Mixed Reality: A Survey

Book Chapter

Author(s):
Constanza, Enrico; Kunz, Andreas; Fjeld, Morten

Publication date:
2009-03-27

Permanent link:
https://doi.org/10.3929/ethz-a-005794342

Originally published in:
Lecture Notes in Computer Science 5440, https://doi.org/10.1007/978-3-642-00437-7_3

Mixed Reality: A Survey

Enrico Costanza[1], Andreas Kunz[2], and Morten Fjeld[3]

[1] EPFL Media and Design Laboratory, Lausanne
`design.epfl.ch`
[2] ETHZ Innovation Center Virtual Reality (ICVR), Zurich
`www.icvr.ethz.ch`
[3] Chalmers TableTop Interaction Lab (t2i lab), Gothenburg
`www.t2i.se`

Abstract. This chapter presents an overview of the Mixed Reality (MR) paradigm, which proposes to overlay our real-world environment with digital, computer-generated objects. It presents example applications and outlines limitations and solutions for their technical implementation. In MR systems, users perceive both the physical environment around them and digital elements presented through, for example, the use of semi-transparent displays. By its very nature, MR is a highly interdisciplinary field engaging signal processing, computer vision, computer graphics, user interfaces, human factors, wearable computing, mobile computing, information visualization, and the design of displays and sensors. This chapter presents potential MR applications, technical challenges in realizing MR systems, as well as issues related to usability and collaboration in MR. It separately presents a section offering a selection of MR projects which have either been partly or fully undertaken at Swiss universities and rounds off with a section on current challenges and trends.

Keywords: Human-computer interaction (HCI), Mixed Reality, Displays, Sensors, Information Visualization, Usability, Switzerland.

1 Introduction

The ready availability of large amounts of computational power in small devices and their constantly decreasing cost paved the way for the concept of “Ubiquitous Computing” [1]. In Weiser’s vision, the goal was to make computational power available to people wherever and whenever they need it, not only at the desktop. This could be in meeting rooms where one might need to retrieve information in order to better contribute to discussion. Other places may include the car, to help us drive more efficiently and safely, a surgeon’s operating room, or a designer’s drawing desk.

How can we integrate this new group of computational devices into the environment? A number of different paradigms have been proposed to answer this question and to move interaction from the computer box into the world. This chapter presents an overview of the Mixed Reality (MR) paradigm, which proposes to overlay our real-world environment with digital, computer-generated

D. Lalanne and J. Kohlas (Eds.): Human Machine Interaction, LNCS 5440, pp. 47–68, 2009.

objects. It presents example applications and outlines limitations and solutions for their technical implementation.

MR was derived both conceptually and historically from Virtual Reality (VR). VR systems are computer systems in which users are immersed in a virtual, computer-generated world. The very first examples were originally developed in the 1960s [2]. Immersion is generally achieved through visual, auditory, and sometimes tactile displays. All these displays isolate users from their familiar surroundings, giving the illusion that the only objects existing around them are those rendered by the computer. In MR systems, users perceive both the physical environment around them and digital elements presented through, for example, the use of semitransparent displays. Imagine a system that indicates the name and provenance of items around you by displaying virtual labels overlaying the objects, or a system that guides your way by showing virtual arrows, or a system that displays people's names and affiliations on virtual badges. The information could be displayed in the native language of each user or could be customized to be most relevant to their individual profile; for example, when browsing food products, specific information could be provided according to the user's allergies.

MR systems are designed to give their users the illusion that digital objects are in the same space as physical ones (Figure 1). For this illusion of coexistence, the digital objects need to be precisely positioned into the real environment and aligned with the real objects in real time [3]. In fact, the precise real-time alignment or registration of virtual and real elements is a definitive characteristic of augmented reality systems [3], and it constitutes a difficult technical challenge for its realization. Augmented reality is often considered to be a branch of MR. According to the definition of Milgram et al. [4], MR is "subclass of VR related technologies that involve merging of real and virtual worlds". MR includes systems in which the virtual aspects are dominant as well as those in which the physical reality is dominant. Within this range, augmented reality has more physical elements than virtual elements.

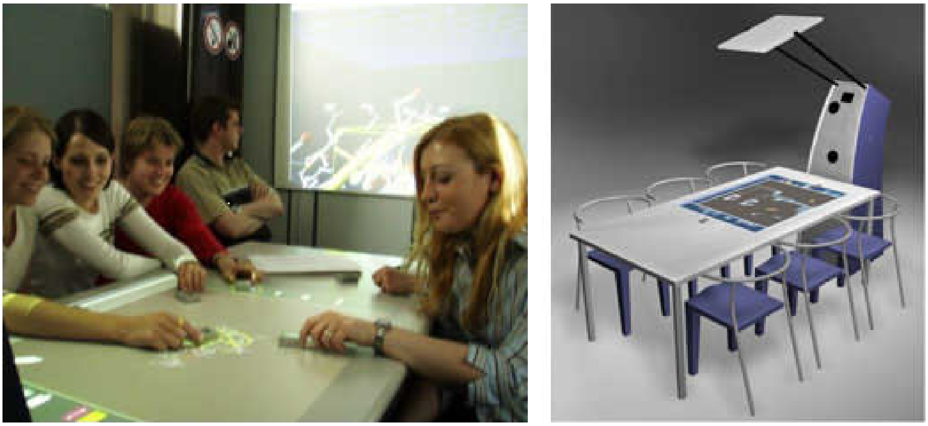

Fig. 1. The BUILD-IT system, an example of a collaborative tabletop MR application

The following section presents a section on potential MR applications, followed by a section on technical challenges in realizing MR systems. The next section presents issues of usability and collaboration related to AR. A separate section offers a selection of MR projects which have either been partly or fully undertaken at Swiss universities. The chapter rounds with a section presenting some current challenges and trends.

2 Applications

By its very nature, Mixed Reality (MR) is a highly interdisciplinary field engaging signal processing, computer vision, computer graphics, user interfaces, human factors, wearable computing, mobile computing, information visualization, and the design of displays and sensors. MR concepts are applicable to a wide range of areas including the automotive industry, surgery, and office environments. Other examples include the maintenance and repair of machinery; instruction notes could be displayed next to the relevant location, as if they were real, physical labels. Steve Feiner's team at Columbia University was the first to demonstrate such a scenario in 1993 [5] by developing one of the earliest MR prototypes: a system to guide end-users in basic maintenance operations of a laser-printer. Through a monochromatic, semitransparent, head-worn display, users see wire-frame computer graphics highlighting specific components of the printer, and text labels indicating how to disassemble the device and replace parts. Recently, Lee and Rhee [6] presented a collaboration-oriented, distributed MR system for car maintenance. Their system includes mobile as well as desktop terminals, connected to a server and an ontology-based context recognition system to render the information in the format appropriate to the client and the situation. Other examples in the field of manufacturing include a system to support the task of car door assembly [7] and a tool to evaluate the placement of new machinery or workstations inside an existing manufacturing plant [8]. In the latter case, the main advantage offered by MR is that the position of new pieces of industrial equipment can be visualized on real images of an existing plant, and the suitability of the placement can be evaluated by visual inspection, determining whether the new tools are within reach or conflict with older ones. Thus it is not necessary to create a virtual model of the entire production plant, but only of the new items.

The Magic Book [9] is a system built to visualize virtual three-dimensional (3D) models on the pages of a physical book. This book acts as a handle for the virtual models: by moving the book, users can move the models and look at them from different viewpoints. Proposed applications for the system are the visualization of interactive 3D children stories and geological data, as well as architectural models. Klinker et al. [10] applied the magic book paradigm to the visualization of new car prototypes in their Fata Morgana proof-of-concept MR system. Fata Morgana was developed in collaboration with an automobile manufacturing company and the system was evaluated by professional car designers.

In the medical field, MR systems can be used to visualize medical imaging (such as CAT scans, MRI, or ultrasound) directly on the patient's body in order to guide the surgeon's action [11,12,13]. Medical images are already available in digital formats, and they are currently displayed on standard monitors in the operating room. A user study of needle biopsy on mockups showed that MR can improve accuracy compared to traditional methods [14].

Remote collaboration is another application area of MR. In typical scenarios, this involves an operator in the field receiving guidance from a remote expert. The operator uses a mobile MR system to capture the scene around her and send it to the expert's system. The expert can see the scene the operator is in and give her instructions using an audio channel or visual annotations displayed on the operator's MR system. Initial examples were also developed at Columbia University [15]: in a demonstrative system, one user is free to roam freely on the university campus, while someone else can add and manipulate virtual objects in the visualization in specific locations. In the medical domain, Welch et al. [16] proposed a system that uses multiple cameras to capture a patient's body which could then be visualized for a remote expert using a high resolution static display or PDA.

MR systems have been proposed to provide navigation guidance. In this scenario, users can see virtual signs anchored to the physical world. Similar to a compass, the signs indicate the correct direction regardless of the device's orientation. A potential application would guide soldiers in an unfamiliar environment [17,18] and provide information about known sources of danger. Yet another would guide tourists in a city [19] or visitors inside a building [20].

A number of entertainment applications were proposed, in which users have to interact with virtual characters or devices appearing in their physical environment. In general, MR games can increase collaboration or competition among players, who can be co-located or remote. Examples include a MR version of the Quake videogame [21], in which users see monsters from the game as well as virtual walls appearing in the physical environment. Players can shoot at the monsters as they would do in the normal game. The Human Pacman game [22] is a version of the popular arcade game transposed into a real city. Players are equipped with wearable computers and Head-Mounted Displays (HMDs). They have to roam the city searching for physical and virtual items to collect and chasing one another. In the Moon Lander game [23], players have to land a virtual spaceship on a real outdoor location. Examples of MR games in indoor settings include a MR Mah-Jongg game [24] and MonkeyBridge [25], in which players have to place virtual objects on a physical table in order to guide the path of virtual characters.

3 Technical Challenges

Mixed Reality (MR) poses a number of demanding technological requirements for its implementation. One challenge is related to the display technology, which must visualize digital objects at high resolution and high contrast. Precise position tracking constitutes another significant challenge. In order to give the illusion that virtual objects are located at fixed physical positions or attached to physical items, the system must know the position of relevant physical objects relative to the display system. In some cases, depending on the type of display being used, the user's point of view (in terms of their position and the direction of their gaze) is also of interest. The following two subsections provide an overview of display and tracking technologies used to implement MR systems and their known limitations.

Most of the MR technologies require the system to know the location of the objects to be mixed and the location and orientation of the display, or, at least, the location of the objects relative to the location of the display. It is important to emphasize the need for both the position and the orientation of the display in all 6 degrees of freedom. In some situations, the tracking system can be physically attached to the display, and so the user wearing such a display can also be tracked.

3.1 Displays

This section gives an overview of displays most commonly used in MR environments. These are Head-Mounted Displays, hand-held displays, ambient projections, and hand-held projectors.

Head-Mounted Displays. Head-Mounted Displays (HMDs) are probably the most common type of displays used in MR. HMDs were originally developed for Virtual Reality (VR) systems. They consist of one or two visual display units together with optically compensated systems that form a perspectively correct virtual image, even though the display is very close to the user's eyes. HMDs developed for VR let the user perceive only what is shown on the display and so do not provide any see-through capability. However, for MR the virtual imagery needs to be mixed with imagery of the surrounding environment. This can be achieved by means of a video camera physically attached to the HMD. The camera's captured image is electronically combined with the synthetic images to create a MR. Another technical solution is to use semi-transparent mirrors for an optical combination of physical and virtual elements. The first type of HMD using a camera is known as a video see-through HMD, while the latter is called an optical see-through HMD. A special technical realization of an optical see-through HMD uses two consecutive LC-panels: one for image generation, i.e. for displaying the virtual objects, and the second for blanking out the real world (non-see-through) or showing the real environment (optical see-through).

Current off the shelf [26] HMDs allow a field-of-view of 45 degrees diagonally (36 degrees horizontally and about 27 vertically), a resolution of 1280 by 1024 pixels, and a weight of about 750 grams. The display produces the impression of an 80" screen positioned at about 2 meters from the user. In general, critical features of HMDs are their weight, resolution, and field of view. The use of an HMD requires the tracking of the user's head position and orientation so that virtual images can be rendered from the correct viewpoint. HMD prices vary widely, depending on their features.

Hand-held Displays. Hand-held displays are used for MR by using the metaphor of a magic lens, through which a reality can be seen that is enriched by

Fig. 2. The operator views the machine operation through the holographic optical element (HOE), which is illuminated with stereoscopic images from the projectors driven by a PC. The setup allows 3D annotation to appear in the workspace, augmenting the operator's view of the process with relevant information [33].

virtual elements. The mix of real and virtual images is achieved using cameras attached to the displays (video see-through). Similar to HMDs, the position and orientation of hand-held displays must be known in order to correctly generate virtual images. Hand-held displays are normally less expensive than HMDs as there is no need for optical compensation. An early example of hand-held MR was presented by Rekimoto [27] using custom-built hardware, while recent examples employ commercially available mobile phones and PDAs [28], which creates a great potential for mass adoption of this type of display. Such use of mobile phones is shown in an interactive road map application [29] (Figure 3).

Ambient Projectors. Rather than addressing a user's perception through a display, be it head-mounted or hand-held, an alternative is to project computer generated images directly onto the environment using standard video-projectors. The projection can be confined to a specific area, such as a desk [30,31], or it can cover an entire room using an actuated mirror to direct the video beam [32]. In both cases, the system needs to track the position of objects in the mixed environment to be able to display virtual information next to or onto them. Projecting on an entire room or onto special objects requires a 3D model of the entire space. This allows the distortion of the projection in order to fit the images to projection surfaces that are typically not perpendicular to the projector [33] (Figure 2). Unlike HMDs and hand-held displays, the positions of the projectors are fixed or controlled by the system. This reduces the tracking requirements but, typically, also the user's freedom of movement.

Hand-held Projectors. The recent miniaturization of video projectors suggested their use as hand-held MR displays [34]. Users could use these projectors to directly point at objects of interest. This allows the direct projection of the computer-generated information onto the object or next to it. These types of displays require information about the position of objects in the environment relative to the projector and also the orientation of the surfaces onto which the information should be projected. With this information, the computer-generated image can be projected perspectively correct onto the objects of the environment. Compared to hand-held displays, these systems require more complex and expensive hardware, but they can create a larger display surface and allow multiple users to interact more easily with the system.

3.2 Registration

In principle, a tracking system is a device that can determine the position and orientation of a body and interpret it. In order to create a realistic virtual environment, the computer must utilize these systems to acquire this information about the user. Tracking systems can be classified as either active or passive. Within a passive tracking system, the object to be detected does not need any special device, but, rather, it is surveyed by sensors from a distant location. These types of tracking systems very often either have a limited resolution or the effort required for a precise detection is great. Thus, the advantage of being unhindered by cable connections must be paid for by the high installation costs for such a system. Because of these reasons active tracking systems are very often used. Within these systems, the object to be tracked must be active, that is, a sensor is directly attached to the object. Active tracking systems use very different working principles, of which the most important ones will be described here. This section presents an overview of tracking systems most commonly used in MR environments: Global Positioning System (GPS), visual markers, acoustical tracking systems, magnetic and inertial sensors, and hybrid systems.

Global Positioning System. Global Positioning System (GPS) receivers use radio signals broadcasted by a number of medium earth orbit satellites to calculate their location [35]. Each satellite continuously transmits messages about its position, the position of other satellites in the system, and the time when the message was sent. Receivers use the difference in the messages' time of arrival from 4 or more satellites to calculate their location. GPS was originally developed by the US Ministry of Defense. Today, it is still used for military purposes, but also for the navigation and guidance of civilian vehicles like airplanes, ships, and cars, as well as in outdoor mobile MR systems, in combination with another system to provide orientation. Since the system is based on the timing of radio signals, the sensitivity and accuracy of the receivers can have a big influence on the resolution of the positioning [23]. Local radio transmitters can be used in addition to the satellites to improve accuracy. However, this requires expensive installations. GPS signals propagate in line-of-sight and they are highly attenuated by buildings, making the system generally non-functioning when the receiver does not have a clear connection with a minimum amount of satellites, perhaps, indoors or near high buildings. The radio reception near buildings can also vary depending on the time of the day [36], making the situation even more problematic.

Visual Markers. Visual markers, sometimes referred to as fiducial markers, are graphic symbols designed in combination with a computer vision recognition algorithm to yield high probability of recognition and low probability of misclassification [20,37,38]. They can be read using a standard video camera connected to a computer. Then, generally, algorithms enable calculation of the markers' positions and orientations with respect to the camera or vice versa: the position and orientation of the camera with respect to the markers. In MR systems, visual markers are often used with a camera as an integral part of the display - for example, attached to HMDs, hand-held displays, or projectors - so that virtual elements can be rendered at the correct position. Typically, these virtual elements are rendered directly in front of the markers hiding them from the viewer, and they can be visualized best using video see-through displays (head-mounted or hand-held). The same camera can be used for video see-through and for recognizing the markers. Disadvantages of visual markers are that they clutter the scene and require preparation of the environment. However, they have the advantage of being inexpensive and, generally, being usable both indoors and outdoors (within constraints due to ambient illumination and contrast).

An alternative use of visual markers which limits the amount of clutter is to place them out of user's field of view, for example on the ceiling of a room, and to have a camera pointing at them. Knowing the exact location of each marker, the system can then triangulate the position and orientation of the camera based on which markers are visible. However, in this case, the same camera cannot be used for video see-through, so a second camera is required (or an optical see-trough display).

While most systems use the visible part of the spectrum, a number of prototypes use infrared (IR) cameras in conjunction with IR light sources and markers printed on special materials that reflect only the IR portion of the spectrum. In this case, there is less interference by the lighting conditions. However, this requires a more complex installation.

Marker-less Tracking. Computer vision techniques can be used to recognize and track typical features of the environment such as faces or objects with specific textures or contours. These systems normally require a training phase in which the objects to be tracked are presented to the system from one or more viewpoint angles [39]. Compared to marker recognition, marker-less systems do not require the placement of extra objects or labels into the environment. However, this is at the expense of being significantly more computationally expensive, having a higher risk of misclassification, or higher latency. The LightSense system [29] tracks the LED on commercial cell phones, enabling them to be used as spatially aware handheld devices (Figure 3). The outside-in approach tracks the light source and streams the data to the phone over Bluetooth.

Acoustical Tracking Systems. Acoustical tracking systems can be distinguished between runtime (time-of- flight, TOF) and phase shift trackers. In the first system, multiple subsonic sources (approximately 40 kHz) are attached to the object to be tracked. At a certain time, the sources emit a subsonic pulse to the receivers, which are mounted to remain stationary. Since the subsonic pulse has different propagation times, depending on the distance between the source

Fig. 3. Outside-in approach tracking the phone light source and streaming the data to the phone. The spatially aware device augments a physical map with a detailed interactive road map of the area of interest [29].

and the receiver, the exact position of the tracked object can be calculated from this. Depending on the required position detection (required degrees of freedom) a different amount of emitters and receivers is needed. The largest amount is needed if the orientation of the object is required in addition to its position. One of the major problems of TOF trackers is the limited update rate caused by the propagation speed of sound in the air. Additionally, the propagation speed of sound depends on parameters of the air such as humidity, temperature, air pressure, and wind. However, these problems can be overcome by continuously measuring the propagation speed of sound with a second set-up. This presumes that the propagation speed is constant within the complete working volume of the tracker.

The other principle of acoustical tracking is the measurement of the phase shift between two signals with the same frequency. Within this phase shift tracking, the signal from the source of the tracked object is superimposed with the signal of a fixed signal source. If only sinusoidal waveforms are used, the position of the tracked object can be determined by the phase shift between the two signals. The two receivers measure the phase difference between the emitted waves and a reference oscillation. Since a phase shift of 360° is equivalent to one wavelength, the difference between the two consecutive measurements can be expressed as the travelling distance of the emitter between these two measurements. This presumes that this distance is less than one wavelength. In order to meet this requirement, the receivers have to measure the phase difference very quickly. If an acoustical tracking frequency of 40 kHz is assumed again, the accuracy is at best around 8 mm.

Magnetic and Inertial Sensors. Magnetic sensors rely on the Earth's magnetic field or artificially generated fields. The simplest example is a digital compass which measures orientation (one degree of freedom) using the Earth's magnetic field. More complex systems can also measure position [40,41]. These systems are typically used in many VR applications and allow tracking of all six degrees of freedom. However, the latter can be distorted by other electronic devices such as monitors and even passive metallic objects.

Inertial sensors generally do not rely on external references, and they measure movement related properties such as velocity and acceleration. The most common inertia sensors are accelerometers and gyroscopes. Theoretically, knowing the initial conditions, it would be possible to calculate a body's position from a consecutive integration of the measured forces. However, in reality, there are measurements errors caused by friction in the accelerometer's bearings, which result in drift errors increasing quadratically over time.

Hybrid Systems. Multiple tracking techniques can be combined to leverage the advantages of each system. As described above, inertial sensors such as accelerometers have drift errors. However, they can be combined with other types of sensors, such as ultrasonic beacons [42] or visual markers [43,44], which can periodically re-calibrate the absolute position of the device. Because ultrasonic and optical systems need a free line-of-sight, they can be ideally complemented by an inertia system for the moments when there are optical shadowing effects caused by the user or other obstacles. As another example, it is common to couple GPS receivers with digital compasses to obtain orientation information with inertial sensors to approximate the device's position whenever the satellite information is temporarily unavailable [23,15].

In order to further increase the calculation speed of such tracking systems, it was shown in [45] that acceleration information is best suited to feed Kalman filters that can predict the trajectory of an object and thus can reduce the lag of tracking systems.

4 User Studies of Mixed Reality

In the design process of an Mixed Reality (MR) application, a series of questions related to human-computer interaction (HCI) demands attention. First of all, who are the users and what are their needs? How can a system be designed to work effectively and efficiently for these users? How are effectiveness and efficiency measured in MR applications? Do users prefer an MR system or an alternative tool to go about their work? And finally, with what types of tasks and alternative tools should the usability of MR applications be tested? A set of perceptual issues, mostly related to the user's visual and auditory capacities, call for further attention. Embodiment and embodied interaction must also be considered as it has been recently pointed out by Dourish [46]. In his understanding, users create and communicate meaning through their interaction with a system. Lastly, issues related to the work context, the task at hand, and collaboration call for additional investigation.

A survey by Swan and Gabbard [47] shows that between 1998 and 2004, less than 10% of a representative sample of MR scientific publications reported studies with real users. The survey groups the studies into 3 categories. The first one is the most popular and includes studies that look at low-level issues in perception and cognition in AR. They examine issues such as perception of virtual objects' depths using different display technologies or rendering algorithms. Within this category, the ability of users to acquire targets under varying degrees of system lag is also studied. The second category covers six higher level evaluations of MR applications. Here, we find comparative studies of different MR interfaces applied to the same task and studies assessing the overall usability of an MR system. In the third category, the survey reports three studies about user interaction and communication in collaborative MR applications, looking, for example, at how communication asymmetries or different MR technologies influence users' behavior and performance. A special journal issue (IJHCI, 2003) on usability and collaborative issues of MR touch upon most of these questions and topics [48]. From the MR papers it presents, some are more visionary and focus on novel enabling technology for collaboration, while others offer solid empirical work presenting experimental studies with alternative applications. Two samples from this special issue follow.

The need for studies evaluating the effect of computerized tools on human cooperation and communication is well justified and documented in a paper offered by Billinghurst et al. [49]. The authors reported on two experiments: the first involving collaboration with MR technology as compared to more traditional unmediated and screen-based collaboration (Figure 4), and the second, the comparison of collaboration with three different MR displays. In both experiments, the authors used process and subjective measures in addition to more traditional performance measures. Process measures captured the process of collaboration through the number and type of gestures used and deictic phrases spoken. Using these measures to analyze communication behavior, it was found that users exhibited many of the same behaviors in a collaborative MR interface as they did in a face-to-face, unmediated collaboration. However, user communication behavior changed with the type of MR display used. The experimental task used was well suited to elicit collaboration and allowed for different styles of interaction to be evaluated within a single experiment. The authors then describe implications of the results for the design of collaborative MR interfaces and present plans for future research. The variety of relevant measures they use contrasts with most MR research which typically focuses on easily quantifiable aspects of task performance such as task completion time and error rate.

In another paper from the same issue, Wiedenmaier et al. showed how MR for assembly processes can be a new kind of computer support for a traditional indus-

Fig. 4. Billinghurst et al. [49] compared collaborative work under three alternative conditions: face-to-face, AR, and projector

trial domain [50]. The article concisely links MR to the real-world task of assembly. The new application of AR-technology is called ARsembly. The article describes a typical scenario for assembly and maintenance personnel and how MR might support both. For this purpose, tasks with different degrees of difficulty were selected from an authentic assembly process of the automotive industry. Two other kinds of assembly support media (a printed manual and a tutorial by an expert) were examined in order to compare them with ARsembly. The results showed that the assembly times varied according to the different support conditions. MR support proved to be more suitable for difficult tasks than the paper manual, whereas for easier tasks, MR support did not appear to be significantly more advantageous. As assumed, tasks done under the guidance of an expert were completed most rapidly. Some of the information obtained in this investigation also indicates important considerations for improving future ARsembly applications. The authors made a valuable contribution in presenting empirical results comparing different types of support for assembly processes. They also showed some evidence that a particular MR system in some situations can have advantages over traditional, printed assembly manuals. The authors have invested significant resources into building their systems and running controlled studies, greatly furthering scientific knowledge of MR and HCI. Their work shows where MR is both suitable and unsuitable. To achieve wide spread application for MR, it is important to take MR out of the lab and into the real world.

5 Mixed Reality Research in Switzerland

In Switzerland, research activities in the field of Mixed Reality (MR) take place at the two federal institutes of technology as well as at other institutions of higher education. This section reports some of the Swiss contributions.

The Virtual Reality Lab at EPFL and the MIRAlab at the University of Geneva were both involved in a project for the augmentation of the archaeological site of Pompei [51], working on several issues from tracking to the creation of virtual actors. Also at EPFL, the Computer Vision Lab developed vision-based markerless tracking techniques for MR [39].

At ETH Zurich, the Computer Vision Lab was involved in projects investigating MR and haptics, [52] as well as calibration techniques [53], mostly related

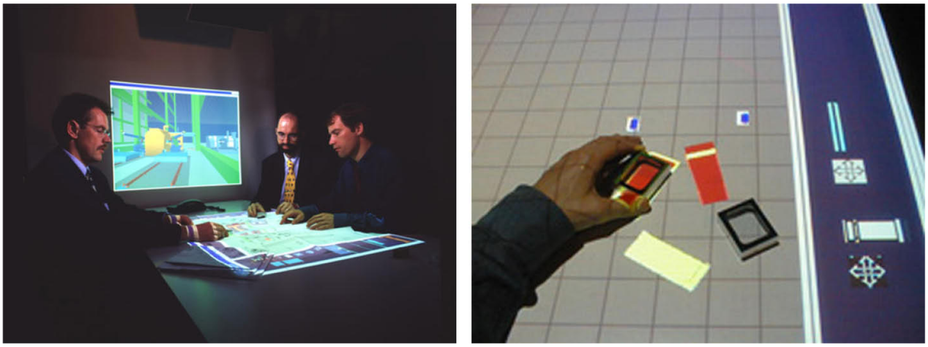

Fig. 5. The BUILD-IT system: Collaborative production plant layout combining digital, physical, and printed media (left); multi-pointer interaction of a furniture scenario (right)

to medical applications. BUILD-IT was a research project involving four ETH Zurich departments (ARCH, MAVT, MTEC, and ANBI) during the period from 1997 to 2001. The resulting BUILD-IT system (Figures 1 and 5) is a planning tool based on computer vision technology with a capacity for complex planning and composition tasks [54,55]. The system enables users, grouped around a table, to interact in a virtual scene using physical bricks to select and manipulate virtual models. A bird's eye view of the scene is projected onto the table. A perspective view of the scene, called the side view, is projected on the wall. The plan view contains a storage space with originals, allowing users to create new models and to activate tools e.g. navigation and height tools. Model selection is done by placing a brick at the model's position. Once selected, models can be positioned, rotated, and fixed by simple brick manipulation.

At the Innovation Center Virtual Reality (ICVR) at ETH Zurich, an AR-system for tabletop interaction was developed which uses typical office components such as pens, rulers, notepads and erasers. to enable interaction with the computer during group work [56,57]. Multiple users can work simultaneously with real objects that are augmented by a back-projection onto the tabletop. The objects are tracked through the screen via an IR-system, and additional information is displayed next to it, such as colour, virtual notepad and measurement results of the ruler. (Figure 6).

The same group carried out the research projects, blue-c and Holoport, focusing on an extension of the real environment into a virtual one (Figure 7). In the blue-c project [58], markerless optical tracking was used to track the user and their action in controlling the system and interacting with the remote 3D avatar. In the Holoport project [59,60], a real table was extended by a virtual one, allowing team meetings in an MR environment with the impression of sitting at the same table.

Also at ICVR (ETH Zurich), a MR application for education in architectural design was realized. Although many 3D models already exist in this field (real and virtual), it had so far not been possible to use such models within a

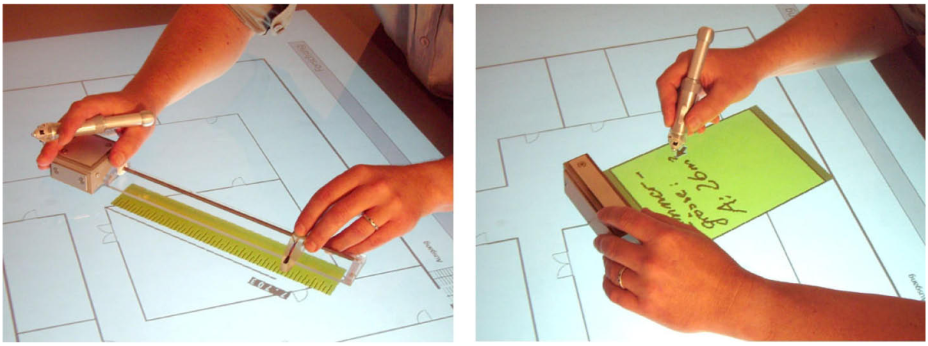

Fig. 6. A desktop AR-system for intuitive interaction with the computer

Fig. 7. The real table is extended by a virtual one, using a holographic projection screen

collaborative team session and to look at such models from different viewpoints. Using common libraries from AR-Toolkit [61], a collaborative 3D viewer was developed which allowed collocated and remote team members to simultaneously inspect the model of a building and perform simple actions like selecting, moving, rotating, scaling, and defining viewing planes (Figure 8).

Finally, the Sensory-Motor System Lab (SMS) Lab at ETH Zurich is investigating how athletes execute and learn the complex rowing movement. In order to do this they have built a rowing simulator based on virtual reality and MR technology. This simulator was required to convey the impression of realistic rowing, provide customizable, augmented feedback, and thus, optimal training conditions for rowing. The participant sits in a shortened racing boat (Figure 9) and holds one or two shortened oars that are virtually complemented in the

Fig. 8. MR application in architectural design

Fig. 9. The MR rowing environment of the SMS lab

computer generated image. The oars are connected to a rope robot. Depending on the oar pose and movement, forces are generated to simulate water resistance. The participant is surrounded by three screens (dimensions 4.44m 3.33m each) onto which three projectors display a river scenario.

Augmented Chemistry (AC) is an application that utilizes a tangible user interface (TUI) for organic chemistry education (Figure 10). First developed at HyperWerk FHNW Basel [62] and later together with IHA, ETH Zurich and Chalmers TH in Gothenburg [63]. An empirical evaluation compared learning effectiveness and user acceptance of AC versus the more traditional ball-and-stick model (BSM) [63]. Learning effectiveness results were almost the same for both learning environments. User preference and rankings, using NASA-TLX

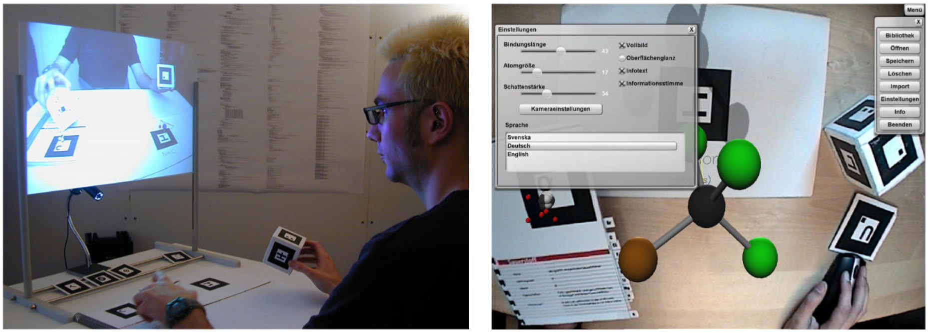

Fig. 10. The original Augmented Chemistry set-up from HyperWerk FHNW in Basel in 2002 (left) and the later version from t2i Lab at Chalmers TH in Gothenburg in 2005 (right)

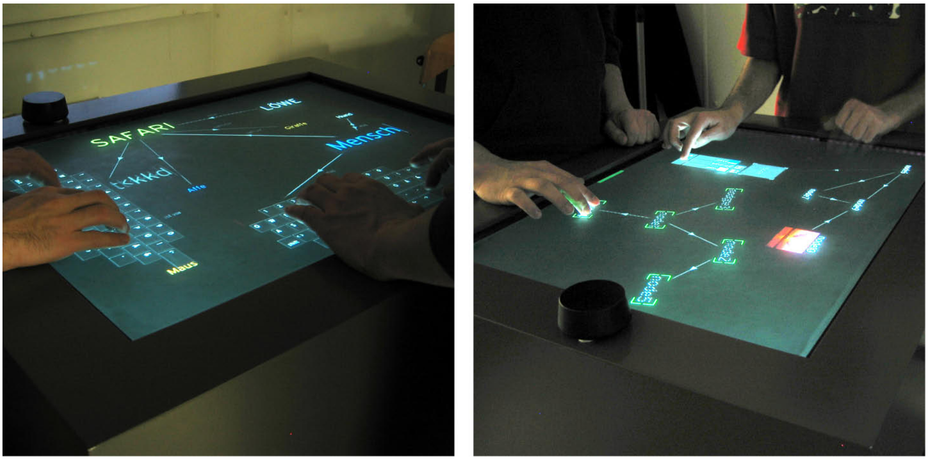

Fig. 11. Tangent: an early multi-touch tabletop framework, realized by Christian Iten and Daniel Lüthi, Interaction Design, ZHdK in Zurich [64]

and SUMI, showed more differences, for example in ease of use and in ease of learning the system. It was therefore decided to focus mainly on improving these aspects in a re-design of the AC system. For enhanced interaction, keyboard-free system configuration, and internal/external database (DB) access, a graphical user interface (GUI) were incorporated into the TUI. Three-dimensional rendering was also improved using shadows and related effects, thereby enhancing depth perception. The re-designed AC system (Figure 10, right) was then compared to the original system by means of a small qualitative user study. This user study showed an improvement in subjective opinions about the systems ease of use and ease of learning the system.

The size of the human finger makes it difficult for users to precisely manipulate small elements on touch screens. Christian Iten and Daniel Lüthi from the ZHdK presented a tabletop framework called Tangent including a tool called Digital Tweezers [64]. This tool enables its users to point, select, and drag interactive elements the size of a few pixels. The tool consists of a cursor with a fixed offset controlled by the thumb and index finger of one hand. Based on the Tangent framework, ConceptMap was realizedin collaboration with Emanuel Zgraggen und Simon Brauchli from IFS at HSR. ConceptMap is a multi-touch application for creating and editing semantic nets (Figure 11).

The Real-Time Coordination and Sustainable Interaction System Group at EPF-L, the Pervasive Artificial Intelligence Group at the University of Fribourg, and the Multimedia Information System Group at University of Applied Sciences of Western Switzerland, Fribourg were all involved in the 6th Sense project which is also presented in this book. The project aims to improve the user experience in mobile MR and the context-aware interaction between real environments and virtual augmentations. The Computer Vision and Multimedia Lab at the University of Geneva and the Laboratoire d'Informatique Industrielle at the University of Applied Sciences of Western Switzerland, Geneva were both involved in the See ColOr project aimed at adding sound to images in order to provide an interactive aid for visually impaired individuals.

6 Current Challenges and Trends

Most of the systems and technical solutions described in the previous sections require prior preparation of the environment to run the Mixed Reality (MR) system in. For example, ultrasound or IR beacons need to be installed and powered, or visual markers and RFID tags need to be placed in specific locations. Even in outdoor environments, GPS requires the installation of local transmitters in addition to the existing satellites in order to achieve high positioning accuracy. Additionally, the systems require an accurate digital model of the real environment and a complete mapping of the sensors' locations to precisely position the virtual elements of the interface within the real space.

A number of prototypes presented in research papers use only a loose connection between the virtual elements and the physical space. For example, from the description of the AR Quake game [21], as well as the Moon Lander AR game [23], it is not clear how the features of the game are aligned to the physical world, and, in fact, there seems to be few compelling reason for a specific mapping of the real and virtual game spaces. Similarly, in the Magic Book [9] and Fata Morgana [10] projects, the system is able to render virtual models aligned to the physical pages of a book, but it is unclear what role the physical object plays in the application beyond being a handle or a controller - the MR features seem to be used solely to rotate the models and to allow users to look at them from different viewpoints. This would also be possible with a virtual reality system or even with just a desktop graphics workstation. In other words, it seems that in a number of MR prototypes the paradigm reverts to VR, the importance is placed solely on the virtual elements and not on the real ones, and the MR features are used only as interactive controllers, like handles. Even though the Augmented Reality definition by Azuma et al. [3] still applies to these systems in the sense that they are interactive and display virtual objects registered to the real world, it can be observed that the connection between the virtual elements and the physical reality is relatively weak.

Future investigation into MR could therefore target specific domains of application, such as the medical field, in which specialized digital data is inherently aligned with the physical world, as in the case of medical image overlays onto a patient's body in surgery. Also, future MR research could consider larger scale, more general applications, reduced registration requirements, and thus allowing an easier implementation onto consumer hand-held devices. Today's mobile phones are ubiquitous and already embed considerable computational capabilities (yet not enough for most registration techniques) - initial exploration suggests that they may have great potential for MR. The relationship between digital content and physical space could then become less tight in terms of resolution, but more stringent in terms of relevance in the direction of location-based services and applications.

Based on the experiences of MR research in Switzerland, we see a trend towards more application-specific projects, typically industrial and educational. Industrial applications are directed towards support planning tasks. The fact that projects are becoming more application-oriented may indicate that MR technologies are becoming more mature. While early MR systems were mostly single-user, more recent applications are collaborative, both for co-located and net-based use.

References

1. Weiser, M.: Some computer science issues in ubiquitous computing. Commun. ACM 36(7), 75–84 (1993)
2. Sutherland, I.: A head-mounted three dimensional display. In: Proceedings of the Fall Joint Computer Conference, vol. 33, pp. 757–764 (1968)
3. Azuma, R., Baillot, Y., Behringer, R., Feiner, S., Julier, S., MacIntyre, B.: Recent advances in augmented reality. IEEE Comput. Graph. Appl. 21(6), 34–47 (2001)
4. Milgram, P., Kishino, F.: A Taxonomy of Mixed Reality Visual Displays. IEICE Transactions on Information Systems 77, 1321–1329 (1994)
5. Feiner, S., Macintyre, B., Seligmann, D.: Knowledge-based augmented reality. In: Communications of the ACM, vol. 36, pp. 53–62. ACM Press, New York (1993)
6. Lee, J., Rhee, G.: Context-aware 3D visualization and collaboration services for ubiquitous cars using augmented reality. The International Journal of Advanced Manufacturing Technology 37, 431–442 (2008)
7. Reiners, D., Stricker, D., Klinker, G., Muller, S.: Augmented Reality for Construction Tasks: Doorlock Assembly. In: 1st International Workshop on Augmented Reality (IWAR 1998), San Francisco (1998)
8. Doil, F., Schreiber, W., Alt, T., Patron, C.: Augmented reality for manufacturing planning. In: Proceedings of the workshop on Virtual environments 2003, pp. 71–76. ACM Press, New York (2003)
9. Billinghurst, M., Kato, H., Poupyrev, I.: The MagicBook: a transitional AR interface. In: Computers & Graphics, vol. 25, pp. 745–753. Elsevier, Amsterdam (2001)

10. Klinker, G., Dutoit, A., Bauer, M., Bayer, J., Novak, V., Matzke, D.: “Fata Morgana” A Presentation System for Product Design. In: Proceedings of the International Symposium on Mixed and Augmented Reality (ISMAR 2002). IEEE Computer Society, Washington (2002)
11. State, A., Livingston, M.A., Garrett, W.F., Hirota, G., Whitton, M.C., Pisano, E.D., Fuchs, H.: Technologies for augmented reality systems: Realizing ultrasound-guided needle biopsies. In: SIGGRAPH, pp. 439–446 (1996)
12. State, A., Hirota, G., Chen, D.T., Garrett, W.F., Livingston, M.A.: Superior augmented reality registration by integrating landmark tracking and magnetic tracking. In: SIGGRAPH, pp. 429–438 (1996)
13. Lorensen, W.E., Cline, H.E., Nafis, C., Altobelli, D., Gleason, L.: Enhancing reality in the operating room. In: IEEE Visualization, pp. 410–415 (1993)
14. Rosenthal, M., State, A., Lee, J., Hirota, G., Ackerman, J., Keller, K., Pisano, E., Jiroutek, M., Muller, K., Fuchs, H.: Augmented reality guidance for needle biopsies: An initial randomized, controlled trial in phantoms. In: Medical Image Analysis, vol. 6, pp. 313–320. Elsevier, Amsterdam (2002)
15. Höllerer, T., Feiner, S., Terauchi, T., Rashid, G., Hallaway, D.: Exploring MARS: developing indoor and outdoor user interfaces to a mobile augmented reality system. In: Computers & Graphics, vol. 23, pp. 779–785. Elsevier, Amsterdam (1999)
16. Welch, G., Sonnenwald, D., Mayer-Patel, K., Yang, R., State, A., Towles, H., Cairns, B., Fuchs, H.: Remote 3D medical consultation. In: Proc. BROADMED Conf. 2005, University College of Borås. Swedish School of Library and Information Science, pp. 103–110 (2005)
17. Julier, S., Baillot, Y., Brown, D., Lanzagorta, M.: Information filtering for mobile augmented reality. IEEE Computer Graphics and Applications 22(5), 12–15 (2002)
18. Thomas, B., Piekarski, W., Hepworth, D., Gunther, B., Demczuk, V.: A wearable computer system with augmented reality to support terrestrial navigation. In: ISWC 1998: Proceedings of the 2nd IEEE International Symposium on Wearable Computers, Washington, DC, USA, p. 168. IEEE Computer Society, Los Alamitos (1998)
19. Feiner, S., MacIntyre, B., Hollerer, T., Webster, A.: A touring machine: Prototyping 3d mobile augmented reality systems for exploring the urban environment. In: ISWC 1997: Proceedings of the 1st IEEE International Symposium on Wearable Computers, Washington, DC, USA, p. 74. IEEE Computer Society, Los Alamitos (1997)
20. Rekimoto, J.: Navicam: A magnifying glass approach to augmented reality. In: Presence, vol. 6, pp. 399–412 (1997)
21. Thomas, B., Close, B., Donoghue, J., Squires, J., Bondi, P., Piekarski, W.: First Person Indoor/Outdoor Augmented Reality Application: ARQuake. In: Personal and Ubiquitous Computing, vol. 6, pp. 75–86. Springer, Heidelberg (2002)
22. Cheok, A., Goh, K., Liu, W., Farbiz, F., Fong, S., Teo, S., Li, Y., Yang, X.: Human Pacman: a mobile, wide-area entertainment system based on physical, social, and ubiquitous computing. In: Personal and Ubiquitous Computing, vol. 8, pp. 71–81. Springer, Heidelberg (2004)
23. Avery, B., Thomas, B., Velikovsky, J., Piekarski, W.: Outdoor augmented reality gaming on five dollars a day. In: Proceedings of the Sixth Australasian conference on User interface, vol. 40, pp. 79–88. Australian Computer Society, Inc, Darlinghurst (2005)
24. Szalavári, Z., Eckstein, E., Gervautz, M.: Collaborative gaming in augmented reality. In: Proceedings of the ACM symposium on Virtual reality software and technology 1998, pp. 195–204. ACM Press, New York (1998)

25. Barakonyi, I., Weilguny, M., Psik, T., Schmalstieg, D.: Monkeybridge: autonomous agents in augmented reality games. In: Lee, N. (ed.) Advances in Computer Entertainment Technology, pp. 172–175. ACM, New York (2005)
26. Bungert, C.: Hmd/headset/vr-helmet comparison chart (2008), `http://www.stereo3d.com/hmd.htm`
27. Rekimoto, J., Nagao, K.: The world through the computer: computer augmented interaction with real world environments. In: Proceedings of the 8th annual ACM symposium on User interface and software technology, pp. 29–36. ACM, New York (1995)
28. Moehring, M., Lessig, C., Bimber, O.: Video See-Through AR on Consumer Cell Phones. In: Proc. of International Symposium on Augmented and Mixed Reality (ISMAR 2004), pp. 252–253 (2004)
29. Olwal, A.: Lightsense: enabling spatially aware handheld interaction devices. In: ISMAR, pp. 119–122 (2006)
30. Wellner, P.: Interacting with paper on the DigitalDesk. In: Communications of the ACM, vol. 36, pp. 87–96. ACM Press, New York (1993)
31. Fitzmaurice, G., Ishii, H., Buxton, W.: Bricks: laying the foundations for graspable user interfaces. In: Proceedings of the SIGCHI conference on Human factors in computing systems, pp. 442–449. ACM Press/Addison-Wesley Publishing Co., New York (1995)
32. Pinhanez, C.: Augmenting Reality with Projected Interactive Displays. In: Virtual and Augmented Architecture (Vaa 2001): Proceedings of the International Symposium on Virtual and Augmented Architecture (Vaa 2001), Trinity College, Dublin, June 21-22 (2001)
33. Olwal, A., Gustafsson, J., Lindfors, C.: Spatial augmented reality on industrial cnc-machines. In: Proceedings of SPIE 2008 Electronic Imaging. The Engineering Reality of Virtual Reality 2008, vol. 6804 (2008)
34. Raskar, R., van Baar, J., Beardsley, P., Willwacher, T., Rao, S., Forlines, C.: iLamps: geometrically aware and self-configuring projectors. In: International Conference on Computer Graphics and Interactive Techniques. ACM Press, New York (2006)
35. Getting, I.: Perspective/navigation-The Global Positioning System. In: Spectrum, vol. 30. IEEE, Los Alamitos (1993)
36. Steed, A.: Supporting Mobile Applications with Real-Time Visualisation of GPS Availability. In: Proceedings of Mobile HCI. Springer, Heidelberg (2004)
37. Kato, H., Billinghurst, M., Poupyrev, I., Imamoto, K., Tachibana, K.: Virtual object manipulation on a table-top AR environment. In: Proc. IEEE and ACM International Symposium on Augmented Reality (ISAR), pp. 111–119 (2000)
38. Costanza, E., Robinson, J.: A region adjacency tree approach to the detection and design of fiducials. In: Vision, Video and Graphics (VVG), pp. 63–70 (2003)
39. Lepetit, V., Fua, P.: Keypoint Recognition Using Randomized Trees. In: IEEE Transactions On Pattern Analysis And Machine Intelligence, pp. 1465–1479. IEEE Computer Society, Los Alamitos (2006)
40. Polhemus: Polhemus “fastrak” commercial system (retrieved, June 2008), `http://www.polhemus.com/?page=Motion_Fastrak`
41. Ascension: Ascension “flock of birds” commercial system (retrieved, June 2008), `http://www.ascension-tech.com/products/flockofbirds.php`
42. Foxlin, E., Harrington, M., Harrington, M., Pfeifer, G.: Constellation: a wide-range wireless motion-tracking system for augmented reality and virtual set applications. In: Proceedings of the 25th annual conference on Computer graphics and interactive techniques, pp. 371–378. ACM New York, NY (1998)

43. Yokokohji, Y., Sugawara, Y., Yoshikawa, T.: Accurate image overlay on video see-through hmds using vision and accelerometers. In: VR 2000: Proceedings of the IEEE Virtual Reality 2000 Conference, Washington, DC, USA, p. 247. IEEE Computer Society, Los Alamitos (2000)
44. Kotake, D., Satoh, K., Uchiyama, S., Yamamoto, H.: A hybrid and linear registration method utilizing inclination constraint. In: Proceedings of the Fourth IEEE and ACM International Symposium on Mixed and Augmented Reality, pp. 140–149 (2005)
45. Azuma, R., Bishop, G.: A frequency-domain analysis of head-motion prediction. In: SIGGRAPH 1995: Proceedings of the 22nd annual conference on Computer graphics and interactive techniques, pp. 401–408. ACM, New York (1995)
46. Dourish, P.: Where the action is: the foundations of embodied interaction. MIT Press, Cambridge (2001)
47. Swan, J.E., Gabbard, J.L.: Survey of user-based experimentation in augmented reality. In: Proceedings of 1st International Conference on Virtual Reality (2005)
48. Fjeld, M.: Special Issue on Usability and collaborative aspects of augmented reality. Interactions 11(6), 11–15 (2004)
49. Billinghurst, M., Belcher, D., Gupta, A., Kiyokawa, K.: Communication Behaviors in Colocated Collaborative AR Interfaces. International Journal of Human-Computer Interaction 16(3), 395–423 (2003)
50. Wiedenmaier, S., Oehme, O., Schmidt, L., Luczak, H.: Augmented reality (ar) for assembly processes design and experimental evaluation. Int. J. Hum. Comput. Interaction 16(3), 497–514 (2003)
51. Papagiannakis, G., Schertenleib, S., Ponder, M., Arévalo-Poizat, M., Magnenat-Thalmann, N., Thalmann, D.: Real-time virtual humans in ar sites. In: 1st European Conference on Visual Media Production (CVMP), pp. 273–276. IEEE Publisher, Los Alamitos (2004)
52. Bianchi, G., Knörlein, B., Szèkely, G., Harders, M.: High precision augmented reality haptics. In: Eurohaptics (2006)
53. Bianchi, G., Wengert, C., Harders, M., Cattin, P., Székely, G.: Camera-marker alignment framework and comparison with hand-eye calibration for augmented reality applications. In: ISMAR (2005)
54. Fjeld, M., Lauche, K., Bichsel, M., Voorhorst, F., Krueger, H., Rauterberg, M.: Physical and virtual tools: Activity theory applied to the design of groupware. In: Computer Supported Cooperative Work (CSCW), vol. 11, pp. 153–180 (2002)
55. Fjeld, M., Morf, M., Krueger, H.: Activity theory and the practice of design: evaluation of a collaborative tangible user interface. Inderscience 4, 94–116 (2004)
56. Ganser Schwab, C., Kennel, T., Kunz, A.: Digital support for net-based teamwork in early design stages. Journal of Design Research 6(1), 150–168 (2007)
57. Schwab, C.G., Kennel, T., Kunz, A.: Digital support for net-based teamwork in early design stages 6(1/2), 150–168 (2007)
58. Gross, M.H., Würmlin, S., Näf, M., Lamboray, E., Spagno, C.P., Kunz, A.M., Koller-Meier, E., Svoboda, T., Gool, L.J.V., Lang, S., Strehlke, K., Moere, A.V., Staadt, O.G.: Blue-c: a spatially immersive display and 3d video portal for telepresence. ACM Trans. Graph. 22(3), 819–827 (2003)
59. Kuechler, M., Kunz, A.: Holoport - a device for simultaneous video and data conferencing featuring gaze awareness. In: Proceedings of the IEEE conference on Virtual Reality (IEEE VR 2006), Washington, DC, United States, pp. 81–87. IEEE Computer Society, Los Alamitos (2006)

60. Kunz, A., Fadel, G., Taiber, J., Schichtel, M.: Towards collaboration in engineering of tomorrow - building highly interactive virtual collaboration platforms. In: SAE 2006 World Congress & Exhibition, United States, SAE International (2006)
61. Kato, H., Billinghurst, M.: Marker tracking and hmd calibration for a video-based augmented reality conferencing system. In: IWAR 1999: Proceedings of the 2nd IEEE and ACM International Workshop on Augmented Reality. IEEE Computer Society, Los Alamitos (1999)
62. Voegtli, B.: Augmented collaboration. M.sc. thesis, HyperWerk FHNW (2002)
63. Fjeld, M., Fredriksson, J., Ejdestig, M., Duca, F., Bötschi, K., Voegtli, B.M., Juchli, P.: Tangible user interface for chemistry education: comparative evaluation and redesign. In: CHI, pp. 805–808 (2007)
64. Iten, C., Lüthi, D.: (2006), `http://www.zima.ch/tangent/`