Chapter 1
Introduction to Virtual and Augmented Reality

Ralf Doerner, Wolfgang Broll, Bernhard Jung, Paul Grimm, Martin Göbel, and Rolf Kruse

Abstract What is Virtual Reality (VR)? What is Augmented Reality (AR)? What is the purpose of VR/AR? What are the basic concepts? What are the hard- and software components of VR/AR systems? How has VR/AR developed historically? The first chapter examines these questions and provides an introduction to this textbook. This chapter is fundamental for the whole book. All subsequent chapters build on it and do not depend directly on one another. Therefore, these chapters can be worked through selectively and in a sequence that suits the individual interests and needs of the readers. Corresponding tips on how this book can be used efficiently by different target groups (students, teachers, users, technology enthusiasts) are provided at the end of the chapter, as well as a summary, questions for reviewing what has been learned, recommendations for further reading, and the references used in the chapter.

1.1 What Is VR/AR About?

Let us first look at the ideal conception of a *Virtual Reality* (*VR*): What is a perfect VR? In this extreme case the underlying ideas of VR become particularly clear. Then we will look at why perfect VR cannot be achieved today (and why one would not want to achieve it, e.g., for ethical reasons) and show how a *virtual environment* can still be created. We introduce the concept of *Augmented Reality* (*AR*). Finally, we motivate what VR and AR can be used for today and why these topics are being dealt with intensively.

Dedicated website for additional material: vr-ar-book.org

R. Doerner (✉)
Department of Design, Computer Science, Media, RheinMain University of Applied Sciences, Wiesbaden, Germany
e-mail: ralf.doerner@hs-rm.de

R. Doerner et al. (eds.), *Virtual and Augmented Reality (VR/AR)*,
https://doi.org/10.1007/978-3-030-79062-2_1

1.1.1 The Perfect Virtual Reality

Humans perceive the world through sensory impressions. If, for example, light is reflected by a real object, such as a tiger, and enters a person's eye, photochemical processes are triggered in special sensory cells located in the retina. The light acts as a stimulus for these sensory cells. The light stimuli set off nerve impulses, which are modified via nerve cells that are connected in a complex way. These signals are then transmitted throughout the brain and processed further. Various areas of the brain that contribute to visual perception have already been identified. The perceived image is not created in the eyes, but rather in brain regions, mainly in the back of the head. The processes in the brain can be divided into several stages. At first, fast parallel processing of the visual sensory impressions takes place during which, for example, the yellow and black areas and also the pattern on the fur of the tiger are identified. Based on this, slower sequential processing follows, e.g., the composition of the colored surfaces to objects (as for example a paw or the teeth of the tiger) with the support of the person's memory. If the human being has already seen a tiger before, this can lead to recognition. We call the whole apparatus, from the sensory cells, via the visual nerves to the visual centers in the brain, the *visual system* of the human being. So, in our example, the human being sees the tiger thanks to the visual system and can draw conclusions about reality from this, e.g., that a real predatory cat is standing nearby and it would be a perfectly suitable time to start running away.

The connection between reality and what people perceive about it through their visual system is anything but simple. The same reality can cause different perceptions in different people. A wall that reflects light with a wavelength of 630 nm triggers the color perception "red" in many people – but some people have a different perception. Because they are in the minority, these people are called colorblind – after all, about 9% of men and 1% of women perceive colors differently than the rest of the population. Color, a term people use to describe visual perception, is therefore not a term that objectively describes reality. Color is not a physical property of the real wall but rather stands for a subjective perception that is indirectly triggered in people by the wall through reflected light.

Even in a single individual there is no simple connection between reality and visual perception of reality. If you look at Fig. 1.1, you can see black squares arranged on a grid. At the intersections of the grid, one can see alternating, partly flickering dark and bright points. But this does not correspond to the properties of the grid points in reality. All grid points are identical and always reflect the light in the same way (if this text is being read with an e-book reader, be assured that there is no trickery here). A number of such phenomena have been described in perceptual psychology, showing how the visual system combines, amplifies, filters out or recombines responses to external stimuli originating from the sensory cells during the complex process of perception. The same stimuli can lead to different perceptions in the same individual at different times, for example depending on whether the individual is concentrating on something or not – or whether the individual has

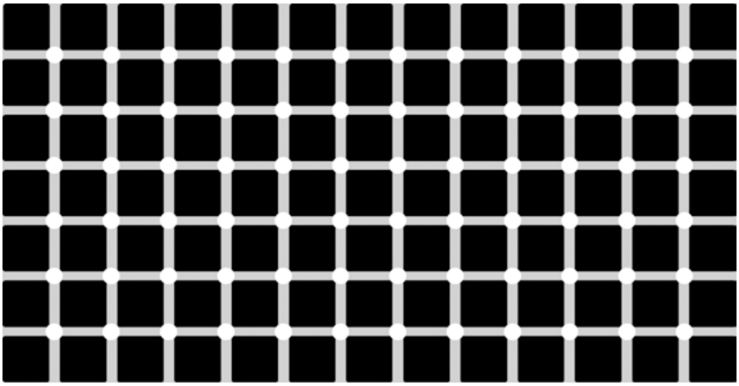

Fig. 1.1 A Hermann grid. Although in reality all grid intersections always reflect light to the same extent, a person sometimes perceives dark spots there. The dark spots disappear as soon as you try to look at them directly

just had a glass of vodka or not. A remarkable characteristic of the visual system is that it can also change its mode of operation over time, adapting itself. The psychologist George M. Stratton made this clear in an impressive self-experiment at the end of the nineteenth century. Stratton wore reversing glasses for several days, which literally turned the world upside down for him. In the beginning this caused him great difficulties: Just putting food in his mouth with a fork was a challenge for him. With time, however, his visual system adapted to the new stimuli from reality and he was able to act normally in his environment again, even seeing it upright when he concentrated. As he took off his reversing glasses, he was again confronted with problems: He used the wrong hand when he wanted to reach for something, for example. Fortunately for Mr. Stratton, an adaptation of perception is reversible, and he did not have to wear reversing glasses for the rest of his life. For him, everything returned to normal after one day.

We can conclude that there is no fixed, unambiguous and objective connection between (1) reality with the light stimuli it exerts on a human being and (2) the visual perception by the human being of this reality. This creates some leeway for manipulating the human visual perception of reality. A simple way is to replace a stimulus emanating from a real object with a similar, artificial stimulus. If the human visual system, stimulated by this artificial stimulus, comes to a similar perception as it would have done with a real object, the human being may even be under the mistaken impression that this object actually exists in reality. Images are a typical example of this approach. If one wishes to cause the visual perception "tiger" in a human being, then one does not need to inconvenience a real predatory cat. One can show the person a photograph of a tiger. Of course, this photograph of a tiger – a sheet of paper printed with pigments reflecting light in a certain way – is a fundamentally different object than a flesh and blood tiger. But both have something in common: They reflect light in a similar way, stimulate the visual system in a similar way and evoke similar visual perceptions in the human being.

Typically, a person will not be deceived so simply. People are usually able to distinguish a real tiger from a photo of a tiger. Therefore, let us assume that we could bring the light stimuli that emanate from a real tiger perfectly into the visual system of a human being, e.g., by playing in the impulses of sensory cells resulting from outside stimuli via a "socket" implanted into the brain. Let us go a step further in our thoughts and not limit ourselves to visual perception alone. Visual perception is the most important source of information about a person's environment – more than 130 million sensory cells (about 70% of all human sensory cells) and more than four billion neurons, i.e., more than about 40% of the cerebral cortex, are involved in seeing. "Man is an eye animal" as Leonardo da Vinci put it. However, the human perception of reality is also based on other sensory impressions. For example, in addition to the cone cells in the retina that react to light, there are special sensory cells, such as Merkel cells, which respond to pressure, or the Pacinian corpuscles, which are stimulated by acceleration. Therefore, let us further assume that we could also transfer the reaction of all these other sensory cells directly to the brain via the imaginary "socket". Besides seeing (visual perception) we would thus also manipulate

- hearing (auditory perception),
- smelling (olfactory perception),
- tasting (gustatory perception),
- feeling (haptic perception),
- and, as part of feeling, touch (tactile perception),
- sense of balance (vestibular perception),
- body sensation (proprioception),
- the sensation of temperature (thermoception),
- and the sensation of pain (nociception).

Would we then be in a position to have the stimuli emanating from a tiger calculated by a computer and played into the brain of a person in such a way that this person would be convinced that there was a real tiger nearby? Would we be able to put a human being into an apparent reality, a virtual reality, that the human being could no longer distinguish from the "real" reality? Can we create a perfect illusion of reality?

These are fascinating questions that the Wachowskis, for example, have vividly dealt with in their film *The Matrix* and its sequels. Other films, such as *Vanilla Sky* and science fiction novels by Stanislaw Lem, for example, also address this question. It also touches on philosophical questions such as those raised by Plato over 2400 years ago with his allegory of the cave. Plato wondered how people would react who had been trapped in a cave since childhood with their heads fixed in such a way that they never see objects behind them but only perceive the objects' shadows cast on the cave wall visible to them. According to Plato's Theory of Ideas, we do not directly recognize reality – the true being – but are only able to perceive indirectly "shadows", images of reality in our "cave", our world limited by the realm of sensual experiences. Similar ideas can also be found, for example, in Indian mythology. Here, Maya, the goddess of illusion, prevents people from directly recognizing reality. Instead, Maya makes us experience only a projection of the world created by ourselves and our perception.

The French philosopher René Descartes went a step further. He stated that our perception of reality might not only be an imperfect image but a complete illusion and that all knowledge about reality is to be doubted. Descartes introduces the figure of the Genius Malignus, an evil spirit, who makes people believe in a reality that does not exist. So, you are not reading a book, but an evil spirit makes you believe that you have eyes and can read a book that does not exist in reality. In fact, the spirit is even so evil that it is a textbook about Virtual Reality.

The philosophical direction of skepticism doubts that there is such a thing as reality or such a thing as fundamental truths at all. With the "Brain in a Vat" experiment, a thought experiment similar to our considerations, the followers of skepticism justify their position. In this experiment, it is assumed that a brain extracted from a human being floating in a vat of nutrient solution is supplied by a computer with impulses that simulate an apparent reality. They answer our question of whether the consciousness in this brain can distinguish the faked reality from real reality, namely the disembodied brain floating in a tub, with a firm "No". Therefore, the argument goes, we can never be sure whether we are in a Virtual Reality – just as most people in the feature film *The Matrix* never realize what their actual reality looks like.

1.1.2 The Simulation of the World

In order to realize a perfect Virtual Reality, at least to some extent, sensory stimuli must be generated that make a person perceive this alternative world. In the first flight simulators, a video camera was attached to a linkage and moved over a physical landscape model similar to a model railway. The images captured by the camera were displayed to the pilot in the flight simulator, who could thus perceive an image of the world when looking out of the cockpit. A more modern approach would be to use computer graphics to generate images or light stimuli for Virtual Reality.

But the generation of the stimuli is only one task on the way to the perfect Virtual Reality. People not only want to see and feel the world but also to act in it. For example, if a person perceives a ball in Virtual Reality, he or she might want to be able to kick the ball and run after it. This requires that the virtual world is simulated, that the actions of the person are known to the simulation, and that these actions can influence the simulation. The results of the simulation in turn have an effect on the generation of the stimuli – if a person moves in Virtual Reality, the generation of stimuli must also take the new position into account. The task of simulation can be performed by a computer system that must have a simulation model of the world at its disposal. The simulation model of the world determines the behavior of the Virtual Reality. Consequently, the reactions of the virtual world in response to the actions of users must be simulated, as well as changes in the virtual world that do not depend on human actions. For example, a day-night cycle in the virtual world could be simulated that cannot be influenced by people.

One can strive to build the simulation model of the world in such a way that the behavior of the virtual world corresponds as closely as possible to that of reality. If a person kicks a virtual ball, the world simulation would move the ball according to the well-known laws of physics – the ball would have a virtual mass and a virtual frictional resistance, and would continue to roll on sloping virtual terrain until it reached a rest position. In Virtual Reality, however, one is not bound by the laws of reality. A kick against a virtual ball, for example, could also cause the ball to move along a serpentine path – or to turn it into a chicken. In this way you can create fantastic virtual worlds, virtual worlds that play in an imaginary future, or virtual worlds that recreate past times.

Being tasked with the recognition of human actions, the simulation of the virtual world, and the generation of stimuli for humans, the VR system can become highly complex. The simulation of a single virtual human being – which includes the generation of realistic images of skin and clothing, speech synthesis, and the simulation of human behavior, emotions, irony and willpower – is a major challenge today. The challenge is further increased by the requirement that this computer system must operate in real time, i.e., it has to keep pace with human beings. This implies that calculations must not take up arbitrary time but must adhere to strict time constraints. For example, a large number of images for Virtual Reality must be generated per second so that the human observer perceives movements in the virtual world as continuous and natural. The required number of images per second depends on the viewers and their current situation – typically 60 images per second are needed to meet the demand for real time (if the viewers have large amounts of alcohol in their blood, however, four images per second may be sufficient). This means that the computer system may not take more than 16 ms to generate images. Real-time requirements are even more demanding for haptic feedback. Typically, the VR system must generate haptic stimuli 1000 times per second in order to create a convincing sensation of touch.

We call a *VR system* a computer system consisting of suitable hardware and software to implement the concept of Virtual Reality. We call the content represented by the VR system a *virtual world*. The virtual world includes, for example, models of objects, their behavioral description for the simulation model and their arrangement in space. If a virtual world is presented with a VR system, we speak of a *virtual environment* for one or more users.

1.1.3 Suspension of Disbelief

The Matrix in the feature film of the same name and the Holodeck in the television series *Star Trek* both transport a person into Virtual Reality. There is one crucial difference: In the Matrix, people do not know that they are in Virtual Reality at all. On the contrary, people enter the Holodeck on the starship *Enterprise* consciously. They go through a door into the virtual environment and know that it is a simulation, but in reality, they are still in a large hall. Nevertheless, people seem to perceive the Holodeck as very real. Does it not bother you to know that you are in Virtual Reality? Can the illusion of a virtual world be achieved at all if you are aware of being in Virtual Reality?

Let us consider the following experiment. We put a helmet on a person, in which two small monitors, one for each eye, are attached. The person can no longer perceive the environment visually, but only the images in the monitors, which are fed in from outside. A sensor is built into the helmet which can determine how the person is turning their head and where the person is located. This information is used to adjust the generated images to the current head position: If the person looks up, images from the sky are shown; if the person tilts their head downwards, then he or she sees the ground; and if the person takes a step forward, then images from this new position are shown. We use a computer to create images of the roof of a virtual skyscraper and want to give the impression that the person is standing at a dizzy height on the edge of a huge building. If you observe people in this situation, you often see that they move forward very slowly and carefully. The closer they get to the edge of the building, the faster their pulse and breathing become. Their hands get wet. These are typical fear reactions that are caused by a danger such as an abyss in reality. The people are always aware that the building is only virtual, that in reality there is no abyss at all, and that they are standing safely in a room. Nevertheless, they succumb to the illusion of Virtual Reality and react to it as if it were the real world.

In certain situations, people possess the ability to blank out the obvious contradiction between a fictitious world and reality. Besides, people want to do this. The philosopher Samuel T. Coleridge coined the expression "*willing suspension of disbelief*". For entertainment purposes, people are prepared to accept the figure of Scrooge McDuck and his virtual world Duckburg as existing, even if it is known that this character consists only of hand-drawn lines and that in reality older drakes do not bathe in money. In dubbed films, one fades out the fact that James Bond as an English agent obviously does not always speak perfect Japanese or German. However, this "suspension of disbelief" is not easy to describe and is sometimes selective. Cartoonist Gary Larson describes the indignation of his readers about the fact that in one of his cartoons a polar bear is surrounded by penguins. Readers criticized that this is impossible since polar bears live at the North Pole, but penguins live at the South Pole. However, at the same time readers are not in the least bothered by the fact that the penguins in the cartoon talk to each other and the polar bear has disguised himself as a penguin.

For the creation of Virtual Reality, this human characteristic of blanking out disbelief means that one does not have to resort to drastic measures. Fortunately, there is no need to drill holes in the top of someone's skull and directly manipulate the brain in order to put people into a virtual environment in which they feel present. In this way, Virtual Realities can be created at different stages of technological advancement, where the ultimate stage would allow the creation of the perfect Virtual Reality discussed above. In fact, highly believable virtual environments can already be created today with relatively little effort.

1.1.4 Motivation

What is the point of all this? Why would you want to build a virtual environment at all and put people into it? What are the advantages of dealing with Virtual Reality? There are many answers to these questions. We will consider some of them in the following.

If the world simulation is performed by a computer, then Virtual Reality is the interface between the computer system and the human being. Under this perspective, every Virtual Reality implements a human–machine interface. This interface can be characterized as being particularly natural and intuitive. For example, instead of using a mouse and keyboard, the use of a steering wheel and foot pedals for a car racing game is a step towards Virtual Reality that makes the operation of the virtual car and its navigation through the virtual world more natural. A perfect Virtual Reality can then be understood as a perfect user interface for software. Users can simply act as they are used to doing in the world. Ideally, they are completely unaware of the fact that they are interacting with a computer program. In this respect, the engagement with Virtual Reality can be understood as a methodical approach to finding new forms of human–computer interaction by working towards a vision of a perfect Virtual Reality. Even though this vision may never be achieved (or one may not even want to achieve this, because extensive manipulation of humans is ethically questionable), valuable new ideas can emerge along the way and innovative user interfaces can be designed to make it easier for humans to handle computer systems.

By exploiting its sophisticated visualization capabilities, Virtual Reality can also make it easier for people to absorb and understand data. For example, through years of study and experience, architects have acquired the ability to imagine a building in their minds by looking at 2D construction plans – many real-estate investors do not have this ability. Virtual Reality can also visualize the data in the construction plans for clients in such a way that they can get a good impression of the building and make more informed decisions regarding alternative design choices. Complex results of computer simulations, e.g., the calculation of how air would flow around a newly planned vehicle, can be visualized directly on a virtual vehicle. Engineers and designers can work together in the virtual world to develop aesthetically pleasing body shapes that avoid air turbulence and reduce the vehicle's air resistance. Even completely abstract data can be displayed in Virtual Reality. In this way, an analyst can be transported to a virtual world of financial data.

Virtual realities offer researchers tools to find out more about human perception. For example, experiments can be conducted in Virtual Reality that help to gain insight into how people orient themselves in three-dimensional space. In addition to gaining knowledge in science, Virtual Reality can also offer a very practical use with tangible financial benefits, as case studies show, e.g., on the use of VR in construction (see Chap. 9).

Hardly any car is built today without using methods from Virtual Reality. For example, designs can be visualized more realistically, and prototypes can be created more cost-efficiently than in traditional model-making (see Chap. 9). How the robots in production lines of automobiles are adjusted to a new car model can be simulated in a virtual world and presented in Virtual Reality before the start of production. The analysis of the planning and the elimination of planning errors in a virtual plant or a virtual factory is much easier and more cost-efficient than performing it in the real world.

Pilots take advantage of Virtual Reality during their training in flight simulators. By not using a real aircraft, the airline saves money. But training in Virtual Reality does not only have financial advantages. As no kerosene is burned, less CO_2 is emitted, which benefits the environment. In comparison to a real aircraft, the pilots can rehearse extreme situations without danger. In addition to flight simulators, simulators of ships, trams, trains, and trucks are also commonly used. German air traffic control even operates a virtual airport where air traffic controllers can train. Another example is the training of personnel for complex systems, such as operating the control center of a coal-fired power plant or maintaining aircraft. Virtual Reality allows training to take place even before the real object is completed, so that well-trained personnel are already available at the time of commissioning. In addition to training in the civilian sector, Virtual Reality also has application potential in the military. For example, crews of fighter jets or tanks are trained in virtual environments.

Interested people can buy tickets to an attraction that allow them to drive a high-speed train through a virtual landscape sitting in a highly realistic mock-up of a locomotive. This is an example of how Virtual Reality is used for entertainment purposes in simulation games. Other game genres also benefit from the use of Virtual Reality, so players can experience adventures in fantastic worlds in adventure games. Very close to reality, tourists can experience historical cities such as ancient Rome by visiting them in Virtual Reality. Museums can offer engaging sensual experiences in virtual environments. Artists use Virtual Reality for installations. Virtual Reality arouses interest and can serve as an eye-catcher – accordingly, it offers potential for marketing, for example at trade show booths.

In medicine, there are possible applications in the field of training. Doctors can practice and plan operations in Virtual Reality without any risk for their patients. Nursing staff can train in the handling of patients. Virtual Reality can even be used for treatment. As already described, people can be positioned at a virtual abyss. In this way, people with a fear of heights can be confronted with critical situations and their phobia can be treated. In Virtual Reality, the factors that cause fear can be safely controlled and dosed during treatment.

The range of possible applications of Virtual Reality can be significantly expanded by trying not to completely cut people off from reality when placing them in a virtual environment. Instead, one can try to integrate parts of a virtual world into reality. Let us look again at the example already described where we have placed a person on the edge of a virtual abyss. Would it not be more effective if we did not have to put a helmet on the person and instead could place him or her on a large glass plate? An image from the virtual world would be projected onto this glass plate from below instead of showing it on the small monitors in the helmet. If the person looks down, he or she can see not only the virtual edge of the building but also their own real feet. So the person still perceives reality, but additionally, at some points, integrated parts from a virtual world that fit into reality. The idea of augmenting images from reality in real time by exactly fitting virtual partial images opens up a whole field of new application possibilities for VR technologies. Another example is the use of special binoculars, similar to the well-known coin-operated binoculars, that are permanently installed at viewing points. When looking through the binoculars, the user sees not only reality but also parts of a virtual world that are displayed according to the area of reality currently being viewed. For example, if the observer is looking at the tower remains of an old castle ruin, the binoculars can display a virtual tower at exactly this point, just as it might have appeared several centuries ago. In this case, one no longer speaks of Virtual Reality (VR) but of *Augmented Reality* (*AR*). The virtual and real portions of an image can vary. In fact, there is a smooth transition. One speaks of AR when the real parts predominate. In Sect. 1.2, we look at VR in more detail, while AR is the subject of Sect. 1.3.

So, there are many reasons to learn more about the theoretical foundations of VR and AR as well as the practice of creating convincing virtual and augmented worlds. If one embarks on this endeavor, one is confronted with many questions. What do you have to consider if you want to put people into a virtual world? What makes it believable? What is conducive to achieving suspension of disbelief – and what can destroy it? What effort must be made in a particular application area to achieve this? How is the transmission of different stimuli from a VR technically realized? Which devices are there to make it easier for a person to immerse him or herself in Virtual Reality? How is a computer system structured that generates the corresponding stimuli, e.g., generates images from a VR close to reality? What is the system architecture of a VR system? Which interfaces are there, which norms, and which standards? How do you build simulation models for the world simulation of VR? How does the simulation get information about the actions of people? How can people move in a virtual world? Which algorithms are used in VR? What is the runtime of these algorithms? How can the VR system meet real-time requirements? When looking at AR in comparison to VR, additional questions arise: Which technology is used to include parts of a virtual world into reality? What is the relationship between virtual and real objects? Can they occlude each other? How is a virtual object illuminated with a real light source? How does a virtual object cast a shadow on a real object? How can a virtual object be placed on top of a real object?

In science, but also in practical implementation, many people have already dealt with such questions and contributed to finding answers. In this textbook, basic scientific knowledge in the field of VR and AR is compiled and its practical application is illustrated with case studies. The knowledge conveyed in the book is a solid foundation for all those who want to use VR and AR practically, but also for those who want to actively contribute to the vision of a perfect Virtual Reality through research and development in the field.

1.2 What Is VR?

As should be clear from the introductory remarks above, one can approach the field of VR in very different ways. At the visionary end of the spectrum, e.g., in science fiction movies and popular culture, "perfect VR" is presented as a comprehensive simulation which is no longer distinguishable from human reality. At the practical end of the spectrum, VR has long been established as a tool for product development in many industrial sectors. In this section, we examine how the scientific and technological field of VR is characterized by the members of the research community.

VR is a relatively young field of science and its development is strongly driven by rapid advances in the underlying hardware. In view of this, it may come as no surprise that the scientific discipline of VR has so far not produced a uniform definition of "Virtual Reality". Nevertheless, there is very broad agreement on the essential or desirable features of VR. The following characterizations of VR take different perspectives to differentiate VR systems from traditional human–computer interfaces: the focus on technological aspects, the classification of VR as a new form of human–computer interaction, and the emphasis on the mental experience of VR.

1.2.1 Technology-Centered Characterizations of VR

> "The ultimate display would, of course, be a room within which the computer can control the existence of matter. A chair displayed in such a room would be good enough to sit in. Handcuffs displayed in such a room would be confining, and a bullet displayed in such a room would be fatal. With appropriate programming such a display could literally be the Wonderland into which Alice walked." (Sutherland 1965)

An iconic feature of VR in many photos or other visual depictions is the special input and output devices worn by the users such as head-mounted displays (HMDs), stereo glasses, spatial tracking devices or data gloves. Accordingly, one way to characterize VR is by highlighting its technological components. However, there is a certain danger with technology-centered approaches that such definitions of VR may refer too much to specific input and output devices (e.g., "wired data suits"), which become quickly outdated by technological progress. "Future-proof" definitions of VR should also be compatible with visionary ideas like Sutherland's Ultimate Display or the Holodeck from *Star Trek*. The following technology-oriented characterizations from the early years of VR still apply to today's VR systems:

> "Virtual Reality (VR) refers to the use of three-dimensional displays and interaction devices to explore real-time computer-generated environments." (Steve Bryson, Call for Participation 1993 IEEE Symposium on Research Frontiers in Virtual Reality)
>
> "Virtual Reality refers to immersive, interactive, multi-sensory, viewer-centered, three-dimensional computer-generated environments and the combination of technologies required to build these environments." (Carolina Cruz-Neira, SIGGRAPH '93 Course Notes "Virtual Reality Overview")

These characterizations of VR can perhaps best be understood in contrast to "traditional" computer graphics, as the science and technology field from which VR evolved. VR builds on 3D content from computer graphics but focuses in particular on *real-time* computer graphics. Matching the 3D content, *three-dimensional displays* are used for its presentation. In the case of the sense of vision, this is achieved using stereoscopic displays. Moreover, 3D content is often presented in a multi-sensory manner by addressing further senses such as hearing or touch, for which spatial audio and haptic feedback devices are employed. Besides 3D presentation, VR systems also facilitate *3D interaction*. 3D interaction devices are input devices whose position and orientation can be tracked in 3D space. Whereas in desktop systems the classic mouse and other "pointing" devices such as trackpads only provide 2D positional information, VR systems make use of 3D tracking to realize, for example, natural pointing gestures. By tracking body and finger movements, grasping of virtual objects can be simulated. *Interactivity* includes users receiving sensory feedback on their inputs, e.g. by mapping hand movements directly onto a virtual hand model. The tracking of the user's position and orientation, particularly *head-tracking*, is the basis for another characteristic of VR systems: Viewer-dependent image generation. If a VR user moves in real space, the 3D environment is automatically displayed from her new perspective. Steve Bryson (2013) succinctly summed up this quintessential property of VR: "*If I turn my head and nothing happens, it ain't VR!*"

Immersion is often considered as an essential feature to distinguish VR from other kinds of human–computer interfaces. Unfortunately, the term immersion is used in non-uniform ways in the literature. Following Slater and Wilbur (1997), we take a technology-centered view of immersion. According to Slater and Wilbur (1997), immersion is based on four technical properties of display systems: *Inclusive* (I) indicates the extent to which the user's sensory impressions are generated by the computer, i.e., the user should be largely isolated from the real environment. *Extensive* (E) refers to the range of sensory modalities accommodated. *Surrounding* (S) indicates the extent to which the presentation is panoramic rather than limited to a narrow area. *Vivid* (V) indicates the resolution, fidelity and dynamic range of stimuli within a particular sensory modality. Immersion is therefore a gradual characteristic that is achieved to different degrees by different displays. For example, HMDs are usually considered highly immersive displays, since the visual sensations of the user are exclusively computer-generated. However, an HMD with a small field of view is less immersive than an HMD with a wider field of view. Similarly, multi-wall projections like CAVEs (see Sect. 9.2) are more immersive than single-screen projections.

The goal of total immersion is achieved by today's VR displays to varying degrees. The terms *immersive VR* or *fully immersive VR* usually refer to VR systems based on HMDs or CAVEs. Desktop systems that provide stereoscopic displays and head-tracking are sometimes referred to as *non-immersive* and large single-screen or table-top displays as *semi-immersive VR*.

Besides the use of the term immersion as a technical property of VR displays, some authors also use the term to describe a mental quality of the VR experience (e.g., Witmer and Singer 1998). To differentiate between the two uses, one also speaks of *physical immersion* and *mental immersion* (Sherman and Craig 2003) and sometimes also of *physiological* and *psychological immersion* (Sadowski and Stanney 2002).

Table 1.1 summarizes the distinguishing features of VR as compared to conventional computer graphics.

Table 1.1 Features of VR as compared to conventional computer graphics

3D Computer Graphics	Virtual Reality
Visual presentation only	Multimodal presentation (i.e., addressing several senses, e.g., visual, acoustic and haptic)
Presentation planning/rendering not necessarily in real-time	Real-time presentation planning and rendering
Viewer-independent image generation (exocentric perspective)	Viewer-dependent image generation(egocentric perspective)
Static scene or precomputed animation	Real-time interaction and simulation
2D interaction (mouse, keyboard)	3D interaction (body, hand and head movements and gestures) + speech input
Non-immersive presentation	Immersive presentation

1.2.2 VR as an Innovative Kind of Human–Computer Interaction

> "The promise of immersive virtual environments is one of a three-dimensional environment in which a user can directly perceive and interact with three-dimensional virtual objects. The underlying belief motivating most virtual reality (VR) research is that this will lead to more natural and effective human–computer interfaces." (Mine et al. 1997)

Another way to characterize VR is to emphasize the goal of creating human–computer interfaces, which, in comparison to traditional user interfaces, enable much more natural or intuitive interaction with the three-dimensional simulated environment (see Fig. 1.2).

Graphical user interfaces (GUIs) and the associated WIMP (Windows, Icons, Menus, Pointing) interaction style represent a paradigm of human–computer interaction that has been dominant for several decades. The WIMP paradigm, which was originally developed for document-processing tasks, however, turns out to be rather inefficient when manipulating 3D content. For example, the task of repositioning an object in 3D space can be naturally achieved in VR by grasping and moving the object. In 2D GUIs, however, this task usually has to be broken down into several subtasks, e.g., first move the object in the *xy*-plane, then move in the *z*-direction. Besides the additional motor effort (e.g., two 2D mouse movements instead of one hand movement in 3D space), this also requires an additional cognitive effort for remembering when and how to change the system control state (e.g., how do you tell the computer that the next 2D mouse movements should be interpreted as translation in the *z*-dimension of 3D space?). As a prerequisite for successfully completing the task, the user must first learn how the 3D task can be broken down into a sequence of 2D subtasks, i.e., there is also a greater learning effort.

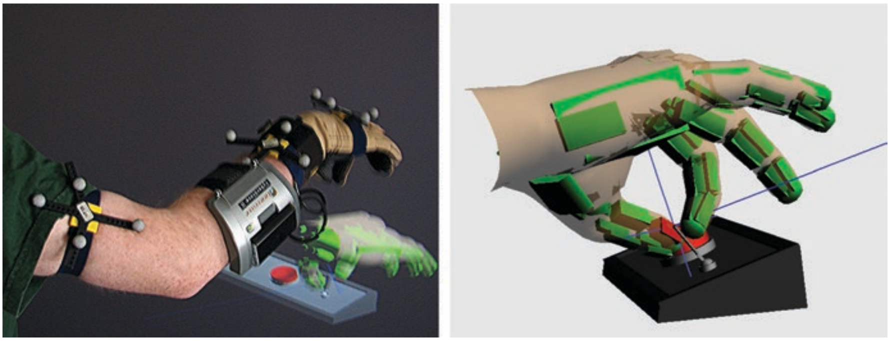

Fig. 1.2 Example of natural interaction: A virtual switch is turned like a physical switch using one's hand

Virtual and Augmented Reality, along with other innovative forms of human-computer interaction, are examples of so-called *post-WIMP interfaces*. Post-WIMP interfaces build on interaction techniques that exploit prior knowledge and skills that the human user has already learned from everyday interactions with physical objects. For example, a person knows from everyday experience how one can use one's body to manipulate objects and has expectations of how these objects will typically behave as a consequence of this interaction. By using this knowledge, learning and further mental effort in natural interaction techniques may be greatly reduced when compared with WIMP techniques.

The following quote from Robert Stone explains the goal of *intuitive user interfaces* in the context of VR systems:

> "An intuitive interface between man and machine is one which requires little training … and proffers a working style most like that used by the human being to interact with environments and objects in his day-to-day life. In other words, the human interacts with elements of his task by looking, holding, manipulating, speaking, listening, and moving, using as many of his natural skills as are appropriate, or can reasonably be expected to be applied to a task." (Stone 1993)

Compared to other innovative forms of human–computer interaction, VR offers great potential for an especially thorough realization of intuitive human–machine interfaces in the sense of Robert Stone. However, the goal of completely natural forms of interaction has arguably not yet been achieved nor is it always aimed for in today's VR systems. Nevertheless, through the use of 3D input and output devices, interactions in existing VR systems are typically much more natural than is the case with conventional 2D interfaces.

> "The primary defining characteristic of VR is inclusion; being surrounded by an environment. VR places the participant inside information." (Bricken 1990)

Metaphors represent another important aspect in the design of human–computer interfaces. Metaphors aim to explain the user interface through analogies with everyday life experiences. Within the WIMP paradigm, well-known examples of metaphors are the desktop, folders with documents in them, or cutting and pasting for transferring parts of one document to another. In the case of VR, the term Virtual Reality itself is a metaphor that makes the analogy to reality as such. The VR metaphor conveys to the user that the objects in the simulated world behave realistically and that natural forms of interaction are supported. Another aspect of the VR metaphor is that the user is situated within the simulated world and experiences it "from the inside" instead of looking at the simulated world "from the outside" through a

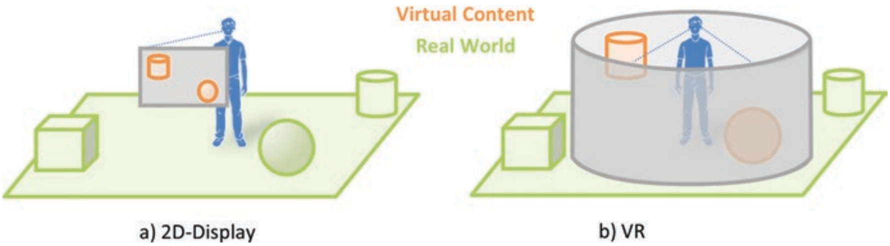

Fig. 1.3 Interaction models for desktop computers and VR: (**a**) When looking at the 2D display of a desktop computer, the user perceives both the real world and the computer-generated environment. (**b**) According to the VR metaphor the user is completely situated within the computer-simulated virtual world and fully isolated from the physical world. (c.f. Rekimoto and Nagao 1995)

window as with conventional desktop computers. According to the VR metaphor – which could be implemented using future perfectly immersive systems – the user is totally isolated from physical reality so that all sensory impressions are computer-generated. Fig. 1.3 contrasts the interaction models of conventional desktop computers with 2D displays and VR.

1.2.3 Mental Aspects of the VR Experience

> "At the heart of VR is an *experience* – the experience of being in a virtual world or a remote location" (Rheingold 1991)

In perfect VR, all of the sensory impressions of the user would be generated by the computer, in the same quantity and quality as people are used to in the real world. Human actions in VR would have the same effects and virtual objects would affect people as they do in the real world. Today's VR systems are by no means perfect, but the development of VR technology is aimed at ever more realistic experiences. But if the computer-generated sensory level is no longer (or hardly) distinguishable from physical reality, what effects does this have on higher-level processes of human perception? Does the user perceive the pixels of the visual displays as pictures or does the feeling of being at another place emerge? What other properties characterize the mental experience of VR? How can you measure or otherwise quantify these properties? How does this inform the design of virtual worlds and the setup of VR systems?

In VR research, these and similar questions regarding the mental experience of VR have played an important role right from the start. The fact that these questions are still the subject of research shows on the one hand their relevance for the research area of VR, but on the other hand that no generally accepted answers have yet become established. Unfortunately, the relevant terms in the literature such as "immersion" and "presence" are sometimes used with different meanings. As noted above, we reserve the term "immersion" in this book, consistent with much of the research community, to exclusively describe the technical properties of VR systems. In contrast, some authors also use the term to describe the mental sensations of VR experiences. When reading different texts on VR, it is necessary to pay close attention to how key terms such as immersion are used. The following presentation of the most important concepts for the analysis of the mental experience of VR essentially follows the terminology of Slater (2003, 2009).

Presence is the central concept for describing the mental aspects of the VR experience. In a broad sense, it refers to the feeling of being within the virtual environment that is displayed by an immersive VR system ("being there"). The concept of presence was originally developed in the context of telerobotics. The aim was to provide the operator with the most realistic impression possible of the robots' environment during remote control of robots, in particular using immersive VR technologies such as HMDs and data gloves. In the early 1990s, the concept of presence was transferred to VR (Held and Durlach 1992; Sheridan 1992). Evidence for (the feeling of) presence is, for example, when VR users react to the virtual environment as if it were a real environment. The concept of presence can be further decomposed to involve three different components:

First, the *place illusion* refers to the feeling of being in the location presented by the VR system (Slater 2009). The place illusion is the human response to a given level of immersion. It tends to arise naturally in highly immersive systems, but is more difficult to achieve with desktop systems (Slater 2009). Particularly important is the ability of the immersive VR system to display the scene from the perspective of the viewer. If the user turns their head, then the virtual environment should still be visible, just from a different perspective. If this is not the case, e.g., due to a single-screen setup, a *break in presence* may occur.

Second, the *plausibility illusion* arises when the events of the simulated environment are perceived as if they are really happening (Slater 2009). While the place illusion is largely induced by how the virtual world is presented, the plausibility illusion has to do with the content of the simulated world. The plausibility illusion relates in particular to events that affect the user but were not initiated by the user him or herself (e.g., a projectile suddenly flying towards the user or a virtual person who appeals to the user). The *believability* of the virtual environment seems to be more important than sensory realism for the emergence of the plausibility illusion. For example, a visually perfectly represented virtual person who communicates only in simple phrases would lead to a break of the plausibility illusion.

Third, *involvement* refers to the level of user attention or interest in the simulated world (Witmer and Singer 1998). Involvement, like the plausibility illusion, is mainly related to the content of the virtual environment. For example, in an immersive VR system, users might feel strongly that they are part of the simulated world (convincing place illusion), but may still get bored (low involvement).

To test whether and to what degree the feeling of presence arises with users, experimental studies with test persons are necessary. Different users may experience different levels of presence in one and the same VR application. One way to record presence is to use special questionnaires (e.g., Witmer and Singer 1998). Furthermore, the behavior of the experiment's participants can be observed, for example movements (e.g., a user ducks away when an object comes flying towards them at high speed) and emotional expression such as fright. Other studies measure physiological parameters such as heart rate or skin resistance, which are often interpreted as signs of stress. In Slater et al. (2010) a "VR in VR" scenario is proposed as a further possibility for quantifying presence, in which the user can configure a VR system in the simulated world that generates a maximal level of presence.

Finally, the feeling of presence is not limited to VR, but may also arise with other media, such as books, movies or arcade machines, though perhaps not equally intensely. A further discussion on this can be found in Sherman and Craig (2003), for example.

1.3 What Is AR?

In the literature, a large number of different and sometimes contradictory definitions of AR exist. While Ivan Sutherland was the first to create an AR system in the late 1960s (Sutherland 1968), the definition according to Azuma from 1997 is widely used in science.

> "Augmented Reality (AR) is a variation of Virtual Environments (VE), or Virtual Reality as it is more commonly called. VE technologies completely immerse a user inside a synthetic environment. While immersed, the user cannot see the real world around him. In contrast, AR allows the user to see the real world, with virtual objects superimposed upon or composited with the real world. Therefore, AR supplements reality, rather than completely replacing it." (Azuma 1997)

According to Azuma (1997), an AR system (see also Sect. 1.6) has the following three characteristic features: (1) It combines reality and virtuality. (2) It is interactive in real time. (3) The virtual contents are registered in 3D. While the second feature is also found in VR, the other two aspects differ significantly from VR. The combination of reality and virtuality is typically achieved by overlaying reality with (artificial) virtual content. That is, an observer (the AR user) simultaneously perceives the real environment and the virtual objects within it as a coherent whole. The virtual content allows for real-time interaction. Furthermore, the virtual content is registered in 3D (i.e., geometrically). This means that in an AR environment, a virtual object has a fixed place in reality and, as long as it is not changed by user interaction or changes itself, e.g., by animation, it remains there. In other words, from the user's perspective, it behaves exactly like a real object that would be in that location. As registration in 3D space and visual superimposition occur in real-time, this does not change even if the user changes their perspective and therefore perceives a different part of the environment.

In the domain of popular science, the term AR is often used to refer to examples limited to the first of the features described by Azuma (i.e., the augmentation of reality by virtual content), while interactivity, real-time capability and especially 3D registration are frequently ignored.

More generally, AR may be defined as follows:

> Augmented Reality (AR) refers to the immediate and seamless perception of the real environment enriched by virtual content in real-time, the latter resembling reality to the largest extent possible regarding its characteristics, appearance, and behavior, so that (if desired) sensory impressions from reality and virtuality may become indistinguishable (for any senses).

Implicitly, this definition also includes the aspects of interactivity and real-time capability, though it considers AR from the perceptual perspective. While AR today (as in much of this book) is mostly limited to the augmentation of visual perception, it can, just like VR, extend to any other form of sensory experience, including auditory, olfactory, gustatory, haptic (including tactile), vestibular, proprioceptive, thermoceptive and nociceptive perception. In contrast to VR, it is not intended to replace the sensory impressions completely by virtual ones. Rather, real and virtual sensory impressions are mutually superimposed.

In addition to AR, the term *Mixed Reality* (MR) is often used, indicating that real and virtual content are mixed together. Although MR and AR are often used interchangeably, MR, unlike AR, represents a continuum. The MR taxonomy of the reality–virtuality continuum introduced by Paul Milgram et al. (1995) is widely accepted in the research community (see Fig. 1.4).

> *Reality–Virtuality Continuum (according to Milgram): Mixed Reality (MR)* is a continuum that extends between reality and virtuality (Virtual Reality), whereby the share of reality continuously decreases while that of virtuality increases. As far as the share of virtuality is prevailing here, without the environment being completely virtual (Virtual Reality), one speaks of *Augmented Virtuality*. If on the other hand the share of reality is larger, then we are talking about AR.

While Azuma sees AR as a special case of VR, Milgram et al. define AR as one representation of MR, whereas MR and VR are disjunct. Thus, while using the AR definition from Azuma, we will apply the taxonomy from Milgram throughout the remainder of this book. Furthermore, although the term *XR* as an abbreviation for *eXtended Reality* goes back to a patent application by the photographer Charles

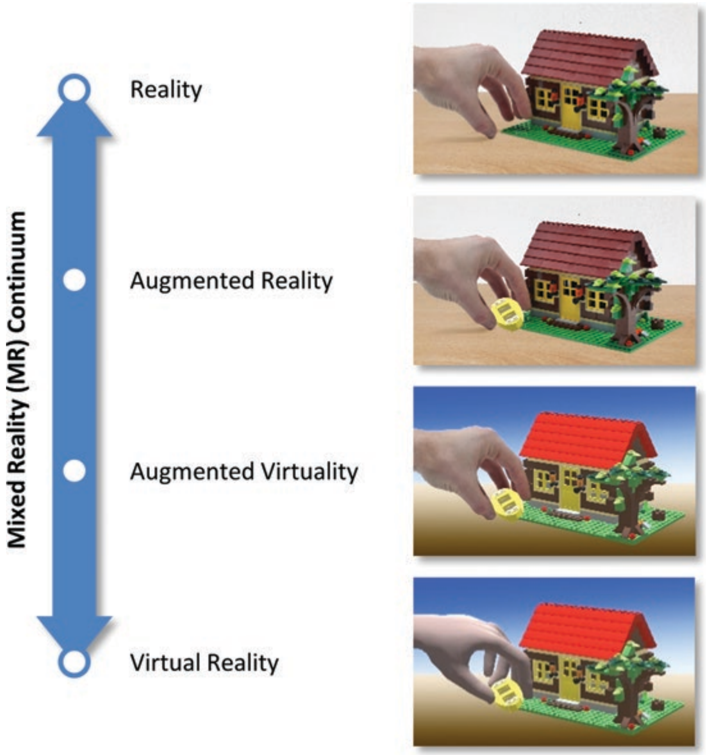

Fig. 1.4 Reality–Virtuality Continuum. (According to Milgram et al. 1995)

Wyckhoff in 1961, it has also been used since then by Sony, for example, to describe their X-Reality technology, or by Paradiso and Landay (2009) and others to describe types of *Cross-Reality,* i.e., a crossbreed between a Virtual Reality and ubiquitous sensor/actuator networks placed in reality. In this book, however, we will use the currently most common variant, namely XR as a generic term for VR and MR (and by that also AR). In this sense, the "X" may also be considered as a placeholder for "V", "A" or "M". As the "X" resembles a cross, XR is also sometimes referred to as Cross Reality (not to be confused with the concept of Cross-Reality mentioned above).

> "Virtual Reality (VR) replaces the user's perception of the real environment by a virtual world. In contrast, Augmented Reality (AR) augments or enhances the perception of reality by virtual content – Diminished Reality (DR) removes parts from the real environment. Augmenting, enhancing, deliberately diminishing, or otherwise altering the perception of the real environment in real time is referred to as Mediated Reality" (Mann 1999)

VR replaces the perception of the user's real environment by that of a virtual world. AR enriches the user's perception of the real environment by virtual content (see Fig. 1.5). In *Mediated Reality* the perception of the real environment is augmented, enriched, consciously diminished or otherwise changed in real time (Mann 1999). If the perception of reality is consciously reduced, i.e., real contents of the environment are deliberately removed from the perception of the user in real time, this is called *Diminished Reality* (DR). While not necessarily following the extended taxonomy of Mann et al. (2018), we use will use their definitions of Mediated Reality and Diminished Reality in this book. Further, we will consider eXtended Reality (XR) to be a subset of Mediated Reality. For clarification of the taxonomy as used in this book, refer to Fig. 1.6.

Comparing AR with VR (see Table 1.2), it becomes obvious that many basic characteristics are very similar, if not identical. Both use a multimodal presentation, in that both interaction and simulation take place in real time, both visualize virtual 3D objects, and both use the egocentric perspective, i.e., the visualization is (at least conceptually) correct in terms of perspective for the respective viewer (although this

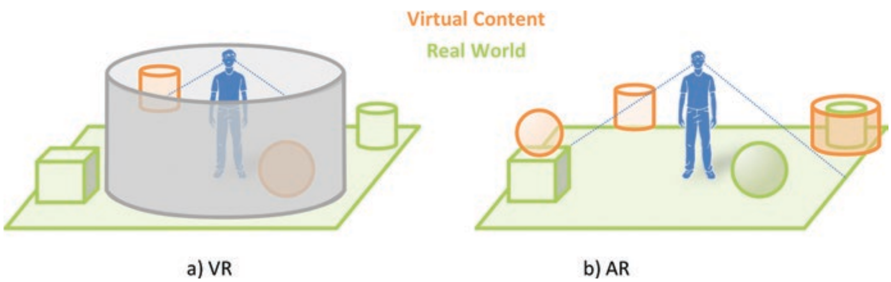

Fig. 1.5 AR compared to VR. In contrast to VR, the user interacts with the virtual content as well as with the real environment. Furthermore, an interaction between the real environment and the virtual content can take place. Virtual content and the real environment are not strictly separated from each other, but can overlap, be superimposed or penetrate each other

Fig. 1.6 Euler diagram showing the relationships between Augmented Reality (AR), Augmented Virtuality (AV), Mixed Reality (MR), Virtual Reality (VR), eXtended Reality (XR), Diminished Reality (DR) and Mediated Reality

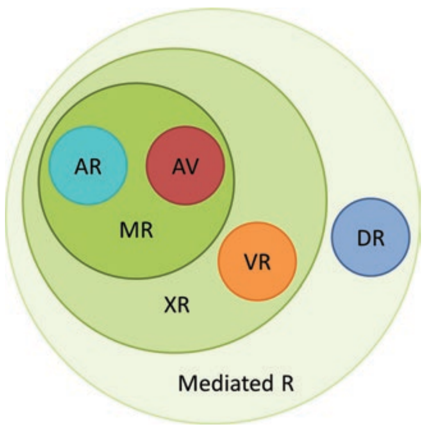

Table 1.2 Features of AR as compared to VR

Virtual Reality	Augmented Reality
Multimodal presentation	Multimodal presentation
Real-time presentation planning and rendering	Real-time presentation planning and rendering
Viewer-dependent image generation(egocentric perspective)	Viewer-dependent image generation(egocentric perspective)
Real-time interaction and simulation	Real-time interaction and simulation
Virtual 3D objects	Virtual 3D objects
All content purely virtual	Combination of reality and virtual content
Immersive presentation (central aspect)	Immersive presentation (open issue)
Tracking	Tracking and geometric (3D) registration
Implicit (restricted) and explicit navigation	Implicit (unrestricted) navigation
Stationary	Stationary or mobile
Indoor	Indoor and outdoor
Virtual illumination	Mutual influence of real and virtual illumination
Arbitrary scaling of the user perspective	User perspective always unscaled (virtual models may have limited scalability)

is not always the case with actual VR and AR systems). However, there are also a number of differences: The most obvious difference is that in VR all content is purely virtual, whereas in AR the virtual content is embedded in the real world. Accordingly, there is no real immersion in AR like there is in VR. For its application to AR, the concept of immersion would have to be significantly expanded. In AR, the focus is rather on the correct superimposition or fusion of reality and virtuality. This is achieved by registration. VR and AR also differ with respect to navigation. While in VR implicit navigation (the user moves in the virtual world analogous to movement in reality) is limited due to the inherent limitation of the dimensions of the room, the tracking area, the cable length of the HMD or the dimensions of the CAVE (see Sect. 1.4), navigation in AR is often unrestricted. For this purpose, VR additionally enables explicit navigation, in which the user changes their point of view by changing the camera position using specific interaction techniques. This allows, for instance, the user to fly through a virtual world, which is obviously not possible in AR at all. VR takes place primarily in closed rooms and these are usually stationary (location-bound) systems. Although there are many AR applications for indoor use, AR is generally not limited to these. Many AR applications are mobile and used outdoors. Also, the lighting and scaling of the virtual contents are fundamentally different. While in VR only the virtual lighting is of importance, in AR there is a mutual influence of the real and virtual lighting situation, although this is only rudimentarily or not at all considered by many applications. In VR, content can be scaled as desired. A user can, therefore, move between molecules or microbes as well as holding the entire Milky Way in their hands. With AR, in contrast, the real environment always provides a frame of reference, so that virtual objects usually have to be on a scale of 1:1. Of course one could also superimpose the Milky Way in AR in such a way that the user is holding it in their hands. However, the perception would be fundamentally different. While in VR the users get the illusion that they have shrunk to the size of a microbe or grown to the size of a galaxy, in AR the users have the impression of holding a model of the Milky Way, since their own size remains unchanged in relation to the real environment.

Sometimes you may hear the question: Which one is better: VR or AR? This question cannot be answered because VR and AR are aimed at different application scenarios. There will rarely be a situation where you have a choice between VR and AR when it comes to implementing them. Rather, the application scenario usually determines the type of system to be used. This, however, does not mean that VR and AR cannot complement each other – in fact, quite the opposite! Thus, for example, in a purely virtual environment (VR), the details of a complex machine can be explained to trainees, problem and danger scenarios can be simulated and options can be tested that do not exist in reality (at least not on site). By using AR, the acquired knowledge can then be tested and further consolidated on the real machine with virtual support. For instance, it is possible to look into a component using virtual X-ray vision, etc. Basically VR, in contrast to AR, has no limitations: neither in content nor in physics (in a VR you can define your own physics!). On the other hand, the continuous use of VR is – at least currently – limited to rather short periods of time (minutes rather than hours). Since you always have to leave the real world for VR, this will not change fundamentally (unless we will live in the matrix one day). AR, on the other hand, has the potential to be used always and everywhere (24/7), although this potential currently cannot be fully exploited due to limitations in software and hardware.

1.4 Historical Development of VR and AR

The history of VR as a field of science and technology began in the 1960s. As part of his research on immersive technologies, Ivan Sutherland (1965) wrote “The Ultimate Display”, in which he described the vision of a room “within which the computer can control the existence of matter”. In his pioneering work, Sutherland took the first step towards connecting the computer with the design, construction, navigation, and experience of virtual worlds, even before the personal computer (PC) was invented (1970). In 1968, Sutherland created a *Head-Mounted Display System* consisting of a data helmet and a mechanical and alternatively ultrasound-based tracking system (see Fig. 1.7a). This system (Sutherland 1968) is often erroneously called the “Sword of Damocles” in the literature, although this was only the name of the mechanical tracking component of it. It enabled the viewer to view a simulated, albeit simple, 3D environment in the correct perspective. This system can also be regarded as the first AR system due to its see-through property.

The VIEW project (Virtual Environment Interface Workstations) of the NASA Ames Research Center in the mid-1980s had the goal of developing a multi-sensory workstation for the simulation of virtual space stations.

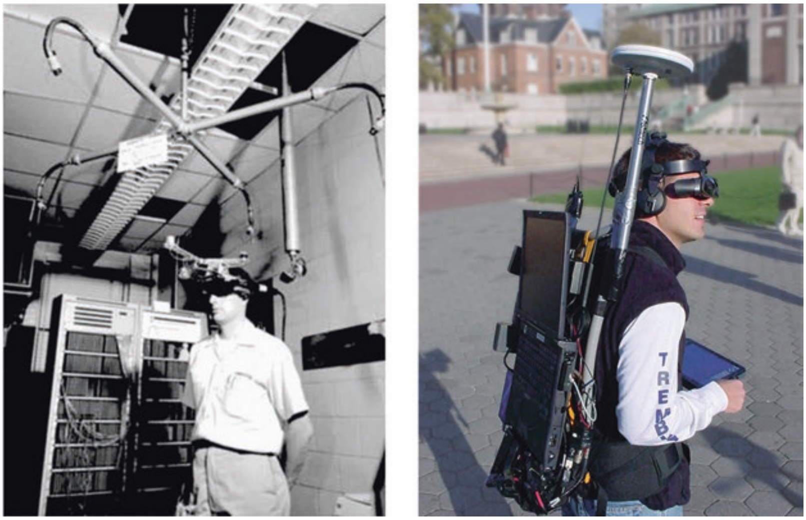

Fig. 1.7 Pioneering work in the field of VR/AR. (Left) Sutherland's data glasses with 6-DOF ultrasound tracking; image courtesy of © Ivan Sutherland, all rights reserved. (Right) Replica of the MARS system of 1997 (Bell et al. 2002). (Image courtesy of © Steve Feiner, all rights reserved)

Around 1987 Thomas Zimmermann described the "DataGlove", a glove that was equipped with glass fibers on the top of the hand to capture finger flexion. He and Jaron Lanier jointly founded the company VPL. Lanier is often credited as the first scientist to use the term "virtual reality". Besides selling the "DataGlove", VPL also developed the "EyePhone" data helmet, a continuation of Sutherland's Head-Mounted Display from the 1960s. The LX version of the EyePhone offered a resolution of 442 × 238 pixels, while the HRX version offered 720 × 480 pixels.

Another milestone was the commercialization of electromagnetic trackers by Polhemus 3Space in 1989. This made it possible to control or determine a target at a certain distance from a computer.

Around the same time, the "BOOM" (Binocular Omni-Orientation Monitor) was developed by Fake Spaces Labs, a 3D visualization device with two monochrome cathode ray tubes, which received NTSC signals generated by a Silicon Graphics Workstation VGX380 (8 RISC processors, 33 MHz per processor, 1280 × 1024 pixels at the graphics output). This workstation was able to generate 800,000 small, transformed and shaded triangles per second that were also clipped at the border of the drawing area. One of the first applications to take advantage of this feature was the "Virtual Wind Tunnel" in the aerospace field by Steve Bryson in 1991.

Around 1988, various high-quality workstations for graphics were introduced to the market. These included Ardent, Stellar, Silicon Graphics (SGI) and HP, of which the SGI Reality Engine from Silicon Graphics prevailed on the worldwide market for high-end graphics systems around 1995. Commercial VR software systems were also introduced to the market. These were "RB2 – Reality built for two" by VPL, "dVS" by the English company Division and "WorldToolKit" by Sense8 (1990–1995).

The term “Augmented Reality” was coined in the early 1990s by a pioneering project at Boeing, which used information superimposed on the visual field to make it easier for workers to lay aircraft cables (Caudell and Mizell 1992).

In 1993, a student of the Massachusetts Institute of Technology (MIT) founded SensAble Technologies Inc., a company that developed and commercially distributed haptic devices. The “PHANTom” facilitated the experience of force feedback – a great innovation at that time.

At the beginning of the 1990s, groundbreaking research was undertaken in the field of Virtual Reality. For the first time, these made projection-based representations possible. The main representatives of these are the “Powerwall”, which consisted of a stereo screen, the “CAVE” (Cave Automatic Virtual Environment), which had four screens (developed at the University of Illinois in 1992), the “Responsive Workbench”, which arranged a screen horizontally analogous to a table surface (developed by GMD in 1993), and “iCONE”, which used semicircular screens.

With “MARS” (see Fig. 1.7b), the first mobile AR system was presented at Columbia University in 1997 (Feiner et al. 1997). The publication of ARToolkit in 1998 (Kato and Billinghurst 1999) made computer vision-based tracking for AR available and triggered a huge wave of research around the world.

After the development of electromagnetic tracking systems, ultrasonic tracking systems came on the market, which in turn were replaced by optical tracking systems based on infrared light around the year 2000. PC clusters also replaced the SGI Reality Engine II, reducing the price for the user to about one fifth. This made more extensive research possible. Founded in 1993, the company Nvidia released their GeForce graphics chips as a successor to the RIVA chip family in 1999. Introducing advanced features to consumer-level 3D hardware, the GeForce is a milestone in graphics hardware.

On the software side, Silicon Graphics developed a toolkit named OpenInventor (originally IRIS Inventor) to support application development that also benefitted VR applications in 1988. It was based on the ANSI standard PHIGS that introduced the concept of a scene graph. The Open Graphics Library (OpenGL) debuted in 1992. With the success of the World Wide Web, VRML, a dedicated markup language for VR, was developed and became an ISO standard in 1997. It would later evolve into X3D. This was also the time when dedicated VR software companies emerged and basic application areas were explored. For example, Henry Fuchs investigated telepresence applications as well as medical applications with VR/AR (Fuchs et al. 1998).

There is a regular exchange of information on the subject of VR throughout the world. In the USA there have been VRAIS Symposia since 1991 and in Europe EuroGraphics VE Workshops since 1993. In Japan the ICAT workshops have also taken place since the beginning of the 1990s. In 1999 the IEEE VR Conference was established as the successor to the VRAIS, which attracts about 500 participants from all over the world every year. Similarly, dedicated conferences on the topic of AR were introduced, e.g., ISMAR, the IEEE International Symposium on Mixed and Augmented Reality, which started in 2002 as a merger of the International Symposium on Augmented Reality (ISAR) and the International Symposium on

Mixed Reality (ISMR). Moreover, VR and AR have been featured in trade shows such as the consumer electronics show (CES).

For several decades, access to VR and AR technology was limited to research institutions, large industrial companies and government agencies, not least because of the sometimes astronomical prices for the necessary hardware. This changed abruptly with the introduction of the first high-end low-cost data glasses, Oculus Rift, in 2013. Since the delivery of the consumer version in 2016 and the market entry of numerous comparable displays (HTC Vive, Playstation VR, Microsoft's "Mixed Reality" displays, etc.) VR has experienced a boom. Approaches to AR glasses have not yet been able to achieve this success. For example, Google Glass has not yet prevailed in the market and Microsoft's Hololens is considered a technical masterpiece but has not achieved widespread use quickly. A new phase in the evolution of AR applications started in 2017 with the introduction of several major software platforms for mobile AR on smartphones and tablet computers. Apple presented ARKit and Google presented ARCore, two modern frameworks that have started to strongly influence the commercial development of AR applications.

1.5 VR Systems

If we summarize the previous requirements for a VR system, we get the following situation: We need a computer system that recognizes the actions of users, simulates the world under this influence, and lets users perceive a virtual world via appropriate stimuli. Technically, three parts can be distinguished: input devices, output devices, and the world simulation. As simply as the tasks of a VR system can be broken down into these three parts, each subsystem can become rather complex: Which sensors can detect a user's actions? What coverage and resolution do these sensors have in terms of space and time? What range of actions do these sensors allow? Do the sensors restrict or limit the user? How can sensor data be passed on to the simulation of the world? How can knowledge about the world be made available to the simulation? How can stimuli be generated in a suitable way for all perception channels of the user? What is the quality of these stimuli? In what radius of action can the user sensibly perceive these stimuli? How can it be ensured that the response time of the overall system keeps pace with the actions of the user?

To demonstrate the importance of the individual subsystems of a VR system, let us revisit a prior experiment and examine it in more depth. In that experiment, we had placed a user in VR on the edge of a virtual abyss to observe the user's reactions to images of the user's surroundings. The user's position and viewing direction must be tracked by the input devices all the time to be able to calculate the correct perspective for the user in the virtual environment. In the first variant of the experiment, it was assumed that a sensor was built into the user's helmet to provide this position and orientation data. What does such a sensor look like? Is only the orientation of the head recognized or also the direction of the eyes? What distances of movement does such a sensor allow? In addition to tracking the head's orientation, is it possible to also track the position of the head so that bending forward is possible in the virtual environment? Can you approach the virtual abyss by taking a step or two? Is it possible to walk on the entire roof of the virtual skyscraper? In addition, is it possible to track the whole body with all limbs to visualize the user's body as an avatar to support self-perception? Would this body tracking recognize only roughly the limbs or also individual finger movements, e.g., is it possible to press the elevator button with one finger to leave the roof of the virtual skyscraper by elevator?

In the early days of Virtual Reality, it was common to attach many of the sensors required here, and the input devices were mostly connected by cable (called *wired clothing*). Examples of this are helmets with monitors or data gloves to recognize finger movements. Electromagnetic and ultrasound-based devices have also been developed over time. Such systems usually consist of transmitter(s) and receiver(s), so that users had always something attached to their bodies. These days the trend is towards optical processes based on one or more cameras, whereby a distinction is to be made regarding the use of so-called markers or markerless systems. Markers are patterns known to the VR system that can be detected automatically with high reliability. Markers can be used to enable or stabilize the camera-based detection, as they are designed in a way that they are easy to detect in camera images and less prone to detection errors due to factors such as occlusions or unfavorable lighting situations. In addition to RGB cameras, markerless systems often use so-called depth cameras, which support the extraction of foreground objects and background. By using multiple cameras, accuracy can be improved and situations can be avoided, where tracking fails in single-camera setups due to occlusion.

Multiple, possibly redundant, input devices are often used at the same time to ensure the best possible recognition of user actions. An example of this is the combination of precise position tracking within a large action space in combination with hand/finger tracking and voice input. Here, the sensor data must be aggregated in a suitable form (sensor fusion) in such a way that overall plausible and non-contradictory data are provided reliably by combining sensor data of different types, even if single sensors fail due to occlusions.

When designing or configuring a VR system, one should always focus on the actual task and analyze which input devices are necessary. It is not always advantageous to include as many sensors as possible in a setup if this results in restrictions for the user. In our abyss experiment, it could be possible to measure the pressure distribution of the sole to infer whether the user is leaning forward or backward. This could be done using pressure-sensitive mats, which would require that the user may only stand on the mat, and thus the user's location would be fixed. This would be counterproductive in relation to other objectives, e.g., that the user should be able to move freely.

The output devices are the counterpart to the input devices. They serve to present the virtual world to the user in multiple modalities. This conversion of the virtual world model to sensory stimuli for the user is called rendering. According to the different sensory modalities through which humans perceive the real world, it is helpful to address as many of them as possible in Virtual Reality. Regarding our experiment, the visual output is of course highly important. Should the user be able to look around freely, as would be possible with a tracked helmet? Is it enough for the user to look down only, as in the second variant of the experiment, in which the image is projected onto the floor? Is it important for this use case that the user can turn and look around?

Which action space should be provided where the user can perceive the virtual environment? In which visual quality should the virtual world be presented? Is it important to recognize moving cars or pedestrians from the skyscraper? In addition to visual stimuli, other sensory modalities of the user can also be addressed. Should the noise of road traffic be perceived louder when you get closer to the edge of the building of the skyscraper? Should the user be able to perceive wind, and should it also change at the edge of the building? As already discussed, the time requirements for the stimulus calculation and rendering for the individual sensory modalities also differ. For the visual system, many new images must be calculated every second. In contrast, it is enough to determine the strength of the wind from the example once or twice per second. It is advisable to analyze exactly what is important for the actual application, instead of implementing everything that is technologically possible.

The task of world simulation is performed by a computer system that relies on an appropriate world model. This model determines the behavior. Depending on the application, physically based simulation models (e.g., for simulating flow behavior) or models based on artificial intelligence (AI) may be appropriate. The world simulation responds to data from the input devices. In addition to the question of the granularity in which the world is or can be modeled, which was dealt with in Sect. 1.1.2, there are questions relating to technical issues: Which delays occur from recognition by an input device to rendering in all output devices? To reduce this response time (which is called end-to-end latency), it may be helpful or even necessary to use pre-calculated simulation data instead of calculating everything in real time. For our experiment, the movements of road traffic can be calculated as well as the flow simulation for the winds between the skyscrapers. It may even be necessary

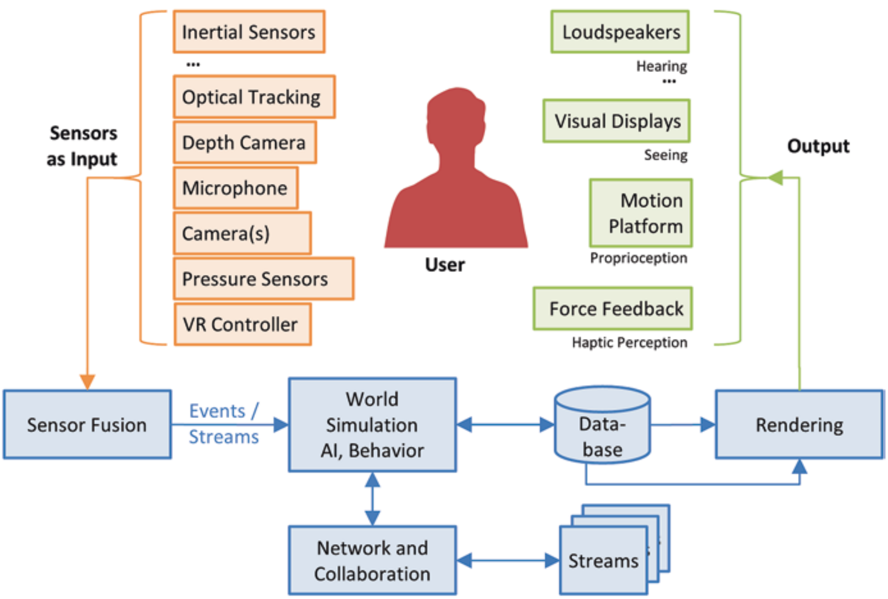

Fig. 1.8 Overview of the subsystems of a VR system

to make major simplifications to keep the delays to an acceptable level. It may also be necessary to distribute the calculation of the world simulation or the rendering to several computers. Does the world simulation rely on locally available data only or does it depend on remote data (e.g., current flight data for a simulator for air traffic controllers or data from VR systems that enable collaboration in virtual space)? Such data can be made available to the world simulation via network connections.

An overview of a VR system is shown in Fig. 1.8, with sensors, which can serve as input devices (in orange), output devices that address the various sensory modalities (in green), and all remaining subsystems of the VR system (in blue).

1.6 AR Systems

We define the term *AR system* by analogy with the already introduced term *VR system*.

> We call an *AR system* a computer system that consists of suitable hardware and software to enrich the perception of the real world with virtual content as seamlessly and indistinguishably as possible for the user.

Even though an AR system typically looks different, its basic composition comprising subsystems is very similar to that of a VR system. Consider the requirements for an AR system: Again, we first need a computer system that performs a simulation depending on user activities. However, this simulation only affects certain parts of the world. One might be inclined here to limit the simulation of an AR system to the virtual part of the world perceived by the user. However, this is by no means sufficient for AR. Since the real and virtual contents are closely intertwined, i.e., there is an interdependency between the two, the parts of the real world that are influenced by the virtual content or, respectively, influence the virtual content, must also be simulated. In AR, the stimulus is generated in such a way that the real and virtual contents complement each other. Many aspects relating to sensors and stimuli apply in a similar way to AR systems. However, in contrast to VR systems, AR systems are usually not restricted to a specific location. This means that factors such as the operating range are omitted, but questions regarding the usability in certain environments have to be added. Can I use my AR System indoors or only outdoors? Will it still work in the subway? What if I am in a room with smooth white walls? Will the display work in sunlight? So, does an AR system have higher or lower technical requirements than a VR system? There is no general answer to these questions, but in a non-stationary system the amount of hardware is naturally limited. Thus, AR systems use on average fewer devices (sensors, output devices, computers, etc.) than VR systems. Nevertheless, the baseline requirements are rather high. While in the above example in the VR system we had a variety of configurations with more or

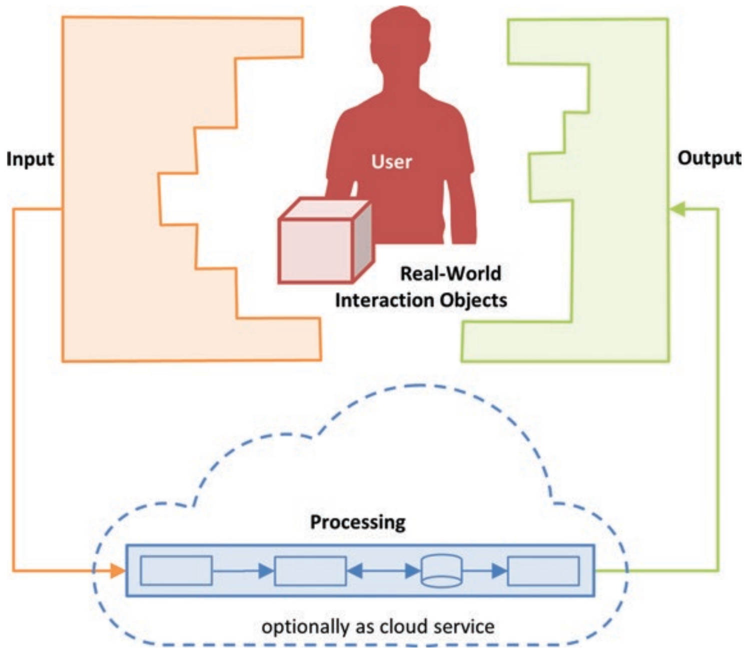

Fig. 1.9 Overview of the subsystems of an AR system. (See also Fig. 1.8)

less sophisticated sensor technology, an AR system must always guarantee the correct superposition of the real and virtual worlds with the proper perspective. On the other hand, many components of VR systems are not required. Through the awareness of reality, self-perception is always guaranteed. Also, navigation in the virtual world is not necessary, because users change their point of view by moving in their natural environment, the real world. While in VR systems the sensors, world simulation and stimulus generation are often distributed over a number of computer systems to ensure the required performance of the overall system, most AR systems are confined to a single computer system. This can be a mobile device such as a smartphone or tablet or it is sometimes completely integrated into AR data glasses (such as the Microsoft Hololens). However, there are also approaches where optical tracking or rendering are outsourced to external systems to improve quality.

The overall view of an AR system is shown in Fig. 1.9: By analogy with Fig. 1.8, the sensors for input are shown in orange, output devices in green and the other subsystems of the AR system in blue.

1.7 Using the Book

In the following, you will find information on how the book is structured and suggestions on how the book can be used by different target groups for different purposes. Recommendations for use in academic courses are also given.

1.7.1 Structure of the Book

Following this introduction, the next chapter (Chap. 2) describes the basics of spatial perception. Starting from the human visual system, the theory of "depth cues" is presented, which describes the basic theory of spatial perception. The physiological aspects of stereoscopy are considered as well as supporting recommendations to enhance spatial perception. In addition to visual perception, the importance of other perceptual channels is discussed. The chapter on virtual worlds (Chap. 3) describes typical concepts employed to build them. Starting from data structures like the scene graph, advanced modeling concepts for virtual worlds are presented: Examples are animation methods, behavior descriptions and event models. In the chapters about VR input devices (Chap. 4) and VR output devices (Chap. 5), the characteristics of sensors and displays are described. After the introduction of underlying properties, methods for the tracking of user actions are shown as well as realization alternatives addressing the different sensory modalities of the user. Based on individual technologies, typical setups with VR hardware are also presented. Concepts and techniques for interactions in virtual worlds are presented in Chap. 6. Basic techniques for navigation and selection are described as well as the iterative approach to creating user interfaces based on user testing. Chapter 7 describes the requirements for the real-time capability of VR systems and presents solutions to meet these requirements. Based on fundamentals such as the importance of latency and efficient representations of large scenes, procedures for typical problems like synchronization and collision detection are discussed. Chapter 8 is dedicated to the topic of Augmented Reality. In addition to special input/output devices, the focus is on geometric and photometric registration as well as on the question of how authenticity or believability can be increased. Chapter 9 contains a series of small case studies that provide insights into the practice of VR/AR and illuminate the many facets of the topic. Software and tools for the practice of VR/AR development are the subject of Chap. 10, while Chap. 11 contains an introduction to the basic mathematics relevant to VR and AR.

1.7.2 Usage Instructions

Each further chapter of this book presumes having read in this chapter. For example, to work through Chap. 6, it is not necessary to read Chaps. 2, 3, 4, and 5 but only the first chapter. This means that the book can be used modularly and selectively – it does not have to be worked through from front to back. All the necessary basic knowledge has already been addressed in this chapter. Although the individual chapters of the present book differ considerably in the complexity of the material dealt with and thus in their scope, all chapters are structured according to a similar basic pattern. This enables the readers to find their way around the individual chapters quickly and to work through them similarly.

Chapters always start with an abstract that summarizes the most important content in a concise form. This enables readers who already have prior knowledge in individual areas or are only interested in certain topics, i.e., who do not want to work through the book sequentially, to quickly identify and select the chapters relevant to them. The most important topics are then dealt with in the respective subchapters. The individual chapters are concluded with a list of questions on the topics covered and a list of recommendations for more in-depth or supplementary literature.

1.7.3 Target Groups

This book is primarily an academic textbook, i.e., it is intended to offer teachers and students a comprehensive and structured treatment of the topic of VR/AR. Therefore, fundamental aspects of VR and AR are covered. Prior knowledge in this field is therefore not necessary, but mathematical basics and basic knowledge of computer graphics are useful. Chapter 11 contains a summary of the most important mathematical elements of VR. A comprehensive and in-depth treatment of all topics relevant to VR/AR would go far beyond the scope of a single book – this book can serve here as an introduction and preparation for the study of specialist literature.

The book has a modular structure – each chapter only requires the reading of in this chapter. This allows students and teachers to adapt the order in which they work through the subject matter to the requirements of their course. It is also possible to select individual chapters and to omit other chapters (except in this chapter) without any problems, as it is not a prerequisite for understanding that all previous chapters have been read.

The creation of interactive virtual worlds is also one of the foundations of modern 3D computer games. Although the present book deals with these topics and there is considerable overlap with the realization of computer games, the book is not primarily aimed at developers of computer games, as game-specific aspects are not considered.

Lecturers in the Field of VR/AR

The book can be used directly as a basis for lectures and seminars in the field of VR/AR. Due to the modular structure of the book it is easy to vary the order of the different topics and thus to adapt to the individual requirements of the respective teaching unit. The individual chapters conclude with a collection of comprehension and transfer questions. These can be used directly as a basis for corresponding examinations or the preparation for them.

In the following, some typical combinations for individual courses are shown as examples. However, this can and should only serve to illustrate and in no way replaces the individual selection based on the respective curriculum and scope.

Introduction to VR/AR

Chapter 1
Sections 2.1, 2.2, 2.3, 2.4
Sections 3.1–3.3, optional 3.5
Sections 4.1, 4.2, 4.3, 4.6
Sections 5.1, 5.2, 5.3, 5.4
Sections 6.1, 6.2, 6.3, 6.4, 6.5
Sections 7.1, 7.2, 7.3
Section 8.1, 8.3, 8.4

3D User Interfaces

Chapter 1
Sections 2.1, 2.2, 2.3, 2.4, 2.5.2
Sections 4.1, 4.2, 4.3, 4.6
Chapter 6: all subchapters
Section 7.1
Section 8.5

Applications of Virtual Reality

Chapter 1
Sections 2.4, 2.5
Chapter 3: all subchapters
Sections 5.1, 5.2, 5.3
Chapter 6: all subchapters
Section 7.2
Section 8.6
Chapter 9 (VR examples)
Section 10.1, 10.2/10.3

Graphically Interactive Systems

Chapter 1
Chapter 2: all subchapters
Chapter 4: all subchapters
Chapter 5: 5.1
Chapter 6: all subchapters
Chapter 9: all subchapters
Chapter 10: all subchapters

Augmented Reality

Chapter 1
Chapter 3
Sections 4.1–4.4
Sections 5.1, 5.2, 5.3
Chapter 6
Chapter 8
Chapter 9 (AR examples)
Chapter 10

Students

The book offers students a universal companion and reference reading for courses on VR, AR and XR. In addition, it enables the self-study of the subject matter. The book is suitable for students of courses of study who might want to develop or extend VR/AR systems themselves, implement VR/AR applications or just use VR/AR applications. While the first aspect particularly appeals to students of Computer Science, Media Computing, Computational Imaging and Media Technology, the other aspects cover a wide range of natural and engineering sciences, humanities and social sciences, as well as creative and artistic fields.

Users and Those Who Want to Become Users

Potential users of new technologies such as VR and AR often have only a vague idea of the potentials and limitations as well as the resources required for their use. This leads to the fact that such technologies are often not used at all or are used too late. Or even worse, many introductions fail in the end. One of the main problems is that often extensive investments are made in hardware before it is clear whether and how it will be used afterward. Who are the users? Who benefits? How are the users trained? How is the infrastructure maintained and developed? Which applications should be created or used? How are they integrated into a production process or adapted to it? This book should help potential users of VR and AR to better assess these issues in advance and thus prevent or at least reduce misplanning. For users from both research and industry, the book enables them to deal with the topic in detail and thus to assess whether and to what extent the use of VR and AR appears to be sensible and what resources are required for this.

The Technology-Savvy

Ultimately, the book reflects the current status quo in the field of VR/AR and thus gives the technologically interested reader an insight into this fascinating world. New techniques and technologies that are currently still primarily used in research or research-related prototype and application development are presented, as well as those that are already an integral part of the production chain today, for example in the automotive industry.

1.8 Summary and Questions

There is no single generally accepted definition of VR today. One can approach the term from a technology-centered perspective and understand it to mean computer systems that build immersive and interactive environments using appropriate hardware, such as stereo displays. But VR can also be described as a methodology to give users the experience of inclusion in an alternative reality. The goal is not necessarily to achieve a perfect Virtual Reality that can no longer be distinguished from reality. Peculiarities of human perception and cognition such as the suspension of disbelief can be exploited to successfully create virtual environments for people and give them the feeling of presence in a VR. This can serve different purposes: research (e.g., about human perception), education, entertainment, communication support, visualization of simulation results or economic goals (e.g., prototyping to increase efficiency or save costs). The basic purpose of VR is to create an innovative interface between humans and computers. The idea of leaving users present in reality, but extending it with parts from a virtual world, leads to Augmented Reality. For the realization of virtual or augmented environments a virtual world and a VR/AR system are required. The virtual world provides the content to be shown in the environment (e.g., description of the geometry, appearance, and behavior of the virtual objects occurring in it). With regard to the VR/AR system, a computer system needs to be implemented that comprises the essential components for the collection of information about the users and their interactions (e.g., by tracking), the generation of stimuli for the user (e.g., images and sounds) as well as the simulation of the virtual world. Despite its more than 50 years of existence, VR/AR is still a young science. Four generations can be distinguished in its development, which can be characterized by the hardware used: (1) HMD and data glove, stereo projection and optical tracking, (2) high-resolution displays and low-cost tracking without the use of artificial markers, (3) consumer HMD including tracking and controllers, and (4) AR on smartphones and tablets.

Check your understanding of the chapter by answering the following questions:

- What would your definitions of the terms “virtual reality”, “virtual world”, “virtual environment”, “augmented reality”, “mixed reality”, “immersion”, “presence”, “simulation”, “tracking”, “user”, “human-machine interaction” and “suspension of disbelief” be?
- The text describes a scenario in which a user stands on a glass plate that is used as a projection screen. This gave the user the impression of standing on a virtual high-rise building where the user could see their real feet. Is this scenario an example of VR or AR?
- Suppose you want to create a jogging app where you run against other runners (or even yourself the day before). Would you implement this with VR or AR? What might this depend on? What would your environment look like? Which hardware would you use for this?
- What can VR and AR be used for? Which application examples do you know, or can you imagine? Why are you interested in VR/AR?

Recommended Reading

Angel E, Shreiner D (2015) *Interactive computer graphics: a top-down approach with WebGL*. Pearson Education, Harlow – *Textbook covering the basics of computer graphics, e.g., discussing the generation of images with the computer. It also introduces OpenGL and WebGL, a programming library for computer graphics, and discusses the possibilities of using graphics processors (GPUs) in the form of so-called shaders.*

Rabin S (2009) *Introduction to game development*, 2nd edition. Charles River Media, Boston – *a standard work on computer games. Due to the manifold points of contact between VR and computer games, literature from the field of computer games is also relevant.*

Original scientific literature can be found in specialist journals and conference proceedings which can be researched and accessed in digital libraries (e.g., dl.acm.org, ieeexplore.org, link.springer.com) or via search engines (e.g. scholar.google.com). In the field of VR the IEEE VR Conference (ieeevr.org) takes place annually. Moreover, there is the Eurographics Symposium on Virtual Environments (EGVE) as well as the VR Conferences of euroVR, which are partly jointly organized as Joint Virtual Reality Conference (JVRC). With the focus on AR, ISMAR, the IEEE Symposium for Mixed and Augmented Reality, is held annually. In addition, there are special events that focus on aspects of user interfaces of VR and AR, such as the ACM VRST conference or the 3DUI, the IEEE Symposium for 3D User Interfaces. There are also further events dealing with special applications of VR, for instance in the industrial sector (e.g., VRCAI – ACM International Conference on Virtual Reality Continuum and Its Applications in Industry). Some scientific journals also focus on VR and AR, e.g., Presence – Teleoperators and Virtual Environments by MIT Press, Virtual Reality by Springer Verlag or the Journal of Virtual Reality and Broadcasting (jVRb) as an open access e-journal.

In addition to conference proceedings and professional journals that deal primarily with VR and AR, literature is also recommended that deals with essential aspects of VR and AR, such as Computer Graphics (e.g., ACM SIGGRAPH and the ACM Transactions on Graphics), Computer Vision (e.g., IEEE ICCV) or Human–Machine Interaction (e.g. ACM SIGCHI).

References

Azuma R (1997) A survey of augmented reality. Presence Teleop Virt 6(4):355–385

Bell B, Feiner S, Hoellerer T (2002) Information at a glance. IEEE Comp Gr Appl 22(4)., July/August, 6–9

Bricken W (1990) Virtual reality: directions of growth. Notes SIGGRAPH '90 Panel (HITL Technical Report R-90-1), University of Washington, Seattle

Bryson S (2013). Virtual Reality: a definition history – a personal essay. *ArXiv*, abs/1312.4322

Caudell TP, Mizell DW (1992) Augmented reality: an application of heads-up display technology to manual manufacturing processes. In: Proceedings of 25th Hawaii International conference on system sciences, Vol. 2, 659–669
Feiner S, MacIntyre B, Höllerer T (1997) A touring machine: prototyping 3D mobile augmented reality systems for exploring the urban environment, digest of papers. In: First International Symposium on Wearable Computers, pp 74–81
Fuchs H, Livingston MA, Raskar R, Colucci D, Keller K, State A, Crawford JR, Rademacher P, Drake SH, Meyer AA (1998) Augmented reality visualization for laparoscopic surgery. In: Wells WM, Colchester A, Delp S (eds) Medical image computing and computer-assisted intervention — MICCAI'98, LNCS, vol 1496. Springer, Berlin/Heidelberg
Held RH, Durlach NI (1992) Telepresence. Presence Teleop Virt 1(1):109–112
Kato H, Billinghurst M (1999) Marker tracking and HMD calibration for a video-based augmented reality conferencing system. In: 2nd IEEE and ACM international workshop on augmented reality (IWAR), pp. 85–94, IEEE
Mann S (1999) Mediated Reality. Linux Journal, Article No 5, Issue 59
Mann S, Furness T, Yuan Y, Iorio J, Wang Z (2018) All Reality: Virtual, Augmented, Mixed (X), Mediated (X,Y), and Multimediated Reality. https://arxiv.org/abs/1804.08386
Milgram P, Takemura H, Utsumi A, Kishino F (1995) Augmented reality: a class of displays on the reality-virtuality continuum. Proc SPIE 2351:282–292
Mine MR, Brooks Jr FP, Sequin CH (1997) Moving objects in space: Exploiting proprioception in virtual-environment interaction. In: Proceedings of SIGGRAPH 1997, pp 19–26
Paradiso JA, Landay JA (2009) Guest editors' introduction: cross-reality environments. IEEE Perv Computing 8(3):14–15
Rekimoto J, Nagao K (1995) The world through the computer: computer augmented interaction with real world environments. In: Proceedings of UIST '95, pp 29–36
Rheingold H (1991) Virtual reality. Summit Books, New York
Sadowski W, Stanney KM (2002) Presence in virtual environments. In: Stanney KM (ed) Handbook of virtual environments: design, implementation, and applications. Lawrence Erlbaum Associates Inc, Mahwah
Sheridan TB (1992) Musings on telepresence and virtual presence. Presence Teleop Virt 1(1):120–125
Sherman W, Craig A (2003) Understanding virtual reality. Morgan Kaufmann, San Mateo
Slater M (2003) A note on presence terminology. Presence Connect 3:3
Slater M (2009) Place illusion and plausibility can lead to realistic behaviour in immersive virtual environments. Phil Trans R Soc B 364(1535):3549–3557
Slater M, Wilbur S (1997) A framework for immersive virtual environments (FIVE): speculations on the role of presence in virtual environments. Presence Teleop Virt 6(6):603–616
Slater M, Spanlang B, Corominas D (2010) Simulating virtual environments within virtual environments as the basis for a psychophysics of presence. ACM Trans Graph 29(4):92
Stone RJ (1993) In: Earnshaw RA, Gigante MA, Jones H (eds) Virtual reality systems. London, Academic
Sutherland IE (1965) The ultimate display. Proceedings of the IFIP congress, 506–508
Sutherland I (1968) A head mounted three dimensional display. In: Proceedings of the AFIPS fall joint computer conference. Thompson Books, Washington, DC, pp 757–764
Witmer BG, Singer MJ (1998) Measuring presence in virtual environments: a presence questionnaire. Presence Teleop Virt Environ 7(3):225–240