Towards natural user interfaces in VR/AR for Design and Manufacturing

Conference Paper

Author(s):
Kunz, Andreas; Wegener, Konrad

Publication date:
2013

Permanent link:
https://doi.org/10.3929/ethz-a-007611572

Towards Natural User Interfaces in VR/AR for Design and Manufacturing

Kunz, A.; Wegener, K.

ETH Zurich, Institute of Machine Tools and Manufacturing

Abstract

VR/AR is used in many application fields in engineering, but also in medicine, architecture or entertainment. It proved its ability to shorten development cycles, to reduce errors in development and to visualize complex data. However, it still is a domain of specialists who know how to handle a complex VR system and to work with devices that are not adapted to the human behavior. To overcome this problem, research realizes more intuitive VR/AR systems. This paper gives an overview on these approaches and shows how new application fields could be opened by realizing a more natural and intuitive interaction.

1 Introduction

Over the past decades, VR/AR matured from a questionable toy of computer freaks to a tool that can be used in many application fields. The acceptance also changed from a mystic technology to a tool that is used in business. This is due to the fact that increased computer power allows a more comfortable operation, displaying self-explaining icons instead of command line operators. The computer power follows Moore's law [1], and now computers are fast enough to solve problems in a reasonable time and to display the results even in a realistic way. Thus, virtual worlds have an increasing reality in their appearance as well as in their behavior. This increased realism motivates users to treat virtual objects like real ones, but here they fail, since user interfaces did not undergo the progress as the other computer components and thus are the bottleneck for many VR/AR applications. Thus, this paper gives an overview on advances in user interface development in order to overcome the problems mentioned above. While the main chapter of this paper focuses on research results of the recent years, the remainder of the paper gives an outlook on future research and closes with a summary and outlook.

2 Motivation

As mentioned, there is a discrepancy between increasing computer power and stagnant user interface capabilities. Due to this, the immaturity of user interfaces is compensated by complex GUIs (Graphical User Interface), which make programs too complex for daily business processes. Mapping a 3D virtual object's multi degrees of freedom onto a 2D pointing device requires some effort for adjusting the desired degree of freedom, as well as an in-depth knowledge of handling the program. This is provided by an operator, while the user's interaction capabilities are limited. Hence, such installations are difficult to integrate in regular work processes and result in the fact that many CAVEs were dismantled.

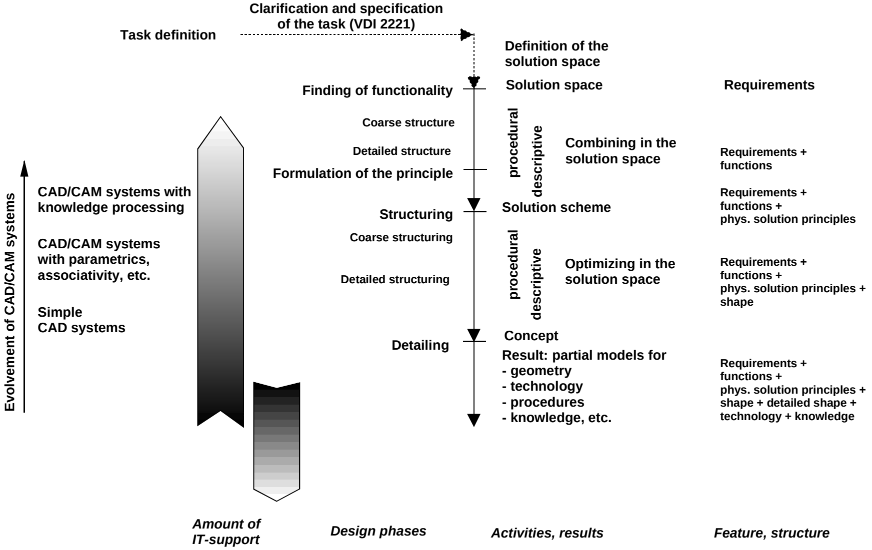

Figure 1: Trends in VR/AR-support

Following Vajna et al. [3], VR/AR technology supports the design phases, but also expands to the ideation and the usage phase (Figure 1). Addressing these two phases involves different disciplines. This superposes two challenges to existing interface solutions:

- Highly efficient and intuitive interfaces for untrained users
- Allowing multiple users and/or multi-device interaction

Although these are obvious goals, they were neglected for a long time due to technical reasons, but also due to the fact that specialized interfaces only cover a small application field. It took more than 6 decades for the mouse to evolve to intuitive user interfaces, which also take into account human behavior like postures, gestures, gazes, etc. The following chapter emphasizes highlights of natural and intuitive user interfaces in VR/AR for design and manufacturing.

3 Milestones in natural user interface design

This chapter will give examples of interfaces for three main application fields: ideation phase, geometric design phase and the usage phase of a product of production line. Since the geometric design phase is the most well-known one, this chapter will start with this, while the other two phases will follow later.

3.1 Natural interfaces in the geometric design phase

When handling complex geometries, it is obvious that a mouse is not the tool of first choice [4]. However, the mouse is still used in VR-environments, in which operators guide users through the virtual world. Here, users are passive, while a next step to increase intuitiveness would be to allow an exploratory behavior, using a so-called spacemouse. This interface allows exploring the virtual environment, but lacks immersion since the perception of walking is not addressed. Thus, an estimation of size can only be done visually by a comparison with known objects.

An obvious next step to more natural user interfaces is to integrate locomotion. However, this poses the problem that virtual worlds are typically larger than a physical room. To overcome this problem, stepping-in-place solutions were proposed by Bouquila et al. [6] and Kobayashi et al. [7]. Here, the user steps in place to navigate through the virtual environment. However, this still differs from walking from the ergonomic point of view, and thus the idea on an omnidirectional treadmill – originally introduced by Darken et al [8] and Iwata [9] – was utilized by the CyberWalk project [10]. With these concepts, real walking in an unlimited virtual environment was possible, but only for a single user.

In summary, navigation devices for virtual environments in the design phase matured from a mouse to a natural input device, but obviously to the cost of huge installations that are currently not applicable to industrial business processes, even for huge global players. Moreover, the second requirement of a multi-user capability is not yet fulfilled, thus leaving sufficient possibilities for future research in this field.

3.2 Natural interfaces in the ideation phase

Following [11], virtual reality is "*an artificial environment in which one's action partially determines what happens in the environment*". This does not only imply CAVE-like installations, but also holds true for e.g. sketching on active surfaces. Although it is not visible at a first glance, there are many similarities in technology and algorithms, and thus there is also the need for an intuitive user interface. Moreover, it is required to detect the most natural interfaces, which are human gestures, postures or facial expressions. These natural interfaces are important in the ideation phase, since they carry important information. Thus, realizing natural user interfaces becomes even more complex and is not completely solved yet.

IT-support started with the “Panaboard” digitizer [12], which used normal pens for interaction. Later, electronic whiteboards detect the pen’s position to generate virtual artifacts. In all systems, the user touches the surface only in one single point. This is a suitable solution for vertical surfaces, but it is unusable when realizing interactive tables. Here, user interfaces are more complex since they have to support the work in multiple layers (Figure 2).

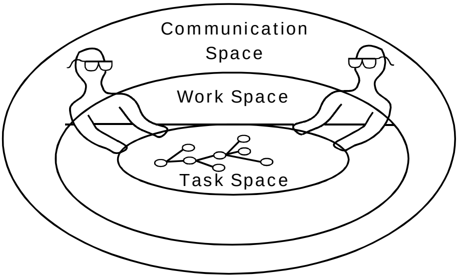

Figure 2: Tabletop systems integrate communication, work and task spaces [13]

In order to realize intuitive interfaces, it is not sufficient to just use technology from vertical interactive workspaces, since the user’s behavior is different on horizontal surfaces. When writing and sketching for example, the user touches the interactive surface not only with the pen’s tip, but also with the ball of the hand. Now, the system has to decide which action is consciously performed, and which one is not. Further, it could occur that multiple users interact simultaneously with the system, or that one user performs a bi-manual interaction, like using a ruler and a pen to draw a straight line. Even this should be unambiguously detected by the system.

Fitzmaurice et al. [14] coined the term ‘GUI’ (Graspable User Interface), which was consciously similar to GUI (Graphical User Interface) to underline that this new generation of interfaces does not require a mouse anymore, but is an object that explains its inherent functionality by its shape. Ullmer et al. [15] introduced ‘metaDESK’, which uses infrared light to detect objects, but also an electromagnetic tracking system for devices like a magnifier lens. Since the interfaces showed what they can be used for, they introduced the term “TUI” (Tangible User Interface), which was consciously the same syllable as in the word “intuitive”.

So far, the systems only could detect which device is where on the tabletop. To further increase the interface’s intuitiveness, they should have multiple states. A ball-pen for example might be switched on or off, or it might be used for writing or pointing. Moreover, no cable-bound devices should be used anymore since they hinder intuitive work. Ganser et al. [16] [17] addressed the problem with interfaces for brainstorming sessions on tabletops. Here, typical devices exist like rulers for measuring distances (caliper metaphor) or drawing lines (ruler metaphor), and an ink dwell, in which the pen can be dipped to change its color. Moreover, a notepad was realized (notepad metaphor). Pens exist that can distinguish between writing and pointing gestures (pen metaphor) (Figure 3). By using infrared light for signalizing position, state and ID, cables are not required anymore. Since active devices are used, other objects like books, laptops, etc. do not irritate the system.

Figure 3: Typical interaction devices for a brainstorming session

All systems for the ideation phase only offer tools to interact on the surface. However, the most intuitive interface – the human gestures, postures and gazes – is not very well supported. Although products such as MS Surface and PixelSense [18] offer touch detection and object identification using non-repetitive patterns (so-called fiducials), they fail to detect multiple device states. To address this problem, Hofer et al. [19] introduced a system which was capable of tracking multi-state devices and finger touches using IR-sensors underneath the LC-matrix (Figure 4).

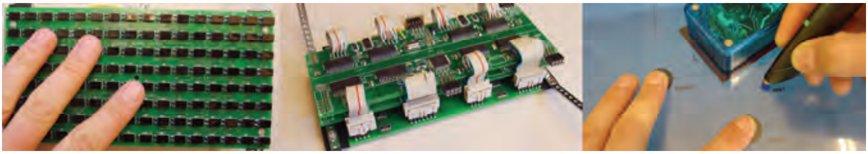

Figure 4: TNT: left: IR sensor array; middle: pre-processing electronics; right: touches and TUIs

The system was capable to handle the devices from [17]. It could also handle pointing gestures using the pen. However, deictic gestures were not supported. The problem of detecting deictic gestures, postures and gazes in front of an interactive surface was addressed by Kunz et al. [20] and Nescher et al. [21], who used a camera to capture a person in front of an LC-screen and overlaid this video image over the content of the digital whiteboard at the remote side (see Figure 5). The system allows transferring deictic gestures, postures and gazes, while having the digital content editable for both sides. It also could detect multiple devices (IDs), but no states. Unlike “VideoArms” [22], the CollaBoard addressed all three spaces from Figure 2.

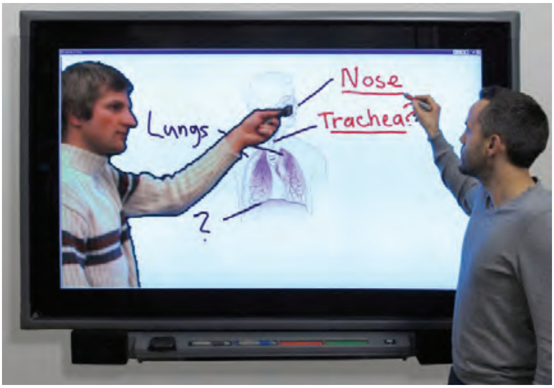

Figure 5: Local and remote users edit the same digital content

To summarize the development of intuitive user interfaces for the early stages of product and production design, there is also a tendency from a non-intuitive single-user interface to a multi-user, multi-state and multi-device system. Moreover, the special application field requires also capturing, transferring and displaying the other collaboration spaces (Figure 2), which is only partly possible today.

3.3 Intuitive User Interfaces for the Usage Phase

Since there is the tendency to use natural interfaces in VR/AR from the very first steps, design and manufacturing planning phases are performed digitally from the very beginning. Thus, it is obvious to use intuitive and natural interfaces also for the usage phase. Unlike the other two application fields for natural user interfaces, other requirements become important in this sector:

- Robustness for daily use even in harsh environments
- Hands-free operation is required very often
- Using off-the-shelve devices

An early approach in designing an intuitive user interface for a service application was introduced by Feiner et al. [23]. The user wore a helmet to see a wireframe that helps maintaining a laser printer. The orientation of the HMD (Head-mounted Display) was captured by an acoustic tracking system. However, the system never left the laboratory stage, since it was quite some effort to prepare the environment with the necessary tracking system. Boeing introduced the “Wire Bundle Assembly Project” [24], trying to reduce the amount of formboards to preshape wire bundles. The system used an optical tracking system and a see-through HMD. By this, the environment preparation, which is the application of white markers on a reusable formboard, can be neglected. It allows a hands-free operation, which is crucial for bending the cables and for placing them in the correct position. However, the drawback of this system was the weight and the fact that it was cable-bound. More recently, the ARVIKA [25] and the ARTESAS [26] projects addressed this problem by providing a more lightweight HMD. However, the systems were still cable-bound and required landmarks in the environment to achieve a reliable tracking.

While the above systems were designed for manufacturing and maintenance, other approaches of intuitive interfaces focus on the customer, who also could benefit from the digital data coming from design and construction. However, it can neither be expected that users would have specialized HMDs, nor that the environment is equipped with fiducials. Thus, researchers worked on solutions how to use PDAs. Beier et al. [27] used a PDA as a modular instruction manual, which does not need any specialized landmarks in the environment. Due to the limited CPU power of PDAs, the feature detection was done on a server, which received an image from the PDA and then returned a matching overlay. However, the limited available bandwidth led to delays between image acquisition and server response. With the appearance of powerful smartphones and other handheld devices, together with a high bandwidth, mobile AR-systems reappear in research. As an example, Kahn et al. [28] use an i-Pad device for a maintenance support.

Summarizing the tendencies in this field, there are approaches of using a HMD and of using handheld devices instead. Both use off-the-shelf components and can be used in even harsh environments, but still differ in hands-free operation.

4 Summary and Future Trends

Within this overview it was shown that VR/AR technologies started to support the whole lifecycle of a product, starting from the very first ideas over production, up to the usage phase of a product. Such a wide range of applications also asks for more natural user interfaces, since many different disciplines and also very different working procedures are involved. Moreover, the complexity of products and production requires a teamwork, which also should be supported by intuitive user interfaces, for collocated situations as well as for net-based work.

While for a long time all interfaces followed the mouse metaphor, meaning that only one single device can be used per time interval, more recent research matured these interfaces to not only match to established work procedures, but also to be used outside the laboratory. There is a still ongoing research in all three application fields mentioned in the above, which will be briefly sketched in the next paragraphs.

4.1 Trends in the Geometric Design Phase

For a long time, CAVEs and Powerwalls were used to visualize complex geometries. However, many institutions abolished these installations due to the high effort to run such systems, and because of the large space requirements.

Newer research in this field focuses more on the haptic sensation of walking and replaced the projection systems by HMDs [29], [30], [31]. However, this requires an unnoticeable compression of an ideally unlimited virtual world into a physically limited real room. While the required hardware is quite usable already, current research addresses more the 'compression' algorithms to reach a high immersion in the virtual environment. Besides significantly lower costs for the installation and a higher immersion, such new system can be installed in almost any room, since the fiducials on the ceiling are the only static installation.

Figure 6: Setup for a real walking in virtual environments

While in current systems laptops are used for generating the virtual environment and for processing the tracking system's measurements (see Figure 6), a future vision will be that this could also be performed on smartphones that already have a powerful CPU. A powerful marker-less feature detection could even replace the fiducials and thus the preparation of spaces. Such systems together with sophisticated compression and redirection algorithms could allow a walkthrough even for large virtual environments such as shipyards and large production lines.

4.2 Trends in Natural Interfaces for the Ideation Phase

To support ideation sessions, more devices with multiple states must be detected together with multi-touch. Some systems are already realized, but so far they suffer from an increasing latency if the amount of interaction devices is increased. A first solution is proposed by Alavi et al. [32], who realized a low-latency interaction in the task space (Figure 2). Further, a reliable tracking is required for the work space to detect explanatory gestures. Here, depth cams or structured light approaches like MS Kinect could contribute, but interference problems among multiple Kinect and active tables like Pixelsense need to be solved.

So far, the communication space is not very well supported, since it was regarded as less relevant for collaboration. However, in particular for the ideation phase, postures, gazes, or the ‘transportation’ of artifacts to other surfaces (e.g. a digital card from the horizontal workspace to the vertical workspace) must be better supported. Here, multi-sensor systems will be required to unequivocally detect and interpret human behavior and to evoke the correct action by the digital environment.

4.3 Trends in Intuitive Interfaces for Usage Phases of a Product or Production

The main drawback of the existing systems is weight and a missing hands-free operation. If such capabilities could be combined with the high CPU power of today’s smartphones, new promising systems could result from this combination. A very first step into this direction are ultra-lightweight glasses that are combined with a camera, as it was introduced by Google’s Project Glass [33]. If such hardware will be combined not only with location-based information (which could easily overload the visual channel), but with more refined data based on prediction algorithms, a powerful tool for providing additional operation information could result from this.

Employing such lightweight devices will allow using VR/AR technology to support the worker even on large construction sites, giving him e.g. positioning information for exactly placing large sub-assemblies. Moreover, the same marker-less tracking technology could be used to support maintenance personnel by providing refined data at the correct position in a shop floor environment. Unlike in previous work, these will not be assembly or disassembly instructions, but additional “invisible” data from machine controls or databases, such as operating time of components, current temperature, long-term comparisons of measured electrical data, virtual product labels, etc.

References

[1] Moore, G.: Cramming more components onto integrated circuits; Electronics, Vol. 38, No. 8; 1965

[2] Nicca, C.: Unterstützung für Zugänglichkeitskontrollen und Maintenance durch Hilfsmittel der virtuellen Realität; Master Thesis ETH Zurich; 2001

[3] Vajna, S.; Podehl, G.: Durchgängige Produktmodellierung mit Features. In: CAD-CAM Report. 1998, (3)

[4] Bannon, L. J.: The Pioneering Work of Douglas C. Engelbart. In: Phylyshyn, Z. W.; Bannon, L. J. (Eds.): Perspectives on the Computer Revolution, 2nd edition, 1989, pp. 301-306

[5] Gross, M.; Würmlin, S.; Naef, M.; Lamboray, E.; Spagno, C.; Kunz, A.; Koller-Meier, E.; Svoboda, T.; van Gool, L.; Lang, S.; Strehlke, K.; VandeMoore, A.; Staadt, O.: blue-c: A Spatially Immersive Display and 3D Video Portal for Telepresence. In: Proc. of ACM Siggraph 2003, 2003, pp. 819-827

[6] Bouguila, L.; Ishii, M.; Sato, M.: Realizing a New Step-in-place Locomotion Interface for Virtual Environment with Large Display System. In: EGVE'02 Proc. of the Workshop on Virtual Environments 2002, pp. 197-207

[7] Kobayashi, M.; Shiwa, S.; Kitagawa, A.; Ichinase, S.: Tilting Disk: A Real Scale Interface for Cyberspace. In: Proc. of SID '98, 1998, pp. 333-336

[8] Darken, R.; Cockayne, W.; Carmein, D.: The Omni-directional Treadmill: A Locomotion Device for Virtual Worlds. In: Proc. of UIST '97, 1997, pp. 213-221

[9] Iwata, H.: The Torus Treadmill: Realizing Locomotion in VEs. In: IEEE Journal of Computer Graphics and Applications, 1999, 19(6):30-35

[10] CyberWalk; http://www.cyberwalk-project.org [21.10.2012]

[11] Merriam-Webster Online Edition; http://www.merriam-webster.com [25.10.2012]

[12] Panaboard; http://www.panaboard.com [25.10.2012]

[13] Poppe, R.; Brown, R.; Johnson, D.; Recker, J.: Preliminary Evaluation of an Augmented Reality Collaborative Process Modeling System. In: Proc. of Cyberworlds 2012, 2012, pp. 77-83

[14] Fitzmaurice, G.; Ishii, H.; Buxton, W.: Bricks: Laying the Foundation for Grapable User Interfaces. In: Proceedings of CHI 1995, 1995, pp. 442-449

[15] Ullmer, B.; Ishii, H.: The metaDESK: Models and Prototypes for Tangible User Interfaces. In: Proc. of UIST '97, ACM, New York, NY, USA, 1997, pp. 223-232

[16] Ganser, C.; Kennel T.; Birkeland, N.; Kunz, A.: Computer-supported Environment for Creativity Processes in Globally Distributed Teams. In: Proc. of International Conference on Engineering Design (ICED), 2005, pp. 109-110

[17] Ganser Schwab, C.; Kennel, T.; Kunz, A.: Digital Support for Net-based Teamwork in Early Design Stages. In: Journal of Design Research, 2007, 6(1/2):150-168

[18] MS Pixelsense; http://www.pixelsense.com [31.10.2012]

[19] Hofer, R.; Kunz, A.: TNT: Touch 'n' Tangibles on LC-displays. In: Proc. of the 8th International Conference on Entertainment Computing

[20] Kunz, A.; Nescher, T.; Küchler, M.: CollaBoard: A Novel Interactive Electronic Whiteboard for Remote Collaboration with People on Content. In: Proc. of Cyberworlds (CW 2010), 2010, pp. 430-437

[21] Nescher, T.; Kunz, A.: An Interactive Whiteboard for Immersive Telecollaboration. In: The Visual Computer: International Journal of Computer Graphics, 2011, pp. 311-320

[22] Tang, A.; Neustaedter, C.; Greenberg, S.: VideoArms: Embodiment for Mixed Presence Groupware. In: People and Computer XX-Engage, 2007, pp. 85-102

[23] Feiner S.; Macintyre, B.; Seligmann, D.: Knowledge-based Augmented Reality. In: Comm. of the ACM – Special Issue on Computer Augmented Environments: Back to the Real World, 1993, 3(7):53-62

[24] Mizell, D.: Boeing's Wire Bundle Assembly Project. In: Barfield, W.; Cardell, T. (Eds.): Fundamentals of Wearable Computers and Augmented Reality, 2001, pp. 447-468

[25] ARVIKA; http://www.arvika.de [01.11.2012]

[26] ARTESAS; http://www.artesas.de [01.11.2012]

[27] Beier, J.; Fruend, J.; Matysczok, C.; Reimann, C.; Rosenbach, W.; Stichling, D.: AR-PDA: Ein mobiles Produktpräsentationssystem. In: Mensch & Computer 2004: Allgegenwärtige Interaktion, 2004, pp. 311-313

[28] Kahn, S.; Olbrich, M.; Engelke, T.; Keil, J.; Riess, P.; Webel, S.; Graf, H.; Bockolt, U.; Picinbono, G.; Beyond 3D 'as-built' Information Using Mobile AR Enhancing the Building Lifecycle Management. In: Proc. of the International Conference on Cyberworlds (CW 2012), 2012, pp. 29-36

[29] Steinicke, F.; Bruder, G.; Kohli, L.; Jerald, J.; Hinirchs, K.: Taxonomy and Implementation of Redirection Technologies for Ubiquitous Passive Haptic Feedback. In: Proc. of the 8th International Conference on Cyberworlds (CW 2008), 2008, pp. 217-223

[30] Razzaque, S.; Kohn, Z.; Whitton, M.C.: Redirected Walking. In: Proc. of Eurographics (EG 2001), 2001, pp. 289-294

[31] Nescher, T.; Kunz, A.: Analysis of Short Term Path Prediction of Human Locomotion for Augmented and Virtual Reality Applications. In: Proc. of the International Conference on Cyberworlds (CW 2012), 2012, pp. 15-22

[32] Alavi, A.; Kunz, A.; Sugimoto, M.; Fjeld, M.: Dual Mode IR Position and State Transfer for Tangible Tabletops. In: Proc. of Interactive Tabletops and Surfaces 2011, 2011, pp. 278-279

[33] Google’s Project Glass; https://plus.google.com/+projectglass [accessed 2.11.2012]