Thesis in Master's Degree Programme in
Industrial and Strategic Design

SHOW ME WHAT YOU MEAN

Gestures and drawings on physical objects as means for remote collaboration and guidance

Andrei Savolainen

Aalto University | School of Art and Design

2011

Abstract

This thesis presents findings based on the study of remote projected interaction and guidance on physical objects. First, the results are based on the study of literature and previous research in the fields of ubiqutious computing and environment, augmented reality, remote collaboration and guidance. Second, the results are based on findings through testing projector technology in remote interaction and guidance with users with the help of prototype.

Previous studies indicate that guidance on physical objects is seen as valuable and in such interaction, the focus should be shifted to the actual object. This thesis contributes to previous research and suggest better integration of hand gestures and drawings into remote collaboration and guidance.

Projected interaction model, described in this thesis, enhances the feeling of togetherness between remote users (expert and novice), and provides critical help in conversational grounding in remote collaboration and guidance with physical objects.

Keywords

Pico projector (also handheld projector, micro projector, mobile projector), remote interaction, remote collaboration, remote guidance, augmented reality, contextual sensitivity, context-aware, ubiquitous computing, ubiquitous environment, gestures in remote collaboration and guidance.

Acknowledgements

I would like to express my appreciation to my tutor, professor Turkka Keinonen from Aalto University and Lutz Abe, a supervisor from Nokia. Thanks for the valuable and constructive feedback and advises during this process. Also thanks goes to Jussi Mikkonen, Laboratory Manager at Aalto University for the technical expertise with pico projectors and valuable insights.

I would also like to express my gratitude to my family and friends, and most of all to my girlfriend - Taru for the unconditional support, patience and feedback. This would not have been possible without you.

Table of contents

Table of contents

Chapter 1:

“A picture is worth a thousand words”
(unknown)

Introduction

Pico projector technology has been on the market for years where display technology is the dominating technology. While pico projector technology is constantly being developed and improved, it still remains to be a niche product. This thesis tackles this problem by bringing forward advantages of pico projector technology in remote interaction, collaboration and guidance with physical objects.

Real-time remote interaction and collaboration is based on voice and video conferencing and methodologies lack the conventions of physicality. "In such approaches interactions are limited to visual and auditory media, and shared environments are confined to digital world." [1] While projection lacks the ability to be touched, it can be used in remote collaboration and guidance on physical objects. As previous research suggests, the focus in such interaction should be on the object, rather the facial expressions of participants. Remote collaboration with the physical world and with physical objects needs different approaches than what simple video conferencing can offer. While people use digital environments for a number of activities, the physical environment should not be forgotten. People use everyday appliances and objects, experience problems with them and seek for help to fix these problems (e.g. by calling to a service). This thesis focuses to on addressing this problem, by providing a projected interaction model with gesture support to be used in remote guidance.

This paper presents a novel approach to enhance remote collaboration and guidance especially with the help of pico projecting technology.

Background

This thesis was done in collaboration with Nokia. I have been working at Nokia since 2007 as an User Interface Graphic Designer and I was excited when I found out that there was an opportunity for thesis collaboration with pico projector technology as a theme. After initial discussions and mutual agreement that pico projecting technology provides a good ground for research in Master thesis - I begun to explore possibilities to approach this technology.

The main goal at first was to find a suitable approach which would provide the "big picture", so that possibilities for this technology would be seen as a part of a larger context. Initial inspiration and a big spark for this thesis was found through ubiquitous computing and ubiquitous environment. These subjects provided a high level of theoretical ground for this thesis.

Theories and concepts behind ubiquitous environment provide approaches to the smarter use of technology as an enabler for more natural and meaningful interaction between the human and the environment.

Remote interaction and collaboration has been a subject for research for years, where the most interesting results and concepts in terms of this thesis are related to tangible interfaces - decreasing the gap between digital and physical or merging these two.

This Thesis emphasizes tridimensionality of the physical environment. The main goal is to present the physical environment with its objects as a a layer on top of which the interaction is built on, in a similar way as augmented reality does. Remote interaction model as it is presented in this paper is built upon pico projecting technology and it takes place in a physical environment where it augments physical objects with digital content.

This model provides a new way to participate and guide remotely in physical environment and with physical objects. While co-located collaboration is by far most effective, the model presented in this thesis provides an immersive approach to make the remote collaboration and guidance less 'remote'.

Framing the topic

Pico projecting technology in ubiquitous environment

I believe that pico projecting technology has significant potential but that potential can be found only by taking a fresh look at this technology with keeping in mind 'the big picture'. In this thesis, this means looking at this technology as a part of a ubiquitous environment. The 'vision of ubiquitous environment' [2]provides wireframe for better understanding of relationships between user, environment and the role of technology.

Remote (tangible) interaction and guidance

Current systems for real-time interaction and collaboration contribute to great extent to voice and video conferencing alone and provide access to shared digital environments. Such interfaces fail to address our sense of touch and physicality and gesture based interaction. Our life in physical environment is based on haptic senses, we manipulate objects directly, by moving and touching objects and use pointing and other gestures to help out in communicating with others, when we collaborate locally on shared physical object.

Local collaboration, taking place in single, real-world over a physical object relies not only on direct manipulation of an object by all users, but also on various sets of gestures. When students are working in a group task involving constructing a mechanical structure for instance, not all of them are working simultaneously. The process requires for some of them to step back, observe, give feedback along direct manipulation of pieces of mechanical structure. Giving feedback is not only based on verbal input, but it is also supported by hand gestures: "take this", "put it there", "replace this piece by that".

When collaboration takes place remotely, current systems are not supporting hand

gestures, which is after all seen a natural part of human interaction.

This thesis addresses physicality of remote interaction by providing a platform, a layer to place gesture-based information on. This platform provides additional value in interaction where the person with expertise 'the expert' guides a person with lesser knowledge 'the novice', providing clues and instructions for the task or a problem related to physical object. While the expert does not have direct access to the object (not able to touch) and is dependent on the actions of novice, however, he is able to feel less remote because of possibility for enhanced communication of his intentions toward the physical object.

The goal of this thesis is therefore to prove that pico projecting technology enhances remote interaction and collaboration with physical objects and in physical environment.

Theoretical wireframe

The Vision of Ubiquitous Environment with three technological development paths: Ubiquitous Computing, Advanced Interaction and Algorithmic Intelligence with its sub-categories provides a theoretical wireframe for evaluating pico-projecting technology from Ubiquitous Environment perspective.

The focus within this thesis will be on three sub-categories from the Vision of Ubiquitous Environment: Mobile Technology, Natural Interaction and Contextual Sensitivity.

Chapter 2:

Ubiquitous computing
Ubiquitous environment

“The most profound technologies are those that disappear.”
Mark Weiser 1991

Ubiquitous computing

Ubiquitous computing or ubicomp is based on the idea of "integrating computers seamlessly into the world"[4]or information which is integrated everyday objects and activities. Simply put, the idea behind is to take away centralized role from the desktop computer which it had at the moment when ubicomp was born and rather to support our activities no matter where we are with computational power which is everywhere and thus always available to us. This is a computing and technology which: "does not live on a personal device of any sort, but is in the woodwork everywhere" [4], and is "essentially invisible to the user" [3]. More formally Ubiquitous computing is defined as "machines that fit the human environment instead of forcing humans to enter theirs." [5]

A desktop computer is traditional example of paradigm change in ubicomp. Taking a look back at how people used personal computers years ago was very different than today. Desktop computer had a dedicated place for itself, the place which was reserved only for the operation of computer and usually nothing else. The arrival of laptops changed this paradigm as people became able to use computer for anything from work to entertaining themselves from any place they could have a sit on: sofas, floors, cafes, waiting rooms and airport lounges - any public spaces. The emergence of smartphones changed the setting once again as these "computers" provided even higher mobility – now we could use these devices whenever and wherever. While such a natural shift, this is only an example of how technology changes our lives. Instant access to the devices, and through them the access to networks, appliances, things, people, information has become a natural thing, almost a necessity. Such paradigm change is also taking place with others, "smart" technologies as well.

Our everyday life is saturated with sensors, computational power and wireless information network – all those tiny bits which make our life a little bit easier in one way or another. These sensors switch on lights, open doors for us, keep our homes at optimal temperature and wirelessly connect our devices with one another. This technology is ambient and it is all around us already. And as many evangelists of ubiquitous computing suggest – there will be few places in the future which will be left out of networked information processing. Ubiquitous computing is therefore "nothing less than the colonization of everyday life by information technology." [6]

What ubiquitous computing promises is new way to use information technology and predicts changes in design paradigms and approaches which are required in designing complex systems and environments. From interaction design point of view, ubicomp predicts changes where it will be more important **what** we want to do rather **what tools** we use to achieve those goals. Ubiquitous computing is not there only to "be used instrumentally to achieve a discrete task", instead "it's simply **is** in a way that personal computing is not, and that quality necessarily evokes an entirely different kind of experience on the part of those encountering it". [6]

Ubiquitous computing challenges how we interact with everyday objects and it challenges how we interact with the environment we live in. It questions the role of the technology in the interaction itself and proposes on shifting the focus onto what

is really important – actions and interaction. “the focus upon humans detracts from support for the support for the activities themselves (..) second, too much attention to the needs of the users can lead to a lack of cohesion and added complexity in the design.” [7]

This paradigm change affects the design for ubiquitous computing as well, as “it is a radically new situation that will require the development over time of a doctrine and a body of standards and conventions-starting with the interfaces through which we address it.” [6]

Ubiquitous Environment

" There are many ubiquitous computings." (Adam Greenfield)

In previous chapter ubiquitous computing was described as something not restricted to a single object or a room. Instead there are many ubiquitous computings in our environment and they can work on different levels. This chapter will be concentrating on what makes the environment ubiquitous – where ubiquitous computing plays a part of bigger picture an environment - and what are those characteristics of this environment.

Ubiquitous environment represents a complete paradigm of many ubiquitous computings and it can appear in many forms and areas: "Wearable computing, augmented reality, locative media, near-field communication, body area networking". [6] While ubiquitous computings has many forms, it is "indistinguishable from the user's perspective and will appear to a user as aspects of single paradigm". [6]

Definition of ubiquitous environment

EU's Information Society Technologies Advisory Group touches the subject of Ubiquitous Environment in the report written in 2006 Shaping Europe's Future through ICT in the following way: "Building on and extending the ambient intelligence vision, technology developments are proceeding along well characterized paths. We note four main trajectories for this next generation of ICT. Systems and services that are:

1) Networked, mobile, seamless and scalable, offering the capability to be always best connected anytime, anywhere and to anything;

2) Embedded into the things of everyday life in a way that is either invisible to the user or brings new form-fitting solutions;

3) Intelligent and personalized, and therefore more centered on the user and their needs;

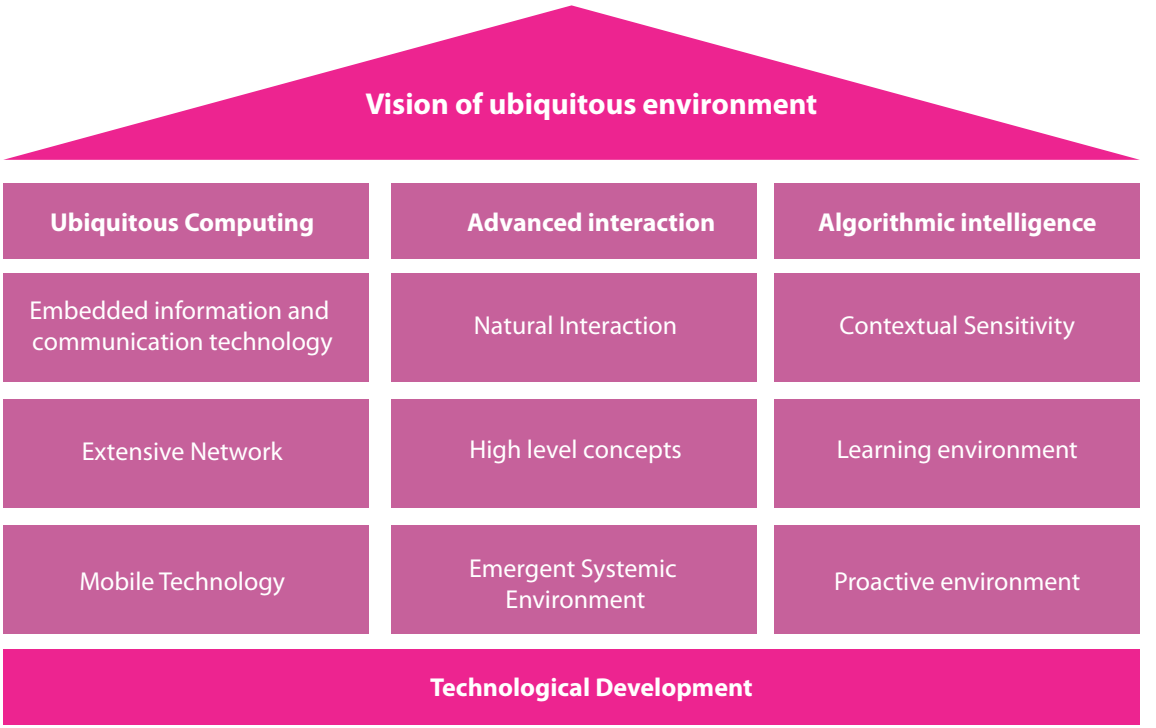
4) Rich in content and experiences and in visual and multimodal interaction." [8]

The vision of ubiquitous environment

"The vision of ubiquitous environment" [2] (figure 1) is built upon three technological development paths: Ubiquitous computing, advanced interaction and algorithmic intelligence.

Ubiquitous computing

Ubiquitous computing as described in previous chapter, promotes the information technology to merge in unnoticeable



* translation from Finnish to English by: Andrei Savolainen

Figure 1: The vision of ubiquitous environment requires three technological development paths (reference: 9): Ubiquitous computing, Advanced interaction and Algorithmic intelligence.

way into physical environment and into habitual products. This way, information technology spreads out from specific information-technological devices into our environment and becomes a natural part of functioning and living within this environment through the properties it enables. In such environment objects and parts of environment can communicate between each other and users locally and globally.

Advanced interaction

Advanced interaction represents factors which improve the interaction between user and informational technology. The interaction between the user and environment is more frequent and versatile. In such environment the the information technology appears to be invisible and the focus is on everyday objects – which offer wider possibilities for interaction.

In such environment the interaction between the user and technology happens on many other ways and levels than via buttons and displays. Natural interaction techniques develop further the interaction habits between the environment and the user. These interaction habits are applied from interaction and communication patterns between people. The interaction in ubiquitous environment broadens from

mere usage of the technology into shaping the technology. Therefore environments are being built gradually and the user has big role in configuring technological abilities and shaping those.

Algorithmic intelligence

Third path of algorithmic intelligence means that the environment recognizes the context of use and adapts accordingly. Because of this it is possible to leave larger 'content- and time-based entities' for technology to be responsible of. "Context sensitive environment anticipates personal and situation-dependent needs of user and responds to them. The environment can learn based on functions of its own, based on feedback from the users and based on the information of other partial systems or based on adopted examples. The environment can anticipate upcoming events and to prepare for those." (free translation from Finnish) [2]

Despite long period of time, the vision of ubiquitous environment is mostly unfulfilled in everyday products, appliances and applications. The development of ubiquitous environment is gradual and it is driven by three technological development paths described above. The first path, the ubiquitous computing exists already and is developing based on existing technology. Second path has lots of studies and results but yet without implementation in real life. The third path of the vision of ubiquitous environments is a target for research work and results remain unobtrusive.

Summing up:

Now we have described what technological development paths are on general level. Each of the paths has a sub-category which is important in terms of this thesis. These sub-categories are: mobile technology, natural interaction and contextual sensitivity and they will be described in detail next.

Mobile technology in ubiquitous environment

"The movement into the ubiquitous computing realm will integrate the advances from both mobile and pervasive computing." [9] Here ubiquitous computing is represented as an integration of two different terms which are often used interchangeably. What mobile computing underline is moving computing services with us, however the computing model itself "does not considerably change while we move". A significant limitation for mobile computing is its incapability to sense and obtain the environment or a context it is situated in and change the computing accordingly. And this is a factor in which per-

vasive computing is in theory at it's best. The idea of pervasive computing revolves around the computer's capability capability to obtain the information from the environment in which it is embedded and utilize it to dynamically."[9] Such environment is populated with sensing technology such as sensors, pads, and badges. In this model, mobile computing is presented as somewhat inferior to pervasive, environmental computing. This model suggests that mobile computing should sense and respond to the environment better than it does today. This is an important goal for development of mobile computing in the future. As mobile computing is one part of ubiquitous environment (also from the perspective of this thesis) it should shift into sensing and responding model in order to contribute to the whole development of ubiquitous environment.

Natural Interaction in physical environment

Natural interaction in terms of ubiquitous environment is described as an interaction where the user does not notice the information technology; instead he notices an interface or the offering – something which indicates that this is a gate for actions. A natural interaction in terms of ubiquitous environment is supported by the offerings of the environment i.e. in the same way as one sits on the log while collecting berries in woods. That is a completely natural thing to do. The log in that case represents an offering of the environment we move about. When we see that piece of log, we know, without giving it much of a thought – that we can use it as a place to rest. In that same way our environment provides same offerings, a door knob to open or close the door and the switch to turn the lights on.

In addition to the previous, natural interaction also means to "interact directly through physical artifacts rather than traditional graphical interfaces and interfaces operated by e.g. mouse." in environments with are supported by computational activity. "While in the everyday world we can manipulate many objects at once, using both hand and three dimensions to arrange the environment for our purposes and the activities at hand. A child playing with blocks engages with them in quite different ways than we could provide in a screen-based virtual equivalent; so tangible computing is exploring how to get the computer "out of the way" and provide people with a much more direct-tangible-interaction experience. [10] The direct-tangible approach or direct manipulation is especially valuable in terms of learning how we deal with real world and physical objects and how we can transfer this information into the developing of future digital or digitalized interaction. While digital world can function as enabler for great things and new interaction models, natural, physical environment and interaction with physical objects will always be an inspiration to build upon.

Context-sensitive interaction

Context-aware computing was first discussed by Schilit and Theimer in 1994 to be software that "adapts according to its location of use, the collection of nearby people and objects, as well as changes to those objects over time." [11] Context-aware can also be referred to with synonyms such as adaptive, reactive, responsive, situated, context-sensitive and environment-directed. Main two catego-

ries of context-aware are: ones which are using context and others which are adapting to context. A wider definition of a system as context-aware is if the system uses "context to provide relevant information and/or services to the user, where relevancy depends on the user's task." [12]

Especially from this thesis point of view, where mobile pico projector technology is used is essential to underline the mobility in context-aware computing. Mobile distributed computing faces challenge in exploiting changing environment "with a new class of applications that are aware of the context in which they are run." [13] In order to be fully context-aware, mobile technology needs to adapt according to "the location of use, the collection of nearby people, hosts, and accessible devices, as well as to changes to such things over time. A system with these capabilities can examine the computing environment and react to changes to the environment". [13]

In terms of this thesis and the vision of ubiquitous environment, context-aware interaction means to offer most relevant information to users and build the interaction upon this information.

Challenges of ubiquitous environment

Victoria Bellotti's (at PARC)paper on "rarely have to worry about questions of the following sort" when designers design conventional systems:

When I address a system, how does it know I am addressing it?

When I ask a system to do something how do I know it is attending?

When I issue a command (such as save, execute or delete), how does the system know what it relates to?

How do I know the system understands my command and is correctly executing my intended action?

How do I recover from mistakes?

Ubiquitous environment is complex and multi-faceted subject and it meets various challenges of future. As mentioned previously, the vision of ubiquitous environment on everyday product, appliance and application level is mostly unfulfilled. This thesis provides an approach to remote collaboration and through this approach contributes to the vision of ubiquitous environment. Following claims summarize what is considered important in the vision of ubiquitous environment.

Summing up:

We have described ubiquitous environment both in general and through the vision of ubiquitous environment, where we have chosen sub-categories (mobile technology, natural interaction and contextual sensitivity) to help to approach pico projector technology in this thesis. Next, we will define and summarize what is considered important in this thesis from ubiquitous environment approach.

Claim 1: Ubiquitous environment is problematic because it is hard to see literally. Therefore concrete examples (prototypes and applications) are needed.

Claim 2: Mobile computing “does not considerably change while we move” and contextual sensitivity of mobile computing needs to be enhanced.

Claim 3: Ubiquitous environment should enhance natural interaction by providing users with possibility to interact with the environment when user needs it (contextually) but in a subtle way at the same time.

Chapter 3:

Pico projecting technology as enabler for augmenting environment

Remote collaboration

Pico projecting technology as enabler for augmenting environment

Figure 2: Pico projector front angle (Microvision)

Figure 3: Pico projector connected to a phone (Microvision)

Pico Projector

A Pico projector (also known as a handheld, pocket, micro, mobile projector micro or mobile projector) is a technology that applies the use of an image projector in a handheld device. Pico projector has same front projection ability as any standalone projector where the image is projected onto the screen or wall from the front. This technology has become a response to the emergence of compact portable devices such as mobile phones, digital cameras and personal digital assistants, which have sufficient storage capacity to handle presentation materials but little space to accommodate an attached display screen.

There are two main projector technologies currently on the market, LCD (Liquid Crystal Display) and DLP (Digital Light Processing). DLP projector is used in a prototype of thesis and therefore will be briefly described next.

"DLP technology is based on an optical semiconductor, called a Digital Micromirror Device (DMD), which uses mirrors made of aluminum to reflect light to make the picture. The DMD is often referred to as the DLP chip. The chip can be held in the palm of your hand, yet it can contain more than 2 million mirrors each, measuring less than one-fifth the width of a human hair. The mirrors are laid out in a matrix, much like a photo mosaic, with each mirror representing one pixel. The number of mirrors corresponds to the resolution of the screen. DLP 1080p technology delivers more than 2 million pixels for true 1920x1080p resolution, the highest available." [14]

The system of pico projector "comprises five main parts: the battery, the electronics, light source, the combiner optic, and

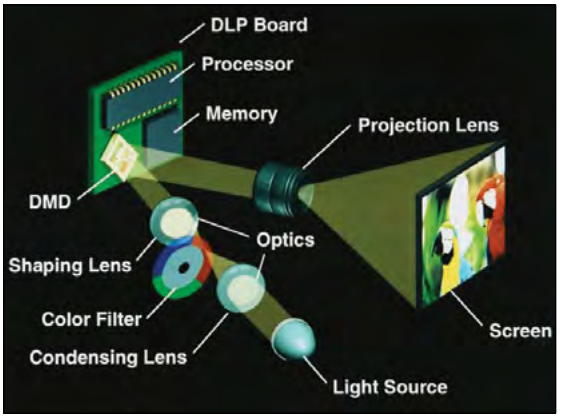


Figure 4: DLP projection technology

Figure 5: Presenting in office (Microvision)

Figure 7: Spontaneous presentation (Microvision)

the scanning mirrors. First, the electronics system turns the image into an electronic signal. Next the electronic signals drive laser light sources with different colors and intensities down different paths. In the combiner optic the different light paths are combined into one path demonstrating a palette of colors. Finally, the mirrors copy the image pixel-by-pixel and can then project the image. This entire system is compacted into one very tiny chip. An important design characteristic of a handheld projector is the ability to project a clear image, regardless of the physical characteristics of the viewing surface.

Figure 6: The view of the Terminator (Science fiction movie)

Augmented Reality

Most of us are familiar with science fiction movies such as “The Terminator”(1984), where a character sees real world with graphical information placed on top of it (augmenting the world), in relation to objects, environment and people in that environment.

Augmented reality is described to be derived from virtual and, where virtual reality means replacing real environment with a simulated one. Virtual and “mixed reality” [15] have been successfully been used in simulated training (e.g. military) for years, and the challenge has been to move away from laboratory-like (and synthetic) environment to real physical environments and to move interaction onto real physical objects and spaces. Such transition is seen to help to use our natural habitat in more useful way; “Wherever possible, we should look for ways of associating electronic information with physical objects in our environment. This means that our in-

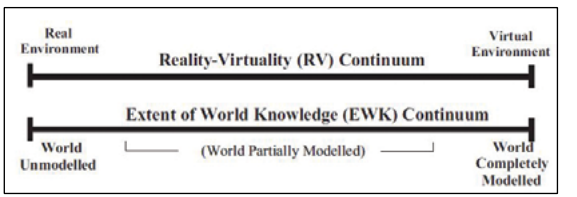


Figure 8: Reality-Virtuality (RV) Continuum, in parallel with Extent of World Knowledge (EWK) Continuum.

formation spaces will be 3D. (..) Our goal is to go a step further by ground and situating the information in a physical context to provide additional understanding of the organization of the space and to improve user orientation." [16]

Narrow interpretation of augmented reality is where "the viewer observes a direct "see-through" view of the real world (e.g. head-up-display), either optically or via video coupling, upon which is superimposed computer generated graphics". [17] Another broader definition refers to "any case in which an otherwise real environment is 'augmented' by means of virtual (computer graphic) objects."[15] or even broader definition: "any mixture of real and virtual environments." [17]

Augmented reality also emphasizes the difference in the interaction with synthetic, such as virtual reality (VR) and real world. "VE technologies completely immerse a user inside a synthetic environment. While immersed, the user cannot see the real world around him. In contrast, AR allows the user to see the real world, with virtual objects superimposed upon or composited with the real world. Therefore, AR supplements reality, rather than completely replacing it. Ideally, it would appear to the user that the virtual and real objects coexisted in the same space." [18]

Figure 9: Augmented reality application on iPhone

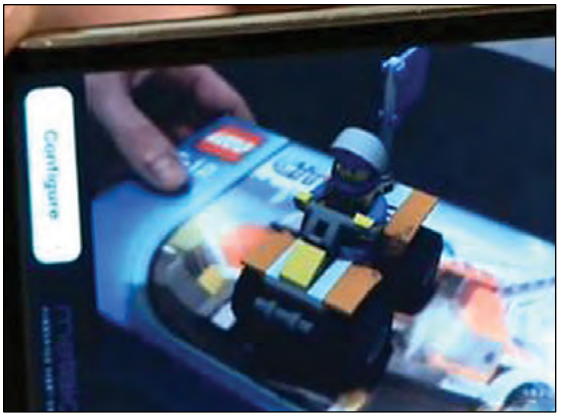


Figure 10: Augmented digital image of lego on top of real lego box, when seen through mobile display (Metaio)

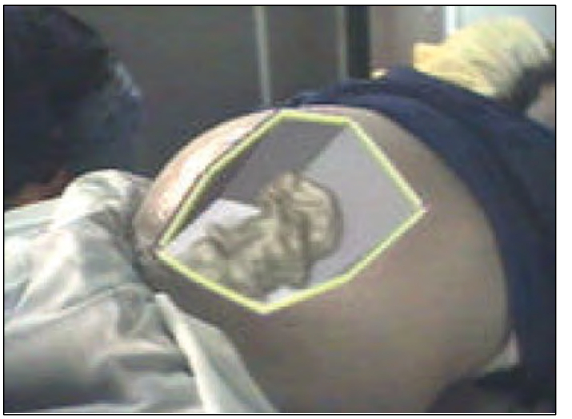


Figure 11: Augmented reality provides interesting prospects also in medical field [18]

Definition of AR

1) Combines real and virtual
2) Interactive in real time

3) Registered in 3-D" [18]

Augmented reality is an exciting issue in terms of interaction between user and ubiquitous environment. After all, as our environment, enriched by context sensitive content, will allow us to use this environment in completely novel ways.

The technology used in augmented reality systems has been mostly display-driven. From this perspective, projecting technology plays an interesting and exciting role providing new ways to build interaction between user and physical world in this thesis.

Augmented reality and projectors

Projector technology in terms of augmenting our natural habitat - physical environment has been widely studied as the projector technology provides interesting platform for interaction with physical surroundings in a completely different way in comparison to display or a head mounted display technology.

"Many researchers have seen the possibility of augmenting objects using combinations of video projectors and video cameras."[19] The studies have however been concentrating on fixed setups with where augmented working area has usually been a desk, with objects such as papers or books are augmented with digital information and where display (or monitor) technology has been used in interaction with the content. Such setup overcomes an important flaw of display in conventional computer systems: "the fact that a monitor can not augment real objects, without using HMDs. Yet, all of these setups are fixed ones. The projectors (as well as cameras) are fixed and calibrated to the desks they project onto. Accordingly the objects (papers) to be augmented have to be brought onto these desks." [19]

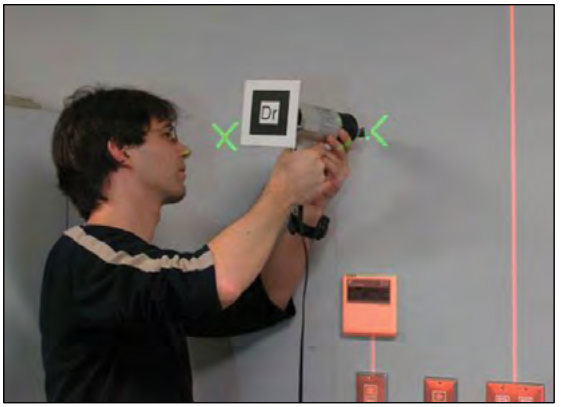


Figure 13: Drill-Guide: This example shows the user where to drill holes and where it's dangerous because of cables [19].

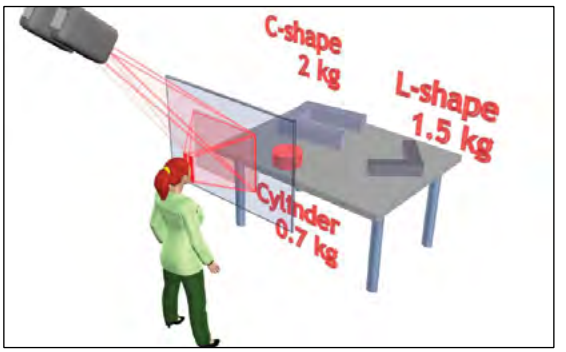


Figure 12: The principle of the display in ASTOR system used for our autosteroscopic optical see-through augmented reality system [21].

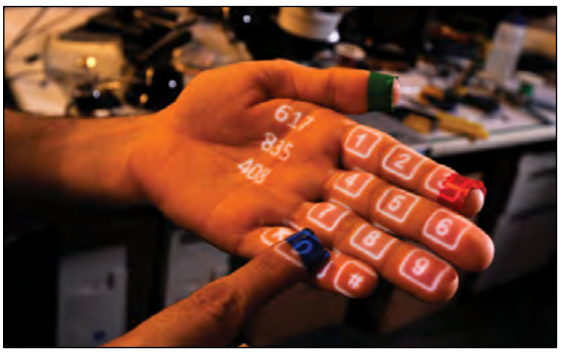


Figure 14: Sixth sense [22].

Different attempts have been made to break away from fixed setup, related to the use of projector technology in interaction with physical world. The advantage of "ASTOR" system [21], while it is not pure projector driven setup, is that the content and images changes "depending on the point of view." [21] This setup as well as setup with multiple rotatable video projectors [19] shows that it is important to provide the user the ability to move in the environment and ability to interact with the environment from various, rather than from fixed perspective.

Projector technology can also be used as a wearable device. "SixthSense" is one example of using projector technology in combination with video camera and this combination allows user to take this devices anywhere: "SixthSense bridges this gap by augmenting the physical world with digital information, bringing intangible information into the tangible world. Using a projector and camera worn as a pendant around the neck, SixthSense sees what you see and visually augments surfaces or objects with which you interact."[22]

Summing up:

Previous research shows that projected interaction gives novel possibilities in terms of interaction with physical environment we live in and with the objects, situated in this environment. Next we will summarize main points of the relation between augmented reality and the projector technology through following claims:

Claim: Pico projector technology contributes to new ways in interaction with physical environment.

Claim: Pico projector technology is advantageous in terms of mobility.

Claim: Projected augmented reality provides new possibilities for providing real time information about the physical environment in forms of instructions and cues.

Remote collaboration

Remote collaboration is an ambiguous subject and one way to approach it is through the CSCW (Computer supported cooperative system) Matrix first introduced in 1988 by Johansen. The CSCW itself can be explained as follows: "CSCW addresses "how collaborative activities and their coordination can be supported by means of computer systems." [23]
CSCW matrix explains remote collaboration through the context of system's use, which considers two dimension: first, whether collaboration is co-located or geographically distributed, and second, whether individuals collaborate synchronously (same time) or asynchronously (not depending on others to be around at the same time). Our focus is remote interaction, therefore it is distributed but takes place at the same time.

Remote collaboration based on previous research is multifaceted subject. Next, remote collaboration will be described in four different parts: (1) previous research related to video-conferencing, (2) remote collaboration with physical objects, (3) remote expertise and guidance and the (4) use of hand gestures and drawings.

Video conferencing

Technologies that provide possibility for remote collaboration have been available for years. Many commercial products and applications for the use in organizations have introduced sophisticated video conferencing technologies (e.g., HP Halo, Cisco Telepresence, Tandberg). These technologies "led to enhancements in the collaborative user experience over traditional video conferencing technologies."

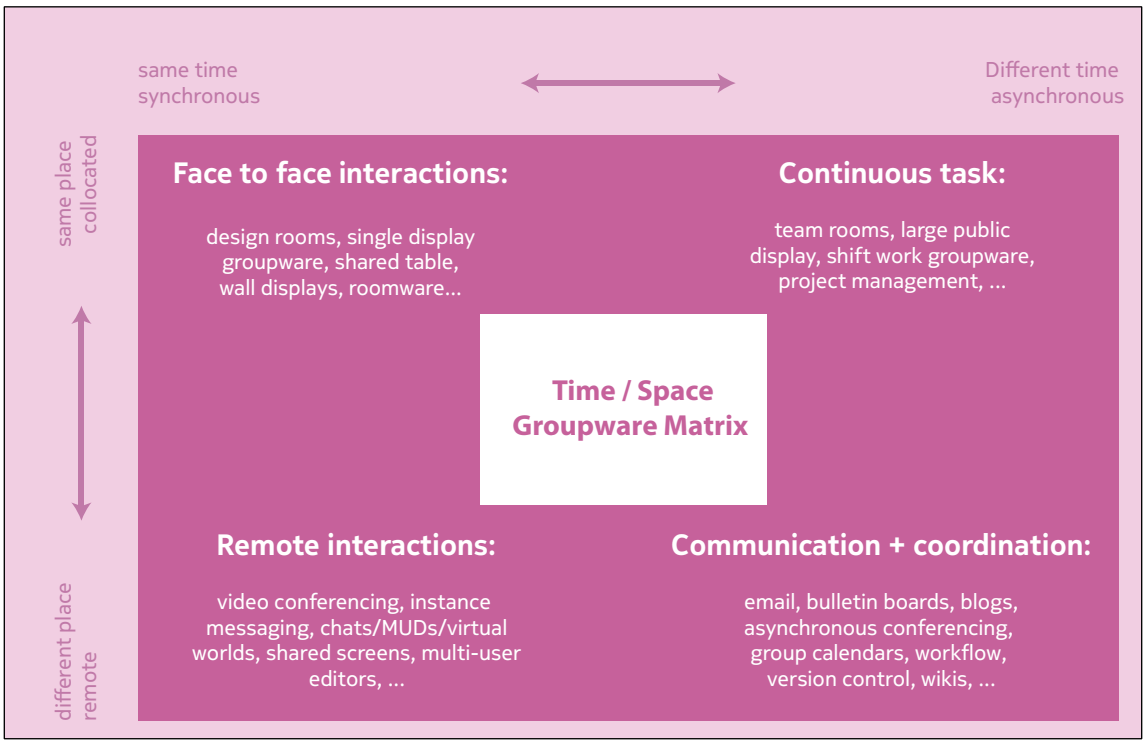


Figure 15: CSCW Matrix

[24] Biggest advantage in such systems is “appropriate shared spatial geometries”, where “distributed spaces become ‘blended’”. [24] This means that spatial geometries continue across distributed boundary from local to remote space in a coherent way, “providing the illusion of a single unified space.” Such setup supports small meetings where the focus of the meeting is conversation with nonverbal communication aspects and the conversation would is too important to be held on mere voice-only (telephone) call.

While video conferencing technologies are important in ‘face-to-face’ video type communication, they fall short in tasks related to physical objects, where the focus is not on the faces of users but instead on the objects at hand.

Remote collaboration on physical objects

Video conferencing tools are not directly applicable in situations where the focus is not on people but on physical objects. This is because video telephony system is using “talking heads model, in which the cameras broadcast pictures of the people in conversation rather than the task they are working on.” [25]

Interacting with real-world, where touch, and physical manipulation play a key role in understanding this environment needs different approach. “All of these actions...” (enabled by computers) “...are abstract and arbitrary compared to the real, physical manipulation of objects, which is where the power of real and perceived affordances lies.” [26]

Remote collaboration with physical or tangible objects can be explained in simplified form through “Interface Techniques for HCI and CSCW.” (figure 16) The focus of this thesis is on physical - tangible interface techniques, as they are different from traditional Graphical User Interface techniques. The difference between GUI and TUI is not only in the physical characteristics of the TUI, but also in the fact that when people interact with physical objects, the object becomes the focus and facial impressions - so important in video conferencing - become less important. “When collaborants were working on a shared object, they spent most of their time looking at the video feed of that object rather than at each other’s faces or the wider context. [27] Nardi et al. found that “nurses monitored video feeds of surgeons operating procedures to anticipate what instruments and supplies they would need next, reducing the need for explicit communication.” [28]

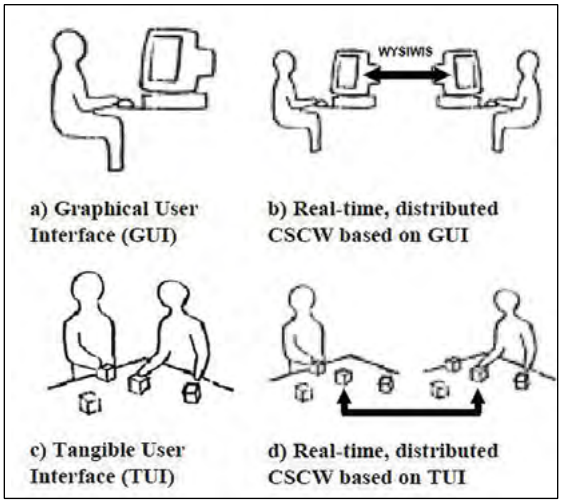


Figure 16: Interface Techniques for HCI and CSCW.

Fussell et al. in the research conducted on “cooperative work—collaborative repair of complex devices”, tested how experts and

novices used visual information remotely in bicycle repairment test. (figure 17) "Results demonstrate the value of a shared visual work space, but raise questions about the adequacy of current video communication technology for implementing it". [25]

Previous research shows that technology as it is an important and enabling part of remote interaction and collaboration with physical tasks, needs further refinement and testing. A set of suggestions and recommendations has been brought forward by Fussell to address the "design of future video-based systems to support collaborative remote repair:

Provide workers with better feedback on what is in the helper's field of view, to clarify what is in the shared visual space.

Provide helpers with a wider field of view, thereby increasing the shared visual space.

Provide helpers with feedback on the worker's attentional focus.

Provide support for helper gestures within the shared visual space." [25]

Figure 17: Worker wearing collaborative system.

Remote expertise and guidance

Robert E. Kraut et al., conducted an "empirical study of people using mobile collaborative systems to support maintenance tasks on a bicycle" with the results showing "that field workers make repairs more quickly and accurately when they have a remote expert helping them." [29] At the same time, the results of remote expertise are found valuable, but they can hardly be compared to co-located interaction which has been seen as optimal and most natural situation. This rises specific need for further development of such systems in the future: "designers should be better able to develop systems that meet the needs of collaborative workers." [25]

As previous reseach implies, simply linking remote spaces through audio-visual video links (as opposed to audio-only) does not improve collaborative performance (..) this has potentially then been the motivation for the development of remote gesturing systems." [30] And that to "facilitate performance in object-focussed tasks a representation of gesture between the spaces should be provided." [31]

Use of gestures and drawings in remote guidance and help

Empirical studies on remote collaboration on physical tasks suggest that people's speech and gestures are intricately related [33]. People according to Wickey[32] point at object of interest with finger to describe ("take this one"), to show where it needs to go ("this goes here"), to describe assembly tasks or procedures through motion ("turn this bit this way") and to convey

Type of Gesture	Definitions	Possible Functions
Deictic (Pointing)	Orienting a finger or hand toward a point in the environment	Reference to object and locations
Concrete repre-sentational Iconic representa-tions	Forming a hands to show what a piece looks like, or to show how two pieces should be positioned relative to one another	Reference to objects, procedural instructions (particularly orientation), descriptions of task status
Spatial/Distance	Indicating through use of one or both hands how far apart two objects should be or how far to move a given object	Procedural instructions, descriptions of task status
Kinetic/Motion	Demonstrating through use of hands what action should be performed on a task object	Procedural instructions

Figure 18: Definitions and possible functions of gestures used in collaborative physical tasks. [33]

a form of the outcome of the assembly task ("this should look like 'L' shape"). In addition gestures can be used to monitor understanding (ok gesture or a thumb up).

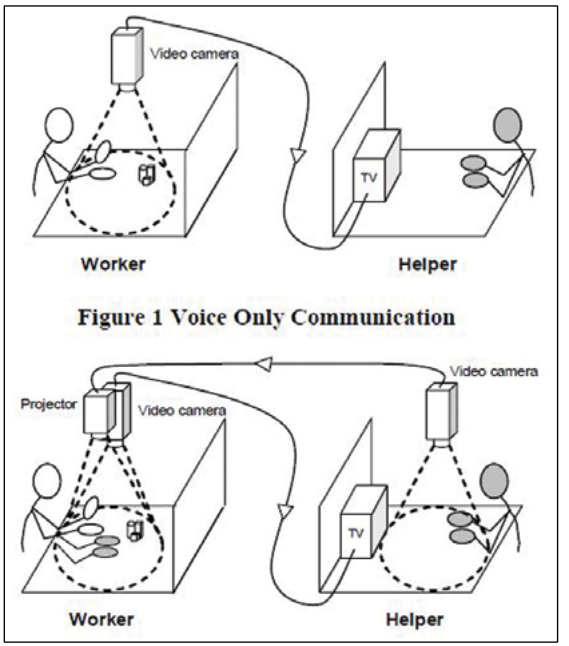


Figure 19: Voice + Projected Hands (Kirk et al)

Remote gestures help in establishing conversational grounding, i.e. "establishing mutual knowledge, belief, attitudes and expectations." [30] Therefore supporting gestures in remote collaboration on physical objects is not only essential and natural, but it provides collaborating users in the establishing of the common ground in conversation e.g.: "this is the object we are talking about". Gestures according to Fussell (figure 18) can be categorized as follows; simple (1) deictic gestures, such as pointing, iconic representation (2), via hand forming, showing the spatial distance (3) by indicating distance with hands and to show an showing the action through kinetic/motion (4) with hands.

Wickey's experiments, based on prior work by Kirk) show that user prefers to point at the screen (figure 21) (instead of workspace) even though that gesture was not transmitted to the other end. "This suggests that it was more intuitive to point

to the screen where the objects were displayed rather than down at the workspace where no objects were displayed or physically present." [32] This suggests the importance of linking gestures, physical objects or representations of physical objects.

Wickey [32]presents design recommendations for remote interaction with gestures:

1. Remote gesture systems should support easy composition of gestures

2. Remote gesture systems should integrate visual control functions

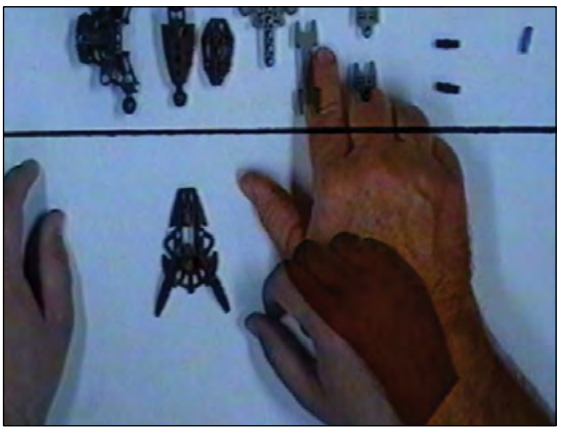


Figure 20: Helper's hand (top right) pointing and tapping at a piece

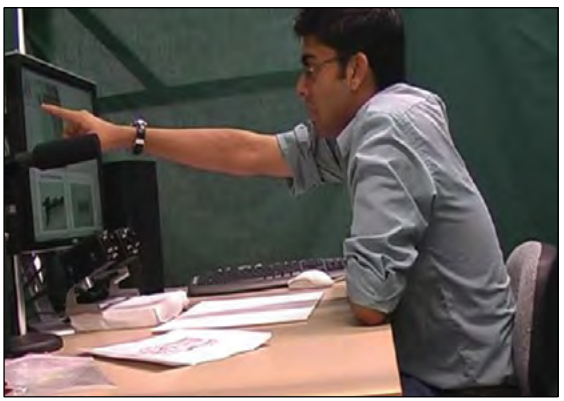


Figure 21: Helper's hand pointing and tapping at a piece on the screen, instead of the workspace (table)

3. Gesture input in remote gesture systems should be located where the objects are displayed

4. Gesture output in remote gesture systems should be semi-transparent"

Summing up:

Research done in the field of remote collaboration shows that technology is both a limiting and enabling factor in terms of interaction. In addition to this it is apparent that technological solutions are highly contextual - there's no single technology or interaction model for every situation.

As interaction with physical objects is the focus of this thesis, it needs to be emphasized that remote collaboration with physical objects is very different than simple video-conferencing as the focus shifts from face into objects. Remote collaboration with physical objects can be enhanced with the use of gestures or drawings.

Each of four parts of remote collaboration described in prior is alone fascinating in terms of this thesis, even though some more than others. Next, each of the part will be summarized through the claim. The goal of this is to bring forward how each of the part is considered valuable in this thesis.

Claim 1: Current remote application products and applications are focused on nonverbal communication in fixed space. It can be considered rewarding to focus on mobility and what happens outside of one fixed space.

Claim 2: Remote collaboration around physical object draws attention to the object, rather than to the face(s) of people involved in collaboration. Focusing on objects rather than faces can be considered advantageous in specific tasks.

Claim 3: Remote expert help make people perform more quickly and accurately in specific tasks and can bring great value in interaction with physical objects.

Claim 4: Prior research in remote guidance shows that the use of hand gestures and drawings provides value through conversational grounding during remote collaboration with physical objects but it should be applied directly on physical objects.

Chapter 4:

Research questions
Research methods

Defining research questions

The vision of ubiquitous environment

In previous chapters we have described what ubiquitous environment is. We also took a closer look at some of the categories which are especially interesting in terms of this thesis (mobile technology, natural interaction and contextual sensitivity). As previous research shows, concrete examples are needed to fulfill the vision of this environment. Previous research also shows that mobile computing "does not considerably change while we move" and therefore needs to support contextual sensitivity. These issues, related to mobile computing and contextual sensitivity will form the first research question.

Projected augmented reality in physical environment

We also took a look at concepts such as augmented reality situated in real world or physical environment. Augmented reality offers fascinating approach when it comes to altering the real-world we live in by the means of digital or computer-generated sensory. Previous research, related to the use of projector technology as a way to augment physical environment with its objects shows that projector technology can be used for augmenting physical environment we live in. In addition, augmented reality is context sensitive, when it uses tags, markers or beacons situated in the environment.

Pico-projectors or projecting in general can offer a unique approach in terms of absence of displays (traditional AR) in between the viewer and the object and also a different approach as the content can be projected exactly on the surface of physical objects or environment. Previous studies also show that projected augmentation is used mostly in fixed, lab-like environments. Pico projector techonology could therefore provide more flexible and novel approaches in exploring physical environments and to be used in interaction with physical objects. These findings will help in definition the first part of second research question.

Remote collaboration and guidance with physical objects

The second part of the second research question concentrates on the use of gestures as means of remote expertise. Prior research in remote guidance shows that remote expertise with physical tasks (objects) makes people perform more quickly and accurately. As previous research also shows, the representation of physical object in expert's (helper's) view in remote guidance is mostly separated from the working area (computer display to observe object and table to be used as actual working area for gesture output) (figure 21). This thesis addresses this problem by combining the representation of the object and working area with the help of projector technology in a cohesive way and allows more direct interaction with the object.

Also, as such combination of real and projected image could result with unexpected outcomes, such concern defines the last research question.

Therefore, with the help of this thesis in, I hope to address remote guidance with physical objects by enhancing the interaction familiar from our everyday lives and by making remote guidance between physical environments less remote and more natural.

Research Questions

1.Do the interactions enabled by pico projectors contribute to the vision of ubiquitous environment ?

2.How gestures and drawings are used when matched with physical object in remote collaboration and guidance?

3.How could a projected image be matched with reality in remote guidance?

Research methods

1. Observations during user tests

2. Users' verbal and written comments (post-it notes) based on review of observation videos

3. Evaluation of observation videos by designer group and arranging post-it notes according to affinities.

4. Evaluation of gestures and drawings used in remote guidance.

Chapter 5:

Experimental setup

Prototype

Test objectives, goals and structure

Task description

Stakeholder description

User tests resuls

Prototype

The technology and preparing it for tests was a pivot - an important part of the project. The prototype used was a combination of off-the-shelf electronics, such as web cameras, pico-projectors, pc's and parts which were purchased from hardware store to build the stands. First tests were made at the Aalto University, School of Art and Design with the help of Jussi Mikkonen – Laboratory Manager with the aim to find out how a web camera and pico-projector should be combined and how the technology works. After those initial tests, the projectors and web cameras which were used in actual tests, were purchased and two stands were constructed for the equipment to be attached to.

Figure 22: Pico Projector, used in tests (Aiptek Pocket Cinema V50

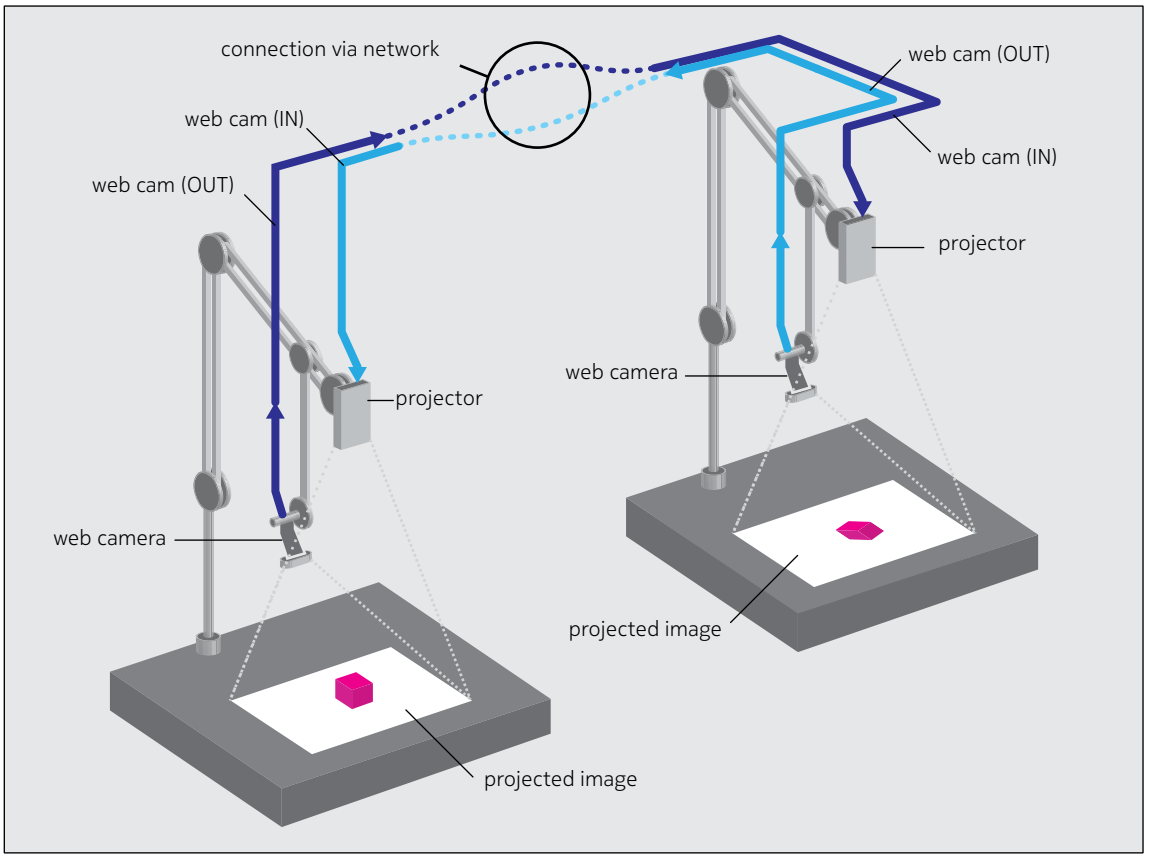


Figure 23: The image from local projected interaction set (left), and the cube within the projected area is sent to distributed projected interaction set (right). The image from the right is then sent back to left.

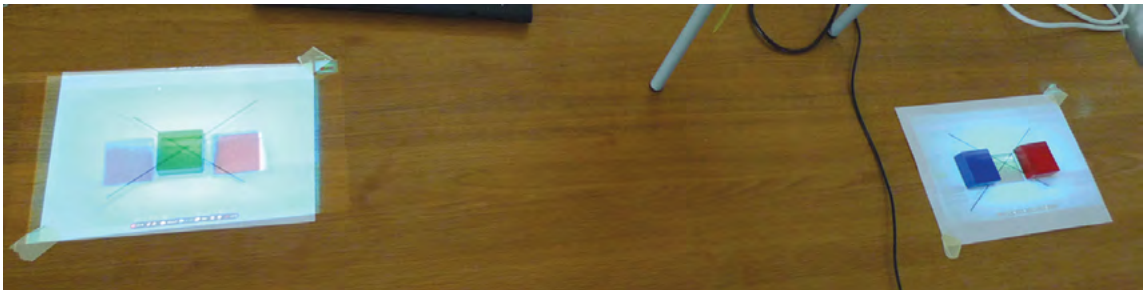


Figure 24: Testing and calibrating the equipment. The green block is physically present only on the left and blue and red blocks are physically present only on the right. The graphical representations of objects are shown next to real physical objects.

Figure 25: Testing and calibrating equipment in one of two rooms.

Aiptek Pico Projector	x 2
Logitech Web Camera	x 2
Laptops	x 2
Logitech VidHD software with Web camera controller software	
Stands	x 2
Video cameras for observation	x 2
Tripods	x 2

Test objectives, goals and structure

The first objective of user tests was to study projected interactive system especially from a physical, tridimensional point of view. The goal was to see how users would interact with physical objects in general and how they would use gestures and drawings in interaction.

The second objective was to study remote guidance on physical objects, with attention to different roles of users, who were working in pairs. Users had four different tasks to accomplish. During the first task users had equal roles, and following tasks required division into expert and novice roles.

Roles

Division into expert and novice roles means diffirent responsibilities. Users with expert roles were given paper instructions in relation to each task (with specific details) which they used to instruct novice users according to these intructions. The expert was supposed to guide the novice, as he or she had a "superior" knowledge regarding specific task. The idea behind such a division between roles was to simulate real-world scenario where someone would be helping another person (by phone, as in conventional scenario) or with projected interface to solve specific problem of physical environment.

Description of the structure of test setup

The tests had two general parts. Total duration was set to two hours, with short break in the middle. The first part had an (1) introduction to tests, user tests (2) reviewing and analyzing videos together with users (3) and posting users' written comments on the wall (4). Users where introduced to the subject by brief explanation of the purpose of tests and through use scenarios – set of images which explained in a chronological order possible use situations for the technology which was tested (figure 30 and 31).

After first part with all four tasks was over, a discussion with the users based on video recordings from tests was held. The purpose of reviewing the videos was to let users see and comment on their own behavior and thus to get first-hand knowledge. Users were also able to see each other's behavior in addition to their own one. They could also comment out loud and write down notes, which were used later on for analysis. At the end of the session, users had to place post-it notes on the wall and explain their writings and to evaluate the difficulty of tests according to semantic differential scale.

Structure:

1. Introduction to user tests through use scenarios

2. User tests with 4 tasks

3. Reviewing observation videos togeth-

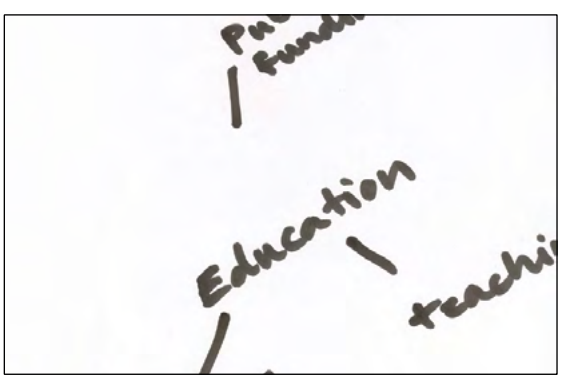


Figure 26: Mind map was made upon generic subject

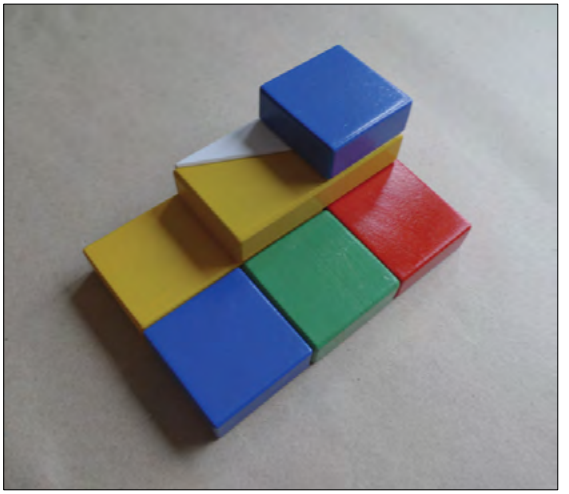


Figure 27: One of two block test tasks

Figure 28: Tilted plug table

er with users.

4. Posting users' written comments on the wall

Mind map

The objective of this task was to see how users use the projected space and how they communicate (verbally and visually) during projected interaction. The topic given was quite general: 'education'. This topic was broad enough and was thought as 'fitting' for the average age of users - 28 years.

Block test

The main objective of this and the following tasks was to test tridimensionality of projected interaction with simple and colorful blocks. Second objective was to observe how different colors and shapes of these objects affect the task.

Tilted plug table

The goal of the plug table test was to experiment with two things. First one was to test how well users see visual instructions when they are projected on surfaces with different brightness (white surface in comparison to black). In addition to this, the second objective was to test out how users cope with physical objects which are presented toward them with the angle which might be considered as "less optimal" and how they would cope in such situation.

Origami

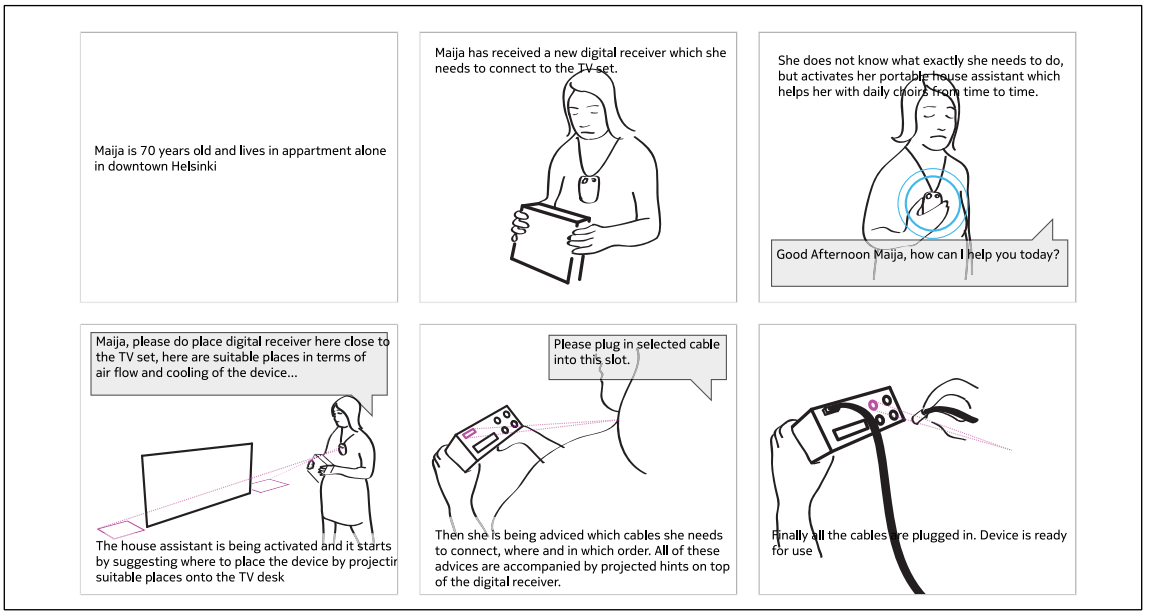
This task supposed to be most difficult as it contained small details, more than the

Figure 29: Origami task

preceding tasks and therefore was the last one in order. As an origami is a three-dimensional form, I expected that projected interaction with the help of visual instructions would bring an edge in building such complex three-dimensional objects.

User scenarios

The goal of these use scenarios was to introduce user test users to pico projecting enabled interaction and to give them an idea of possible use situations.



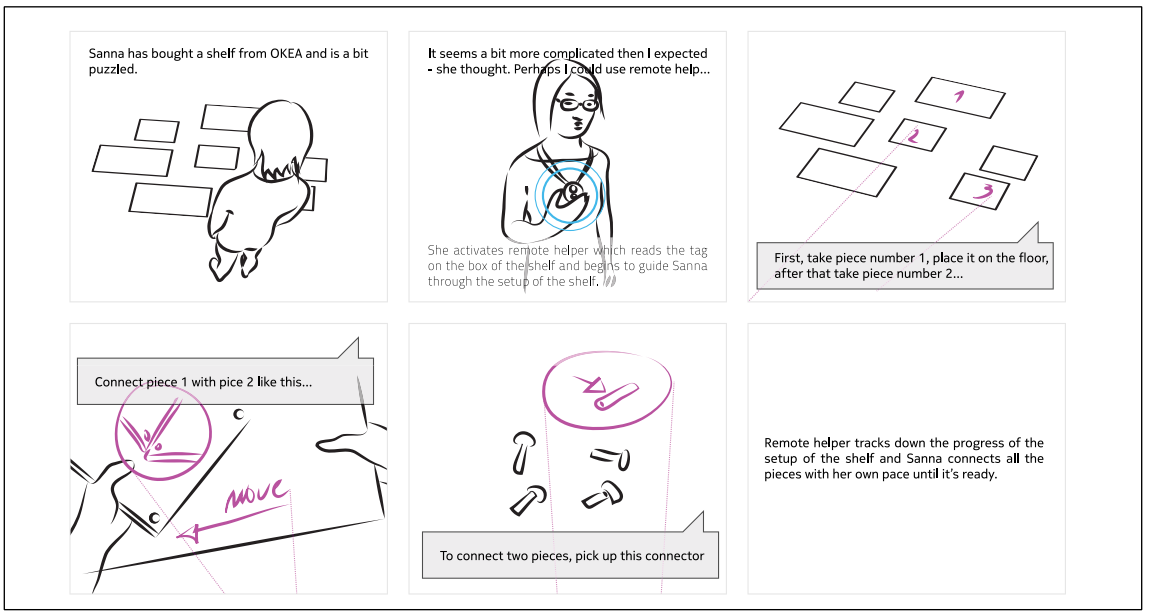


Figure 30 and 31: Scenario 1 - connecting digital receiver to the TV set. Scenario 2: Assembly of the shelf.

User tests results

Stakeholder description

Eight people were recruited for user tests. The amount of users was deliberately limited to this amount, because of the nature and goals of user tests. The point was not to provide statistically scientific proof, rather to provide hints of latent user needs, hints of possibilities of the technology and applicability this technology in remote interaction, also in relation to ubiquitous environment. The goal was to collect qualitative over quantitative data, to study phenomena experienced during user tests carefully with the help of video footage analysis, where also users were involved.

Four couples (8 users: 4 females and 4 males) took part in user tests. Majority of users were students from Aalto University and all users were given a compensation for participation (a gift card worth 20 e). Participants are described in the matrix (figure 32).

Half of couples used English and half used Finnish language for communication. Average age of users was: 28,1 years.

*All names of users are fictional

User's description	Role distribution based on tasks:				Language:
	Mind map	Color blocks	Tilted plug table	Origami	
Maria student (Media Lab) from Turkey, age: 24. (Phone: iPhone)	Equal roles	Novice	Novice	Novice	English
Kai student (Graphic design, MA), age: 30 (Phone: iPhone)		Expert	Expert	Expert	
Miia student (Haaga-Helia) age: 28 (Phone: Nokia 2760)	Equal roles	Expert	Expert	Expert	Finnish
Lauri civil cervant, age: 28 (Phone: Nokia 6700 slide)		Novice	Novice	Novice	
Satu student (Textile design, MA) age: 32 (Phone: Nokia-6100)	Equal roles	Novice	Novice	Novice	Finnish
Pekka student (Industrial design, BA) age: 22 (Phone: Nokia 6210 navigator & iTouch as mp3 player)		Expert	Expert	Expert	
Anna student (PhD) age: 27	Equal roles	Novice	Novice	Novice	English
Mark student (PhD) age: 34 (Phone: Samsung (touch))		Expert	Expert	Expert	

Figure 32: Users participating in tests in relation to distribution of roles (expert vs. novice) and language used.

Results from user tests are presented with division by tasks, (1) mind map task (2) block test (3) tilted plug board (4) origami and include comments from users.

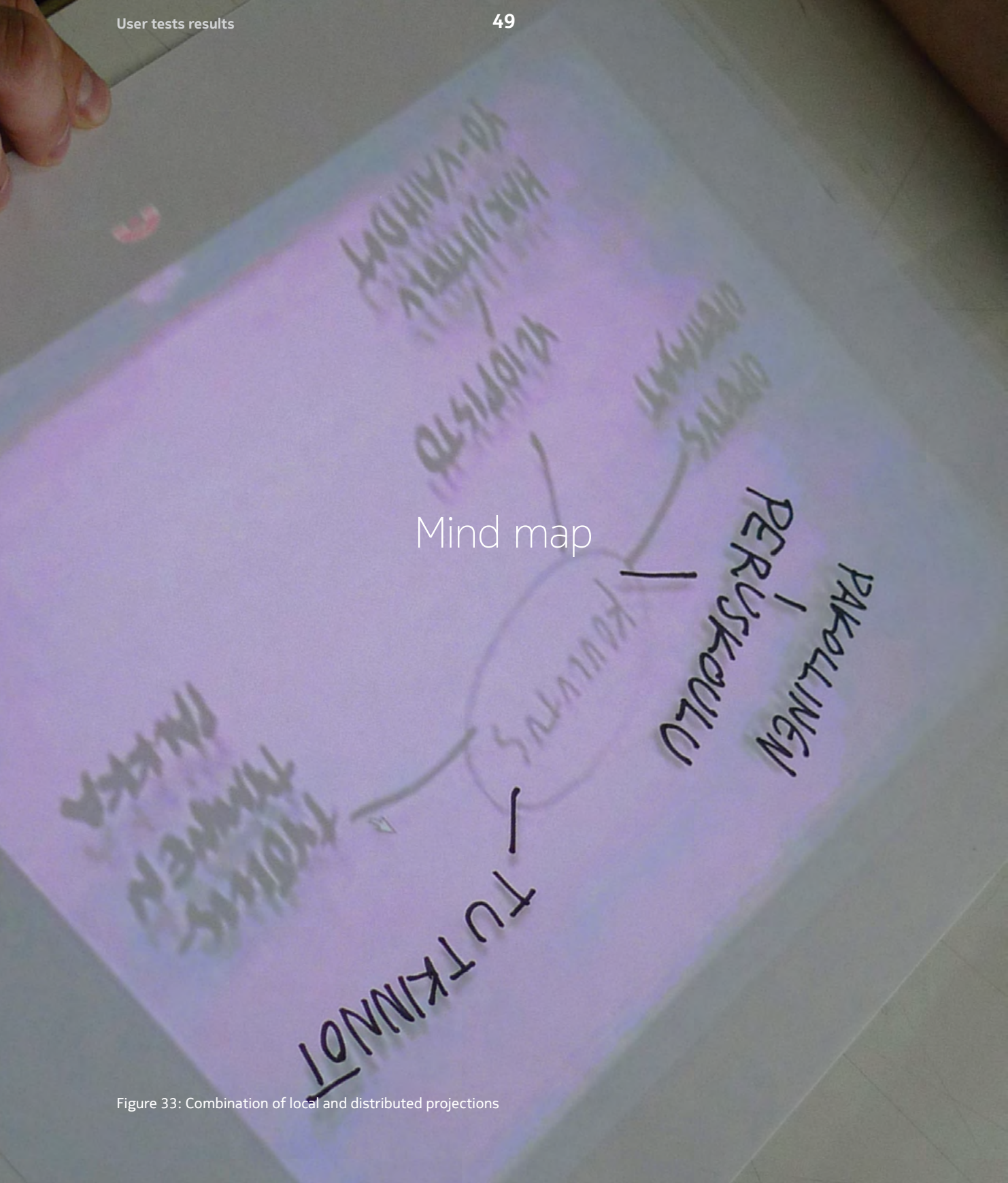


Figure 33: Combination of local and distributed projections

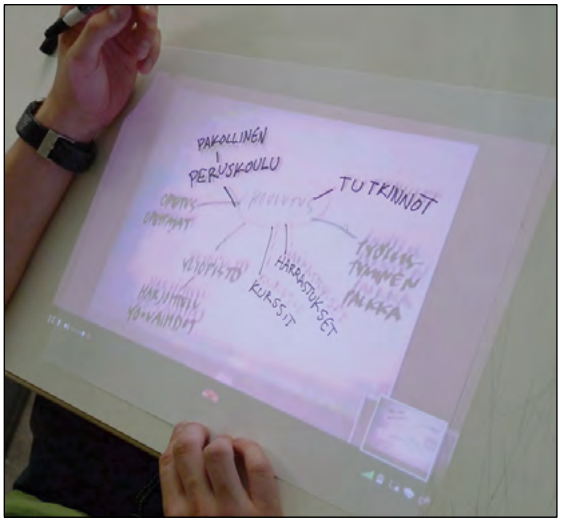


Figure 34: The view of projected area

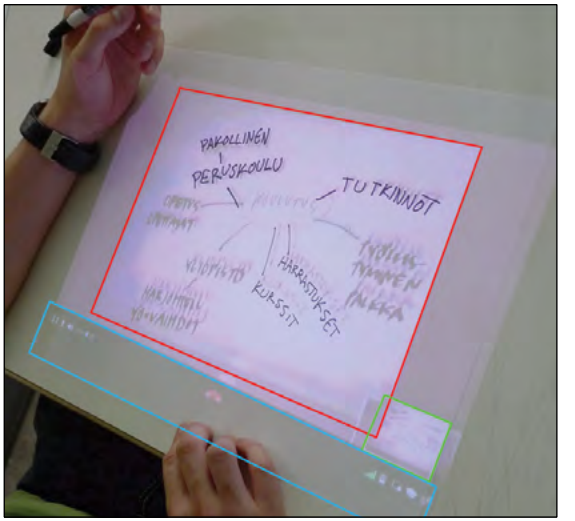


Figure 35: Active projected area (red), The view of own camera (green), VidHD controll bar(blue).

Projection area and brightness:

The projector was placed 70 cm above the active projected area (21 x 30 cm). The active projected area in this case means the area where interaction is seen in both ends. One of the users commented, that "the rectangle was too small."

The brightness produced by pico projectors used in the test was "up to 50 peak Lumens" (40 ANSI lumens) according to the specs. The amount of light in the room also affected the interaction. As one of users, Lauri suggested: "It would be good, if this device could adjust the lightness of the environment by itself..."

The task in general was quite simple, straightforward and did not involve that much of active conversation in relation to the topic. sers wrote down things on paper in turns. As the projection quality was not optimal for the use of thin markers (first users used thin ones (1-2mm), the rest of users - thick ones (3-5mm)) first users, Kai and Maria spoke up occasionally when they needed to verify what they saw. "What does it say" – was repeated quite often in this task. Besides this they spoke out loud what they were writing down as well.

Resolution quality:

Even though the resolution of Aiptek projectors, used in this test was sharp when single projector was tested (The resolution of each projector was 854 x 480 (WVGA)),

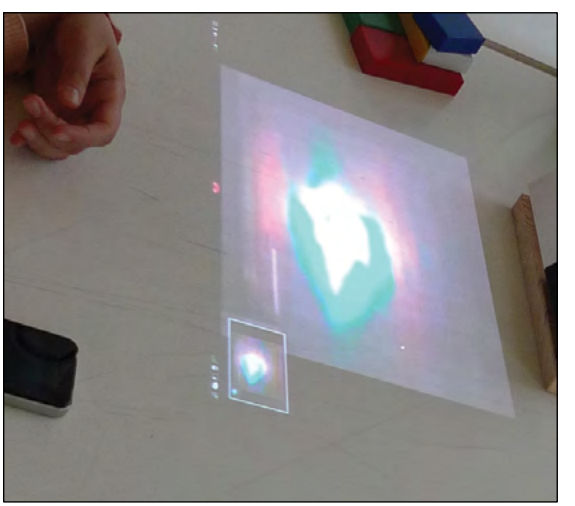


Figure 36: Spectrum in the middle of projected area, when the area was empty (white).

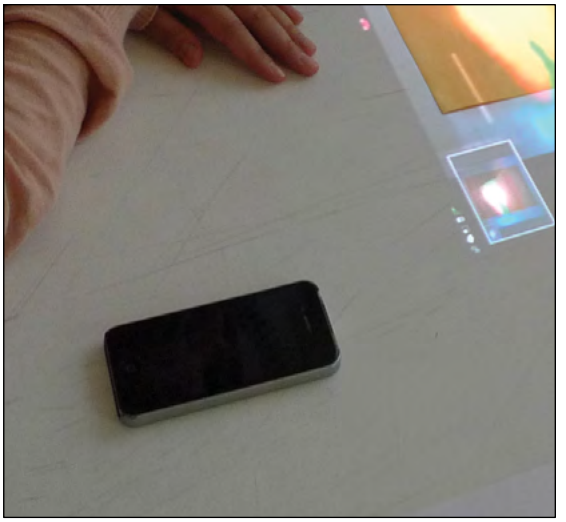


Figure 37: Phone speaker was used in most cases for audio communication

the problem appeared, when two projected images were placed on top of each other in each end, creating blurred effect when displaced. Kai, using thin pen for writing: "You can't really rely on what you read, because sometimes it's difficult to see."

Kai: "Also, this flare of light in the center. (..) Better resolution encourages interaction. And also voice quality is necessary." The projected image seen in each end had a spectrum in the middle as it was the result from double projection. One projection came from a projector situated in the same room and another projection came from another (remote) room. Maria: "Projection kills the projection and makes it like a phone conversation."

The image within the projected area was not sufficiently focused and produced duplication of the lines or 'shadows'. One the of users tried to move the paper to remove duplicated lines.
Miia: "When papers were moved the system did not realize this and did not focus papers. Papers were not placed on top of each other." Lauri: "It would be nice, if there was some kind of image stabilizer, which would focus the image. Another thing which messed up the thing was the shadow",

Sound quality

Headsets were planned to be used for audio communication, however in 3/4 of the cases they did not work and we had to use alternative audio communication - speaker on regular phones, placed close to each user.

Video lag

Maria: "It was hard for me to follow his hand movements."
Miia: "Another thing was the lag in time", you just waited, the hand was here there and then the text just appeared. I couldn't see each letter being drawn."
Pekka: "I wasn't sure whether you were about to
write or just wrested the hand, so I had to wait."

Benefits of interaction

Maria: "In general I can say it creates a feeling, of not being alienated. Because of the projection I felt that we were working on the same paper." Kai: "Yeah. There was like interaction going on. I can imagine if we were to design something and another person could improve it: 'how about this'."

Pekka:"It was fun, when the feeling disappeared that we were working on two separate sheets of paper." Satu: "Maybe we could do it without speaking, not as if you're fine if I write this? Instead we could

just be writing, continuing the writings of each other. It could be quite efficient for let's say four people meeting to share one flap board and work together."

Miia and Lauri saw this interaction as a way for remote learning. Miia: "It would be convenient for following lectures remotely. I could hear the voice of the teacher and other students and if the teacher was making notes here (shows a place in front of her) I could make notes if I was at home or at the other corner of the world. I would not miss the lecture if there was this kind of system."

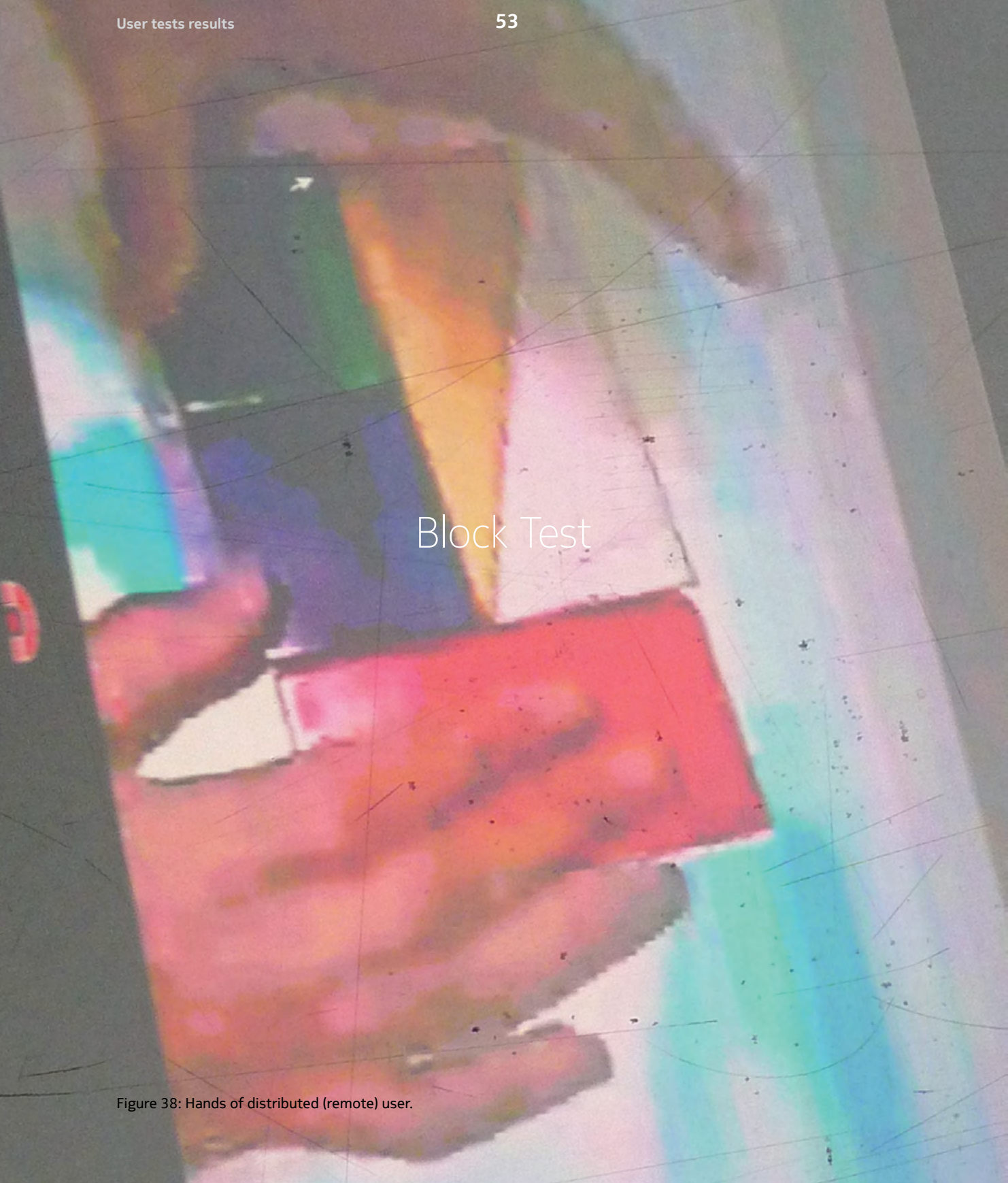


Figure 38: Hands of distributed (remote) user.

The following task of color blocks, raised two problems or challenges which affected the interaction. First one was the conversational grounding [30]. Two out of four user tests was communicated in English, which was not the mother tongue for any of users and caused problems in providing correct instructions and understanding them.

The second challenge or problem in this particular task was the novelty of interaction. It took time to get familiar with what projected interaction can and can not enable. Because of this novelty users first tried to rely on a conventional method of interaction – verbal communication. When that for one reason or another was found insufficient, users took projected interaction into the use.

Size of objects:

Working with color blocks or large objects in general was experienced positive. Users did comment that the prototype works worse the higher the amount of details becomes. Lauri: “Probably will not be that good with lace making for instance”.

Colors:

White, green and blue colors were problematic in terms of visual recognition. Kai: “Because of poor projection it was really difficult to see which was the white triangle, because there was not enough contrast.”

The hand did ‘blend in’ into the projected area: Kai: “Was it difficult to see my fingers when I was drawing, was it happening in real time or was it…” Maria: “What can I say, you put your hand on it and you also get the projection on it and I get the projected image and it’s the same colors with your hand.” Kai: “So it’s difficult to see. So maybe if you could draw with a black marker on a transparent and … (shows placing something over something with his hands) then it …” Maria: “Yeah maybe.”

Miia: “What was good in these block tests was that they had different colors, they were much easier”.

Drawings

Pekka, one of the users was an experi-

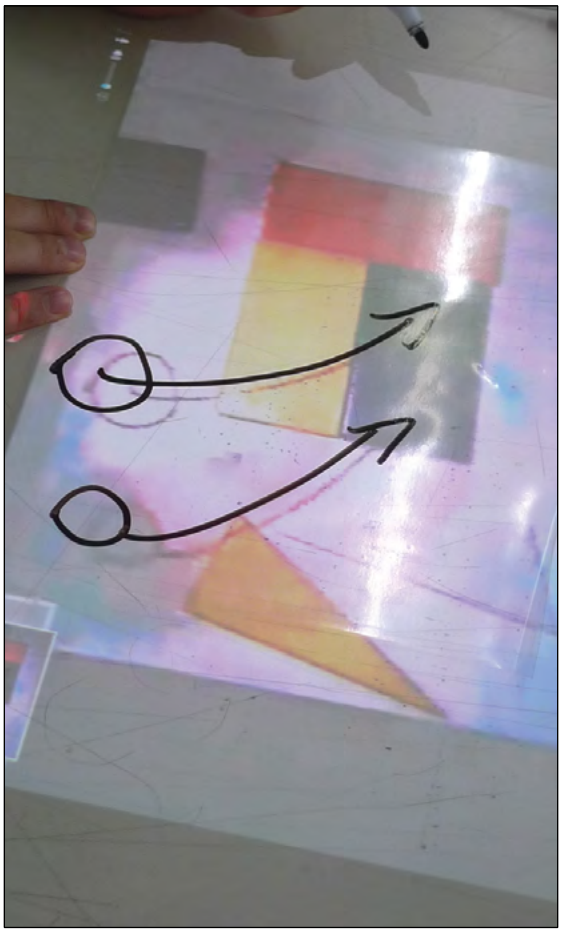


Figure 39: Drawing instructions

enced user of drawing tablet and therefore picked up the pen as a tool for giving instructions already in block test, even though some problems appeared with the use of pen on top blocks. Pekka: “The previous task (mind map) was two-dimensional and here if one draws circle part of it drops down from the block.”

Helping systems

Miia: “In this block test, it could have been projected where each block goes. You could see what kind of blocks you have and it would say yellow goes here and green goes here.”

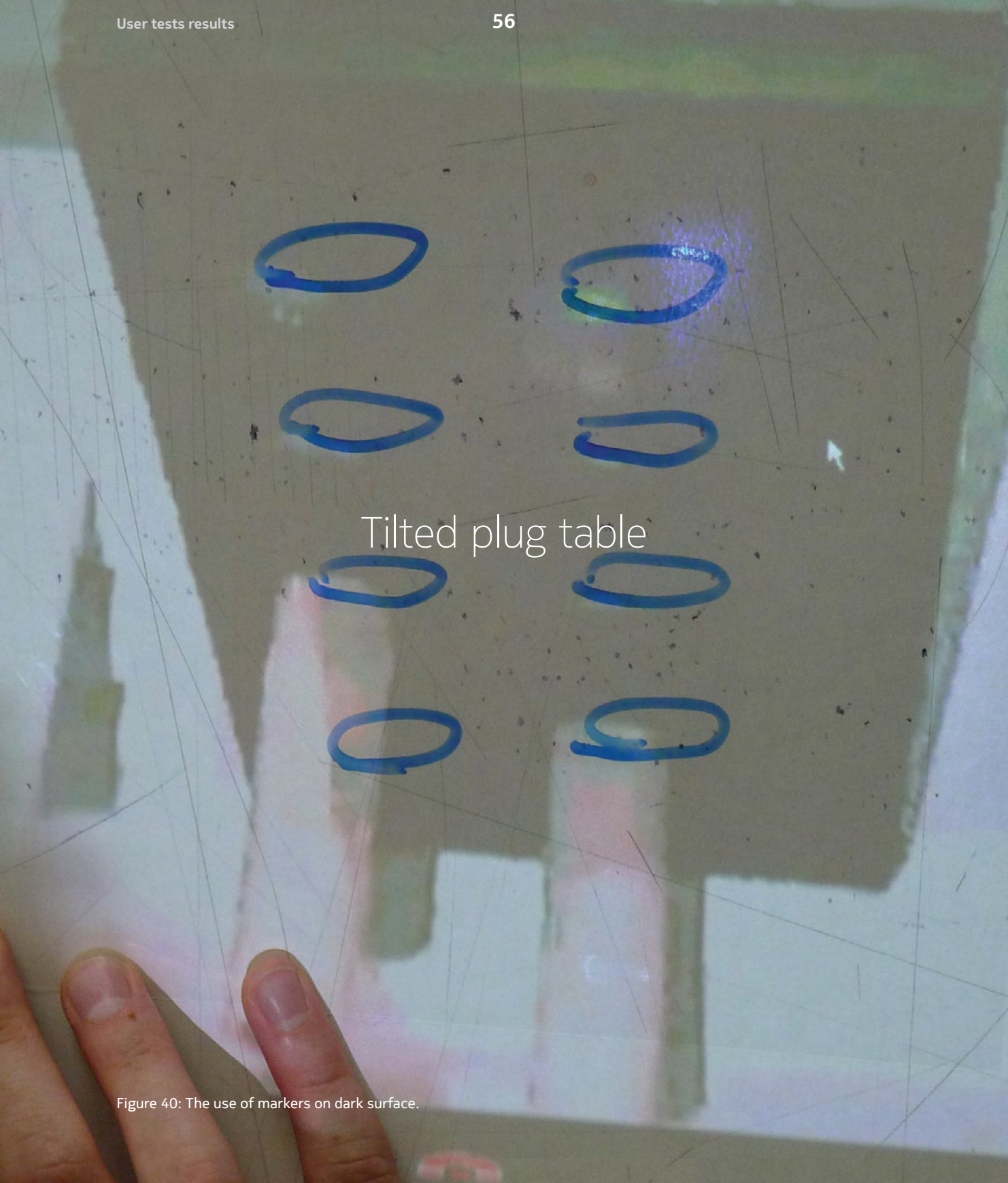


Figure 40: The use of markers on dark surface.

Tilted surface was challenging for the users as the plug board was tilted (45 degress), it appeared smaller in expert's view and tilted surface made projection more inacurate than projection on flat surface. Half of the experts asked to turn the surface so that it faces them as 'flat' (figure 42). Pekka: "the projection needs to be 1:1. It needs to be accurate where you draw or where you think you're drawing."

Also the black side of the board made the use for the projection of hand gestures and drawings difficult as expected. Some users came up with interesting solutions.

Black side:

The black side of tilted the plug table caused problems as it was difficult to use it for drawings or for projection of hand gestures. Majoritiy of users solved this by relying on verbal explanation, while one user used drawing on white side, to show actions, intended for black side (figure 43), to show wher plugs should go.

Miia: "I wasn't sure if there was one more row below. In the case of black color, you could not make sense of it from the top view."

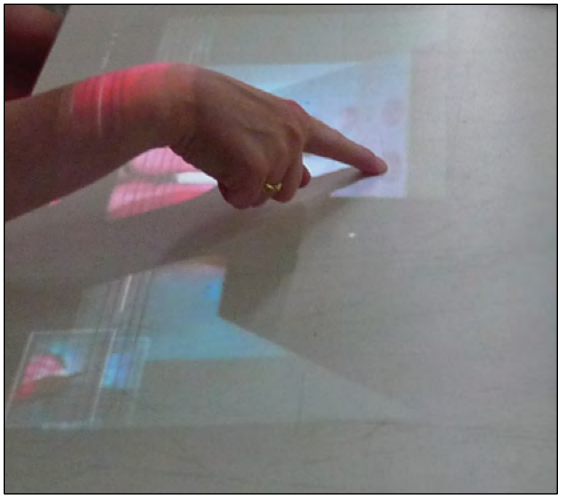


Figure 41: Pointing was mainly used in this task to show where plugs should go

Figure 42: Half of users turned tilted plug table so that plugs would face the expert directly

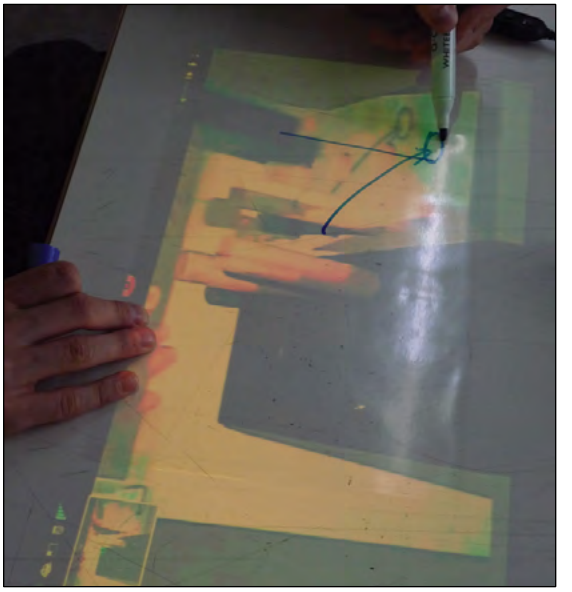


Figure 43: One of users used drawing as a mean for communication of instructions

Pekka: "I wasn't a hundred percent sure about the black side, but when I saw one plug and then I saw that the other one is lower, I could judge it from that."

Control of the image:

Kai: “It would be good if you could adjust the area which is being projected. Of course if you go higher up, it darkens the image.”

Pekka: “If there was some kind of home helper device, the projection could be changed in perspective, so it can adapt to where you can see the instruction.”

Other representations of hands:

Lauri: “And .what would be quite wild, if one could make the program, if there was the possibility, to place your hands on what you made in art class for instance if you were to repeat what you were doing, your hands were kind of a transparent hologram. That would be pretty cool.”
Miia: “It would be like this - do this” and shows turning something with her hands.

Voice only:

One of users, Mark used a matrix as an approach in this task: “If you look at it as a matrix, position 1-1 would have blue, position 2-2 would have white, position 3-3 would have yellow, position 4-4 would have the other yellow.”

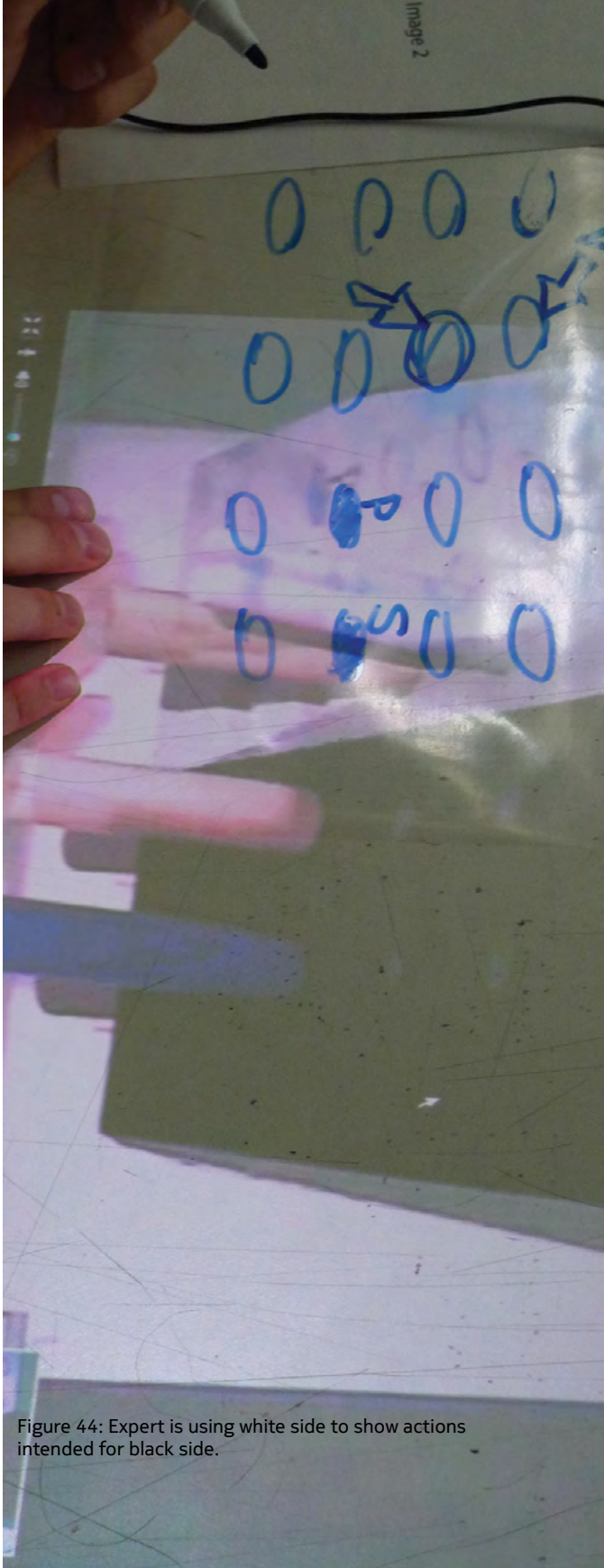


Figure 44: Expert is using white side to show actions intended for black side.

Origami

First half of users failed in completing origami task within set time(~30min). As the origami "peace dove" seemed to be too difficult, I decided to seek for origami where folding lines would follow simpler geometry (foldings according to 180, 90, 45 degrees in comparison to "peace dove" with degrees as small as 20). Simpler origami "water bomb" was then used with the rest of users.

Peace dove

Water bomb

Figure 45: Origami task required change as "peace dove" origami was too difficult.

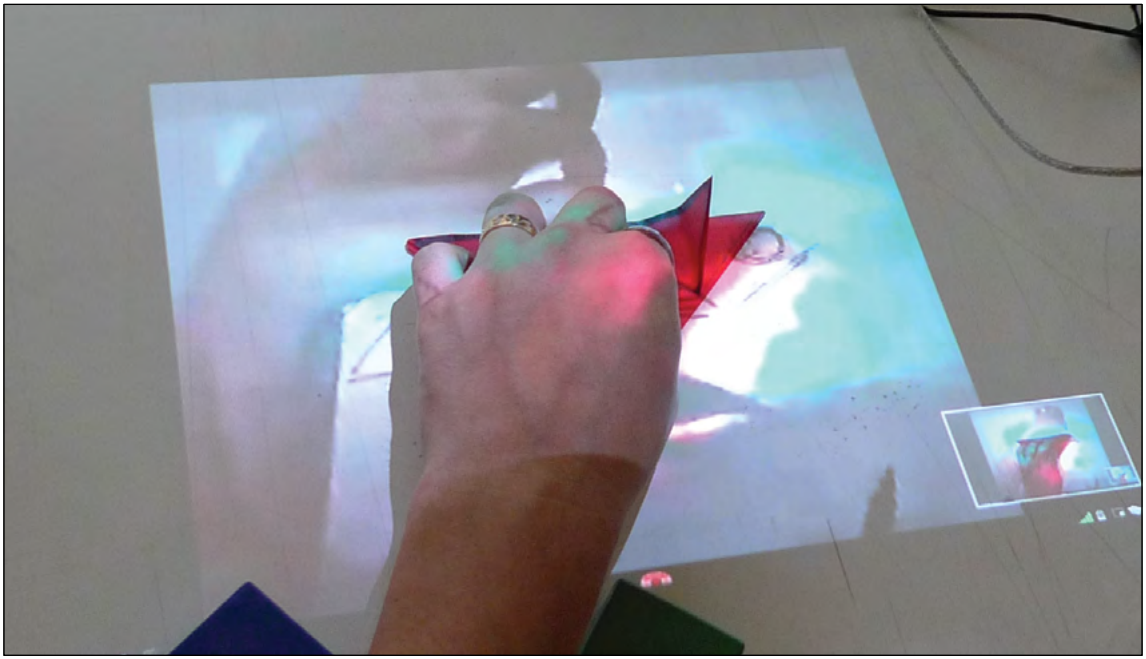


Figure 46: Novice folding according to the lines drawn by an expert

Origami was the most difficult task as expected. While there were problems with readability due to poor quality of resolution, sharpness, light, projection, contrast and focus, also problems related to language did appear. Most surprising was the notion that one of the users was not sure or needed to verify if his or her hand movements were seen by the other person. Kai asked surprising question at the beginning of origami task. He placed his hand within projected rectangle, moved it a bit and asked Maria: “Do you see what I’m doing here?” When Maria confirmed that she sees it, Kai asked “You do”? Surprising here, was that he needed to clarify the visibility of his hand as he was not sure about whether the other person sees it.

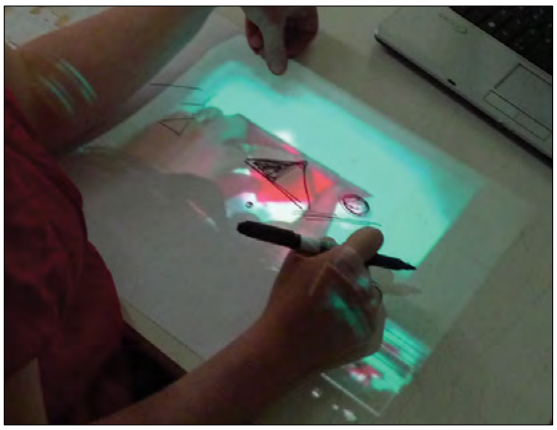


Figure 47: Expert drawing the lines

Different colors of paper were tested as well in following order (with the goal to test whether the color affects the task): Pair 1 (yellow), Pair 2(Red), Pair 3(Light Pink), Pair 4(Red). While the red color was the darkest in terms of projection in comparison to yellow and pink, results suggest that no clear link between the color of paper and the success of the task was noticed. Half of users failed the task (paper: Pink and Red) and half of users succeeded in completing the task (paper: Yellow and Red). Results suggest, that this was mostly due

to the complexity of the task, as those users who failed the task were working on "piece dove" which was more complex.

Showing simultaneously

Maria: "I think it would be easier if you did it at the same time. It would be easier to follow." Kai: "I told her first that I had a piece of paper, so I could show her. So it could be ..." Maria: "More helpful." Kai: "Sometimes it's difficult to find the words, to make it understandable."

Perspective

"Hold on. - says Miia. Now that I see this from different angle, this is so much easier to see what's going on." And points at the video where Lauri is folding the origami. This angle is different than what she saw during tests, as it was from a video camera which was recording the whole interaction from the side. "So if it was a perspective from the room it would be easier" – adds Lauri.

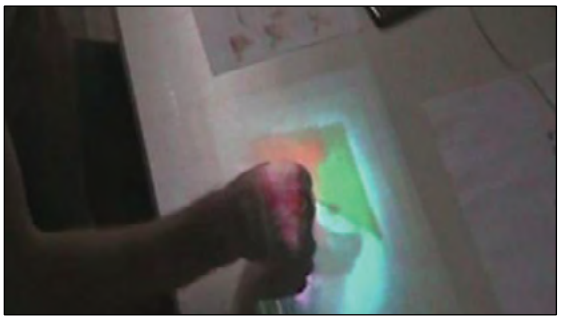


Figure 48: The expert draws an outline of the triangle with finger

Figure 49: Novice sees the movement

Pekka said while watching the video footage, that he saw the origami originally from the top view and now that he saw the origami from an angle he saw more than originally. He said that at one point the origami looked as flat, the way he originally saw it but then he noticed that it was bit opened after all. "It was projected from the top so I did not see it, if it was projected from the side..." said Pekka.

Mark: "I could not see it because it is 2D and I could not see a shadow, so I could not see how you did it."

Removable drawings

When asked if drawings were ever in the way, Satu said no they were not, as she moved the sheet away while working on it and the drawing was still there. She said though, that the reason for taking the paper sheet away to fold it was not the drawings which were projected, instead it was simply easier to fold the paper when it was closer to her.

Transparency by request

As Anna, one of the users explained, hands were sometimes covering valuable information: "I could not tell which point, because your hand was.." (shows palms over something)
Lauri: "Also the hologram-thingy, could be perhaps used, so that your hand would be slightly transparent or with outlines."

Chapter 6:

Findings

Analysis methods

Practical limitations

Evaluation by designer group

Gestures and drawings

Analysis methods

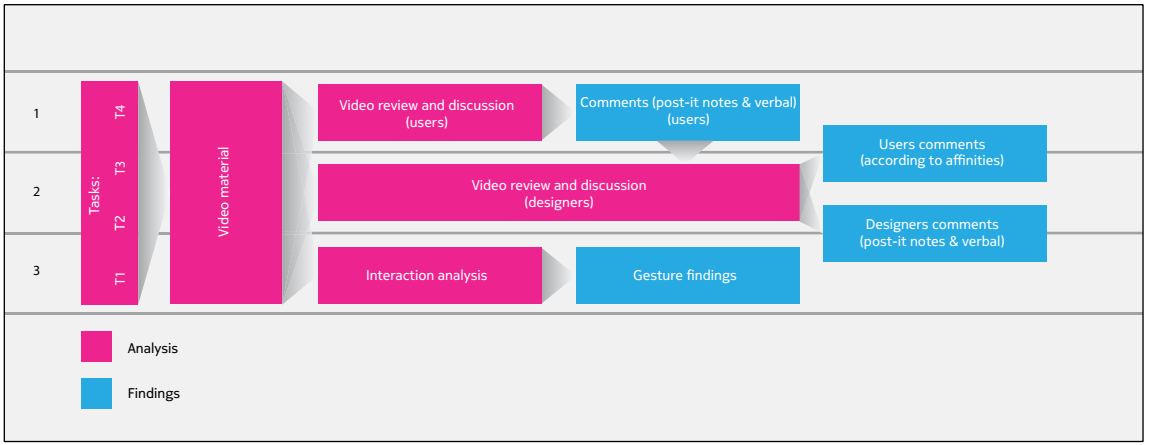


Figure 50: Data analysis

Data analysis (figure 50) was based on (1) user comments (post-it notes and verbal comments and query according to semantic differentiation) and on reflecting those comments and video material with (2) a designer evaluation group and seeking for patterns. During evaluation and reflection with the designer group, some suggestions were added by designers as well and such will be presented separately. In addition user videos were analyzed from the perspective of the use of (3) gestures and drawings during interaction.

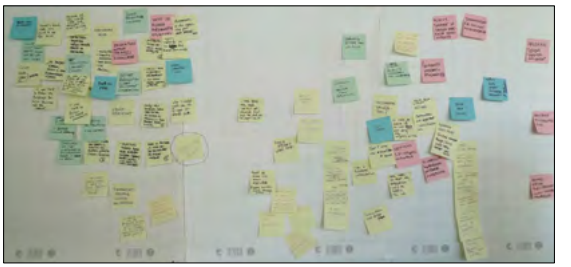


Figure 51: Designer evaluation group assisted in arranging user's post-it notes according to affinities

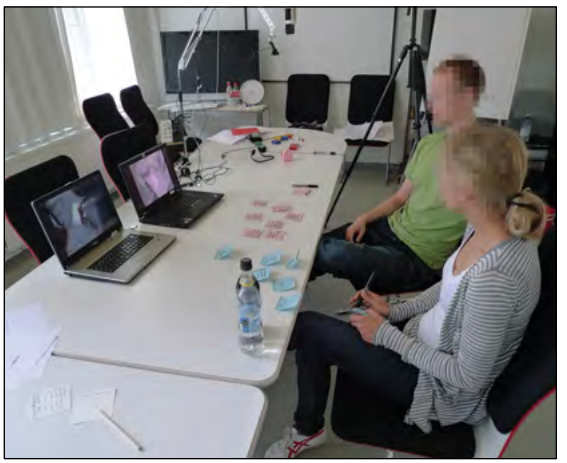


Figure 52: Users reviewing observation videos

Semantic differentiation:
1(very easy) 5(very difficult)

Mindmap	2.5
Block test	2.4
Tilted plug table	2
Origami	4

Figure: Results of semantic differentiation query.

Practical limitations

To address practical limitations which were raised during user tests and to address prior research related to pico projectors a meeting with Jarkko Mattila from Nokia was arranged after user tests. During this discussion, pico projector technology was discussed and a big part of the discussion was related to the technology itself. As a results of this discussion, Mattila's feedback was used to generate proposals to solve some of the practical limitations experienced during user tests. Solutions and proposals are included in the following text.

Resolution and focus

"Better resolution encourages interaction." (one of users) The goal of the projection was to provide sufficiently accurate interface so that users could draw directly on physical objects. The resolution and focus became more important - the smaller the details of objects became (origami) or when the object was positioned in a disadvantageous angle toward projector (tilted plug table). While projection was relatively accurate at the center with few millimeters of inaccuracy (e.g. shifted lines in mind map task), inaccuracy grew towards edges of projection up to 15 mm. Because of this users also saw "shadows" or "ghost lines" - local line transmitted to remote place was send back and was not projected onto exactly same place as originally.

As in this setup, the aspect ratio in one of the pc's was different than in the rest of equipment, one solution to fix the inaccuracy fix in the future would be to align aspect ratio throughout the whole equipment (Pc's, projectors and cameras).

DLP projector, used in this test requires manual focusing and is sufficient when projected area is flat. However when the projected area is uneven (an object), a laser projector would provide a clear advantage, because it has a far superior "dynamic area", which is always in focus.

Double projection

Double projection means that each user saw not only the projection from another user, but his or her own projection as well projected back to him. These two projections were then placed on top of each other resulting in double lines, blurring and enhancing the light concentration at the center of projected area i.e "the flare" or "spectrum of colors".

Alternative solution for this problem would be to replace the projected interface at expert's end with touch-display device, so that the expert would interact through display rather than projected area. Another solution would be to test whether double projection works better with laser technology as it is sharper and does not necessarily create a spectrum effect.

Video & refresh rate

Wlan connection used for communication between two pc's was somewhere between good and very weak during user tests. As commented by one of the users who preferred a pen, the refreshment rate could be faster and if it was faster, perhaps that would make him use hands more. In this sense drawings were not only valuable as a subject of giving instructions in this thesis, they also provided longer lasting effects than movement of the hand where the movement disappeared if it was made too fast.

The importance of a sufficient network connection and refreshment rate needs to be taken into consideration in future tests.

Audio

Headsets proved to be unreliable and regular phone connection was used in 3 out of 4 cases.

As quality of sound is extremely important, this needs to be taken into consideration in following tests.

Approach of Augmented reality

As user tests show, the misalignment of projected and real lines of objects result as a blur in workspace. Markers used in AR, (figure) on page 24, could help in focusing , and aligning two remote workspaces when placed within each workspace.

Similar approach with markers on the fingertips of the expert could also provide a short-term solution for the use of hands to create drawn lines which would not stay permanently in novice's view.

Evaluation by designers group

Figure 53: Rearrangement of post-it notes according to affinities and adding designers post-it notes

User tests were followed by reviewing user videos, discussing and reflecting findings with a group of industrial designers. Designers (D1, D2, D3) were not familiar with the subject before this, therefore they had 'fresh' approach in relation to the findings made. The overall goal of evaluation by the designer group was to get feedback on user tests and reflect ideas on a higher level. Designers had a chance to test the technology and interaction themselves. The review consisted out of reviewing some of the videos from user observation, writing down comments, verbally commenting inbetween of videos, attaching written comments on white board and helping in arranging users post-it notes according to affinities.

Voice and projecting:

Projected interaction model was seen as beneficial in general in "Assembling simple things"(D3) and that "Voice is enough for simple task, but then a projector can give specific details. You can really show: I mean this one (shows by pointing down with a finger). You can also see what's the situation, because using only voice you don't know what's happening."(D1)

Remote access and guidance:

"Reminds of 'sametime' (communication/ chat tool), allows a tool to take over and control. Projector would allow to take it to any place... Does it have to be matched into both ends? A tracking movement pen? Almost an educational tool as I see it."(D2)

Education:

"I can see this as an education sort of thing. If the teacher is in the class and he is watching at projection in front of him, first he shows something and then he can switch between students and see how everyone does it. And then explain what they've done wrong. It's kind of an exiting idea. So either students can be physically present or be in any part of the world."(D2)

When discussing the difference between watching a simple video tutorial and receiving guidance with the projector, it was commented that the projector "feels concrete and practical" and as the connection is 'live' there is no gap such as in tutorial watching situation, where user needs to "switch in between" the task and tutorial.
"And if you do something wrong, what then? The tutorial doesn't know. The tutorial goes on and on."(D1)

Projecting on real physical objects:

Designers commented that the projector could be used as "pointing device"(D3), "Augmented finger"(D1) or "smart laser pointer (..) could almost be like a feature in the phone, built in."(D2) Such device, could point at critical information and work in light condition, in the same way as "saw laser" works: "it shows where it's going to cut, so you cut it right there. It works in light conditions also."(D2)

The perspective/viewpoint:

Because the prototype used in user tests was fixed on the table, it did not give the most optimal point of view for the expert to provide accurate instructions. It was emphasized by designers, that users need to share same viewpoint: "The most optimal thing is, I see what you see (..) I know exactly what you're looking at."(D1) Or to be able to move and adjust the viewpoint: "In some user cases e.g. construction workers have inbuilt things in their helmet or something, but if we talk about more regular (activities), you have to take the device somewhere around the objects and that person sees that. And the person at the other end somehow has to be able to point out with whatever projector or laser at things – what he wants to explain. That the image streamed to the other end has to be interactive, so I could point out things to the receiving end."(D2)

Drawing:

The importance of different drawing skills in relation to conversational grounding was mentioned as well, as "explanations are important" (D3) and "lines and arrows are quite natural, you could draw them but after a while they could just vanish after few seconds." (D2) At the same time it was mentioned that "people have different skill levels"(D3) and that this should also be addressed in the future.

Future Research:

Also the extent of implementing projecting technology was raised: "Probably if you talk about near future you might want to take small steps and if you talk further away you could talk in a more interactive manner – full picture."(D2)

Gesture and drawing analysis

Both hand gestures and drawings were used during interaction. Not all of the users realized immediately that hand gestures are transmitted to the other or were unsure how to use the projected interface. One couple realized this during the tilted plug table task. The novice user suggested to the expert user: “you can use your hands to point”. Two other expert users seemed to be unsure about whether their hand movement is seen at the other end during the last task. One expert user asked at the beginning of the origami task: “do you see what I’m doing here?” Another expert user asked in the middle of the origami task: ”can you see my finger?” at the point, when he had problems in explaining instructions verbally. Despite some difficulties, uncertainty among couple of users and slow starts, following images show the extensive use of both hand gestures (1) and drawings (2).

1. Hand gestures

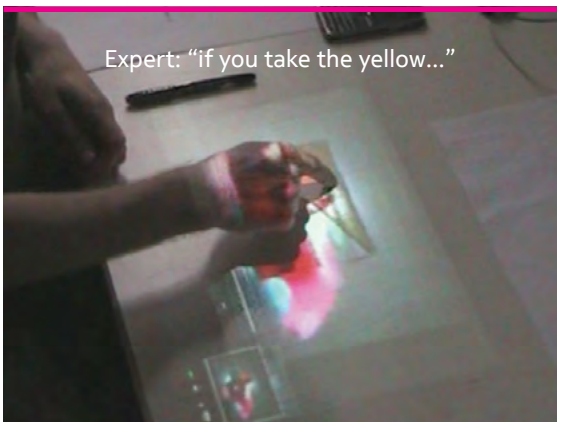


Figure 54: Pointing (expert view)

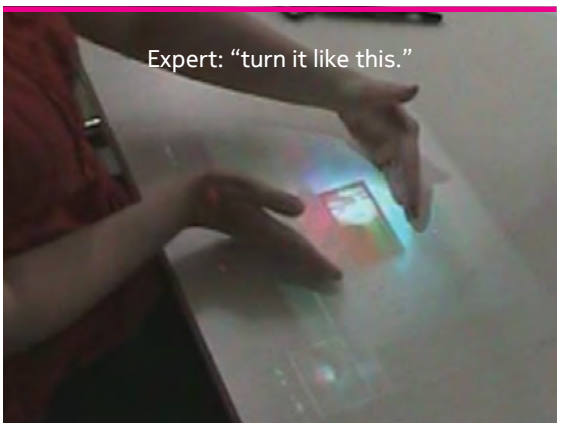


Figure 55: Showing orientation and place of object

Figure 56: Non-transmitted gesture (user does not realize his hand is not seen or is not aware that he can use his gestures for interaction).

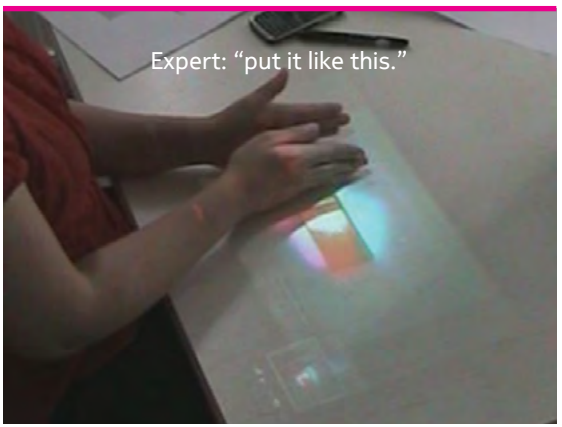


Figure 57: Showing orientation of object (here: vertical)

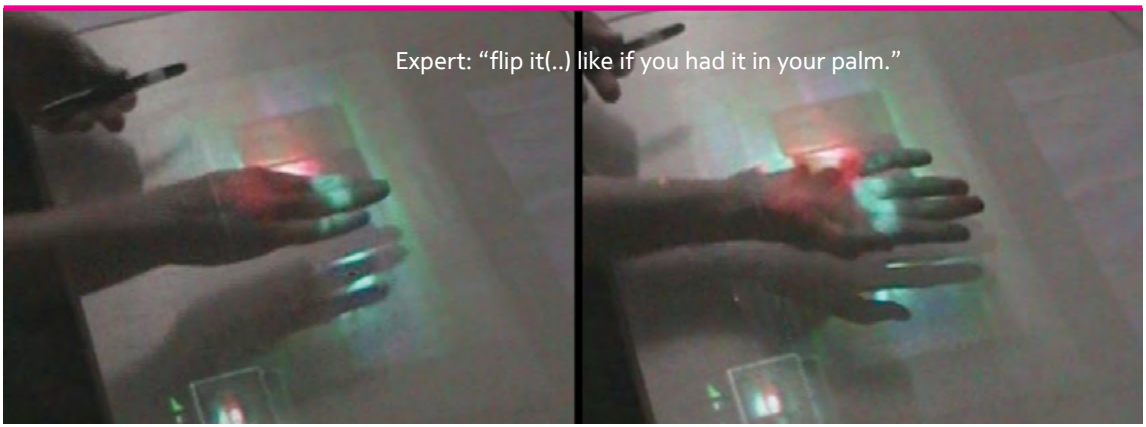


Figure 58: Indicating the action (here: flipping the object)

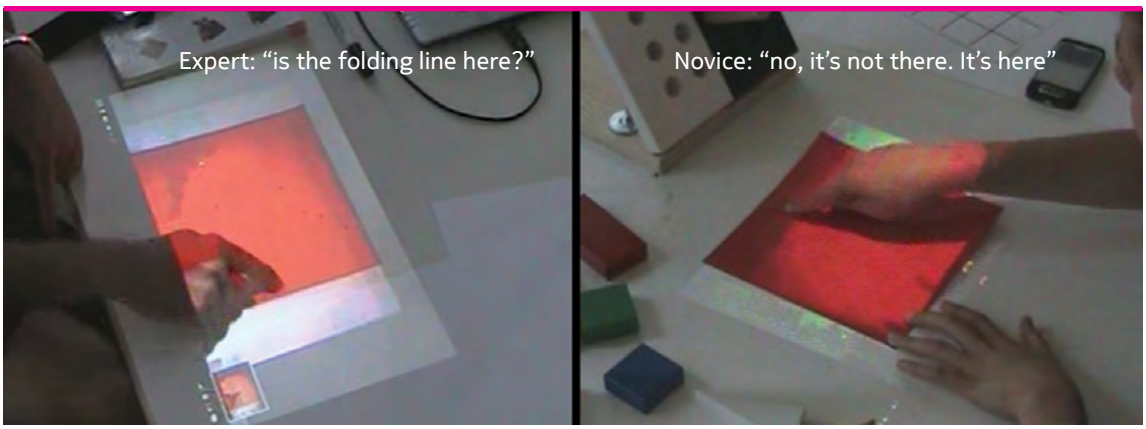


Figure 59: Showing the line with hand movement

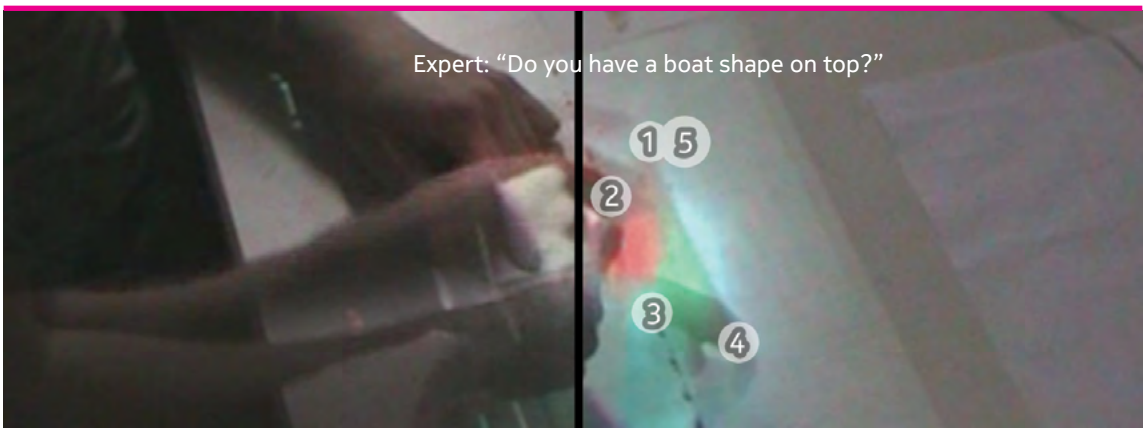


Figure 60: Showing the outline by movement of the finger (order: 1,2,3,4,5) (here: a “boat shape”)

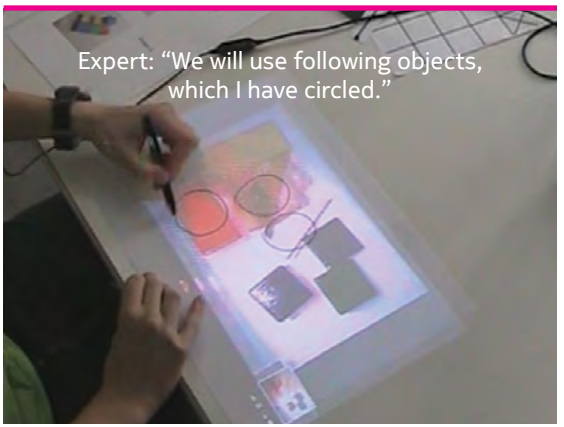


Figure 61: Circling objects (expert view)

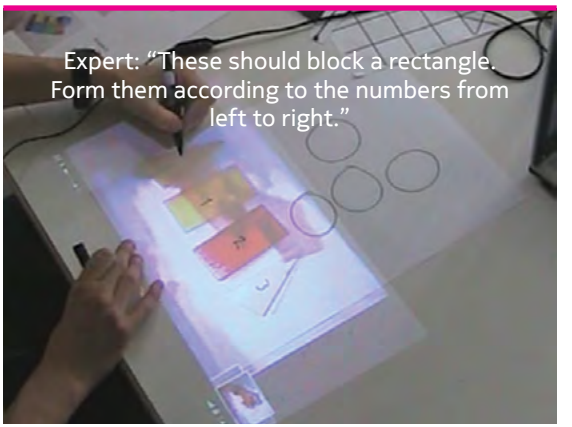


Figure 62: Numbering objects (expert view)

2. Drawings

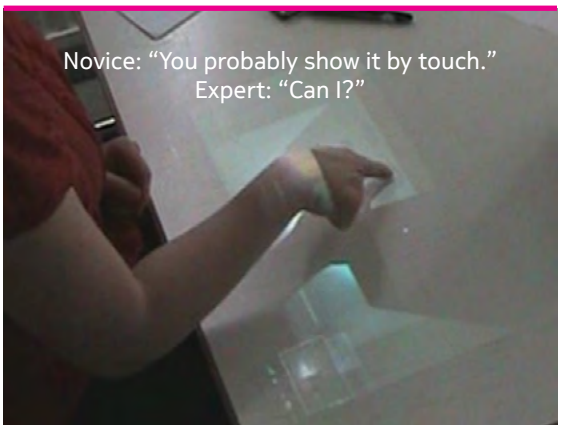


Figure 63: Pointing where plugs should go

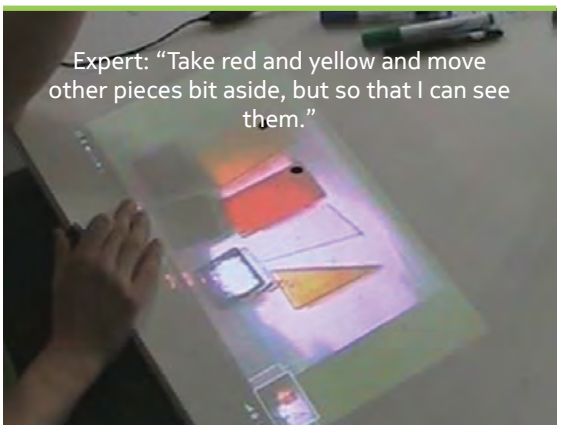


Figure 64: Marking objects

Figure 65: Using arrow to indicate movement (expert view)

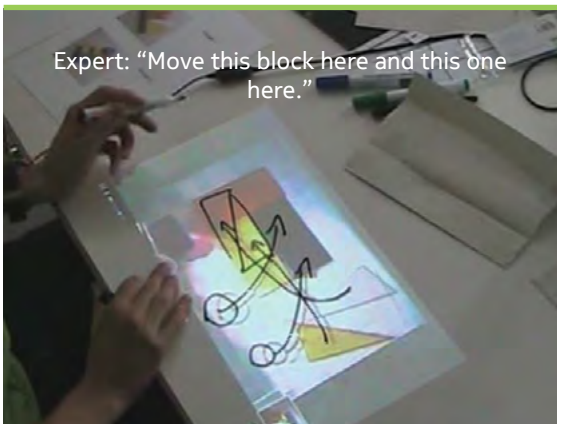


Figure 66: Using object outlines to indicate placement (expert view)

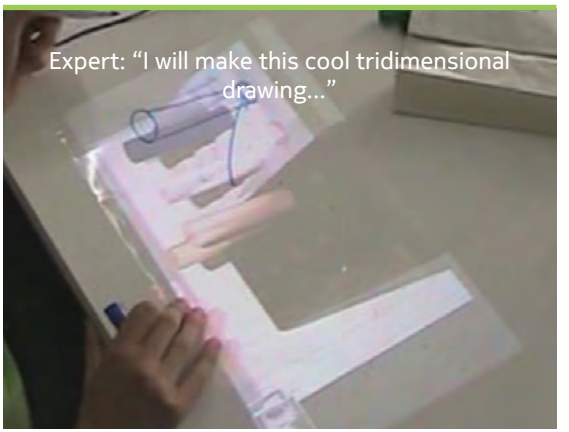


Figure 67: Tridimensional drawings

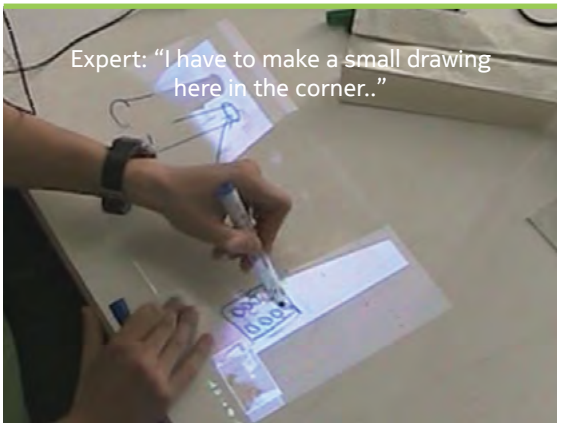


Figure 68: Using misplaced drawing when surface (here: black) is not optimal for projection

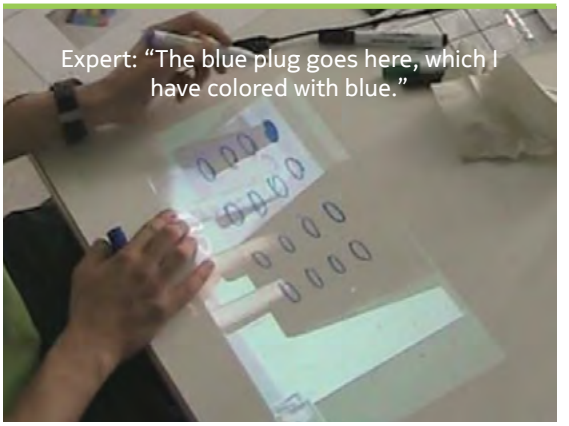


Figure 69: Drawing a grid on object

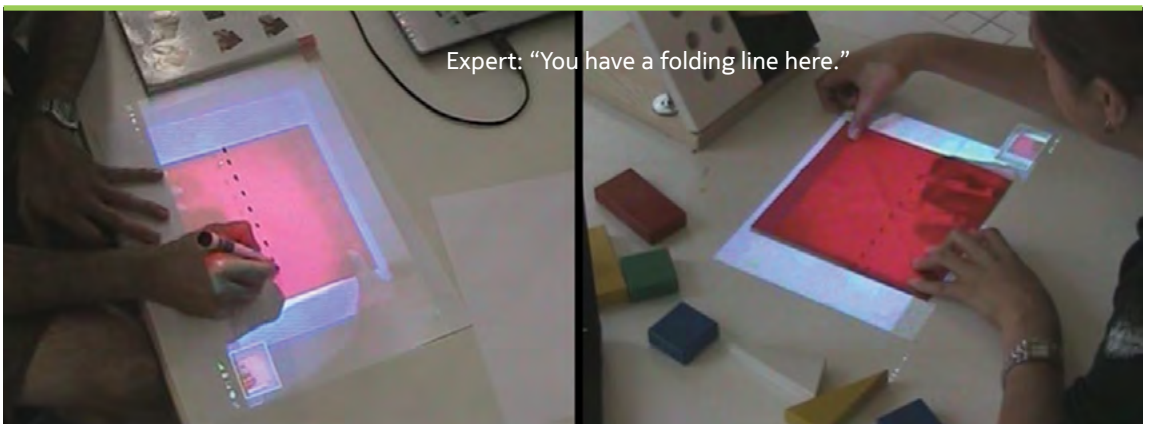


Figure 70: Drawing the outline where paper should fold

Figure 71: Drawing boundaries of objects with folding lines

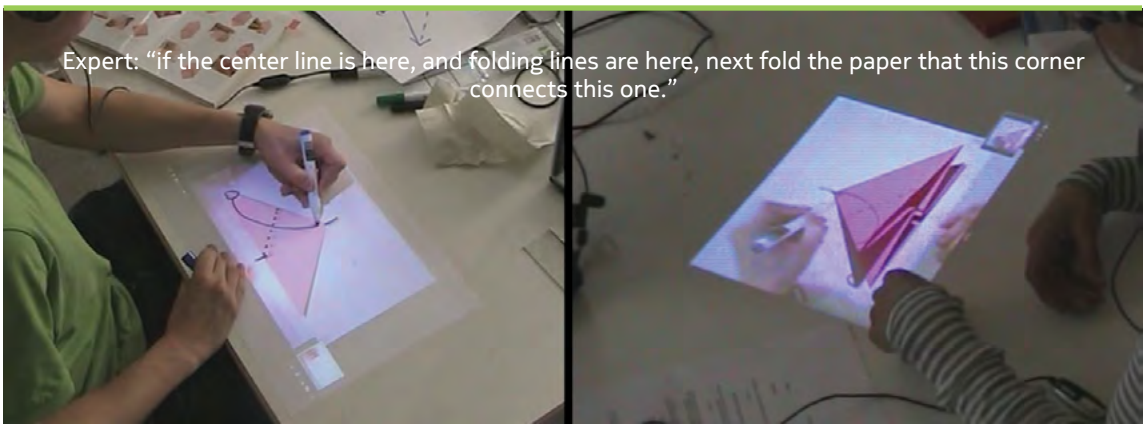


Figure 72: Drawn folding and movement in novices view.

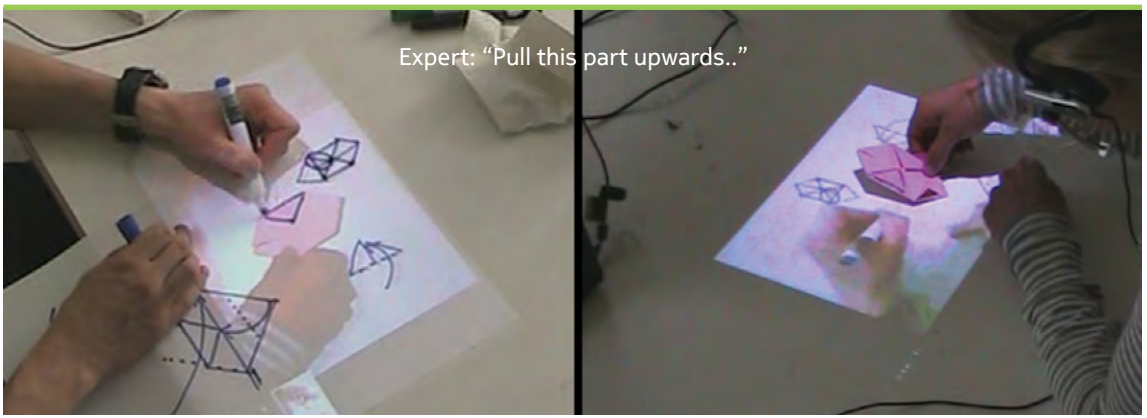


Figure 73: Instructions for high level details

Hand gestures and drawings from video analysis are categorized according to two categories. First, gestures are arranged in matrix according to the preferences of each expert user (figure 74) between gestures or drawing, in relation to tasks. Second, examples of gestures used by each expert user are shown in relation to the type of gesture (figure 75) [33].

The matrix with preferences of experts (hands vs. drawing) shows that most of users were using hand gestures up until origami task where they swapped hand gestures to the drawing at some point. Only one of expert users, Pekka used drawing throughout the whole set of tasks, as he felt that using hands was "clumsy way". Clearly while he was using drawing from the first task till last one, other expert users relying on hand gestures at first had to change the strategy with the last task as they noticed the hand gestures were not enough in origami task, which was more complex than previous tasks and included higher level of details. Therefore origami task was clearly critical where simple hand movements, which "blend in" with objects and were seen in brief glimpses rather as gradual movement were replaced by drawings which had clarity, provided instructions as long as needed (until the expert removed them) and could communicate movement with arrows and movement lines.

"When I try to point at something, it takes at least five frames till the hand moves there. But when I draw, I can directly explain: 'fold according to this line' and the

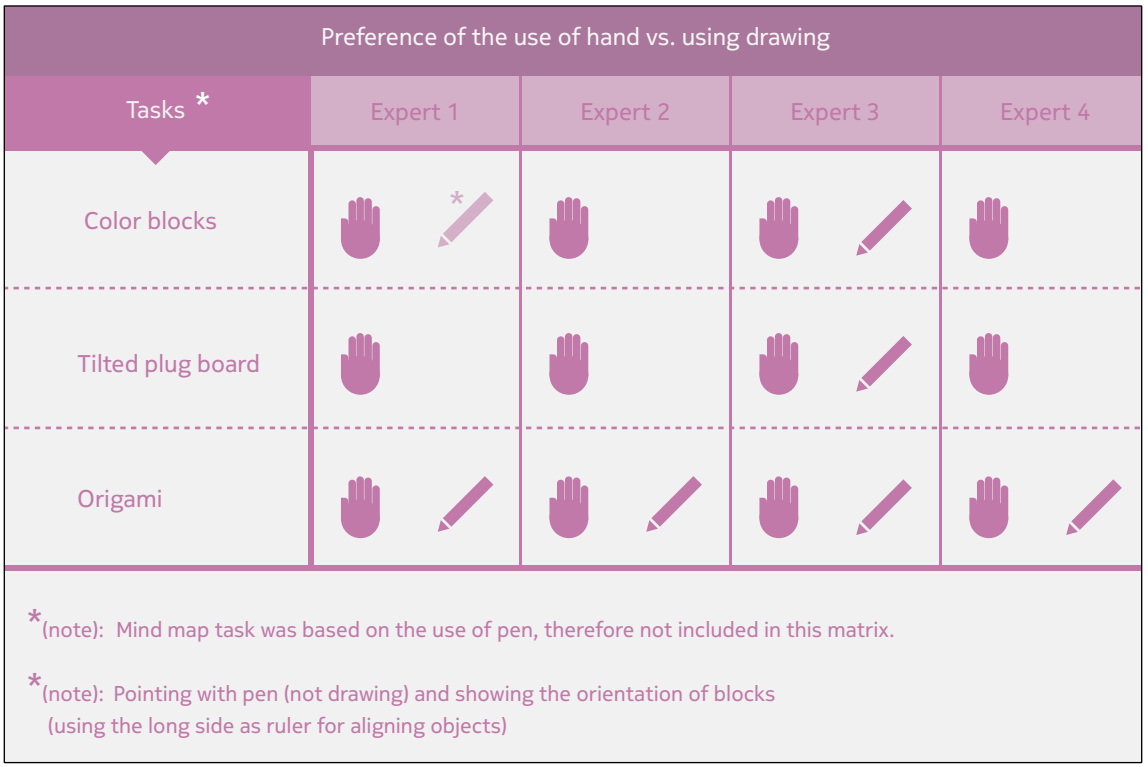


Figure 74: Preference of hand gestures versus using the drawing

Type of Gesture	Definitions
Deictic (Pointing)	**Expert 1:** OR Uses finger and pen (not drawing) to point at objects. **Expert 2:** TPT Uses finger to point at where the plug needs to go. OR Uses finger to show at corners and uses drawing for more complex instructions. **Expert 3:** CB TPT OR Uses drawing (marking objects with dots or circling them) **Expert 4:** OR Uses pen to point and to draw.
Concrete representational Iconic representations	**Expert 1:** TPT OR Shows orientation (placing hand where the object needs to be), rotation or flipping the object by using the hand as the representation of the object. Uses long side of pen (not drawing) as an 'aligning' tool. **Expert 2:** CB Uses hand to show where the block should be placed. OR Shows outlines with finger and pen. **Expert 3:** CB TPT OR Uses drawing to indicate appearance of the object (outlines (silhoutette) and foldings). **Expert 4:** TPT Uses hand to show the formation of pieces. OR Uses drawing to show edges, critical parts and foldings.
Spatial/Distance	**Expert 1:** CB Uses hand and pen (not drawing) to show placement of the object in realation to other objects. OR Uses finger to show corners. **Expert 2:** TPT Shows with hand where should be an empty row by movement of the finger. OR Draws a curve between two corners of paper to show that they need to be connected. **Expert 3:** CB TPT OR Draws arrows to indicate starting and end point of object to be moved or two corners which need to be connected during origami task. **Expert 4:** OR Shows by hand and by drawing, which corners need to be connected.
Kinetic/Motion	**Expert 1:** TPT Shows flipping tilted plug table, so that he sees the surface 'flat'. OR Shows flipping the paper with the hand movement. **Expert 2:** CB Shows how to turn the object with hand movement. **Expert 3:** CB TPT OR Uses arrows to indicate actions. **Expert 4:** OR Shows rotation (also flipping), grabbing and manipulation of paper.

CB Color blocks
TPT Tilted plug table
OR Origami

Figure 75: Mapping gestures and drawings according to Fussell's four types of gestures.

line stays there all the time and you can wait till it refreshes." (Pekka, expert user) "The drawing part was a key point here, otherwise I would never got which point he was talking about."(Anna, novice user) Drawing had advantages as:"it's small area where the origami is done, so it's easier to point at one small corner with pen than with hand which covers..." (Satu, novice user)

Second matrix where types of gestures are presented (figure), shows that each expert user used hand gestures or drawing or both in interaction at some point for indication from simple (pointing) to more complex (tridimensional hand movements e.g. flipping) and using drawings for detailed instructions. This indicates that such gestures were useful, each type of gesture was used by every expert user at some point (with variation depending on tasks), which on other hand indicates that gestures and drawings were natural way to provide guidance remotely on physical objects.

While the Fussell's matrix provides general guideline to catogorize gestures according to types, and most of gestures or drawings noticed during user tests fall into these categories well, some on other hand are might be considered less clear in relation to existing categories. Such gestures are tridimensional movement (showing movement not only on two axes, but three i.e. as tridimensional e.g. (figure)), drawing outline with finger or pen, circling critical objects and numbering them.

Advantages of gestures or drawings based on users comments

The relevance of the use of gestures or drawings remotely was commented by users with suggestions of possible use scenarios, where the focus was on educational and collaborational aspects: "Idea cloud - work around the world" and "You could work together and add things". On providing instructions remotely: "One could take an overall image of the situation. Another one showing how it can be fixed. " or "the person teaching could instruct, don't do like this. Instead do like this." And as "people don't know how to explain (..) even if it's a simple thing. (..) they could just show it, that would ease the thing a lot."

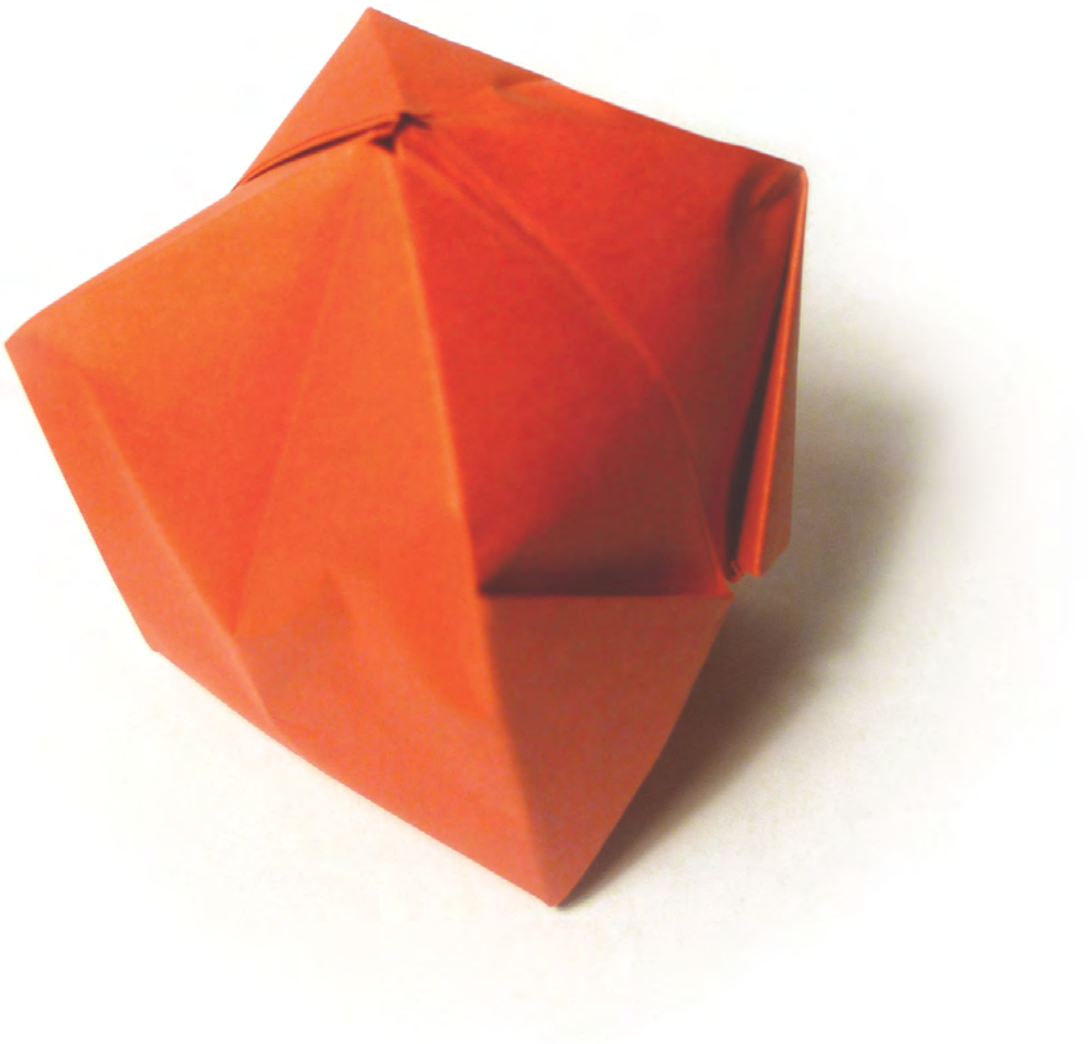


Figure 76: Photo of “water bomb” origami, folded by a user.

Chapter 7 :

Answers to research questions
Discussion
Future work and research

Answers to research questions

1.Do the interactions enabled by pico projectors contribute to the vision of ubiquitous environment ?

2.How gestures and drawings are used when matched with physical object in remote collaboration and guidance?

3.How could a projected image be matched with reality in remote guidance?

1. The vision of ubiquitous environment

"....you feel like you're simply working on document and not on a representation of a document." [6]

Pico projector technology in a way it was used in user test in this thesis, fits well into following sub categories: mobile technology, natural interaction and contextual sensitivity. Pico projector technology contributes to the vision of ubiquitous environment as it is mobile technology and could allow the use of projected interaction model, described in this thesis also on the go. In relation to two other sub categories of the vision of ubiquitous environment, projected interaction model is also especially beneficial in terms of natural interaction. According to one of the visions of ubicomp, the interaction with our everyday objects should be direct, where users could interact directly through physical artifacts. Results of this thesis are aligned with such previous research.

Previous results in augmented reality [18] [19] show that our physical environment can be augmented with digital information based on context and thus providing additional value to the user. While the prototype used in this thesis was fixed on the table and was tested in a controlled environment, it does show the possibilities of remote guidance on physical objects already, where an expert (whether that is human or system) can provide contextual guidance not only in fixed setups (as in most setups of previous research) but also in improvised environments and with improvised objects .

2. The use of gestures and drawings

Gestures were seen as a natural and inseparable part of verbal communication. Some of the users did not pick up immediately that they can use gestures to guide another person through the tasks on physical objects. Especially at the beginning users used gestures as in a normal conversation, without realizing that their hands are not seen by another user (hands were outside of projected area). Even though these were not seen by another user, they were seen as spontaneous and natural.

One of the users (novice) commented on the origami test: "Sometimes I felt that my hands were sort of an obstacle, when I was folding the origami I felt that my hand was in a way." Both of the users (expert for giving instructions and novice for working on origami) in this couple used their hands which blended in time to time as they were placed on top of the same object. Most likely, because the novice needed to hold down the origami in this case, he probably did not see that well instructions coming from the expert. Although hand gestures are seen as natural, the visual representation of the hand can be something else than a hand itself: "Previous results suggest that simple surrogate tools can be used to convey gestures from remote sites, but that the tools need to be able to convey representational as well as pointing gestures to be effective."[33]

Even this might lead to the question, whether the representation of the hand is necessary, user tests in this test underline the contextual importance of both hand and detailed instructions made by draw-

ing. Drawings were used especially in critical situations, where experts had difficulties in explaining instructions verbally or even when using hand gestures was not enough. As one of the users - an expert repeated many times: “It is really difficult to explain” and then added at some point “I will make a drawing or something”. Drawing was seen as positive and even though in some cases it was not used by all users, some expert users used hand gestures as if they would had been drawing. One of the users tried to show the outline of origami by ‘drawing’ its outline by repetitive movements of moving the finger along this outline. Such behavior could be replaced by drawn lines and symbols, which would stay as long as it needed and then erased. The couple who used mainly hand gestures mentioned that gestures could have been replaced by a “black marker”. Replacing gestures with graphical symbols has certain benefits as long as they can be erased when not needed any more [33]. This was also proposed during evaluation session with group of designers: “Lines and arrows are quite natural, you could draw them but after a while they could just vanish after few seconds.”

Expert users also used drawing as an instructions for themselves: “I’ll draw marks here for myself where are the corners of your paper”, which indicates that drawing can function also as reassurance of situation for both of users.

Semitransparent representation of hands or simple outlines could provide solution for two problems commented by users. First, it appeared as hands might have covered something important during interaction (e.g. when the hand of user was above the object, another particiapant could not see under this hand) and second, because the hands of users did “blend in” with objects, outlined version could have been seen better.
Drawn lines were seen not only well, but were also critical and “key point” moments in communication.

One of the issues in using drawing in remote guidance should address different levels of drawing abilities of individuals. These user tests suggest that it greatly depends on a person how pro-active he or she can be with picking up the pen to draw. While hand gestures are very natural for us, picking up a pen can be a threshold question and might be less natural for some of us than others. This should be addressed by merging drawing abilities and using natural hand-driven interaction in the future.

3. How the projected image matches reality in remote guidance

The prototype described in this thesis provides a simple approach in matching projected image with physical objects. The projector and web camera were positioned above the table to project directly downwards. This was a deliberate decision, in order to simplify the focusing issues.

Drawing on 3D objects with this setup was experienced as “challenging” by some expert users because of the focusing issues. The projection/image did not “fall onto the 3D too easily”. The need for autofocus was mentioned also after the first, mind map task, as when the papers were moved, writings became misaligned (local vs. remote paper). The tilted plug table was especially difficult to be projected on as half of the experts asked the novice to

turn the table so that it faces the camera at the direct angle. While this was likely not only because the tilted surface was difficult to project to, but also because when facing camera, it appeared narrower in comparison to when it was put ‘flat’.

As these focusing issues arised during user tests with fixed setup, they need to be addressed as remote interaction on physical objects can also take place with non-fixed setups, where projectors and cameras are not fixed and are moved by the users.

Misalignements of projection and physical object, experienced during fixed setup general could be addressed by solutions described in practical limitations chapter (page 65). The use of markers familiar from augmented reality or tags, could provide short-term solution - a tool to align expert’s and novice’s working spaces in fixed and non-fixed interactions.

Discussion

As previous research shows, remote collaboration needs contextual approach. Results from remote collaboration with physical objects show that the focus should be on the objects rather than the faces. Results from research on remote guidance and expertise show that remote expertise increases speed and accuracy of users[29]. Results from the use of hand gestures and drawings show that such information is seen as an additional and helpful. Based on this information, the goal of this thesis was to contribute to remote collaboration with physical, tridimensional objects through interaction based on hand gestures and drawing.

At the same time, high level theoretical part of the thesis - the vision of a ubiquitous environment - was used to reflect findings from user tests to and discuss whether the projected interaction model based on gestures and drawings with physical objects contributes to the idea of a ubiquitous environment. To reflect on the findings, mobile technology, natural interaction and contextual sensitivity were picked as sub categories from the matrix of the vision of a ubiquitous environment.

Sufficient level of projection

Although the prototype used in the user tests of this thesis was robust and the technology was not refined, the real focus is on benefits of the interaction experienced during user tests. While our prototype proved to be challenging with high detailed tasks (origami), its technical capabilities were sufficient in interaction on larger physical objects (color blocks and white section of tilted plug board).

Conversational grounding

Previous research has shown that "interpersonal communication is demonstrably more efficient when people share greater amounts of common ground—mutual knowledge, beliefs, goals, attitudes, and so on. [33] Results of users tests described in this thesis indicate that gesture-based communication is very natural even when it happens remotely. Both users, experts and novices used simple (pointing) and more complex gestures (indicating movement) to communicate and establish commong ground ("this is the object we are talking about"). User tests of this thesis showed that remote collaboration with physical objects can be difficult when users lack proper conversational grounding. In such situations gestures and drawings can provide valuable and supportive information in multimodal interfaces [34]and help to avoid errors.

Control of the view and projection

Previous research by Ranjan [36] indicates that there is a need for the user "to control the view of the workspace." Feedback from the users in this thesis is aligned with such findings. Users commented that perspective view (comments during review of observation videos) makes them able to take another angle and to see better. "Now that I see this from a different angle, this is so much easier to see what's going on." (expert user) "So if it was a perspective from the room it would be easier"(novice user). "As the image was projected straight from the top, there was not sense of depth (..) if you'd sat beside the person who was working on the origami, it would be easier to see." (expert user)

Similar ideas were brought forward during evaluation with the designer group. Users also commented on the lack of the sense of depth in projection. As this was due to the direct projection from the top it could be corrected with giving users better control over projection and camera.

Results from user tests, discussions with designers during evaluation session imply that experts should have better control over camera and projector. Experts should share the same view as novices and they should project guidance from the same direction of the eye-level of the novice i.e by using orthographic projection.

Showing by doing

Most likely some tasks would have been easier to show directly how it is done by example (e.g. expert and novice folding origami at the same time) and that "It would be easier to follow" and "helfpul" as commented by some of the users. Even though this was not the focus of this thesis, it might be considered valuable to address this need in the future development of these systems.

Pair remote collaboration and group collaboration (education)

Previous research indicates that current strategies on collaboration on phones (with displays)can be categorized as follows: "(1) one person controls the phone and verbalizes information for others, (2) one person holds up the phone for others to view, (3) the phone is passed around, and (4) someone shares a link or reference and others view the information on their own phones."[35] In respect to this, projector technology provides a completely different approach in terms of scalability of the projected area and in terms of usage of physical environment for collaboration.

In this thesis, remote collaboration and guidance with projector technology was tested between two people. Users' feedback and feedback during evaluation with the designer group suggest that this kind of interaction could be used in education and "remote lectures". While the voice is seen as enough for simple tasks, projector can be used to show details "I mean this one".

In relation to education, projector technology could especially enable simultaneous guidance of several people, where the expert works in a lecture mode and switches between users to check the progress and provide guidance. Another natural direction would involve group work with physical objects with comparing the progress and assisting one another during the task. Therefore, future research should address how remote collaboration with physical objects would function between larger groups, what such collaboration would mean in terms of roles (expert versus novice) and how gestures would work when more than one person might provide guidance at the same time.

Advantages of projected interaction

"It was fun, when the feeling disappeared that we were working on two separate sheets of paper." Another user: "Kind of Nintendo wii feeling" And: "There was like interaction going on. I can imagine if we were to design something and another person could improve it: 'how about this'."

Users also commented on projected interaction as “a fun way of working”, where they felt as if they workded on the same task and did not therefore feel alienated. Users appreciated, in general, visual explanations (gestures and drawings) as they helped them to communicate through difficult tasks.

Design recommendations for projected interaction model

Results of the user tests show that projected interaction model when compared to design recommendations by Wickey [32] does not match all the requirements yet, however it already supports (1) “easy composition of gestures”, (3) “gesture input in remote gesture system which is located where objects are displayed”. The second recommendation (2) where “remote gesture systems should integrate visual control functions” and fourth (4), where “gesture output in remote gesture systems should be semi-transparent” were not met with our prototype, but the feedback from users was in alignment with these recommendations. The extension of these recommendations in relation to Fussell’s matrix [33] would also be to focus on the extension of hand gestures; to provide the possibility to replace representation of hand with drawings or simplified (graphical) representation of hand or fingers. As the drawing lasts until it is either automatically or manually erased, this would provide a certain advantage in comparison to rapid hand movements.

Future work and research

Remote help by human or system

User tests described in this thesis did concentrate on remote human-to-human collaboration and guidance. Future research in remote collaboration and guidance with physical objects should also address system-to-human guidance. User tests communicated latent needs for such system, as some of the users suggested, that “you would not need necessarily the second person.” It was implied that remote collaboration could work in pre-recorded mode, where projector could provide automatic instructions.

In order for a system to project the content on physical objects, it needs to “understand” what is the context, what are the objects and what is the environment like. Augmented reality could provide an approach for the system recognition of the environment (with markers, tags, beacons or sensors) in short-term but in order for the system to understand more complex environment this would either mean to increase the amounts of sensors, tags etc. or to take different approach through virtual, 3D models.

In order for a system to provide contextual information of the environment to the user, the system needs to be aware of this environment as a tridimensional space. One way to achieve this would be to build a virtual 3D model of the physical environment or an object and establish a live link between digital and physical environments, where changes made to physical object would result in changes to digital object. When the link between digital and real spaces is achieved, the system can use physical environment with the objects in it to provide contextual information with the help of projector(s).

Early tests conducted with the help of Sami Sorvali from Mural Media (full text in appendix) showed that the pico projector has sufficient projecting abilities to use simple physical objects and their tridimensional forms to be used not only as a single surface but as separate surfaces so that the user but also a system could point at, highlight etc. separate facets or parts of the object and so that digital information is projected without distortion or misplacement. This provides interesting insights on possibilities for automatic context sensitive instructions and guidance by a system, taking place on actual physical appearance, form of the object or environment.

Voice-only and projected interaction only - interactions

Remote communication and interaction described in this thesis is based on a combination where voice communication complements visual interaction and vice versa. The idea of the need of control groups such as voice only and projected interaction without voice was raised during evaluation with designer group. While the establishment of such control groups was not possible in this thesis, this idea should be addressed in future work. The communication and interaction, based on visuals alone, might however require totally new approaches, as it will become extremely important that "gestures are represented and commonly understood in a mixed ecology."[34]

Support for tridimensionality of gestures

Users used gestures which had a sense of tridimensionality i.e users moved their hands not only on x and y axes (flat surface), but on all three x,y,z axes (flipping, turning the object, opening up the fold of origami). Support and future research in relation to the tridimensionality could provide interesting insights and findings.

Conclusions

The ubiquitous environment was described in the theoretical part of this thesis with attention to technological development paths (the vision of ubiquitous environment). Sub-categories from these development paths (mobile technology, natural interaction and contextual sensitiveness) were picked to help to review projector technology and results from user tests in relation to ubiquitous environment. Also, pico projecting technology was reviewed in relation to the previous research in augmented reality and remote collaboration.

User tests conducted and described in this thesis showed that pico projector technology contributes to the building up of a ubiquitous environment. The pico projector is mobile technology, and it enables natural interaction with physical objects through the support of gestures. Prototype used in this thesis provides an interaction model, where gestures and drawings are used according to the context (the object).

The prototype described in this thesis, helps in establishing live link for remote collaboration and guidance with physical objects between two remote users (expert and novice) with the help of projected interface. Results based on user tests of this thesis show that the use of gestures and drawing in general can provide valuable help in remote interaction. Gestures are helpful not only in providing additional information, but they can help in establishing conversational grounding between remote users and help in critical situations which users if relying on verbal communication alone, would likely fail to overcome. Results show that gestures and drawings can be beneficial in remote collaboration with physical objects. The use of hand gestures and drawings provides additional value as users can rely also on natural interaction rather than only on verbal. Interaction with the help of projected interaction is seen as 'fun' and it increases the feeling of togetherness - all valuable points in terms of further studies of this interaction.

The future and research could address the aspects of mobility in projected interaction in contrast to fixed setup, used in this thesis. Also, another natural step for further studies would be to focus on a machine-based or system-based guidance, as live link with another person might not always be possible or necessary (simple tasks) and where machine or a system could provide reasonable, alternative approach.

Whle the prototype used in this thesis was not technologically sophisticated (built from off-the-shelf technology), it does already provide sufficient support as such for remote communication with gestures and drawings on physical objects. The prototype described in this thesis is based on existing technology and can already be integrated as an additional tool in existing mobile phones.

The contribution to Nokia through this thesis is to provide an overview of possibilities of projected interaction with gestures on physical objects and to raise attention and excitement towards pico projecting technology in future research and product development processes.

Appendix
Prototype specifications
Mapping test
References

Prototype specifications:

Pc 1:
Lenovo W510
OS: Windows 7
Aspect ratio of display 16:9.
Pc 2:
Fujitsu Siemens Amilo
OS: Windows XP
Aspect ratio of display 3:2.

Projectors (x2):

Specification:
Aiptek Pocket Cinema V50
Optical technology: DLP
Light source: RGB LED
Brightness: Up to 50 Peak Lumens (ANSI Lumens: 40 Lumens)
Aspect Ratio: 16:9
Resolution: 854 x 480 (WVGA)
Projection Image Size: 25.2cm ~ 216.3cm (9”~85”diagonal)
Projection Distance: 35cm ~300cm
Contrast: 1000:1

Web cameras:
Logitech HD Webcam C270
Logitech HD Webcam C310

Communication Software:
Logitech VidHD

Headsets (x2):
Logitech (USB)

Mapping test

The findings from user tests, in combination with discussions on projecting and three-dimensional objects raised following concerns and possibilities:

1.Projected lines not matching with object and falling off the boundaries of the object require solutions.

2.Hands were sometimes “in the way”, thus more graphical approach could be more advantageous.

Also due to the “lag” in video, graphical input was seen as benefit.

3. Tridimensionality of both objects and gestures provides interesting possibilities

Partly based on earlier expectations and also greatly based on results from user studies, simple mapping test was done with the help of Sami Sorvali from Mural Media in order to test whether mapping the texture onto physical object could provide a more accurate way for placing instructions on top of objects with the help of pico projector.

As a result we found out that each facet of the object can be used for the projection of graphical symbols (e.g. letters,colors, symbols etc.) separately and that pico projector provides sufficient luminance. We found out that small simple objects (here: 4 x 4 c 4 cm) can also be used in such a way that their visual appearance can be augmented or changed. This finging can be especially interesting in terms of augmenting physical appearance of the object with affordances, for instance to indicate where the action should be applied (e.g.

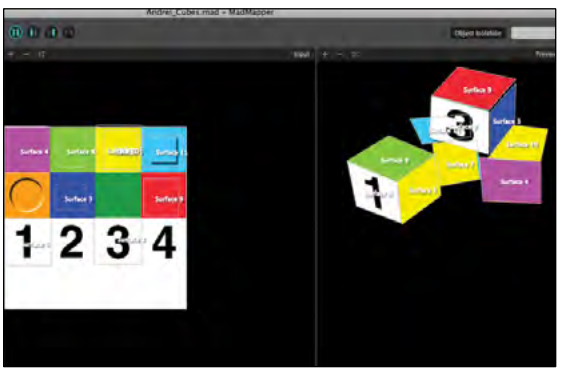


Figure 77: Software screenshot (mad mapper)

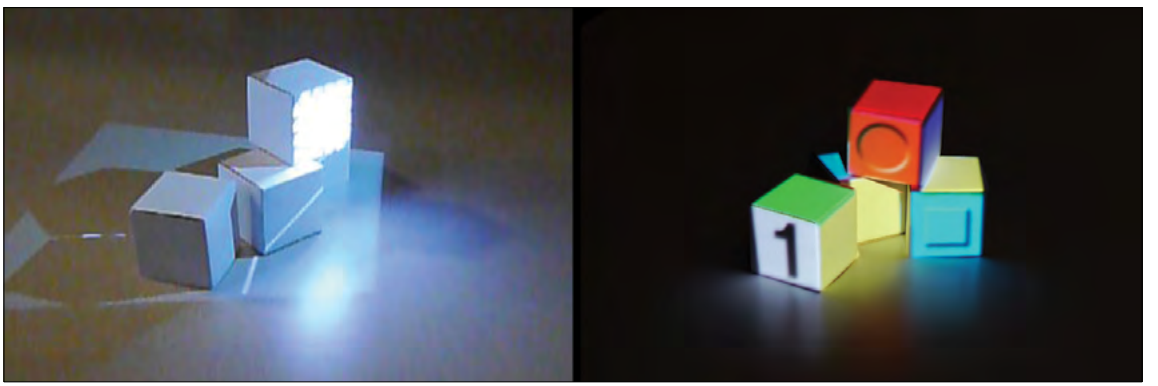


Figure 78: Beginning of process(left). Final result (right) *note that image was not taken from orthographic angle (same as projection) with the result of visible shadows (left corner of the red cube).

push this). This can work on an otherwise empty surface such as wall, where the action could be hidden and revealed with the entrance of someone into a room, or the action could depend on the other preferences or rules.

Therefore, with this setup the system can suggest and guide the user by projecting graphical symbols exactly where they are needed. So instead of suggesting "take this cube and rotate it...", the system can suggest "place this facet" (points at facet) so that it is facing upwards". Or the system can show the affordance (a facet in a form as button) to be pressed. This mapping test suggests, that such abilities of the system to 'understand' and provide contextual information based on real properties and appearance of physical objects and physical environment can be especially valuable when it comes to help in remote guidance.

As the mapping test described here was done on a rather simple level with simple objects - to fully understand the capabilities of this approach, it needs to be addressed with further studies.

References

[1] Tangible Interfaces for Remote Collaboration and Communication, Scott Brave, Hiroshi Ishii, and Andrew Dahley, MIT Media Laboratory, 1998

[2] Älykkäiden ympäristöjen suunnittelu: Kohti ekologista systeemiajattelua
Eija Kaasinen ja Leena Norrsos (toim.) Tammer-paino Oy, Tampere 2007

[3] Some computer science issues in ubiquitous computing, Mark Weiser, ACM, 1993

[4]The Computer for the 21st Century. Mark Weiser. 1991

[5] Human–computer interaction issues for mobile computing in a variable work context," J. York, P.C. Pendharkar, Int. J. Human-Computer Studies 60 (2004) 771–797)

[6] Everyware, Adam Greenfield, 2006, New Riders

[7] Human-Centered Design Considered Harmful. Donald A. Norman. 2005

[8] Shaping Europe's Future through ICT " ISTAG, 2006

[9] Issues and Challenges in Ubiquitous Computing, Lyytinen and Youngjin Yoo,

[10] Where the Action Is: the foundations of embodied interaction, Paul Dourish, 2001.

[11] Disseminating Active Map Information to Mobile Hosts, Schilit, B., Theimer, M. IEEE Network, 8(5) (1994) 22-32

[12] Towards a Better Understanding of Context and Context-Awareness, Anind K. Dey and Gregory D. Abowd, Georgia Institute of Technology,

[13] Context-Aware Computing Applications, Bill Schilit, Norman Adams, Roy Want, Computer Science Dept, Columbia University, Palo Alto Research Center, Xerox Corporation)

[14] http://electronics.howstuffworks.com/ dlp1.htm

[15] Taxonomy of Mixed Reality Visual Displays IEICE Transactions on Information and Systems, P. Milgram and A. F. Kishino, E77-D(12), pp. 1321-1329, 1994.

[16] Situated information spaces and spatially aware palmtop computers. George W. Fitzmaurice. Published by ACM 1993

[17] A Taxonomy of Real and Virtual World Display Integration. Paul Milgram, Herman Colquhoun Jr.

[18] A Survey of Augmented Reality, Ronald T. Azuma, Hughes Research Laboratories, 1997)

[19] Projected Augmentation - Augmented Reality using Rotatable Video Projectors. Jochen Ehnes, Koichi Hirota, Michitaka Hirose. 2004

[34]The role of voice in human-machine communication. Philip R. Cohen and Sharon L. Oviatt, Department of Computer Science and Engineering, Oregon Graduate Institute of Science and Technology. 1995

[21] ASTOR: An Autostereoscopic Optical See-through Augmented Reality System. Alex Olwal, Christoffer Lindfors, Jonny Gustafsson, Torsten Kjellberg, Lars Mattsson. 2005

[22] http://www.media.mit.edu/research/groups/1461/sixthsense

[23] Computer supported cooperative work: new challenges to systems design. Carstensen, P.H.; Schmidt, K. (1999). Retrieved 2007-08-03.

[24] Blended Interaction Spaces for Distributed Team Collaboration, Kenton O'Hara, Microsoft Research, Jesper Kjeldskov and Jeni Paay, Aalborg University.

[25] Coordination of Communication: Effects of Shared Visual Context on Collaborative Work, Susan R. Fussell, Robert E. Kraut and Jane Siegel, Proceeding CSCW'00 Proceedings of the 2000 ACM conference on Computer supported cooperative work, 2000)

[26] Affordance, Conventions and Design. Donald A. Norman. 1999

[27] One is not enough: Multiple views in a media space. Gaver, W., Sellen, A., Heath, C., & Luff, P. Interachi '93 (335-341). NY: ACM Press.

[28] Turning away from talking heads: The use of video-as-data in neurosurgery. Nardi, B., Schwarz, H., Kuchinsky, A., Leichner, R., Whittaker, S. & Sclabassi, R. (1993). T. Proceedings of Interchi '93 (327-334). NY: ACM Press.

[29] Collaboration in Performance of Physical Tasks: Effects on Outcomes and Communication. Robert E. Kraut et al., 1996.

[30] Turn It This Way: Grounding Collaborative Action with Remote Gestures. David Kirk,

Tom Rodden and Danaë Stanton Fraser. CHI 2007 Proceedings

[31] Gestures over video streams to support remote collaboration on physical tasks. Fussell, S. R., Setlock, L. D., Yang, J., Ou, J., Mauer, E. M., & Kramer, A. Human-Computer Interaction, 19, (2004). 273-309.

[32] Analysis of Hand Gestures in Remote Collaboration: Some Design Recommendations. Aiden Wickey, Leila Alem. 2007

[33] Gestures Over Video Streams to Support Remote Collaboration on Physical Tasks, Susan R. Fussell et al. HUMAN–COMPUTER INTERACTION, 2004, Volume 19, pp. 273–309

[34] Ways of the Hands. David Kirk, Andy Crabtree and Tom Rodden, School of Computer Science & IT, University of Nottingham, Nottingham, UK. 2005

[35] ShadowPuppets: Supporting Collocated Interaction with Mobile Projector Phones Using Hand Shadows. Lisa G. Cowan and Kevin A. Li. CHI 2011 Session: Interaction on Mobile Devices

[36] An Exploratory Analysis of Partner Action and Camera Control in a Video-Mediated Collaborative Task. Abhishek Ranjan, Jeremy P. Birnholtz, Ravin Balakrishnan (2006): In Proceedings of CSCW 2006, Banff, Canada, pp 403-412.