Interaction between users and Augmented Reality systems:
Human-Computer Interaction of the future

By Wouter Alexander de Landgraaf,

an essay for HCI 2003/2004, Vrije Universiteit Amsterdam

Email: wadlandg@cs.vu.nl

Stdnr: 1256033

Table of Contents

Preface

For years, the computer was something mystical locked up in the basements of large companies and universities. Then came home computers, desktops and video consoles, bringing both entertainment and an increase in productivity to users. In the last twelve years, the Internet has become the largest source of information on the planet, linking together computer users from all over the world. Then came the surge in mobile computing; laptops, personal digital assistants and especially mobile phones have become common-place in todays Information Society. Computers are getting smaller and smaller and users want access to information and communication everywhere they go.
Users would like to have direct access to information and communication, where ever they might be.

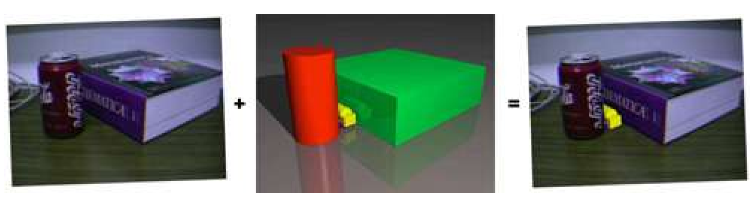
Instead of having increasingly smaller and increasingly more difficult-to-use displays, information and interfaces should be superimposed directly onto the vision of users. This is the idea behind Augmented Reality (AR). In this paper, the current state of AR will be shown, with a number of real-world applications in use today. The focus will be on the interaction between the user and the AR and the types of interfaces users will see in the near future. These new ways of interaction are sometimes called “Post-WIMP”: AR isn't concerned about Windows, Icons, Menus and Pointers, but in new ways for users to interact.

Augmented Reality, what is it?

The idea of incorporating the computer transparently into our daily lives isn't new. More than a decade ago, the computerscientist Mark Weiser termed this “ubiquitous computing”[1], where there are small computers all around us, but not directly visible to or interesting for the user. Currently, we are in the personal computing era, where the user and computer stare at each other and try to get along. Ubiquitous Computing is the next step in the evolution of computing and Augmented Reality is part of this new wave (the third wave. The first being the mainframe era, the second being the personal computing era).

Ubiquitous Computing is the opposite of Virtual Reality (VR). In VR the user is surrounded by a virtual environment. However in Ubiquitous Computing the user lives his/her life in the real world, while being aided transparently by computers.

Augmented Reality focuses on including both Virtual Reality and real-world elements together, while being interactive in real time. Instead of having programmers build the environment for interaction with the user, AR uses a person's own visual and spatial abilities, normally used for interaction with the real world, for interaction with an invisible computer. These new interfaces need to be well-integrated with the normal actions of the users. In a UR context, AR should be just as natural to interface with as a user would use an extension of the body.



Augmented Reality and visualization

Even though Augmented Reality itself seems relatively new, Sutherland and Sproul [3] already experimented with AR in the 1960's, when computers were not just cumbersome, but filled whole rooms for even trivial uses. Their idea was that “the more natural the interface is for users, the more useful the computer would become” [4]. In 1968, he developed the first see-through HMD (Head-mounted display). His system had a pair of stereo displays worn on the head, with an image which was combined with the virtual image using mirror beamsplitters. His system let a user interact with the system by changing the position and orientation of their head via a 3D tracking system. The system would change the view depending on which direction the wearer was looking at. The only problem was that there weren't any computers powerful enough to use it in any functional manner.

Of these head-mounted displays, there are currently two techniques[5] used to create Augmented Reality:

- Video see-through
- Optical see-through

The first blocks all the users vision and instead uses multiple cameras on the outside of the goggles to project the surroundings onto the inside of the goggles. Added to the videostream are computer-generated graphics, creating the AR. The largest problem of this type of AR is that there is a small amount of time for the system to react to changes while it's processing the video data. Thus, the images shown happened in the past. This is most notable when the user moves his head. Even so, it is cheaper and often easier to process than information using optical see-through devices.

The optical see-through technique doesn't have the latency problem. Devices that use optical see-through paint light directly onto the retina by rapidly moving the light source across and down the retina. This way, the user views his surroundings like usual without delay and extra information is projected over his normal vision. This seems much like the flight data that is projected onto jet windshields for years, with the exception that the data is projected directly onto the retina and is interactive with the users actions.

Sony has a range of video see-through devices (dubbed their Glasstron series), but they were sold for watching movies and have been discontinued since then. There are quite a few major entertainment companies selling these devices. They are relatively cheap. The optical see-through devices are very expensive, available from a number of other specialist companies that have been working on perfecting them. Microvision has a number of their Nomad devices, one shown in the image below. Their devices weigh about half a kilogram, but unfortunately only provide mono-color output. However, these devices are commercially available and usable for specialists. MotionResearch is pionering with head-mounted displays for use in sportsequipment, where speed, time and position are displayed onto a users eye using cheaper manufacturing techniques than normal and thus should be affordable by the general public. They are also developing displays for non-sports applications. MotionResearch is expecting to launch their *SportsVue*[tm] products in the second quarter of 2004. These are interesting steps in augmenting the users vision, but are still a long way from having true Augmented Reality. [6][7][8]

HMD Nomad Augmented Vision optical see-through system (mockup) [6]

Augmented Reality Tracking

Besides displaying the actual information, tracking the users movements and knowing how his head is orientated are the major problems. The information displayed is not something static, like in the jetfighters as stated above: AR is dynamic because it reacts to what the user is looking at and overlays information on the right position of the users vision. Tracking is done to receive information on all 6 dimensions that are necessary for AR to work: the usual 3 dimensions and pitch, roll and yaw. This is much like how planes are piloted, the similarities are clearly visible between AR tracking and aviation.
Currently, the use of GPS (Global Positioning System) is most common in outdoor tracking systems, however it is fairly inaccurate: it has an accuracy of about 20 meters. There are a number of ways to overcome this problem: the use of multiple GPS signals, used by such devices as car navigation systems and gives an accuracy of one or two meters, differential GPS, which gives sub meter accuracy and real-time kinematic GPS, which should prove accuracy on the centimeter scale and is most promising for AR. These devices are however very expensive and are still a couple of years off.
For use indoors, there are more precise solutions that have been developed for AR purposes. The most successful is the UNC Tracker project, that has a commercial spin-off

called the HiBall Tracking System. This system provides sub-millimeter position accuracy, but the biggest drawback is that it can only be used in a specially adapted room with strips on the ceiling for calibration of the system. [9] [10]

HiBall Tracking System, strips and sensor [10]

Mobile Augmented Reality

A new type of Augmented Reality is under development which could solve the tracking problem, using a device that nearly everyone is already accustomed with: the mobile phone. The use of AR to complement the ever-decreasing size of mobile phones offers the possibility to use GPRS/UMTS for tracking and as a means of using computing power that the user doesn't need to carry around with him/her.

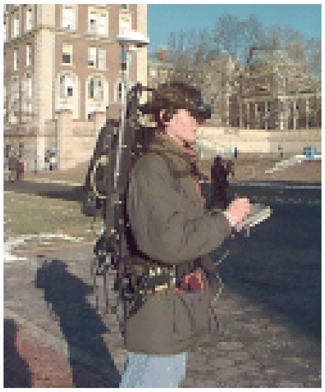
Mobile Augmented Reality was initially started by the Columbia University at their Computer Graphics and User Interfaces Laboratory. The first system was dubbed the “Touring Machine”, a wink to Alan Turing with his Turing Machine, which determines what computations are feasible using a computer. The Touring Machine, of which a paper was published in 97 and 99 [11][12], was first of its kind to bring AR out of the laboratory and into the outdoors. It used GPS and thus suffered from its inaccuracy, but did pave the way to the use of AR together with mobile devices like GPRS. [13]

Applications of AR

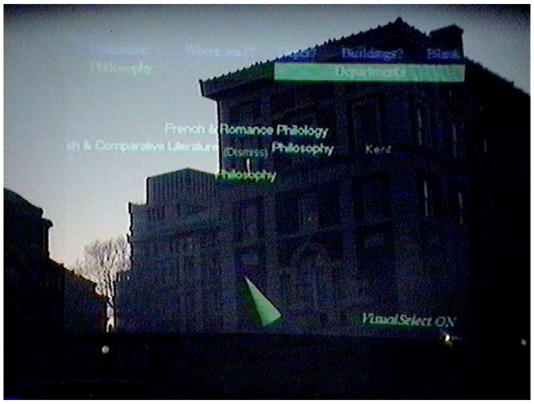
Even though a number of companies are working on optical overlaying devices and there has been research for a number of decades, commercial applications of true AR aren't available. All the problems stated above are solved or solvable, however truly combining real with virtual scenes also requires object and pattern recognition, which is primarily being done in university and research & development departments.

The Touring Machine in use

Of the prototypes the Columbia University has developed, the Touring Machine, briefly stated above, is the most known. It is made around the idea of delivering information about the university campus, with the user using an optical see-through HMD to get the augmented vision. The computer is in his backpack, as his interface device he uses a hand held computer and stylus. Labels of the building are displayed while the user walks around the campus, this hides the fact that GPS is fairly inaccurate. The user sees the pointer that comes from the position of the stylus on the touch pad of the hand held computer and is able to select various choices from the menu's which are also overlaid ("Where am I?", "Buildings?" etc). The user is also able to disable the overlay with an option, so that he can view the campus unaugmented. Different shades of colors are used to show which items are selected in the menu structure. When a building is selected, a list of URL's are shown on the hand held computer in the users hands, from which the user can browse to obtain further information. Information on the whereabouts of the selected building is shown on the HMD, a compass is shown so that the user knows he's looking at the correct building, however for the actual information the user must look at his hand held display. All buildings and menu items selected by the user are added to the collection of links on the hand held, so that the user can browse further on buildings he has seen previously. When a menu item is selected, a quick animation is executed on the HMD showing the names of the department links that have been added to the hand held device. [12]



The Touring Machine [12]



And the resulting view [12]

Is the Touring Machine a practical application? The weight of the whole system was 20 kilograms, it only worked on the campus (it was extended to a number of locations in the vicinity of the campus later on) and from a Human-Computer Interaction point of view it was more a hybrid to obtain information: The information the user would be interested in is not displayed on the HMD itself. This was one of the objectives of the researchers, to experiment with hybrid ways of using AR. Also, the resolution of the overlaid display was low and the use of GPS limited the accuracy of the AR. Nonetheless, it was a first step into making a mobile AR application using, compared to technology 6 years later, reasonably primitive means.

The team experimented with overlaying 3D architectural information directly onto the HMD. Also, the team working on the Touring Machine worked together for 2 years with the department of Journalism to create virtual documentaries using multimedia content. They integrated these two, giving the user an overlaid view of the building in the past over the current building, while giving a short spoken introduction to the user via the HMD's headphones [14] [15].

The people behind the Touring Machine also worked further on the development of user interfaces for their Augmented Reality. They ended up with a combination of the labels, flags that represent important information and the virtually overlaid buildings. [16]

3D model of "Bloomingdale Assylum" [16]

The interfaces that are always visible at the same position on the display are both the menu bar at the top and the green cone-shaped pointer at the bottom. In conjunction with the HMD, they also developed a 3D map which was visible on the handheld. This way, a user can use the map to orientate himself, while seeing points of interest. However, no information on the advantages/disadvantages of different user interfaces are given in the various papers [12] [16].

Interactive Paper

A whole different application area for augmenting is the use of interactive paper. Mackay et. al. [17] [18] have done a considerable amount of research on the use of paper in a diverse number of areas. One of their projects, Cameleon, is focused on the use of paper strips in Air Traffic Control. Although not true AR, the project did a lot of research on the use of these paper strips and a large number of previous automatize projects that all seem to have failed. From the developers there were often remarks about a reluctance to use computerized tools from the side of the controllers, but Mackay discovered that it came from the undervalued flexibility of simple paper: it is tangible and easy to annotate. These properties were too difficult or even impossible to duplicate reliably enough for the controllers. Instead of relying only on computers, Mackay built his system to add to the use of the paper strips, not replacing them. Even though not Augmented Reality in the true sense of the term as defined above, it does continue to show that there is a long way to go for AR to replace current systems. However, if AR is used to augment to previous, already working, solutions, there is a much higher chance for the system to become useful to its users.

Augmented Learning of Music

Of a lighter subject is the use of Augmented Reality as a learning assistant for learning to play a musical instrument. Cakmakci et. al. [19] used both the visual and auditorial senses to *accelerate the association of musical notes with the fingerboard of an electric bass guitar*. By overlaying markers on the strings of the guitar for playing the chords, the system helps the user in making the connection between notes and chords. The interface is very simple, as it just uses markers on the fingers and strings to determine if the user is pressing down the strings necessary, but it does show a simple interface in an area with practical use.

Augmented Entertainment

The University of South Australia has been doing research for a number of years on various forms of Augmented Reality in their Wearable Computer Lab, using both optical and video see-through devices. In particular, they have integrated their mobile AR system to play a familiar computer game, Quake. ARQuake, although still primitive, does show the possibilities of AR for use in entertainment. The input devices for the game consist of special gloves for the user and a prop gun, giving haptic feedback to its user when used. Piekarski and Thomas note in their paper: *Moving and looking around in the game world is simple, the user just has to walk in the appropriate direction for the required distance, or look in the correct direction. To shoot the weapon, the user presses the trigger on the gun prop*. [21]

ARQuake [20]

AR does have quite a future in entertainment, if done correctly. However, from their papers it seems that the developers first had to model the outside world (the campus) as black boxes, so that the system knows what to let through and where to display the opponents. This is tedious work and something AR should cope with. As such, ARQuake does offer an interesting view upon the future of gaming and entertainment, but it's far off from the goal of having a system that knows how to interact and respond to it's environment.

Multi-user Augmented Reality

Another interesting application of AR is the use of AR for collaboration means. Regenbrecht [22] as experimented with a workspace in which multiple users are working together on a single model, collaborating by making changes and annoting different parts of the model. His team used a *Cake platter*, a turnable plate-shaped device, as the location for the 3D models. Using handhelds, they were able to add models to the platter. Even more, they were able to use a plane in their hands in order to view subsections of the model and use a real flashlight as a means to provide lighting on their virtual model. Using all these different ways of interaction, the users can work together on modeling virtual objects together.

User-Augmented Reality interaction

From the previous chapters, we can summarize the following means of user interaction:

- Motion sensors, changes in 6 dimensional space
- Vision focus
- Markers and props (gun, flashlight)
- Audible interaction and feedback
- Additional hand held computer, stylus
- Gloves, hand movement

However, there is a near limitless amount of different ways a user could interact with an AR system. Normal ways of human-computer interaction, like WIMP (Windows, Icons, Menus, Pointers) can be supplemented with anything a users senses can take as input and with any action a user could come up with. A goal of AR is coming up with interaction that feels as natural as possible to the user, which is often particularly complex.

Tan et. al. [23] have done research in AR interaction by using different types of cards in engineering tasks. A special help card could be placed or held next to another card and a small animation would show help information about the card (*Tangible Bubble Help*). For a short piece of information, the system automatically showed the name of an object once it enters the users workspace area, defined to be within arms length (*Tangible Tool tips*). It appeared to be a very natural way for users, in this case engineers, to quickly gain extra information without having to resort to extra interaction with the system. There are tests underway within a large automobile company to use these results in trial applications.

Interaction between AR and users is such a complex subject, that it is linked to a broader but strongly related subject, called NUI: Natural User Interfaces. This paradigm is based on the assumption that users should interact with the virtual world just as they would with the real world (speech, handwriting, 3D projection). At the last Human Interaction symposium in 1999, there were a number of projects presented that focused on this topic: DigitalDesk projects, about the use of a real desk and the adding of virtual tools to annotate papers on each others desks, the use of retrieving computer files by using real-world objects and the use of real playing chips on a virtual playing board for casino games. It appeared that the NUI/AR-enabled games gave the players a higher chance in winning in comparison with using conventional interfaces (command line, mouse+GUI, touchscreen+GUI). Familiarity with interfaces clearly helps users in this project [24]. Their conclusion was that NUI/AR *have many advantages over traditional interaction styles and Virtual Reality*, with an emphasis on non-verbal communication (motor movements).

Conclusion: the Future of Augmented Reality

Although promising, Augmented Reality isn't there yet. There are a large number of technical challenges that need to be solved (pattern and 3D object recognition, optical character recognition, image processing, speech recognition as input, Registration, alignment issues for providing a usable output, next to a decreasing of the size of the systems), but especially the input with the user is an open book. Should there be a default set of interfaces for users to communicate with, or should this depend on the application at hand? Simply put, the field of AR has only in the past years begun experimenting with a large variety of interfaces with the user.

What if all these problems are solved and Ubiquitous Computing becomes a reality? What if we always could be connected with a virtual world, able to control computers with a wave of our hand, or by thought, or without even by thinking at all? The possibilities are endless, yet there could be a large number of risks and social problems connected to this pandora's box. Privacy, security are but a few of the topics concerned. Having computers assist us with our daily work could even cause unrest among workers: If all you have to do is follow orders, aren't you just being used as a robot?

Still, having AR augmented to our lives should make the use of computers a lot easier for users. Instead of having mobile phones, desktops, laptops, PDA's and wristwatches you could make do with a single, consistent way of controlling the mirad of computers we know today using a natural interface.

In this paper, we have looked at what Augmented Reality is, or should be, we have seen a number of Augmented Reality research projects showing the various forms possible, we have looked at the different ways humans and Augmented Reality systems interact with each other. We have also seen a large number of limitations and problems. Still, AR seems to be the only direction in which the next generation of Human-Computer Interaction can evolve. There are many scientists working on solving separate issues with AR, in the not-so-distant future AR will be reality.

References

[1] "The Computer for the 21st Century," by Mark Weiser; Scientific American, September 1991

[2] Augmented Reality at the TU Munich, http://wwwbruegge.in.tum.de/projects/lehrstuhl/twiki/bin/view/DWARF/AugmentedReality

[3] I.E. Sutherland, "Head Mounted Three Dimensional Display," in Proc. of the Fall Joint Computer Conference, 1968, Vol. 33, pp. 757-764.

[4] Howard Rheingold. Virtual Reality. Touchstone Books, 1991.

[5] How stuff works, http://www.howstuffworks.com

[6] Microvision, Nomad Augmented Vision System, www.mvis.com/nomad

[7] Sony Glasstron, www.sony.com

[8] MotionResearch, http://www.motionresearch.com

[9] The UNC Tracker Project, http://www.cs.unc.edu/tracker

[10] HiBall Tracker, http://www.3rdtech.com/HiBall.htm

[11] Feiner S, MacIntyre B. and Webster A., A Touring Machine: Prototyping 3D Mobile Augmented Reality Systems for Exploring the Urban Environment (1999)

[12] S. Feiner, B. MacIntyre, T. Höllerer, and T. Webster, A touring machine: Prototyping 3D mobile augmented reality systems for exploring the urban environment. (1997)

[13] Anders Henrysson and Mark Ollila, Augmented Reality on Smartphones (2003)

[14] T. Höllerer, S. Feiner, J. Pavlik, Situated Documentaries: Embedding Multimedia Presentations in the Real World (1999)

[15] Mobile Journalist's Workstation, http://www1.cs.columbia.edu/graphics/projects/mars/mjwSd.html

[16] T. Höllerer, S. Feiner, T. Terauchi, G. Rashid, D. Hallaway, Exploring MARS: Developing Indoor and Outdoor User Interfaces to a Mobile Augmented Reality System (1999)

[17] W. Mackay, A. Fayard. Designing Interactive Paper: Lessons from three Augmented Reality Projects (1999)

[18] W. Mackay. Is Paper Safer? The Role of Paper Flight Strips in Air Traffic Control (1999)

[19] O. Cakmakci, F. Bernard, J. Coutaz. An Augmented Reality Based Learning Assistant for Electric Bass Guitar (2001)

[20] ARQuake, http://wearables.unisa.edu.au/projects/ARQuake/www

[21] Piekarski, Thomas. ARQuake: the Outdoor Augmented Reality Gaming System (2002)

[22] H. Regenbrecht, T. Wagner. Interaction in a Collaborative Augmented Reality Environment (2002)

[23] D. Tan, I. Poupyrev, M. Billinghurst, H. Kato, H. Regenbrecht, N. Tetsutani. On-demand, In-place Help for Mixed Reality Environments (2001)

[24] M. Rautenberg. New directions in User-System Interaction: Augmented Reality, Ubiquitous and mobile computing (1999)