Videodraw it overcomes some of the limitations of VideoDraw. Likewise a trade off takes place concerning the transmitting of mimic. The observation of VideoWhiteboard in use describes that "it feels like the remote collaborator is located on the other side of the screen, instead of in a remote location" [6]. This illusion can be seen as an advantage and might possibly add great value to the enhancement of collaborating more successful over distance. Finally it can be reasoned that both systems provide drawing activity that people are familiar with from co-located human-to-human collaboration.

3.2.3 Office Central

With "Office Central" [9] a further interesting concept was found while doing research on collaboration over distance. It is a project by Sun Microsystems that was created to make "remote workers advertise their presence in public spaces within offices" and "encourage informal, unplanned interactions between remote workers" [9]. The system's technology is based on a screen in combination with a camera and audio speakers. It is setup in a public place of a company, for instance inside the cafeteria. The identification of local people is carried out by RFID technology. As soon as a local worker enters the system's radius which can be tracked by the RFID antenna, he is being registered to the system and recognizeable by remote colleagues. Distant people can see names and photos of all the people in the environment of the Office Central system as well as a video stream. Furthermore two high-quality microphones provide a stereo broadcast of the attendant people's voices which are part of establishing the conversation over an audio channel. The system also displays an allusion to further remote users in case local people are chatting with a physically remote colleague. Hence the remote user percepting it the system offers the opportunity to him to immediately join the chat-conversation.

3.2.4 t-Room

t-Room is a distributed multiparty tabletop system introduced by Yamashita et al. [8]. It consists of multiple cameras, displays, a projector and a working table. The cameras take high definition pictures which are displayed at the remote t-Room's eight 65-inch LCD panels that are mounted to the walls around the table. The table is located at the center of the room. The cameras are installed above the screens and a projector is arranged atop the center of the table (see Figure 3 and Figure 4).

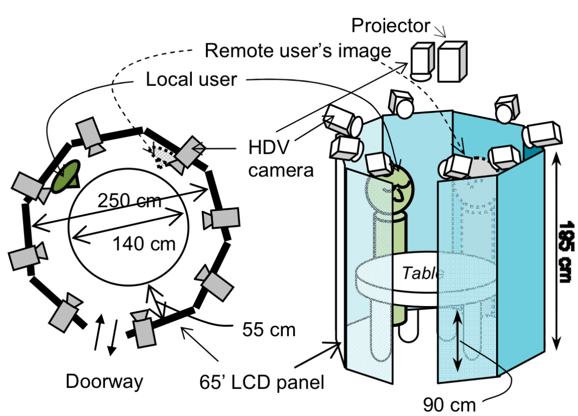

Fig. 3. Hardware Design of t-Room: Top-view (left) Bird's eye view (right) [8]

The LCD pannels provide a 360 degree view of the remote participants surrounding the distant table in their room. Furthermore the projecter displays the remote hand gestures which are made above the table and its re-arrangement of the remote working objects placed on

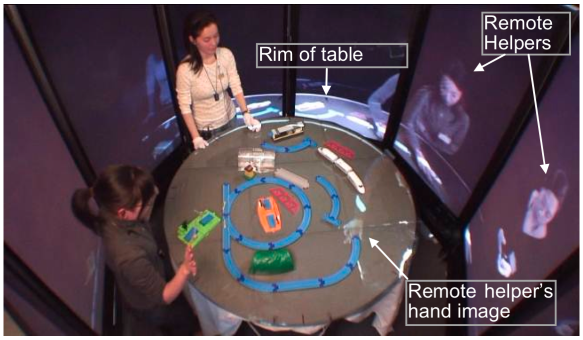

Fig. 4. Four collaborators working around a shared tabletop with upper body view [8]

the table. A space between the table and the wall allows participiants to move to all possible positions around the table. The study of setting up two differently located t-Rooms and actively letting different groups of participants use them was made to examine two things:

- How much does upper body view improve collaboration over distance and speed up the working process?
- How position participants themselves around the table and might upper body view interfere with positioning?

Referring to the question of how much upper body view enhances task performance, two different setups were established: one using the eight surrounding screens and one with turning them off to achieve an environment without upper body view. The result states that "Participants in the upper body view condition completed the task considerably faster than those without upper body view" [8] and it was determined that participants often looked up to to glance at the remote colleagues. Asking them after the participation was ended, mostly they answered that confusion and even anxiety was experienced while using the setup without upper body view. Concerning the positioning it was examined that helper-worker pairs selected overlapping positions for sharing the same perspective onto the table with the remote collaborator.

Even if the study shows that upper-body view in this case enhances collaboration over distance, in comparison to a co-located environment tasks were achieved significantly slower. Many different factors can be responsible for this result and need to be evaluated in further studies.

3.2.5 NESPOLE!

NESPOLE! is a speech translation system [4]. It is a collaborative project between three european research labs and two industrial partners and provides the translation of four languages: Italian, German, French and English. The practical use scenario for NESPOLE! is a website through which the user/customer can get live support in case of detailed or special questions which are not answered on the website itself. An online video-conferencing functionality connects the customer to a human agent. The video-conferencing system provides speech, video and gesture. In addition it contains a real-time translation functionality for the audio and the video channel (in the form of subtitles). In case the user and the agent are not speaking a common language this functionality is the key for compensating language barriers between the involved parties. As a further feature the system integrates a whiteboard application which allows the agent to send web pages to the customers screen and both making annotations to it. The system is developed assuming a minimal hardware and software setup on the user's side, comprehending standard audio and video hardware, a web browser software as well as a webcam and a commercially available internet video-conferencing software. This assumption of available hardware and software components does not include a speech translation software locally on the users PC. Instead the NESPOLE! system is built on a server-type architecture to provide speech recognition and translation.

As afore described, NESPOLE! is a very extensive prototype which offers a lot of functionalities for bridging communication barriers between different native speakers. It provivides speech, video, gesture and also offers a whiteboard application for sharing files and creating marks/annotations on them. All functionalities are allocated with a minimal supply of hardware and software components by the user which could play an important role for acceptance criterias in case of industrial use.

3.2.6 Remote Control to improve Customer Service

Stockburger and Fernandez [5] describe how the Amercian University Help Desk uses remote utilities to give active support on distance to the university's students and employees. Remote control help allows a faster response to technical problems which originally requires an onsite technician to make a visit. Many remote control systems deserve a local setup on the users machine to allow access for the technician. With Lotus Sametime a possibility is used by the American University Help Deks that does neither require a software installation, nor a network configuration on the distant users computer to get started. The only thing a user has to do is to grant permission to a helpdesk analyst. That he can do by clicking on a meeting link at a website and in the next step clicking a button to allow access which includes screen sharing and remote control. The article also mentions how different attitudes can be on letting someone else remotely control your computer:

- The positive attitude towards remote control is appreciation on not having to wait for hours or even days until a technician visits to help and that they just have to sit back and watch how their cursor fixes the problem by itself.
- The negative attitude is based on suspicion of security issues in terms of the computer being watched and controled by others that are distantly based.

In case of disregarding possible social inacceptance for security reasons the presented way of using a remote control in combination with screen sharing functionality over distance is a great opportunity to solve problems without a need of co-location. Especially Lotus Sametime offers an easy to use possibility without a need of complicated software setup on a users' machine carves this technology very attractive due to great usability without technical circumstances.

4 PRACTICAL AND VISIONARY USE

After describing currently available technologies and visionary prototypes of systems to enhance collaboration over distance, this chapter describes how combining different aspects of the introduced sytems can (combining available technologies) and could (combining prototypes) lead to an enhancement of distant collaboration by considering the described problems mentioned inside the chapter "Introduction and Motivation".

From my own practical experience on the use of available technologies to collaborate over distance, collected while working for three years for a company which offers web-based services around 32 countries worldwide and which contains different offices and partners around the globe, a combined use of technologies in a real world scenario will be described at this point.

The main channels for distant-communication, colleague data management and file sharing are electronic mail plus telephone and the available technologies mentioned inside the Related Work section: Skype, WebEx, JIRA, Wiki, Shared Calendars and WebDAV. Skype and WebEx are used for Instant Messaging, Phone Conferencing and Screen sharing. Experiences made by working as part of the Marketing-Design team, JIRA is used as a ticketing system to manage, comment and share files among tasks that are preset by the supervisors. A Wiki is setup which contains a list of all employees. The page of each employee provides an avatar thumbnail, information on where the person is geographically based and for which department he is working and which role he plays inside it. In addition contact information is also provided like E-Mail and Instant Messenger names. Furthermore each employee can publish information what projects he is currently working on and give information of how the process develops. Hence the wiki can be seen as a digital and easy to update ”employee-book” which helps getting an overview of (remote based) colleagues, their competences and contact data. Shared calendars are used to enable the possibility of coordinating (virtual) meetings more efficiently and knowing when a person is available and/or not available during the week. Furthermore to share digital media and files between distant colleagues, WebDAV and FTP Servers are used for sharing and providing online access from whereever an internet connection can be established.

The combination of all these technologies enables great flexibility for differently located people to communicate with each other and to have access to information on colleagues, states and processes of tasks and projects. This offers new and already present forms of working independently from where someone is geographically based.

As described before and in articles that were found during research on this topic, the fact encounters that distributed virtual meetings by using device-to-device communication are not as productive as original face-to-face meetings (for example [8]: ”Task performance was still significantly lower than face-to-face tabletop collaboration”). The question is how much visions and prototyped systems could possibly contribute to an enhancement of communication and effectiveness in the future while collaborating over distance.

Relating to the problems introduced in chapter ”Introduction and Motivation”, this part tries to sketch how those could be counteracted by making use of different parts of prototypes in combination with each other. In general less communication takes place between remote located co-workers, which leads to less trust and again leads to less effective business outcome and productivity [9]. More concretely the following problems were described:

- Less face-to-face meetings. Instead communication over technical channels which often are limited compared to co-located meetings (loss of nonverbal ways of communication: for instance no gestures, mimics)
- Less common interaction possibilities in case of working in distant environments
- Organisational problems like not knowing who to respond to on distance in case of an issue
- Cultural differences, language barriers and different timezones

One of the biggest challenge in collaboration over distance is to provide same conditions for working remote as if people were working co-located. All regarded studies on the topic of technologies for distant collaboration lead to the same conclusion: making remote collaboration as effective as co-located collaboration with any existing technology or prototype is not achievable. The question is to which degree it is possible to enhance current methods of distant collaboration, or if more innovational methods are needed in this field for achieving a higher degree of effectiveness. Some recently described prototypes that were developed on the basis of earlier research on this topic offer advanced technological possibilities.

Parts of related work prototypes can be aligned with the introduced problems:

- Almost all described prototypes contain a video channel to transmit nonverbal ways of communicating. An exception is made by Office Central which only provides one-sided video transmission. VideoWhiteboard only captures shapes of participants which automatically excludes a transmission of mimic.
- To counteract the loss of common interaction possibilities with VideoDraw a shared sketchbook environment is presented. VideWhiteboard is an application providing a shared whiteboard environment over distance. In addition t-Room describes a concept of setting up equal working environments over distance which enhances visual perception on one after another steps in a process. The advanced visualisation techniques simplify receiving help by remote colleagues. One sided remote access to a distant person’s computer is also a way of direct interaction between distantly located working environments and is introduced by the paper about the American University and it’s use as a help desk application. Furthermore the NESPOLE! system either implements a whiteboard application that enables opportunity of remotely working together on a virtually shared object.
- Office Central could be used as it is to convey communication and awareness between distant colleagues.
- To bridge language barriers, the concept of NESPOLE! offers great opportunity with it’s server-sided real-time translation functionality.

As demonstrated many existing systems and prototypes were already invented and (partly) realized to counteract the problems caused by collaboration over distance. Depending on the class of business a company belongs to, requirements might differ. Concerning the described situation of a global player in distributing online web services, advanced technologies for shared workspace access could be a valuable advancement for different departments of such a company (for example for a succesful cooperation between marketing and design teams). Moreover a real time translation system can be useful in a globally operating company, even if all employees are educated in speaking english. The educational degree, experience and different accents are challenges that are essential to master while communicating through a foreign language at work.

A combination of real time language translation provided by NESPOLE! with features of VideoDraw (see Figure 5) for advanced common interaction possibilities would lead to a high degree of measuring the requirements for more effective collaboration over distance by fighting three of the four described main problems:

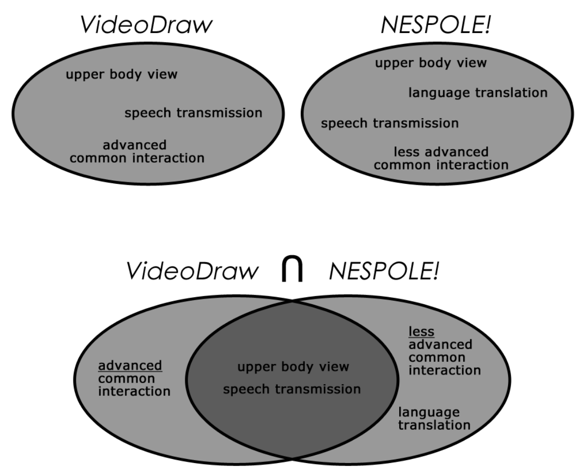

Fig. 5. Combining features of VideoDraw and NESPOLE!

Both systems provide audio channels for the transmission of speech as well as video conferencing functionality that provides an upper body view of remote colleagues. This functionality enables nonverbal communication features. Both systems also provide functionality for common interaction between distributed working environments. With NESPOLE! a software for a shared whiteboard on a desktop computer is contained. By combining the functionalities of the two systems, this feature could be replaced by the more advanced common interaction functionality of VideoWhiteboard which allows direct marking on a screen using a digital pen. This can be interesting for collective planning, for instance in software development processes. It would be even more interesting for cooperated work of a distributed design team. Visionary is running Adobe Photoshop or Illustrator for instance on distant synchronized drawing tablets with built-in screens to directly draw onto. Direct markup inside the sketches by different parties would also allow incorporating supervisors to give feedback which could be realized by the designers in real time. From my experience this can be very efficient in case many people are involved in the process of deciding what the final version has to look like. The more people are involved into such a process, the harder it is to bundle feedback and design a solution that satisfies all attendants. In case you bring all participants to one table, decisions commonly are made which expectations of each feedback need to be realized. As soon as people agree to each other and changes are quickly done (due to it is the finalisation), direct feedback can be given and further changes be made in one to anothers feedback loops. This way of working is much more efficient. Otherwise feedback loops by electronic mail or the planning of further conference calls lead to less productivity by reason of "wasting time" to get forward in processing.

Not to forget, by combining features of VideoDraw and NESPOLE!, the language translation is also provided in case of many different native speakers are involved in a project. Synchronized live sketching functionality in combination with live language translation might be envisioned as a setup to enhance collaboration over distance. Functionality for recording meetings conversation is also conceivable. The recorded audio file could be automatically stored on a shared drive together with the graphical result that was sketched during the collaboration.

5 CONCLUSION

This paper analyzes problems being caused by collaboration over distance. It figures currently available technologies that are already used by the industry to downsize problems, occuring due to collaboration over distance, and lead to negative impacts on productivity. In addition prototype systems are pictured including advanced technological functionalities against distant collaboration problems. Finally an idea is depicted by combining visionary features of two prototypes in relation to the requirements for enhanced collaboration over distance.

More concrete examinations are necessary to build a realistic set of concepts for real use in industry and a more and more globalized world. This work does not accomplish analysis on costs. Remember: one of the main reasons for distributed work (for instance due to outsourcing) is saving costs. Replacing face-to-face meetings by device-to-device meetings is done for the same reason in the next step. For that cause, costs need to be considered. A technological advanced system has to be affordable which means it costs have to provide a well performing cost-benefit.

Moreover acceptance criteria and usability of new technologies can also play a key role in succeesfully enhancing collaboration over distance. Further studies on that topic might be elementary and should not be disregarded.

REFERENCES

[1] J. D. Herbsleb, A. Mockus, T. A. Finholt, and R. E. Grinter. Distance, dependencies, and delay in a global collaboration. In *Proceedings of the 2000 ACM conference on Computer supported cooperative work*, CSCW '00, pages 319–328, New York, NY, USA, 2000. ACM.

[2] S. Kinkel and G. Lay. Fertigungstiefe: Ballast oder kapital? stand und effekte von out- und insourcing im verarbeitenden gewerbe deutschlands. Technical report, 2003.

[3] E. Koh. Conferencing room for telepresence with remote participants. In *Proceedings of the 16th ACM international conference on Supporting group work*, GROUP '10, pages 309–310, New York, NY, USA, 2010. ACM.

[4] A. Lavie, C. Langley, A. Waibel, F. Pianesi, G. Lazzari, P. Coletti, L. Taddei, and F. Balducci. Architecture and design considerations in nespole!: a speech translation system for e-commerce applications. In *Proceedings of the first international conference on Human language technology research*, HLT '01, pages 1–4, Stroudsburg, PA, USA, 2001. Association for Computational Linguistics.

[5] S. Stockburger and T. Fernandez. Virtual onsite support: using internet chat and remote control to improve customer service. In *Proceedings of the 30th annual ACM SIGUCCS conference on User services*, SIGUCCS '02, pages 143–147, New York, NY, USA, 2002. ACM.

[6] J. C. Tang and S. Minneman. Videowhiteboard: video shadows to support remote collaboration. In *Proceedings of the SIGCHI conference on Human factors in computing systems: Reaching through technology*, CHI '91, pages 315–322, New York, NY, USA, 1991. ACM.

[7] J. C. Tang and S. L. Minneman. Videodraw: a video interface for collaborative drawing. In *Proceedings of the SIGCHI conference on Human factors in computing systems: Empowering people*, CHI '90, pages 313–320, New York, NY, USA, 1990. ACM.

[8] N. Yamashita, H. Kuzuoka, K. Hirata, S. Aoyagi, and Y. Shirai. Supporting fluid tabletop collaboration across distances. In *Proceedings of the 2011 annual conference on Human factors in computing systems*, CHI '11, pages 2827–2836, New York, NY, USA, 2011. ACM.

[9] N. Yankelovich, M. Wessler, J. Kaplan, J. Provino, N. Simpson, K. Haberl, and J. Matejka. Office central. In *Proceedings of the 2005 conference on Designing for User eXperience*, DUX '05, New York, NY, USA, 2005. AIGA: American Institute of Graphic Arts.

Collaboration over distance via Telepresence Robots

Lukas Höfer

Abstract— Telepresence goes beyond mere tele-operation. Essentially the operator should be able to feel present in the remote environment. Therefor not only an easy to use interface should be provided but also a form of avatar which enables the user to interact with his surroundings. In this paper it is shown, how current telepresence systems do not match that criteria and how scientists are trying to overcome the essential problems by creating robotic avatars. Those come in very different sizes and shapes to fit the demands of their environments, but main goal is to rebuilt the human body for telepresence operations. It is shown how researches are able to mimic the humans abilities and maybe leading the way in a new age of robotics, as predicted by Marvin Minsky in 1980.

Index Terms—Over Distance, Telepresence, Robot, Robotic Aid, Avatar

1 Introduction

> You don a comfortable jacket lined with sensors and muscle-like motors. Each motion of your arm, hand, and fingers is reproduced at another place by mobile, mechanical hands. Light, dexterous, and strong, these hands have their own sensors through which you see and feel what is happening. Using this instrument, you can 'work' in another room, in another city, in another country, or on another planet. Your remote presence possesses the strength of a giant or the delicacy of a surgeon. Heat or pain is translated into informative but tolerable sensation. Your dangerous job becomes safe and pleasant. - Telepresence by Marvin Minsky, OMNI magazine, June 1980 [37]

Marvin Minsky is a Professor at the Massachusetts Institute of Technology [36]. In 1980 Minsky wrote an article about the past and the future of teleoperation systems. The title of this article, suggested by his friend Patrick Gunkel, gave a whole field of studies its name: Telepresence [37].

Back in the 1980th - an era of peace in the USA - there haven't been much advancements in teleoperating systems. The military had some early drones, some teleoperated 'hands' where installed in nuclear power plants and other industrial facilities. Minsky wanted to break this stagnation and promoted a change from mere teleoperation systems to easily usable telepresence systems.

In todays science the terms teleoperation and telepresence can be used interchangeable. However, some argue that a line has to be drawn due to the meaning of the word presence. Therefore telepresence can only be achieved if the operator feels present in the distant or virtual system [45], which coincides with Minskys vision.

This paper will provide a look at current telepresence systems as used by the military, but although in the private sector. Furthermore selected projects will be introduced, which seem to hold a great value for the future development of telepresence robotics. The subsequent section presents some projects using robotic avatars, which lie in the near future. Finally a conclusion is drawn whether or not robotic telepresence will provide any practical use.

- *Lukas Höfer is studying Media Informatics at the University of Munich, Germany, E-mail: hoefer.lukas@campus.lmu.de*
- *This research paper was written for the Media Informatics Advanced Seminar 'Over Distance', 2011/2012*

2 State Of The Art

2.1 Common Telepresence Applications

Telepresence can be defined as the Operators feeling, to be in an distant environment. Therefor even a voice only telephone can be considered a telepresence system [46]. However the sense of being presence rises with the depth of the virtual world. [46]. It seems as if a telepresence system to fit the definition by an industrial standard should be at least supported by a live video transmission - the bigger, the better. Those 'TelePresence' systems are currently promoted by Cisco [14] and Telekom [49]. As advertised those systems already provide a lifelike representation of the communication partner and thus safe time and traveling costs. Marvin Minsky - the first person to use the term telepresence in writing - postulated for such systems to be not only easy to use, but also portable. Using a laptop or smartphone in combination with Skype and other programs as Google VoIP are currently closer to fulfill this postulation. In some cases those systems have proven to be very successful. A dancer from England, reports in an special issue of the journal "Performance Research" his very positive experience with the Software. Unable to travel to every rehearsal in New York, he used Skype instead and was still able to perform on stage with his colleges from the other side of the Atlantic Ocean [41]. Also in 2005 Godwin-Jones describes the positive effect of Skype in association with learning a language [20]. But still, others argue, whether or not the quality of service is of a level high enough to comply with the first demands of minsky [12].

Another flaw of those systems in general is the fact that no real physical interaction is possible, other than the collaborative work on digital documents.

Fig. 1. One of Ciscos TelePresence systems

2.2 Common Applications for Telepresence Robots

Today quad copters with cameras can be easily purchased, even in toy stores. That might be a good indication of how common telepresence robots already are. But besides the fun factor, there are more important ranges of use for such devices. The initiativly excerpt of Minskys article shows that even back in the 1980th the main fields of applications for telepresence robots where known.

Minimizing danger for humans Long before the term telepresence was established, the military was using remote controlled vehicles and aircrafts [19]. Those drones are still in use. Where it is to dangerous for soldiers to go or the human abilities are not enough for a mission, telepresence robots are the solution. But most robots are only able to perform one specific task. The TiaLinx Cougar 10 is designed to listen through walls and determine the presence of humans and their level of fear by analyzing their breath patterns, but it is very slow. The Lockheed Martin-made Squad Mission Support System (SMSS) on the other hand is a transportation unit, designed to carry packs and loads, which is much faster than the TiaLinx Cougar 10. So ingenious soldiers combined both for a mission to get an superior robot which combines the listening skills of the TiaLinx Cougar 10 with the speed of the SMSS. But still, even this new robot is only fit to serve in only a few scenarios [35].

Another example for the usage of telepresence robots again addresses danger, and again was already predicted by Minsky in 1980. In March 2011 due to a series of earthquakes the cooling system of the Japanese nuclear power plant Fukushima Daiichi was crippled and radiation made it to dangerous for human scientists to enter the station. A bomb-disposal robot was the first to be sent in. Not only did the robot confirm the radiation levels, but also cleaned up some of the radioactive water and debris, making it possible for technicians to enter some sectors for reparations [23].

Exploring the humanly unreachable But not only when danger is the factor, telepresence robots can be necessary. The eponymous syllable 'tele' derives from the Greek language and means distance, distant, far off. Those terms also imply unreachable. One example of those for humans unreachable regions is the sea ground. The high pressure below 2 miles depth makes it even more hostile, than the surface of mars [37]. In both cases, deep in the ocean and up on the mars, telepresence robots are able to collect valuable information, where (so far) no human is able to. Another important and unreachable source of information was the grave chamber in the pyramid of Gizeh. But again, a specially designed crawling robot, was able to enter 'air shafts' and supplied the scientists with pictures of the inside [7].

Superhuman precision Yet another factor, considered by Minsky, is the otherwise unreachable precision provided by telepresence robots. The most famous example is the surgical robot 'DaVinci' *(see figure 2)*. Since 1999 when the robot was first revealed to the public, only a few changes have been made. The robot consists of two parts. The controlling unit - a so called master - has two monoculars for the eyes, which provide a three dimensional view and it has three handles to control the robots arms. The current version of the actual robot - the slave - is equipped with three arms and the necessary camera to provide the 3D view [24].

2.3 Rising Applications for Telepresence Robots

Section 2.1 shows current applications for communication on an widely accepted level of telepresence. Whereas section 2.2 provides an insight in commonly used tools for remote operation. But there is a third group of telepresence applications. The combination of both of the previous sections. By taking the ability to communicate via audio and motion picture - or better -, and the mobility of the teleoperated robots, a more dynamic system for telepresence could be build.

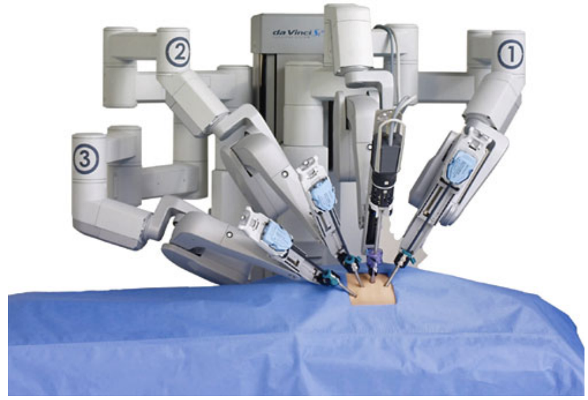

Fig. 2. The DaVinci surgical robot

Fig. 3. The state of the art: Anybot

One concept which is spreading through modern companies is the usage of a very basic concept. A frame, approximately the size of a human, is build on a wheeled platform. Attached to that frame are a screen with a camera, a microphone and speakers. One of those commercially available robots can be seen in *5*. This slave unit is controlled by the remote operator, which also has a camera, speakers of any kind, a screen and an input device for movement - usually a joystick or the mouse. Using a setup like this, the operator is now able to move nearly without boundaries - as far as ramps exist - and communicate not only in an given conference room, but throwout the whole complex as long as there are people around to push the buttons on the elevator. One company using such devices is Mozilla Corp. [34].
In single cases those telepresence robots are already used to help individuals, like in the case of 12 year old Stepan Supin, who is to sick to attend school with his real body [50].
But the study of Johannsen et al. reveals, that such communication is far from perfect [27]. Different approaches exist to improve the current modules. A selection of promising technology will be discussed in the next section.

3 Rebuilding The Human Body

When developing a new product it seems to be common practice for scientists to observe and copy nature. Even the genius Leonardo da Vinci spend a lot of time observing and analyzing birds in the hope of copying their ability of flight. But while talking about telepresence robots, which of natures creations would be better to use as a blueprint than the human body itself.

3.1 Grip And Sensitivity

One of the main differences between primates and other species, which is by some believed to have played a main part in the evolution of the homo sapiens is the thumb [8]. The thumb is what makes it possible to hold, grasp and manipulate objects the way only we humans and our near relatives do.

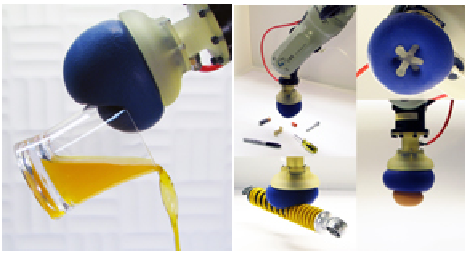

Fig. 4. A grasper using vacuum and granulate

3.1.1 The Robotic Human Hand

A big milestone in the history of robotic development was the first robotic hand by Raymond Groetz in 1951 [11]. Unlike modern robotic hands this hand did not use any motors, but was in direct physical connection to the operators hand *(see figure 5)*. Using todays digitalized data processing, it is possible to use small motors to simulate muscles and their contractions.
Another important aspect of such a complex device is its robustness. One project in particular is determined to build a nearly unbreakable hand. A team of the German Aerospace Center built a robotic hand which is not only strong, but flexible. It is able to withstand the full hit of a baseball bat [21].

Fig. 5. Raymond Groetz and his mechanical hand

3.1.2 Robotic graspers

Long before the robotic hand some graspers where developed. Some approaches where using simple clamps, others a three finger system to hold objects [37]. Even though nowadays it is possible to copy human hands way more sophisticated than in the early days of teleoperating systems, using a complex replica is not always the best solution. The human hand is by design very complex, which means a replica is more expensive to produce and has more need of computation power than a simple clamp.
Besides those clamps, which are still in use in robots like the SmartPal VII [17] and the RP2 named 'graspy' [52] there are a few other approaches worth mentioning.
One system tries to copy the grasping abilities of squids by creating

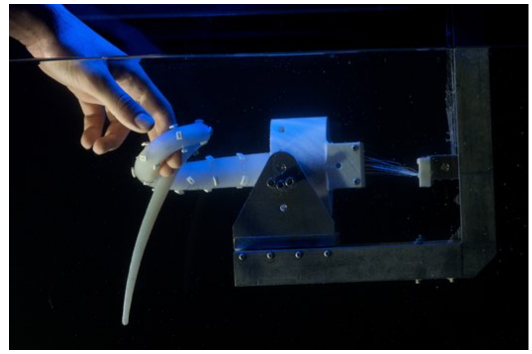

Fig. 6. Robotic Tentacle

an tentacle *(see figure 6)*. This very simple system enables the operator to grasp primitives and hold them tight [10].
An optical similar idea is it to rebuild the trunk of an elephant *(see figure 7)*. The Festo Corp. has developed a special design, which is also used to replicate the movements of penguin and seagulls. The special geometrical design of the individual elements leads to a very unique grasping technique, which is not only fast and sensitive, but also very simple [18].
Yet another approach is it to use the strength of vacuum. Without any need for fingers the Cornell Creative Machine Lab has created an universal grasper. In their design a sealed bag, filled with granulate is put on any graspable object *(see figure 4)*. By applying a vacuum to the bag the granulate hardens and therefor the object in question is fixed to the hand. Thus complicated feedback sensors are unnecessary [9].

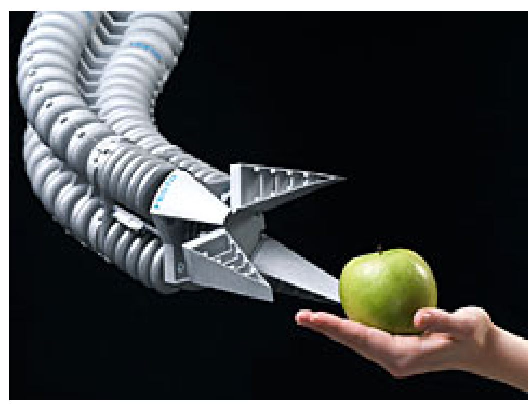

Fig. 7. Festos Tripod is the copy of an elephants trunk

3.1.3 Adding Sensitivity

"When any job becomes too large, small, heavy, or light for human hands, it becomes difficult to distinguish the inertia' and elasticity of the instrument from what it's working on." - Minsky
As mentioned above, the robotic hand developed by Raymond Groetz had an absolute physical connection to the operator. Therefore a direct link of physical forces from the hand to the operator was established [37]. Furthermore Minsky reports of Danny Hillis one of his Students, who developed a thin material which enables the operator to 'feel' tactile surfaces over distance.
Today there are several kinds of sensors which allow the user to feel the remote stimuli. One of them, primarily designed for couples, is the

a combination of a transmitting Koala doll [39] and a receiving west, which simulates pressure. Another system, the Telesar 5, transmits not only touch, but also heat. Furthermore the touching sensation is so sophisticated, that it is possible to transmit the surface of a lego brick [47].
But how does the robots touch reach the operator? There are several projects with different approaches involved to answer that question. The Bao research group for example has developed a skin-like film based on carbon nanotubes [33] *(see figure 8)*. The sensors are able to detect touch - up to the landing of a fly - , or chemical compounds. Those films can be combined with stretchable solar cells providing energy for the sensors. The material is elastic up to 150%, therefor it will not break if for example the robots fingers move.

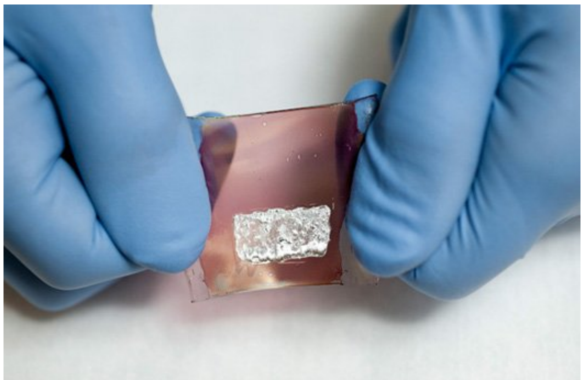

Fig. 8. A stretchable skin-like film developed at the Stanford University

3.2 Body Language

According to Jan Bouwen, a residential research director at Bell Labs, there is more to immersive discussions than audio and video. It is not only important to hear how people react, but to see them react. Those reactions can be very subtle like turning towards the current speaker to signalize approval, or turning away while silently disagreeing [48]. The following projects are partitioned in very basic elements of human interaction and therefor communication.

3.2.1 Mobility

A study of Nakanishi et al. has shown, that the ability of the operator to move the robotic avatar towards or from an object, increases the feeling of presence in the distant location [40]. Such findings are also confirmed by the MIT study of MeBot [38] and the interviews conducted in the Mozilla Corp. [34].
But how does it work? The most simple and most common solution is the usage of wheels. Memorable projects are the GestureMan of Kuzuoka et al. [31], MITs MeBot and common telepresence robots as used by Mozilla Corp. Even remotely controlled vehicles on the moon use this basic technique. Other ways of making an avatar move are wings [15], blimps [51], fins [18], legs [6] and pretty much any form of movement inspired by insects [32] [1] [6]. For the purpose of building a human avatar, at least on day legs will probably be irreplaceable. Projects like Asimo [25], DARwIn-OP [42] and Mahru [22] will lead the way. "We might then adapt designs and concepts from the arm to make legs, yielding a system able to work wherever people can, not only on carefully prepared floors." - Marvin Minsky [37]

3.2.2 Pointing

Another important aspect of non verbal communication is the basic task of pointing. In current telepresence robots as used by the Mozilla Corp. this important concept of communication is not implemented. But of course one could argue, that the visual representation on the screen allows the operator to point. The ability to move the robot towards the element, which should be emphasized, could be interpreted as pointing, too. But a precise pointing in the three dimensional space, which can easily be interpreted from nearly every point of view can not be established. Especially for tasks where the remote operator needs to guide a persons actions, a pointing device is very helpful [29].

Teroos is a creation of the Anzai-Mai Lab *(see figure 9)*. This tiny robot is mounted on the shoulder of the collaboration partner [4]. Due to the pivoted camera the operator is able to look at any given object in the near space around the carrier. The collaborator is able to easily follow the line of sight of the robot and therefore a sort of pointing is established.

Fig. 9. The shoulder mounted robot Teroos

The Gestureman is a telepresence robot built by the Japanese scientist Kuzuoka. In his paper the author emphasis the necessity of a pointing device [31]. Instead of using a robotic arm, the Gestureman is equipped with a laser pointer with three degrees of freedom of movement. The operator is now able to precisely mark elements in the collaborators physical space.

Using an arm as a pointing device might be the most obvious solution. Using modern technologies as the Kinnect, an arm has proven to be easy in usage [2]. But robotic arms are still expensive and thus are far off from becoming a standard.

3.2.3 Posture

Another aspect of human communication is posture. Some projects rely on virtual avatars, to transmit and amplify the emotions of the user. Others like the MeBot [38] (*(see figure 10)*) use robotic substitutes to simulate posture. But the MeBot is only interpreting facial expressions and translating them into body movements. One can argue whether or not the display of posture can be accomplished this way.
The Geminoid, as a life like human replica, is able to move its head, shoulders and arms. This makes it possible to impersonate posture on a higher level as the models previous mentioned [43]. It is sitting upright, which is currently limiting the range of postures. The Telenoid is another creation of the Hiroshi Ishiguro Laboratory. It is able of minimal 'arm' movement, to transmit hug-like postures [26].
The Korean robot Mahru is able to copy the movement of the human body. One can ether use motion tracking using a camera, or wear a motion capture suite, which enables to transmit the movements in real time [22].

3.3 Face to Face

3.3.1 Screens in Telepresence Robots

As discussed above, screens are the state of the art in displaying data for telepresence systems. Since the main technologies - a screen and a camera - have not only proven themselves over the past decades, but also have been improved, the necessary technologies for simple telepresence robots is cheap, stable and widely available. Tools using

Fig. 10. Mebot

screens are the the 'TelePresence' systems of Telekom [49] and Cisco [14] *(see figure 1)*. Commercially available telepresence robots using screens are for example Anybot [3] and VGo [54].
A slightly different approach using a screen was designed by Tobita et al. which created an floating avatar. A blimp is used to lift their telepresence robot from the ground. The balloon involved is also used as a screen, on which the face of the user is projected thus creating a sort of three dimensional image [51]. A more live like version of this method is used in Mask-bot [30], which will be discussed in section 3.3.3.

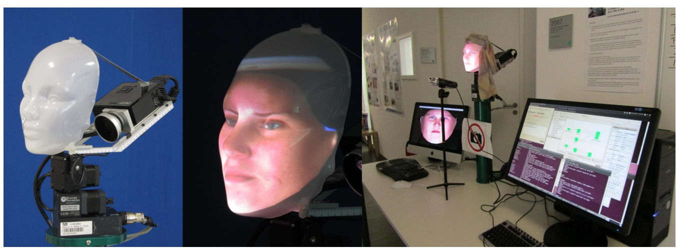

Fig. 11. The face of Mask-Bot

3.3.2 Robotic substitutes

For a human it is not always necessary to interact with a real human face. One big flaw of the human mind is to personify animals and objects. A phenomena which is known as pareidolia [16]. But this flaw can be used to transmit inter human communication through relatively simply objects. One of those simplified transmitters is, as mentioned above, MITs MeBot [38]. Not only can the face in the screen be substituted by a smilyface avatar, but also does the structure of the robot reassemble the human gestures, just by analyzing its facial structure. MeBot will translate a smile into a gesture, which mimics open arms - a friendly gesture.
Another impersonation is Teroos [4]. Besides the ability to look in different directions, which can be considered as pointing, a simple face like structure is build around the camera. Through an computer interface this structure can be easily altered by the operator, enabling the carrier to directly interpret the emotions of its communication partner.

3.3.3 The human face

Most of the projects mentioned use a simplified version of the human face. But two projects are trying to get close to the real face. One design developed in a cooperation of the technical university in Munich and Japanese scientists is named the Mask-Bot [30] . It uses a semi-transparent mask in the shape of a human face, in which a three dimensional image is projected. At the moment the Mask-Bot is getting its input by a computer program, but in future it could also be used to transmit a three dimensional image of a remote operator *(see figure 2)*.
The other project is the Geminoid series. Whereas other projects of the Hiroshi Ishiguro Laboratory try to minimize the reassembles of their telepresence robots with the real human face, by standardizing it, the Geminoids are copies of real humans. The robotic skeleton is covered by a latex skin. Real hair and realistic teeth and eye replicas are used to complete the illusion *(see figure 12)* [44].

Fig. 12. The Danish Geminoid

4 The (near) Future Of Telepresence Robots

So far only a few of Minskys predictions have been fulfilled. But there are some promising projects in progress, which might take us one step closer towards his vision.

4.1 Exploring Space

"One small step for man, one giant leap for tinman kind." - Robonaut2 [5]
Those famous words were among the first tweets of @astroRobonaut as he was unpacked and switched on orbiting the earth in the international spaces station (ISS). The Robonaut2 is the first humanoid robot in space. The current version as stationed on the ISS does not have any legs, but is a waste up humanoid. He is designed to help the human astronauts performing otherwise riskfull tasks such as repairs in the hostile environment on the outside of the space station. The information collected on this long term mission will help the scientists of NASA to improve the robonaut in aspects of working in zero gravity and collaboration with the human astronauts.
A team of NASA scientists has announced the possibility of sending an improved robonaut to be stationed on the moon. The project is calculated to take only about 1000 days to prepare and could therefore be accomplished until the year 2015. The advantages of an humanoid over his more primitive ancestors like the mars rover lie in its flexibility towards future tasks. At the moment the tools necessary to perform tests on other planets are mounted on the remote controlled vehicles. But if those tools would become hand held, one robot would be enough to perform a variety of tasks, or even reuse the tools in a way they were

Fig. 13. The Robonaut2 currently relies on wheels

not designed to perform. Thus adapting to new tasks, just like a human would *(see figure 2)*.
Marvin Minsky the Professor at the Massachusetts Institute of Technology who was the first person to use the term telepresence in writing wrote in 1980: ”why not use cheap, Earth-based labor via telepresence to build moon factories?” [37]. So it seems as if this NASA mission is long over due.
Another of Minskys predictions contains service robots in earth orbit to repair satellites. At the moment without the space shuttle around some broken satellites cannot be repaired in space and are therefore expensive junk. Students of the Johns Hopkins University are currently working on a project, which will fulfill Minskys vision using the surgical robot DaVinci *(see figure 2)*. In cooperation with NASA a new kind of service satellite will be build to refuel and repair satellites [55].

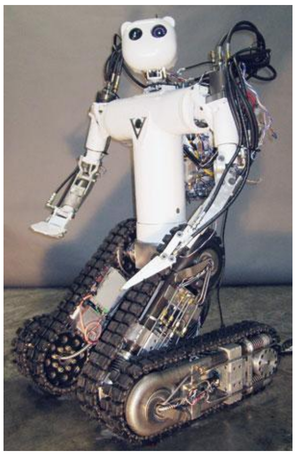

Fig. 14. Vecna Robotics' Bear

4.2 Saving Lives

Hostile conditions can not only be found in space. As mentioned in section there are a variety of dangerous situations and places on earth, where telepresence robots seem to be the best choice. One example mentioned above is the Japanese nuclear plant Fukushima Daiichi which in march 2011 began to emitted dangerous radiation. Teleoperated robots were the first to enter critical areas, measuring radiation levels and cleaning up debris that engineers later where able to access the areas with less exposure to radiation. Other robots where send to help with the cleanup of the polluted environment to assist and replace human work in the hostile environment where possible.
Those robots, who also helped finding and rescuing people in other parts of japan after the natural disaster, are around for quite a while. Some of those models where already in use by the rescue operation on 9/11 in New York. But with ongoing technological advancements and some lessons learned from previous operations, new models are created to assist in future missions.
One way to assist humans in rescue operations is to build stronger robots, to remove debris and carry injured humans to safety. The Bear (Battlefield Extraction-Assist Robot) is a product of Vecna Robotics *(see figure 14)*. This robotic medic is designed to lift and transport weights over 200kg. The newest version is equipped with a fire and explosion prove batteries. The individual treads function leg like, enabling the robot to overcome obstacles, stand upright or kneel [53].
Another approach is to make the robots smaller, so that they can be used as scouts in humanly unreachable dimensions. There are different projects based on very different ideas. The most outstanding idea, when it comes to telepresence, is a concept of the University of Michigan [28]. In their concept real bugs are equipped with a camera and remote control units attach to their nerval system *(see figure 15)*. Big batteries are to heavy to lift and the duration of small ones is not long enough for a real life mission, thus the energy is generated by the wing movement of the bug. Those remote controlled insects are very small and therefor able to crawl and fly through collapsed buildings, long before any human could pass.

Fig. 15. Cyborg Bug

4.3 Becoming Human

The team of the Hiroshi Ishiguro Laboratory has different plans for telepresence robots [26]. The plan is not to build their robots superhuman, but just as human as possible. One of the main goals is to understand human presence, and to transfere it, if possible. To achieve this goal the Geminoids are used, which are as mentioned above very realistic replica of humans. A Geminoid consist of a robotic basis, where the face, the head movement and the arms are controllable. All Geminoids reassemble a real human person by using a latex mask with human hair. The danish version mostly works in an office like laboratory, the Asian versions give interviews and the female even performs in theater plays.
Minsky wrote 1980: ”The biggest challenge to developing telepresence is achieving that sense of 'being there.' Can telepresence be a true substitute for the real thing?” [37]
By recreating the feeling of presence through robots human closeness could be transfered. For example could a psychiatrist using such technology be able to work over distance or families could be brought together, without the need of travel.
Another very human aspect of telepresence is the possibility to recreate lost abilities for disabled people. Through specialized software such as lipreading programs deaf-dumb people could better communicate over distance than by using the telephone. Their avatar on the other side of the connection could for example fluently speak their words, whereas the avatar on their side could use sign language or screens to represent the spoken word.
Marvin Minskys term telepresence was inspired by the science fiction novel Waldo by Robert A. Heinlein from the year 1948. The main character Waldo is not able to move in normal gravity. Therefor he inhabits a space station in earth orbit, controlling robotic hands on the surface through a very intuitive interface [37]. This vision, which might be considered the base of telepresence, is not that far off from modern science. As shown in 2010 by Christoforou et al. it is possibly to navigate a distant telepresence robot in real time using the brains E.E.G.-signals. In their presentation it was shown, that even untrained users are able to perform such a task [13]. Even today 12 year old Stepan Supin, is using a telepresence robot to attend to school. The robots helps him to leave his home, which would otherwise be to dangerous, due to his leukemia [50].

5 Conclusion

The current age of information will one day be replaced by an age of robotics. At least this is what Hiroshi Ishiguro is predicting [26]. And we won't even see it coming, because robotics will slowly be introduced to society until we accept them as we do planes and cars. Marvin Minsky seems to believe likewise [37]. He foresaw many of todays robotic inventions and predicted a level of telepresence which even today seem unimaginable.

”Right now, the ability to work from home is a little dicey - the connection is not something you can depend on and the technology is not particularly great - but I think that all those things are just around the corner.” - Bell Labs’ residential applications research director Jan Bouwen [48]
Indeed they are. This paper presented not only the state of the art, but also the great inventions of modern scientists, waiting to be ready for commercial use. Superhuman strength and robustness, artificial self powering skin, which can even distinguish between chemical compounds, quite realistic human masks and the abilities to walk, to fly, to swim, to listen and to talk, are to prove the real use of robotic telepresence.

REFERENCES

[1] A. Hoover, E. Steltz, R.S. Fearing. Roach: An autonomous 2.4g crawling hexapod robot. IEEE Int. Conf. on Intelligent Robots and Systems, Nice France Sept. 2008, 2008.

[2] E. Ackerman. Kinect-Powered robot lets you clean up your house remotely. Website, 2011. Available online at `http://tinyurl.com/kinect-powered`; visited on December 8th 2011.

[3] Anybot Inc. Anybots. Website, 2011. Available online at `www.anybots.com`; visited on December 8th 2011.

[4] Anzai-Mai Lab. Teroos: The telepresence robot you wear on your shoulder. Website, 2011. Available online at `http://www.diginfo.tv/2011/03/22/11-0071-r-en.php`; visited on December 8th 2011.

[5] AstroRobonaut. @astrorobonaut: A humanoid robot designed to work side by side with humans, or go where the risks are too great for people. Website, 2011. Available online at `http://twitter.com/AstroRobonaut`; visited on December 8th 2011.

[6] R. Becker. High-tech spider for hazardous missions. Website, 2011. Available online at `http://www.fraunhofer.de/en/press/research-news/2011/november/high-tech-spider.html`; visited on December 8th 2011.

[7] S. Belless. Website, 1999. Available online at `http://www.cheops.org/`; visited on December 8th 2011.

[8] G. Boeree. Website, 2002. Available online at `http://webspace.ship.edu/cgboer/humanevol.html`; visited on December 8th 2011.

[9] Brown, E. Rodenberg, N., Amend, J., Mozeika, A., Steltz, E., Zakin, M., Lipson, H., Jaeger, H. Universal robotic gripper based on the jamming of granular material, 2010.

[10] M. Calisti, M. Giorelli, G. Levy, B. Mazzolai, B. Hochner, C. Laschi, and P. Dario. An octopus-bioinspired solution to movement and manipulation for soft robots. *Bioinspiration & Biomimetics*, 2011.

[11] J. Chen. What is a robot? Website. Available online at `http://users.cecs.anu.edu.au/~chen/teaching/Robotics_ENGN4627_2005/lectureNotes/engn4627-Part01.pdf`; visited on December 8th 2011.

[12] K. Chen, C. Huang, P. Huang, and C. Lei. Quantifying skype user satisfaction. page 399. ACM Press, 2006.

[13] C. Christoforou, N. Mavridis, E. Machado, and G. Spanoudis. Android tele-operation through Brain-Computer interfacing: A real-world demo with non-expert users. In *Proceedings of the International Symposium on Robotics and Intelligent Sensors (IRIS)*, 2010.

[14] Cisco Systems, Inc. Cisco TelePresence system user guide - getting to know the cisco TelePresence systemÂ [Cisco TelePresence administration software] - cisco systems. Website. Available online at `http://www.cisco.com/en/US/docs/telepresence/cts\_admin/1\_7/userguide/cts\_ug\_know.html`; visited on December 8th 2011.

[15] DARPA. Website, 2011. Available online at `http://www.darpa.mil/NewsEvents/Releases/2011/11/24.aspx`; visited on December 8th 2011.

[16] Dictionary.com, LLC. World english dictionary: Pareidolia. Website. Available online at `http://dictionary.reference.com/browse/pareidolia`; visited on December 8th 2011.

[17] Digitized Information, Inc. Bionic learning network. Website, 2011. Available online at `http://www.diginfo.tv/2011/11/21/11-0248-r-en.php`; visited on December 8th 2011.

[18] Festo AG & Co. KG. http://www.festo.com, 2009.

[19] R. Gockley, J. Forlizzi, and R. Simmons. Natural person-following behavior for social robots. page 17. ACM Press, 2007.

[20] R. Godwin-Jones. Emerging technologies skype and podcasting: Disruptive technologies for language learning. Website, 2005. Available online at `http://llt.msu.edu/vol9num3/emerging/default.html`; visited on December 8th 2011.

[21] Grebenstein. The dlr hand arm system, 2011.

[22] E. Guizzo. Humanoid robot mahru mimics a person’s movements in real time. Website, 2010. Available online at `http://tinyurl.com/mahru-realtime`; visited on December 8th 2011.

[23] D. Hambling. Japan sends robots into fukushima nuclear plant. Website, 2011. Available online at `http://www.newscientist.com/blogs/onepercent/2011/03/japanese-send-robots-into-fuku.html`; visited on December 8th 2011.

[24] N. G. Hockstein, C. G. Gourin, R. A. Faust, and D. J. Terris. A history of robots: from science fiction to surgical robotics. *Journal of Robotic Surgery*, 1:113 to 118, Mar. 2007.

[25] Honda Motor Co., Ltd. Asimo: The honda humanoid robot asimo. Website, 2011. Available online at `http://world.honda.com/ASIMO/`; visited on December 8th 2011.

[26] H. Ishiguro. Hiroshi ishiguro laboratory: The mission. Website, 2011. Available online at `http://www.geminoid.jp`; visited on December 8th 2011.

[27] A. Johannsen, W. van Diggelen, G.-J. de Vreede, and H. Kremar. Effects of video communication and telepresence on cooperative telelearning arrangements. Proceedings of the 33rd Hawaii International Conference on System Sciences, 2000.

[28] K. Najafi, E. Aktakka. Michigan engineering | insect cyborgs may become first responders. Website. Available online at `http://www.engin.umich.edu/newscenter/feature/cyborginsects/`; visited on December 8th 2011.

[29] T. Kurata, N. Sakata, M. Kourogi, H. Kuzuoka, and M. Billinghurst. Remote collaboration using a shoulder-worn active camera/laser. In *Wearable Computers, 2004. ISWC 2004. Eighth International Symposium on*, volume 1, page 62 to 69, 2004.

[30] G. Kuratate, Takaaki; Cheng. Mask-bot: Ein roboter mit menschlichem antlitz. Website, 2011. Available online at `http://portal.mytum.de/pressestelle/pressemitteilungen/NewsArticle_20111107_151556?searchterm=Mask-Bot`; visited on December 8th 2011.

[31] H. Kuzuoka, S. Oyama, K. Yamazaki, K. Suzuki, and M. Mitsuishi. GestureMan a mobile robot that embodies a remote instructor’s actions. In *Proceedings of the 2000 ACM conference on Computer supported cooperative work*, page 155 to 162, 2000.

[32] H. Lin, G. G. Leisk, and B. Trimmer. GoQBot: a caterpillar-inspired soft-bodied rolling robot. *Bioinspiration & Biomimetics*, 6(2):026007, June 2011.

[33] D. J. Lipomi, M. Vosgueritchian, B. C.-K. Tee, S. L. Hellstrom, J. A. Lee, C. H. Fox, and Z. Bao. Skin-like pressure and strain sensors based on transparent elastic films of carbon nanotubes. 2011. Available online at `http://www.nature.com/nnano/journal/vaop/ncurrent/abs/nnano.2011.184.html`; visited on December 8th 2011.

[34] C. Major. Robots at work. Website. Available online at `http://video.nytimes.com/video/2010/09/04/science/1248068965210/robots-at-work.html`; visited on December 8th 2011.

[35] P. McLeary. Robots piggybacking on robots. Website. Available online at `http://www.aviationweek.com/`; visited on December 8th 2011.

[36] M. Minsky. Brief academic biography of marvin minsky. Website. Available online at `http://web.media.mit.edu/{\textasciitilde}minsky/minskybiog.html`; visited on December 8th 2011.

[37] M. Minsky. Telepresence. Website. Available online at `http://web.media.mit.edu/~minsky/papers/Telepresence.html`; visited on December 8th 2011.

[38] MIT Lab: Personal Robots Group. Mebot. Website, 2010. Available online at `http://robotic.media.mit.edu/projects/robots/mebot/overview/overview.html`; visited on December 8th 2011.

[39] F. Mueller, F. Vetere, M. Gibbs, J. Kjeldskov, S. Pedell, and S. Howard. Hug over a distance. In *CHI’05 extended abstracts on Human factors in computing systems*, page 1673 to 1676, 2005.

[40] H. Nakanishi, Y. Murakami, D. Nogami, and H. Ishiguro. Minimum movement matters: impact of robot-mounted cameras on social telepresence. In *Proceedings of the 2008 ACM conference on Computer supported cooperative work*, page 303 to 312, 2008.

[41] P. Petralia. Here, there and In-Between: rehearsing over skype. *Performance Research*, 16:113 to 116, Sept. 2011.

[42] robotis.com. DARwIn-OP. Website, 2011. Available online at `http://www.robotis.com/xe/darwin_en`; visited on December 8th 2011.

[43] D. Sakamoto and H. Ishiguro. Geminoid: Remote-controlled android system for studying human presence. *Kansei Engineering International*, 8(1):3–9, 2009.

[44] H. Scharfe. Geminoid dk. Website, 2011. Available online at `http://geminoid.dk/`; visited on December 8th 2011.

[45] D. W. Schloerb. A quantitative measure of telepresence, 1995.

[46] J. Steuer. Defining virtual reality. Website, 1992. Available online at `http://www-bcf.usc.edu/~kwanminl/courses/comm631/readings/Steuer\%20(1992)\_Defining\_Virtual\_Reality\_JOC.pdf`; visited on December 8th 2011.

[47] Tachi. Telexistence robot avatar transmits sight, hearing and touch - telesar v. Website, 2011. Available online at `http://www.diginfo.tv/2011/11/07/11-0225-r-en.php`; visited on December 8th 2011.

[48] L. Tay. Bell labs builds telepresence 'robots'. Website, 2011. Available online at `http://www.itnews.com.au/News/278794,bell-labs-builds-telepresence-robots.aspx`; visited on December 8th 2011.

[49] Telekom Deutschland GmbH. Website, 2011. Available online at `http://telepresence.telekom.com/`; visited on December 8th 2011.

[50] H. Tim. Robots help sick kids go to school. Website, 2011. Available online at `http://news.cnet.com/8301-17938_105-20030156-1.html`; visited on December 8th 2011.

[51] H. Tobita, S. Maruyama, and T. Kuzi. Floating avatar. page 541. ACM Press, 2011.

[52] University of Pennsylvania. Website, 2011. Available online at `https://www.grasp.upenn.edu/news/pr2_robot_learns_scoop_poop`; visited on December 8th 2011.

[53] Vecna Technologies Inc. The bear : Vecna robotics. Website. Available online at `http://www.vecna.com/robotics/solutions/bear/index.shtml`; visited on December 8th 2011.

[54] I. VGo Communications. Vgo. Website, 2011. Available online at `www.cgocom.com`; visited on December 8th 2011.

[55] L. Whitcomb, T. Xia, and J. Bohren. Medical robotics experts help advance nasa's 'satellite surgery' project. Website, 2011. Available online at `http://releases.jhu.edu/2011/12/05/medical-robotics-experts-help-advance-nasa%E2%80%99s-%E2%80%98satellite-surgery%E2%80%99/`; visited on December 8th 2011.

Haptic Actuator Technology for Virtual Environment Interfaces

Lars v. Aichelburg

Abstract— This paper is about Haptic Actuator Technology in general. It will explain and define the term Haptics as it is used in related literature. It will discuss advantages and drawbacks of those technologies and discuss some of them exemplarily. Further there is an dedicated chapter about technologies used for those interfaces. The paper ends with a brief look to the future

Index Terms—Haptic, Tactile, Interface, Virtual Environment, sense of touch

1 Introduction

This chapter discusses the potential and drawbacks of haptic feedback in Virtual Environments.

1.1 Potential

Haptic feedback is very important to interact with the real world. Even simple tasks (e.g.: opening a door) are very hard to do, if we only have visual or auditory feedback.

Thats is why haptic technology became so important for virtual environments. These kind of interfaces allow the user to feel more present in the virtual world. Espacially fine manipulating and precise interaction is not possible without the sense of touch.

With the sense of touch in a virtual environment we get a better access to human-computer interaction. New application areas that have been too expensive, too dangerous or too impractical so far, can be accessed.

However haptic interface technology is still very limited. But since the 1990s a lot of effort has been put into the developement of that technology, because there is a lot of potential. It will be used (or is already in use) in the nuclear industry, for servomanipulators, in the ceramics industry, for land mine clearing, surgery, gaming industry and many more.

1.2 Drawbacks

Both tactile and force feedback interfaces for virtual environments represent active areas of research and development. There are major limitations of the tactile feedback technology. First of all, it is very hard to represent surface characteristics such as texture, local shape and slip. Most devices are good in imitating one tactile sensation, while not beeing able to imitate others and they are still limited to small areas. Another problem is the lack of models and algorithms for an efficient generation of tactile signals.

Most current force feedback devices can be defined as exoskeleton devices. The limitations of force feedback are similar to those given for tactile interfaces. It is restricted to only a few different VE interactions and to a number of joints. Further we have the intrusive nature of force feedback devices and their constraints of user movement. Similar to tactile feedback we have a lack of common models and algorithms [32].

2 Overview

This section starts with the definition of the terminology used in this paper. The next subsection gives a short timeline of the development of haptic interfaces. Finally a list of studies show the importance of the sense of touch in Virtual Environments.

- *Lars v. Aichelburg is studying Informatics at the University of Munich, Germany, E-mail: lars.aichelburg@campus.lmu.de*

- *This research paper was written for the Media Informatics Advanced Seminar 'Over Distance', 2011/2012*

2.1 Terminology

This chapter defines the technical terms as used in related literature

2.1.1 Virtual Environment Interface

A virtual environment (VE) or virtual reality/synthetic environment is a Computer-simulated reality. It simulates physical presence either in places of the real world or in imaginary places. Interfaces allows us to interact with the VE [14].

We can classify those Interfaces:

- Visual Interfaces
- Tracking Interfaces (position and orientation tracking)
- Auditory Interfaces
- Olfactory Interfaces
- Motion Interfaces
- Haptic Interfaces

[32]

2.1.2 Haptic

This paper adresses the topic of haptic interfaces. The human haptic sense can be divided into three categories [10]: the kinesthetic sense (force motion) and the tactile sense (tact, touch). The third one, the proprioceptive feedback, will not be discussed in this paper.

- Tactile Feedback: User gets response of the system by sensations (e.g.: roughness, rigidity, shape, smoothness and temperature)
- Force Feedback: User gets response of the system by pressure, movements or other stimulus, like vibration. The system also simulating an objects hardness, weight and inertia.
- Proprioceptive Feedback: Gives the user a feel of his posture or body position

2.1.3 Haptic Substitution

Most current VE substitute haptic feedback. For example the user can hear a sound, when he comes in contact with a virtual object [21].

2.2 History of Haptic Interfaces

The first interface was developed in 1954. It used force feedback and served a tele-operation system for nuclear environments. In 1992 the Rutgers University developed a light and portable force feedback glove called the "Rutgers Master", which will be discussed in chapter 4.2. Nowadays the PHANToM arm is the most popular force feedback system. It was introduced in 1994. For more details, see chapter 4.1. It was followed by the the impulse engine in 1995 and the CyberGrasp glove in 1998. The CyberGrasp is an exoskeleton, which adds force feedback to each finger.

On the other hand we have the tactile feedback technologies, which were developed at MIT first. The first technology used voice coils, which will be described later on. In 1990 the ”Sandpaper” tactile joystick was developed. Then in 1993 we had the ”TouchMaster” and the CyberTouch glove in 1995 (offers vibrotactile stimulators on each finger). In 1997 the ”FEELit Mouse” was developed. This desktop two-DOF interface enables the user to feel hard surfaces, rough textures, smooth contours and even rubbery materials [13].

2.3 Role of Touch in Virtual Environments

In VEs the human sense of touch has an important role. Unlike the visual and auditory systems, the haptic sense can provide input and output to a system. [13]

There has been many studies about the Role of Touch in Virtual Environments like the study of E.-L. Sallnäs et al. According to that study, force feedback significantly improved task performance. The participants could do five tasks within 24 minutes, if they had a haptic force feedback. Else they needed 35 minuts. Further the subjects had to answer questions about the perceived virtual presence and perceived task performance and it turned out, that it had been significantly higher in the three-dimensional visual/audio/haptic condition.

Hasser et al. [1998] released a study about a computer mouse with force feedback. They showed, that it improved targeting performance and decreased targeting errors [28].

Further there has been experiments that showed the value of tactile feedback:

- for simple tracking tasks [27]
- for time reduction in target pointing [7]
- in degraded visual conditions [29]

3 Actuator Technology and examples

The following classification is mainly based on [32]

3.1 Actuator Technologies for tactile feedback

The following chapters describes the Actuator Technologies for tactile feedback with expamples.

3.1.1 Piezoelectric crystals

Piezoelectricity means electricity resulting from pressure. This charge is accumulated in crystals most commonly quartz. When these crystals are compressed or pulled, the structure of the molecule changes *(see figure 1)*.

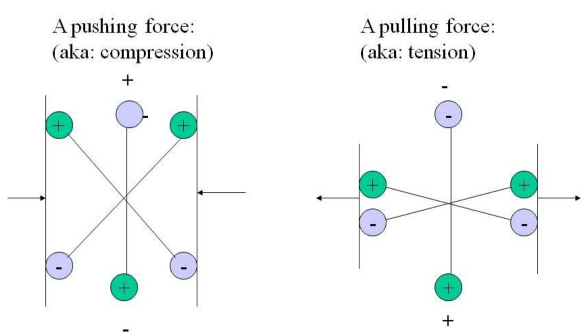

Fig. 1. the piezoelectric effect [1]

The piezoelectric effect is a reversible process. That means that you can provide electrical charge by a mechanical force as well as a mechanical strain by an applied electrical field.

The advantages are the high spatial resultion, but on the other hand the piezos are restricted to resonant frequency [32] page 119 .

This technology is mainly used for Touchscreens, to create a tactile feedback for virtual buttons. It is even possible to make buttons feel more heavy, by longer delays in tactile feedback [16].

3.1.2 Electrostatic Friction

The technique of electrostatic friction exploits the principle of electrovibration, which allows to create a broad range of tactile sensations *(see figure 2)*. The effect of electrovibration was discovered in 1954 by accident. Mallinekrodt et al. noticed a ”rubbery” feeling, when dragging a dry finger over a conductive surface covered with a thin insulating layer. An alternating voltage on the conductive surface, results in an intermittent attraction force between the finger and the surface.This effect was named ”electrovibration”. It has a lot of advantages over other technologies. It is fast, low-powered, dynamic and can be used in a wide range of haptic scenarios. If we talk about electrovibrition, it is important to understand the difference to electrocutaneous stimulation. Electrocutaneous displays stimulate the Finger of a user by an electric charge passing through the skin. With electrovibration tactile feedback however, we have the charge inducted by moving on a conductive surface. Furthermore the stimulation is mechanical, meaning that the electrostatic force is deforming the skin of the moving finger.

The surface can be influenced by the frequency. A low frequency stimuli results in a rougher feeling, like wood, while a higher frequency results in feeling of things like paper or a painted wall. The other factor is the amplitude. An increase of amplitude increases the perceived smoothness. [9]

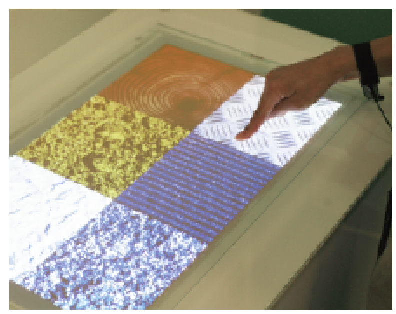

Fig. 2. different textures produce different sensations

3.1.3 Electrocutaneous

Electrocutaneous (also called electrotactile) stimulation creates a feeling of tacile sensations within the skin by passing a local electric current through the skin [17].

3.1.4 Solenoid

A solenoid is a coil wound into a tightly packed helix. Solenoids can create a controlled magnetic field, when an electric current is passed through it. If there is a magnet in the solenoid, there is a force pushing the magnet out in one direction. *(see figure 3)*. This force can be used to provide tactile feedback.

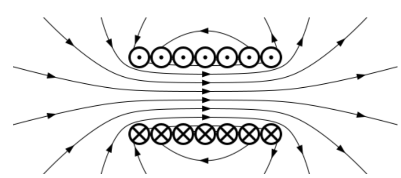

Fig. 3. Magnetic Field created by a solenoid [2]

An example for tactile feedback, using solenoids is the haptic pen. A pressure-sensitive stylus is combined with a small solenoid. This solenoid has several basic actions, like buzz, hold, lift and so on. By clicking on a virtual button you can get a buzz feedback or a feedback how hard it is to push the button [22]. The main disadvantages are, that this technology is realtively heavy and that it is nonlinear, meaning that it can require extra effort to control. It has a better bandwidth than other materials (expect for piezoelectric crystals and voice coils) [**?**].

3.1.5 Voice Coils

A voice coil is a coil in a magnetic field which get deflected by Lorentz force. If we feed voltage through a voice coil, we create a magnetic field. Because of an installed permanent magnetic field, the voice coile reacts and causes a mechanical vibration. Voice coils are mainly used for dynamic Microphones and hard disc drives *(see figure 4)*.

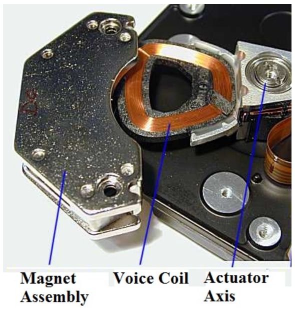

Fig. 4. hard disk using a voice coil as head actuator [20]

The soundTouch is an expample for a voice coil used to provide tactile feedback. The soundTouch can provide the feeling of tapping as well as rubbing. This is optimal for messages that have an underlying connotation, like reminders (tapping) and expressions fo care and comfort (rubbing) [24].

Vibrotactile alerts are loud, for example you can hear the noise of a phone vibrating. Rubbing, in contrast, is silent. Or we can intend to use it in game controllers. Most popular game controllers us imersion's vibrotatcile technology. But vibration is not always the best way to go. It is good for events like being shot or driving off the road, but for positive events, such es picking up a health pack, a rubbing tactile feedback might be better in this case.

Another example for a tactile interface using voice coil is the light-weighted finger haptic device developed by Sato Group *(see figure 5)*. The purpose of that group was to realize the feeling of touching soft virtual bodies by force-feedback to hands or arms and stimulation-feedback to fingers on the same time. By combining two systems they could realize soft material touching [6].

Advatages and disadvatages of the voice coil technology in general are [32]:

- Advantages:
  - High temporal resolution
  - Relatively small, does not obstruct normal movement ranges of the fingers
- Disadvantages:
  - Poor spatial resolution
  - limited scalability

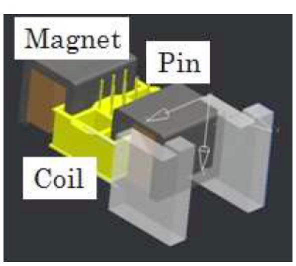

Fig. 5. Construction of a tactile display using voice coils

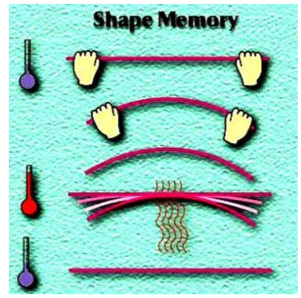

Fig. 6. Shape Memory Alloys [4]

3.1.6 Shape Memory Alloy(SMA)

A shape memory alloy is a material that ”remembers” its original shape. It is getting its original shape by heating. *(see figure 6)*

For Haptic Feedback wires and springs are used to contract by heating and expand again by cooling. These alloys have a good power-to-mass ratio. But they are inefficient during contraction and heat dissipation problems limit relaxation rate of wires [**?**].

3.1.7 Heat Pump

A heat pump is a machine, that transfers thermal energy from one location with lower temperature, to another location. This is normally realized by mechanical work, but also by some source of thermodynamic work. Heat pumps can be used for heating as well as cooling. When we look at the classical thermodynamics, where heat is defined as a movement of energy, the name ”heat pump” is not correct. Classical heat cannot be moved or ”pumped” from one place to another. Instead some of the energy is moved in the form of thermodynamic work or mechanical work (as already mentioned above). But in the end the effect is the same and the device gained its name by that analogy. One commercial product that provides temperature feedback for Virtual Environments is the ”Displaced Temperature Sensing System” *(see figure 7)*, designed for Virtual Environments. This system uses some tracking device and allows a temperature to be sensed fitting the user's location in the VE. The temperature has a range from 10 °C to 45 °C and a temperature Resolution of 0.1 °C. [32]

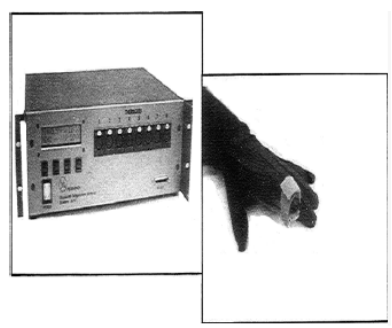

Fig. 7. Displaced Temperature Sensing System

3.1.8 Pneumatic

The following ist directly taken from [**?**]:

Pneumatic takes many forms: As air-jets, provides an array of air nozzles that can be gated to a display pattern. As air-rings(cluffs), like miniature blood pressure cuffs. As bladers (bellows), often the size of a finger pad and held against the finger by a glove or band. As an array of tiny presurize bladders, many to a single finger pad.

This technology has a low mass on hand, but on the otherhand it has poor spatial and temporal resolution as well as a limited bandwidth [**?**].

3.2 Actuator Technologies for force feedback

The following chapters describes the Actuator Technologies for force feedback with expamples.

3.2.1 Solenoids

Solenoids have been described above for tactile feedback. But this technology can also be used for force feedback. The advantage of this technology is, that it is possible to change forces quickly (no big masses that add inertia are involved) and the forces can be transmited directly (without unnecessary friction).

One example is the Haptic SpaceMouse *(see figure 8)* [11] .

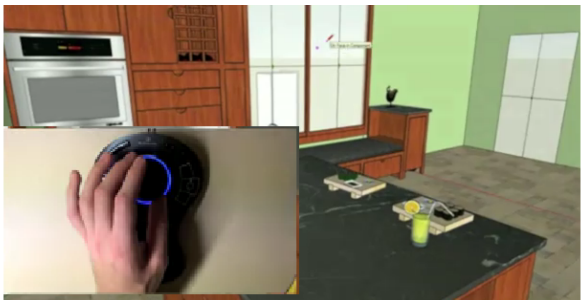

Fig. 8. SpaceMouse

3.2.2 Piezoelectric Motors

Piezoelectricy has been already discussed in the section of tactile feedback. Piezoelectric motors translate the vibration of piezoelectric materials to linear or rotatory motion using frictional forces to produce usable torques or forces at low speeds, without the need for gear reduction. They produce high forces at low speeds in small package. On the other hand these motors require precision machining and the necessary power gating can cause annoying noise. Piezoelectric motors are still the subject of research [32] page 141 .

3.2.3 Electromagnetic Motors

Electromagnetic Motors produce moment of force with two time-varying magnetic fields, cause by tow coils or a coil and a magnet. The major advantages are, that they are easy to control, clean and quiet. On the other hand they consists of heavy components and have heat dissipation problems. Further they have low power densities at small scales and low static force capability.

3.2.4 Hydraulics

Fluid power is the us of fluids under pressure to generate, control and transmit power. Hydraulics using a liquid in contrast to pneumatics (using gas) Hydraulics force capability, power output, stiffness and bandwidth are unmatched by other technologies. But this technology has a high mass. Its design is difficult and expensive. And last it has a tendency for fluid leaks.

3.2.5 Pneumatics

Pneumatics is a technology using pressurized gas to effect mechanical motion.

One advanced system, using Pneumatics for force feedback is the GI-Mentor from Simbionix *(see figure 9)*. The GI-Mentor is an endoscopic medical simulator [19].

Another example using this technology is the Rutgers Master introduced later in chapter 4.

3.2.6 Magnetorestrictive

The following ist directly taken from [**?**]:

Magnetorestrictive materials change shape when subjected to magnetic fields. Magnetorestrictive motors also mechanically rectify small oscillatory motions of the driving element(s). This technic provides high forces at low speeds in small package. On the other hand power gating is necessary and can cause annoying and potenially hazardous noise, depending on the design. Further it has heat dissipation problems and requires precision machining.

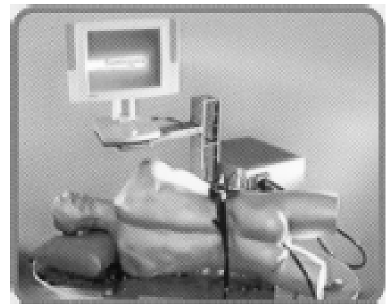

Fig. 9. GI-Mentor from Simbionix

3.2.7 Shape Memory Alloy(SMA)

See Shape Memory Alloy for Tactile Feedback.

4 Further Examples in Detail

There are many examples of Haptic Interfaces in Virtual Environments. From interfaces providing haptic feedback to only one Finger until full body haptic interfaces. I try to give widespread examples of different device types, so that after reading this chapter, the reader can imagine what exists and is possible today. Starting with the most used and common interface at moment: the PHANToM. Then i will present a glove - the Rutgers Master. After that I present an exoskeloton interface, followed by the haptic hand. And finally the most difficult haptic interface to construct: a full body force-feedback interface.

Last chapter categorized the interfaces depending on their technology used. Another categorization would be the following [15]:

- exoskeletons and stationary devices
- gloves and wearable devices
- locomotion interfaces
- full body force feedback

Some products using more then one actuator technology explained in the chapter before.

4.1 The PHANToM

The PHANToM was developed at MIT and is now marketed by SensAble Devices. The first PHANToM provided haptic feedback, by putting the users finger into a thimble. It was designed after three basic criteria [26]:

- Free space must feel free: There should not be any external forces on a user moving through a free virtual space. Static Backdrive friction for the first PHANToM was less then 0.1 Newton.
- Solid virtual objects must feel stiff: PHANToM normally providing a maximum stiffness of about 30-300 Nt/cm. Take in mind, that most Users would be convinced that a virtual surface with a stiffness of 20 Nt/cm represents a solid, immovable wall.
- Virtual constraints must not be easily saturated: That means the user should not fall through a virtual wall leaning against it

The most effort was put in the research of computing reaction forces in response to contact and motion, while exert these forces on the user was relatively easy with the phantom.

Today there are many different PHANToM interfaces providing force feedback and positional sensing. They differ in the count of DOFs (Degree of Freedom). Normally there are up to 6 degrees: Heaving, swaying, surging, pitching, yawing, rolling. Further they differ in weight, resolution, provided forces etc. As an example, the Phantom3.0/6DOF is illustrated *(see figure 10)*.

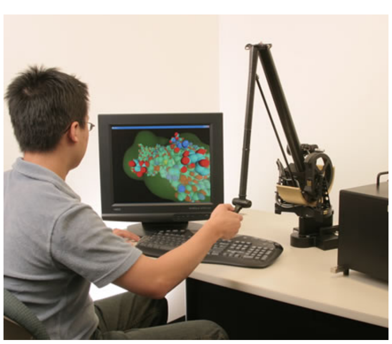

Fig. 10. Phantom3.0/6DOF [3]

Today there are countless fields of application e.g.: Surgical Simulation, haptic painting, etc..

4.1.1 Fakespace

FakeSpace is an example to imitate a 3D environment using PHANToM *(see figure 11)*, introduced here. The system uses a 3D display system calld "Fakespace" for a visual output. Two projectors with stereoscopic glasses provide high-quality, real-time stereo images on the screen. The user now uses special glasses to view 3D models. Secondly it uses a 3DOF PHANToM to provide haptic feedback. The user can now interact with virtual deformable objecs (e.g.: change its shape) and gets real-time force feedback. Every object in this world has a so called haptic interface point (HIP) to touch and feel the object and constraint points. The deformation is controlled by the selected constrained points and the HIP.

The hole system is multithreaded, one haptic and one visual thread. This is being implemented by Microsoft Visual C++.NET and OpenHaptics Toolkit. In the future there will be multiple user and object support implemented. [31]

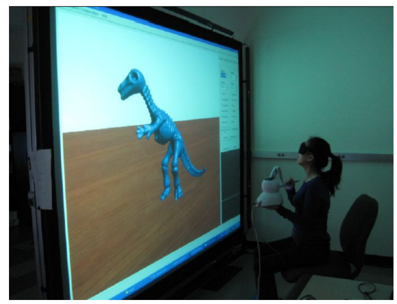

Fig. 11. Deformation of Virtual Objects, using PHANToM as haptic interface)

4.1.2 Voxel Man

The Voxel Man is another example. It is a surgery simulator for training and planning surgical access to the middle ear *(see figure 12)*. Voxel-Man Sinux, an add-on for Voxel-Man is under development.

Fig. 12. Voxel Man (Picture taken from www.voxel-man.de)

4.2 Rutgers Master

The next system introduced here is a glove. The Requirments of gloves are maintained forces to multiple fingers, light weight, safe in use and arm freedom.

The Rutgers Master II-ND *(see figure 13)* is a pneumatic actuators based system. The main advantage over the PHANToM is, that it provides haptic feedback for each finger and can sustain high forces.

There are two Hall-effect sensors to measure the flexion and adduction/abduction angles and an infrared sonsor measuring the translation of the piston inside an air cylinder. They are non-contact, so that there is no friction force in the process of measuring position. This is important for any haptic device, because it affects the sensitivity and dynamic range. To model the virtual 3D-hand the Rutgers Master uses three sensor measurements. And last the "Haptic Control Interface" controlls the force-feedback glove.

This glove is in use in several Virtual Environments applications, ranging from hand rehabilitation to military command and control. [12]

Fig. 13. The Rutgers Master II-ND

4.3 Virtual Car

The Virtual Car developed at Swiss Federal Institue of Technology Lausanne (EPFL) *(see figure 14)* combines direct manipulation with indirect manipulation and is an example for exoskeleton haptic interfaces. The user interacts directly with the mediator, in this case the car. But the mediator world interacts only indirectly with the rest of the world. The direction is given with a higher level interface, which just indicates the direction instead of driving it with precise movements through a steering wheel.

The mediator consists of a pair of 3D handles. The user can now manipulate the virtual handles and receives force-feedback through the Haptic Workstation. [23]

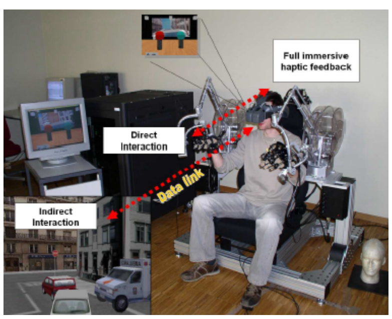

Fig. 14. Virtual Car

4.4 The Haptic Hand

The Haptic Hand developed at the university of North Carolina at Chapel Hill, provides feedback with the non-dominant hand (NDH) in virtual environments *(see figure 15)*.

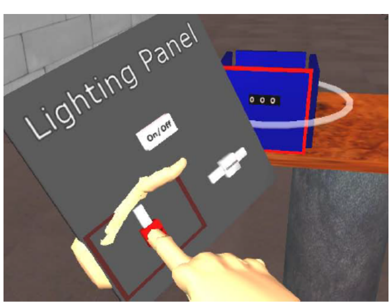

Fig. 15. Interacting with the Virtual Environment

Fig. 16. The NDH Interface

In this system u interact with the VE with hand gestures. There are virtual representations of the users's two hands. The NDH supports two poses (open and closed), while the DH suports a pose with a pointing index finger.

By closing the NDH, the user can choose a panel. The panel, which intersects most with the NDH gets highlighted. Now the user opens the hand and the choosen panel gets snaped as well as centered to the palm. It is now important that the panel now moves and rotates with the NDH, because there is no physical manifastation of the interface panel and it is almost impossible to stabalize that hand.

The next step is the manipulation of the panel, where now the DH comes into play. Buttons are pressed and released using discrete, ballistic motions. Sliders could be changed bei dragging along the surface of the NDH.

The technology used to track the user's hands are Polhemus Fastrak magnetic trackers. To not obstruct the surface of the palm, the trackers are mounted on the back of the user's hand. The main problem is, that the surface of a hand is not planar. So the trackers do not always know, if the DH is in contact with the NDH palm. When the users lifts the DH finger off of the NDH palm it could be that a slider still is active and the value changes from the intended one. To solve this problem there is a slider locking mechanism implemented. A slider gets locked after holding a certain value for at least 200ms.

In a study to find out about the how well the NDH interface works, most participants agreed that it was very intuitive.[18]

4.5 Full Body

Today there are not many full body implementations of haptic feedback interfaces. One of them is the TactaVest using the TactaBoard system *(see figure 15)*. This system incorporates the control of 16 vibrotactile devices into a single interface. The TactaBoard understands a set of simple commands. These commands set an output to the user (i.e. "set motor 1 to level 150") One major problem is the variation in size of wearers. The garment needs to keep the tactors tight against the body, even during movement and it has to fit different sized users. Thats why the garment was made out of five individual pieces of stretch neopren.

The TactaVest was integrated in a Virtual Environment system of the Immersive Simulation Section (ISS) at the U.S. Naval Research Laboratory (NRL). The group is part of a team working to create fully virtual simulator for close-quarter battle. It involves small groups of Marines operating inside buildings, while this buildings may be dark or filled with smoke. So there will be intentional and unintended contact with walls or obstacles. The TactaVest should simulate this as accurate as possible. [25]

5 Conclusion and Future Work

> "The history underlying the development of haptic technologies has, it must be said, benefited from more innovation, enthusiasm and excitement than that of the visual display industry and it is those qualities that have helped to produce the intuitive systems and stunning applications evident today. The best is yet to come!" [30]

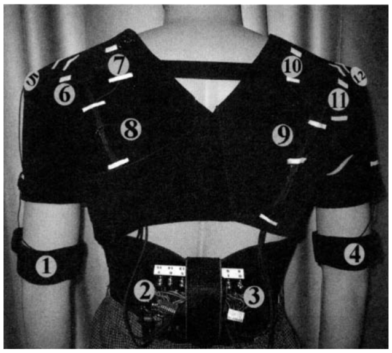

Fig. 17. The TactaVest [25]

Today there is much research done and technology wil most likely progress in the next years. Initial advances will be application-specific, most likely in the medical industry, because there is much interest in supporting surgical procedures. At the moment they are mainly restricted to research applications. The holy grail - free motion and full-body all-senses feedback is far from being realized.

But how could it work? Some ideas have been adressed in [5]. Lamellar corpuscles or Pacinian corpuscles cannot differentiate between increased pressure and lessened pressure. Both deform the corpuscle in opposite ways, but the fibres of the corpuscle deform in the same way. Thats why it is possible to simulate pressure on the skin, by vacuum sucking it. That way a Full Body Haptic Suit could be constructed *(see figure 18)*.

Another way is direct neural stimulation. Instead of trying to deal with thousands of corpuscles, we can try to access the brainstem directly. The brainstem is a nerve bundle where all these signals converge on their way to the brain. Thsese artificial signals tell the body it is being touched, without the actual touching taking place.

What would be possible? Perhaps we could even feel cold wind on our skin, walking down the beach or we could feel water sourrounding our feet, when a wave comes. Or Maybe we get a realistc feeling of the warm sunrays touching our skin. But one thing is sure:

Once the full body haptic feedback interfaces are realized, the applications coming to mind are unlimited. One exciting thing surely is CyberSex with NPCs. Virtual Worlds like the computer game "Second life", would become increasingly attractive. With virtual sex feeling completely realistic, this is certainly a climax in this developement. The question about the moral, people spending more and more hours in those worlds and not having a real life any more, is not being discussed here.

I think I dont have to point out the revolutinary impact in military training, surgery and almost every other field, those haptic interfaces will have. But it is almost impossible to predict, when this technology will arrive [8].

Fig. 18. Full Body Haptic Suits. When will they come? [8]

References

[1] http://de.wikipedia.org/wiki/piezoelektrizit
[2] http://en.wikipedia.org/wiki/solenoid.
[3] http://www.sensable.com/haptic-phantom-premium-6dof.htm last visited: January 2012.
[4] http://www.steelguru.com/article/details/njg=/shape memory alloys.html last visited: January 2012.
[5] http://www.virtualworldlets.net/resources/hosted/resource.php?name=haptics-pressure: Haptics - pressure, last visited: January 2012.
[6] K. Akahane, C. Yu, X. Liu, and M. Sato. The research of light-weighted finger haptic device using voice coil.
[7] Akamatsu. 1994.
[8] M. Anissimov. http://www.acceleratingfuture.com/michael/blog/2007/11/full-body-haptic-suits/ last visited: January 2012. 2007.
[9] O. Bau, I. Poupyrev, A. Israr, and C. Harrison. Teslatouch: Electrovibration for touch surfaces.
[10] M. Benali-Khoudja, M. Hafez, J.-M. Alexandre, and A. Kheddar. Tactile interfaces: a state-of-the-art survey.
[11] A. Bernstein and R. Lenhardt. The haptic spacemouse - an input device with force-feedback through solenoids.
[12] M. Bouzit, G. Popescu, G. Burdea, and R. Boian. The rutgers master ii-nd force feedback glove. 2002.
[13] G. Burdea. Haptic feedback for virtual reality.
[14] S. Ellis. What are virtual environments.
[15] A. Fisch, C. Mavroidis, J. Melli-Huber, and Y. B. Cohen. Haptic devices for virtual reality, telepresence, and human-assistive robotics. 2003.
[16] T. Kaaresojai, E. Hoggan, and E. Anttila. Playing with tactile feedback latency in touchscreen interaction: Two approaches.
[17] K. Kaczmarek. http://en.wikipedia.org/wiki/solenoid. 1991.
[18] L. Kohli and M. Whitton. The haptic hand: Providing user interface feedback with the non-dominant hand in virtual environments.
[19] O. Körner and R. Männer. Haptic display for a virtual reality simulator for flexible endoscopy.
[20] C. Kozierok. http://www.pcguide.com/ref/hdd/op/act actuator.htm last visited: January 2012. 2001.
[21] A. Lecuyer. Simulation haptic feedback using vision: A survey of research and applications of pseudo-haptic feedback. February 2009.
[22] J. Lee, P. Dietz, D. Leigh, W. Yerazunis, and S. Hudson. Haptic pen: A tactile feedback stylus for touch screens.
[23] P. Lemoine, M. Gutierrez, F. Vexo, and D. Thalsmann. Mediators: Virtual interfaces with haptic feedback.
[24] K. Li, P. Baudisch, W. Griswold, and J. Hollan. Tapping and rubbing: Exploring new dimensions of tactile feedback with voice coil motors.
[25] R. Lindeman, R. Page, Y. Yanagida, and J. Sibert. Towards full-body haptic feedback: The design and deployment of a spatialized vibrotactile feedback system.
[26] T. Massie and J. Salisbury. The phantom haptic interface: A device for probing virtual objects.
[27] Patrick, Sheridan, and Massimo. 1990.
[28] E.-L. Sallnas, K. Rassmus-Gruehn, and C. Sjoestroem. Supporting presence in collaborative environments by haptic force feedback.
[29] Sheridan and Massimo. 1993.
[30] P. R. Stone. Haptic feedback: A potted history, from telepresence to virtual reality.
[31] Z. Tang, A. Patel, X. Guo, and B. Prabhakaran. A multimodal virtual environment for interaction with 3d deformable models.
[32] C. Youngblut, R. Johnston, S. Nash, R. Wienclaw, and C. Will. Review of virtual environment interface technology. 1996.

Technology for Tactile Tangibles

Horst-Egon Brucker

Abstract— Interaction between man and machine has long been restricted to using the classic peripherals like keyboard, mouse and joypad/joystick as input devices and the monitor respectively the television as output device. With recent developments in the video gaming industry other means of input, like the Wii Remote(TM), are gaining more and more importance. But the feedback from such devices is still reduced to a mere rumble/vibration or force feedback. However as stated by Robles-De-La-Torre the sense of touch and therefore tactile feedback is at least as important for some applications as the visual or auditory senses [23]. Therefore in the recent few decades researchers have been studying various technologies and techniques to convey such a tactile experience to the user. This paper concentrates on the technology behind the actuators for tactile interfaces. First the technologies will be classified and the functionality of those actuators will be explained. Subsequently there will be some examples of proposed devices for each class.

Index Terms—Tactile, Tangible, Actuator, Haptic

1 Introduction

In recent years advances in Virtual Reality (VR) respectively Augmented Reality (AR) Applications have been made. Although this progress is mainly evident in the visual and auditory sections there has been a lot of research into interfaces that can stimulate the other human senses, like the sense of touch, as well. As a prerequisite for a more realistic experience in interacting with Virtual Environments (VEs) it is important that an interface to the computer is capable of such kind of interaction. Therefore a crucial aspect of Human-Computer Interaction (HCI) is the stimulus of the somatosensory system. Such stimuli present the user with needed information that, according to Robles-De-La-Torre, in some cases, like robotic surgery, is imperative to fulfilling that certain task [23]. The simulation of such a feedback from an application like a Virtual Environment, Teleoperation, Telepresence or, like aforementioned, medicine, is called tactile feedback. In conjunction with newly developed tangible interfaces, that "augment the real physical world by coupling digital information to everyday physical objects and environments" [13], a whole new scope of Human-Computer Interaction is possible.

In the following chapters I will give a brief outline of the development of Tangible User Interfaces (TUIs) and the definitions for tactile and haptic feedback and I will define the Tactile Tangible User Interface (TTUI) as a combination of both tactile and tangible interfaces.

1.1 From classical GUIs to TUIs

Typical Human-Computer Interaction (HCI) is based on classical GUIs (Graphical User Interfaces). The first GUIs were developed in the early 1980s. One idea behind this development was to create a metaphor for the workspace, e.g. the desktop, folders and files. Other design principles, used to create GUIs, were "seeing and pointing" and "what you see is what you get" (WYSIWYG). Using a monitor, a mouse and a keyboard was sufficient to achieve these goals. Nowadays the paradigm of GUIs is pervasive and in almost every device available, whether it is a smartphone, a tablet PC or the classical PC at home or at work [13].

However, by using keyboard and mouse as input devices and a monitor as output device, the interaction with (respectively the control) and the representation of digital data is decoupled from each other. For example to move a folder on the "virtual" desktop we have to use the mouse, and the representation of the folder is a pixel image on the screen. Newer touchscreens in tablet PCs or smartphones may loosen this restraint a little but the representation is still decoupled from the physical world. Moreover the whole interaction is limited to a two dimensional input via mouse or touchpad/-screen and output via flat two dimensional display [13].

To escape these shortcumings Ishii et al. defined a new concept, the "Tangible User Interfaces" (TUIs)[13], which "will augment the real physical world by coupling digital information to everyday physical objects and environments" [13]. This means that in the simplest relization of a TUI a single physical object will represent digital data and by interacting with that physical representation this data can be manipulated. More complex forms of TUIs "will change the world itself into an interface" [13]. The important difference to classical GUIs and the associated in- and output by means of classical peripherals (such as mouse, keyboard or screens) is, that the representation of digital data and the control for the manipulation of that data is coupled into the same physical object.

A good example for a Tangible User Interface is SandScape *(see figure 1)* [12]. Users can manipulate the sand and can introduce various other objects and thereby influence the form of the landscape, that is represented as a computer model. In turn the calculated results, in terms of for example the landscape height, drainage or other aspects, are projected onto the sand surface [12].

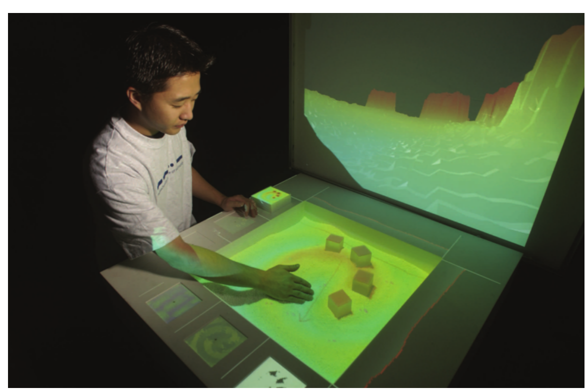

Fig. 1. SandScape: "Users alter the form of the landscape model by manipulating sand while seeing the resultant effects of computational analysis generated and projected onto the surface of sand in real time." [12]

• *Horst-Egon Brucker is studying Informatics at the University of Munich, Germany, E-mail: brucker@cip.ifi.lmu.de*

• *This research paper was written for the Media Informatics Advanced Seminar 'Over Distance', 2011/2012*

1.2 Tactile and haptic feedback

Oakley et al. [19] defines the term haptic as "relating to the sense of touch". The author then subdivides haptics further, which encompasses both cutaneous and kinesthetic senses. Kinesthetic "means the feeling of motion" and is "related to sensations originating in muscles, tendons and joints". A subset of kinesthetic feedback is force feedback which is "related to the mechanical production of information sensed by the human kinesthetic system". Therefore "force feedback devices affect the finger, hand, or body position and movement". On the other side the cutaneous sense "pertains to the skin itself or the skin as a sense organ". It "includes sensation of pressure, temperature, and pain". The tactile sense, as a subset of the cutaneous sense, relates "more specifically [to] the sensation of pressure rather than temperature or pain". "Tactile devices affect the skin surface by stretching it or pulling it, for example." [19] [1]

1.3 Enhancing TUIs with touch

A Tangible User Interface by design has a passive tactile feedback which Ishii et al. [11] define as first interaction loop. This feedback is felt by the user simply by interacting with the object itself, that represents digital data *(see figure 2)*. However, this kind of tactile feedback is not generated by the computer as a result of a computation in response to the user interaction. A second and third interaction loop need to be defined. The second loop, being a simple intangible response by the computer via visual or auditory displays, is negligible for the purpose of this paper. The third loop provides the needed interaction definition to implement active haptic feedback into a tangible user interface. It enables the computer to utilize actuators built into the tangible device to pass haptic feedback to the user either as a result of a user interaction or as haptic cues (depending on the application). Considering that tactile feedback is a subset of haptic feedback, by focusing on actuators that only convey tactile feedback the resulting interface is a Tactile Tangible User Interface (TTUI).
In short: "Tactile Tangibles are physical objects that allow for the manipluation of digital elements and communicate tactile information." [21]
However in this paper I will shift my focus on the tactile actuator technology used in Tactile Tangible User interfaces.

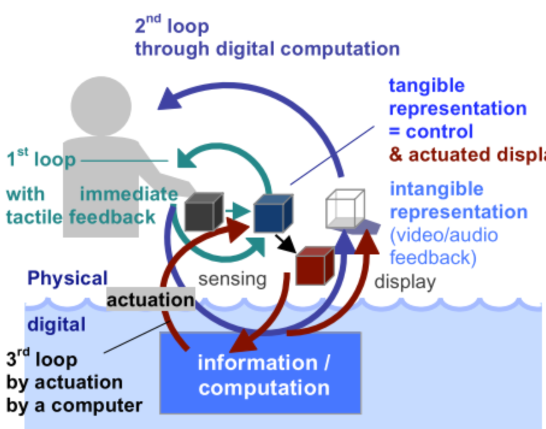

Fig. 2. The different interaction loops for TUIs with haptic/tactile feedback [11]

1.4 Restrictions

There are not many examples that fit into both the tactile aspect and the strict tangible UI definition. Therefore I will loosen up the restrictive definition and consider:

- every tactile feedback as representation of digital data
- the interface as tangible even if it is not representing some specific digital data

[1] All of the definitions for haptic to tactile feedback are taken from Oakley et al. [19].

1.5 Outline

In the following chapters I will classify the actuator technology used for tactile interfaces and explain the functionality of those technologies. Afterwards there will be some examples for each class of actuator. In conclusion there will be a discussion of some of the potentials and drawbacks of the used technology.

2 Actuator Technology

Actuators commonly used to convey tactile information in currently researched applications vary in size, shape and usage modalities. But for this paper it is necessary to classify them based on the technology they employ.
In general they can be assigned to one of the following classes:

- Vibrational
  This class encompasses some examples for devices which can generate mechanical vibration, like vibration motors, voice coils, solenoids and piezoelectric elements.
- Pneumatic
  The common principle of pneumatic actuators is the usage of compressed gas, for example to inflate and deflate a balloon or a shell/reservoir like device covered with a membrane, that can simulate a touch feeling.
- Electrical
  This class is subdevided into two application areas. The first area is utilizing elecricity that is directly applied to the skin and the second is taking advantage of an effect called electrovibration.

2.1 Vibrational

Vibro-tactile actuators generate vibration that can be registered by the skin. Most of the vibro-tactile feedback actuators use one of the following technologies:

2.1.1 Vibration motor

The most common form of vibration motors are the eccentric mass vibration motors, which use an unbalanced weight attached to the shaft of an electric motor. When the shaft rotates this imbalance causes the motor to vibrate. These types of motors are known from their application in cell phones and pagers [18]. Besides the cylindrical vibration motors *(see figure 3)* there are also pancake or coin shaped vibration motors *(see figure 4)* which also allow smaller and thinner application designs [18].

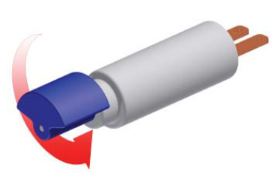

Fig. 3. Cylinder type vibration motor - an unbalanced weight is causing the device to vibrate [18]

2.1.2 Voice coil

Another vibro-tactile actuator that can be used to generate vibration is the voice coil. A voice coil, known for its application in loudspeakers, usually consist of an insulated copper wire which is wound around a cylindrical core or form to create an electromagnet. This electromagnet resides within the magnetic field of a permanent magnet. Older designs are built with permanent magnets, that have an axial magnetization. As soon as a current flows through the coil wire another magnetic field is created, that, in conjunction with the magnetic field of the permanent magnet, either attracts the coil to the permanent magnet

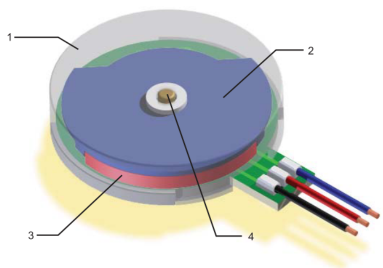

Fig. 4. "Coin-type vibration motor. A relatively flat eccentric weight spins in a protective enclosure. (1) enclosure, (2) rotor base, (3) weight, (4) shaft." [18]

or repells it, depending on the polarity of the electric current. Newer designs use premanent magnets that have a radial magnetic field. The physical law that is used for such a device is called the Lorentz Force Principle. This is the force which is exerted upon an electric charge passing through a magnetic field. Thus, by applying an electric current to the coil, the Lorentz Force will accelerate it in the direction that is at a right angle to both the electric current direction in the coil conductor and the magnetic field of the permanent magnet. Again, by reversing the polarity the acceleration direction is also reversed. Changing this polarity with a certain frequency will cause the voice coil to vibrate with that freqeuency. The force that is exerted by the voice coil onto another device can be influenced by the amplitude of the electrical signal *(see figure 5)*. The voice coil actuator depicted in the schematic is from linear type, that moves in a straight line. However, there also exist rotary voice coil actuators, whose motion is in an arc at a limited angle. These are for example built into computer hard drives to position the read/write heads. For the extent of this paper, however, I will not further address this type of voice coil actuators, since it hasn't been used in any of the examples for Tactile Tangibles described herein [6][8][9].

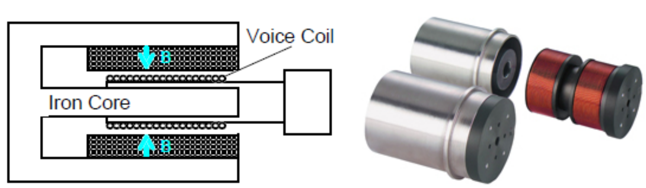

Fig. 5. Voice coil schematic (left); B is the magnetic field of the permanent magnet [26]. Example of a linear voice coil actuator (right) [3]

2.1.3 Solenoid

An electromechanical solenoid is a sort of electromagnet. It consists of a fixed helically wound coil in a steel housing with a movable steel or iron plunger *(see figure 6)* [16]. The plunger's initial position is displaced outwards from the middle of the coil. When the current is turned on, the coil acts as an electromagnet and is causing the steel plunger to be magnetized in such a way that the plunger is attracted by the magnetism and pulled into the coil. After turning off the current the magnetic field of the coil and the magnetization in the plunger will dissipate and by using a spring the plunger can be reset to its initial position. On *figure 6* a push type solenoid is illustrated, which can actuate the application device by pushing the pin. The pull type solenoid is similar to the push type solenoid but has no need for the push pin because the application device, i.e. actuated device, is directly attached to the plunger [16]. Like a voice coil a solenoid can be used to create a vibration or a buzz. However, when applied directly to the skin, it could also be used to exert a certain amount of pressure on the skin [15].

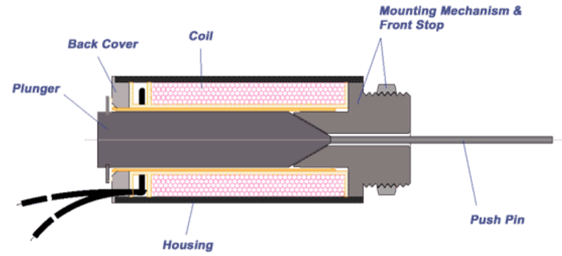

Fig. 6. Electromechanical push type solenoid [16]

2.1.4 Piezoelectric element

Piezoelectric elements are special crystals that can generate a voltage when a mechanical force is exerted on them *(see figure 7 "generator action")*. The polarity of the voltage generated depends on whether it was a compressive force or a tension. When a voltage is applied to the piezo elements with the same polarity and direction of the poling voltage (of the piezo crystal), the crystals will lengthen. A voltage with polarity opposite that of the poling voltage will shorten the piezo element *(see figure 7 "motor action")* [1].

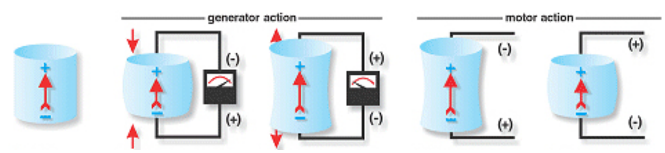

Fig. 7. Piezoelectricity [1]

2.2 Pneumatic

The principle utilized by pneumatic atuators is the usage of compressed gas to inflate or deflate the actuator, which usually "consists of [some kind of] a 'hard' shell/reservoir with a 'soft' membrane covering the opening of the shell" [17] *(see figure 8)*. By applying an oscillating air flow the pneumatic actuator will vibrate. Such air flows are typically generated via solenoid valves which are "connected to either a compressor or pressurized air tank". The vibration of the membrane causes strong tactile feedback [17]. Generally pneumatic actuators work at different controllable frequencies and intensities [25].

Fig. 8. Pneumatic tactile actuator [25]

2.3 Electrical

2.3.1 Electrotactile

This Method is also known as electrocutaneous stimulation. Actuators for this method of stimulation typically use electrodes that are attached to the skin. In some cases they are also attached subcutaneously. The size and placement of such electrodes allows to choose which part of the skin is affected by the electricity. Based on this choice the current could have a stimulating effect on the skin receptors or entire nerve bundles. Depending on the voltage, currents of different frequencies, amplitudes and waveform, the electrode size and material, the skin location and its attributes like thickness and hydration and the contact force of the electrodes that are used, electrotactile stimulation can induce sensations like tingle, itch, touch, pressure, vibration, buzz, pinch or pain [14].

2.3.2 Electrovibration

This effect was discovered in 1954. It uses the phenomenon, that applying an alternating voltage to the finger on one side and an insulated surface with an electrode beneath on the other side creates the two opposing plates of a capacitor. The result of this effect is that opposing charges exert an attractive force, the so-called coulomb force, on the finger which in turn increases the friction between the finger and the touched surface. This can create a feeling of a ”rubbery” surface. With finer detailed signals and charges the impression of suface texture, contours and edges can be simulated. Typical applications range from arrays of electrodes like metal pins insulated by a thin layer of dielectric to one transparent electrode covering the entire display and making it possible to combine input and visual and tactile output in one display [2][24].

3 Examples

In this section I will present some examples which use the actuator technologies described above.

3.1 Vibrational

CyberTouch

A general[2] example for a vibro-tactile application is the CyberTouch from CyberGlove Systems. First released in 1995 the CyberTouch is based on the CyberGlove, which provides sensors for the movement/bending of the fingers, flexion of the hand and wrist. The CyberTouch essentially is an extension for the CyberGlove. It provides 6 vibro-tactile actuators, five for the fingers and thumb and one for the palm *(see figure 9)*. The actuators can generate pulses, sustained vibration, or customized vibration patterns. The application has a frequency range of 0-125 Hz and a vibrational amplitude of 1.2 N peak-to-peak (at 125 Hz) [7].

Fig. 9. CyberTouch: A CyberGlove with vibro-tactile feedback [7]

3.1.1 Vibration motor - Haptic Wheel

One example of a tactile tangible interface that uses a vibration motor to generate vibro-tactile feedback is the Haptic Wheel. Due to its primary application intent, which is the eyes-free entry of pins, it resembles a dial from a safe, which can be rotated clockwise and anticlockwise. The hardware assembly is comprised of a rotational dial with a ”rotary encoder to track its absolute orientation, a binary switch mounted on its top center (for selection input) and a low-cost eccentric vibration motor to generate haptic feedback in the form of vibro-tactile stimuli” *(see figure 10)* [5]. By driving the vibration motor with different frequencies it is possible to provide structured vibro-tactile cues (tactons). Interaction with the Haptic Wheel is performed by rotational input by means of the dial and activation of the top button. Because of its design it is also capable of delivering tactile cues in the user’s passive state, i.e the user only touches the device but does not interact by turning the dial nor pushing the button [5]. A possible implementation which is taking advantage of the Haptic Wheel is a tactile password entry system [4]. The system encodes the password as a sequence of tactons [5]. The password input is accomplished by the user sensing each tacton and selecting the corresponding target angle [4] by turning the dial and clicking the button [5]. Besides, denoted application possibilities also include the usage of the Haptic Wheel as a puck for a tangible tabletop displays [5].

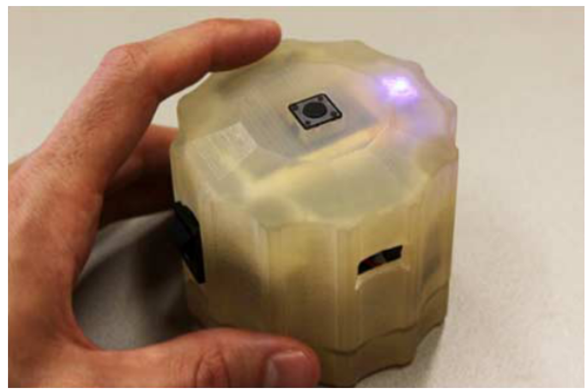

Fig. 10. The Haptic Wheel [5]

3.1.2 Voice Coil - HapTouch

The HapTouch [22] is an in-vehicle information system (IVIS), that was developed by Hendrik Richter et al. at the University of Munich with the support of members of the BMW Group Research and Technology *(see figure 11)*. This application aims at providing benefits in terms of both a lower error rate for the driver’s inputs into the system and a better driving performance due to reduced visual distraction of the driver by the IVIS. These benefits are realized by using a force-sensitive touch display, that is actuated by a voice coil to generate tactile feedback *(see figure 12)*. The usage of these additional force sensors allow the user to touch the display without activating the UI elements displayed (touching force is less than a predefined threshold). In this state the user can move the finger over the display and feel the UI elements (e.g. border edges, surfaces, etc.) via tactile feedback from the voice coil. By pressing harder (force greater than the threshold) the UI elements can be activated (e.g. button click or dragging). When pressing on the display even more a new state is reached, in which the user can for example execute new actions like zooming or resizing, depending on the pressure he applies to the screen. Every user action can cause the display to be actuated by the voice coil to give a tactile feedback to that respective action, for example a button click could trigger a mechanical ”snap”, like it was a real button that was pressed.

3.1.3 Solenoid - Haptic Pen

One of the applications that uses a solenoid as a tactile actuator is the Haptic Pen. It is based on a stylus design pen that is enhanced with a small push-type solenoid. The plunger of the solenoid is attached to the pen while the coil (in its housing) is freely movable, thus representing the actuated mass *(see figure 13)*. The primary force is then generated by applying a current to the coil which accelerates the coil away from the tip. In this scenario for the reset to its initial position the gravity force is used rather than a spring or other mechanical device.

[2] Due to the fact, that no information about the technology used for the actuators is provided, neither in the literature referencing this application nor on the manufacturers webpage, I cannot classify this example further.

Fig. 11. "The HapTouch system is touchscreen device with tactile feedback" [22]

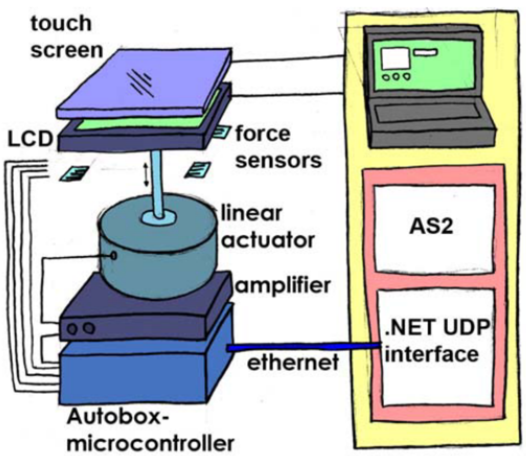

Fig. 12. HapTouch schematic overview [22]

The solenoid's function is triggered by a microcontroller. Furthermore a pressure sensor is used to provide the controller with data about the pressure the user applies to the pen and the surface. This data can be used to operate the solenoid in different ways that reflect the user interaction. Some of the actions defined include a lift, hold, hop or buzz of the actuator. The lift accelerates the mass upwards, the hold keeps the solenoid in the lifted position, the hop causes the actuator to lift and immediately fall to its resting position again, and the buzz generates an oscillating lift and drop action. For each of these actions the strength can be controlled as well. Due to its design, the Haptic Pen is able to provide individualized feedback in a multi-user setting [15].

3.1.4 Piezoelectirc element - The TouchEngine

An example for a piezoelectric actuator is the TouchEngine. It consists of multiple extremely thin layers of piezoceramic with electrodes in between. The physical principle of piezoceramic is that by applying a current to the material it will expand or contract depending on the applied polarity. Because of the opposing polarity on both sides of the layers the layers will contract and expand in a way that the entire structure will bend, thus giving it the name "bending motor" actuator *(see figure 14)*. Depending on the application demands it can be built with a different number of layers [20]. The displacement and the force with which the piezoceramic is bent is directly (linear) related to the voltage applied. "The greater the acceleration of the actuator, the stronger the impulse force becomes [...] and the sharper the tactile impulses can be detected by the user" [20]. There are two basic possibilities to use the TouchEngine: the direct approach, which actuates parts of the device, like buttons or displays, and the indirect method, where the entire device acts as a tactile display. Poupyrev et al. used the TouchEngine in a mobile device *(see figure 15)* with different applications, e.g. browsing

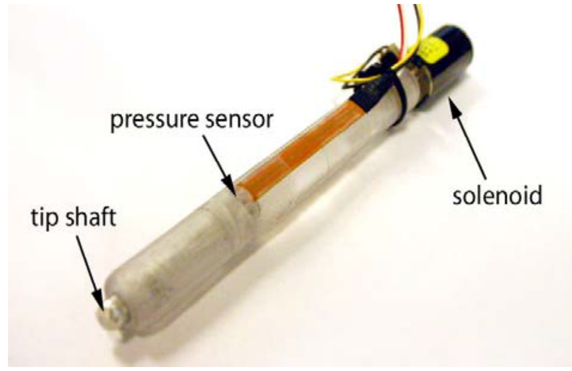

Fig. 13. Haptic Pen: Tactile feedback by means of a solenoid [15]

a subway map. By tilting the mobile phone the map is scrolled in the respecitve direction of the tilt, which causes a scratching tactile feedback. A different tactile pulse is actuated through the TouchEngine on reaching the map's boundaries.

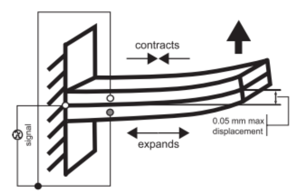

Fig. 14. Piezo bending motor [20]

Fig. 15. Piezo bending motor [20]

3.2 Pneumatic

In an effort to combine the benefits of physical buttons, such as low-attention and vision-free interaction, with advantages of touch screens, such as flexibility, Harrison et al. designed "a technique for creating dynamic physical buttons using pneumatic actuation" [10]. The interface consists of one or multiple pneumatic chambers which are built by using transparent and translucent materials. The chambers are composed of layers of transparent acrylic, that are specially cut to represent the interface buttons. On top of the acrylic pattern there is a layer of translucent latex, which is fixed to the acrylic with adhesive *(see figure 16)*. The interface itself is projected onto this latex layer by a rear projector.
The latex layer allows a convex deformation, when there is a positive pressure applied to the chambers, and a concave or inward deformation for a negatively pressurized chamber. With this method it is possible to dynamically manipulate the pneumatically actuated interface by the controlling software in a way that can "activate" (by inflating) or "deactivate" (by deflating) buttons *(see figure 17)*. This behaviour also poses the tactile feedback of the interface. Depending on the type of application, this interface is cusomized for, it also may consist of one individually controllable pneumatic chamber per button.

The pressure for actuating the interface buttons is created via a fan-based pump. With a solenoid valve it is possible to generate both negative and positive pressure for the chambers. This allows for a rapid pressure change enabling the buttons to pulsate. Furthermore you can ”create buttons that push back when pressed, or ’snap in’ when depressed” [10]. For input sensing the application uses two types of sensors. On the one hand the fingers approaching the device are sensed by an infrared camera behind the display and on the other hand by pushing the buttons the pressure in the chamber changes and is registered by a pressure sensor. The combination of these sensing techniques is necessary because both, if used independently, have some ambiguity issues. For example the camera can not distinguish between a finger that presses the button or just hovers over it whereas the pressure sensor can not identify which button was pressed when more than one buttons share a pressure chamber or the sensor is used for multiple chambers [10].
The tangible aspect of this interface is constituded by the active alteration of the button representation. Further the tactile feedback is ”provided by [the] outward facing features, such as convex deformations and edges”, when the buttons are inflated and in case of deflated buttons by ”depth and curvature” [10].

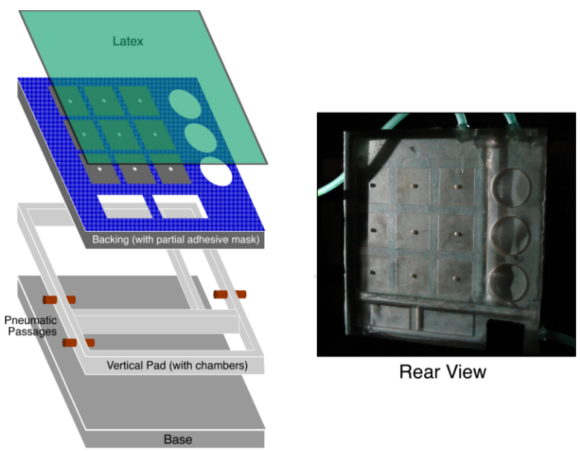

Fig. 16. Design example of a multi-chambered pneumatic tactile display [10]

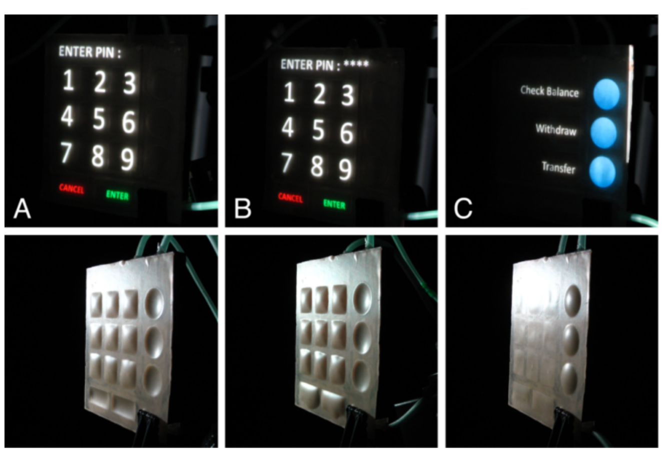

Fig. 17. Application example of the pneumatic display with different tactile configurations [10]

3.3 Electrical

3.3.1 Electrovibration - TeslaTouch

For this application a capacitive based microtouch panel, originally designed for touch sensing, was used. It consists of a transparent electrode on a glass plate and insulated by a thin layer of dielectric. Upon application of a periodic electrical signal to the transparent electrode this electrode and the touching finger act like an electric capacitor whose two opposing charges exert an attractive force, thus changing (increasing) the friction between the finger and the panel surface *(see figure 18)*. Changes in friction an skin deformation vary with the signal amplitude. The deformations of the skin are perceived as vibration or friction. Also the deformations can be controlled by the signal amplitude and frequency. The implementation for the signal output of TeslaTouch uses a sinusoidal waveform that is generated by a sound card and amplified by an operational amplifier and a transformer to a maximum of about 120 Vpp. The charge of the input signal is uniformly distributed across the whole conductive layer of the display, therefore creating a spatially uniform tactile sensation. To increase the intensity of this sensation the user has to be grounded by means of a ground electrode, like an antistatic wristband. In mobile devices this ground electrode could be integrated in the devices enclosure *(see figure 19)* [2].
The technology of electrovibration has some advantages. One of these advantages is the scalability and versatile scope of application. The design of this actuator allows for its usage on many different surfaces whether in size or in shape. Compared to mechanical vibration feedback devices it also doesn’t have moving parts and thus it doesn’t have the problem of wear and tear over time. Additionally this actuator is noiseless, contrary to other mechanical vibration feedback devices [2]. However for the effect to be felt by the user, he has to move his finger(s) over the surface. Fingers that are not in motion do not experience the electrostatic effect. Bau et al. [2] mentions this fact both in an advantageous context and as a drawback since application scenarios like pressing buttons cannot be realized by this technology.

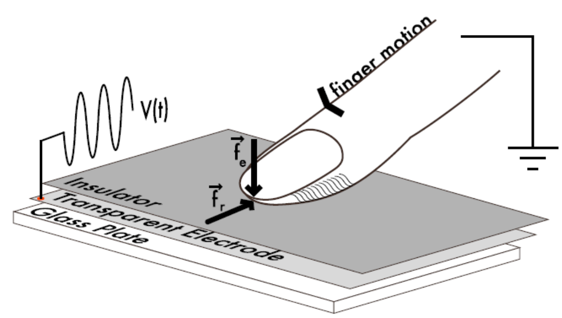

Fig. 18. TeslaTouch: Composition [2]

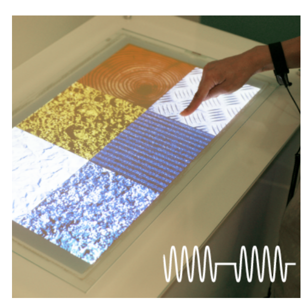

Fig. 19. TeslaTouch: Different texture and friction simulation [2]

3.3.2 Electrovibration - Senseg E-Sense Tixel(TM)

A potentially commercial application using electrovibration was built by Senseg corporation. The name Tixel is composed of ”tactile pixels". In the simplest form such a tactile display consists of a thin conductive layer of tixels between two nonconductive layers. The conductive tixel layer is then controlled by an electronics module, that creates the effect signals and electric charges. Via API the software can react on the users intput and can create tactile feedback like surface texture, physical edges and contours. *(see figure 20)* Due to the design of this display it scales easily with smaller devices like smartphones or tablet computers or even larger devices like tabletop computers. Furthermore it can also be integrated in other surfaces like device covers *(see figure 21)* [24].

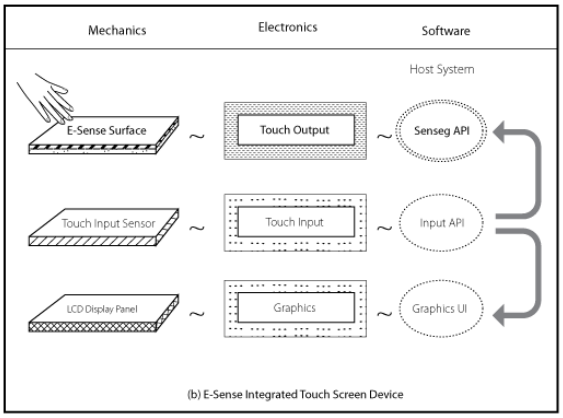

Fig. 20. The Senseg Tixel: System architecture [24]

Fig. 21. The Senseg Tixel: An electrovibration tactile interface [24]

4 Conclusion

4.1 Potentials

In general there exist many ideas and designs for tactile tangibles.

4.1.1 Vibrational

The Haptic Wheel, as described above, can be used for secure password entry in public spaces by making shoulder surfing or concealed recording impossible.

4.1.2 Pneumatic

The pneumatic display example shown before could also be used as a password entry device. It provides vision-free interaction due to intuitive handling. It enhances user experience and improves interaction for those with visual and motoric disabilities.

4.2 Drawbacks

A lot of the presented applications are very complex and unhandy due to their size or the needed wiring.

4.2.1 Pneumatic

Due to the construction of the displays using pressure chambers resulting displays/interfaces are not particularly flexible and are still bound to the one design, they were built with. Pneumatic displays do not fully scale with every possible size of the displays, e.g. it is unfeasible to try to build such a device into mobile phones. Furthermore the solution described in the examples section is relatively complex, using a rear projector, an infrared camera, pneumatic pumps and tubes to the display chambers. Also the used material's durability may be low and prone to leakage [10].

4.2.2 Electrical

In the case of electotactile or electrocutaneous stimulation it is to mention, that it is easier to utilize mechanical vibrational devices instead of attaching electrodes to the skin. Furthermore the skin conditions of the subjects differ and different placement of the electrodes can also change the percepted feeling [17]. Furthermore some of the examples, primarily the electrotactile ones, could be dangerous. If mishandled they could lead to serious injuries and damage of the skin or even death [14].

References

[1] APC International. Piezoelectricity. `http://www.americanpiezo.com/knowledge-center/piezo-theory/piezoelectricity.html`, 2012-02-26.

[2] O. Bau, I. Poupyrev, A. Israr, and C. Harrison. TeslaTouch: electrovibration for touch surfaces. *UIST '10 Proceedings of the 23nd annual ACM symposium on User interface software and technology*, 2010.

[3] BEI Kimco Magnetics. Linear Voice Coil Actuators - Cylindrical Un-Housed. `http://www.beikimco.com/actuators_linear_CYL_unhoused.php`.

[4] A. Bianchi, I. Oakley, J. K. Lee, and D. S. Kwon. The Haptic Wheel: Design & Evaluation of a Tactile Password System. *CHI EA '10 Proceedings of the 28th of the international conference extended abstracts on Human factors in computing systems*, pages 3525–3530, 2010.

[5] A. Bianchi, I. Oakley, J. K. Lee, D. S. Kwon, and V. Kostakos. Haptics for tangible interaction: a vibro-tactile prototype. *TEI '11 Proceedings of the fifth international conference on Tangible, embedded, and embodied interaction*, 2011.

[6] B. Black, M. Lopez, and A. Morcos. Basics of Voice Coil Actuators. *PCIM-VENTURA CA-*, 1993.

[7] CyberGlove Systems. CyberTouch. `http://cyberglovesystems.com/products/cybertouch/overview`, 2012-02-26.

[8] G. P. Gogue and J. J. Stupak Jr. Voice-coil actuators. `http://www.consult-g2.com/papers/paper8/paper.html`, 2012-02-26.

[9] G. P. Gogue and J. J. Stupak Jr. VOICE-COIL ACTUATORS: INSIGHT INTO THE DESIGN. `http://www.consult-g2.com/papers/paper9/paper.html`, 2012-02-26.

[10] C. Harrison and S. Hudson. Providing Dynamically Changeable Physical Buttons on a Visual Display. In *Proceedings of the 27th international conference on Human factors in computing systems*, pages 299–308. ACM, 2009.

[11] H. Ishii. Tangible bits: beyond pixels. In *Proceedings of the 2nd international conference on Tangible and embedded interaction*, pages xv–xxv. ACM, 2008.

[12] H. Ishii. The tangible user interface and its evolution. *Communications of the ACM*, 51(6):32, June 2008.

[13] H. Ishii and B. Ullmer. Tangible Bits : Towards Seamless Interfaces between People , Bits and Atoms. *Proceedings of the CHI97 conference on human factors in computing systems, Atlanta, Georgia*, pages 234–241, 1997.

[14] K. a. Kaczmarek, J. G. Webster, P. Bach-y Rita, and W. J. Tompkins. Electrotactile and vibrotactile displays for sensory substitution systems. *IEEE transactions on bio-medical engineering*, 38(1):1–16, Jan. 1991.

[15] J. Lee, P. Dietz, D. Leigh, and W. Yerazunis. Haptic pen: a tactile feedback stylus for touch screens. *Proceedings of the 17th ACM Symposium on User Interface Software and Technology (UIST)*, 2004.

[16] Magnetic Sensor Systems. Solenoid Construction. `http://www.solenoidcity.com/solenoid/manual/construction/construction.htm`, 2012-02-26.

[17] B. McGrath, A. McKinley, M. Duistermaat, O. Carlander, C. Brill, G. Zets, and J. B. F. van Erp. Chapter 4 TACTILE ACTUATOR TECHNOLOGY. *NATO - OTAN*, pages 1–12.

[18] S. Milano. Fully Integrated Hall Effect Motor Driver for Brushless DC Vibration Motor Applications. *Application Information from Allegro Microsystems Inc.*, (1):1–6, 2007.

[19] I. Oakley, M. McGee, S. Brewster, and P. Gray. Putting the Feel in Look and Feel . In *Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 415–422. ACM, 2000.

[20] I. Poupyrev, S. Maruyama, and J. Rekimoto. Ambient touch: designing tactile interfaces for handheld devices. In *Proceedings of the 15th annual ACM symposium on User interface software and technology*, volume 4, pages 51–60. ACM, 2002.

[21] H. Richter. Tactile Tangibles. *Advanced Seminar Media Informatics 2011/2012, Folien zur Einführungsveranstaltung*, 2011.

[22] H. Richter, R. Ecker, C. Deisler, and A. Butz. HapTouch and the 2+1 State Model: Potentials of Haptic Feedback on Touch Based In-Vehicle Information Systems. In *Proceedings of the 2nd International Conference on Automotive User Interfaces and Interactive Vehicular Applications*, pages 72–79. ACM, 2010.

[23] G. Robles-De-La-Torre. The Importance of the Sense of Touch in Virtual and Real Environments. *Ieee Multimedia*, 13(3):24–30, 2006.

[24] Senseg. The Senseg Tixel. `http://senseg.com/technology/senseg-technology`, 2012-02-26.

[25] Steadfast Technologies. Pneumatic Tactile Display Components. `http://www.tactileresearch.org/rcholewi/files/Pneumatic.pdf`, 2012-02-26.

[26] UNIVERSITY OF TECHNOLOGY, SYDNEY. Voice Coil Motors. `http://services.eng.uts.edu.au/cempe/subjects_JGZ/ems/ems\ch11_ppt.pdf`, 2012-02-26.