Diss. ETH. No. 15739

The Impact of Video Systems on Architecture

A dissertation submitted to the
Swiss Federal Institute of Technology Zurich

for the degree of

Doctor of Technical Sciences

presented by

Silke Berit Lang
Dipl.–Ing. (Univ.) / Dipl. NDS ETHZ in Architecture
University of Technology, Munich (TUM)
Swiss Federal Institute of Technology, Zurich (ETHZ)
born June 15, 1975
citizen of Germany

accepted on the recommendation of
Prof. Dr.–Ing. Ludger Hovestadt, examiner
Prof. Dr.–Ing. Markus Gross, co-examiner

Zurich, 2004

"Die Wissenschaft braucht Zusammenarbeit,
in der sich das Wissen des einen
durch die Entdeckung des anderen bereichert."
José Ortega y Gasset

Acknowledgments

This research work took its origin during my activities as a research and teaching scientist at the Chair of CAAD at the ETH Zurich.

I express my sincerest thanks to my supervisor, Prof. Dr.-Ing. Ludger Hovestadt, for giving me the opportunity to join his group and to work in a highly innovative field of work. His support and stimulating ideas during my doctoral studies have been highly valuable and helpful.

I would like to express my grateful thanks to Prof. Dr.-Ing. Markus Gross for supporting my work and for acting as a co-advisor. His guidance, support, and his capability to motivate the interdisciplinary team and bring blue-c to international success is greatly acknowledged.

I would like to thank Prof. Andrea Deplazes for taking the chair of my examination for the doctorate.

I would like to express my gratitude to all blue-c members. Special thanks go to Edouard Lamboray, Martin Näf, Christian Spagno, Stephan Würmlin, and Andrew Vande Moere for their fruitful cooperation and uncountable inspiring discussions.

During my doctoral studies, I had the pleasure to interact and discuss with a fairly large number of external scientists and experts. Mentioned among them are: Prof. Jan Borchers, Prof. Klaus Diepold, Prof. Dirk Donath, Prof. Urs Hirschberg, Prof. Gerd Hirzinger, Prof. Joachim Sauter, Markus Schäfer, Gabriele Schumann, Prof. Eckehard Steinbach, and Dr. Joachim Taiber. Many thanks to BMW and Prada for providing image material for my investigations.

I am thankful to my colleagues from the Chair of CAAD for a good working atmosphere, and to Sibylla Spycher and Manuela Cavegn for their valuable administrative support.

Last but not least, I thank my parents, my sister, and my friends for their support and encouragement.

Abstract

Today, computers and digital media are common tools and supplies in the field of architecture. Recently, architects also integrate modern information and communication technologies in their projects. The novel opportunities of these technologies enable two fields of activity: Building Intelligence and Global Communication. These two fields allow for the design of spaces that adapt to the changing social and cultural trends.

This dissertation investigates video systems, especially three-dimensional video, as a category of information and communication technologies. The thesis presents and discusses solutions how video systems can be applied to overcome space and time distances in the context of architecture. The dissertation explores the potential of three-dimensional video to interconnect the real and the virtual world by recombining the characteristics of both worlds.

This research is embedded in the historical and architectural context of video systems. Various theses are proposed which are tied up to existing technologies and historical development directions are carried on. The contributions and advantages of video systems with respect to architecture are analyzed and discussed. This research places a strong emphasis on the feasibility and elaboration in an architectural context. The development of prototype applications is the proof of concept, and demonstrates the real world value. To ensure the quality as well as the practicability, the applications and their integration into real locations are discussed with potential users from the industry. In the scope of this research, there is not a full marketable product developed, but an emergent framework of possibilities to integrate video systems into architecture.

This research is closely linked to the interdisciplinary blue-c project at the ETH Zurich. Within blue-c a new generation of an immersive projection and three-dimensional video acquisition environment is developed. blue-c integrates areas such as computer graphics, vision, communication engineering, mechanical engineering, and architecture.

Zusammenfassung

Computer und digitale Medien sind heute auch im Bereich der Architektur Standardwerkzeuge. Architekten integrieren in jüngster Zeit auch zeitgemäße Informations- und Kommunikationstechnologien in ihre Projekte. Die neuartigen Möglichkeiten, die diese Technologien bieten, bringen zwei Gebiete hervor: Intelligente Gebäude und Globale Kommunikation. Diese beiden Bereiche ermöglichen es, Räume zu gestalten, die sich den sozialen und kulturellen Veränderungen anpassen.

Schwerpunkt dieser wissenschaftlichen Abhandlung sind Videosysteme, insbesondere dreidimensionales Video, die eine Kategorie der Informations- und Kommunikationstechnik darstellen. Die vorliegende Dissertation erörtert, wie im architektonischen Kontext Raum und Zeit mit Hilfe von Videosystemen überbrückt werden können. Diese Arbeit behandelt Möglichkeiten, wie sich mit Hilfe von dreidimensionalen Videosystemen reale und virtuelle Welten miteinander verknüpfen lassen. Die charakteristischen Eigenschaften dieser beiden Welten werden dabei vorteilhaft miteinander kombiniert.

Diese Forschungsarbeit ist einerseits in den historischen Kontext von Videosystemen eingebettet und baut andererseits auf deren Gebrauch in der Architektur auf. Die vorgetragenen Thesen basieren auf existierenden Technologien und knüpfen an bestehende Entwicklungen an. Mögliche Beiträge und Vorteile von Videosystemen für die Architektur werden analysiert und diskutiert. Besonderer Wert wird auf die angemessene Umsetzbarkeit im architektonischen Kontext gelegt. Die Praxistauglichkeit der entwickelten Konzepte und Thesen wird mit Hilfe prototypischer Anwendungen überprüft. Zur Untermauerung der Qualität und der sinnvollen Einbindung in reale Szenarien werden die Anwendungen mit potentiellen Nutzern aus der Wirtschaft erörtert. Im Blickfeld dieser Forschungsarbeit wurde nicht ein marktreifes Produkt entwickelt, sondern ein Rahmen für die erfolgversprechende Integration von Videosystemen in die Architektur geschaffen.

Diese Arbeit wurde im Rahmen des interdisziplinären Forschungsprojektes blue-c an der ETH Zürich durchgeführt. Innerhalb von blue-c wurde eine neuartige immersive Projektionsumgebung entwickelt, in der Personen und Objekte mittels dreidimensionaler Videotechniken repräsentiert werden. An diesem Projekt sind vier Forschungseinheiten der ETH beteiligt: Computer Graphics Lab (CGL), Computer Vision Group (CVG), Center for Product Development (ZPE) und die Professur für CAAD.

Contents

1 Introduction

Since ever, architects make use of technologies in multiple ways and promote them. They apply nearly all kinds of natural and synthetic materials for constructing, for building, and for equipping buildings, places and sites, parks, and their environments. Architects also handle water, sanitary, heating, air conditioning, and electricity on the basis of topical technologies to improve the quality of life. Nowadays, they in addition integrate more and more modern information and communication technologies in their projects. Thus, architects contribute to the establishment and improvement of three socially important, geographical and national boarders bridging networks. Recently, communication networks experience the most significant and extensive innovations. For this reason, they are also of prime importance for architects overcoming space and time distances since life is getting more dynamic, flexible, and mobile [Der80]. The integration of information and communication technologies into architecture allows designing spaces that are adapted to the changing social and cultural trends.

What can architects contribute to this effect? How can they design environments that are more flexible and dynamic?

Building Intelligence and Global Communication

In answering these questions using information and communication technologies two fields of activity emerge: Building Intelligence and Global Communication.

The first, relatively new field is looking into a greater architectural approach to ubiquitous computing. With the help of information and communication technologies it will soon be possible to equip and net the main components of a building with computer performance. This allows the establishment of a 'nervous system' for buildings [BIB]. These buildings will provide more individual comfort to their users. The buildings are personalized, and able to react and to adapt to the behavior of their users. As a matter of fact, new interaction mechanisms have to be developed. It is crucial that the user is able to interact with this buildings without going through an extensive learning phase. It is important to address the needs of the users and to develop innovative interfaces that contribute to the quality of their everyday tasks. The ease of use has to be as simple as a light switch so that the users accept to have them in their environment [inLiv].

How can architects design interfaces where people can interact with these buildings in the way they are used to?

With every new technology, from speech to written text, to graphics, to images, to video , and to multimedia dialog, systems have emerged that have broadened the multimodal channel between remote collaborators in order to increase the quality of their interaction with each other and their common environment [Ebe04]. Analyzing recent trends, it can be observed that distant communication becomes more and more important in business and in private life. Information and communication technologies offer new ways to reduce spatial and temporal distances between objects, buildings, facilities, foundations, services, cooperating persons, and institutions. Therefore, finding the right interaction and communication techniques that let people communicate over distances in a natural manner is a challenging task. Video systems offer a promising contribution. One challenge will be bringing them closer to the users and make them part of their physical environment.

In fact, public spaces as well as private spaces are more and more equipped with video cameras. Video systems offer the possibility to connect geographical disparate spaces. Moreover, they provide novel viewing directions to physical distant users. For people to have trust in these spaces a reliable level of privacy and security needs to be established [Lev02].

How can architects contribute in returning an adequate level of privacy to the persons inhabiting these spaces?

Investigations

Based on this background, the focus in this dissertation is on video systems as a category of information and communication technologies. Special notification is on 3D video, their challenges, and their potential for architecture addressing the above-named questions.

The real world as well as the virtual world are known areas. The characteristics of the real world are actual existing scenes that are time and location dependent. As opposed, the virtual world consists of synthetic, computer generated scenes that are time and location independent. Up to now, video, in particular 3D video, is a relatively new field of research. 3D video has the potential to interconnect these two worlds and recombine the characteristics of both worlds (see Figure 1.1). The investigation of 3D video in an architectural context is a novel, not yet explored research area. The investigations described in this dissertation show how video sys-

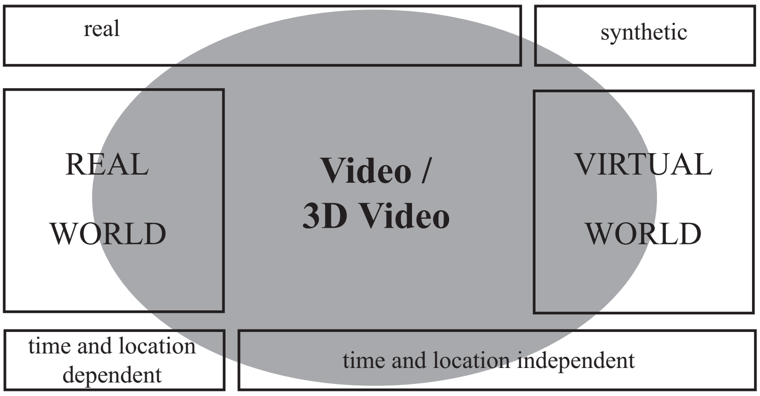


Figure 1.1: Investigate 3D video in order to interconnect the real and the virtual world.

tems with a special emphasis on 3D video can be applied to overcome space and time distances. The traditional understanding of location, space, and time may be redefined.

This scientific paper was carried out within the context of the interdisciplinary research project blue-c at the ETH Zurich. Within blue-c a new generation of an immersive projection and 3D video acquisition environment is developed. blue-c integrates areas such as computer graphics, vision, communication engineering, mechanical engineering, and architecture. As an architect, the essential contribution to the project was to embed 3D video technology into real-world application scenarios.

The excellence and main value of this work is the experimental realization of ideas and concepts. On the basis of hardware and software setups they are evaluated and revised. The development of prototypes and the implementation of practical applications underlie the postulates and the empirical investigations. They provide the basis for a generalization in the context of architecture. Within this dissertation a continuous theory based on facts has been evolved and proven.

Presentations to experts from different branches of the industry and extensive discussions have shown that the proposed concepts and applications are of real world value.

Major Contributions

This research investigates the impact of video systems on architecture. The aim is to bridge the gap between virtual and actual reality applying video systems. Video systems are an advanced form of modern information and communication technologies. They offer new ways to reduce spatial and temporal interdependencies between different activities. Therefore, a central goal of this scientific paper is to use video systems to create and design new forms of reality. Classic reality scenes are characterized by coinciding space and time co-ordinates. The novel Tele Reality scenes created and designed with the help of 3D video make completing new categories of scenes possible where coinciding as well as variable space and time co-ordinates are combined. Furthermore, synthetically created components can be added to the scene.

In this dissertation, several contributions for architecture applying video systems are made. The major contributions can be summarized as follows:

- It is shown that video systems have a significant impact on architecture. To get to this postulate, recent developments in information and communication technologies and their use in the field of architecture are carried out. In doing so, it is realized that the most profound application attributes are camera, environment, immersion, and capability.

- In order to classify video systems into suitable categories an adequate design space has been identified. A space with the following four dimensions is proposed: modality, eye contact, tele-immersion, and user realism.

- The primary interest in this work are motion scenes. In classifying visual perception by space and time, and in classifying the partnership relations between persons and objects it is figured out that Tele Reality will be a novel architectural goal. Tele-presence, communication, integration, adaptivity, personalization, comfort, security, and augmentation are essential for the realization of Tele Reality in architecture. The challenge for architects, in doing so, are the concerns of privacy, translate digital information into visual and auditory interfaces and make them part of the building, and the development of natural and multimodal interaction techniques.

- 3D video technology is investigated as a novel visual medium for architecture and human-computer interaction. It provides a time-varying view, independent from the user's perspective. Furthermore, 3D video acts as an expanding communication channel in the absence of physical contact.

- To investigate the impact of video systems on architecture and to verify the postulates the application IN:SHOP was developed and implemented. Discussions with experts from the industry guaranty the relevance to the real world and help to figure out the requirements to be fulfilled.

Organization of the Dissertation

The dissertation is organized as follows:

- **Chapter 2** presents the motivation and background of the research described in this dissertation.
- **Chapter 3** classifies and illustrates video systems. A comparative analysis of digital video and virtual reality technologies is given.
- **Chapter 4** discusses the impact of information technologies on architecture. After describing the general emergences of information and communication technologies in architecture, the focus is on the use of video systems to redefine the objects of architecture. This chapter concludes with an analysis of previously described systems and their applications, followed by a description of a novel video system.
- In **Chapter 5**, the results of Chapter 3 and Chapter 4 are merged. This chapter shows how architects can generate interactive and dynamic spaces by treating video systems like design materials integrated into the design process. First, the challenges for architects that emerge from video systems and novel technologies are discussed. By means of precise technical facts the impact on architecture is pointed out and the relating advantages for architecture are described. This chapter concludes with major criteria for developing and implementing appropriate scenarios and applications.
- **Chapter 6** describes, by means of the example application IN:SHOP, the use and the possible integration of video systems into buildings. The concept of IN:SHOP and its implementation are explained. The relevance of video systems for novel applications are discussed.
- **Chapter 7** contains the requirements of adequate ways in designing new technologically enhanced environments using video systems. Future contributions to intelligent environments are discussed.
- **Chapter 8** summarizes the dissertation and concludes with directions for future research.

2 Background

In this section, the motivation and context of investigating the impact of video systems on architecture are described. The research reported in this dissertation combines architecture, information technology, and computer graphics. Merged together, they offer new ways to reduce spatial and temporal interdependencies between different activities.

2.1 Motivation

At the Swiss Federal Institute of Technology Zurich (ETH Zurich), scientists of various departments are cooperating in modern multidisciplinary research projects. Within blue-c [Gro03], a new generation of an immersive projection and 3D video acquisition environment for virtual design and collaboration has been developed. blue-c integrates areas such as computer graphics, vision, communication engineering, mechanical engineering, and architecture. Under the direction of Professor Dr. Ludger Hovestadt the Computer Aided Architectural Design Group contributes essential application aspects to this project. The focus is centered on system integration, and the adoption and expansion of existing technologies. Traditional operating methods and new technologies are merged and investigated. In its first phase, blue-c combines simultaneous contributions of multiple live video streams with advanced 3D projection technologies in a CAVE™-like environment, creating the impression of total immersion. The technologies of the first phase of blue-c are consolidated its second phase, and the system becomes more modular and scalable.

2.2 Interdisciplinary Research

The presented work was carried out within the context of the blue-c project. blue-c created a novel hard- and software system that successfully combined the advantages of a CAVE™-like projection environment with simultaneous and real-time 3D video capturing and processing of the user. As a major technical achievement, users can now become part of the visualized scene while keeping visual contact. Consequently, these features make the system a powerful tool for high-end remote collaboration and presentation. This approach is new in the specialized literature. blue-c is currently implemented as two portals with complementary characteristics, networked with a gigabit connection. One portal is located at ETH Rechenzentrum, the second one at ETH Hönggerberg (see Figure 2.1). Various applications prove the concept and demonstrate the usefulness of blue-c.

Figure 2.1: The two blue-c portals in action. The left-hand side shows a three-sided CAVE^{TM}-like blue-c portal with three actively shuttered projection walls. The right-hand side shows a PowerWall-like blue-c portal with one projection wall. Both systems acquire the 3D video inlay with 16 cameras and have stereo projection capabilities.

The blue-c project was organized as an ETH internal research project. It started on May 1^{st}, 2000 and lasted for three years. Four ETH research groups participated in this project: the Computer Graphics Laboratory (CGL) under the supervision of Prof. Markus Gross, the Computer Vision Laboratory (CVL) under the supervision of Prof. Luc Van Gool, the Center of Product Development (ZPE) under the supervision of Prof. Markus Meier, and the chair of Computer Aided Architectural Design (CAAD) supervised by Prof. Maia Engeli at the beginning of the project and later by Prof. Ludger Hovestadt. The CGL directed the project and was responsible for the core software components, including graphics rendering and 3D video processing, and for the computing and networking infrastructure. The CVL took care of the silhouette extraction on the captured images and of the camera calibration. The ZPE was responsible for the hardware and projection setup, including the construction of the first blue-c portal at ETH Rechenzentrum. The CAAD chair investigated applications and interaction techniques, designed the virtual reality installations, and built the second blue-c portal at ETH Hönggerberg.

Figure 2.2 illustrates the framework of the core components of the blue-c project. In the following, these core components are briefly described.

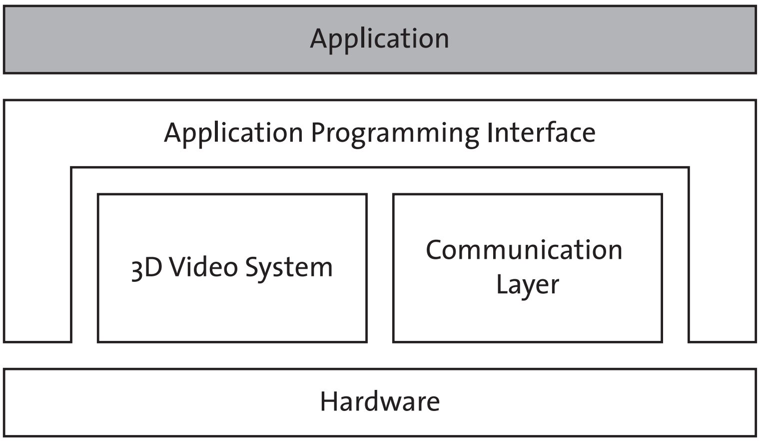


Figure 2.2: Framework of the blue-c system.

Hardware. The hardware includes the projection system, the integration of the acquisition system, the illumination and the synchronization electronics for all the components. The hardware has to satisfy the contradicting needs of projection and image acquisition. The blue-c hardware is described in the thesis of Christian Spagno [Spa03].

3D Video System. The 3D video software computes a 3D representation of the acquired persons from inside a blue-c installation in real-time. The software first builds a 3D point cloud from multiple camera video streams. The point cloud can then be efficiently encoded and streamed to a remotely located blue-c portal where it is rendered into the application. The thesis of Stephan Würmlin deals with the 3D video system [Wue04].

Communication Layer. The acquisition and rendering systems of the distributed installations are connected among each other and thus, allows for tele-collaboration applications. Strict temporal and bandwidth constraints have to be met in order to enable high quality collaboration with live video streams. Furthermore, the system needs to adapt to changing qualities of the network services. The communication layer is detailed in Edouard Lamborays thesis [Lam04].

Application Programming Interface. The blue-c Application Programming Interface (bcAPI) exposes the blue-c system functionality to the application developer. It provides a rapid application development environment which supports tele-collaboration and the integration of multimedia data into the virtual world. The blue-c Application Programming Interface is subject of Martin Naefs thesis [Nae04].

Applications. The applications run on top of the blue-c API and are customized to the requirements of the specific hardware and software layout of the blue-c in order to make use of the full potential of this new system. Two application areas have been researched in the framework of the blue-c project. Infoticles is an novel information visualization metaphor which uses the motion characteristics of particles to explore unexpected data patterns in large, time-varying datasets. This visualization technique is described in detail in the thesis of Andrew Vande Moere [Moe04]. IN:SHOP is the first example application to investigate and analyze the possibilities of integrating the blue-c technology into buildings. A novel approach to distributed shopping in a new interactive space is introduced. The physical shopping floor is connected and extended into virtual and remote spaces. The impact of video systems on architecture is the subject of this dissertation.

3 Video Systems

Video systems are an advanced form of modern information technologies. This chapter gives a comparative analysis of digital video and virtual reality technologies.

3.1 Information Technologies

The interaction of communication, computer, and electronic media technologies offers a continuously increasing number of information services since the invention of Morse Telegraphy. Microelectronics, optical technologies, advanced software solutions, and modern system architectures induce that performance, functionality, and finally the field of applications of information technological systems are rapidly growing. This tendency, on the other hand, is accompanied by a breathtaking cost decay. The scope ranges from small, always present, sometimes mobile, sometimes unconsciously observed systems (ubiquitous computing) to highly complex ones including methods of Artificial Intelligence (AI), Virtual (VR), Augmented (AR), and Mixed Reality (MR). Networked information and communication systems make high-quality multimedia, multimodal dialog, and database access possible - within buildings, but also over very long distances.

Not only engineers and computer experts, but also users with different backgrounds and worktasks are concerned with matters of Information Technology (IT). This term, in the meantime, encompasses all technological aspects of creating, modifying, processing, transmitting, exchanging, storing, requesting, and reconstructing all kinds of coded or non-coded information representations, i.e. data, text, graphics, voice, music, still images, motion pictures, scenes, and multimedia. For display and representation purposes suitable software platforms, application programming interfaces are required. The recently introduced Multimedia Home Platform (MHP), for instance, is going to speed up the convergence between digital television and the Internet [SPM01]. Information of any form at any time and any place to any person is nearly reality.

3.1.1 Historical Developments

This section describes the major developments of information technologies. The following list is not exhaustive.

Communication Technologies	Information Technologies	Media Technologies
		1452 Johannes Gutenberg: Letterpress
		1609 Newspaper
1799 Alessandro Volta: Electrical Battery		
	1822 Charles Babbage: Mechanical Calculator	
1837 Samuel Morse: Telegraph Transmitter and Receiver		
1858 Atlantic Cable		
1877 Graham Bell: Commercial Magneto-Telephone		
1902 G. Marconi: Transatlantic Radiotelegraph		
1916 NVVR: national radio society, Netherlands		
		1927 Philo Farnsworth: Dissector TV Tube
	1936 Conrad Zuse: Z1 Computer	
	1936 Allan Turing: Abstract Computer	
	1946 John von Neumann: Proposal for stored-program computer	
	1948 John Bardeen, Walter Brattain & Wiliam Shockley: Transistor	
	1951 John P. Eckert & John W. Mauchly: UNIVAC Computer	1951 Charles Ginsburg: VTR: first video tape recorder
1953 Transatlantic Cable		
1958 Arthur Schawlow & Charles Townes: Laser	1958 Marvin Minsky & John Mc Carthy: Artificial Intelligence Lab., MIT	
1960 First Digital Telephone Switch, Illinois		
		1962 Telstar 1: First TV-Satellit transmission between USA and Europe
	1969 ARPAnet: The original Internet	
	1971 INTEL 4004: First microprocessor	
	1973 Robert Metcalfe & Xerox: Ethernet (local computer network)	
1978 RSA: Rivest, Shamir, Adelman public-key cipher system		
1979 Cellular Phone: Type of wireless communication		
1980 OSI: Open System Interface International Standard		
		1989 HDTV: High-definition television
	1990 Tim Berners-Lee: World Wide Web /Internetprotocol (HTTP) and WWW language (HTML)	
		1997 DVD video
	2000 Digital TV	
	2001 MHP: Multimedia Home Platform	
	2003 UMTS: Universal Mobile Telecommunications System	

Originally, information, communication and media technologies have been developed independently of each other. The situation changed when the performance of digital information processing, transmission, and storage increased rapidly and became cost-effective. The different market sections grew together. Today networked and interleaved information, communication and media systems are more and more going to cooperate.

3.1.2 Classification

Processing and communication of information between distributed users includes a wide range of application dependent meshed devices such as computers and workstations with associated equipment, printers, plotters, copiers, visual displays, projection systems, cameras, microphones, DVD-players, and sensors. The communication services may be classified by the following different approaches.

1. Information stream:
   - interactive service, e.g. telephone
   - distribution service, e.g. radio broadcast, television
   - request service, e.g. database access
2. Information representation:
   - voice communication, e.g. telephone
   - text communication, e.g. telex message
   - data communication, e.g. multi-computer system
   - fixed-image communication, e.g. individual image transmission
   - moving picture communication, e.g. interactive television
   - multimedia communication, e.g. mixed communication media
3. Transmission and Switching:
   - analog - digital
   - narrowband - broadband
   - wire-bound - wireless
   - circuit switched - packet switched

The focus in this dissertation is on interactive 3D video services. That means, the main interest is in broadband multimedia processing and communication, in other words, a real challenge for IT engineers.

3.1.3 Processing Performance

Computers are getting more and more cheaper, smaller and more powerful. Gordon Moore prophesied more than 25 years ago that the number of transistors on a microprocessor available at a particular price would double approximately every 18 months [Moo65]. Figure 3.1 shows that Moore's law has proven remarkably accurate today. The computer industry follows this trend for already two decades. The technology of electronic grows on an exponential increase of integration in time. The results are a rapid increase in speed, performance, and complexity. Information processing systems, such as communication bandwidth, and storage capacity follow similar rules.

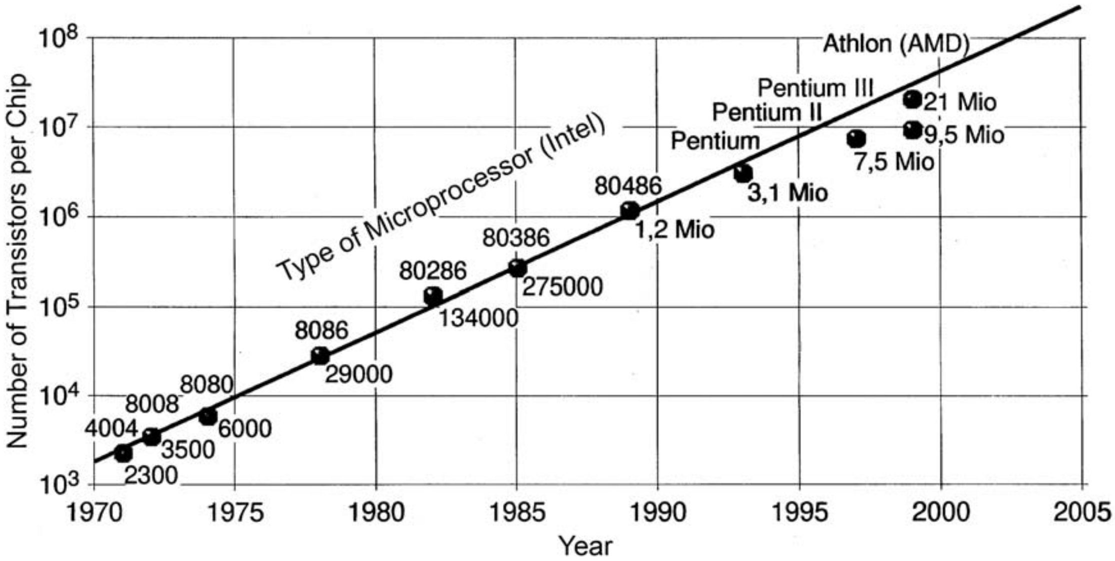


Figure 3.1: Continuous increase of transistor numbers per chip.

Ambient Intelligence

The term *Ambient Intelligence* (AmI) refers to the vision that electronic environments will react in a sensitive, responsive, and adaptive way to the presence of humans and objects [Duc01]. In order to achieve this goal, technology will become invisible, embedded in our natural surrounding, context aware, personalized, adaptive, easy to access, and autonomous acting for providing various services to people. AmI environments will be integrated in our everyday environments. This builds on three key technologies: ubiquitous computing, ubiquitous communication, and intelligent user interfaces. Ubiquitous computing describes the integration of microprocessors into everyday objects. The term was coined by Marc Weiser around 1990 [Wei91]. Ubiquitous communication enables these object to communicate with each other as well as users to communicate with these objects. Current research in

AmI deals with how the environment is able to identify and to model users' activities (see Figure 3.2) [Enc04, Nij04].

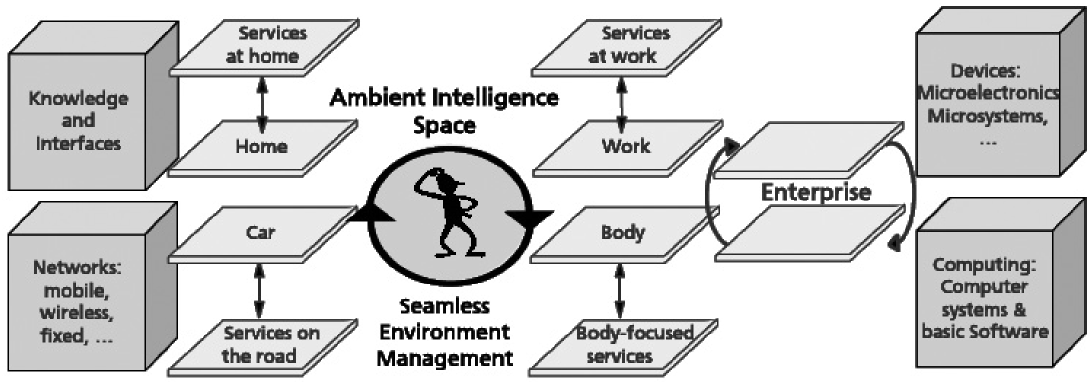


Figure 3.2: Ambient Intelligence context space.

Disappearing Computer

The term *Disappearing Computer* is a slightly modified view of everyday invisible presence of information technology. Up to now, technology is becoming more and more sophisticated, less obtrusive, and less visible. On the one hand, they disappear into building structures and devices. On the other hand, they are omnipresent, but not dominant any more. The *Disappearing Computer* is a futuristic vision and a EU-funded initiative of Future and Emerging Technologies (FET) activity of the Information Society Technologies (IST) research program [DC]. The goal is to explore how everyday objects and places can be infused and augmented with information technology. The technology providing these capabilities is merged with real objects and places.

3.1.4 Media Technology

According to the Lexikon der Physik [Phy03], *Media Technology* is a collective term describing information-media technologies for the overall description of electroacoustics, audio technology, video-, broadcast-, television-, movie- and light technology in conjunction with digital data processing. Media Technology is the basis for modern communication. Due to its rapid development new possibilities of information transmission arise. In the broadest sense media are intermediary, thus everything that enables communication and/or information transmission: speech, mimic, gesture, text, books, telegrams, broadcast and television, computer and their software, movies, photos, letters, images, media technology based devices, cables, satellite, and many more. Media Technology deals with technical bases of multi medial information systems.

According to R. Steinmetz [Ste00] media are differentiated into:

- perception
- presentation
- representation
- storage
- transmission

of different information and data types.

Perception media derives from the human senses and expresses that multimedia systems are human centered (see also Section 5.4). The major aspect in defining perception media, is the way humans receive information. The most important media are therefore visual, auditory, and haptic media. *Presentation media* correspond to input and output (I/O) devices that are available for displaying and recording multi medial information. Most of the I/O devices are human sense oriented. *Representation media* means the description of information in an abstract form, for the use in media systems. Different representation types can be used to present the same content. Therefore, multimedia systems have to support different representation types. *Storage media* are data carrier for the short-term or long-term information storage. The most common digital magnetic, electronic and optical storage media are compact disks, hard drive, RAM, CD-ROM, CD-RW, DVD, and DVD-RW. Whereas *transmission media* relate to time continued information transmission. Information is preferably transmitted from A to B without time delay. It is distinguished between wired and wire-less or mobile transmission (see also Section 3.2.2).

The term *multimedia* is used to denote the property of handling a variety of representation media in an integrated manner. Multimedia is the transmission that combines media of communication.

Multimedia definition by Fluckiger [Flu95]:

> "Digital multimedia is the field concerned with the computer-controlled integration of text, graphics, still and moving images, animation, sounds, and any other medium where every type of information be represented, stored, transmitted and processed digitally."

With the growing ubiquity and mobility of multimedia-enabled devices, Universal Multimedia Access (UMA) is emerging as an important component of coming multimedia applications. UMA deals with the delivery of media resources under different network conditions, user preferences, and capabilities of terminal devices (see Figure 3.3). “Universal” applies to the user location and time, but also to the content to be accessed [Stei03].

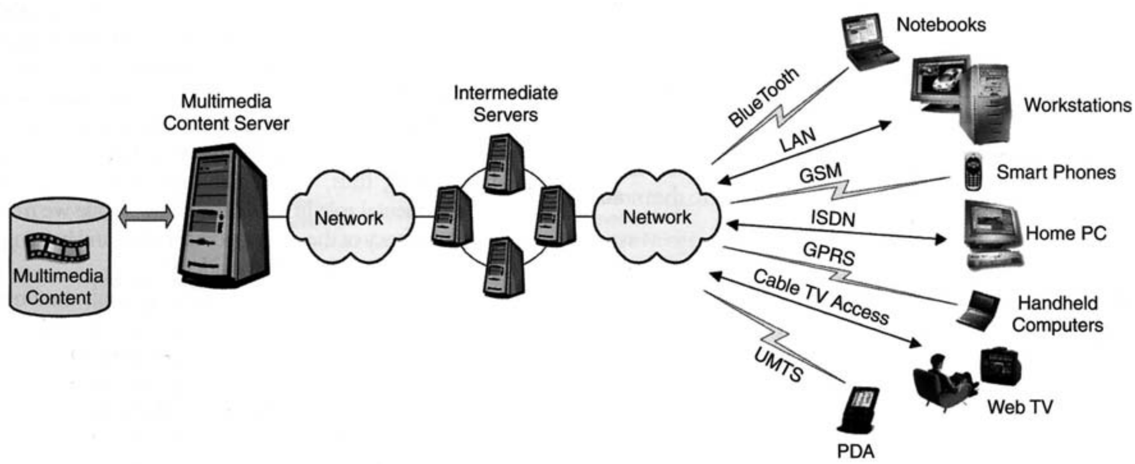


Figure 3.3: Different terminals access rich multimedia content through different networks.

3.2 Digital Video

Digital video is a video image signal represented by computer-readable binary numbers. These numbers describe a finite set of colors and luminance levels. Signals captured from the physical world are translated into digital form by digitization. To achieve smooth motioned accurate color representation digital video technologies are necessary. They are an essential part of multimedia, image communication, and broadcast industry.

The human visual system is able to organize and understand the many complex elements of environmental scenes and situations. It detects, processes, interprets, and stores pictorial information very effectively. With roughly 200 Mb/s the visual channel has the highest receptivity of all human sensory systems. It is, therefore, obvious that methods for computer aided representing, processing, transmitting, and storing of video signals have extensively broad areas of application in different fields of science and technology. They offer also most human adequate solutions. Well-known examples among others are: Evaluation of micrographs and X-ray images in medicine and technology, cell analysis, computer tomography, weather mapping from satellite data, automatic recognition and tracking of objects, pattern recognition and classification of alphanumerics, checks, documents, fingerprints, gesture and mimicry, fault detection and quality control in production processes, remote sensing and video surveillance, military applications, computer vision, robotics, television, image transmission via communication networks, image synthesis, and computer animation.

The enormous amount of information contained in a 2D image or even more in a time varying 3D scene makes image processing methods very powerful on the one hand, but requires considerable high processor performance and transmission bandwidth on the other hand. Real-time processing of an image represented by 1Mb up to 200 Mb would require, for instance, computer operations in the order of 1 Gb/s and more. Today, processing, storing, and transmitting of 2D, 3D, and 4D video signals is mainly carried out with the help of digital algorithms, with digital hardware and software technologies. Also digital television is not far away from its widespread realization. Essential interfaces between the analog optical world and the digital video world are formed by cameras at the signal input side, by viewing or projection screens, printers, and plotters at the signal output side of the system. High resolution and high image quality are required in most applications and are therefore still an item of critical research and development [SPM03b].

3D Video is another promising research area with the goal that the user can interactively choose his viewpoint and/or direction within dynamic real scenes, reconstructed from real captured imagery. There are five main categories of scene representations: omni-directional video, interactive stereo video, interactive multiple view video (free viewpoint video, 3D video objects), free viewpoint video objects, and 3D audio. In Section 4.6.3 3D video is described in more detail.

3.2.1 Video Processing

A 2D digital image is typically obtained by sampling an analog image into a two-dimensional array of picture elements, called pixels. The pixels are charged with quantized luminance and chrominance values. By reducing the number of pixels in an image pattern, the spatial resolution decreases, and image details begin to disappear. A spatial resolution similar to that of an image on a 35 mm film, for instance, is obtained with an array of 1024 x 1024 pixels, and similar to that of a standard TV image with 512 x 512 pixels. By reducing the number of amplitude quantization levels, quantization noise appears, and contour distortions result. Color coding is based on RGB (red, green, blue) primaries and various linear transformations of the standardized RGB space. The television systems YIQ or YUV, for instance, with the luminance signal Y, and the chrominance signals I,Q, respectively U,V, are compatible with existing monochrome receivers. Motion pictures are represented by a sequence of image frames. About 100 frames per second are perceived flicker-free by our visual system.

A number of image processing methods applying arithmetic operations on the pixel or bit level are suitable to improve, restore, modify, manipulate an input image, or to detect details out of it. Some of them are performed in the image space, others in a transformed space (for instance in the Fourier transform space). These methods are called low-level processes. Symbolic operations on the object level or image content level applying computer models, cognitive processes, and context knowledge are called high-level processes. Remarkable data reductions can be achieved with the help of modern image coding technologies. They make use of object based, respectively image contend based pixel structures, predict object changes from frame to frame, and apply motion estimation methods. MPEG-7 (Motion Picture Expert Group) is a metadata standard of this kind.

Static 3D objects and scenes can be digitally represented in a similar manner as 2D patterns. Three dimensional volume elements, so called voxels, replace two dimensional pixels. Another possibility for representing 3D objects is to compose them of thin slices, respectively of a large number of projections of those slices. This method is successfully practiced, for instance, with computer tomography.

Digital processing of 3D motion vision, however, is much more complex than digital processing of static vision. Changing the observers view-position during recording a scene in motion is a serious challenge for recording, processing, and playback. In fact, four dimensional space-time problems have to be solved. Domain dependent motion processing approaches use computer based models of objects, situations, and tasks. Domain independent motion processing approaches extract information from a set of time-varying image patterns. Very promising is composing 3D video by adequately superimposing and processing of 2D video streams from a certain number of synchronized cameras.

3.2.2 Video Transmission

Spatially distributed transmitters and receivers of data streams are interconnected by communication networks. Their transmission lines are physically achieved by copper or glass fiber cables or by free space electromagnetic (radio, infrared) waves. Different kinds of switching technologies exist for interconnecting terminals or communication partners, respectively. The tendency to increasing mobile users implies also increasing wireless transmission lines. Integrating wire-bound and wireless networks on the basis of adequate protocols is going to achieve efficient and flexible communication. Progress in semiconductor technologies looks promising in improving switching and transmitting performances, and reducing electrical power consumption at acceptable costs.

Media	Bit-Rate
Voice Modem	33.4 kb/s
V.34 Modem	28.8 kb/s
ISDN	64 kb/s
T1 (24xISDN)	1.544 Mb/s
Ethernet / Fast Ethernet	10 Mb/s / 100Mb/s
DSL downlink	384 ...2048 kb/s
T- DSL	2,3 Mb/s
Computer Hard Disk	20 ... 40 Mb/s
FDDI	100 Mb/s
GSM	15 kb/s
UMTS (stationary)	384 kb/s
UMTS (mobile)	144 kb/s
Cellular Phone Network	112 kb/s
DVB-T	1 Mb/s
DAB	150 kb/s

Table 3.1: Transmission rates.

Looking at raw data rates of digital video in TV quality it is obvious that without compression it is not possible to transmit video in real-time with today's transmission media (see Tables 3.1 and 3.2). Transmitting high-quality digital video in real-time requires high data rate channels. HDTV (High Definition TV) video without data compression, for instance, would need about 600 Mb/s, digitized regular television without compression about 150 Mb/s. Modern compression methods are able to considerably reduce transmission rates. Digitized regular TV, for instance, can be reduced to about 4 Mb/s with MPEG-2, or to about 0,5 Mb/s with MPEG-4 without remarkable loss of image quality. MPEG refers to standards of the Moving Picture Expert Group. Phone and cable companies are working on high-speed communication networks providing bit rates in the Mb/s up to the Gb/s (ultrawideband, UWB) range. With fiber-optic cables to the curb 20 Mb/s will be realized in 2004 [Spe03]. Business buildings and homes can get the transmission bandwidths they need for videoconferencing, video-on-demand, and multimedia services.

Standard	Year	Description	Example	Bit-Rate
ITU-R 601		International standard for digitising component colour television video, describing the 4:2:2 digital video signal	digital VTRs, TV	166 Mb/s
H.120	1984	Basis of modern video compression		1544kb/s (NTSC) 2048 kb/s (PAL)
H.261	1990	First digital video coding standard	ISDN, video conferencing, picture phone	64-2048 kb/s
MPEG-1	1992	Coding of moving pictures and associated audio for digital storage media	video CD, MP3	1,5 Mb/s
MPEG-2	1994	Generic coding of moving pictures and associated audio informationmedia	digital TV, DVD	2 - 20Mb/s
H.263	1995	The next generation of video coding performance, standard for practical video telecommunication	PSTN , picture phone	< 28.8 kb/s
MPEG-4	1999	Coding of audio-visual objects	digital TV, DVD, Inte ractive graphics application, Interactive	audio: 6 kb/s and 4kHz up to braodcast quality video: 5 kb/s to more than 1 Gb/s
MPEG-7	2001	Multimedia content description interface	digital libraries, multimedia directories, broadcast media selection	
MPEG-21	in development	Multimedia framework		

Table 3.2: Compression standards.

3.2.3 Coding and Encryption

Adequate data reduction and data compression methods are required as long as technical equipments are limited by processing performance, restricted memory capacity, and band-limited transmission channels. Therefore a major coding objective is to represent video information with as few bits as possible while preserving its image quality and maintaining low system complexity [SPM03]. Modern video coders also adapt to characteristics of the human visual system - for instance with respect to its spatial resolution and its flicker-free motion detection capabilities. Data reduction can be achieved by removing redundancy, eliminating irrelevance, taking into account spatial and temporal intraframe and interframe correlations between picture elements (pixels), estimating scene variations by predictive coding technologies and motion compensating algorithms (MC).

Common video coding algorithms can roughly be classified into the following categories and combinations of them (hybrid coding):

- In *Waveform Coders* the luminance and chrominance distribution of the pixel structure itself or a variation of it is coded. To this purpose the video image is subdivided into macroblocks of typically 8 by 8 or 16 by 16 pixels.
- In *Transform Coders* the block organized pixel structures are first transformed into transform domains which significantly differ from the luminance and chrominance distributions. The Discrete Fourier Transform (DFT), the Discrete Cosinus Transform (DCT), and others are typical transformations. Then the transform coefficients are coded. Transform coding makes it possible, for instance, to stepwisely improve the quality of a transmitted video by successively adding higher order spatial frequency components. The DCT is part of nearly every modern video coder.
- *Model Coders* are applied when 2D or 3D scenes or parts of them are to be modelled. Coded are the model parameters of synthesized objects and scenes. Performance and quality depend strongly on the model used. This coding category is worthwhile when natural video objects are combined with computer generated video, graphics, text, and sound.

Practical powerful video coders combine different coding algorithms and apply motion compensation strategies. With object based pixel structures, instead of block organized ones, a video scene may comprise several arbitrarily shaped video objects which are independently coded and transmitted. MPEG-7 is a standard of this kind.

It should be mentioned that also quantization and codeword assignment are elements of a practical coder. To represent video information with a finite number of bits, only a finite number of quantization levels can be used for luminance and chrominance values, transform coefficients, or model parameters. Various methods exist - such as scalar and vector quantization - for assigning of quantization levels and decision boundaries. The codeword assignment is responsible for a uniquely identifying of reconstructed levels.

In many applications, both the data stored within a database and the messages that are transmitted over an open communication network are confidential. In such cases, it is usual to perform special coding operations, so that the data will be safe to an authorized, and incomprehensible to an unauthorized recipient. Coding operations of this type are known as *encryption*, and the corresponding decoding operations as *decryption*. A key, known only by the authorized correspondents, is used for the encryption and decryption process. Essential security requirements for stored or transmitted data are [Kun97a]:

access control:	to prevent unauthorized access
authentication:	to confirm the identities of the communication partners
confidentiality:	to protect data against bugging and to provide traffic flow confidentiality
integrity:	to protect data against loss and manipulation
non-repudiation:	to provide proof of origin and delivery of data

A commonly used key-controlled encryption standard is the Data Encryption Standard (DES, 1977) which enciphers 64 bit blocks of plain-text into 64 bit blocks of cipher-text under the control of a 56 bit key. Increasing the key length improves security, but results in a growth of processing time.

There is an increasing demand for secure real-time video services. Encryption and decryption of video data streams in real-time requires high computing effort and may surcharge the capability of standard workstations and PCs. For this reason special hardware and software solutions and combinations of them have been developed. Another promising fast and cost-effective approach is *partial encryption* where only relevant portions of the video data are application specifically encrypted [Kun97]. Authenticity of documents and copyright information can be provided by *digital watermarking*. This is a technique where copyright notes are included at randomly chosen locations in the data stream.

3.3 Virtual, Augmented and Mixed Reality

The remainder of this section gives an overview of basic principles on virtual, augmented, and mixed reality technologies and its wide spectrum of application possibilities. In 1989, the term *Virtual Reality* (VR) was first used by Jaron Lanier in order to differentiate between his immersive computer generated environments and traditional computer simulations [Pri94]. Typically, VR is defined in terms of technological hardware. The following descriptions of VR, Augmented Reality (AR), and Mixed Reality (MR) are not exhaustive.

In the early 90s, Virtual Reality has been an elitist, expensive technology which is mainly used by laboratories for research purposes and for flight simulators. In the last few years, VR has been developing fast and its applications found their way into the industry and into the public eye at conventions and exhibits. Virtual game environments are part of today's teenagers everyday life. In reality, VR is in the meantime all around us and comes in all shapes and sizes. For example, computer and video games are tools to view VR. With VR technologies, abstract themes and complex contexts can come to life, counter to all rules of space and time.

In the Encyclopdia Britannica [EB], Virtual Reality is defined as the use of computer modeling and simulation to enable a person to interact with an artificial three-dimensional visual or other sensory environment. VR applications immerse the user in a computer-generated environment that simulates reality through the use of interactive devices which send and receive information and are worn as goggles, headsets, gloves. For many years VR inhabited our consciousness and eventually became part of our daily vocabulary. Virtual Reality is used with a couple of different meanings and often in a confusing and misleading manner. In general, VR is understood as a hypothetical three-dimensional visual, simulated world created with advanced technologies, including computer hardware and software using 3D graphics as well as special devices and various multimedia peripherals. The resulting artificial environment is presented to the user in such a way that it appears and feels like a real environment. The user interacts with the aid of special transducers and sensors, such as special goggles, and sometimes fiber optic gloves, with the virtual world moving and manipulating objects. Sometimes, the term Virtual Reality is used to refer to any virtual world represented in a computer, even if it is just a text-based or graphical representation. In general, Virtual Reality is an innovative form of human-computer interaction that opens the door to virtual worlds.

The unique features and flexibility of VR allow users to experience and interact in work-related applications such as life-like models or environments, in safety and at convenient times, while providing a degree of control over the simulation that is usually not possible in the corresponding real-world situation.

Virtual Reality can be classified into two major categories. First, the simulation of real environments. For example, the exterior and the interior of a building, or a surgery with the purpose of training and education. Second, the creation and development of a fictional environment, typically for a game. Over the past years, different application areas of this new field, including those in art, biomedical molecular science, defense technology, education, entertainment, flight simulators, gaming, medicine, prosthetics, robotics, and social interactions, have been developed (see Section 3.3.4). The most promising work-related applications are those that employ VR for visualization and representation of:

- abstract systems
  like magnetic fields, turbulent flow structures, molecular models, auditorium acoustics, population density, information flows, etc.

- real systems
  like buildings, landscapes, underwater shipwrecks, space crafts, archeological excavation sites, human anatomy, sculptures, crime scene reconstruction, etc.

- distance communication and education

- hands-on training
  military, medical, equipment operation, etc.

- orientation and navigation

- conceivable systems
  including artistic and creative work of abstract nature

The Wikipedia [WP] describes *Augmented Reality* as an emerging technology where the environment includes both virtual reality and real-world elements. The user's view of the real-world environment is enhanced with additional computer-based information or behavior. The user can access this additional information or behavior not only via a Head Mounted Display (HMD), but also by means of mobile interfaces. It is important that this information is presented on the correct geometric spot. As opposed to VR that reproduces the real world, AR enhances reality by integrating virtual objects or additional information into the physical world. Virtual worlds used in virtual environments evidence that it is difficult to duplicate the world environment around us in a computer. Therefore, VR worlds are either very simplistic or the systems are very expensive such as flight simulators. While VR in its original meaning represents complete computer immersive generated environments, AR combines real persons, objects, and environments with virtual complements. A composite view of the user is generated. Multimodal interactions between user and system are common practice. For that reason intelligent sensors are of prime importance.

Of *Mixed Reality* is spoken, when real and virtual environments are interleaved in applications and the limits between system, and the user become blurred.

3.3.1 CAVE Technology

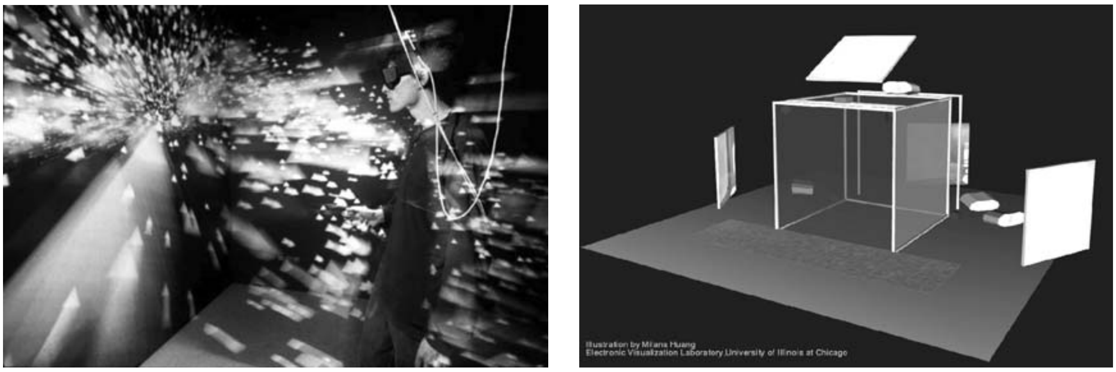
A CAVE is a virtual reality facility designed to display 3D virtual environments. It provides the illusion of immersion by projecting stereo images on the walls and floor of a room-sized cube. Communication links connect the CAVE with outside equipment and other projection walls or projection rooms.

In 1992, the Electronic Visualization Laboratory (EVL) of the University of Illinois, Chicago, invented the *CAVE* (CAVE Automatic Virtual Environment) which is commercially available through Fakespace Systems Inc. The term CAVE is also a reference to Plato's "The Smile of the Cave" in Book VII *The Republic* (ca. 380 B.C.). In his work Plato explored the concept of human perception, reality, and illusion. He used the analogy of men in a cave. A person is facing the back of a cave alive with shadows that are his only basis for ideas of what real objects are.

The CAVE [Cru93] is a projection based virtual reality system that surrounds the viewer with four screens. It was first introduced at the SIGGRAPH conference in 1992 [Cru92]. The theater measures 10x10x9 feet. Four projectors are used to throw full-color, computer-generated images onto three rear-projection screens and a down-projection screen for the floor (see Figure 3.4). In the CAVE all perspectives are calculated from the point of view of the user. A head and a hand tracker provide information about the users position. To experience the stereo effect, the user wears Stereographics' CrystalEyes liquid crystal shutter glasses which alternately block the left and right eye. Since than, more than two dozen CAVEs have been installed worldwide. Up to now, CAVEs are internationally recognized as a display environment for exploration and interaction with spatially environments for computational science, engineering, and art. Many CAVEs are used as a collaborative environment for design, visualization, simulation, education, and training.

3.3.2 Immersive VR

Immersion in the context of media and virtual environments can be described as the feeling of presence, being part of a virtual world, and playing a role in it. The convincing illusion of being part of an artificial world is enhanced by auditory, haptic, and other non-visual technologies. In movies and animations the viewer experiences



(a) Application running inside the CAVE (b) CAVE diagram

Figure 3.4: CAVE System, 1992.

the content via the camera's point of view. Therefore, the viewer is outside of the content watching it. Whereas in immersive VR the user is able to look-around, walk-around, and fly-through in the simulated environment which is presented in full scale and relates properly to the human size. These opportunities of interaction contribute to the feeling of being inside the content. The stereoscopic view enhances the perception of depth and the sense of space. Immersive projection technology characterizes virtual reality display systems. Modern immersive projection technologies let persons freely move in space.

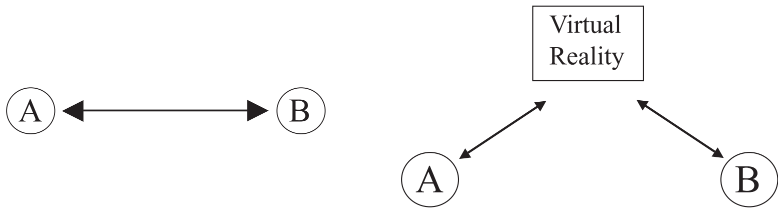
3.3.3 Tele-Presence Systems

The term *tele-presence* was coined by Marvin Minsky in 1980 in reference to teleoperation systems for remote manipulation of physical objects [Min80]. VR can also be defined based on the concept of presence and tele-presence, and therefore put in relation to other media. Based on VR technologies, and high-speed communication networks remote persons and objects collaborate in real-time in a simulated shared environment. Tele-presence refers to the mediated perception of an environment. This environment can be either a real environment, for example viewed through a video camera, or a computer generated world such as in video games.

The traditional communication model [DeF89] may preferably be replaced by the communication model of Steuer relating to Krueger [Steu92] (see Figure 3.5).

Tele-presence systems like tele-collaboration, tele-learning, tele-training, tele-robotic, and tele-medicine immerse the user in a real world that is captured by video cameras at a distant location. Therefore, tele-presence copes with the transmission of real-time information such as video and voice.

Virtual workspace generated by computer graphics and video images make collaboration between remote users possible. The objective is to give the users the



(a) Traditional view (b) Tele-Presence view

Figure 3.5: Two models of mediated communication between A and B.

feeling that they share the same space, even when they ar physically apart and in another environment. The system allows for the same possibilities in interacting with the environment and objects as if they where physically present. Tele-presence is a crucial part in the robotics research.

3.3.4 Applications

There exists a remarkable big potential for VR-AR-MR applications. However, most of them are still in their beginning phase depending on the technologies' degree of maturity, and the cost/performance ratio of necessary computer systems. Prototype systems are in prooftest in the USA, in Japan, and in Europe [Fri04]. Favored industrial areas are the automotive and the airplane industry, power stations, production engineering including robotics, civil engineering, consumption goods, sports, and games. Of great importance is the information and communication industry, and very interesting are also medical applications and tutorial systems.

The following list refers to some examples of VR-AR-MR applications migrating from research to practical and economic services. The selection is typical, but not complete. It is worth mentioning that promising applications combine VR-AR-MR technologies with adequate, preferably multimodal, human-machine interaction technologies.

VR and Robotics

Tele-robotics is commonly used to permit human operators to achieve dexterous manipulation in remote or hazardous environments. Primary, tele-robotic systems are used in conditions such as acquiring and disabling explosives, maintaining nuclear plants, maintaining high-voltage electrical transmission lines, processing biohazardous materials, and cleaning hazardous waste sites, remote controlling of engines in dangerous situations and environments. Replacing astronauts by robots in space missions is also seriously discussed by experts [Hir04]. Interactive and co-operative robot assistants applied in manufacturing processes and in home environments are worldwide successfully investigated [Lay03]. Robots offer also a number of advantages for minimal invasive surgery applications (see AR in Medicine). Visual 3D technologies contribute essentially to a more reliable robot control by humans.

ARVIKA

The goal of this project was the prototype development and investigation of user centered AR-technologies for a broad spectrum of industrial applications. The target areas were development, production, and service processes in automotive and airplane companies, in power stations, and in the machine-tool industry. The visual superposition of real and computer generated virtual objects enables the users and technicians to accomplish their task in the real environment. Application tests were performed in real scenarios with expert users of the respective application areas. Intuitive communication with computers, especially portable ones, multimodal interaction technologies such as speech and gesture recognition, novel visualization technologies, networked access to data bases as well as to various users allows effective workflow and reduces working time and costs. More than twenty research and development groups of leading companies assisted by University experts were incorporated in this project which was supported by the German government from 1999 to 2003. The project, its results, and the novel integral system architecture found international recognition. In the meantime product and service developments based on the achieved results are in action [Fri04].

Virtual Architecture and Historic City Models

Several VR-AR models collect, combine, complement, and visualize research results of archeologists and historians. Usually, the scientific findings refer to a long space of time and are fragmentary. The developed VR-AR computer models have essentially two aims: On the one hand, they offer extensive data bases and tools for further research, on the other hand, they serve as tutoring systems to visually experience history and historic developments. The users of VR based historic city models decide individually about their travelling through space and time. The 3D models of *Troia VR*, for instance, allow virtual historic travelling over a time space of about 2000 years. They are in practical use at the University of Tübingen and meet with remarkable consent by the visitors [Kir02]. Other AR guides are offering historic city sightseeing tours without a personal guide. Using see-through glasses the visitor of a historic place gets additional visual information about its history, about architectural developments, about past persons and stories. Very promising tests were carried out recently in the cities of Heidelberg and Hamburg [Schi04].

AR in Medicine

Main fields of application are education on the one hand, and surgery preparation and support on the other hand. 3D technologies combined with haptic, visual, and sometimes auditive interfaces are important. With the help of software and hardware models students navigate through blood vessels, the digestive system, and organs. They learn detecting and treating abnormalities and deceases. They can do it repeatedly until becoming enough experienced for real situations. Multimodal 3D simulation environments with haptic interaction displays exist, among others, for the clinical operational test of joint functions and for obstetrics, respectively [Rie04]. For minimal invasive surgery, for instance, AR supported navigation systems offer insight into the patient and provide preoperative examination of the individual case as well as 3D control during the physician's intervention. The development of VR-AR supported simulation, visualization, and control systems for orthodontics, and plastic surgery of jaw and face is successful in clinical tryouts [Kue04, Goe04, Schn04].

4 Video Systems in Architecture

Today, computers and digital media are common tools and supplies in the field of architecture. Since the beginning of the sixties the employment of modern technologies are considerably accompanied by developments, research work, and the use in teaching. The use of modern information technologies both as tool and as communication media, and as part of a virtual extension of the built environment leads to profound changes in the discipline of architecture.

Architects integrate more and more modern information technologies in their projects. This chapter presents fundamental projects in the context of architecture using modern information technologies. In particular, video systems and virtual reality in the built environment. First, an overview of the historical development and the classification of technologies used in architecture is given. The theories of well-known media theorists are described. Analyzing the applications discussed within this chapter and Chapter 3 leads to novel architectural goals applying video systems. The chapter ends with the description of a novel approach of a video system.

4.1 Technologies in Architecture

4.1.1 Historical Development

Figure 4.1: Historical Development: From Sutherland's sketchpad to interactive CD-ROMs.

This section describes the major developments of computer-aided architectural design (CAAD). Over time, architectonic representation progresses more and more realistic, precise, and detailed (see Figure 4.1). The history of architectonic representation starts with flat drawings such as plans, sections, and elevations. In the

20^{th} century adequate computer power and software allow for digital representations and animations. Movies, videos, and animated representations make it possible to add a forth dimension time to architectonic representations. In the following, milestones that had an influence on architectonic representation using computers are listed. This list is not exhaustive.

- 1960s/1970s introduction of computer graphics: architects start drawing their designs in computer generated formats
- 1963: fundamental starting points for CAAD (interactivity, modular design, object-oriented modelling); Ivan Sutherlands Sketchpad (thesis at Massachusetts Institute of Technology); Coons, Ross & Rodriquez, Johnson
- 1964: Christopher Alexander's *Notes on the Synthesis of Form*
- 1970: Nicolas Negroponte's *Architecture Machine*
- early 1970s: first installations of interactive computers
- 1977: William Mitchel's *Computer Aided Architectural Design*
- late 1970s: building modelling research established with Athur Britch, Arat Bijl, and Tom Maver
- early 1980s: Skidmore, Owings and Merril (SOM): development of their own software
- 1990s: 3D modelling, rendering, animation, and multimedia presentations
- 1990s: approaches to CAAD based on the ideas computers would represent the whole of a design project

4.1.2 Classification

Commonly, computer aided architectural design is understood as CAD systems and software for architects. It can be observed that CAAD is not only becoming more and more realistic, but also more and more differentiated.

1. CAAD (Computer Aided Architectural Design):
    - CAD for architects
    - visualization
    - simulation
    - modeling

  - virtual reality
  - presentation

2. Facility Management:
  - object management
  - object maintenance
  - space management

3. Computer Technology in Architecture:
  - programming design
  - computer aided manufacturing
  - building intelligence

4.1.3 Media Theory

The following remarks look into video and its use for designing enhanced environments for communication and interaction from a media point of view. *Media Theory* deals with media, its use, and what people can learn from it. It also examines what kind of information can actually be extracted from the media and how effective media are in delivering messages. Media theorists have a more general, but not at all uniform, view of the term media. A standardized definition of media does not exist. Some theories look primarily at technologies, others at functions or contents. The following unestimated extract offers insight into leading recommendations for the classification of media.

The Pross classification [Pro76], shown in Table 4.1, is very common, but does not take new digital and net based media into account. Hence, Roland Burkhart added to this classification *quaternary media* [Bur02]. Quaternary media are characterized by two features. On the one hand the traditional subdivision into sender and receiver is macerated via digitalization and networking. On the other hand all media can be integrated and referred to each other via the so called hybridization and the multi medialisation. As a result the traditional media specification resolves into print, audio, and audiovisual media.

At all times, the question about reality and truth about the content a medium transfers was linked to the concept of medium. Mostly, the function of the medium was to approximate the message of a world not available to the single human. According to Halbach and Faßler [Fas98] it can be distinguished between the following concepts:

- Mediation between reality and illusion as well as between truth and deception. In this context Plato's cave model is often cited.

- Media as a cultural addicted part of social self-characterization. The methodological basic item is part of this concept. This position means that our knowledge arises from our perception and its transfer into awareness and cognition.

- Media as an autonomous system of the pretending creation of reality. The theories that analyze the development, spread, and establishment of the media power of newspapers, movies, radio, and television.

Type	Characteristic	Description	Example
Primary Media	no special facilities for production, transport, and reception, fundamental human contacts	non-verbal language such as posture, leg and head position, mimic art, gesture, and verbal language	theater, dance, oral communication
Secondary Media	special facilities for production but not for reception	semaphores, landmarks, smoke signals or art of writing and printing are sender-sided bounded to devices; receiver require no devices	book, poster
Tertiary Media	special facilities for production and for reception	media which require technical construction phase, technical sender and technical receiver	broadcast, telefon, telegram, television, video, computer

Table 4.1: Media classification according to H. Pross.

Media are at the same time devices, instances, and nubs of world pictures. It is not possible to separate the medium from its application. That is to say, the medium is in principle non-detachable from the social locking up of [Fas97]:

- the applying coding systems such as the alphabet, numerical systems, image based drawing systems, and linguistic symbolic order.

- the time-based methods in which they are used, for example books in schools, compositions in scientific contexts, and narratives in autobiographical traditions.

- precise spatial-time events where mediation is associated with presence, attendance, presence demand or recreation.

Looking at the history of media theory [Klo00], it is realizable that there are different starting points and goals in identifying and analyzing media. Overall, it can be observed that all hypotheses try to define the inner life of the secondary and tertiary media structure. They try to find out why media have such an outstanding relevance. Two major media perceptions and approaches can be figured out in analyzing different media theorists.

1. Media acting as intermediary for communication ranging from the alphabet via the letter press to the computer. This spans both storage and transmission of information. The theories of Flusser, Postman, and Kittler are assigned to this context. Vilém Flusser regards media as uniform technical instruments. His analyzes of information technologies and communication structures range between technical science and humanities [Flus97]. Neil Postman describes culture as the environment of media and media constellations generating special patterns of thoughts and concepts of truth [Pos92]. Friedrich A. Kittler defines media as technical equipment [Kit93].

2. Media as extension or as replacement of body parts and body functions which on their parts retroact to the body. Technology, ranging from the wheel via the steam-engine to the video camera, is generally understood as media and brought in relation to the human body. This is more likely an anthropological approach. Both McLuhan and Virillio agree on looking at media as technical artefact that undergoes "reality" in a certain manner.

Faulstich, however, classifies media theories depending on range, media concept, object understanding, and fundamental theoretical reference as follows [Fau00, Fau03]:

1. Single media theory: Theories or theoretical thoughts about individual media such as movie, radio, theater, television. They are verbalized absolutely or in historical relation to different disciplines.

2. Communication theoretical media theory: Based on general terms "mass communication" or "public communication". Media are described such as with respect to their benefits, and their role for the society. This theory is essential affected by journalism science.

3. System theoretical media theory: Throws a more generalized, overriding glance at media. Interchange between subsystems such as mass media is managed by interaction media such as power, influence, money, or trust.

4. Society critical media theory: Gives special attention to political and economical roles and influences of the media.

5. Post-modern or structural media theory: Considered as the associative succession of McLuhan's global concepts. "Concepts of reality" are understood and universalized as medial concepts which are seen as staging, simulation, or illusion. This applies with different significance to the media concepts of V. Flusser, P. Virilio, J. Derrida, J. Lacan, J. Baudrillard, or J.-F. Lyotard. In Germany, similar positions are represented by F. Rötzer, N. Bolz, and F. Kittler.

Epochs	Dominant Media	Dominant Sense Organ
Oral Tribe Culture	language	ear (acoustic perception)
Manuscript Culture	hand -writing	synesthesia and sense of touch
Gutenberg Galaxy (nation states)	printing	eye (seperation emotions and mind)
Electronic Age (global village)	electricity (electronic network)	activate all senses

Table 4.2: Division of media history in four epoches according to McLuhan.

McLuhan subdivides history into four epochs [McL01]: the oral tribe culture, the manuscript culture, the Gutenberg galaxy, and the electronic age (see Table 4.2). According to McLuhan's well known statement "The medium is the message", constructed, implemented, and used information and communication systems have a selective message as the case may be effect on the social evolution and on all human beings.

Concluding, it can be noticed that the transition to a media society results from three scenarios affiliated with each other: First, the creative development in the field of applied electricity and the cumulative establishment of a electric powered sphere for tele-communication (Morse, telegraph, telegram, telephone, broadcast). Second, the almost complete alphabetization of the respective population and the resulting commercialization of documents such as newspapers, journals, novels, and the politication of the paper sector. Third, the media technical reproduction of images, motions, colors, etc. and the establishment of a *media sphere* that is closely connected to the electric powered sphere.

Looking back at time, two major observations can be stated: First, new media do not replace old media, but rather complement one another. Second, most of media innovations have been developed for other purposes than they are used today in the first place. New media eventually develop their own forms, processes, and contents. The early stages of any media or technology are characterized by the imitation of earlier uses. The new technology needs to have some time before people start to see that it is not like the previous media at all, and offers other possibilities. In a first step, new media seek to solve old and already known problems in a more sophisticated way. Later on, these already familiar problems are substituted with new opportunities using the new media. This process corresponds with Marshall McLuhan's declaration that each media is an extensions of who and what we are [McL94].

For example, the written language did not replace speech. The invention of the film camera did not replace theater, and television did not replace cinema. Media are complementary. The qualities and advantages of the respectively new medium are going to be of use and therefore, complementary and new qualities are created. Sticking at the example of the invention of the camera: Searching for application fields for the camera, people started recording stage plays, from the spectators point of view. After recognizing that this is kind of boring, the idea came up to put the camera on a trolley and zoom creating new perspectives. Unlike recording plays the new films were dynamic. The first television broadcasts, as an other example, were radio programs with pictures.

However, media extend themselves beyond the original problems that gave them rise. Therefore, analyzing media development, four interesting and important questions arise:

What was the purpose of the development of the systems?

What is the use of these systems today?

What kind of forecasts have been made?

What did really happen?

4.2 Video Spaces

Architects, designers, and artists integrate more and more modern information technologies in their projects. This section lists scenarios and projects using surveillance technologies such as video systems in the context of architecture and design.

4.2.1 Controlled Spaces

"Can technology be generative without conventions of productivity and efficiency?"
Elizabeth Diller

The interdisciplinary exhibition *CTRL[SPACE]: Rhetorics of Surveillance from Bentham to Big Brother* [Lev02] at the Center for Art and Media in Karlsruhe (ZKM) from October 12, 2001 to February 21, 2002 was devoted to the debate about the distributed and constantly expanding omnipresence of surveillance. The exhibition explored the growing impact of surveillance from traditional imaging and tracking technologies to the practice of "dataveillance". Dataveillance describes the phenomenon of social control and leaving traces in the daily life, e.g. the use of credit cards, cell phones, surfing the web, and tracking technologies from traditional video to biometric face-recognition systems. Artist and designers increasingly focus on transformations and experiences of surveillance cameras. CTRL [SPACE] dealt with the historicity of surveillance practices in the context of architecture, digital culture, video, painting, photography, cinema, installation work, television, robotics, etc.

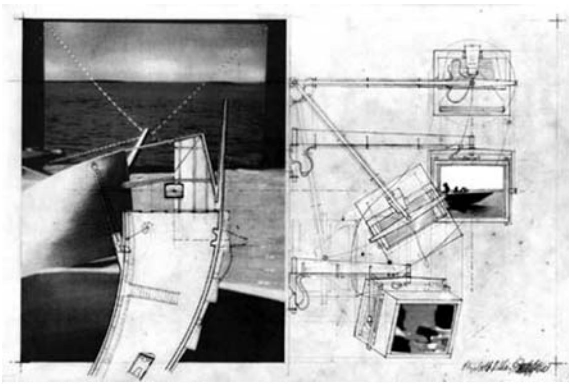
As it is well known, Diller + Scofidio use technical innovations to redefine the objects of architecture. In their work, they cross media boundaries as opposed to the common understanding of architecture (shelter, comfort, and functionality). They unite architectural design, performance, and electronic media with cultural and architectural theory and criticism. Using digital technology and video images they manipulate space and time, and make the invisible visible. Over the years they investigated the technological possibilities and integrated them into the contemporary built environment. The themes of their work include domesticity, daily life, rituals, control, security, and surveillance [Bet03].

Diller + Scofidio use the surveillance technology including cameras and monitors installed in stores, banks, train stations, etc. to illustrate the sense of presence, surveillance, and power as well as voyeurism in an architectural context. Video observation becomes visible for all, instead of being hidden for the public.

By displaying the images of the surveillance cameras they present hidden points of view on our everyday settings. Viewer and the footage occupy the same space as opposed to common cinemas where the observer is separated from the film. The technologies used in their work, for example Parasite (1989), Jump Cuts (1995), Master/Slave (1999), and The Brasserie (2000) do not produce something new. In fact, they reframe the reality by visually and temporally altering it.

The View as Presentation

One project on display is the unrealized vacation residence *Slow House* (1991) on Long Island, New York (see Figure 4.2). It belongs to the category of domestic architecture. The house is designed with a strong emphasis on the view. The view decides on the shape of the building. It features a completely out of scale large glass wall that frames the sea. This ocean view is captured by a video camera that feeds a monitor in front of glass wall. As the camera above the floor is slightly shifted towards the view the perspectives are dis-aligned. The video camera can pan or zoom in the landscape. The monitor plays back either the live video stream or a recorded sequences of the view. The monitors' version of the view is electronically manipulated, technologically duplicated reality. Thus, both views are merged together.



(a) Slow House (1991), Long Island, New York

(b) Jump Cuts (1995): A layer of crystal panels outside faces a row of video projectors inside

Figure 4.2: Diller + Scofidio: The view as presentation.

In *Jump Cuts* (1995) digital technology is used to blur clear boundaries in architectural design (see Figure 4.2). By alternating real view and view transmitted live from the inside of the United Cineplex Theater in San Jose onto liquid crystal display screens at the outside, architectural interior, and exterior are merged.

The use of surveillance methods is the distinguished feature of the redesign of *The Brasserie* (2000), the famous Philip Johnson-designed restaurant in the lower level of the Mies van der Rohe Seagram Building, in 2000. Each visitor entering through the revolving entrance door is captured by a camera. The snapshot is feed to the last of fifteen LCD monitors hanging over the bar (Figure 4.3). The other images shift to the right. Therefore, with each new customer, a new image appears and the oldest image disappears. As the bar is windowless, the monitors act as windows to the outside since the camera also captures the street and sidewalk. Surveillance technology serves as an architectonic design material in order to create transparency using an electronic medium. On the other hand, the event of entering, taking place in a single moment and location, is separated from time and place.

Figure 4.3: The Brasserie: The revolving entrance door consists of a sensor triggering a snapshot that is displayed over the bar

Facsimile (2003), for example, moves displayed text, advertising, or images on building facades one step further by displaying video images from a camera behind the screen pointing into the building. Surveillance, voyeurism, and stage drama is combined switching between live footage and fictional prerecorded vignettes. The recorded episodes augment the live view.

Interactive Theater

Joachim Sauter from ART+COM, together with André Werner, the Büro Staubach, and the Center for Art and Media (ZKM), developed a technically as well as visually amazing generative stage and costumes production for the opera *Der Jude von Malta* (The Jew of Malta) for the Opera Biennale in Munich 2002 (see Figure 4.4). This set works by recognizing movements. In his work, Joachim Sauter emphasized generativity. On the one hand, he creates virtual stage light technique that breaks up spaces, lets them become changeable: on three screens, an interactive, affectable, virtual, and generative architecture. The screens act as clipping planes. They cut the invisible virtual architecture and project it onto the stage. Thus, the virtual and the real space are connected in real-time using tracking technology. This architecture is controllable by the protagonist of the opera. Movements and gestures are registered by the computer in crossing predefined points on the stage. This allows for an appropriate interaction with the six projection screens. Due to its own generative processes, the architecture becomes more and more an acting part. On the other hand, he developed a complicated sensor-light technique for clothing the respective actors in individual color patterns.

Figure 4.4: Virtual Opera (The Jew of Malta): The projected costumes and the stage setting change via tracking the movements of the actors.

4.2.2 Mediatecture

The idea of *mediatecture* arises from the combination of the different disciplines of architecture, design, media design, and marketing in the year 1993 by the company AG4 [AG4]. Mediatecture is understood as the design of the interface between virtual worlds and the specific site.

Media Facade

At all times, the medial task of the architectural outer skin has been to represent the user's identity or to use the facade for communication purposes. Within the modern trend, this differentiation got lost in favor of rational facade construction. In the age of communication, the pressure on architecture increases to get more communicative. Since the beginning of the eighties the topic of media facades was getting into the focus of architecture. The technical development of electronic media made it possible for the first time to display moving images on large surfaces. For example, the exhibition "video sculpture" 1990 in Cologne inspired to stage public spaces atmospherically using digital images. Architecture has always been affected by new technical possibilities which were typical for the respective time. The charm of media facades and the use of new media surely is the fact that they suggest a special trendsetting direction. All creative designers burned-in the image of electronically animated facades from the movie *Blade Runner*.

(a) Blinkenlights: Heart Movie, Berlin, 2001

(b) BIX: Kunsthaus Graz, 2003

Figure 4.5: Media facades.

Blinkenlights by the Chaos Computer Club was an installation running from September 12^{th}, 2001 to February 23^{rd}, 2002 [BL]. The *"Haus des Lehrers"* (House of the Teacher) at Berlin Alexanderplatz turned into the world's biggest interactive computer display (see Figure 4.5). 144 lamps have been arranged and connected behind the building's front windows. A computer controlled each lamp independently. Using *BlinkenPaint* the public was able to create their own images and animations playing on the facade.

However, the time factor is specific for the design of media facades. So far, time is a dimension that plays a secondary role in the architectural design. With the design of media facades, a new dimension comes into architecture and into urban planning. Using software which is aligned to communication, media facades are a suitable medium in order to project the inside of a building outwards. Therefore, individual social systems become perceivable outwards. In the urban space the corporate identity of a company is direct readable on the building. For example, the media facade *BIX* by realities:united [BIX] designed for the Kunsthaus Graz, Austria, can be seen as an "urban screen". Architecture and media technology are merged together to create a communicative display skin. With the help of a digital control system schematic animations, graphics, and alphabets can be displayed. A special software allows artists to develop films, animations, and images for the BIX installation (see Figure 4.5).

The media wall of the *Vattenfall* company headquarters by plex GmbH, Berlin [plex] extends over eighteen separate Priva-Lite glass panes. This panels can be changed from opaque to translucent. Eighteen video projectors are networked with MPEG-2 players and screen five-story-high dynamic movies by internationally acclaimed artists (see Figure 4.6). Within the structure they are controlled from one location.

(a) View of media wall (b) Detail: back projection

Figure 4.6: plex: Vattenfall Media Facade.

4.3 Virtual Reality and Architecture

This section describes the use of Virtual Reality in architecture. Although VR has found good applications in the video gaming and film industry, it is not yet widespread in architectural applications.

4.3.1 Visualization and Simulation

Up to now, virtual reality and digital visualization in architecture has been primarily used as a tool for architects to better view their designs and understand the spatial relationships within them. Virtual reality and digital visualization are mainly known as efficient descriptive models of architecture and urban planning. Most people think of real-time walkthroughs, 3D interfaces, and high-end renderings from buildings that are not built yet. Currently, typical uses of VR in architecture are to simulate and present physical architecture. Actually, these aspects remain important for research work in the field of CAAD. Many architecture offices are using computers for their daily work.

In the 20^{th} century, the computer pushed the evolution of human life more than former inventions. The explosive development of hardware and software technologies in conjunction with increasing price reduction caused a new digital information impulse: computer aided 3D rendering. The computer helped planners and architects to visualize the unimaginable, to collect complex data, and transfer abstract science concepts into tangible visual presentations. Thus, the computer brought virtual reality closer to the planners and architects. This technology generates informative and impressive images. Therefore, the visual capacity of architects has been expanded. These technical potentials allow architects to explore and perceive unbuilt projects in realistic way. With the help of computer, photography, and special

(a) Outside (b) Inside

Figure 4.7: Design studio assignment "Baukunst des Schattens", M. Lorenz, ETH Zurich, 2002.

software it is possible to generate highly realistic, almost surrealistic architectural images (see Figure 4.7). Surrealistic, because they exceed real scenes, because they are not restricted to reality. They use space limitations and reality as frame for designing phantasies. Today, computers which proved to be an omnipresent tool are a crucial part in the education of architects. They are established as a medium for illustrating architecture. Today, architects present their designs and ideas using renderings, animations, and interactive CD-ROMs.

CAVE Technology and Architecture

In architecture the CAVE is mainly used for visualizing the future, planned buildings, or the past -ancient buildings. In his book *Being Digital* [Neg97] Negroponte expected that in the year 1998 more than one in ten people travelling wear head-mounted displays. Although, architectural design visualization has been thought as the ideal field of application for CAVE technology, its impact on architecture is still imperceptible. It still remains a tool for specialized research in the first place. The use of VR in architecture can be subdivided into the following categories:

- walk-throughs for visualization, analytical simulation, and virtual reconstruction
- design decision making
- collaboration
- marketing
- construction

Brian Lingard classifies VR into three stages: passive, exploratory, and immersive [Lin95]. Virtual Reality in architecture is primarily used in passive and exploratory applications. The most widespread applications are architectural walk-throughs and distant collaboration. To achieve the level of realism like in renderings in real-time immersive applications VR systems are still quite cost intensive. Therefore, real-time interactivity and understanding of human-scale spatial relationships within proposed designs become more beneficial than traditional techniques.

(a) Immersive Environment Lab at Penn State University

(b) Desk-Cave at Design Systems, Eindhoven University of Technology

Figure 4.8: VR systems in architectural education.

At the Immersive Environment Lab (IEL), a joint venture between the Pennsylvania State Information Technology Service and the School of Architecture and Landscape Architecture, students explore the potential of immersive environments within a virtual studio. From the beginning the students design in three dimensions. VR is used to help students to understand space, textures, contrast, and color. The IEL is an architectural visualization system using projection-based virtual reality to display student designs. It consists of three panorama displays and a cluster of graphics workstations. One screen is for PowerPoint, Flash, etc. presentations, the second for QuickTime video, and the third screen provides an interactive walk-through of the design (see Figure 4.8). It is also possible to use all three screens for an immersive experience and full scale walk-through of the design. Thus, students are able to position themselves in a variety of viewpoints inside their designs. This helps to reduce the learning process, to develop a spatially understanding, and a sense of scale [Kal02]. Recently, Design Systems at the Faculty of Architecture, Building, and Planning of the Eindhoven University of Technology installed a so-called *Desk-Cave* which is used to wander in VR-style through a design project, but also to develop new design tools [Ach04].

4.3.2 Video Games

In 1952, A.S. Douglas programmed the first graphical computer game, *Noughts and Crosses* - a version of Tic-Tac-Toe (see Figure 4.9a). This game is part of his PhD dissertation at the University of Cambridge on Human-Computer interaction. The first video game *ping pong* was created by Ralph Baer working at Sanders Associates in 1967. In the late seventies and the eighties, the graphical illustrations have been two dimensional. In the beginning of the nineties, the scrolling technology arose. This admitted longer journeys through virtual worlds. The complexity of the cope with a task increased continuously, like in *Super Mario*. With the invention of the 3D perspective games have been created where the player is standing behind the pawn in a game. After all, the ego-perspective came up. The player is able to witness the actions from the point of view of the pawn in a game. 1993, this perspective was widespread via the ego-shooter game *Doom*, which could be played networked using the internet (see Figure 4.10a).

(a) Noughts and Crosses, 1952 (b) Myst,1993

Figure 4.9: Graphical computer games.

Due to the Binary Space Partition (BSP), a new method to sort 3D data, it became possible to generate levels of unknown size. This levels could be played with real-time 3D graphics.

Since 1993, the adventure game *Myst* by Cayan Worlds is considered to be a milestone in the history of computer games (see Figure 4.9b). Up to now, Myst and his succession game *Riven* are the best selling computer games. Myst has set a new standard for computer games. The player is cast to explore five worlds, seeking clues, and gaining knowledge of Myst. The world is animated via images, sound effects, music, and text. By means of a point-and-click interface the player navigates and makes choices in the virtual world. Unlike traditional computer games, analytical problem solving and a multimedia environment in Myst is novel in the design of games.

(a) Doom, 1993 (b) Max Payne 2, 2003

Figure 4.10: Ego-shooter games.

At all times, computer games demonstrated the application of the latest technology in a playful way. Games are significant for design, because they help the development of technology. For a good performance, meaning running in real time and looking real, they need technological advancements in hardware and software. These technologies are instrumental in the development of CAD systems and computer graphic. The CAD systems have a significant benefit from the technologies initially developed for computer games. Today, most computer graphic approaches are hardware accelerated and use the power of the graphic cards.

Video games provide increasingly realistic and lifelike 3D VR environments (see Figure 4.10b). Thus, they are engaging. The most popular and widespread virtual world applications are ego-shooter games. Most of them such as Quake, Unreal

Tournament, and Max Payne offer editors for level design available to everybody who owns the game. These 3D accelerated game engines are able to render more than a million polygons. In comparison, most CAD software, for instance AutoCAD and ArchiCAD are able to render between 10.000 and 50.000 flat shaded polygons per second in their 3D views. As opposed to conventional 3D rendering software like 3DS Max, Maya, and Rino these game creation editors aim for real-time graphics and offer scripting possibilities to create digital dynamic and interactive spaces.

Video games are also interesting for the design of interfaces. In the field of video games advanced and successful interfaces have been developed. They feature fast graphics, artificial intelligence, and comprehensive interaction. Multi-user and collective experiences are an important part in video games.

4.3.3 VR in the Film Industry

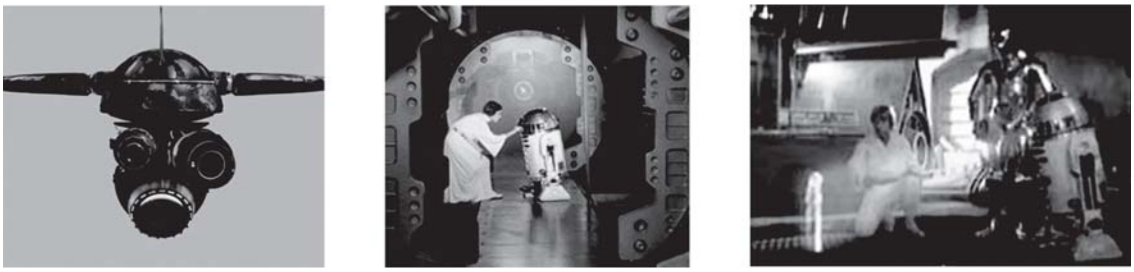
Today, Virtual Reality is an inherent part of the entertainment industry. Past and present, the film industry has portrayed VR in different ways. First, the appearance of virtual reality in films and, second, the use of Virtual Reality technology creating special effects.

Virtual Reality Visions

One of the first examples of a vision of Virtual Reality was the movie *Tron* in 1982 which looked at video games from the other side of the screen, through the eyes of the programs. This movie was also one of the first to use computer generated sequences. *eXistenZ* and *The thirteenth Floor* pose also a scenario questioning the basis of reality. Both movies are about VR player creating a VR game inside a game.

The Holodeck or Holographic Environment Simulator, described in *Star Trek: The Next Generation* shows the most comprehensive view of where VR technology may head in the future. Energy conversion create alternate realities around the participant. The Holodeck is a place that appears so real that it is difficult to distinguish it from reality. In the Star Trek series, several technological inventions such as the ship's computer, sensors and tricorders, etc. are presented. Hologram technology is moved one step further by introducing holocams and holoprojectors within the *Star Wars Episodes*. Holocams come in all manner of shape and configurations. They are mainly used for surveillance and face-to-face contact. Some of them are enriched with intelligence and serve as roving cameras for covering large events. Holocams record three-dimensional images, broadcast, or archive them. Holoprojectors carry the holograms and can be used as independent recorder. This technology enables

3D multimedia communication using 3D images floating in mid-air within the star wars episodes (see Figure 4.11).



Holocam for recording 3D messages

Leia hands R2D2 a 3D message over

Leias 3D floating-in-air image messages

Figure 4.11: Star Wars: 3D multimedia communication, 1977.

Enemy of the State is about surveillance versus personal privacy. The film features surveillance cameras that are able to rotate around an object to see what is there. Camera angles are manipulable after the recording to focus on special sciences.

Movies like Enemy of the State, Minority Report, Mission Impossible, xXx, The Net, etc. illustrate surveillance scenarios that are not yet possible. On the one hand, because the needed technology is not yet invented, and on the other hand, methods and visions are shown that are technologically already possible but outlawed.

Special Effects

Nowadays, computer generated effects become more and more important in the film industry. A lot of actions and performances are not accomplishable without the help of computer technologies. Today, the film industry is able to produce extremely realistic computer generated effects, like in The Matrix, Crounching Tiger Hidden Dragon, Star Wars, Independence Day, Spider Man, etc. Special effects are now a common extension of film making. The line between original shooting and digital effect has become blurred.

The use of computer technology in *The Matrix* (1999) is quite prominent containing more than 450 digital effect shots and about 125 “green screen” shots. The Matrix created a new style of action sequences, especially the use of ‘bullet-time photography’ (see Figure 4.12). In bullet-time photography, an array of 120 cameras that approach 12000 frames per second are setup in a circle around any given scene which then allows for computer manipulation of the speed and trajectories of the objects that comprise the shot, creating the effect of a virtual camera. This allows almost unlimited flexibility in controlling the speed and movement of on-screen elements [TM]. This approach has been copied in several films like Mission Impossible II, Romeo Must Die, and Matrix Reloaded.

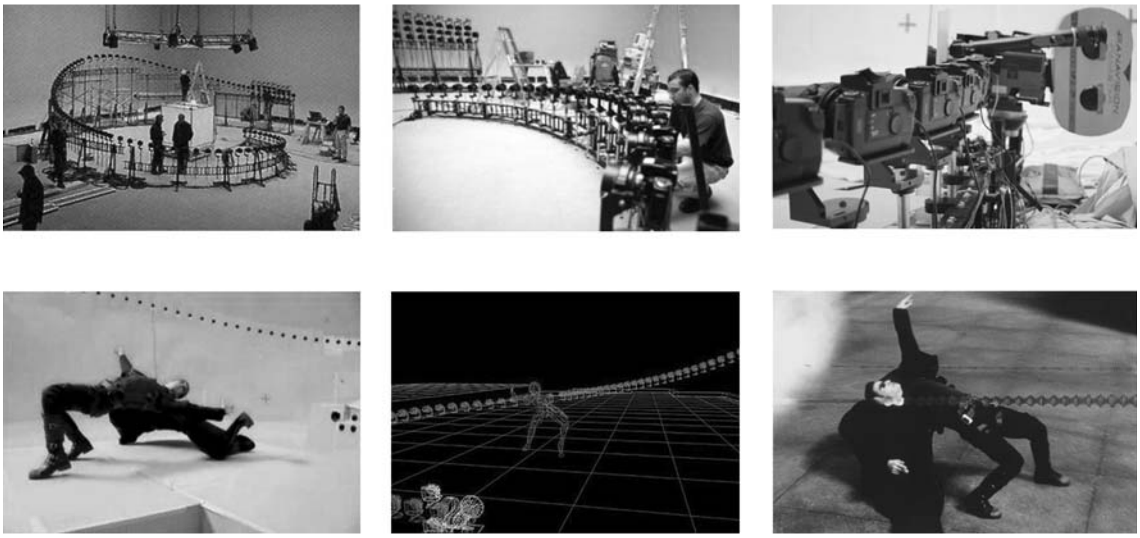


Figure 4.12: Bullet-Time-Photography, 1999.

Looking at today's film industry, it can be observed that no movie, music clip, and advertising movie goes without VR and computer graphics. VR is already an inherent part in the film industry.

4.4 Building Intelligence

"A house is a machine for living."
Le Corbusier

The term *building intelligence* describes a building that is able to acquire and apply knowledge about the user and his surroundings in order to improve his comfort, security, productivity, and efficiency. Today, buildings consist of mechanical, electrical, electronic, computational, and communication devices. As building services are getting more and more sophisticated and cheaply obtained they contain even more sensors, effectors, computer based devices, and networks. Services such as light, projectors, climate-control, multi-media devices, security mechanisms, etc. can be controlled from remote places. It is important to integrate capabilities of these services into the building.

The *Tron House* was the first "intelligent house" completed in 1988 in Nishi Azabu in the Tokyo area of Japan. The key concept was the fusion of humans, nature, and computers. The house contained over one thousand computers and sensors built into it all of them interconnected via the "Tron Architecture", a special computer architecture. The aim of the Tron project was to establish the infrastructure technologies necessary for realizing an ubiquitous-computers age [Tro89]. The Tron House was dismantled in 1994.

The actual research is looking for a greater architectural approach to ubiquitous computing (see Section 3.1.3). Intelligent building components (components equipped with computing power) are networked, and are able to sense and response to their changing environment. Hardware is systematically replaced by software allowing the buildings to become highly adaptable. The goal is to bring all active devices together providing one medium, one protocol, one server, and one user interface. Special services are configured for special spaces connected to databases [RC]. The versatility of space will be enhanced through electronic intelligence and functionality. Therefore, novel directions will be brought to this field. The technology of building intelligence influences the design, the use, and ergonomics of spaces. These intelligent environments affect the perception and appreciation of buildings.

4.4.1 Intelligent Cameras

"There was no way of knowing whether you were being watched at any given moment."
George Orwell: 1984, 1948

Video cameras become smaller, faster, cheaper, and produce higher-quality images. Today, pan-and-tilt cameras are five inches in diameter. It is now possible to squeeze the processing power of a notebook computer into a video camera. Therefore, automation systems are able to see. Public cameras enable remote users to see what they would otherwise need to visit. Such cameras can watch for anything from imperfections on production lines to the number of people on a subway platform.

This cameras are able to analyze data according to programmable criteria. Unlike cameras the human attention is limited. Therefore, the challenge of modern video technology is processing large volumes of images.

The learning camera developed at the Siemens Research Center in Roke Manor, USA, records a previously defined standard process and "learns" it as a normal sequence. As soon as a derivation from the pre-defined pattern occurs an alarm is actuated.

4.4.2 Video Surveillance

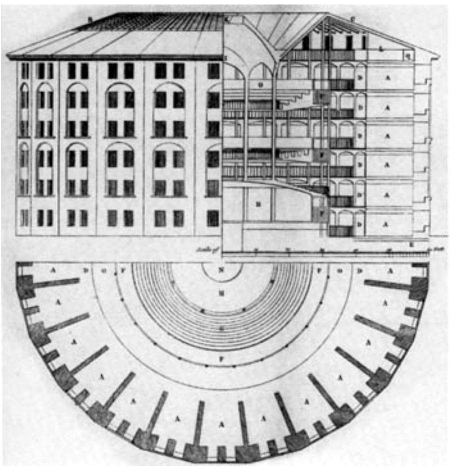
"Sed quis custodiet ipsos custodes?"
Juvenal

Moore's Law states that the price of integrated circuity falls exponentially over time (see also Section 3.1.3). The tools of surveillance are based on integrated circuits. In the last years, there has been a surge of video cameras in public spaces. The performances of this cameras rank from swivel to locate a person to zoom in. They intrude unsuspectingly on personal space. The use of video cameras for surveillance in public spaces is much more prevalent in Europe than it is in the United States. For example, the government of Great Britain placed more than two million cameras throughout public spaces over the last years. Based on rough estimates the average Londoner is caught approximately 300 times by video cameras a day. Table 4.3 lists some public locations and their estimated amount of installed video cameras used for surveillance and operation.

Location	Amount	Purpose
Austria	160 000 cameras	security and operating surveillance
Germany	500 000 cameras	security and operating surveillance
Great Britan	2 million cameras	security and operating surveillance
Switzerland	40 000 cameras	security and operating surveillance
Frankfurt Airport	2000 cameras	security and operating surveillance
Frankfurt Mainstation	120 cameras	security surveillance
Hannover Expo 2000	260 cameras	security surveillance
Las Vegas Casinos	up to 1900 cameras each	security surveillance
New York	5000 cameras	security and operating surveillance
Times Square NY	200 cameras	security surveillance
Zurich Airport	430 cameras	operating surveillance
Zurich Mainstation	100 cameras	security surveillance

Table 4.3: Amount of video cameras in public spaces.

In fact, the idea of surveillance in architecture is not new. In 1785, the British philosopher Jeremy Bentham invented the panopticon, a type of prison. The significant feature of the design was that the inspector can see each of the prisoners at all times, without being seen. The prisoners never knew for sure whether they are watched or not, thus conveying a sentiment of an invisible omniscience. In the center of a circle of the individual jail cells an observation tower has been placed (see Figure 4.13). Today, Bentham's proposed concept of the controlled space of the panopticon is a synonym for the cultures and practices of surveillance of the modern world. As an architectural logic the panopticon had a significant impact on the layout of schools, factories, and hospitals as well.



(a) Panopticon devised by Jeremy Bentham, 1791

(b) Illustration

Figure 4.13: Panopticon, 1785.

Observing Surveillance Project

The *Observing Surveillance Project* [OSP] illustrates the presence of video cameras in the capital city of the United States - Washington DC (see Table 4.4). One goal of the project was to promote a public debate about the presence of video cameras after September 11. Normally, surveillance systems are installed quietly, and hidden from view or disguised. Surveillance cameras are generally hidden so well that they stay undetected by those passing by. Another goal was to explore the use of media to promote public dialogue.

A map, handed out by the Washington tourist bureau to visitors has been modified. 15 locations of 16 cameras installed by the Washington D.C. police department have been drawn in this map. These cameras offer 360-degree views and magnify up to 17 times. Tourists and residents can now avoid these regions, or at least know that they are watched. Figure 4.14 shows an adoption of the information form the D.C. Metropolitan Police Department into a city map of Washington DC.

Camera	Location	View
1	Old Post Office Pavilion 1100 Pennsylvania Ave. NW	Pennsylvania Ave. NW from 14th St. to the Capitol
2	Smithsonian Institution Castle 1000 Jefferson Dr. SW	Entire Mall in both directions
3	L'Enfant Plaza 480 L'Enfant Plaza	Southbound I-395, Pentagon, Reagan National Airport
4	U.S. Department of Labor 2nd St. and Constitution Ave. NW	Capitol, Intersection of Constituion and Pennsylvania Aves. NW, and 3rd St.
5	Voice of America 3rd St. and Independence Ave. SW	Independence Ave. from the Capitol to 14th and 3rd St. north to the Department of Labor
6	Dupont Circle 1350 Conneticut ave. NW	Dupont Circel Area
7	Park Tower 1001 N. 19th St., Arlington	Key Bridge, Potomac River, Kennedy Center and the D.C. shoreline along Potomac
8	Union Station 520 N. Captiol St. NW	Plaza in front of the Station
9	Hotel Washington 15th St. and Pennsylvania Ave. NW	15th St. and Pennsylvania Ave. NW between 12th St. and the White House
10	Banana Republic M St. and Wisconsin Ave. NW	Wisconsin Ave. at M St. NW
11	National Gallery of Art East Wing 3rd St. and Constitution Ave. NW	Only for special events at the request of the building management
12	Columbia Plaza 24th St. and Virginia Ave. NW	Whitehurs Freeway, Roosevelt Bridge and Memorial Bridge
13	Hilton Hotel 1919 Conneticut Ave. NW	Hotel surroundings and Conneticut Ave. down to Dupont Circle
14	World Bank I 19th St. and Pennsylvania Ave. NW	World Bank buildings surroundings and Pennsylvania Ave. between 22nd St. and 17th St.
15	World Bank II 18th St. and Pennsylvania Ave. NW	World Bank buildings surroundings and Pennsylvania Ave. between 22nd St. and Old Executive Office Building

Table 4.4: Location of the cameras and their views (Source: D.C. Metropolitan Police Department).

Casino Surveillance in Las Vegas

Las Vegas, also called Sin City, is a living lab for surveillance technology. The Strip is a testbed for surveillance. The casinos spend millions of dollars on the latest high-

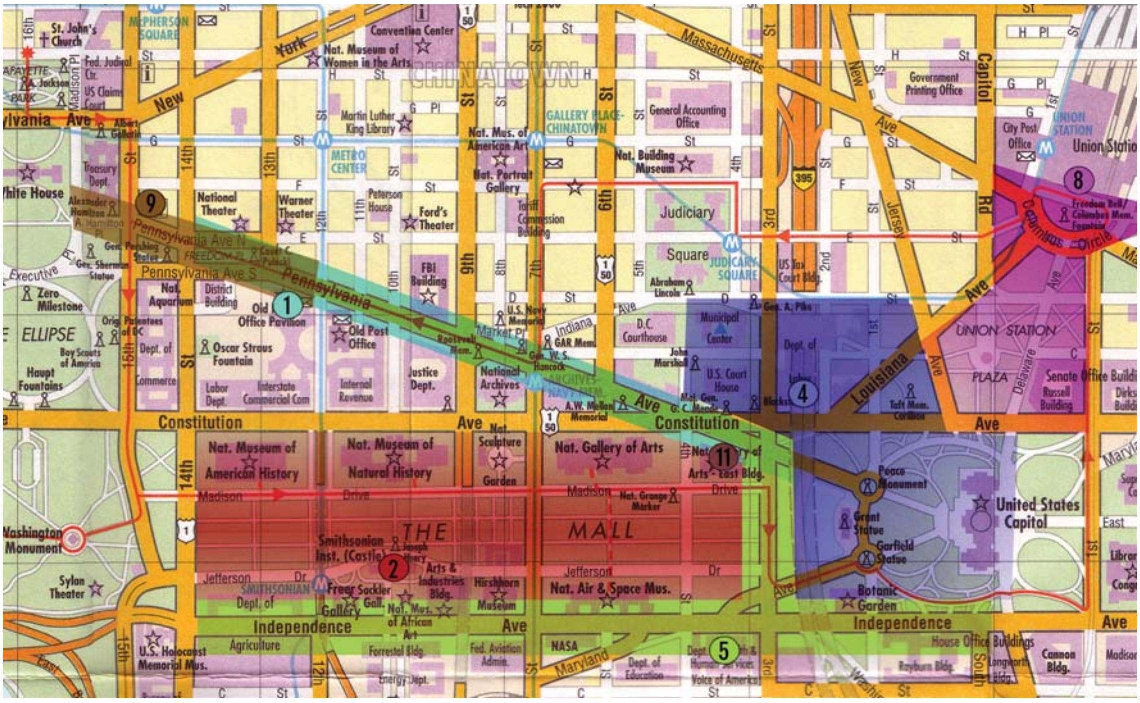


Figure 4.14: Map of selected surveillance cameras and their views.

tech security technology, staying at the cutting edge of security, to track customers and employees. As soon as someone enters a casino he is watched. Everything that happens inside the casino walls is recorded 24 hours a day and at least archived for seven days.

The Bellagio, for example, consists of a surveillance system as sophisticated as the system currently used by the government. The system contains 1900 cameras, facial recognition, backup system, and half a dozen surveillance experts. The cameras tilt, pan, and zoom in on any suspicious activity, or track a person from one point to another. Non-Obvious Relationship Awareness (NORA^{TM}) software developed by Systems Research and Development searches databases detecting hard-to-see links. Information taken from images from the video cameras are mixed and matched with retinal scans, fingerprints, voice patterns, and information from digital ID cards. Speed, networking, and shared databases allow to identify people fast and efficient, even if they change the casinos.

The analog-tape surveillance has moved to digital. One obvious advantage of digital surveillance is that no tape changing is required, and therefore no actions are unrecorded. Digital information is stored on hard drives or optical storage systems. Smart backup systems allow the easy access of crucial information.

Despite of improving surveillance technology, the casinos relay on smart, properly trained people. After all, crooks and criminals always find a way to circumvent the latest technology.

Airport Zurich Kloten

The video surveillance system at Zurich Airport is used for optimizing passenger flows. The video cameras are installed in the following areas: boarder control, security, and waiting zones. The video surveillance at Zurich Airport is not area-wide, but rather certain points are covered. Precisely, these points are the coverage of country transfer as well as areas around the passport, and security control. Persons are always covered from behind. The video footage is only for employees disposition. The video material is only digital archived for 15 to 120 minutes. The cameras installed by Zurich Unique Airport are not used for the surveillance of passengers. There are no video cameras installed in the check-in area. There are approximately 426 cameras located in the airport building, 260 of them within the access area. The airport and security zone consists of 60 cameras and the airport parking is covered by 70 cameras. Further cameras are at the airport tower (10), the customs (12), and within the road tunnel (14). At the moment there are two analog surveillance rooms where several operators working at once watching the action on video screens and monitors. Unique is planning a new digital video surveillance system with only one surveillance room. The custom, police, and airport system are going to be networked. In the future, it is also planed to install cameras at the terminals.

The face recognition system installed for testing purposes in 2002 at Zurich Airport runs by the government of Switzerland and the Zurich Airport Police. The system is located in the travel and transportation areas targeting illegal immigrants from West Africa, Middle East, and Asia. Among other things, video surveillance should identify passengers and ascertain the airline they used to enter Switzerland. After leaving the airplane passengers are captured by a video camera. The face recognition software Facesnap by C-Vis analyses selects facial sections and compares them with database information. The database consists of images of refused asylum seeker.

4.5 Application Attributes

To conclude Chapter 3 and Chapter 4, attributes of the applications illustrated within these chapters are listed in Table 4.5. The most profound attributes are camera, display, environment, immersion, and capabilities. In the following, each of them is described in more detail.

		observing surveillance project	video surveillance	digital effects	video games	animation	architectonic represenation	media facades	interactive theater	controlled spaces	vr and robotics	image ammination	classical cave applications	movie maps
cameras	no camera						X	X						
	real camera recorded											X		X
	real camera real-time	X	X						X	X				
	virtual camera recorded			X		X					X			
	virtual camera real-time				X								X	
display	no	X												
	secure		X											
	presentation			X	X	X	X		X		X	X	X	X
	public							X		X				
environment	real	X	X	X					X	X	X	X		X
	synthetic				X	X	X	X					X	
immersion	no	X	X	X	X	X	X	X		X	X	X		X
	yes								X				X	
capabilities	interaction				X				X		X		X	X
	communication	X						X				X		
	augmentation			X			X			X				

Table 4.5: Application attributes.

Cameras

This attribute classifies the applications by the signal dimension of the use of cameras plus whether the video footage is recorded or real-time streamed. The use of a virtual camera is also taken into account. It is also distinguished between recorded and real-time. The recorded virtual camera allows the user to freely navigate within the application instead of following a predefined camera path. The real-time virtual camera refers to the ability to interpolate in real-time between two video cameras:

no: within the application no cameras are used

recorded: within the application recorded camera images are used

real-time: within the application real-time camera images are used

recorded virtual camera: the application consists of a virtual camera

real-time virtual camera: the application consists of a real-time virtual camera

Display

This attribute indicates how the applications are presented. There are four major categories:

no: the application is not displayed

secure: only authorized users are allowed to view the application, e.g. security staff

presentation: the application is open to everyone, but the user has actively to show an interest in viewing the application, e.g. buying the application

public: the application is displayed to the public without being asked, similar to advertisement

Environment

The environments in the different applications are created using different kinds of media. Realism describes whether an environment has a synthetic or a real representation within the application. There are two major approaches: *synthetic* and *real*.

Synthetic refers to computer generated environments.

Another approach is to capture the *real* existing world. This is called real.

Immersion

Immersion refers to how a system can make users feel like they are experiencing an alternate reality and not just merely observing it. The degree of immersion depends on how good the system supports the user in feeling being part of the application and not only observing it. Immersion can be achieved either through stereoscopic depth or by overlaying the real world. *Yes* means immersion exists, *no* means it does not.

Capabilities

For applications, this attribute describes the primary capabilities the systems provide to the user. These three capabilities are the most relevant within this work:

interaction: describes the possibility for the user to directly interact with the system

communication: describes whether the application is meant for displaying information to the user or even more let him communicate to other users

augmentation: describes the superimposition with other applications or reality

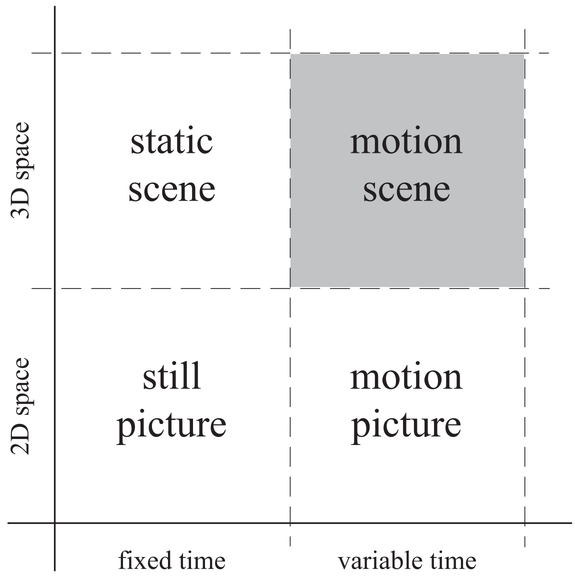


Figure 4.15: Classification of visual perception.

The primary interests in this dissertation are motion scenes. Figure 4.15 classifies visual perception by space and time co-ordinates.

By having a closer look on motion scenes and classifying the partnership relations between persons and objects it becomes obvious that the novel architectural goal applying video systems will be the design and development of *Tele Reality*. Figure 4.16 illustrates these relations using time and space co-ordinates. Chapter 5 makes evident how the idea of Tele Reality can be realized. In the following, the different subdivisions of these co-ordinates are described in more detail.

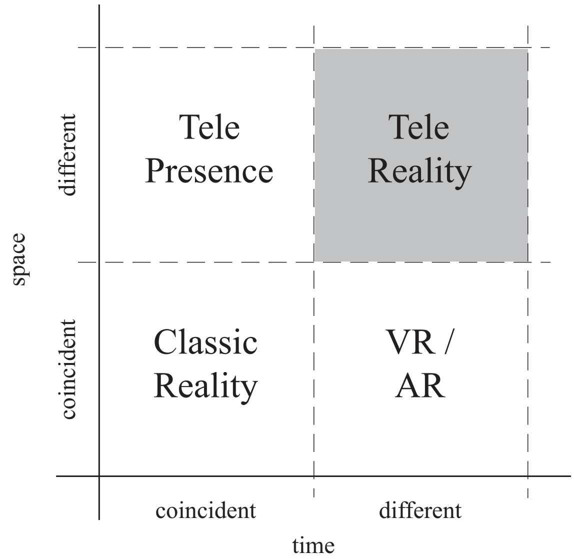


Figure 4.16: Partnership relations between persons and objects.

Space Co-ordinate

coincident space: Persons and objects are at the same location, for instance in a shop, an office, or a private home.

different space: Persons and objects are at different locations, for instance, shops, offices, or private homes are in different buildings, cities, or even countries.

Time Co-ordinate

coincident time: Presence or model of persons and objects happen at the same time.

different time: Presence or model of persons and objects may happen at the same time as well as at different times. A scene, for instance, contains not only real-time persons and objects, but also respectively models of them which were recorded in the past.

4.6 Novel Approaches

The blue-c project (see Section 2.2) provides a collaborative tele-presence environment with simultaneous acquisition of 3D video and projection. Each user finds himself in an immersive virtual environment that displays a 3D representation of the common virtual space and the other user's 3D video representation. Within blue-c a 3D video system has been developed.

In this section a few technical annotations, without entering into details, are given to emphasize from an architectural view how the experimental blue-c system, a novel video system, has been realized.

4.6.1 System Overview

Figure 4.17 displays an overview of the system architecture. The setup is asymmetric. Besides costs, the major reason for this asymmetric design is to demonstrate scalability and the integration into buildings.

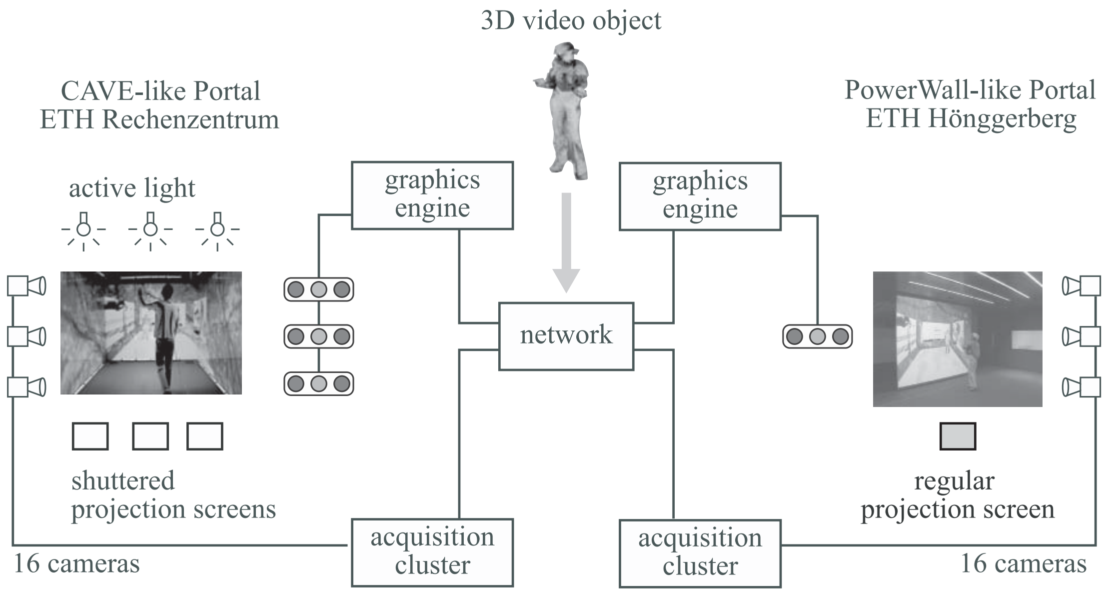


Figure 4.17: Conceptual components of the blue-c system, including hardware and software.

The portal at ETH Rechenzentrum (see Figure 4.18) represents the experimental development of the core blue-c technologies between 2000 and 2003. Using six projectors the virtual world is projected onto three shuttered glass screens. The user experiences the virtual world in full 3D using stereo glasses. Simultaneously, 16 cameras create a 3D video representation of the user in real-time. The resulting video inlays are transmitted to the remote blue-c portal. During the acquisition process, the cameras can "look through" the projection walls by employing screens with switchable opacity. During a short time slot the walls are "opened" for image acquisition. But since the shutter glasses are totally opaque in this phase the user is not aware of the cameras or the transparent walls. The shuttered projection screens, actively shuttered projection system with six projectors and the shuttered stereo glasses are synchronized.

Figure 4.18: Unfolded panoramic picture of the CAVE^{TM}-like portal at ETH Center showing the three projection screens in transparent mode located in the server space of the ETH Zurich.

As opposed to the portal at ETH Rechenzentrum, the portal at ETH Hönggerberg (see Figure 4.19) consists of 16 cameras, a regular projection screen and a stereo projector. The installation at ETH Hönggerberg was especially developed to create a public presence for the blue-c project, and to represent a multifunctional room for the whole Department of Architecture.

Figure 4.19: Panoramic picture of the PowerWall-like blue-c portal at ETH Hönggerberg showing the projection screen integrated into the building.

In each portal spatial audio rendering enables exact location of sound sources of virtual objects. This technology also allows real-time speech transmission of participants.

4.6.2 Acquisition of Large Real-World Environments

While new information and communication technologies continuously advance and the prices of powerful computers decrease (see Section 3.1.3), relatively little work has been devoted into the design of large scaled scenarios which combine the recorded and live streamed video data. ADVISOR [IST99] - Annotated Digital Video for Intelligent Surveillance and Optimized Retrieval - is an example for intelligent transport infrastructure and mobility management. The project aim is to use multi-camera CCTV to monitor and track passengers in subway stations in order to obtain early warning of hazardous situations and to accumulate statistics necessary for the safe and economic operation of the transport network [Val03]. Safety and security of everyday life is improved by MISSY [IST00] where an intelligent, robust, and affordable solid-state microsystem enables fast and parallel acquisition of 3D-information. The Robotics Institute at Carnegie Mellon University (CMU) and the Sarnoff Corporation developed a system for autonomous Video Surveillance and Monitoring under the Video Surveillance and Monitoring (VSAM) project [Val03]. Technically, this approach uses multiple cooperative video sensors to provide continuous coverage of people and vehicles in a cluttered environment. The developed technology will enable a single human operator to monitor activities over a large, complex area using a distributed network of video sensors.

As opposed to these approaches, blue-c-II [Hov03] focuses on how to design, configure, and manage large, complex, and hybrid installations. The background in architecture and building intelligence contributes to these existing surveillance systems. Using the blue-c research descriptive novel real world applications will be investigated and analyzed. Novel methods for interactive view-independent 2D and 3D video, and display technology are able to acquire large, open, and complex physical environments.

Given the results achieved within the first phase of blue-c, the obvious next step is to open the real, physical world to the user. In the first phase of blue-c users at different physical locations are brought together in a shared virtual environment. The system features a small range of operations confined to dense environments. Whereas in the first phase the user interacts in a prefabricated displayed scene, the goal of the next phase is the interaction with large, complex real-world scenes. Combining blue-c-I and blue-c-II, it is possible to create 3D video both of objects and entire scenes.

In blue-c-II the cameras will not only be turned inward, in order to build a model of the user standing inside the cave, but most cameras will actually be pointed outward and possibly quite remotely from the cave, towards different locations within

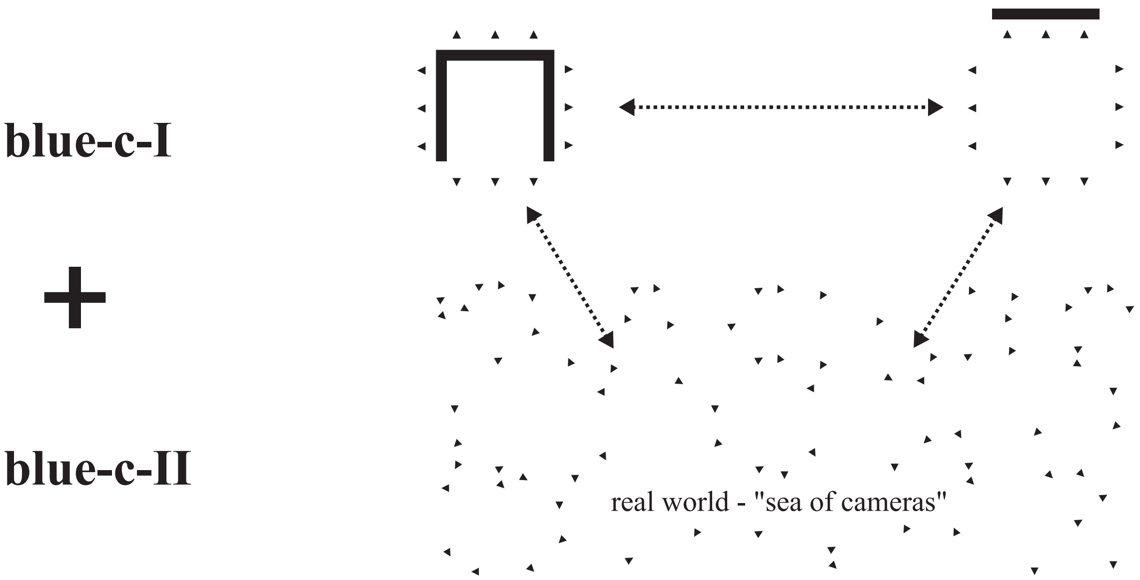


Figure 4.20: Open the real, physical world to the user.

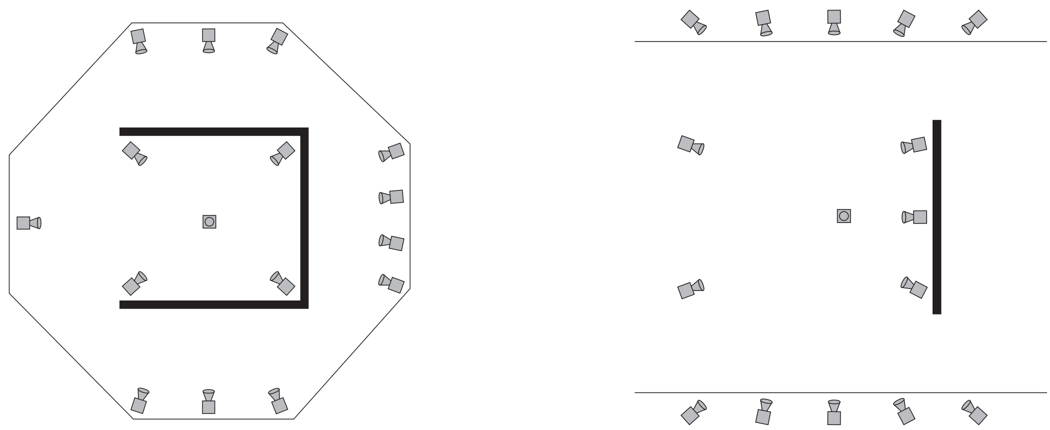
a large-scale environment. In order to cover such large environments, the number of cameras will have to be drastically increased. Furthermore, a scalable and highly modular acquisition setup will have to be developed which has the technical ability to consist of different types of cameras. Figure 4.20 illustrates how merging the results of blue-c-I and blue-c-II opens the real, physical world to the user.

4.6.3 3D Video Technologies

Figure 4.21: 3D video reconstruction of users.

In traditional video, the viewpoint in a scene is the one chosen by the director. The real world is represented as a 2D impression where the camera positions are defined. Therefore, the viewer who is watching the video is not able to change the viewpoint. The *Aspen Movie Map* developed by the Architecture Machine Group (now part of the Media Lab) at the Massachusetts Institute of Technology (MIT) in 1978 can be considered as precursor of free-viewpoint video. Due to videodisc technology it was possible to allow for non-sequential access. Recording of the video footage was done using four cameras, each pointing in a different direction, and mounted on the back of a truck. After recording, the pictures were linked together. Therefore, the user could start ar a particular point in the movie and move forward, back, left or right, and interactively ride through the city of Aspen.

Recent advantages in powerful, low-cost computers and video cameras enabled the prospect of creating 3D video. 3D video refers to stereoscopic motion-picture imaging. 3D video extends conventional video to the third space dimension. It is generated by populating the physical space with multiple cameras. Dynamic visual events in the real world are recorded as they are. The object is recorded in full 3D shape, motion, and precise surface properties. In order to detect motion and precise surface properties in generating 3D videos the physical space is populated with multiple cameras.



(a) CAVE^{TM}-like portal at ETH Center (b) PowerWall-like blue-c portal at ETH Hönggerberg

Figure 4.22: Arrangement of cameras and projection walls of the two blue-c portals.

Both blue-c portals can be used for 3D video acquisition. The portals are equipped with multiple video cameras (see Figure 4.22) to capture and stream video images of the portal volume. In particular, the 3D video system is composed of 16 camera nodes which acquire images and perform 2D image processing. The resulting 3D video inlays are transmitted to the remote site. Camera and reconstruction nodes are at the same physical location and are connected in a local area network. The 3D information of the user is extracted by employing a shape-from silhouettes technique which calculates a visual hull. The 3D video data is streamed to a rendering node which, in a real-world tele-presence application, runs at a remote location. More in-depth explanations can be found in Stephan Würmlin's dissertation [Wue04].

Free Viewpoint Video

In recent years, the research on free-viewpoint video [Wue04, Car03, Mat02] increased. The goal of free-viewpoint video, as a main category of 3D video, is to create a feeling of immersion by giving the viewer the possibility to choose his own viewpoint interactively. From multiple camera views video sequences are generated. The camera settings are ranking from parallel view, over convergent view, to divergent view. In general any camera setting is possible. From these views, a 3D object is generated that can be described by, e.g., polygon meshes, point samples, implicit surfaces, depth images, or layered depth images. Free-viewpoint video can be computed by extracting geometry and texture information from a set of concentric views of the same object. The viewer has the opportunity to interactively change the viewpoint in the scene.

Within the blue-c environment 3D video objects (see Figure 4.21) are represented as a 3D point cloud [Wue03]. Each point sample carries attributes such as position, color, and surface normal. No connectivity between the point samples is necessary which allows for efficient and low latency transmission of 3D video objects. The surface normals allow for re-shading the 3D object according to the lighting conditions of the virtual world. Based on a shape from silhouette method, a three dimensional video fragment representation is calculated, compressed, and streamed to the other sites. These video fragments are integrated into the virtual world using hardware-accelerated point rendering. Only objects in the portals are extracted and transmitted, the portal itself is not represented. Since the approach used for extracting geometric information is not model-based also arbitrary objects can be introduced into the system besides the user. Basically, everything that gets into a portal is three-dimensionally represented. To allow this separation of foreground objects from background a learning phase is necessary. This usually has to be done once per session and takes about 10 seconds.

The performance of the 3D video engines used within blue-c is sufficient for processing objects with less than 30k point samples. In a typical 3D video sequence, processed at 10 frames per second, between 15k and 25k points in the 3D video objects per site are maintained. The blue-c real-time 3D video system has a system inherent latency of 3 frames. In the worst case, the round trip latency sums up to 5 frames. In experiments between the two portals at ETH Zurich the minimum latency is 200 ms and the average latency is about 300 ms. Latencies up to 400 ms were experienced in experiments with an inter-continental connection between ETH Zurich and UC Davis.

5 Contributions and Benefits to Architecture

Normally in time, technologies in architecture are used for being productive, for solving problems, for affording temporal, economic, and manufacturing efficiencies. However, technologies may also be used for creating enhanced environments. Computer technologies are no longer just design and fabrication tools for architects supporting their design tasks. With the help of information and communication technologies it will soon be possible to equip and net the main components of a building with computer performance. Thus, our buildings will have a distributed network of devices that provide the users with information, communication, and entertainment services. These systems adapt themselves to the users and anticipate on their demands. The technologies offer new ways to reduce spatial and temporal interdependencies between different activities. Therefore, they provide new ways for architects in their design task. In fact, treating them like design materials integrated into the building design process, they provide architects with the opportunity to generate interactive and dynamic spaces. Consequently, the role of architecture can be increased by using computers to create new forms of reality.

The advent of information and communication technologies such as video systems have caused a significant shift. They not only have a bearing on the perception of architecture, but also in the way how architecture is conceived and perceived. Like other inventions, for example the radio or the telephone, people succeeded in finding new applications and handlings which have been primary neither intended nor predictable. Architecture belongs to the highly interesting application areas.

This chapter analyzes and investigates the possibilities of merging architecture and video systems. The focus is on video as novel media for designing environments that provide and require novel forms of augmentation, communication between people, and interacting with information. What kind of capabilities information technologies have been able to affect the use of space and place, and the way people perceive and think about their surroundings?

5.1 Challenges

"You really can't separate issues of technology and the space that accommodates it. At this point- you have to think of the two of them together."
William Mitchell

The architectural writer Reyner Banham already took stock of the impact of tradition and technology on architecture in his essays. He was fascinated by the miniaturization and by the reduction of advanced technology which has been styled and then brought closer to the public [Ban81]. On the one hand, he saw architecture as expression and on the other hand, he saw architecture as technology. According to his analyzes, there are two forces affecting contemporary architecture. First, the conservative pressure of tradition. Second, technology which encourages a more progressive, open-ended approach to architectural problems [Ban84]. Banham redefined tradition as the stock of general knowledge and professional history which architects assume as the basis of present profession and future progress. Whereas technology represents the converse of tradition. Technology describes the method of exploring using the instrument of science [Ban80].

One of the great triumphs of urban planning was the separation of noisy, polluted industrial zones from leafy, garden suburban. Earlier telegraph and telephone networks, later the linkage by computer transportation networks enabled spatial separation of management of industrial production. The business workplaces of the 20^{th} century are characterized by high-rise offices made out of steel, concrete, glass, and electrically powered systems providing the means. At the end of the 20^{th} century architecture began to change fundamentally. The solvent of digital information was decomposing traditional building types and familiar forms vanished [Mit97]. According to William Mitchell the occasions for a characteristic new architecture of the 21^{st} century are the intersection of electronic information flows, mobile bodies, and physical places. Invisible, tangible, electromagnetically encoded information establishes new types of relationships among physical events occurring in physical places [Mit03].

5.1.1 Design of Technology Enhanced Spaces

Looking back at the history of architecture, it can be observed that architects since ever, combine different and the latest technological developments in their building design. The interest of the re-use of technology and methods from other industries is not new. Architects have always looked beyond the boundaries, appropriating materials, methods, and processes from other disciplines. Therefore, architects are currently, in their profession and education, challenged by combining applied information technologies in their designs. Due to their novelty and the focus on their development, most of modern information technologies lack of useful and inspiring application possibilities. Modern information technologies provide several opportunities for simplifying and streamlining everyday life. Normally, novel technical developments are not usable for the average people. Bringing computation into the real, physical world and embed them into the built environment brings out a new task in creating spaces that are designed around human beings rather than technological needs. In a world where technical devices abound, the key quality criterion is the easy access for the user. In these interactive, computational spaces the components have to cooperate and support the user. Computation has to be embedded in ordinary environments. These new spaces in which computation is seamless used to enhance ordinary activities offer the possibility of multimodal interactions in a natural and familiar way.

From the historical point of view buildings corresponded and were distinguished by their uses. The design represented social division and structure, and made them visible. Architecture had a profound impact on the traditional city. The buildings and their facades reflected organizations and social groupings. Due to the fact that information technologies increasingly influence the physical world the character of the built space also changes. Since information technologies have radically different spatial implications than traditional arrangements architects are challenged to suit the role of architecture to the altered living conditions (distinctions). Today, architects have to address telecommunications as they do with circulations. One crucial task will be to translate digital information into visual and auditory interfaces and make them part of buildings.

Humans will always have their bodies and possessions. Therefore, the need for built spaces will not disappear. Humans still will require shelter and rely on tangible goods and services. Architects of the 21^{st} century still shape, arrange, and connect places to satisfy human needs [Mit97]. Architects are challenged more than ever in designing technology experiences instead of designing an experience enabled by technology. To make broad use of modern information and communication technologies in order to overcome boundaries between space and time, and to improve the quality of life is an ambitious and fascinating challenge for architects.

5.1.2 Privacy and Security

The biggest moral issue concerning embedded information technologies is that of privacy. The term privacy can bear different meanings. On the one hand, it refers to the desire to keep communication from being seen by anyone other than the intended recipients. On the other hand, it can be used be to ensure that individuals maintain the right to control what information is collected about them, in which way it is used, who uses it, and for which propose it is used. The first definition is related to cryptographic techniques in order to make exchanged data impossible to understand for an interceptor (see Section 3.2.3). The latter definition addresses the infrastructure that deals with information and communication. Privacy is a concept, and not a technology. Therefore, it can be provided in different ways depending on the available technology and the used application.

Since business and personal life depend more and more on computers digital security has been growing in importance. As Section 4.4.2 reveals more and more public spaces are covered by video cameras. Once, the natural condition of the city was opacity. Architects created limited transparency by means of door and window openings, enfilades, open rooms, and public spaces. Toady, it can rather be said that the default condition is electronic transparency. Therefore, architects have to work hard to produce limited zones of privacy today [Bri98]. One key challenge will be to guarantee secure services and to create spaces in such a way that security for humans will be improved without remarkably reducing their demand for privacy.

Spaces under surveillance - whether public or private - are a sensitive issue. On the one hand, one wants to enhance personal comfort (e.g. security or health and care), and on the other hand, people are in general reluctant being controlled. In an architectural context it is distinguished between public and private spaces. The surge of video cameras will annihilate anonymity in public spaces and negate privacy in private spaces. Therefore, it is necessary to think about returning an adequate control of the levels of surveillance to the person tracked by the system. This can be done, for example, using actual techniques of encryption (see Section 3.2.3) in combination with smart control devices.

The following assumptions seem to be important in realizing such systems. Persons in private spaces must have the right of their privacy. This will be realized by allowing the user to encrypt parts of his private space. Video cameras looking into encrypted areas will have to encrypt the portion of the video stream which is showing the encrypted area. Only authorized people will be able to see the whole space by decrypting the video stream.

This allows to give people back the control over their level of anonymity in public spaces. In general, the acquired data of a public space will not be encrypted. If there is a surveillance request on an individual person, the person to be tracked will get a signal on its smart device. This person will be informed of the tracking request and will have the possibility allow or disable the tracking request. In addition, to denying the tracking request, it will be possible for the person that video streams will encrypt the areas were it is possible to identify the person on [Hov03].

It is important that architecture deals with the ways of ‘controlling space’, such as the issues of encryption of spaces and the handling of authentication and authorization of people in private and public spaces. It should not be forgotten that surveillance systems are not always there to catch people doing something they should not be doing. In many cases they are located just in case in places where it is necessary and wanted to have an eye on incidents.

5.2 Impact

The spread of digital media has a bearing on architecture. It is obvious that the novel opportunities of digital media are not restricted to generate almost perfect, with photographic precision operating project visualization or to produce computer aided customized building components. Today, the computer is more than a drawing machine or a technical help at realization. Video is more than a simple recording machine.

5.2.1 Architecture and Information Technologies

Over the past decades architects dealt with the issues of technology in theory, design, and production. Thereby, their approaches varied as the technologies themselves. For functionalist design automobiles, ocean liners, and airplanes acted as metaphors. Le Corbusier wanted to solve the problems of the house using machine technologies, especially industrial mass production and modular fabrication techniques [Cor86]. Banham instead developed a series of pneumatic houses resulting from mechanical systems. Today, architects still explore technology in order to solve social, economic, and design problems. The technology is used in order to be productive, solving problems, affording temporal, economic, and manufacturing efficiencies.

Nevertheless, can the technology also be used for creating highly embedded, interactive spaces? The computer is not only an instrument to rationalize production, processes and work flows, but also a medium for mass communication. Technologies are no longer only design and fabrication tools for architects supporting their design task. In fact, treading them like design materials integrated in the building design they provide architects the opportunity to generate interactive, multi-media based, and communicative spaces. Interactive, because the flows are not only linear. Different kinds of media units, for example images, movies, audio, letter press, and virtual space, can be independent from each other accessed. Networking the spaces and media units makes the spaces communicative.

The focus in finding appropriate applications is centered on system integration, and the adoption and expansion of existing technologies in order to show what is feasible today. Traditional operating methods and novel technologies are merged and investigated.

5.2.2 Augmented Environments

"The most profound technologies are those that disappear. They weave themselves into the fabric of everyday life until they are indistinguishable from it."
Mark Weisser

Today, our environment, created by architects and modified by its users, is more and more about the use of information, its generation, and exchange. Recent innovations in information technology influence our architectural understanding of geographical spaces, urban forms, and immediate habitats. Actors like displays for presenting information and sensors like cameras for capturing information, and network connections are as integral to built spaces as walls, floors, windows, doors, and doorways.

Environments can be considered as intelligent when they are able to learn from the user's behavior, and able to adjust according to the user's preferences. Therefore intelligent environments are sensitive, adaptive, and responsive to the presence of people.

These intelligent environments are equipped with cameras, microphones, displays, sound output systems wireless network, detectors, and controls for physical entities such as curtains, lighting, door look, thermostats, TV, radio, etc.

Intelligent spaces are highly interactive environments. Architects are more and more asked to integrate the capabilities of these systems in their design. These technical inventions allow them to create responsive and intelligent environments. These spaces are able to react and adapt to the behavior of their users. They improve the quality of life of people. Intelligent, personalized inter-connected systems and services are used to create the desired functionality.

Once desktop computers, telephones, and wired networks have been fixed points of presence and therefore, were powerful attractors of human presence and activity. Ambient Intelligence (AmI) and Distributed Computing (see Section 3.1.3) create continuous fields of presence extending buildings, outdoors, public spaces as well as private spaces. By selectively loosing place-to-place requirements these technologies fundamentally alter, fragment, and recombine familiar building types, urban patterns, and space use. This has a profound impact on locations and spatial distributions of human activities depending on resource availability and information access.

Intelligent user interfaces allow users of these environments to control and interact with AmI environments without having to go through an expensive learning process and using the knowledge they already have acquired during their lifetime or their professional training, respectively. Information and content will be available to any user, at anytime, and anywhere.

In a next step, it is important to take peoples desires and behaviors in these environments into account. Further research has to investigate how the environment communicates with the user and how the user communicates with the environment.

Modern technology is constantly becoming faster, smaller, and cheaper while memory and processing power is on the up (see Section 3.1.3). Today, most computers are not seen nor used as such by humans. Rather they are sophisticated, computerized, networked machines integrated in telephones, cars, microwave ovens, cash registers, and a multitude of other fanciful and mundane devices and systems. Loosing the person-to-place congruity requirements, these technologies have created an additional degree of spatial indeterminacy. There is no doubt that this effect will remarkably influence our physical, architectural environment.

Information technology will disappear into the background and become invisible as a pervasive medium, similar to electricity. Real objects and everyday setting come into foreground again. As a consequence, human-centered notions instead of computer-centric ones are important. When the computer disappears the environment will become the interface. This hybrid, a mix of analog and digital data processes, environments provide new ways of interaction with information and communicating. Objects become augmented with new properties and qualities. This leads to "people-friendly" environments which support and enhance people's live, and produce new behavior and functionality. There is a profound need for architects in designing these human-oriented environments.

5.2.3 Mobility and Miniaturization

A lot of research is done to gather physiological and geographical information of humans. The technological inventions are able to collect, store, process, and present data of humans. Environmental devices are connected to sensors, actuators, and appliances. In most cases, they are integrated into physical objects of small dimensions communicating with each other and handheld devices through a network.

Sensors are integrated into the walls of the buildings to track peoples' movements. The sensors are capable to gather information about the people. The information can be subdivided into three categories:

- vital
- inputs/ outputs
- movement

The reduction of computer size, weight, and prize has given the wearable market large potential. GPS, RFID tags, vision sensors, magnetic sensors, ultrasonic sensors, and other sensor devices are available at much lower prices than ever. Computer-based parts are very small. Therefore they can fit in the hand, be interwoven with clothing, be attached to the body, or even implemented in the body. There is an ongoing research on wearable and integrated body gadgets [Eve00]. The electronic components integrated into everyday's clothing should be designed in a functional, unobtrusive, robust, small, and inexpensive way [Dit00]. These transponders are constantly transmitting their locations to the network for precise, close-range tracking. Also biological data can be tracked. Devices such as card readers, RFID tags, people counters, movement sensors, and web cameras can track people and the physical environment. These devices can be used to design perceptual interfaces which can track the real world. Today, computers are integrated into 'shells' where usually the association with a computer is no more visible.

5.2.4 Outlook

The present state-of-the-art and the recognizable running developments of information and communication technologies give rise to a promising view into future human adequate scenic interactions which will bridge spatial and temporal distances. However, some fundamental problems have to be solved prior to a widespread economic propagation. Technical equipments (such as processors, memories, transmission lines, cameras, displays, sensors, and actors) and methods applied (such as image coding, processing and interpretation, data stream networking and management) must carefully be adapted to each other. The overall system integration and optimization is a real challenge to the co-operation between engineers and architects. It would be very creditable to learn more about the correlation which exists between the complexity of a scene, the viewing angles, respectively the number and types of cameras being necessary for recording the scene, and the management of the accessory data streams. Theatrical performances, football games, manufacturing processes, for example, have obviously a much higher scene complexity than empty cubic rooms with plain walls. A reliable benchmark which is suitable for characterizing scene complexity and correlating it unambiguously with the technical equipments and applied methods for scene recording is highly desirable. To find such a benchmark could be a future ambitious scientific as well as an economic advantageous task.

5.3 Advantages

The integration of video systems offer architects the possibility to make the future development of our society experienceable. Glass, for example, got an 'heroic' material for the demonstration of democratic awareness since Mies van der Rohe's transparent high-rise building. Consequential, video systems as well can be understood as intelligent building materials. However, video systems feature a special characteristic: they display their power not until the software is connected. This software is dynamic over time, a factor which has been so far in construction oriented architecture of limited importance. Furthermore, video systems as part of information and communication technologies are a contribution of the decentralization of both people and activities. They expand the architectural space and reduce spatial and temporal interdependencies among humans and their activities [Lan04b].

5.3.1 Enhanced Reality

Real environments are those in which real persons are able to move in real world. As opposed, immersive virtual environments separate and detach the user from the real world and makes him part of a computer generated, artificial world. Therefore, the user is disconnected from what is traditionally called reality.

However, the research has to go "beyond being there" and enhance reality. Figure 5.1 shows the evolution of information sharing using various communication services connecting to the virtual world. Until the latter half of the 1990s, information was received via dial-up connection which were established when necessary. Toady, bridges between the virtual and the real world have been built. The Internet provides the availability of permanent accessibility to get information necessary for daily life and business. In the future, the impact of broadband and high-speed connections allows to get any information anytime and anywhere. Another important feature of technological advancement is that of miniaturization (see Section 5.2.3). Powerful microprocessors can fit into very small areas. The virtual and the real world are merged together to an enhanced reality.

Enhanced Reality can be considered as an information-rich environment where information blend with the environments in various forms every place: home, office, public spaces.

The infrastructure for rich, limit-free communication in the enhanced reality will

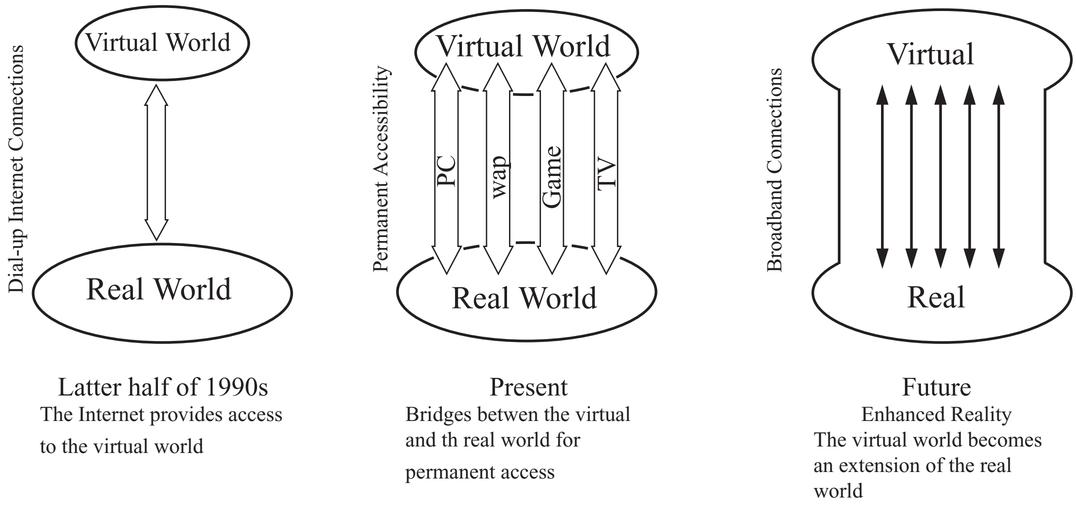


Figure 5.1: Information in the virtual world becomes part of the real world, so called Enhanced Reality.

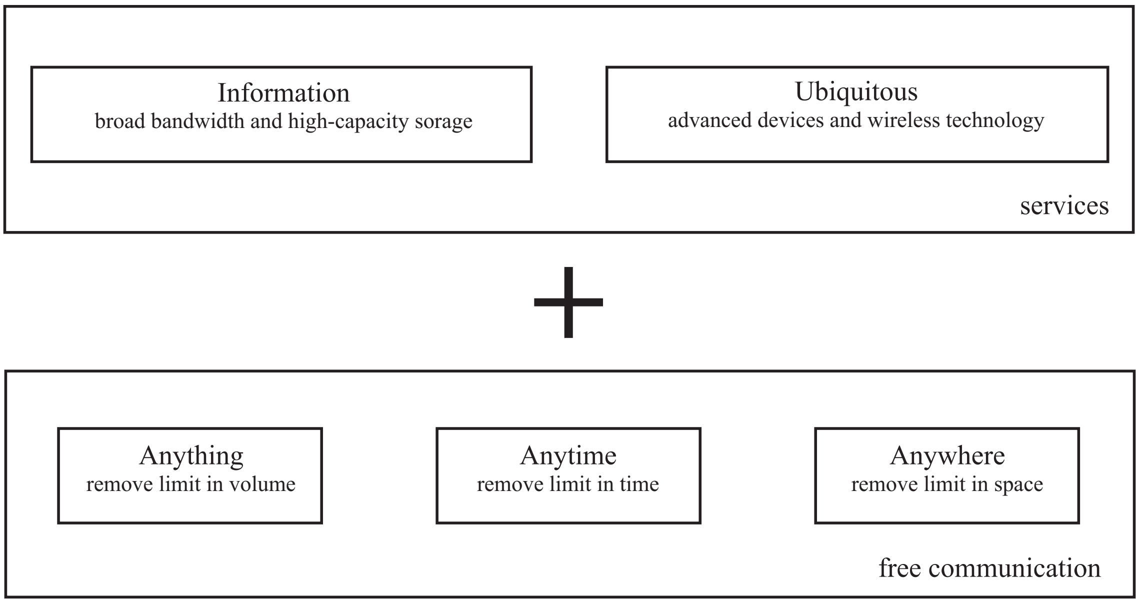


Figure 5.2: Infrastructure for rich, limit-free communication.

be like bringing two main streams together: services and free communication (see Figure 5.2). Information services provide and manage information to share among distributed users. They take advantage of increasing bandwidth and large-storage capabilities. Ubiquitous services means that not only persons but also various items of equipment can access information. These services take advantage of the diversity and omnipresence of networks. Therefore, people can receive information from anywhere and at anytime. The idea that environments will have some degree of intelligence or computer power, relies on miniaturization of equipments and reduction of transmission bandwidth. Miniaturization offers more flexibility in the design of these environments. The size is no longer a dominant factor influencing the design choice.

This means that the space is no longer built around the technology. New technologies provide much more flexibility in space use. The focus of designing spaces is, therefore, no longer creating spaces that meet specific technical requirements rather the challenge is to make spaces human-oriented. The focus is on expanding the availability of human resources.

5.3.2 Interactive Architecture

At present, the user is able to see 3D, but he can only limited interact in digital 3D worlds. The next step is to bridge the gap into virtual environments. The user will be able to alter scenes. Using the capabilities of today's technologies it is possible to create a specific virtual architecture that is only possible in virtuality. Objects, for example, can maintain a behavior over time or the moment of interaction. An object is able to obtain infinitely information.

In the book *Hamlet on the Holodeck* by J. H. Murray [Mur98] interactive, digital environments are described as follows:

> "Digital environments are procedural, participatory, spatial, and encyclopedic. The first two properties make up most of what we mean by the vaguely used word *interactive*; the remaining two properties help to make digital creations seem as explorable and extensive as the actual world, making up much of what we mean when we say that cyberspace is *immersive*." (p. 71)

The four parts of this description can be used to define interactive immersive environments. *Procedural* refers to smart environments where the system identifies the user and adjusts to his profile. Environments are procedural when they remember the users' preferences and other personal user settings. The common term describing this function is *personalized* environment. An environment is *participatory* when it gives users the possibility to actively control it or to communicate with from disparate locations. The *spatial* property deals with the navigation in digital environments. *Encyclopedic* refers to the possibility of data retrieval. That is to say that the user has full access to recorded and stored data.

The interaction between humans and virtual objects in 3D spaces will be an integral part of human communication. Interactive architecture arises new aspects to design rooms and space interfaces. It concerns a dynamic concept which revamps the way humans interact with the world as they do now.

5.3.3 Interaction and Navigation in Large Scaled Installations

On the one hand, building automation and building security make new ways of communication and tele-presence possible. On the other hand, they represent not only new ways for architects to create space, but also for programming how the space itself can and will react. The use of wearable computers and immersive communication will invoke a dynamic and interactive understanding of space, place, and time redefining our awareness of traditional architecture. Digital cameras will shrink in size and prize. Therefore, they will be usable in more situations and will be inevitable. They will also offer better quality and can often be applied wireless.

blue-c-II enhances the research from blue-c-I by introducing a sea of cameras in large-scaled public spaces (see Section 4.6.2). Therefore, whole moving scenes can be gathered. By recording video from multiple cameras in large scaled spaces and make a reconstruction of them, the observer can move everywhere in the 3D space and watch events from completely new 3D positions. Architectural applications are described as large, spatially differentiated and hybrid systems with a high degree on user interaction and a long service life of approximately 20 years. Hybrid data sources have to be integrated into a consistent navigation within large captured spaces. Therefore, it is necessary to provide wearable devices to users. Thus, essential and novel criteria, qualities, and quantities of video systems regarding scalability, interoperability, modularity, user interaction, and user acceptability are enabled.

Usually monitoring facilities and video surveillance technologies are used only for varied security, law enforcement, civilian environment monitoring, health care, wildlife preserve management, traffic measurement, and catastrophe response management. Combining different video types and integrating off-the-shelf cameras in large scaled spaces allows to transform the one-way monitored view of a space into a novel communication space. At the moment the monitored views are a collection of fragmented views from surveillance cameras. Typically, in large control centers, human operators interact with dozens of data sources using several separated monitors for visualization. The operator is easily overwhelmed with the task of understanding the relationship between these separated images and how the situation changes in the scenes. A consistent navigation enables a global view of the large-scaled space and provides an accurate and comprehensive picture of an entire scene.

3D graphics hardware is becoming inexpensive, and visual technologies are rapidly advancing. By optimizing obvious efficiency issues such as cost, time, and resources it is possible to design appropriate applications. The integration of hybrid data sources including 2D video, 3D video, wearable video, and high definition video allows to deal with a large amount of visual data generated by a sea of cameras by introducing a consistent navigation. This requires an interface to mange and configure the different video processing lines. The long tradition and experience in architecture to create and improve the human environment and architectural principles make architects capable to design human-computer interfaces. Architects are by nature trained to consider the needs and desires of those who use their designs. In order to provide modularity and scalability, and to bring the technology closer to the public it is important to develop applications that adapt to different end clients ranking from CAVE to PDA. Therefore, the user is able to roam large spaces, in real-time as well as in 2D or 3D. These applications provide a more flexible and consistent view to large-scaled spaces than traditionally monitoring systems. The dimension time is relatively new or even unknown to applied architecture. The combination of recorded, slow motion, time-lapsed, and real-time or live streamed video data enables radically different spatial implications. Thus, not only place-to-place connection, but also place-to-time relations are abolished.

5.3.4 Interaction via 3D Video

Interaction describes a mutual and reciprocal action. Mostly, interactivity is associated with novel information and communication technologies. Communication between people, or the actions of people that affect others can also be considered as interaction. Interactivity is, therefore, a communicating bridge between humans, and between humans and machines. 3D video makes a contribution to a more efficient, effective, and satisfying human-computer-interaction.

Spatial cues such as position and gaze are essential for a natural interaction in immersive spaces. Since 2D video just provides one viewpoint those cues are distorted. Whereas 3D video preserves spatial cues by providing multiple views.

It is distinguished between three different levels of interactivity [Jan02]. Participation refers to the user's interaction with the content. As opposed to traditional video the user can freely navigate through the video content. The user can participate by interacting with the 3D objects and influence the scene by choosing between different options. Communication specifies the amount of users participating in the application. This can be either immediacy in the broadest sense real-time or time delayed. Using 3D video within architectural settings enables synchronous as well as asynchronous communication. The degree of complexity varies from one-to-one, one-to-multiple to multiple-to-multiple participants. Creation refers to the results of the participation of the user, e.g. providing reactive content.

Until now, video in architecture is limited to behavior and presentation studies, as part of the design and execution of construction projects, and for collaboration within virtual design studios. The moving image can be considered as the engine of the imaginary, the pivot of desire, and the hub of transparency that characterizes our lived environments [Hog00]. 3D video represents a technical approach and affirmation of these powers. It offers a promising contribution to the emergence of a information technology culture of architectural enquiry [Lan04a]. 3D video has the ability to make someone feel being part of a scene instead of merely looking at it.

Interactive 3D Video

Interactive 3D video refers to the possibility that the user or the audience can control the outcome of the scenes they are experiencing. Due to the 3D video technologies described in Section 4.6.3 it is possible to create video scenes in such a way that the user can influence it. Therefore, video becomes truly interactive and immersive. By offering the user the possibility to access, recorded and stored live video content, and additional information interactive as well as choosing the path of navigation 3D video is encyclopaedic.

The most familiar navigation through video content are the functions characteristic for the video player. The most common interaction features known from video players are play back, pause, variable speed fast forward and reverse, and slow motion. The same functionalities are provided for interacting with 3D video content. Conventionally 2D video systems commonly are only capable of capturing temporal changes of scenes. As opposed thereto, the extension of video to the third dimension enables spatial variations and novel interaction possibilities [Lan04]. Since 3D video captures dynamics and motions of the object during recording, panning, zooming, rotating, and tilting are now possible. Moreover, the combination of traditional features and these novel possibilities allow for spatio-temporal effects such as arbitrary scaling and pause-and-rotate [Wue03]. Therefore, 3D video allows to realize digital effects used in the film industry (see Section 4.3.3) faster and at lesser cost. 3D video now enables user participation by interacting with representations of real world objects, and by controlling the point of view.

Besides navigating through 3D content it is also possible to directly interact with the 3D representations. Since the objects are a digital representation they can be modified and deformed. At the moment sine wave operations through objects, explosions of objects, dissolving and morphing of objects, and beaming of objects from one location to another are implemented [Wue04a]. Beaming describes the process where the user can change the location of the object to a different location in the virtual space. In addition, the user can freely place and change the orientation of 3D video objects in the virtual environment. The user is able to introduce arbitrary instances of objects into the virtual scene and combine them together. Thus, overlapping of objects is allowed. For example, 3D video makes it possible for a dance or ballet student to fly around is own 3D representation and see his back and side views. Furthermore, he can move this representation into his remote or prerecorded teacher's 3D object. Thus, he can directly check out his movements. This is a huge advantage for all applications where postures are important.

As opposed to traditional 3D object modelling using common CAD software 3D video allows the 3D representation of any real world object. The extensive modelling effort of already existing objects ceases to apply. Using 3D video, objects are no longer simplified or a out-of-date digital reproduction of it. This allows for much greater accuracy of the objects. In former times, architects modelled 3D worlds by creating simple plain objects. Today, they use digital images of the real object and texture map them on the digital model, in order to coming drawn near reality. The obvious next step will be working with any real world object, without modelling efforts. Therefore, 3D video technology is considered as a novel visual medium for architecture as well as for communication and human-computer interaction.

Bi-Directional Communication

3D video technology supports face-to-face dialogic interaction between multiple remote users that share a common visual environment. The use of a fully 3D remote user representation greatly enhances the sense of presence and allows for a natural interaction between the users. Natural describes the human-to-human communication based on contextual information, such as gestures, body language, and voice. People speak, gesture, touch, and sense in their interactions with other people. In a conversation between humans a lot of information is only exchanged implicitly. The behavior of the participants supplies valuable information that is often crucial for understanding the message. Another issue are the emotive aspects, as a major distinction between humans and artificial representations. Emotions are a crucial part of the human communication capabilities. They are part of every communication. Using 3D video technology the user can communicate in a natural and intuitive way, comparable to the traditional situation where all participants are physically present. The user's speech is transmitted similar to conventional audio and video-conferencing tools. Audio input is an intuitive medium for communication with other users. This kind of bi-directional communication creates a natural experience for the users in the form of daily human-to-human communication. They see and interact with each other at the same time (see Figure 5.3). The embedded cameras are for the eyes and the microphones for the ears.

Another aspect is that the user can freely navigate within the video content and see objects and remote persons from arbitrary viewing directions. Therefore, opposed to traditional 2D video systems the user do not have to tell the geographical distant users to turn around, they can just fly around each other. Traditional video-conferencing systems force the users to directly gaze at each other. 3D video allows the user to feel more comfortable within a remote conversation. Similar to the physical face-to-face dialog the participants can freely move during the dialog. The combination of these technologies allows for a much better sense of reality within the immersive experience.

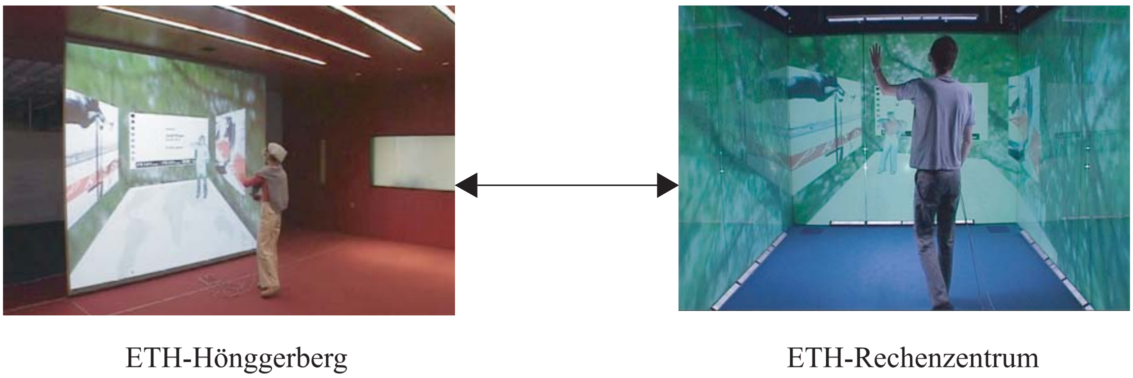


Figure 5.3: Bi-directional collaboration and communication in a natural and intuitive way between the user at ETH-Hönggerberg and ETH-Rechenzentrum.

Physical face-to-face situations remain still enormous important. 3D video is not seen as a substitute for traditional face-to-face communication, but rather an augmentation, a way of expanding communication. 3D video supports physical interaction between geographical disparate persons and virtual environments in the way that the interaction will be more like the way people interact in the real physical world. In fact, 3D video acts as a communication hinge in the absence of physical contact. The blue-c setup can be considered as contribution of increasing audiovisual capabilities. The architectural settings to support face-to-face interaction and communication are important. Thus, architects are asked to structure face-to-face communication in architectural settings and enrich them by overlaying with digital technology [Mit03].

Mad Eye Moody Effect

This effect refers to the Auror Alastor Moody, a the fictional character in Harry Potter and the Order of Phoenix [Row03]. Moody's main characteristic is his magical eye. This nickname "Mad Eye" results from his eye which is capable of seeing through most surfaces and which is free to move.

Unlike traditional video, free-viewpoint video (see Section 4.6.3) allows the user to interactively explore the video from arbitrary viewing directions (virtual camera). Free viewpoint video enables the creation of a video image that would be impossible to obtain by actual shooting with a conventional system. Therefore, the action of a dynamic real-world scene can be observed from any viewpoint. For a user it is now possible to look through the eyes of another person within the scene. This effect is similar to the capabilities of Moody's magical eye where the view can be removed from the physical eye height of the users. Therefore, 3D video offers novel viewing directions. 3D video provides a time-varying view independent from the user's perspective.

Video Hyperlinks

Video hyperlinks [Wic01] are the combination of 3D video sequences with additional content. The video sequences contain selectable objects attached with so called video hyperlinks. Hyperlinks are known from hypertext systems like the World Wide Web. Objects in an electronic document can be linked to another place or to each other. Video hyperlinks provide the user with annotations and additional information, for instance visual or acoustic presentations. The user selects the video object by clicking on it, activates it, and gets additional information. 3D video objects have a spatial and time characteristic. This offers an other interaction access to the user.

Interaction Metaphors

Multiple cameras are placed around the actor or scene and capture everything in a pre-defined area. These multiple 2D video streams can be used for vision-based object recognition. Detected hand and head positions can be used as standard tracking device sensors thus allowing novel interaction metaphors in order to reduce additional input devices (e.g. 3D mouse, keyboard) in immersive applications. An example application illustrating the potential of this technology has been realized. A user can point a virtual flashlight into the virtual scene. The software calculates a soft flashlight cone and only objects in this cone get illuminated. The rest of the scene is dark.

Vision-based object recognition offers a contribution in combining the physical world with virtuality by interacting with the physical space. The respective environment, therefore, consists of virtual objects which are projected onto the screens and real objects like the projection screens. Using collision detection algorithms a touch screen is implemented. For example, the user can press virtual buttons by physically touching the projection walls. Thereto, collisions between the user and the environment are detectable and the necessary collisions are detected by analyzing the 3D video data. This can be both performed on recorded video sequences and in real time.

Summary

Recapitulating, 3D video technology brings out a new kind of dynamic and flexibility by offering the following advantages for immersive applications:

3D video technology enables an overall impression of a scene, because of its independent view.

With 3D video a viewer can interactively choose the viewpoint in 3D space to observe the action of a dynamic real-world scene from arbitrary perspectives.

Substantial user's freedom of decision due to the possibility to freely navigate.

3D video provides the user the opportunity to watch a scene from completely new 3D positions, respectively from an other user's perspective. Therefore, he can move everywhere in the 3D space and change the way he watches the scene.

Unlike 2D video, 3D video facilities offer a spatial sense. Given that the camera moves in traditional 2D video, the users are observers. Whereas 3D video indicates the users to an active interaction and to swing.

Vision-based object recognition enables intuitive interaction mechanisms with the environment and the system.

5.4 blue-c Design Space

In order to classify the blue-c system (described in Section 4.6.1) into suitable categories for tele-presence systems, tele-collaboration systems, video system, and applications an adequate design space is introduced [Lan04]. The design space approach is similar to that used in Human Computer Interaction (HCI) since the 1980's [Nig93]. This four dimensional space is based on modality, eye contact, tele-immersion, and user realism. Before describing each of these attributes in more detail a short general remark on dimensions and systems is given.

Persons and objects in space and time are to be described by three space co-ordinates and one time co-ordinate, i.e. as 4D representation (see also Section 3.2.1). Motion implies time-dependent variations. In the case of a dancer, for instance, time dependent variations may refer to changes of his position in space as well as to his physical appearance.

Physical representations of information are called signals. It is spoken of 1D, 2D, 3D, and 4D signals depending on the number of co-ordinates the signal value depends on. A speech signal, for instance, is represented as a 1D time-signal, a still image pattern by a 2D space signal, a static object by a 3D space signal, a motion picture by a 3D signal described by two space co-ordinates and one time co-ordinate. In fact, 3D video in general implies 4D signal processing.

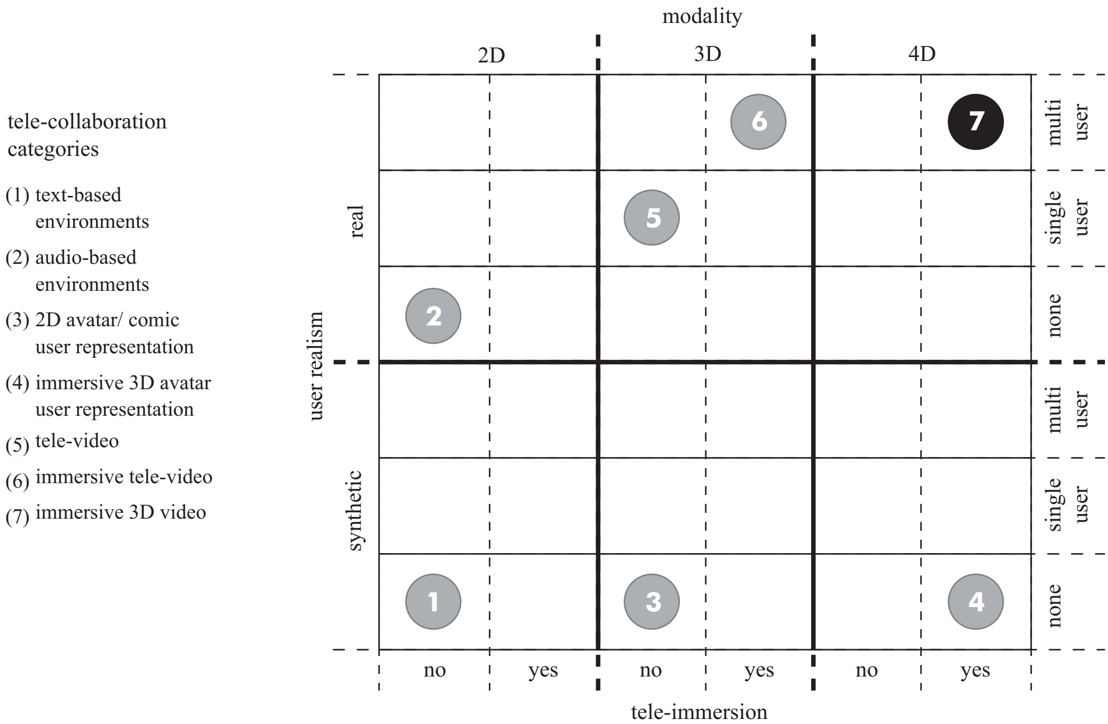


Figure 5.4: blue-c design space for tele-collaboration systems.

Modality

The term *modality* refers to the sensory and actuatory communication channels with which humans are able to interact with their environment, i.e. with other humans, with animals, with machines, with different kinds of objects. Multimodal systems represent and manipulate information from different human communication channels at multiple levels of abstraction. They can automatically extract the meaning from multimodal, raw input data, and controversially they produce perceivable information from symbolic abstract representations.

In the blue-c design space, illustrated in Figure 5.4, modality describes the appearance of the user. Space and time dependencies are explicitly pointed out in order to distinguish between motion and no-motion situations. The following three classes, written in italic letters, are defined:

- *2D modality*: the appearance has 1D space-characteristic and motion, e.g. online text or audio; or 2D space-characteristic without motion, e.g., still image, static 2D avatars (icons, comics).
- *3D modality*: the appearance has 2D space-characteristic and motion, e.g. video, animated 2D avatars; or 3D space-characteristic without motion, e.g., static 3D avatars.
- *4D modality*: the appearance has 3D space-characteristic and motion; e.g. 3D video, animated 3D avatars.

Eye Contact

This attribute indicates whether a system provides face-to-face communication between one or more participants. Eye contact is experienced when two people look into each others eyes. In visual communication they also transfer information by doing so. Figure 5.4 distinguishes between *none, single user, and multi user* eye contact.

Tele-Immersion

Tele-immersion refers to how a system can make users feel like they are experiencing an alternate reality and not just merely observing it. Moreover, this technique enables users at different physical locations to meet and collaborate in a simulated, shared environment in real-time. The degree of immersion depends on the one hand, on the feeling of presence, of being there, surrounded by the space, and capable of interacting with all available objects. On the other hand, it depends on the perceptional reaction whether the communication between the users is experienced like a real face-to-face communication. Figure 5.4 points out whether this is the case or not.

User Realism

In tele-presence systems the users are represented with different kinds of media. User realism describes whether a user has a synthetic, respectively a symbolic, or a real representation in the virtual world. There are two major approaches: synthetic and real.

One approach is the identification of handles, aliases, or nicknames in text-based virtual communities. Avatars, also called characters, players, virtual actors, icons, or virtual humans are commonly used in virtual environments. These representations are called *synthetic*.

Another approach is photorealistic representation of the users such as 2D video images or 3D reconstruction of real persons and objects employing vision algorithms. This describes *real* representations.

Tele-Collaboration Categories

A subset of interaction and communication techniques that let people communicate over distant is briefly described. Numbers in brackets refer to the tele-collaboration categories listed in Figure 5.4.

In text-based environments (1), such as chat rooms or multi-user domains (MUDs), the user representations have a *2D modality*. Audio-based environments (2), e.g. telephone conferences, have a real user representation. The appearance of the user has a 1D characteristic and motion. Both systems do not allow for eye contact between the users. 2D avatar or comic representations (3) as used in The Place [TP] have a *3D modality*, e.g. the avatars can change their expressions. MASSIVE [Gre95] and ACTIVEWORLDS [AWs] for example are immersive teleconferencing systems with 3D avatar user representations (4). These systems have synthetic user representations with a *4D modality*. Eye contact between the users is not possible. Videoconference systems (5), for instance GAZE-2 [Ver03] and desktop video conference (DVC) provide the most natural communication channels between remote sites using 2D video. Eye contact with multiple users is still a inherent problem. Immersive tele-video systems (6), such as the Office of the Future [Ras98], the National Tele-Immersion Initiative [Sad01] and VIRTUE [Kau02] provide real representations of the users with a *3D modality*. A significant difference to immersive 3D video systems (7) is that they only allow for single user eye contact. The blue-c system provides an immersive 3D video representation of the users. The actual visual appearance of the users is rendered in the immersive space of the remote users. Therefore, the users have a real, *4D modality* representation. As opposed to other systems blue-c provides arbitrary viewpoints and eye contact between multiple users.

5.5 Design Principles

This section describes criteria for creating embedded, interactive, and dynamic spaces that contribute to building intelligence. The motivation for designing and implementing real-world application scenarios is to bring computation, especially video systems, as part of information and communication technologies into the physical, architectural environment. This emphasizes the importance of bringing the technology closer to the user. The goal of these applications is to support what is traditionally considered as non-computational activity among humans. This includes the investigation to design experiences for end-users based on the possibilities of emerging and enabling technologies. There is an influence of technology on design. Today, technology changes very fast and issues of usability and accessibility accelerate this process. In the following, some design principles for designing environmental experiences enabled by technology are proposed.

The first step for an application developer is to identify appropriate scenarios where the novel technology can be used in an innovative way. The major factors that determine the choice of applications and the development of scenarios are the interaction possibilities for the end-user, content enhancement, and usage.

Application development starts from the user point of view. Therefore, it is important to observe people, how they use technology, and do things. It is important to take behavioral concepts into account. The concept of interactivity specified in Section 5.3.4 identifies different ways in designing interactive applications and to control different levels of interactivity.

Tim Brown cites two classes of experience design, “top down” and “bottom up” [Bro04]. Normally, designers use the top down method and conceive the whole design. Everything is controlled and scripted by the designer. Examples of this method are practiced by the companies iTunes, Disney World, and Prada’s New York Store. But in most situations the designer is not able to control all aspects. Whereas the bottom up method emerges and evolves out the actions of the users. The companies eBay, NTT (Nippon Telegraph and Telephone Corporation), and DoCoMo are engineered using the bottom up experience.

The first goal of designing applications using novel technologies is to get people to understand what possibilities this technology can provide and to figure out a way how it can be integrated into their lives. In order to get user interested interactivity and the ease of use are at the heart of these applications. The concept of interactivity has to be analyzed and decomposed from the perspective of the end-user, content, and application. The number of options and choices for the user have to be kept at such a level that they can handle them. Different kinds of levels include: personalization (setting profile), composing own viewing experience (navigation), and influencing scene composition (participation). To keep the user attracted and to motivate him to use the application more than once it is important to create a non-linear experience and let the user influence the story line. This can be achieved by introducing a virtual camera. This allows the user to skip parts, to choose from available scenes, to make changes in the scene composition, and to go back in time storage. This is very important for the design of interesting and attractive applications for users. Therefore, it is commendable to enhance the applications with on demand data, personalized content, and fictional content.

Good application design requires joints between different disciplines. At this point in time architects and technology engineers must work together from the very beginning of the process to form an interdisciplinary team.

In order to illustrate the realization of these design principles a prototype application is employed to proof the relevance of the technologies described before for architecture. The focus is to find attractive applications based on the integration of technologies by taking a consumer point of view. The concept and implementation of this application is described in the following chapter.

6 IN:SHOP

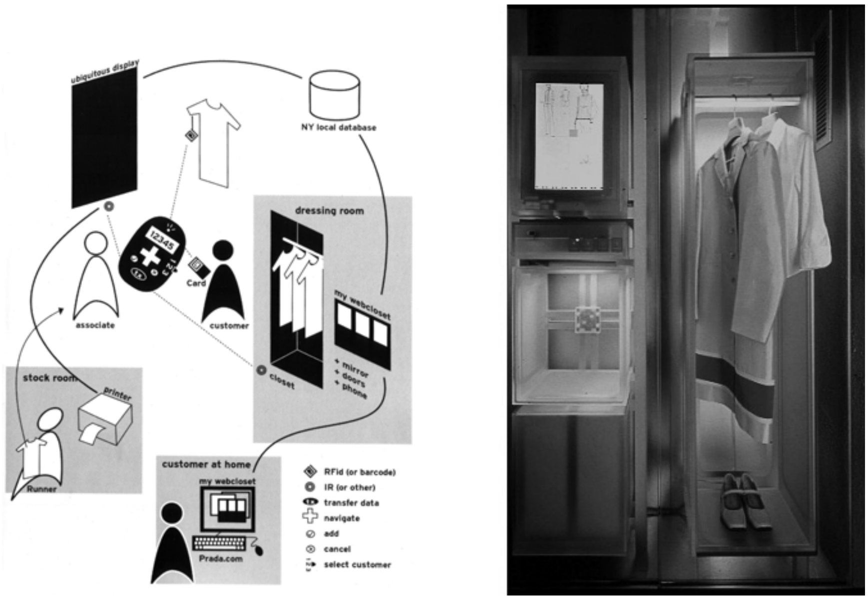
The term *IN:SHOP* merges the concept of a virtual shop and a physical shopping location. In effect, the word IN:SHOP denotes a shop-in-the-shop concept. IN:SHOP illustrates an approach to enhance reality in shopping areas and connects geographically distant persons. Therefore, it is an enhancement of the traditional, physical shopping, and the shop itself. IN:SHOP goes beyond the development of an isolated application to showcase video systems such as blue-c, but aims to demonstrate their use in real-world scenarios.

This chapter describes in detail the concept and implementation of the application IN:SHOP. IN:SHOP combines traditional shopping and marketing structures with 3D computer graphics, tele-presence, spatially immersive displays, and internet shopping paradigms to create a flexible and adaptable commercial environment. This scenario underlines the practicality of the developments made, and serves as a test-bench for their evaluation.

6.1 Preface

Digital technology accompanies shopping at Prada™ stores in New York as well as later in Los Angeles, San Francisco, and Tokyo [Koo01]. Prada's epicenter in New York is fully wired. The new Prada store is equipped with technical innovations from the IT sector (see Figure 6.1). Rem Koolhaas and his architecture and research company OMA/AMO integrated innovative display technology, intelligent mirrors, and interactive dressing rooms into the shop. Within a research project, AMO developed the virtual part of the shop. A shoe section that converts to a theater for performances and other "non-shopping events", an electronic customer-identification and service system that tracks shoppers and their needs, and smart dressing rooms are provided. The dressing rooms feature simultaneous, digitally-produced front, back, and side-views of the customer, phones for requesting assistance from the sales personnel or for communicating with people outside, and walls that can be switched from transparent - so the customer can model for his friends - to opaque for privacy. In the near future, the customers' profiles will also provide the input data for customized web pages to support shopping at home.

This IT concept for the in-store technology developed by AMO for the Prada



(a) Service system (b) Interactive dressing room

Figure 6.1: Prada store IT concept in New York.

Store in New York was the center motivation for developing IN:SHOP. In fact, the active glass panels used for the dressing room doors use the same technology as the projection screens used for the CAVE^{TM}-like portal at ETH Rechenzentrum. The video-based Magic Mirror concept already allows users to see themselves from different angels. Ad hoc developed RFID tags give users the possibility to directly access information that relates to a particular garment. 75 staff devices, designed, engineered, and manufactured by IDEO were introduced to the New York Store. This wireless staff device serves as communication and interaction device for the sales associates. It provides information to the sales associates and functions as a remote control for the ubiquitous displays scattered within the store. The combination of these technologies enable an in-store sales experience choreography.

By analyzing the use of this technologies and the integration of different media within this IT concept, it became clear how blue-c technology can contribute in designing interactive spaces. The idea and concept of IN:SHOP was developed by focusing on the added value of integrating 3D video systems. In finding the right concept it was important to see the technology as a phenomenon and to consider how it can be applied in creating something novel. The key thought for finding the right application scenario was that something novel can only be found by combining existing and emerging technologies.

6.2 Related Work

Internet shopping as well as tele-immersion technology is an active field of research. The IN:SHOP concept aims at integrating the best of both worlds to create a new interactive shopping experience which bridges the gap between purely virtual and real shopping spaces. This section list the investigations in architecture that have a significant influence to the IN:SHOP concept. On the one hand, an insight of the most important tele-presence systems and internet shopping systems is given. On the other hand, the influence of branding in architecture is shown by parenting the most relevant projects.

6.2.1 Collaborative Virtual Environments

Immersive collaboration will invoke a dynamic and interactive understanding of space and place, redefining our understanding of traditional architecture. The emerging technology of tele-immersion will change the way we communicate and will have an impact on our everyday life. Tele-immersion is an enhanced form of communication [Sad01]. At present, the user perceives 3D, but is restricted in the actual interaction with virtual 3D worlds. The next step is to bridge this gap into virtual environments. Several systems that support Collaborative Virtual Environments (CVE) have been developed, such as NPSNET [Mac94], RING [Fun95], DIVE [Fre98] and Avango [Tra99]. These systems focus on large-scale virtual environments with synchronization happening at the application level as opposed to the geometric representation. CVEs are a computer-based, distributed, virtual space or set of places. In such places, people can meet and interact with others, with agents or with virtual objects. They vary in their representational richness from 3D graphical spaces, 2.5D and 2D environments, to text-based environments. MASSIVE [Gre95], Teleport [Lal98], the Office of the future [Ras98], and recently the National Tele-Immersion initiative are systems using tele-presence. A detailed overview of these and similar systems can be found in [Chu01]. These systems use either simple texture mappings to billboards or provide simple 3D vision using stereo cameras. As opposed to these approaches blue-c as a novel video system (see Section 4.6) provides a full real-time 3D acquisition of the user that allows the others to freely navigate around the user. This is unique for the blue-c system. The main difference to other similar systems is that real persons and objects are represented instead of virtual humans (avatars).

6.2.2 Internet Shopping Systems

Using virtual reality techniques to support distributed collaboration of people has been explored in different applications. The idea to support designers, buyers, and manufacturers in the design of a garment to reduce the production cost has been researched in Fashion Pilot [Gra98]. Virtual Try-On [Div04] is motivated by the growing sales-volume of online shopping and the further development in multimedia. Avatar technology is used to enable online shopping. A three dimensional animated electronic representation of the customer wears the product before the customer buys it. As opposed to IN:SHOP, Virtual Try-On focuses on VR technologies for simulating and visualizing of garments put on by virtual counterparts of customer avatars. British Telecom (BTexact) also investigates the use of avatars for fashion applications [Dav01]. These approaches are complementary to IN:SHOP as their primary target is desktop digital product presentation and customization for internet shopping, whereas IN:SHOP concentrates on the social communication channels, collaborative and distributed sales process, and the use of highly immersive display devices.

6.2.3 Architecture and Branding

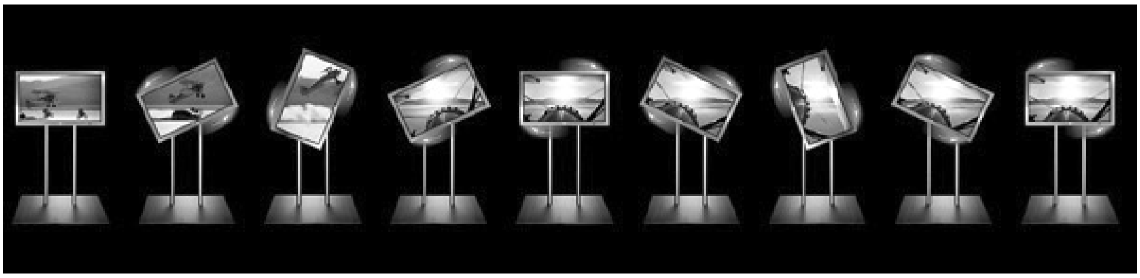


Figure 6.2: Bogner B Vision - a rotating large-sized screen.

Branding and Corporate Design are getting more and more popular in the field of architecture. Originally, branding comes from the Wild West. In order to make the possession visible, the cattle were provided with brand mark. Nowadays, branding is understood to provide products with additional immaterial values which characterize the manufacturer. Each medium - whether analogue or digital, whether changeable or static - can be used as carriers of mark contents. Commercially a pronounced identification of the customer with the product is promoted. In former times, advertisement and customer information was centered on the presentation of goods and their qualities. Today, modern shopping worlds produce shopping experiences which address the customer personally. In Nike's flagship store *Niketown* the center of attention is sporty activities and their cultural social environment. Prada combines reality and product with culture and commerce in its *epicenter* in New York. Bogner developed for their stores a rotating large-sized screen where the movements of the screens are synchronized with the image content. *Bogner B Vision* (Figure 6.2) is an audio-visual communication system. On April 1^{st} 2004, a new shoe shop, *Limited Experience* [LE], opened in Amsterdam. The Italian shoe brand Roberto Botticelli developed a novel, interactive concept for its sport collection. In this futuristic store there are no shoes displayed. The customer composes his own shoe using one of the five plasma screens installed within the store. He can make a number of choices: color, model, and type of sole. The system displays the shoe from the collection which fits best the choices the customer made. All models of the collection are in stock. Figure 6.3 shows novel shopping concepts in the fashion industry as well as in the automotive industry.

(a) Limited Experience, Amsterdam (b) Transparent Factory, Dresden

Figure 6.3: Interactive shopping experiences.

In the automotive industry a car's purchase becomes increasingly a shopping experience. Volkswagen, for instance, established the *Transparent Factory* in Dresden to distribute its latest luxury designs. The customer not only selects and orders his car, but he also follows its production in real time. In BMW's new distribution center, as another example, customers can experience the brand by visual impressions. An individual preferential treatment as well as a world -wide openness are most important. ART+COM [ART] designed the "Customer Interaction System" (2003) for Mercedes Benz. The system serves as additional information for buyers and sales aid for sales associates. It is used in Mercedes Benz outlets. In 1999, ART+COM already developed the "Virtual Vehicle", an interactive presentation system, for Daimler Chrysler. Virtual Vehicle allows customers to view and configure the entire model range of a Mercedes exposed in a showroom or at a trade fair.

Companies create absorbing experiences that engage corporate customers as well as consumers in an inherently personal way [Gil99]. How can architects contribute to the design of such shopping experiences? They are challenged to not only design static spaces, but also to design accompanying worlds of dynamic arrangements. The role of architecture is to make brands visually present. Clients want even more than implementing the latest technology into a building. They want a complete architectural setting, a building which vividly reflects the latest technology.

6.3 Concept

Multimedia elements such as large projection screens, video walls, and media installations, as well as information technologies such as networked information terminals, represent state of the art in designing modern public buildings. As the enabling technologies improve and decrease in price, these design elements achieve greater use. At the moment, this displays do not provide interactivity, merely playing uninterrupted video streams without allowing user interaction. IN:SHOP takes this trend one step further by combining traditional shopping and marketing structures with internet paradigms, 3D video technology, and tele-presence. An adaptable commercial environment which extends and connects the physical shopping floor into virtual and remote spaces is created. The application is not intended to supplant physical shopping or the shop itself, but it redefines the experience and architecture of commercial spaces. IN:SHOP enhances the shopping experience as well as the customer support. Utilizing blue-c technology an additional integrated virtual shop infrastructure is provided, which facilitates the presentation, assortment, and trade of personalized goods at the point-of-sale. Instead of replacing the current shopping experience with online shopping, IN:SHOP aims at enhancing the customer support and decision making through extending customer services.

Figure 6.4 shows the distributed application concept of IN:SHOP. Connecting several sites allows remotely located users to meet, communicate, and collaborate in the virtual shopping space. Already existing display technologies and glass walls with switch-able transparencies form the portals into the computer generated shopping world. These portals make distributed shopping in a real space possible and allow connecting different shops and locations. Geographically disparate persons can meet, communicate, and collaborate in real-time in a shared virtual environment. IN:SHOP allows pooling a worldwide network of branch offices and subsidiaries in a single virtual space. This concept also provides a strong sense of corporate identity [Lan03c].

The shop-in-the-shop concept, for instance, connects geographically separated customers, sales associates, and experts. They communicate with each other and interact with 3D representations of real objects in real-time. By using physical data acquired within the portal, data gathered from previous sessions, and from the customers database profile, the customer has direct access to his individual shopping environment. The system identifies the customer and loads his personal profile. This profile includes personal data, bought items, interests, demands, and so on. The virtual environment is customized based on recorded customers shop-

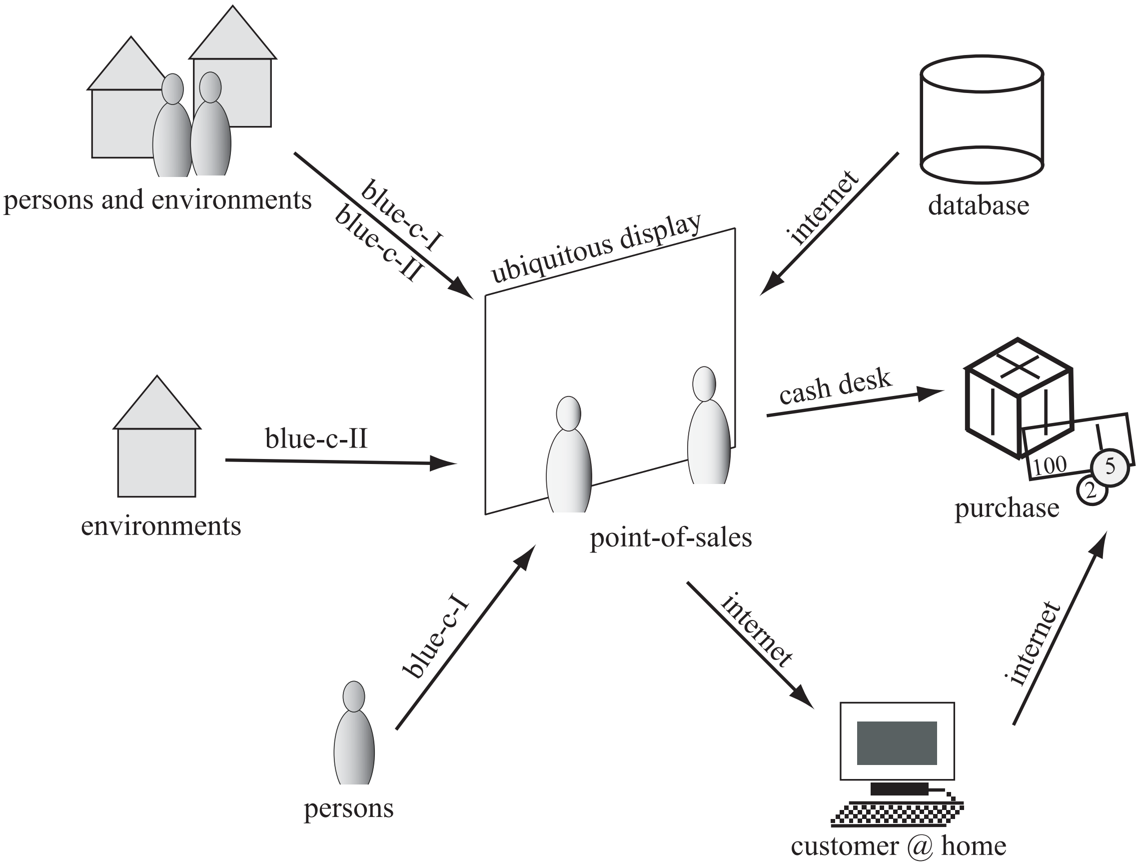


Figure 6.4: Distributed application concept. Using a shared database, tele-presence, and networking technology a unified shopping space among distant locations is created.

ping habits. Therefore, the presentation of information is personalized within the environment and is used to inform the customer and the sales associate as well as to support their communication. Also his personal sales associate from his patronized store is virtually included. The customer and the sales associate browse together through the personalized catalogue of requirements and offers. If the customer has indicated interest in certain items, the sales associate selects and presents them using a variety of traditional and virtual reality based media. Different forms of data can be presented in the immersive space, ranging from pre-recorded 3D representations of humans and objects, 2D movies and images, to hand drawn graphics, and text. During the shopping process the customer sees the object gradually becoming more detailed and interactively approaching the finally agreed purchase.

This concept allows presenting a large, multifaceted product range. The real shopping floor persists clearly arranged, the shelves are not glutted, and the storage space is slashed. Undoubtedly, it is also interesting from an economical point of view to create a unified shopping environment among distant branches using tele-presence and networked systems. For example, travelling customers are able to contact their familiar sales associate from miles around. Also experts of different areas of expertise can be hooked up linked to an item of interest. Technological and cost expenditures, of course, mainly aim at specialized expensive and custom-made products where additional information about the customer, the product, and its use by the customer is important. Other examples than fashion and cars might be jewellery, furniture, musical instruments, and yachts.

The appearance of a fully realized 3D representation of a person allows for a great sense of presence. This makes to a large extent, a natural interaction possible. The sales associate has direct access to databases including detailed product information and their visual and text based descriptions. He retrieves, visualizes, and modifies this data according to the customers' demands. In the sales conversation the new product is generated interactively and presented accordingly.

In the following the concept of IN:SHOP is described for a fashion shop and a car seller in more detail.

6.3.1 FashionShop

FashionShop is the first implementation of the IN:SHOP concept for the fashion industry. By browsing through an interactive catalog the customer is able to shortlist interesting pieces of clothing. To support this pre-selection the desired garments are presented in a fully immersive personal fashion show (see Figure 6.5). 3D video recordings of mannequins are used to present selected items. The customer can freely navigate in 3D around these 3D video objects, select them, and get additional information. Although, thereby, the customer gets a first impression of how the garments look like when they are worn, this preview does not provide enough information regarding sizing and combinations. Therefore, in the case the customer wants to have a detailed look at a favored combination of clothes she can get further information such as production, washing guidance, and fitting items by selecting this 3D video objects. Together with the sales associate the customer selects some base models for further customization. Customizable items are presented with additional information, such as matching items, availability, and number of sold items on the main information screen. The customer configures the product assisted by the sales associate. They discuss parameters using detailed design sketches. The customer decides on style, accessories, fabric, color, etc. An overview of the configuration decisions is shown as two dimensional detail images. Previous customization can be changed at any time. Before agreeing to buy a product the customer is able to interact with a 3D representation of the real object, to put it into his virtual closet or to discard the selection.

Figure 6.5: Application FashionShop: Concept of the virtual personal fashion show.

Back in the physical shopping space, the selected product is delivered to the customer. This can be either “prêt à porter”, or as a custom manufactured item created from the consumer’s data profile or from measured body data acquired within the portal itself. At the point-of-sales the customer tries-on the real physical garment. Using blue-c technology the idea of the Magic Mirror built in the Prada Store in New York can be brought one step further. By implementing the MIRROR:3D application (see Section A.1) the customer can not only view herself in slow motion from all angles like in the Prada dressing rooms, furthermore she can freely navigate around herself and see herself from arbitrary viewing directions in real-time. The 3D customer representation can be integrated into a prerecorded immersive situation where the selected cloth combination usually is worn. The adjustment of the lighting of the physical environment to the virtual environment allows for an almost realistic experience.

6.3.2 CarShop

CarShop is another scenario for selling luxury cars to prove the applicability of IN:SHOP. Today, prospective buyers go to their car dealer and get a pile of printed catalogs containing all vehicle specific information. At home, the customer studies the illustrated catalog. It is known, that most people have difficulties in imagining how the car will look like they are ordering. Internet based car configurators already help visualizing the huge alternatives. Anyhow, these 2D configurators do not have the ability to visualize a three dimensional model, including exterior and interior configurations, of a customized vehicle. Using CarShop, customers are not only choosing color and accessory varieties at the push of a button, but rather experience the optical effect in full dimensions in real-time.

Figure 6.6: Application CarShop:Presentation of the 3D model of the virtually customized vehicle.

The product presentation and the databases are adapted to the slightly different selling process of FashionShop. The presentation and underlying media database is optimized towards the somewhat different sales process. The potential buyer and the sales associate commonly configure and customize the desired vehicle. The configuration process runs through different selection steps: series, base model, paint color, interior design such as type of upholstery, extras such as rim variety, and financial services. Base model information, such as design, measurements, extras, and available colors, is presented on the main information screen. Further technical details, such as engine configurations, board computers, air conditioning, car stereo systems, etc., are available on data sheets. The individually configured car is presented on a rotating platform as a 3D polygon model (see Figure 6.6). The car becomes experienceable for the customer in its real dimensions. The customer can inspect the car from the exterior, and “fly” into the interior for a view from the inside. Taking the experience one step further, due to an audio system the customer is able to perceive the engine sound, too. The traditional shopping basket metaphor is visualized using a virtual garage. The customer puts and stores his virtually configured cars in the garage and is able to compare them later. Configuration details for the purchase are sent directly to the manufacturer.

CarShop allows potential buyers to interactively configure and experience in three dimensional mode a rendered representations of the vehicle they are purchasing. CarShop generates the desired car model so realistic allowing the customer to freely navigate around the representation. All details can be surveyed from all viewing angles. In addition, the potential buyer gets a detailed overview of the brand, and a preview of vehicle options and features that might not be available in the showroom. Thus, the customer becomes the designer of his own car of desire. CarShop allows presenting a wide range of models and their accessory varieties in a real showroom.

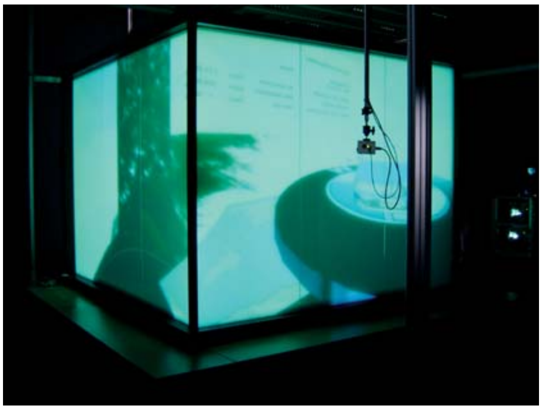
6.4 Technologies

This section describes the technologies used to implement the IN:SHOP application.

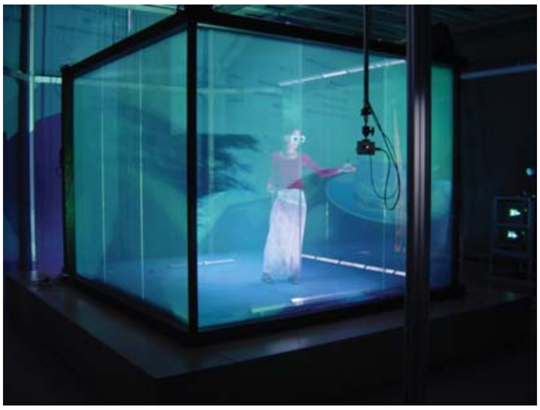
6.4.1 3D Video Acquisition

The application takes advantage of research that combines the rendering of real objects in real time within networked virtual environments described in Section 4.6.1.

Real-time 3D user reconstruction and streaming is the distinctive feature of the blue-c system. This technology supports concurrent projection of the virtual shop environment and acquisition of the sales associate and the customer. Based on standard glass walls with switchable transparency , camera images are captured through the projection walls of the CAVE^{TM}-like portal at ETH Rechenzentrum [Spa03a]. Almost all cameras are hidden behind the projection walls. The cameras can “look through” the projection walls by employing screens with switchable opacity (see Figure 6.7). During a short time slot the walls are “opened” for image acquisition. But since the shutter glasses are totally opaque in this phase the user is not aware of the cameras or the transparent walls. The walls are switched at 62.5 Hz. This is well above the fusion frequency of the human visual system. The same glass walls are already used as part of interactive presentation booths at Prada fashion stores, making the installation of blue-c technology in a real world application very feasible. The user herself enters a blue-c portal without going through any initialization phase. The video acquisition components are hidden from the user so that he is not aware of it. From 16 video streams a 3D representation of the user or any other object inside the two blue-c portals, except for objects which are too close to each other, is rendered. Since 3D video inlays are available as three-dimensional representations they can be easily combined with the virtual world projected onto the display. Thus, a seamless and natural integration of inlay and virtual world is perceived. Figure 6.12 gives an impression of the visual appearance of the perceived display. The multi-pass renderer provides blending and filtering to reduce aliasing and acquisition noise. More in-depth discussion of the 3D video pipeline can be found in [Wue04].



(a) During projection the walls are opaque



(b) The cameras "look through" the projection walls while the screens are transparent

Figure 6.7: 3D video acquisition: Based on standard glass walls with switchable transparency, camera images are captured through the projection walls of the CAVE^{TM}-like portal at ETH Rechenzentrum.

The 3D video technology described in Section 4.6.3 allows presenting the products as interactive 3D video objects. The recording technology is used to play back sequences of mannequins presenting the latest fashion on the virtual catwalk. The user can freely navigate in 3D around these objects. In addition, static 3D objects can be acquired using a multi-camera scanning system, represented and edited as surfed objects [Zwi01], and introduced into the scene using a multi-pass point renderer. This allows to quickly integrate real-world 3D objects into the scene without a major modelling effort.

6.4.2 blue-c API

The *blue-c application programming interface (API)* [Nae04] is a software toolkit for media-rich, collaborative, immersive virtual reality application. It provides an application development environment that offers flexible access to all software and hardware features, including graphics and sound rendering, device input, 3D video, and scene distribution for collaborative work. It integrates real-time 3D video technology, and uses the blue-c network layer for all communication tasks. The subsystems are provided as services and managed by the core. Figure 6.8 illustrates an overview of the blue-c API.

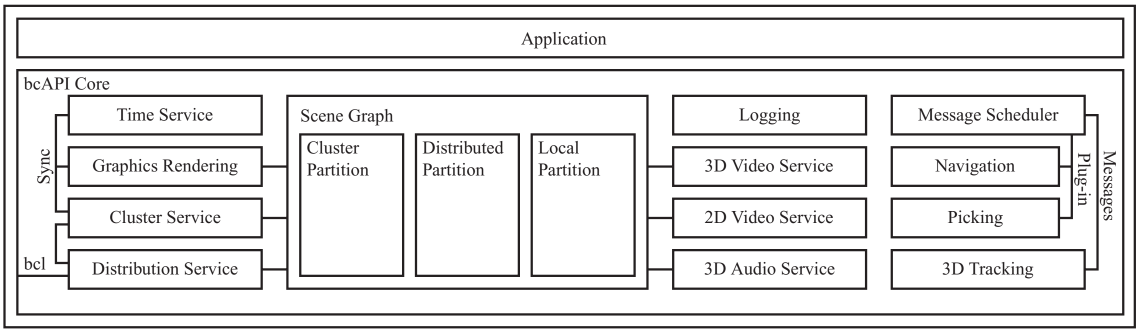


Figure 6.8: blue-c API overview.

FashionShop as well as Carshop are media-rich applications and integrate different media beyond the traditional 3D geometry. The blue-c API is actually a collection of services which all operate on the scene graph as the main data structure. Services can be activated on demand and customized using configuration scripts. Services include graphics rendering, 3D audio, the scene synchronization system which operates only on the shared partition of the scene, and others. The application directly accesses the scene graph data structure and communicates with the services through provided interfaces.

The blue-c API integrates different media types beyond the traditional 3D geometry. All media-relevant data in the scene graph which is traversed concurrently by the different rendering services are kept. This guarantees timing accuracy between the different services in the order of a single frame. 3D video is integrated into the scene by the use of proxy nodes which can be treated almost like regular geometry. Traditional 2D video files can be used as animated textures. The 3D audio server is controlled by active nodes in the scene, therefore allowing adding "sound attributes" to geometry objects. Last but not least, active nodes enable key-frame based animation.

The programming environment as well as the resulting code run on SGI IRIX and Linux operating systems. The blue-c API is built on top of the SGI OpenGL Performer API, which offers a scene graph and real-time rendering system [Roh94]. The OpenGL Performer toolkit is extended to provide a distributed scene graph maintaining full synchronized down to vertex and texel level. Therefore, the application code of all IN:SHOP applications is written using the C++ programming language and the SGIOpenGL Performer software library.

More in-depth technical explanations of the blue-c API functionalities and implementation issues are described in Martin Naef's dissertation [Nae04].

6.5 Implementation

Creating a distributed shopping experience requires both the design and the programming of a convincing storyboard, a specially modelled virtual environment, interaction metaphors, and intuitive user interfaces [Lan03b].

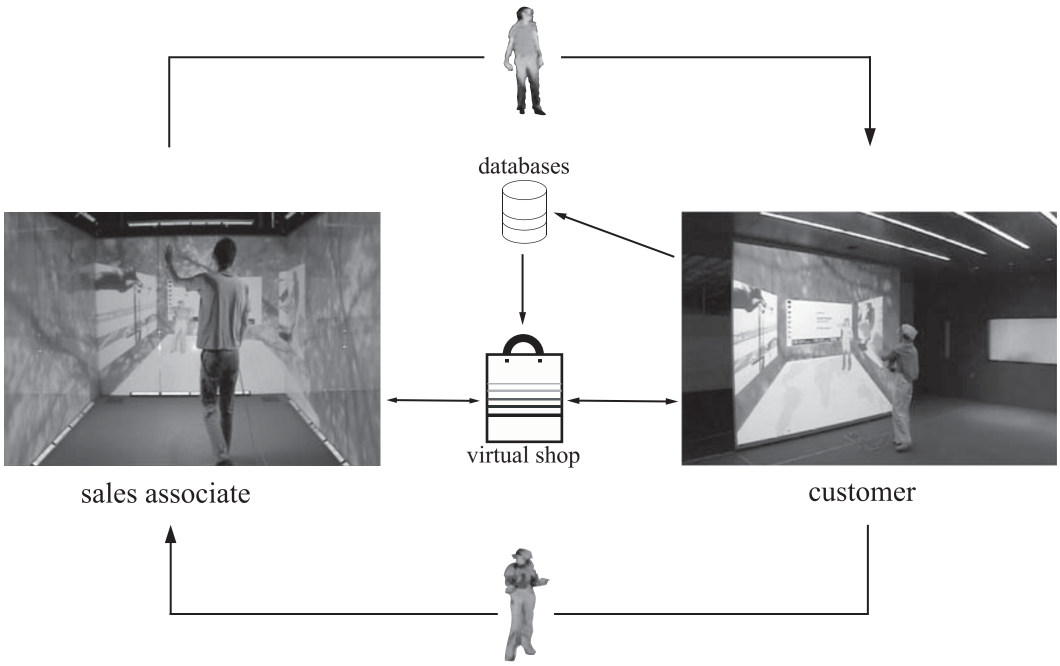


Figure 6.9: Distributed application concept. Using a shared database, tele-presence and networking technology a unified shopping space among distant locations is created.

Figure 6.9 illustrates the distributed application concept. The customer and sales associate interact in the same shared virtual shopping environment. In the virtual world, the customer gets direct access to a personalized shopping environment by entering one of the portals. A unified shopping space among distant locations using a shared database, tele-presence, and networking technology is created. The implementation of the FashionShop and the CarShop environment mainly differ by the underlying product information. The implementation of both applications was done in close cooperation with Martin Naef who has been member of the Computer Graphics Laboratory at the ETH.

6.5.1 Framework

A complex system like blue-c needs an easy to use framework to support the design of architectural applications. Therefore, a framework [Lan03a] to hide the complexity of the blue-C API was designed. This framework is built on top of the blue-c API and supports architects with basic programming skills in creating spatial scenarios using the blue-c environment. It concentrates on the integration aspect of different media and data types. Using this framework a single virtual space combining multiple media can be created. The term multimedia describes the use of text, graphics, animations, pictures, and sound to present information (see Section 3.1.4). Generally people remember 10% of what they read, 20% of what they hear, 30% of what they see, 50% of what they both hear and see, and 80% of what they personally experience [Tre67]. These media have different usefulness and speed in the process of architectural design. In this context it was realized that images, movies, 2D drawings, and animations are quick media. As opposed to these media 3D models, renderings and 3D representations are slow media. Nowadays, the pragmatic use of different available media in architectural design is inevitable.

This framework provides patterns that are easy to understand and fit a wide range of creating virtual shop environments using blue-c features without knowing all technical details behind it.

6.5.2 Storyboard

The virtual shopping environment consists of different presentation modes (see Figure 6.10). The customer can freely navigate between these modes. The following scenario describes a typical way a customer experiences CarShop at the point of sales.

Enter

A customer decides to buy a new car in a showroom. He enters CarShop, attracted by commercial campaigns. He sees the newest advertisements and new features.

Welcome

The system identifies the customer as John Morgan. His personal profile is loaded and the virtual shop environment is customized accordingly. The profile consists of personal data, buying history, items of interest, vital statistics, and personalized suggestions from in-stock and coming inventory. Based on the customer's history, the shopping environment is customized, appropriate new items and personalized advertisement are presented. John is fully immersed in his personalized environment where he meets the sales associate from his patronized shop from Switzerland and another subsidiaries.

Catalog

John browses the personalized product catalog assisted by the remote sales associate, Bob. As an expert user, Bob has direct access to the full content database and can, therefore, present additional items on John's demand. The sales associate has access to additional details to answer the customer's questions.

Selection

John expresses an interest in a certain car model series and selects it. The selection is stored and the content for the following stages is processed accordingly. John can freely navigate between different presentation modes.

Customization

John and Bob configure the selected car. Base model information such as design, measurements, and available colors is presented on the main information screen. Further technical details are available on additional data sheets.

Configuration

The customer and the sales associate discuss the extras. They decide on interior color and style, wheels, board computer, air conditioning, etc. Detail information is presented on sheets, the main screen always reflects the current selection and configuration stage.

Preview

An overview of the configuration decisions is shown as two dimensional detail images, allowing direct access to change previous customizations.

Presentation

Finally, the three dimensional model of the personalized car is presented on a rotating platform. John gets a briefing of his configuration and a statement of cost. Now he can decide to order the car, put it into his personal shopping garage, or discard the selection.

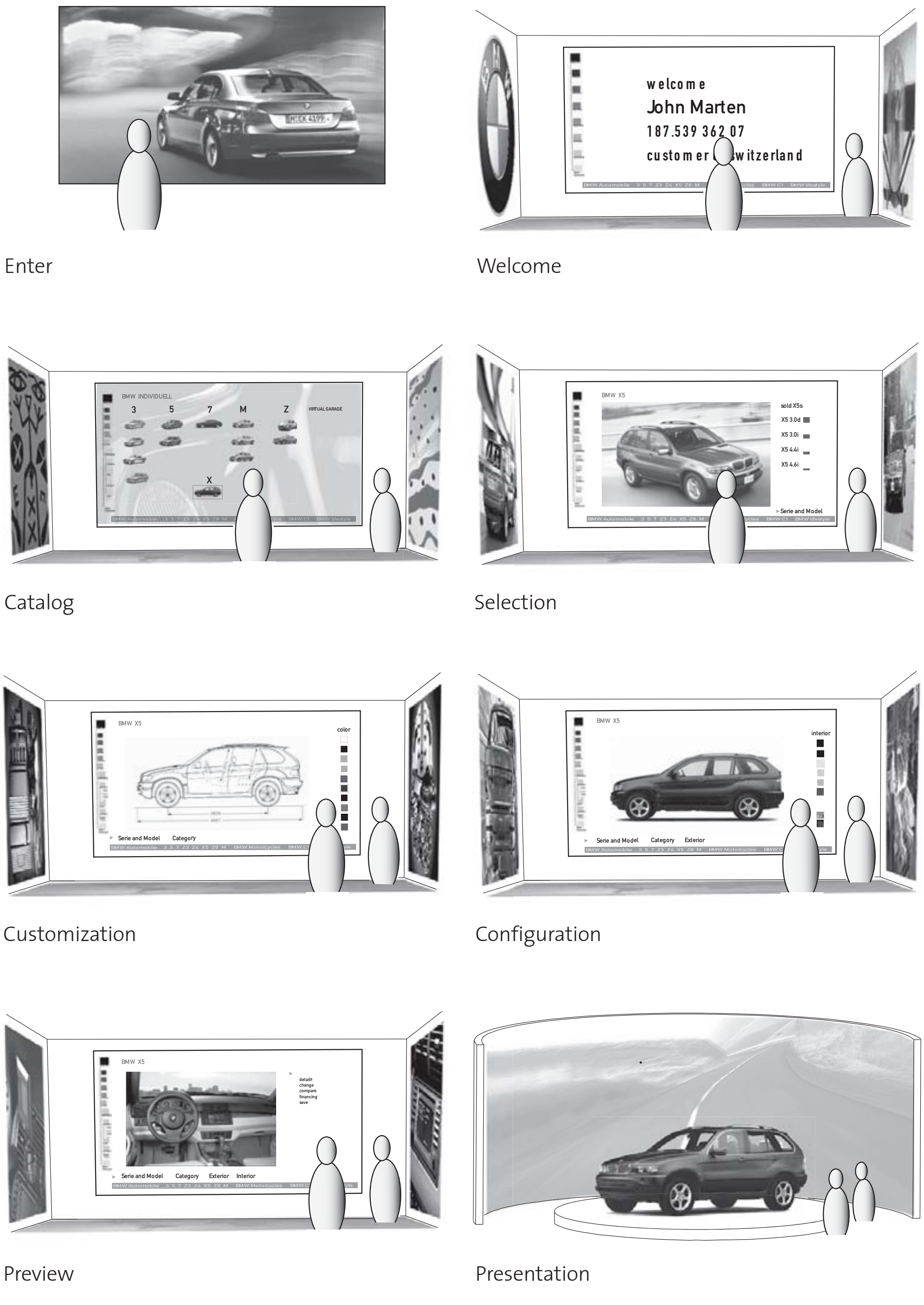


Figure 6.10: CarShop: Presentation modes.

6.5.3 Virtual Environment

The experimental virtual shop consists of three information displays and a ground platform (see Figure 6.11). The platform is an interactive map that illustrates the possible geographical disparate locations the customer can access. The main screen is for presentations and for advises to the customer. It is connected to the customer and the content database. Therefore, it displays all kinds of information ranging form the product catalogue to product suggestions. The side screens provide the customer with background information such as campaigns, production processes, technical explanations, history, and so on. The screens can display different media types such as images, slide shows, and movies. In a next step, it will be possible to feed the screens with live and recorded footage from video cameras. Text and image based data sheets can additionally be introduced in the virtual scene supporting the counselling interview. Customer and sales assistant use them for discussing details in the configuration process. The virtual shop is surrounded by a large cylindrical video wall. The shop environment changes according to different configuration stages. Depending on the configuration step the ground platform moves like an elevator to the relating stage. On the one hand, this supports the users orientation, and on the other hand, demonstrates the spatial changeability that is only possible in virtual space. At any point in time the customer can jump between different stages. Sound is used to generate the appropriate acoustical ambiance as well as to support the user interface. Sound sources can be placed anywhere in 3D space. Sound service are also used as a voice communication channel between the participants.

6.5.4 User Interface

The visual interface is a very important component of the virtual environment in giving the users the feeling of presence. IN:SHOP consist of a stage that is dynamically created based on the customer. The stage is generated and adapted according to preferences and history that can be specified by the user. Therefore, the navigation within the application is customized and personalized. The adaptive user interface allows for a personalized presentation of the significant content to the customer. It combines the customer profile by the user model according to his interests and the meta-content. The meta-content includes the actual content, yielding a personalized presentation of the content, and individual dynamic navigation structures for its exploration. For a good interface design it is important that the user is able to switch the profile at any point in time. The interface allows both a target search on specific topics as well as quarrying in the thematic structure during the shopping process.

Complex interfaces and navigation are unsuitable for most users of the virtual shop environment of IN:SHOP. Customers use the application rarely and sporadic, and sales associates are faced with several tasks while assisting the customer. Therefore, the interface has to be as simple and naturally as possible.

Navigation Structure

Navigating through the virtual shop takes place via the main information display. This display is reasoned based on graphic user interfaces (GUIs) of web-based interfaces and consists of two navigation menus, and the content frame (see Figure 6.11). The goal of this interface is to allow a user to navigate through the virtual space, and to find their way through the databases in an interactive and intuitive way. The consistent placement of the navigation system is important for the user to find his way through the shop, shows them where they are, and therefore prevents them from getting lost. The information display is divided into three panes. The areas are devoted to distinct functions: one area displays customized navigation information, the other area provides the general store orientation while the third area is devoted to displaying the main content for the node.

The bottom navigation bar is the main navigation through the virtual shop where general information can be accessed. It concerns a menu-tree navigation whereby the information is structured under major topic headings and sub-headings. The user navigates down a particular path and back out of that path using visible menus. The menu and path remains visible and is evident at any point in the tree structure.

The left navigation bar is a personalized tab-stop navigation whereby information is again structured under major headings and sub-headings, where the user accesses that information using a file-folder tab-stop metaphor.

Product Presentation Possibilities

The sales associate presents the selections to the customer using a variety of traditional and virtual based media. Different forms of data can be presented in the immersive space, ranging from pre-recorded 3D representations of humans and objects, 2D movies, images, hand drawn graphics, to simple text (see Section 6.5.1).

In addition to the main information display, data sheets are used for the counselling interview to discuss details and to configure the product. These data sheets fade into the virtual environment whenever they are needed and fade out as soon as they fall into disuse. Figure 6.11 shows the use of a data sheet displaying additional technical details relevant to the car configuration process.

Since the approach used for extracting geometric information is not model-based also arbitrary objects can be introduced into the system besides the user. Figure 6.12 shows a typical scene within FashionShop where the sales associate presents a sweater to the customer, a prime example where model-based 3D video approaches typically fail. Thus, the objects within the scene are no longer modelled, and therefore they are not simplified or simulated.

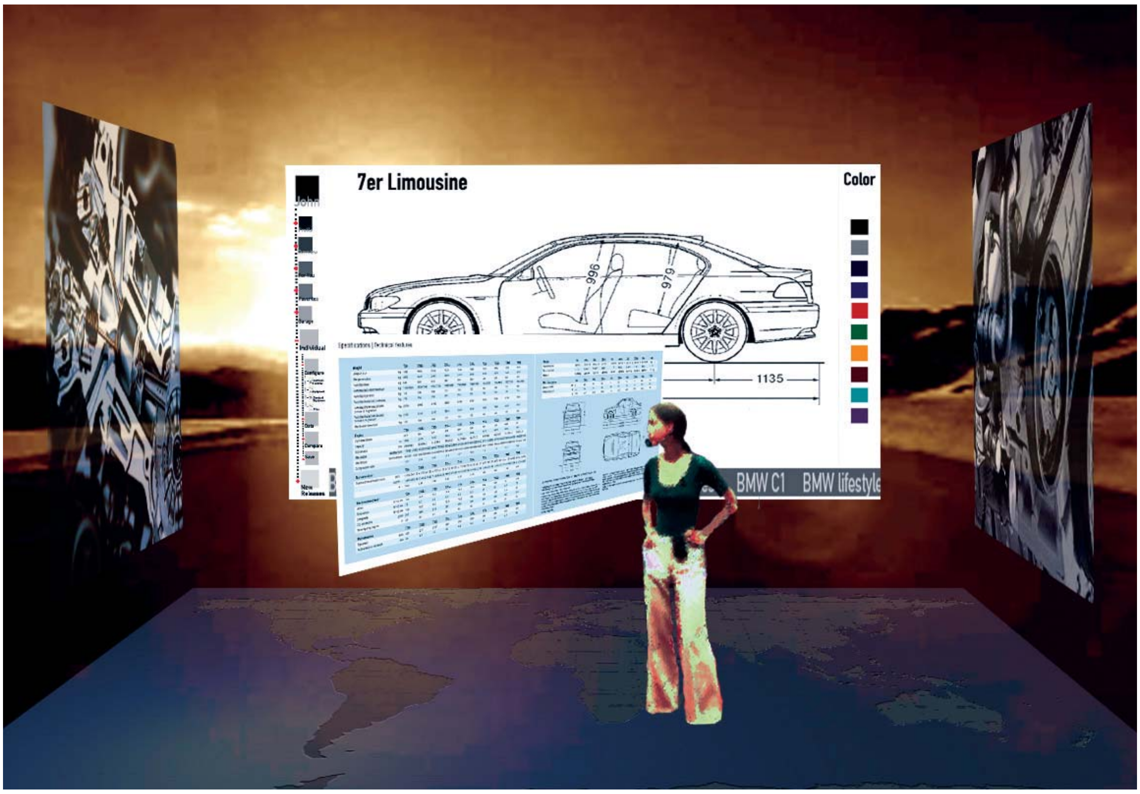


Figure 6.11: CarShop: Navigation via the main information display and data sheet for providing technical details.

Auditory Interface

The auditory interface is an important component in a multiuser virtual environment for collaboration. Audio is a very effective means of communication. The input audio data can derive from pre-recorded audio, live input, networked input, and synthesized audio.

Directional sound technology can be used to render an immersive sound environment, where the user is able to tell where the sound comes from within the environment. A sound rendering system is used for background theme music to enhance the virtual shop as well as sound events to support the user interface. Sound sources can be placed anywhere in the virtual space. Localized sound sources are used to playback sound whenever a user approaches one of the side screens. Recorded sound belonging to the displayed image or video occurs. The playback stops when the customer departs. This makes the virtual world more realistic and increases the sense of presence. Sound services are also used to enable voice communication channels between the separated users. The live audio input is important for the natural communication between the co-located users of remote portals.

All sound sources are rendered by the audio rendering system described in [Nae02].

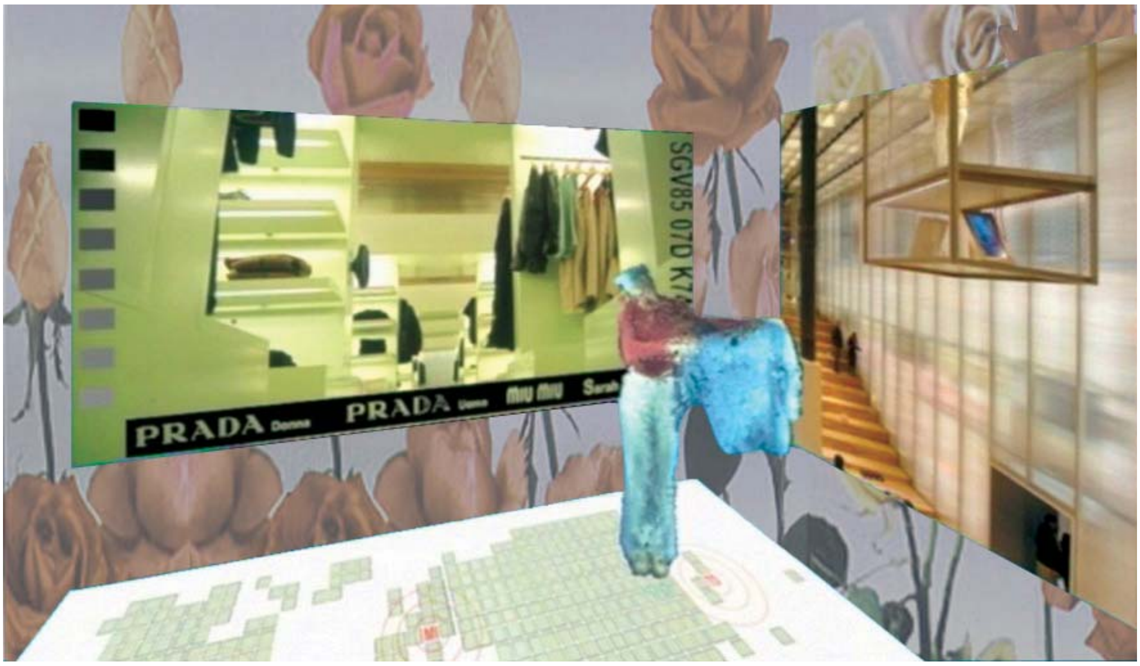


Figure 6.12: FashionShop: Sweater presentation using 3D video technology.

Video Hyperlinks

The prerecorded 3D video sequences of mannequins are selectable objects. By clicking on the video objects additional information is introduced to the customer. The additional information is presented according to how and when the user wants to receive it. The information is available according to his personal profile, e.g. language, level of detail.

Input Devices

The development of suitable interaction techniques is crucial for navigating inside an immersive virtual environment. Common input devices such as mice and keyboards can not be used within the virtual environment, and they prevent a natural and intuitive movement within the immersive experience.

Therefore, the primary input device used for navigating and for inputting commands is a 3D mouse. The customer and the sales associate interact with the virtual shopping environment using a standard six degree of freedom mouse with a small joystick and three buttons (Fakespace Wand). Figure 6.13 illustrates the basic navigation functionalities necessary to navigate through the virtual shop environment. The joystick is used to move forward, backward, or to rotate left and right. The left button is for navigating through the main information screen and functions similar to traditional mice as computer input devices. The middle button allows user to teleport to predefined viewpoints within the virtual space. This prevents the user from getting lost in the scene. Users who loose their spatial orientation while navigating in 3D world can therefore easily go back to the starting point or another significant point within the scene. The right button is for fading in and out additional information such as the data sheets. As an expert user, the screens are controlled by the sales associate.

To aid application development and debugging at the desktop, mouse, and keyboard input is supported.

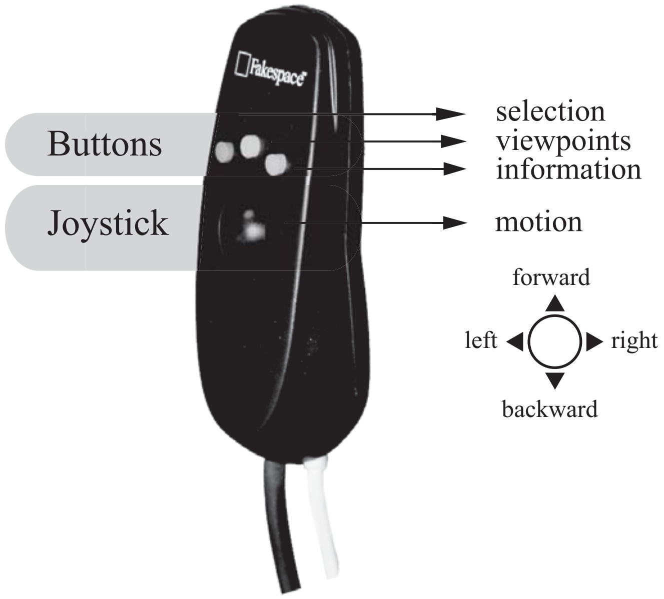


Figure 6.13: Wand device and assigned navigation.

Future Devices

Creating a compelling and engaging interactive virtual environment, video and audio are the most fundamental elements. They not only acting as ambient elements as described in Section 6.5.3, even more they are important as input methods. Video can not only be used for recording and replay, but also for gesture recognition and analyzing of motion (see Section 5.3.4). Collisions of the user with his environment are detected and will be used for interaction with the virtual shop environment in two ways. First, for pressing the virtual buttons on the main information display. Pushing a button is identical to detect a collision with the projection screen at the coordinates of the button. This allows using the projection screens as touch screens. Second, collisions with objects within the virtual environment will be used to interact with these objects. Applying a force as soon as a collision with one object is detected enables grabbing and moving of this object. Therefore, the main navigation and interaction tasks using the Fakespace Wand device can be replaced by natural interaction mechanisms. The future goal is to reduce as many as possible additional input devices and reduce them by interaction techniques people are already familiar with like gesture, voice, and movement. Video and audio are media that enable controlling the environment intuitively without using special devices.

6.6 Evaluation

In the following some experiences derived from inserting and interacting with 3D representations of the users in the application IN:SHOP are described.

Test Plan

Before starting the real test run, the test persons are introduced to the functionality of the system and the shopping environment. The use of natural speech and gestures as well as the common navigation tasks discussed in Section 6.5.4 are explained. In order to assess whether the presence of a real-time 3D representation improves the quality of communication, a task in which participants would benefit from having visual feedback is needed. Therefore, the applications CarShop and FashionShop are ideal for evaluating the applications as well as the experimental system.

It was pointed out that this is not an acting task, and that all participants should be themselves and should feel free to improvise. This helped to find out if in same-room interaction, factors such as interpersonal distance, posture, and gesture might have an influence.

User Experience

The fundamental variable of interest is the suitability of 3D video for user interaction. It is expected that 3D video improves the communication within Tele Reality applications. The questionnaires are subdivided into the following five indicators, similar to [Gar01].

1. *Eye Contact*: The extent to which the shopping process was experienced as being like a real face-to-face client counselling.

2. *Immersion*: The extent to which the test users experienced the virtual shop and felt as a part of it.

3. *Presence*: The extent of presence of the remote sales assistant, i.e. meaning the sense of presence of the remote user.

4. *Communication*: The extent to which the interaction and communication were comparable to the traditional situation, and the extent to which the experience was enjoyable.

5. *Learning-time*: The extent to which effort is necessary to use the system.

(a) The test person gets instructions to functionality of the system and the application CarShop

(b) The test person uses the application FashionShop

Figure 6.14: Evaluation of IN:SHOP: test persons with different backgrounds were taking part in the experiment.

Preliminary Results

15 persons with different backgrounds were visiting the blue-c system and taking part in the experiment (see Figure 6.14). All of them were technically skilled, nearly two thirds stated to use tele-presence systems. Half of them already had experiences with virtual reality systems and immersive environments.

The response of the variables described above were obtained by means of a questionnaire with each participant. It was noticed that all users were hesitant and reserved using the application in the first minutes.

All of the users tried more than once to grab 3D objects within the scene or to touch the 3D representation of the remote user. All users found it very important for the counselling process that they not only hear the sales associate, but also see his gestures. They felt that this greatly enhances the sense of presence and reality within the virtual shop. They liked being able to point on information displayed in the virtual shop. All users found that they were able to talk to the remote user like in the traditional situation where both are in the same physical space. They all enjoyed configuring their personal car with a remote sales associate. Even though our stereo glasses minimize real eye contact between the customer and the sales assistant, nearly all user experienced the shopping process closely to a face-to-face client counselling. They considered to use smaller glasses comparable to the size of sun glasses. Users familiar with tele-presence systems liked this concept because they felt that the 3D representation gave them a better understanding of the remote user. Persons who never used such systems before were focused on and occupied with one tasked. Either they experienced the virtual environment or were busy with the navigation. In the beginning they even forgot that they can talk to the remote user. By having a skilled user at the remote location who took over the navigation and started the conversation this initial inhibitions could be overcame.

Although the users were generally enthusiastic about the potentials of the system, they felt that some improvements have to be done. Half of the users would appreciate using the screens as virtual touch screens. Some of the users proposed to use PDAs as input devices. A third of the users wish to be able to put the remote user their real jacket on.

Overall, the interviewed users were enthused and felt that the concept would be helpful in increasing the realism in immersive tele-presence applications, although the quality of the 3D video representations has to be improved.

6.7 Discussion

Normally in time, technologies in architecture are used for being productive, for solving problems, for affording temporal, economic, and manufacturing efficiencies. However, technologies may also be used for creating communicative spaces. Technologies are no longer just design and fabrication tools for architects supporting their design tasks. In fact, treating them like design materials integrated into the building design process, they provide architects with the opportunity to generate interactive and communicative spaces. Over the last few years, it can be observed that displaying of objects and situations is becoming more and more important. Acceptance, attractiveness, and performance are leading marketing goals. Production is no longer the pivot of our economic and cultural life. Looking at the Prada Soho store in New York, architecture and design are used to frame, state, and represent the objects to be sold. As a result, the role of architects extends beyond designing and constructing buildings. They are also going to create places and scenarios for display. Display takes place in stores, restaurants, bars, theaters, sport and vacation facilities, and many more (see Section 4.2). People come to different kinds of places in order to see, to be seen, to purchase, and to get compressed information. Since architects are concerned with appearances and staging they are designers of urban environments where display takes place. In other words, architects are also engineers of display [Bet03]. They define appearance of goods and control the flow of people.

Combining real and virtual worlds allows designing enabling interfaces that build on the best affordances [Nor02] of everyday reality and vitality. Video streams of advanced information technologies offer novel ways to reduce spatial and temporal interdependencies between different activities. Video systems help, therefore, to bring information closer to the users and make them part of their physical environment. Thus, they offer a promising contribution to bridge the gap between virtual and actual reality.

With IN:SHOP it is possible to introduce a novel approach to distributed shopping. The application allows geographically disparate customers and sales assistants to communicate and interact with fully realized 3D representations of real objects in real-time. The 3D representation of real persons or any other real world object is a novel visual presentation media for communication and man-machine- interaction. Architecture, information technology and computer graphics are combined.

Future Scenario

The core feature of the blue-c system, the video acquisition and rendering of a user into a computer generated world, has been exploited by the application IN:SHOP. Users at different physical locations are brought together in a shared virtual environment. This application demonstrates that 3D representations of any real world object constitute an important component for interpersonal communication and for conveying the intention of the designer.

The obvious next step is to bring physically distant persons together with their environment in a virtual space. Virtual refers to the situation where people are not in one physical location, but in remote, distributed locations. For example, people of a virtual space are still people in real physical spaces - only in a different location. Covering larger areas with novel data presentations for 3D video of complex scenes allows to reconstruct complete physical shopping floors as well as persons (see Section 4.6.2). Therefore, it is possible to extend and connect physical spaces into remote spaces. In the future people are going to communicate and share information independent of the physical location. Sales associate, customers product suppliers, and experts no longer have to be brought together in the same location. They have the possibility to establish novel forms of contact and interaction. In order to achieve this there is a need to open the physical space, to reduce the cost of installation, and to develop robust and scalable setups.

Towards an Enhanced Reality (see Section 5.3.1) stores will not only provide products, but also customer related services. The stores of tomorrow will have the ability to understand and to be responsive to the needs of their customers. Spatial and temporal linkages among different subsidiaries, producer, manufacturers, experts, and consumer are selectively loosened. Multimodal and multimedia interaction enable an advanced customer centric counselling and a novel concept of shopping. This again translates to a novel store architecture. Taking the privacy and security issues discussed in Section 5.1.2 in account, information on persons, object, and locations afford to broaden the product range and to lower costs of stocking. The introduction of video systems into shopping environments liberates from known commercial pressures, leads to novel communication channels between customers and goods, and combines one-to-one services. Networked shops equipped with cameras, sensors, and actors to acquire information offer the customer more comfort and an adaptive advisory service. It is crucial that the customer is aware of the inbuilt technology. The customer has to be able to access these additional services without special input devices and additional learning time.

Feedback from the Industry

Several presentations and talks with experts from the car industry have shown that CarShop as a prototype application is a promising contribution to future showrooms. There is no doubt that the application is not ready for marketing and that by all means the quality of 3D video has to be improved. Different aspects occurred while talking to persons from the research sections and persons from the marketing section.

The research departments are more concerned with the highly iterative and time consuming design process. For them it is important to improve the cooperation between the different specialists, and to reduce development and production costs. Bringing research, data from the product development and the design process to the point of sales may increase the product range by 40%. Thus, the amount of provided information to the customer goes up, the range of products expands, and a better personal customer support can be provided.

For the marketing departments three major contributions are interesting:

1. The interactive 3D car configurator where the potential buyer is able to experience himself together with the representation of his new car in different environments. The lack of typical VR input devices such as data-cloves and the effort to use natural interaction mechanisms such as gestures and voice instead is an important factor offering the application to their car sellers.

2. The potential for their car sellers presenting a wide range of models and their accessory varieties in spatially limited physical showrooms without purchas-

ing several expensive demonstration cars. This is, especially, relevant for the European sector where customers order their car ex factory.

3. The modularity and scalability of the hardware components allow to adapt CarShop to different, already existing showrooms.

CarShop will contribute to the long-term goal of the car manufacturer towards the development of a market place to encouraging a stronger market identification. They are aware that architects play a significant role towards the ideas of selling products of multifunctional venues. Leading car manufacturers in Germany started in using architecture for creating places of experience where purchase and handover of the vehicle become an unforgettable experience.

For the fashion industry the quality of 3D video is more crucial and a key factor for bringing FashioShop to realization. Specially, presenting fashion motion is an important feature. Dress material is effective until it is worn and one can see how it drapes. Therefore, the fashion industry is heavily interested in capturing high resolution motion scenes. The biggest concern for them are issues for protecting privacy. In general they are positive about enhancing the already existing IT concepts by bringing them to the third dimension and connecting different locations. Integrating the idea of IN:SHOP in a real fashion store would only be possible at the moment by using mono projection. The customer of today would not feel comfortable wearing big stereo glasses. FashionShop, however, can contribute to the idea towards making the act of shopping central to lifestyle and also matter of a cultural debate [Chun01]. The notion behind these new shopping locations is to reshape the concept and function of shopping by encouraging for meshing consumption and culture.

7 Requirements for Video Systems in Architecture

This chapter describes and repeats the consequences that emerge from video systems on architecture. This research has shown that there is a great potential in designing technological enhanced environments by applying video systems as an alterable and agile design material. There is an increase in popularity to use architecture as a sign for brands and principle offices (see Section 6.2.3). Given that architecture is traditionally durable and sedate whereas 'venture worlds' are dynamic and agile, it is a great challenge for architects to meet the requirements of communication the companies ask for. Merging architecture with information and communication technology enables to catch up with the dynamics of the merging, purchasing, and selling practices of companies. The implementation, analysis, and discussions with experts have shown that there are still some factors for improvements towards the realization of Tele Reality. In the following, the requirements to be fulfilled are discussed.

7.1 Goals

This research demonstrates how the combination of video systems as part of communication and information technology with architecture is exploited to generate augmented environments. The goal in doing so is to bridge the gap between virtual and actual reality, so called Tele Reality (see Section 4.5). To overcome the substantial space and to carry on the human senses beyond built walls is a primary characteristic for architecture. Video systems open up unimagined design scopes towards overcoming substantial limits. Towards the design and development of Tele Reality by applying video systems an application space approach (Figure 7.1) is provided for the first time. On the basis of the breadth of systems applications are pinpointed and marked as areas for further investigations. As proposed in Section 4.5 the five attributes camera, display, immersion, realism, and capability describe the application space.

The fundamental objectives are to address real problems, to offer technical, economic, and sustainable feasibility, and to guarantee adequate usability by people. A main contribution for architects will be applying current shifts in the development

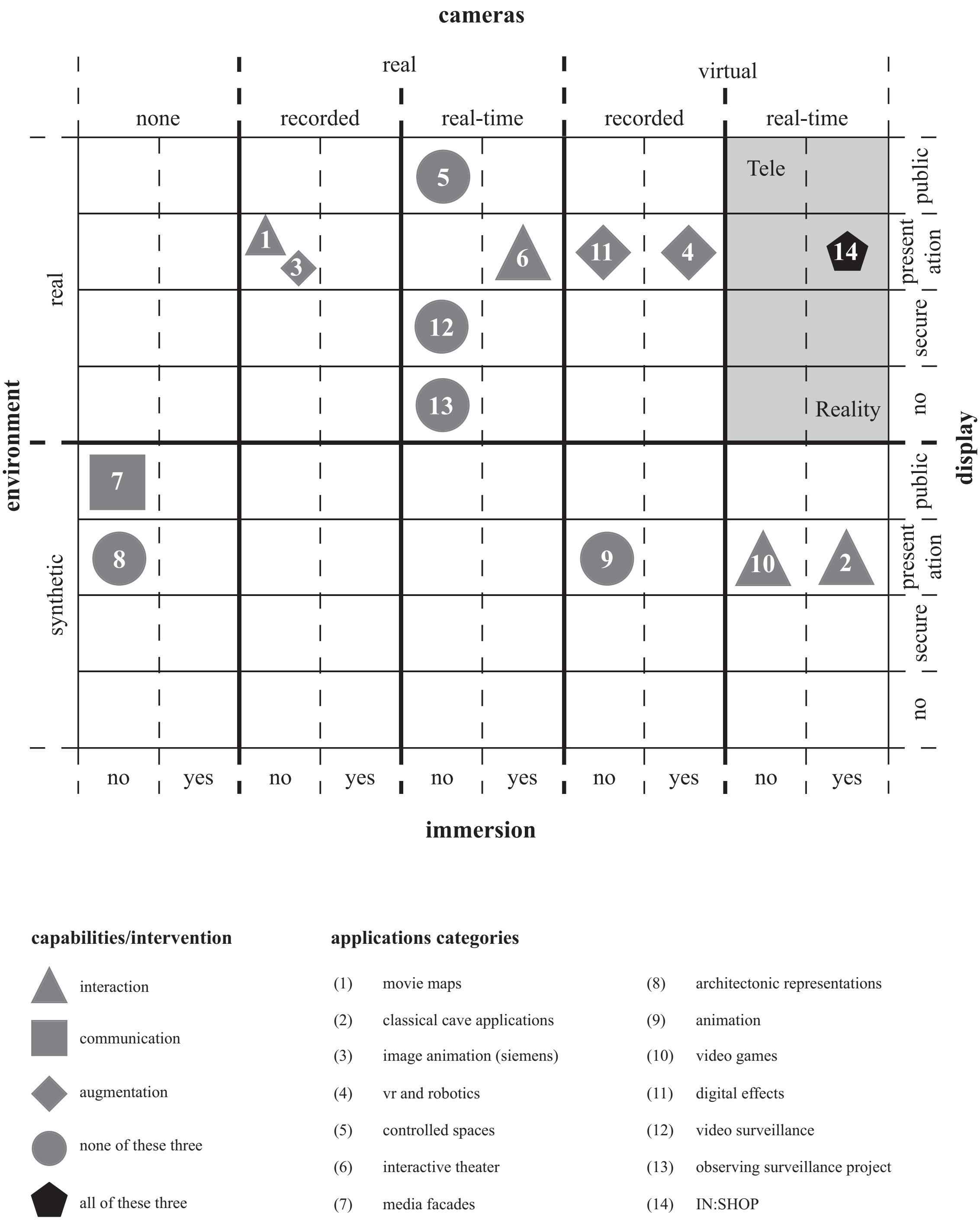


Figure 7.1: Application Space.

of information technology and experiences in interactivity. They are asked to design spaces including computational systems where people can interact in an intuitive way. The goal thereby is creating highly interactive and visual experiences using the latest information technologies amalgamating with architecture. It is important that the technologies and their system's complexity remain invisible and hidden to the user.

7.2 Technologies and Performances

For a seamless integration of video systems into buildings and for treating them as design materials, technology and performance issues have to be accomplished. But the success of a novel system, however, still strongly depends on user acceptance.

7.2.1 Gaps to Fill

Whether the vision of Tele Reality comes true depends on many enabling, facilitating, constraining, driving, or hindering factors. These factors are both of technical and of human nature.

Technical Factors

Technical gaps and the co-dependencies between the developments of information technology are arbitrative factors.

For people to feel comfortable with and to trust digitized and connected environments, an acceptable level of privacy and security needs to be established. The challenges in terms of trust, confidence, and security are enormous. It is fundamentally difficult for computers to understand trust, since it is an intangible asset. Since enhanced environments are not only for specialists, but rather for non-technical people it is constraining necessary that they are user-friendly. In this context, dependability, reliability, robustness, and predictability are important factors. The technical infrastructure has to be upgradable, scalable, and flexible. For a technology acceptance by the public it is crucial to achieve a sufficient usability and reliability.

For integrating video systems into enhanced environments in general practice, global standards are an important factor. Suitable definitions are crucial.

Video transmission and processing depend on high bandwidth needs. In order to depict with video systems dynamic large scaled environments, many 3D images must be computed very fastly. Rendering dynamic scenes at high spatial resolution requires high processing requirements as well as a great volume of data. The success of real-time Tele Reality applications depends primarily on the time delay between distant cooperating partners. The delay from sender to receiver and backwards (feedback loop) is therefore a critical factor, especially regarding interaction possibilities. The quality of 3D video in relation to the processing speed must be satisfactory. To achieve the goal that users will interact in real-time with free viewpoint video captured from real scenes and objects, a high level of quality and accuracy has to be guaranteed. Regarding visual impressions the human eye is very precise and hard to please. As soon as the quality of 3D video goes along with the high resolution images we are used to, such systems are marketable and will be applied.

Internet and mobile devices expand distant human-to-human communication. Broadband IP networks will lead to an advanced style of telecommunication and move Tele Reality a great step further. Broadband networks enable new services that were not possible within narrowband networks. In broadband networks the frequency range is divided into multiple independent channels for simultaneous high speed data transmission of signals such as voice, data, files, videos, music, photos, graphics, x-rays, and 3D. The further development of the networks enables a novel high reality multimodal distant human-to-human communication style. The most fundamental effect expected of broadband digital telecommunications on architecture is replacing physical circulations by networked digital connections. This enables to interconnect constituent, physically distant spaces, and to combine and interlock them according to novel logics.

One assumption to realize mobile services for spatial, temporal, and social distributed activities are mobile end devices. However, the miniaturization of high technology devices and interaction tools (see Section 5.2.3) is not sufficient for a human friendly design. Besides still existing transmission and processing gaps of 3D video the development of natural and multimodal interaction technologies is crucial for promising applications. At the moment most systems lack of simple and satisfactory interaction procedures. In many cases the operation panel is not easy to understand. Making video systems part of everyday environments is going to offer human interaction possibilities which are similar to the real world's human-to-human communication using gestures, mimicry, voice, and context. Traditional input devices are mice and keyboards. They do not really support the vision of an enhanced reality. Therefore, novel interaction metaphors are desirable. They should be adapted to the human needs and their requirements. Context awareness and automatic emotion recognition are still problems hard to please, but it can lead to novel interaction methods.

Human Factors

Social and cultural values and trends framing and shaping people's life are crucial for a successful integration of video systems, because they ultimately condition the acceptance or refusal of novel technologies.

For a high user acceptance it is important that the scenarios contribute to the quality of everyday life. Everyday tasks have to be better supported and enhanced by the introduction of information technologies. Moreover, it is important that people feel comfortable in using these technologies.

The following future trends have considerable relevance in finding appropriate applications, the requiring and stimulating technologies, and innovations. The most influencing trends are that the amount of single households rises as well as the amount of working women and of tele-working. The elderly population increases and the population attaches more importance to individuality, mobility, comfort, security, and well-being. Not to be underestimated are the technical trends of increasing automation and networking.

One big challenge is to develop holistic and humanistic solutions. Holistic refers to solve research questions addressing problems as well as persons. Humanistic describes using information technology to enhance the quality of life. In doing so, important factors are honoring values such as new trends in social and cultural life, and security and privacy issues of individuals.

The long-term goal of Tele Reality is to enable the remote and nevertheless effective participation of geographical separated participants to community life. Distant cooperations play a more active role so that new forms of participation can emerge. A world wide diffusion of modern telecommunication services would also encourage the global social and economic well being.

However, there is still the not underestimated fact that unfortunately architecture has remained close to its ancient origins. At this point in time, drawings and models are done using the computers and computer controlled machines. This can be already considered as a kind of automation, but buildings are still constructed in a non automated way. Information technology can not only help to make building construction more efficient, it allows architects to design spaces where all necessary infrastructures are invisible.

8 Conclusion and Outlook

This chapter summarizes the results achieved in this dissertation and gives an outlook on further research directions and possible extensions. This dissertation deals with the seamless integration of information and communication technology into buildings. It is explored how architects can contribute in designing environments that are more scalable, flexible, and dynamic. The results offer a promising contribution to building intelligence.

Principle Contributions

Throughout this dissertation it is assessed how video systems might transform society and their architecture. Video systems, particulary 3D video, will bring people closer to each other. Since humans are social by their nature they will naturally try to establish relationships among them. However, the central question will be whether such relationships will have the same quality as those formed in person and reality.

At this point, it is time to consider the greatest transformation and impact video systems have to offer to architecture. To get to this point, the key discoveries within this work are reconstructed. This work begins with a description of recent developments in information technologies. It goes on to explore the influence of these technologies on the field of architecture. It is found out that technology, since ever, had an impact on architecture and that architects always use the latest inventions from different disciplines in their designs. The second major discovery is that information technologies will dramatically affect people on a wide scale, and therefore, change spatial needs and traditional building structures.

Putting these both aspects in perspective, it is realized that there are different factors that influence our life. Each of them brings consequences for architects, and enables novel architectural settings. To better understand the architectural impact of video systems the value and its consequences is demonstrated in the application IN:SHOP. IN:SHOP as a prototype demonstrator illustrates that dynamic scenes such as motion and kinematics are essential parts of future application scenarios. The benefit of 3D video allows for innovative and novel application fields. Unlike traditional videoconferencing systems persons and any other objects are fully recorded. Promising operational areas are, therefore, virtual replays, sport events, sport colleges, sport studios, and rehearsals (opera, ballet, musical, stage performances). For a satisfying interaction it is crucial for the future that video objects are changeable. This leads to a few troublesome gaps and requirements: the issue of privacy, bandwidth, video quality, and interaction methods. To leverage video systems, their use needs to be considerably easier than it is today. With increased transmission rates and customization, it can be looked forward to a larger array of better, cheaper, and more customized services that will reach people even faster than before.

Another major aspect is that 3D video will increase the quality to bring people together across space and time. Many social consequences, good and bad ones, will arise as this new technology distributes powers of control from central authorities to many eyes of the worlds people.

Video systems will not be a substitute for the physical world. Many of human most valued actions and decisions, emotions, and relationships including trust, love, and fear do not pass through these systems.

Directions for Future Work

Moreover, people use new technologies in ways that are very different from their intended use by their inventors (see Section 4.1.3). There is no typical, uniform user and use, but rather a diversity of users and uses. Dertouzos already stated 1997 in his book *What WILL be* [Der97] that the real challenge is to discern plausible applications. Whether they are exiting or not, it is important that they can be reliably developed from today's technological trends and are economically feasible. It is mandatory that these applications fulfill some useful human needs.

Successful innovation of enhanced environments will be the result of the right thing, in the right place, at the right time, and in the right combination. In fact, an important area of information technologies involves shelter. Information technologies increasingly have an effect on the physical world. The need for built space might be reduced but not disappear. Therefore, with the help of information technology architects are now able to achieve the challenge of creating buildings that serve people. In earlier decades architects where restricted in their designs to technical needs such as circulations, and were forced to built spaces around the technology. The integration of communication and information technologies into building design will become as integral as walls, doors, and windows. Given that network connections offer the possibility to connect people over space and time, they are as essential to buildings as doorways. Designing intelligent environments requires to center the human, because without him the integrated technology does not make sense.

To achieve the goal of building intelligence and to ensure the quality, it is important to put architects and technology engineers together at the very beginning of a project. At this point in time, the issues of technology and space that accommodates can no longer be separated. These two have to be seen as a couple. Applying developments and experiences in information technology to the field of architecture enables designs where people's needs are a priority. New technology makes designing spaces more challenging, because architects can now focus more than ever on designing spaces around human's needs where people are exited to be in. They do not have to take into account specific technical requirements.

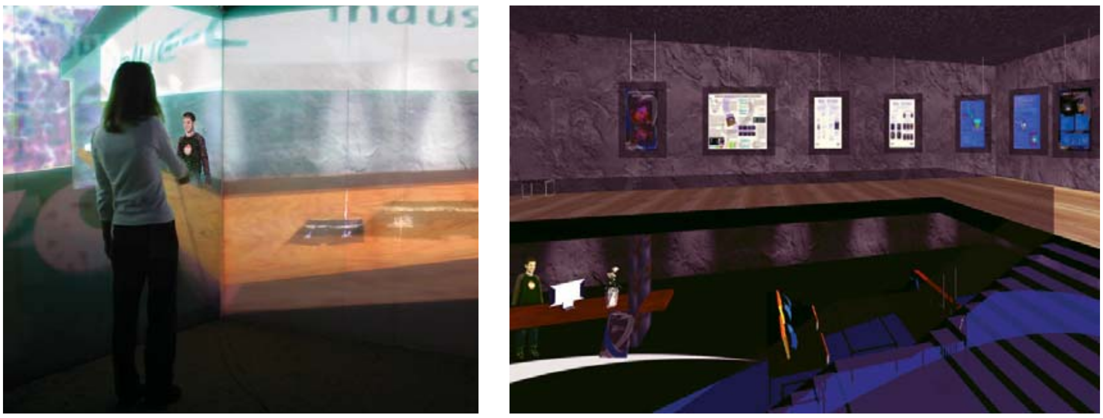
Just like the telephone changed the architecture of its day, and made new connections, spaces, and urban relationships possible, video systems together with further developments in information technology will introduce a novel sense of speed and mobility to social and business life. Technology changes very quickly, and becomes also very quickly obsolete. Architecture instead is a long-term thing and it is necessary to focus on the fundamentals of architecture and to see what happens over the long run.

A Appendix

A.1 FashionShow

FashionShow was the first application to create a world that relies on the blue-c technology. It demonstrates the functionalities and possibilities of blue-c, shown at the blue-c Industry Meeting in June 2002. The implementation was done with Martin Naef.

FashionShow is the development of a virtual architecture where the users experience the immersion and stereoscopic depth of the system. The final goal was to create a space that is intuitive and explorable. The visualization and virtualization of the technology tour into a digital architecture drew from various places and events within the storyline. Digital architectural models are enriched with textures, billboards, images, animations, movies, and 3D video. Sound is used for background theme music as well as sound events to support the user interface. The virtual world consists of two major places: the conference lobby and the fashion show [Lan03a].



(a) Demo at the Industry Meeting starting at the welcome area

(b) Screen shot conference lobby including the areas registration and poster session

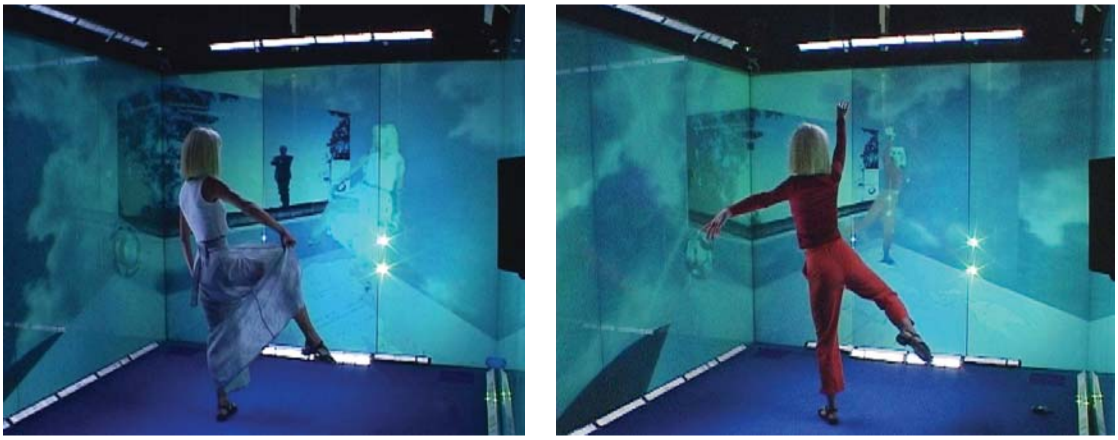
Figure A.1: The application FashionShow demonstrates the functionalities and possibilities of blue-c, shown at the blue-c Industry Meeting in June 2002.

The technology tour starts at the conference lobby divided into three areas according to different activities: welcome and registration, information, and poster session (see Figure A.1). The guest is welcomed by a prerecorded 3D video sequence at the reception desk where he gets an overview of the meeting program. A TV screen located on the ground floor of the lobby displays the logos of all meeting participants. The logos change every 15 seconds. This allows the guest to get some information about the other guests backgrounds. The poster session is located on the gallery of the lobby on the first floor. A wide staircase is leading to the posters hanging from the ceiling. The posters are high-resolution texture maps of blue-c posters presented at international conferences. These posters are not only used to explain technical details of the project, but also demonstrate the projection quality.

Figure A.2: Screen shot of the application FashionShow: The virtual fashion show is enriched with different kinds of media.

At the end of the gallery a gradually widening door leads to a new room. The backstage of a fashion show is located behind this door. The user starts to explore the completely new and unexpected environment. Finally, he finds himself on the catwalk of a crowded fashion show. The audience is represented by billboards to keep the geometry low for a better performance. Flat screens hanging down from the ceiling play several blue-c movies (see Figure A.2). Once the guest turns around he is pitted against himself on two large projection screens. The acquired images of the user from the front cameras are projected onto these screens. Instead of barging into a fashion show, he realizes being part of it. This element of surprise is made possible using the 2D video capabilities by displaying network video streams of acquired and segmented camera images. After understanding what technically happened he leaves the catwalk entering the backstage zone. From here he leaves the room through the exit door that leads back to the first floor of the lobby hall vis-à-vis the poster session. An image gallery consisting of the blue-c group members' logos are attached to the wall.

As the 3D video technology improves the 3D user representation is put onto the catwalk. Now, he is able to experience his movements from all viewing directions in real-time. This feature is also called MIRROR:3D and is similar to a regular 2D mirror (see Figure A.3). The significant difference is the possibility to freely navigate and see oneself in full 3D from arbitrary viewing directions.



(a) Experience of movements

(b) Dancing inside the CAVE

Figure A.3: Application MIRROR:3D: The user is able to see herself from arbitrary viewing directions.

Today, FashionShow is used to showcase the multimedia capabilities of the blue-c system and the blue-c API. Especially the performance of the 2D video system can be demonstrated. Without any visible performance impact several stored and live videos can be streamed simultaneously. FashionShow is not only a testbed for blue-c technology, but also plays with features only possible in virtual space.

In addition, this application served to test the format for loading texture-mapped model files. The best way is doing the whole modelling in 3DS Max and save it in vrml1.0. It is important that the textures are in RGB format. Using a script the vrml1.0 is first converted into the Inventor file format (iv), and than into the Performer native binary format (pfb). The loading times are dramatically reduced using the pfb file format. The application was also used to implement a basic navigation. The navigation features the classical elements like walking forward, backward, left, right, up, and down. Beside this, it is possible to define positions. This prevents the user from getting lost within the virtual space.

Therefore, it is now possible to introduce virtual models into the blue-c environment. This allows to visualize unbuilt architecture in an immersive environment. Classical architectural walk-throughs can be presented, including the 3D representation of a geographical distant collaborator, for example from the building industry.

A.2 Abbreviations and Acronyms

1D	One–dimensional
2D	Two–dimensional
3D	Three–dimensional
4D	Four–dimensional
AI	Artificial Intelligence
AmI	Ambient Intelligence
API	Application Programming Interface
AR	Augmented Reality
ARPA	Advanced Research Projects Agency
CAD	Computer Aided Design
CAAD	Computer Aided Architectural Design
CAVE	CAVE Automatic Virtual Environment
CD-ROM	Compact Disc Read-Only Memory
CD-RW	Compact Disc Rewritable
DAB	Digital Audio Broadcasting
DC	Disappearing Computer
DCT	Discrete Cosinus Transform
DES	Data Encryption Standard
DFT	Discrete Fourier Transform
DSL	Digital Subscriber Line
DVB-T	Digital Video Broadcasting - Terrestrial
DVC	Desktop Video Conference
DVD	Digital Versatile Disc
EU	European Union
ETH	Swiss Federal Institute of Technology
FET	Future and Emerging Technologies
FDDI	Fiber Distributed Data Interface
Gb/s	Giga bits per second
GPRS	General Packet Radio Service
GPS	Global Positioning System
GSM	Global System for Mobile Communication
GUI	Graphic User Interface
HCI	Human Computer Interaction
HDTV	High-Definition Television
HMD	Head Mounted Display
HTML	Hypertext Markup Language
HTTP	Hypertext Transfer Protocol
I/O	Input and Output

IP	Internet Protocol
ISDN	Integrated Services Digital Network
IST	Information Society Technologies
IT	Information Technology
kb/s	kilo bits per second
LAN	Local Area Netwok
Mb/s	Mega bits per second
MHP	Multimedia Home Platform
MPEG	Motion Picture Expert Group
MR	Mixed Reality
MUD	Multi-User Domain
NTSC	National Television System Committee
OMA	Office of Metropolitan Architecture
OSI	Open system Interface
PAL	Phase Alternating Line
PC	Personal Computer
PDA	Personal Digital Assistant
pfb	Performer binary format
PSTN	Public Switched Telephone Network
RAM	Random Access Memory
RFID	Radio Frequency Identification
RGB	Red, Green, Blue
SGI	Silicon Graphics, Inc.
TCP	Transmission Control Protocol
T-DSL	Telecom Digital Subscriber Line
TV	Television
UMA	Universal Multimedia Access
UMTS	Universal Mobile Telecommunications System
VR	Virtual Reality
VRML	Virtual Reality Modeling Language
VTR	Video Tape Recorder
WWW	World Wide Web

List of Figures

List of Tables

Bibliography

[AG4] ag4 | mediatechture company: http://www.ag4.de [accessed: July 12, 2004]

[ART] ART+COM: http://www.artcom.de [accessed: July 12, 2004]

[AWs] activeworlds: http://www.activeworlds.com [accessed: July 12, 2004]

[BIB] Building Intelligence: http://www.caad.arch.ethz.ch/CAAD-Extern/891 [accessed: July 12, 2004]

[BIX] BIX: http://www.bix.at [accessed: July 12, 2004]

[BL] Blinkenlights: http://www.blinkenlights.de [accessed: July 12, 2004]

[DC] The Disappearing Computer: http://www.disappearing-computer.net [accessed: July 12, 2004]

[EB] Encyclopdia Britannica Online: http://www.britannica.com [accessed: July 12, 2004]

[inLiv] interLiving: http://interliving.kth.se [accessed: July 12, 2004]

[LE] Limited Experience: http://www.botticellilimited.com [accessed: July 12, 2004]

[OSP] Observing Surveillance: http://www.observingsurveillance.org [accessed: July 12, 2004]

[plex] PLEX: http://cox.plexgroup.com [accessed: July 12, 2004]

[RC] RaumComputer: http://www.raumcomputer.de [accessed: July 12, 2004]

[TM] Matrix Revolution: http://whatisthematrix.warnerbros.com [accessed: July 12, 2004]

[TP] The Place: http://www.theplace.com [accessed: September 30, 2003]

[WP] Wikipedia: http://en.wikipedia.org [accessed: July 12, 2004]

[Ach04] H. Achten, J. Jessurun, and B. de Vries, "The Desk-Cave, A Low-Cost Versatile Virtual Reality Design and Research Setup Between Desktop and CAVE," In *Proceedings of Architecture in the Network Society - 22nd International eCAADe Conference*, Kopenhagen, Denmark, pp. 142–147, September 2004.

[Ban80] R. Banham, *"Theory and Design in the First Machine Age,"* ISBN 0262520583, MIT Press, Cambridge, 1980.

[Ban81] R. Banham, *"Design by Choices,"* ISBN 0847803848, Rizzoli, New York, 1981.

[Ban84] R. Banham, *"The Architecture of the well-tempered Environment (Second Edition),"* ISBN 0851397492, The Architectural Press, London, 1984.

[Bet03] A. Betsky, K. Michael Hays, and L. Anderson, *"Scanning: The Aberrant Architectures of Diller + Scofidio,"* ISBN 0874271312, Whitney Museum of Art, New York, 2003.

[Bri98] D. Brin, *"The Transparent Society: Will Technology Force Us to Choose Between Privacy and Freedom?,"* ISBN 0738201448, Perseus Publishing, Boulder, 1998.

[Bro04] T. Brown, "The Future of Designing Experiences," *Clossing Planary at the ACM CHI 2003 Conference on Human Factors in Computing Systems*, Vienna, Austria, 29. April 2004.

[Bur02] R. Burkhart, *"Kommunikationswissenschaft. Grundlagen und Problemfelder. Umrisse einer interdisziplinären Sozialwissenschaft,"* ISBN 3825222594, UTB für Wissenschaft: Uni-Taschenbücher, Stuttgart, 2002.

[Car03] J. Carranza, Ch. Theobalt, M. A. Magnor, and H.-P. Seidel, "Free-Viewpoint Video of Human Actors," In *SIGGRAPH 2003 Conference Proceedings*, ACM SIGGRAPH Annual Conference Series, San Diego, USA, pp. 569–577, July 2003.

[Chu01] E. F. Churchill (ed.), *"Collaborative Virtual Environments,"* ISBN 1852332441, Springer-Verlag, London, 2001.

[Chun01] Ch. J. Chung (ed.), *"The Harvard Design School Guide to Shopping,"* ISBN 3822860476, Taschen, Köln, 2001.

[Cru92] C. Cruz-Neira, D.J. Sandin, T.A. DeFanti, R. Kenyon, and J.C. Hart, "The CAVE, Audio Visual Experience Automatic Virtual Environment," In *Communications of the ACM*, pp. 64–72, June 2002.

[Cru93] C. Cruz-Neira, D.J. Sandin, and T.A. DeFanti, " Surround-screen projection based virtual reality: The design and implementation of a Cave," In *SIGGRAPH 1993 Conference Proceedings*, ACM SIGGRAPH Annual Conference Series, Anaheim, USA, pp.135–142, August 1993.

[Col00] R. Collins, A. Lipton, T. Kanade, H. Fujiyoshi, D. Duggins, Y. Tsin, D. Tolliver, N. Enomoto, and O. Hasegawa, *"A System for Video Surveillance and Monitoring,"* Robotics Institute, Carnegie Mellon University, Technical Report CMU-RI-TR-00-12, Pittsburgh, USA, 2000.

[Cor86] Le Corbusier, *"Towards a New Architecture,"* ISBN 0486250237, Dover Publications, New York, 1986.

[Dav01] J. Davies, "Planet fashion." *Insight Interactive*, July 2001.

[DeF89] M. L. DeFleur and S. Bale-Rokeach, *"Theories of Mass Communication (5th Edition),"* ISBN 0582998700, Pearson Allyn & Bacon, 1989.

[Der80] M. Dertouzos and J. Moses (ed.), *"The Computer Age: A Twenty-Year View,"* ISBN 0262540363, MIT Press, Cambridge, 1980.

[Der97] M. Dertouzos, *"What WILL be: How the New World of Information Will Change our Lives ,"* ISBN 0062514792, HarperEdge - HarperSanFrancisco, 1997.

[Dit00] D. Ditlea, "The PC goes ready to wear," *IEEE Spectrum*, October, pp. 35–39, 2000.

[Div04] A. Divivier et al., "Virtual try-on: Topics in Realistic, Individualized Dressing in Virtual Reality," In *International Status Conference for Virtual and Augmented Reality*, Leipzig, Germany, February 2004.

[Dro97] P. Droege, *"Intelligent Environments,"* ISBN 0444823328, Elsevier Science B.V., Amsterdam, 1997.

[Duc01] K. Ducatel, M. Bogdanowicz, F. Scapolp, J. Leijten, and J.-C. Burgelman, *"ISTAG - Scenarios for Ambient Intelligence in 2010,"* Final Report, IPTS-Seville, February 2001.

[Ebe04] J. Eberspächer and H.-P. Quadt (Hrsg.), *"Breitband-Perspektiven,"* ISBN 3540221042, Springer Verlag Berlin, Heidelberg, 2004.

[Enc04] J. L. Encarnação, "Ambient Intelligence - Visions and Trends for a Human-Centered, Multimedia and Multimodal Interaction," *Invited Lecture at International Status Conference for Virtual and Augmented Reality*, Leipzig, Germany, 19. February 2004.

[Eve00] D. Eves, J. Green, C. van Heerden, J. Mama, S. Marzano, and L. Traldi, *"New Nomads. An exploration of wearable electronics by Philips,"* ISBN 9064504229, 010 Publishers, Rotterdam, 2000.

[Fau00] W. Faulstich, *"Grundwissen Medien,"* ISBN 3825281698, UTB für Wissenschaft: Uni-Taschenbücher, Stuttgart, 2000.

[Fau03] W. Faulstich, *"Einführung in die Medienwissenschaft,"* ISBN 3825224074, UTB für Wissenschaft: Uni-Taschenbücher, Stuttgart, 2003.

[Fas97] M. Fassler, *"Was ist Kommunikation?,"* ISBN 3825219607, UTB für Wissenschaft: Uni-Taschenbücher, Stuttgart, 1997.

[Fas98] M. Fassler and W. Halbach (Hrsg.), *"Geschichte der Medien,"* ISBN 3825219844, UTB für Wissenschaft: Uni-Taschenbücher, Stuttgart, 1998.

[Flu95] F. Fluckiger, *"Understanding Networked Multimedia: Applications and Technology,"* ISBN 0131909924, Prentice Hall, New Jersey, 1995.

[Flus97] V. Flusser, *"Medienkultur,"* ISBN 3596133866, Fischer Taschenbuch Verlag, Frankfurt/Main, 1997.

[Fre98] E. Frécon and M. Stenius, "DIVE: A scaleable network architecture for distributed virtual environemnts," *Distributed Systems Engineering Journal*, 5:91-100, Special Issue on Distributed Virtual Environments, September 1998.

[Fri04] W. Friedrich (Hrsg.), *"ARVIKA: Augmented Reality für Entwicklung, Produktion und Service,"* ISBN 3895782394 Publicis Corporate Publishing, Erlangen, 2004.

[Fun95] T. A. Funkhouser, "RING: A Client-Server System for Multi-User Virtual Environments," In *Computer Graphics (1995 SIGGRAPH Symposium on Interactive 3D Graphics)*, Monterey, USA, pp. 85–92, April 1995.

[Gar01] M. Garau, M. Slater, S. Bee, and M. A. Sasse, "The impact of eye gaze on communication using humanoid avatars," In *Proceedings of the ACM CHI 2003 Conference on Human Factors in Computing Systems*, Seattle, USA, pp. 309–316, April 2001.

[Gil99] J. Gilmore and B. Pine, *"The Experience Economyl,"* ISBN 0875848192, Harvard Business School Press, Boston, 1999.

[Goe04] G. Goebbels et al., "Development of an Augmented Reality System for intra-operative navigation in maxillo-facial surgery," In *International Status Conference for Virtual and Augmented Reality*, Leipzig, Germany, February 2004.

[Gra98] S. Gray, "Virtual reality in virtual fashion," *IEEE Spectrum*, February 1998.

[Gre95] C. M. Greenhalgh and S. D. Benford, "MASSIVE: A virtual reality system for teleconferencing," *ACM Transactions on Computer Human Interfaces (TOCHI)*, volume 2 (issue 3), pp. 239–261, September 1995.

[Gro03] M. Gross, S. Würmlin, M. Naef, E. Lamboray, C. Spagno, A. Kunz, E. Koller-Meier, T. Svoboda, L. Van Gool, S. Lang, K. Strehlke, A. Vande Moere, and O. Staadt, "blue-c: A spatially immersive display and 3D video portal for telepresence," In *SIGGRAPH 2003 Conference Proceedings*, ACM SIGGRAPH Annual Conference Series, San Diego, USA, pp. 819–827, July 2003.

[Hin02] T. Hine, *"I Want That!: How We All Became Shoppers ,"* ISBN 0060185112, HarperCollins, New York, 2002.

[Hir04] G. Hirzinger, "Robotics and Virtual Worlds," *Invited Lecture at International Status Conference for Virtual and Augmented Reality*, Leipzig, Germany, 19. February 2004.

[Hog00] G. Hogben, "Digital Video and Architectural Design," In *Proceedings 18th International eCAADe Conference 2000*, Weimar, Germany, pp. 201–204, September 2000.

[Hov03] L. Hovestadt, M. Gross, L. Van Gool, M. Meier, S. Würmlin, K. Strehlke, T. Svoboda, S. Lang, A. Kunz, and M. Küchler, *"blue-c-II,"* ETH Zurich - Internal Research Proposal, Zurich, Switzerland, May 2003.

[Hov96] L. Hovestadt, V. Hovestadt, and K. Friedrichs, *"Das Kooperative Gebäude,"* Institute for Industrial Building Production, University Karlsruhe, Research Note, Karsruhe, Germany, November 1996.

[1516] "Standard for Modeling and Simulation (M&S) High Level Architecture (HLA) - Framework and Rules," IEEE Standard 1516, September 2000.

[1993] "IEEE standard for information technology - protocols for distributed simulation applications: Entity information and interaction," IEEE Standard 1278-1993, 1993.

[IST99] "Information Society Technologies project ADVISOR (Advisor Digital Video Storage and On-line Retrieval System)," http://advisor.matrasi-tls.fr [Accessed: 12.July 2004].

[IST00] "Information Society Technologies project MISSY (3D- Information based Safety and Security by Solid State Microsystems)," http://w4.siemens.de/ct/en/technologies/ms/missy/index.html [accessed: May 10, 2003].

[Jan02] M. Janse, *"3 Applications: End-user Scenario's, Storyboards and Specifications: ICE-CREAM project,"* Philips Electronics Nederland B.V., Nat.Lab. Unclassified Report 2002/832, Netherland, 2002.

[Kal02] L. Kalisperis, G. Otto, K. Muramoto, J. Gundrum, R. Masters, and B. Orland, "An Affordable Immersive Environment in Beginning Design Studio Education," In *Proceedings of Thresholds Conference of the Association for Computer Aided Design In Architecture*, Pomona, USA, pp. 49–56, October 2002.

[Kau02] P. Kauff and O. Schreer, "An immersive 3D video-conferencing system using shared virtual team user environments," In *Proceedings of Collaborative Virtual Environments (CVE) 2002*, Bonn, Germany, pp. 105-112, September 2002.

[Kir02] S. Kirchner, "Virtuelle Archäologie," In *International Status Conference for Virtual and Augmented Reality*, Leipzig, Germany, November 2002.

[Kit93] F. Kittler, "Geschichte der Kommunikationsmedien," In *J. Huber, and A. M. Müller: Raum und Verfahren*, ISBN 3878774141, Stroemfeld Verlag, Frankfurt/Main, 1993.

[Klo00] D. Kloock and A. Spahr, *"Medientheorien. Eine Einführung,"* ISBN 3825219860, UTB für Wissenschaft: Uni-Taschenbücher, Stuttgart, 2000.

[Koo01] R. Koolhaas, M. Prada, and P. Bertelli, *"Projects for Prada 1,"* ISBN 8887029180, Fondazione Prada, Milano, 2001.

[Kue04] U. Kühnapfel et al., "HapticIO : Haptic Interface-Systems for Virtual-Reality Training in Minimally-Invasive Surgery," In *International Status*

Conference for Virtual and Augmented Reality, Leipzig, Germany, February 2004.

[Kun97] Th. Kunkelmann and R. Reinema, "A Scalable Security Gateway for Multimedia Communication Tools," In *Proceedings of the International Conference on Multimedia Computing and Systems (ICMCS)*, Ottawa, Canada, pp. 660–662, June 1997.

[Kun97a] Th. Kunkelmann, R. Reinema, R. Steinmetz, and Th. Blecher, "Evaluation of Different Video Encryption Methods for a Secure Multimedia Conferencing Gateway," In *Proceedings of the 4^{th} International COST 237 Workshop*, Lisboa, Portugal, pp. 75–89, December 1997.

[Lal98] V. Lalioti, C. Garcia, and F. Hasenbrink, "Virtual meeting in cyberstage," In *Proceedings of the ACM Symposium on Virtual Reality Software and Technology (VRST) 1998*, Taipei, Taiwan, pp. 205–212, November 1998.

[Lam04] E. Lamboray, *"A Communication Infrastructure for Highly-Immersive Collaborative Virtual Environments,"* PhD Thesis No 15618, ETH Zurich, 2004.

[Lan04] S. Lang, S. Würmlin, J. Borchers, L. Hovestadt, and M. Gross, "blue-c: Using 3D Video for Immersive Telepresence Applications," In*CHI2004 Fringe: CHIPlace Fringe Forum* , Vienna, Austria, http://www.chiplace.org/chifringe/2004/655.pdf, April 2004.

[Lan04a] S. Lang and L. Hovestadt, "Interaction in Architectural Immersive Virtual Environments using 3D Video," In *Proceedings of Architecture in the Network Society - 22^{nd} International eCAADe Conference*, Kopenhagen, Denmark, pp. 74–81, September 2004.

[Lan04b] S. Lang, "Interactive Spaces for Advanced Communication using 3D Video," *IJAC International Journal of Architectural Computing*, Issue 01, vol.02, pp. 109–122, March 2004.

[Lan03a] S. Lang and L. Hovestadt, "An Architectural Framework within a Spatially Immersive Environment for Advanced Communication," In *Proceedings of DIGITAL DESIGN - 21^{st} International eCAADe Conference*, Graz, Austria, pp. 3–10, September 2003.

[Lan03b] S. Lang, M. Naef, M. Gross, and L. Hovestadt, "IN:SHOP - Using Telepresence and Immersive VR for a New Shopping Expirience," In *Proceedings of VISION, MODELING, AND VISUALIZATION 2003*, Munich, Germany, pp. 37–43, November 2003.

[Lan03c] S. Lang, "IN:SHOP: Virtuelle Shopping Räume," *archithese*, vol. 06.2003, pp. 42–45, December 2003.

[Lay03] K. Lay et al., "MORPHA: Interactive and Cooperative Robots - at the frontier to a new generation of robots," In *Proceedings of BMBF Human Computer Interaction Status Conference*, Berlin, Germany, June 2003.

[Lev02] Th. Y. Levin, U. Frohne, and P. Weibel, *"CTRL[SPACE] - Rhetorics of Surveillance from Bentham to Big Brother,"* ISBN 0262621657, MIT Press, Cambridge, 2002.

[Lin95] B. Lingard, *"Human Interfacing Issues of Virtual Reality,"* http://www.cs.wpi.edu/˜matt/courses/cs563/talks/brian1.html, [accessed: June 8, 2004].

[Mac94] M. R. Macedonia, M. J. Zyda, D. R. Pratt, P. T. Barham, and S. Zeswitz, "NPSNET: A Network Software Architecture for Large Scale Virtual Environments," *Presence*, 3(4), 1994.

[Mah00] M. L. Maher, S. Simoff, N. Gu, and K. H. Lau, "Designing Virtual Architecture," In *Proceedings of CAADRIA 2000*, Singapore, pp. 481–490, May 2000.

[McL94] M. McLuhan, *"Understanding Media: The Extensions of Man,"* ISBN 0262631598, MIT Press, 1994.

[McL01] M. McLuhan, Q. Fiore, and J. Agel, *"The Medium is the Massage,"* ISBN 1584230703, Gingko Press Inc., 2001.

[Mat02] T. Matsuyama and T. Takai, "Generation, visualization, and editing of 3D video," In *Proceedings of 1st International Symposium on 3D Data Processing Visualization and Transmission (3DPVT02)*, Padova, Italy, pp. 234-245, July 2002.

[Min80] M. Minsky, "Telepresence," *OMNI magazine*, pp. 45–51, May 1980.

[Mit77] W. Mitchell, *"Computer–Aided Architectural Design,"* ISBN 0442262485, Van Nostrand Reinhold Company Inc., New York, 1977.

[Mit97] W. Mitchell, "Recombinant Architecture," In *Peter Dröge: Intelligent Environmetns*, ISBN 0444823328, Elsevier Science B.V., Amsterdam, pp. 551–582, 1997.

[Mit03] W. Mitchell, *"Me++: The Cyborg Self and the Networked City,"* ISBN 0262134349, MIT Press, Cambridge, 2003.

[Moe04] A. Vande Moere, *"Infoticles: Immersive Information Visualization using Data-Driven Particles,"* PhD Thesis No. 15388, ETH Zurich, 2004.

[Moo65] G. Moore, "Cramming more components onto integrated circuits," *electronics Magazine*, Volume 38, Number 8, 1965.

[Mur98] J. H. Murray, *"Hamlet on the Holodeck,"* ISBN 0262631873, MIT Press, Cambridge, 1998.

[Nae04] M. Naef, *"blue-c API: Software Support for Collaborative Virtual Reality,"* PhD Thesis No. 15632, ETH Zurich, 2004.

[Nae03] M. Naef, E. Lamboray, O. Staadt, and M. Gross, "The blue-c distributed scene graph," In *Proceedings of the IPT/EGVE Workshop 2003*, Zurich, Switzerland, pp. 125–133, May 2003.

[Nae02] M. Naef, O. Staadt, and M. Gross, "Spatialized audio rendering for immersive virtual environments," In *Proceedings of the ACM Symposium on Virtual Reality Software and Technology (VRST) 2002*, Hong Kong, China, pp. 65–72, November 2002.

[Neg97] N. Negroponte, *"Total Digital,"* ISBN 3442127211, Goldmann Verlag, München, 1997.

[Nig93] L. Nigay and J. Contaz, "A design space for multimodal systems: Concurrent processing and data fusion," In *Proceedings of INTERCHI 1993*, Amsterdam, The Netherlands, pp. 172–178, April 1993.

[Nij04] A. Nijholt, Th. Rist, and K. Tuijnenbreijer, "Lost in Ambient Intelligence," In *Proceedings of the ACM CHI 2004 Conference on Human Factors in Computing Systems*, Vienna, Austria, pp. 1725–1726, April 2004.

[Nor02] D. A. Norman, *"The Design of Everyday Things,"* ISBN 0465067107, Perseus Publishing, Boulder, 2002.

[Phy03] *"Lexikon der Physik,"* ISBN 3827414628, Spektrum Akademischer Verlag, Heidelberg, 2003.

[Pos92] N. Postman, "Sieben Thesen zur Medientechnologie," In *W. Fröhlich, R. Zitzlsperger, and B. Franzmann: Die verstellte Welt. Beiträge zur Medienökologie.*, ISBN 3407349017, Beltz, Weinheim, pp. 9–23, 1992.

[Pri94] K. Primentel and K. Teixeira, *"Virtual Reality: Through the New Looking Glass,"* ISBN 0070501688, 2nd edition, McGraw Hill Osborne Media, Emeryville, 1994.

[Pro76] H. Pross and H. Beth, *"Einführung in die Kommunikationswissenschaft,"* ISBN 3170025473, Kohlhammer Verlag, Stuttgart, 1976.

[Ras98] R. Raskar, G. Welch, M. Cutts, A. Lake, L. Stesin, and H. Fuchs, "The office of the future: A unified approach to image-based modeling and spatially immersive displays," In *SIGGRAPH 1998 Conference Proceedings*, ACM SIGGRAPH Annual Conference Series, Orlando, USA, pp. 179–188, July 1998.

[Rie04] R. Riener, R. Burgkart, R. Merzendorfer, M. Frey, T. Pröll, F. Regenfelder, M. Ulmer, and E. Lusczak, "VOR - Virtual Orthopedic Reality: Scientific Results and Commercial Realisation," In *International Status Conference for Virtual and Augmented Reality*, Leipzig, Germany, February 2004.

[Roh94] J. Rohlf and J. Helman, "IRIS Performer: A high performance multiprocessing toolkit for real-time 3d graphics," In *SIGGRAPH 1994 Conference Proceedings*, ACM SIGGRAPH Annual Conference Series, Orlando, USA, pp. 381–395, July 1994.

[Row03] J. K. Rowling, *"Harry Potter and the Order of the Phoenix ,"* ISBN 043935806X, Scholastic, New York, 2003.

[Sad01] A. Sadagic, H. Towles, J. Lanier, H. Fuchs, A. van Dam, K. Daniilidis, J. Mulligan, L. Holden, and B. Zeleznik, "National tele-immersion initiative: Towards compelling tele-immersive collaborative environments," *Presentation given at Medicine meets Virtual Reality 2001 Conference*, Newport Beach, USA, January 2001.

[Schi04] A. Schilling and D. Holweg, "3D GIS for Virtual City Models," *Computer Graphik*, Volume 16, Number 1, 2004.

[Schm] D. Schmidt, "Real-time CORBA with TAO (The ACE ORB)," http://www.cs.wustl.edu/schmidt/TAO.html.

[Schm99] G. Schmitt, *"Information Architecture,"* ISBN 3764360925, Birkenhäuser, Basel, 1999.

[Schn04] M. Schnaider, S. Röddiger, H. Seibert, and B. Schwald, "Implementation and Evaluation of an Augmented Reality System Supporting Minimal Invasive Interventions," In *International Status Conference for Virtual and Augmented Reality*, Leipzig, Germany, February 2004.

[Sin99] S. Singhal and M. Zyda, "Networked Virtual Environments: Design and Implementation," *ACM Press - SIGGRAPH Series*, Addison-Wesley, 1999.

[Spa03] C. Spagno, *"Immersive (VR) Projection System with Simultaneous Image Acquisition Using Active Projection Screens,"* PhD Thesis No. 15273, ETH Zurich, 2003.

[Spa03a] C. Spagno and A. Kunz, "Construction of a three-sided immersive telecollaboration system," In *Proceedings of the IEEE Virtual Reality Conference 2003 (VR 2003)*, Los Angeles, USA, pp. 37–44, March 2003.

[Spe03] "What's Next for Home Networking," *IEEE Spectrum*, September 2003.

[SPM03] "Universal Multimedia Access," *IEEE Signal Processing Magazine*, vol.20, No.2, March 2003.

[SPM03a] "Universal Multimedia access," *IEEE Signal Processing Magazine*, vol.20, No.2, p.65 Figure1, March 2003.

[SPM03b] "Super-Resolution Image Reconstruction," *IEEE Signal Processing Magazine*, vol.20, No.3, May 2003.

[SPM01] "Immersive Interaction Technologies," *IEEE Signal Processing Magazine*, vol.18, No.3, May 2001.

[Ste00] R. Steinmetz, *"Multimedia-Technologie,"* ISBN 3540673326, Springer-Verlag, Berlin, 2000.

[Stei03] E. Steinbach, "Course Notes: Media Technology ," *Media Technology Group, Technical University of Munich*, WS 03/04, 2003.

[Steu92] J. Steuer, "Defining Virtual Reality: Dimensions Determining Telepresence," *Journal of Communication*, Autumn, pp. 73–93, 1992.

[Tra99] H. Tramberend, "Avocado: A distributed virtual reality framework," In *Proceedings of the IEEE Virtual Reality Conference 1999 (VR 1999)*, Houston, USA, pp. 14–21, March 1999.

[Tre67] D. -G. Treichler, *"Are you missing the boat in training aid? Film and A-V Communication,"*, United Business Publications, New York, 1967.

[Tro89] "TRON Project bee monthly," Tron Association, No. 9, Tokyo, May / June 1989.

[Val03] M. Valera Espina and S.A. Velastin, "An approach for designing a real-time intelligent distributed surveillance system," In *Symposium on Intelligent Distributed Surveillance Systems*, London, Great Britan, pp. 6/1–6/5, February 2003.

[Wei91] M. Weiser, "The Computer for the 21st Century," *Scientific American 265 (3)*, pp. 94–101, 1991.

[Wic01] R. Wichert, M. Finke, and D. Balfanz, "An Interactive Video System for Live Coverage Transmission using Real Time Video Hyperlinks," In *Proceedings of International Conference on Multimedia, Internet, Video Technologies 2001 (MIV'01)*, Malta, 2001.

[Wue04] S. Würmlin, *"Dynamic Point Samples as Primitives for Free-Viewpoint Video,"* PhD Thesis No 15643, ETH Zurich, 2004.

[Wue04a] S. Würmlin, E. Lamboray, and M. Gross, "3D Video Fragments: Dynamic Point Samples for Real-time Free-Viewpoint Video," *Computers & Graphics 28 (1)*, pp. 3–14, 2004.

[Wue03] S. Würmlin, E. Lamboray, O. G. Staadt, and M. H. Gross, "3D Video Recorder: A System for Recording and Playing Free-Viewpoint Video," *Computer Graphics Forum 22 (2)*, pp. 181–193, 2003.

[Ver03] R. Vertegaal, I. Weevers, Ch. Sohn, and Ch. Cheung, "GAZE-2: Conveying eye contact in group video conferencing using eye-controlled camera direction," In *Proceedings of CHI 2003*, Fort Lauderdale, USA, pp. 521–528, April 2003.

[Zwi02] M. Zwicker, M. Pauly, O. Knoll, and M. Gross, "Pointshop 3D: An interactive system for point-based surface editing," In *SIGGRAPH 2002 Conference Proceedings*, ACM SIGGRAPH Annual Conference Series, San Antonio, USA, pp. 322–329, July 2002.

[Zwi01] M. Zwicker, H. Pfister, J. VanBaar, and M. Gross, "Surface splatting," In *SIGGRAPH 2001 Conference Proceedings*, ACM SIGGRAPH Annual Conference Series, Los Angeles, USA, pp. 371–378, July 2001.

Copyrights

Image Credits

Figure 3.1 ...... Picture courtesy of Intel Research
Figure 3.2 ...... [Enc04]
Figure 3.3 ...... [SPM03a]
Figure 3.4 ...... http://www.evl.uic.edu
Figure 4.2 ...... [Bet03]
Figure 4.3 ...... Picture courtesy of Michael Moran
Figure 4.4 ...... Picture courtesy of Art+Com, Prof. Joachim Sauter, 2004
Figure 4.5a ...... http://www.blinkenlights.de
Figure 4.5b ...... http://www.bix.at
Figure 4.6 ...... http://cox.plexgroup.com
Figure 4.7 ...... Picture courtesy of M. Lorenz
Figure 4.8a ...... http://gears.aset.psu.edu/viz/projects/vr/iel
Figure 4.8b ...... http://www.ds.arch.tue.nl/Education/DSLab
Figure 4.9a ...... http://www.adit.co.uk/html/noughts_and_crosses.html
Figure 4.9b ...... Picture courtesy of Cyan Worlds, Inc., Myst
Figure 4.10a ...... Picture courtesy of id Software, Doom
Figure 4.10b ...... Picture courtesy of Rockstar Games, Max Payne2
Figure 4.11 ...... http://www.kinokunst.de, http://www.starwars.com
Figure 4.12 ...... http://whatisthematrix.warnerbros.com
Figure 4.13 ...... http://en.wikipedia.org, http://www.ctheory.net
Figure 6.1 ...... Picture courtesy of OMA/AMO, Markus Schäfer, 2001
Figure 6.2 ...... http://www.bogner.com

Curriculum Vitae

Personal Data

Name:	Silke Berit Lang
Date of Birth:	June 15, 1975
Place of Birth:	Munich, Germany

Education

11/2001 – 10/2004	Doctoral Program at the Swiss Federal Institute of Technology (ETH), Zurich, Switzerland.
10/2000 – 10/2001	Postgraduate Studies "Computer Aided Architectural Design" with main focus on Visual Communication, Swiss Federal Institute of Technology (ETH), Zurich, Switzerland.
10/1999 – 02/2000	Scholarship , Department of Architecture, Swiss Federal Institute of Technology (ETH), Zurich, Switzerland.
10/1994 – 07/2000	Studies of Architecture, Technical University of Munich (TUM), Munich, Germany.

Professional Data

current	Research Assistant at the Computer Graphics Laboratory, Institute of Computational Science, Swiss Federal Institute of Technology (ETH), Zurich, Switzerland. Cooperation with Visual Computing & Virtual Reality Lab., Computer Science and Engineering Department, Ewha Womans University, Seoul, Korea.
11/2001 – 10/2004	Teaching Assistant at the Chair of CAAD, Institute of Building Technology, and Research Assistant at the interdisciplinary blue-c project, Swiss Federal Institute of Technology (ETH), Zurich, Switzerland.
09/1998 – 09/1999	Student Associate at the Office of Architecture and Urban Planning Peter D. Bergler, Munich, Germany.
04/1998 – 08/1998	Student Teaching Assistant at the Chair of Building Realization + Computer Science, Faculty of Architecture, Technical University of Munich (TUM), Munich, Germany.