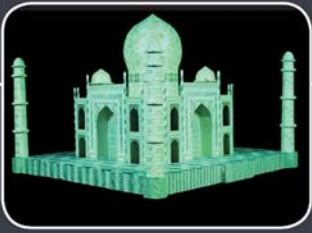
Oliver Bimber
Ramesh Raskar

Spatial Augmented Reality

Merging Real and Virtual Worlds



Spatial Augmented Reality

Spatial Augmented Reality

Merging Real and Virtual Worlds

Oliver Bimber
Bauhaus-University, Weimar

Ramesh Raskar
Mitsubishi Electric Research Laboratory,
Cambridge, MA

A K Peters
Wellesley, Massachusetts

CRC Press
Taylor & Francis Group
6000 Broken Sound Parkway NW, Suite 300
Boca Raton, FL 33487-2742

Version Date: 20150213

International Standard Book Number-13: 978-1-4398-6494-4 (eBook - PDF)

Visit the Taylor & Francis Web site at
http://www.taylorandfrancis.com

and the CRC Press Web site at
http://www.crcpress.com

To Mel

—O. B.

To my parents

—R. R.

Contents

Preface xi

1 A Brief Introduction to Augmented Reality 1
- 1.1 What is Augmented Reality . . . 1
- 1.2 Today's Challenges . . . 4
- 1.3 Spatial Augmented Reality . . . 7
- 1.4 Outline of the Book . . . 8

2 Fundamentals: From Photons to Pixels 13
- 2.1 Light in a Nutshell . . . 14
- 2.2 Geometric Optics . . . 17
- 2.3 Visual Depth Perception . . . 30
- 2.4 Rendering Pipelines . . . 44
- 2.5 Summary and Discussion . . . 66

3 Augmented Reality Displays 71
- 3.1 Head-Attached Displays . . . 72
- 3.2 Hand-Held Displays . . . 79
- 3.3 Spatial Displays . . . 83
- 3.4 Summary and Discussion . . . 90

4 Geometric Projection Concepts 93
- 4.1 Geometric Model . . . 93
- 4.2 Rendering Framework . . . 98
- 4.3 Calibration Goals . . . 104
- 4.4 Display Environments and Applications . . . 108
- 4.5 Summary and Discussion . . . 108

5 Creating Images with Spatial Projection Displays 111
5.1 Planar Displays 111
5.2 Non-Planar Display 126
5.3 Projector Overlap Intensity Blending 129
5.4 Quadric Curved Displays 133
5.5 Illuminating Objects 142
5.6 Summary and Discussion 147

6 Generating Optical Overlays 149
6.1 Transparent Screens 150
6.2 Mirror Beam Combiners 152
6.3 Planar Mirror Beam Combiners 152
6.4 Screen Transformation and Curved Screens 163
6.5 Moving Components 166
6.6 Multi-Plane Beam Combiners 166
6.7 Curved Mirror Beam Combiners 174
6.8 Summary and Discussion 206

7 Projector-Based Illumination and Augmentation 213
7.1 Image-Based Illumination: Changing Surface Appearance . 214
7.2 Creating Consistent Occlusion 220
7.3 Creating Consistent Illumination 227
7.4 Augmenting Optical Holograms 244
7.5 Augmenting Flat and Textured Surfaces 255
7.6 Augmenting Geometrically Non-Trivial Textured Surfaces . 267
7.7 Summary and Discussion 276

8 Examples of Spatial AR Displays 279
8.1 Shader Lamps 280
8.2 Being There 286
8.3 iLamps: Mobile Projectors 288
8.4 The Extended Virtual Table 291
8.5 The Virtual Showcase 297
8.6 The HoloStation 302
8.7 Augmented Paintings 308
8.8 Smart Projectors 314
8.9 Summary and Discussion 320

9 The Future 321
9.1 Displays 321
9.2 Supporting Elements 326
9.3 Summary and Discussion 329

A Calibration of a Projector (or a Camera) 331
A.1 Source code for Calibration of a Projector (or a camera) . . 331
A.2 Quadric Image Transfer 334

B OpenGL's Transformation Pipeline Partially Re-Implemented 337
B.1 General Definitions 337
B.2 Projection Functions 337
B.3 Transformation Functions 338
B.4 Additional Functions 340

Bibliography 343

Preface

Spatial Augmented Reality is a rapidly emerging field which concerns everyone working in digital art and media who uses any aspects of augmented reality and is interested in cutting-edge technology of display technologies and the impact of computer graphics. We believe that a rich pallet of different display technologies, mobile and non-mobile, must be considered and adapted to fit a given application so that one can choose the most efficient technology. While this broader view is common in the very established area of virtual reality, it is becoming more accepted in augmented reality which has been dominated by research involving mobile devices.

This book reflects our research efforts over several years and the material has been refined in several courses that we taught at the invitation of Eurographics and ACM SIGGRAPH.

Who Should Read This Book

In order for a broad spectrum of readers—system designers, programmers, artists, etc— to profit from the book, we require no particular programming experience or mathematical background. However, a general knowledge of basic computer graphics techniques, 3D tools, and optics will be useful.

The reader will learn about techniques involving both hardware and software to implement spatial augmented reality installations. Many Cg and OpenGL code fragments, together with algorithms, formulas, drawings, and photographs will guide the interested readers who want to experiment with their own spatial augmented reality installations.

By including a number of exemplary displays examples from different environments, such as museums, edutainment settings, research projects, and industrial settings, we want to stimulate our readers to imagine novel

AR installations and to implement them. Supplementary material can be found at http://www.spatialar.com/.

About the Cover

The images at the top of the front cover show a rainbow hologram of a dinosaur (Deinonychus) skull (found in North America). It has been augmented with reconstructed soft tissue and artificial shading and occlusion effects. The soft tissue data, provided by Lawrence M. Witmer of Ohio University, were rendered autostereoscopically. A replica of the skull was holographed by Tim Frieb at the Holowood holographic studio in Bamberg, Germany. The hologram was reconstructed by projected digital light that could be controlled and synchronized to the rendered graphics. This enabled a seamless integration of interactive graphical elements into optical holograms.

The image at the bottom show an example of Shader Lamps: an augmentation of a white wooden model of the Taj Mahal with two projectors. The wooden model was built by Linda Welch, George Spindler and Marty Spindler in the late 1970s. In 1999, at the University of North Carolina, Greg Welch spray painted the wooden model white and Kok-Lim Low scanned it with a robotic arm to create a 3D model. The wooden model is shown illuminated with images rendered with real time animation of a sunrise.

The art work on the back cover was created by Matthias Hanzlik. The sketches show early concepts of SAR prototypes. They have all been realized and are described in the book.

Acknowledgements

The material presented in this book would not have been realized without the help and dedication of many students and colleagues. We want to thank them first of all: Gordon Wetzstein, Anselm Grundhöfer, Sebastian Knödel, Franz Coriand, Alexander Kleppe, Erich Bruns, Stefanie Zollmann, Tobias Langlotz, Mathias Möhring, Christian Lessig, Sebastian Derkau, Tim Gollub, Andreas Emmerling, Thomas Klemmer, Uwe Hahne, Paul Föckler, Christian Nitschke, Brian Ng, Kok-Lim Low, Wei-Chao Chen, Michael Brown, Aditi Majumder, Matt Cutts, Deepak Bandyopadhyay, Thomas Willwacher, Srinivas Rao, Yao Wang, and Johnny Lee.

We also want to thank all colleagues in the field who provided additional image material: Vincent Lepetit, Pascal Fua, Simon Gibson, Kiyoshi Kiyokawa, Hong Hua, Susumu Tachi, Daniel Wagner, Dieter Schmalstieg, George Stetten, Tetsuro Ogi, Barney Kaelin, Manfred Bogen, Kok-Lim Low, Claudio Pinhanez, Masahiko Inami, Eric Foxlin, Michael Schnaider, Bernd Schwald, and Chad Dyner.

Special thanks go to those institutions and people who made this research possible: Bauhaus-University Weimar, Mitsubishi Electric Research Laboratories, Jeroen van Baar, Paul Beardsley, Paul Dietz, Cliff Forlines, Joe Marks, Darren Leigh, Bill Yerazunis, Remo Ziegler, Yoshihiro Ashizaki, Masatoshi Kameyama, Keiichi Shiotani, Fraunhofer Institute for Computer Graphics, José L. Encarnação, Bodo Urban, Erhard Berndt, University of North Carolina at Chapel Hill, Henry Fuchs, Greg Welch, Herman Towles, Fraunhofer Center for Research in Computer Graphics, Miguel Encarnação, Bert Herzog, David Zeltzer, Deutsche Forschungsgemeinschaft, and European Union.

We thank the Institution of Electronics & Electrical Engineers (IEEE), the Association for Computing Machinery (ACM), Elsevier, MIT Press, Springer-Verlag, and Blackwell Publishing for granting permission to republish some of the material used in this book.

Many thanks to Thomas Zeidler and to Alice and Klaus Peters for helping us manage the process of writing this book alongside our everyday duties.

Oliver Bimber
Weimar April 2005

Ramesh Raskar
Cambridge, MA April 2005

1

A Brief Introduction to Augmented Reality

Like *Virtual Reality* (VR), *Augmented Reality* (AR) is becoming an emerging edutainment platform for museums. Many artists have started using this technology in semi-permanent exhibitions. Industrial use of augmented reality is also on the rise. Some of these efforts are, however, limited to using off-the-shelf head-worn displays. New, application-specific alternative display approaches pave the way towards flexibility, higher efficiency, and new applications for augmented reality in many non-mobile application domains. Novel approaches have taken augmented reality beyond traditional eye-worn or hand-held displays, enabling new application areas for museums, edutainment, research, industry, and the art community. This book discusses *spatial augmented reality* (SAR) approaches that exploit large optical elements and video-projectors, as well as interactive rendering algorithms, calibration techniques, and display examples. It provides a comprehensive overview with detailed mathematics equations and formulas, code fragments, and implementation instructions that enable interested readers to realize spatial AR displays by themselves.

This chapter will give a brief and general introduction into augmented reality and its current research challenges. It also outlines the remaining chapters of the book.

1.1 What is Augmented Reality

The terms *virtual reality* and *cyberspace* have become very popular outside the research community within the last two decades. Science fiction movies, such as *Star Trek*, have not only brought this concept to the public, but have also influenced the research community more than they are willing to admit. Most of us associate these terms with the technological possibility to dive into a completely synthetic, computer-generated world—sometimes referred to as a *virtual environment*. In a virtual environment our senses, such as vision, hearing, haptics, smell, etc., are controlled by a computer while our actions influence the produced stimuli. *Star Trek*'s Holodeck is probably one of the most popular examples. Although some bits and pieces of the Holodeck have been realized today, most of it is still science fiction.

So what is augmented reality then? As is the case for virtual reality, several formal definitions and classifications for augmented reality exist (e.g., [109, 110]). Some define AR as a special case of VR; others argue that AR is a more general concept and see VR as a special case of AR. We do not want to make a formal definition here, but rather leave it to the reader to philosophize on their own. The fact is that in contrast to traditional VR, in AR the real environment is not completely suppressed; instead it plays a dominant role. Rather than immersing a person into a completely synthetic world, AR attempts to embed synthetic supplements into the real environment (or into a live video of the real environment). This leads to a fundamental problem: a real environment is much more difficult to control than a completely synthetic one. Figure 1.1 shows some examples of augmented reality applications.

As stated previously, augmented reality means to integrate synthetic information into the real environment. With this statement in mind, would a TV screen playing a cartoon movie, or a radio playing music, then be an AR display? Most of us would say no—but why not? Obviously, there is more to it. The augmented information has to have a much stronger link to the real environment. This link is mostly a spatial relation between the augmentations and the real environment. We call this link *registration*. R2-D2's spatial projection of Princess Leia in *Star Wars* would be a popular science fiction example for augmented reality. Some technological approaches that mimic a holographic-like spatial projection, like the Holodeck, do exist today. But once again, the technical implementation as shown in *Star Wars* still remains a Hollywood illusion.

Some say that Ivan Sutherland established the theoretical foundations of virtual reality in 1965, describing what in his opinion would be the *ultimate display* [182]:

> 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. Hand-

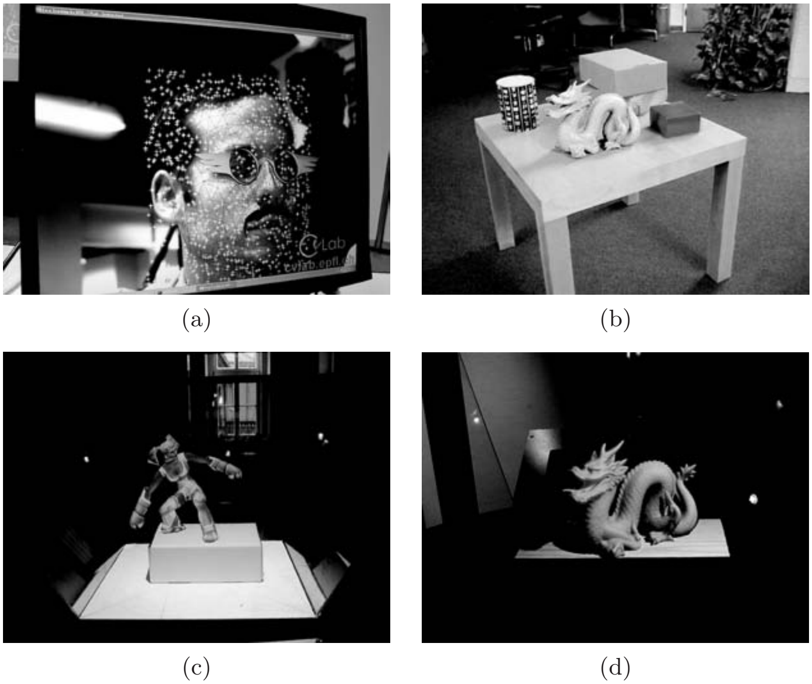


Figure 1.1. Example of augmented reality applications. The glasses, mustache, dragons, and fighter figure are synthetic: (a) and (b) augmenting a video of the real environment; (c) and (d) augmenting the real environment optically. (*Images: (a) courtesy of Vincent Lepetit, EPFL [87]; (b) courtesy of Simon Gibson [55], Advanced Interfaces Group © University of Manchester 2005; (c) and (d) prototypes implemented by the Barhaus-University Weimar.*)

> cuffs 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.

However, technical virtual reality display solutions were proposed much earlier. In the late 1950s, for instance, a young cinematographer named Mort Heilig invented the *Sensorama* simulator, which was a one-person demo unit that combined 3D movies, stereo sound, mechanical vibrations, fan-blown air, and aromas. Stereoscopy even dates back to 1832 when Charles Wheatstone invented the *stereoscopic viewer*.

Then why did Sutherland's suggestions lay the foundation for virtual reality? In contrast to existing systems, he stressed that the user of such an ultimate display should be able to interact with the virtual environment. This led him to the development of the first functioning *Head-Mounted Display* (HMD) [183], which was also the birth of augmented reality. He used half-silvered mirrors as optical combiners that allowed the user to see both the computer-generated images reflected from *cathode ray tubes* (CRTs) and objects in the room, simultaneously. In addition, he used mechanical and ultrasonic head position sensors to measure the position of the user's head. This ensured a correct registration of the real environment and the graphical overlays.

The interested reader is referred to several surveys [4, 5] and Web sites [3, 193] of augmented reality projects and achievements. Section 1.2 gives a brief overview of today's technical challenges for augmented reality. It is beyond the scope of this book to discuss these challenges in great detail.

1.2 Today's Challenges

As mentioned previously, a correct and consistent registration between synthetic augmentations (usually three-dimensional graphical elements) and the real environment is one of the most important tasks for augmented reality. For example, to achieve this for a moving user requires the system to continuously determine the user's position within the environment.

Thus the *tracking* and *registration* problem is one of the most fundamental challenges in AR research today. The precise, fast, and robust tracking of the observer, as well as the real and virtual objects within the environment, is critical for convincing AR applications. In general, we can differentiate between *outside-in* and *inside-out tracking* if absolute tracking within a global coordinate system has to be achieved. The first type, outside-in, refers to systems that apply fixed sensors within the environment that track emitters on moving targets. The second type, inside-out, uses sensors that are attached to moving targets. These sensors are able to determine their positions relative to fixed mounted emitters in the environment. Usually these two tracking types are employed to classify camera-based approaches only—but they are well suited to describe other tracking technologies as well.

After *mechanical* and *electromagnetic tracking*, *optical tracking* became very popular. While *infrared* solutions can achieve a high precision and a high tracking speed, *marker-based tracking*, using conventional cameras, represent a low-cost option. Tracking solutions that do not require artificial markers, called *markerless tracking*, remains the most challenging, and at the same time, the most promising tracking solution for future augmented reality applications. Figure 1.1(a) shows an example of a markerless face tracker.

Much research effort is spent to improve performance, precision, robustness, and affordability of tracking systems. High-quality tracking within large environments, such as the outdoors, is still very difficult to achieve even with today's technology, such as a *Global Positioning System* (GPS) in combination with relative measuring devices like *gyroscopes* and *accelerometers*. A general survey on different tracking technology [164] can be used for additional reading.

Besides tracking, *display technology* is another basic building block for augmented reality. As mentioned previously, head-mounted displays are the dominant display technology for AR applications today. However, they still suffer from optical (e.g., limited field of view and fixed focus), technical (e.g., limited resolution and unstable image registration relative to eyes) and human-factor (e.g., weight and size) limitations. The reason for this dominance might be the long time unique possibility of HMDs to support *mobile AR applications*. The increasing technological capabilities of *cell phones* and *Personal Digital Assistants* (PDAs), however, clear the way to more promising display platforms in the near future. In addition, not all AR applications require mobility. In these cases, *spatial display configurations* are much more efficient.

The third basic element for augmented reality is *real-time rendering*. Since AR mainly concentrates on superimposing the real environment with graphical elements, fast and realistic rendering methods play an important role. An ultimate goal could be to integrate graphical objects into the real environment in such a way that the observer can no longer distinguish between real and virtual. Note that not all AR applications really make this requirement. But if so, then besides perfect tracking and display technologies, *photo-realistic real-time rendering* would be another requisite. Graphical objects, even if rendered in a high visual quality, would have to be integrated into the real environment in a consistent way. For instance, they have to follow a consistent *occlusion*, *shadow-casting*, and *inter-reflection* behavior, as Figure 1.1 demonstrates.

Realistic, non-real-time capable *global illumination* techniques, such as ray-tracing or *radiosity*, can be used if no interactive frame rates are required, . But for interactive applications, faster image generation methods have to be used to avoid a large *system lag* and a resultant *misregistration* after fast user motions. The improving *hardware acceleration* of today's

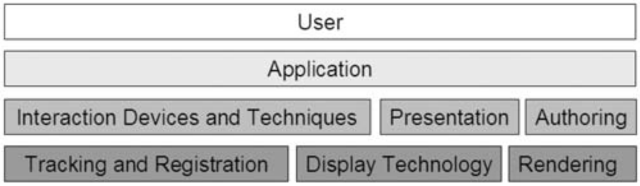


Figure 1.2. Building blocks for augmented reality.

graphics cards make a lot possible, as is shown throughout the following chapters. The ongoing paradigm shift of the computer graphics community from the old *fixed function rendering pipelines* to *programmable pipelines* strongly influences the rendering capabilities of AR applications. Examples of consistent rendering techniques for augmented reality have been discussed in different scientific publications and in the course of several conference tutorials [176, 17].

Figure 1.2 illustrates some general building blocks for augmented reality. As we can see, the previously discussed challenges (tracking and registration, display technology and rendering) represent fundamental components. On top of this base level, more advanced modules can be found: *interaction devices* and *techniques*, *presentation*, and *authoring*. If we take a comparative look at virtual reality again, we can see that the base technology of today's VR is much more mature. In contrast to VR where a large portion of research is now being shifted to the second layer, the AR community still has to tackle substantial problems on the base level.

Ideas and early implementations of presentation techniques, authoring tools, and interaction devices/techniques for AR applications are just emerging. Some of them are derived from the existing counterparts in related areas such as virtual reality, multimedia, or digital storytelling. Others are new and adapted more to the problem domain of augmented reality. However, it is yet too early to spot matured concepts and philosophies at this level.

The third layer, the *application*, is finally the interface to the user. Using augmented reality, our overall goal is to implement applications that are tools which allow us to solve problems more effectively. Consequently, augmented reality is no more than a human-computer interface which has the potential to be more efficient for some applications than others. Although many ideas for possible applications of this interface exist, not many have actually become applicable today. One reason for this is the immature base layer. With a stable core technology, augmented reality does have the potential to address many application areas more effectively. Some virtual reality applications, for instance, have already managed to become real tools outside the research community. It is also clear, that broader base levels will lead to a broader application spectrum.

Some *software frameworks* (e.g., [165]) are being realized that comprise several of these parts. Good software engineering will be important for the efficient handling of an increasing pallet of new tools and techniques. Finally, user studies have to be carried out to provide measures of how effective augmented reality really is.

1.3 Spatial Augmented Reality

The roots of virtual reality and augmented reality are not that far apart. After almost forty years of research and development, however, they do not follow the same technological paths anymore. In the early 1990s, *projection-based surround screen displays* became popular. One of the most well-known is the *CAVE* [35]—a multi-sided, immersive projection room. But there are other examples of *semi-immersive* wall-like and table-like displays or *immersive* cylindrical and spherical *spatial displays*. In general, spatial displays detach the display technology from the user and integrate it into the environment. Compared to head- or body-attached displays, spatial displays offer many advantages and solve several problems that are related to visual quality (e.g., resolution, field-of-view, focus, etc.), technical issues (e.g., tracking, lighting, etc.), and human factors (e.g., cumbersomeness, etc.), but they are limited to non-mobile applications..

The virtual reality community has oriented themselves away from head-mounted displays and towards spatial displays. Today, a large variety of spatial displays make up an estimated 90% of all VR displays. Head-mounted displays, however, are still the dominant displays for augmented reality. The reason for this might lie in the strong focus of mobile AR applications—requiring mobile displays.

Video see-through and optical see-through head-mounted displays have been the traditional output technologies for augmented reality applications for almost forty years. However, they still suffer from several technological and ergonomic drawbacks which prevent them from being used effectively in many application areas. In an all-purpose context, HMDs are used in many non-mobile AR applications. This affects the efficiency of these applications and does not currently allow them to expand beyond laboratory demonstrations. In the future, other mobile devices, such as cell phones or PDAs might replace HMDs in many mobile areas. Head-mounted displays will also be enhanced by future technology, leading to a variety of new and different possibilities for mobile AR.

Furthermore, we believe that for non-mobile applications a rich pallet of different spatial display configurations can be as beneficial for augmented reality, as they have been for virtual reality. Novel approaches have taken augmented reality beyond traditional eye-worn or hand-held displays enabling additional application areas. New display paradigms exploit large spatially-aligned optical elements, such as mirror beam combiners, transparent screens, or holograms, as well as video projectors. Thus, we call this technological variation *spatial augmented reality* (SAR). In many situations, SAR displays are able to overcome technological and ergonomic limitations of conventional AR systems. Due to the decrease in cost and availability of projection technology, personal computers, and graphics hardware, there has been a considerable interest in exploiting SAR systems in universities, research laboratories, museums, industry, and the art community. Parallels to the development of virtual environments from head-attached displays to spatial projection screens can be clearly drawn. We believe that an analog evolution of augmented reality has the potential to yield a similar successful factor in many application domains. Thereby, SAR and body-attached AR are not competitive, but complementary.

1.4 Outline of the Book

This book provides survey and implementation details of modern techniques for spatial augmented reality systems and aims to enable the interested reader to realize such systems on his or her own. This is supported by more than 200 illustrations and many concrete code fragments.

After laying foundations in optics, interactive rendering, and perspective geometry, we discuss conventional mobile AR displays and present spatial augmented reality approaches that are overcoming some of their limitations. We present state-of-the-art concepts, details about hardware and software implementations, and current areas of application in domains such as museums, edutainment, research, and industrial areas. We draw parallels between display techniques used for virtual reality and augmented reality and stimulate thinking about the alternative approaches for AR.

One potential goal of AR is to create a high level of consistency between real and virtual environments. This book describes techniques for the optical combination of virtual and real environments using mirror beam combiners, transparent screens, and holograms. It presents projector-based augmentation of geometrically complex and textured display surfaces, which along with optical combiners achieve consistent illumination and occlusion effects. We present many spatial display examples, such as Shader Lamps, Virtual Showcases, Extended Virtual Tables, interactive holograms, apparent motion, Augmented Paintings, and Smart Projectors.

Finally, we discuss the current problems, future possibilities, and enabling technologies of spatial augmented reality.

Chapter 2 lays the foundation for the topics discussed in this book. It starts with a discussion on light. The atomic view on light will give some hints on how it is generated—knowing about its properties allows us to realize how it travels through space.

From electromagnetic waves, a more geometric view on light can be abstracted. This is beneficial for describing how simple optical elements, such as mirrors and lenses, work. This chapter will describe how images are formed by bundling real and virtual light rays in a single spatial spot or area in three-dimensional space. Furthermore, it is explained that the structure and functionality of the human eye (as the final destination of visible light produced by a display) is as complex as an optical system itself, and that the binocular interplay of two eyes leads to visual depth perception. The depth perception can be tricked by viewing flat stereo images on a stereoscopic display. The principles of stereoscopic vision and presentation, as well as a classification of stereoscopic and autostereoscopic displays, will be discussed in this chapter as well. We will illustrate how images that are presented on stereoscopic displays are computed. Basic rendering concepts, such as components of traditional fixed function rendering pipelines and techniques like multi-pass rendering, but also modern programmable rendering pipelines will be described.

Chapter 3 classifies current augmented reality displays into head-attached, hand-held, and spatial displays. It gives examples of particular displays that are representative for each class and discusses their advantages and disadvantages. Retinal displays, video see-through and optical see-through head-mounted displays, and head-mounted projectors are presented first in the context of head-attached displays. For hand-held displays, personal digital assistants, cell phones, hand-held projectors, and several optical see-through variations are outlined. Finally, spatial AR displays are presented. First examples include screen-based video see-through displays, spatial optical see-through displays, and projector-based spatial displays. The following chapters will explain how to realize such displays, both from a hardware and software point of view.

Chapter 4 reviews the fundamental geometric concepts in using a projector for displaying images. A projector can be treated as a dual of a camera. The geometric relationship between the two-dimensional pixels in the projector frame buffer and the three-dimensional points in the world illuminated by those pixels is described using perspective projection. The chapter introduces a general framework to express the link between geometric components involved in a display system. The framework leads to a simpler rendering technique and a better understanding of the calibration goals. We describe the procedure to calibrate projectors and render images for planar as well non-planar displays along with issues in creating seamless images using multiple projectors.

Chapter 5 expands on the geometric framework introduced in the previous chapter and describes the concrete issues in calibration and rendering for various types of display surfaces. The techniques use parametric as well as non-parametric approaches. For planar surface, a procedure based on homography transformation is effective. For arbitrary non-planar displays, we outline a two-pass scheme, and for curved displays, we describe a scheme based on quadric image transfer. Finally, the chapter discusses the specific projector-based augmentation problem where images are projected not on display screens, but directly onto real-world objects.

Chapter 6 explains spatial optical see-through displays in detail. An essential component of an optical see-through display is the optical combiner—an optical element that mixes the light emitted by the illuminated real environment with the light produced with an image source that displays the rendered graphics. Creating graphical overlays with spatial optical see-through displays is similar to rendering images for spatial projection screens for some optical combiners. For others, however, it is more complex and requires additional steps before the rendered graphics are displayed and optically combined. While monitors, diffuse projection screens, or video projectors usually serve as light emitting image sources, two different types of optical combiners are normally used for such displays: transparent screens and half-silvered mirror beam combiners. Rendering techniques that support creating correct graphical overlays with both types of optical combiners and with different images sources will be discussed in this chapter. In particular, Chapter 6 will discuss rendering techniques for spatial optical see-through displays which apply transparent screens, as well as planar and curved mirror beam combiners, in many different configurations. We describe how optical effects, such as reflection and refraction, can be neutralized by hardware-accelerated rendering techniques and present building blocks that can easily be integrated into an existing software framework.

Chapter 7 outlines interactive rendering techniques for projector-based illumination and augmentation. It starts with an overview of methods that allow augmenting artificial surface appearance, such as shading, shadows, and highlights, of geometrically non-trivial objects. Calibrated projectors are used to create these effects on physical objects with uniformly white surface color. We then describe how to use calibrated projectors in combination with optical see-through configurations (described in Chapter 6) for the illumination of arbitrary real objects. This allows digitization of the illumination of the real environment and synchronization with the rendering process that generates the graphical overlays.

The final goal is to create consistent occlusion effects between real and virtual objects in any situation and to support single or multiple observers. The surface appearance of real objects with non-trivial surface color and texture can also be modified with a projector-based illumination. Such an approach, for instance, allows the creation of consistent global or local lighting situations between real and virtual objects. Appropriate rendering techniques are also described in this chapter. A projector-based illumination also makes it possible to integrate graphical augmentations into optical holograms. In this case, variations of the algorithms explained previously let us replace a physical environment by a high-quality optical hologram. Finally, this chapter presents real-time color-correction algorithms that, in combination with an appropriate geometric correction, allow an augmentation of arbitrary (colored/textured) three-dimensional surfaces with computer generated graphics.

Chapter 8 brings together the previous, more technical chapters in an application-oriented approach and describes several existing spatial AR display configurations. It first outlines examples that utilize the projector-based augmentation concept in both a small desktop approach (e.g., Shader Lamps) and a large immersive configuration (e.g., the *Being There* project). In addition, an interactive extension, called *iLamps*, that uses hand-held projectors is described. Furthermore, several spatial optical see-through variations that support single or multiple users, such as the *Extended Virtual Table* and the *Virtual Showcase* are explained. It is shown how they can be combined with projector-based illumination techniques to present real and virtual environments consistently. A scientific workstation, the *HoloStation*, is presented which combines optical hologram records of fossils with interactive computer simulations. Finally, two configurations (*Augmented Paintings* and *Smart Projectors*) are presented. They use real-time color correction and geometric warping to augment artistic paintings with multimedia presentations, as well as to make projector-based home-entertainment possible without artificial canvases.

Potential application areas for the display configurations described in this chapter are industrial design and visualization (e.g., Shader Lamps, iLamps, Extended Virtual Table), scientific simulations (e.g., HoloStation), inertial design and architecture (e.g., Being There), digital storytelling and next-generation edutainment tools for museums (e.g., Virtual Showcase and Augmented Paintings), and home-entertainment (e.g., Smart Projector). However, the interested reader can easily derive further application domains, such as those in an artistic context.

Another goal of this chapter is to show that spatial augmented reality display configurations can be applied successfully and efficiently outside research laboratories. The Virtual Showcase, for instance, has been presented to more than 120,000 visitors at more than 11 exhibitions in museums, trade shows, and conferences. Unattended running times of four months and more are an indicator for the fact that it is possible to make the technology (soft- and hardware) robust enough to be used by museums and other public places.

Chapter 9 postulates future directions in spatial augmented reality. Many new opportunities are based on emerging hardware components, and they are briefly reviewed. The chapter discusses innovative optics for displays, new materials such as light emitting polymers, promising developments in sensor networks including those using photosensors, and the excitement surrounding radio frequency identification tags. These technology developments will not only open new possibilities for SAR, but also for other AR display concepts, such as hand-held and head-attached displays.

2

Fundamentals: From Photons to Pixels

This chapter lays the foundation for the topics discussed in this book. It will not describe all aspects in detail but will introduce them on a level that is sufficient for understanding the material in the following chapters. We strongly encourage the reader to consult the secondary literature.

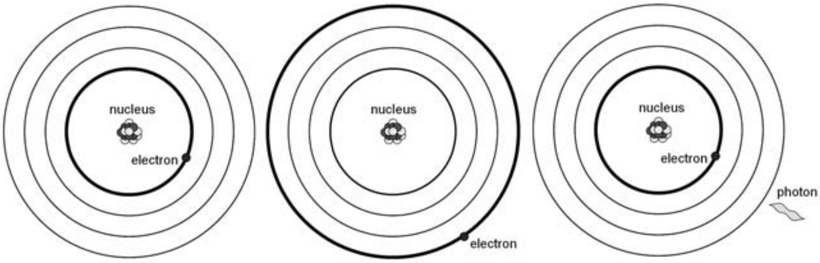
We start our journey at the most basic element that is relevant for all display technology—light. The atomic view on light will give us some hints on how it is generated. Knowing about its properties allows us to understand how it travels through space.

Starting from electromagnetic waves, we abstract our view on light to a more geometric concept that is beneficial for describing how simple optical elements, such as mirrors and lenses, work. We see how images are formed by bundling real and virtual light rays in a single spatial spot or area in three-dimensional space.

In addition, we learn that the structure and functionality of the human eye (as the final destination of visible light produced by a display) is as complex as an optical system itself, and that the binocular interplay of two eyes leads to visual depth perception.

Depth perception, however, can be tricked by viewing flat stereo images on a stereoscopic display. The principles of stereoscopic vision and presentation, as well as a classification of stereoscopic and autostereoscopic displays, will be discussed.

From stereoscopic displays, we will see how images presented on these displays are computed. Basic rendering concepts, such as components of traditional fixed function rendering pipelines and techniques like multi-pass rendering will be described. Modern programmable rendering pipelines will be discussed as well.



(a) (b) (c)

Figure 2.1. Planetary model of atom: (a) electron orbiting the nucleus in a non-excited state; (b) excited electron after quantum leap; (c) electron releasing a photon while dropping back into a non-excited state.

2.1 Light in a Nutshell

To explain what light is, we want to start at a very low level—at an atomic level. Illustrated by the well-known *planetary model* by Niels Bohr (1913), atoms consist of a *nucleus* and *electrons* that orbit the nucleus (Figure 2.1). They are held in the orbit by an electrical force. The nucleus itself consists of *protons* (having a positive electrical charge) and *neutrons* (having no electrical charge). The electrons have a negative electrical charge and can move from atom to atom. This flow is called *electricity*.

The level at which an electron orbits the nucleus (i.e., the distance of the electron to the nucleus) is called its *energy state*. By default, an electron exists at the lowest energy state—that is the orbit closest to the nucleus. If excited by external energy (e.g., heat) the electron can move from lower to higher energy states (i.e., further away from the nucleus). This shift from a lower to a higher energy state is called *quantum leap*. Since all energy is always preserved (it might be converted, but it is never lost), the electrons have to release energy when they drop back to lower energy states. They do so by releasing packages of energy. Albert Einstein called these packages *photons*.

Photons have a frequency that relates to the amount of energy they carry which, in turn, relates to the size of the drop from the higher state to the lower one. They behave like waves—they travel in waves with a specific *phase*, *frequency*, and *amplitude*, but they have no mass. These *electromagnetic waves* travel in a range of frequencies called electromagnetic (EM) spectrum, which was described by J. C. Maxwell (1864–1873).

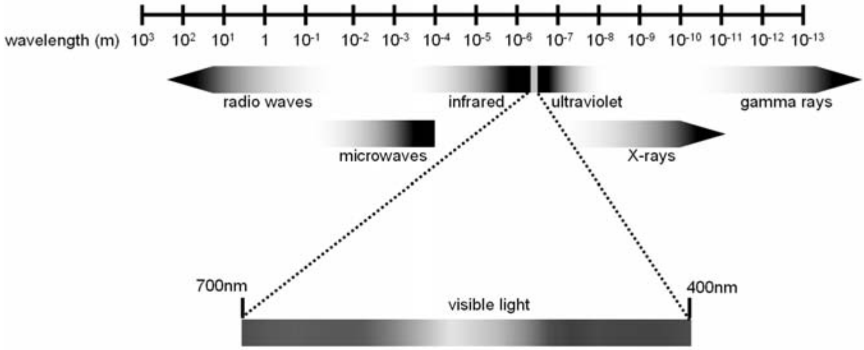


Figure 2.2. EM spectrum and spectrum of visible light. (*See Plate I.*)

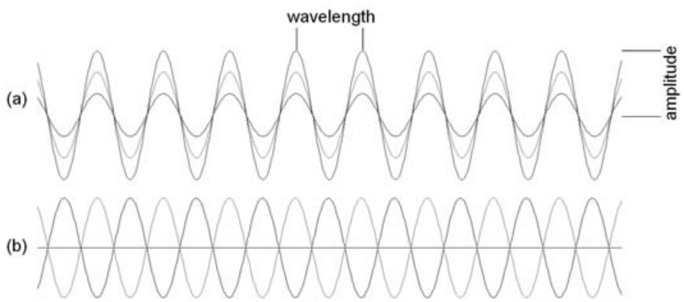


Figure 2.3. Interference of light waves: (a) amplification with zero phase shift, (b) cancellation with π phase shift.

A small part of this spectrum is the EM radiation that can be perceived as *visible light*.

Since light behaves like waves, it shares many properties of other waves. Thomas Young showed in the early 1800s that light waves can interfere with each other.

Depending on their phase, frequency, and amplitude, multiple light waves can amplify or cancel each other out (Figure 2.3). Light that consists of only one wavelength is called *monochromatic light*. Light waves that are in phase in both time and space are called *coherent*. Monochromaticity and low divergence are two properties of coherent light.

If a photon passes by an excited electron, the electron will release a photon with the same properties. This effect—called *stimulated emission*—was predicted by Einstein and is used today to produce coherent laser

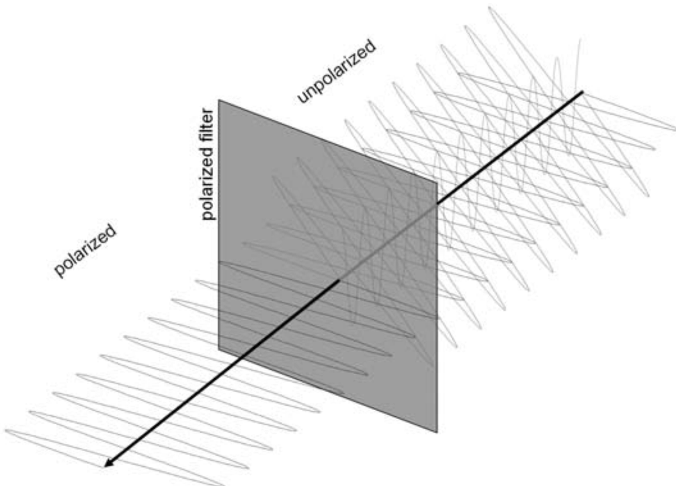


Figure 2.4. Polarization of light: only light waves with a specific orientation pass through the filter.

light. In general, "normal" light consists of an arbitrary superimposition of multiple incoherent light waves that create a complex *interference pattern*.

Light travels in a composition of waves with a variety of different orientations. It can be *polarized* by selecting waves with a specific orientation. The filtered portion is called *polarized light*. Augustine Fresnel explained this phenomenon in the 19th century.

Depending on the material properties, light can be *reflected*, *refracted*, *scattered*, or *light!absorbed* by matter. If reflected, light is bounced off a surface. Imperfections on the reflecting surface causes the light to be scattered (*diffused*) in different directions. Light can also be scattered when it collides with small particles (like molecules). The amount of scattering depends on the size of the particle with respect to the wavelength of the light. If the particle (e.g., a dust particle) is larger than the wavelength, light will be reflected. If the particle (e.g., a gas molecule) is smaller, then light will be absorbed. Such molecules will then radiate light at the frequency of the absorbed light in different directions. John Rayleigh explained this effect in the 1870s, thus this process is called *Rayleigh scattering*. Light can also be absorbed and its energy converted (e.g., into heat). Refraction occurs when light travels across the boundaries of two mediums. In a vacuum, light travels at 299.792 km/s (the speed of light). If travelling through a denser medium, it is slowed down which causes it to alter its direction.

The amount of refraction also depends on the wavelength of the light—as described by Isaac Newton who showed that white light splits into different angles depending on its wavelength.

2.2 Geometric Optics

In general, optics refers to all appearances that are perceived by the human visual system. The physical reason for these appearances, the light, was analyzed at an early time, and the basic principles of geometric optics and wave optics were outlined in the 19th century. In geometric optics, the light is represented by individual rays that, in ordinary media, are represented by straight lines. An ordinary media is *homogeneous* (the same at all points) and *isotropic* (the same for all directions). One of the basic hypotheses of geometric optics, the principle of P. de Fermat (1657), allows the representation of light rays within isotropic media that are independent of the light's wave nature. Today this hypothesis is known as the *principle of the optical path length*, and it states that the time that light travels on the path between two points is minimal.

2.2.1 Snell's Laws

The following laws were discovered by W. Snellius in 1621. They can also be derived from Fermat's principle. They describe the reflection and refraction behavior of straight light rays at the interfacing surface between two *homogeneous media*.

Figure 2.5 illustrates the interfacing surface that separates two homogeneous media with the two indices of refraction η_1 and η_2. A light ray r intersects the surface at i (with the normal vector n) and is refracted into r'.

The vectorized form of Snell's' law is given by $\eta_2 r' - \eta_1 r = an$, where r, r' and n are normalized and a is real.

Laws of refraction. We can derive the following *laws of refraction* from the vectorized form of Snell's law.

Theorem 2.2.1 (First refraction theorem.) *Since $r' = (\eta_1 r + an)/\eta_2$, the refracted ray r' lies on the plane that is spanned by r and n. This plane is called the* plane of incidence*.*

Theorem 2.2.2 (Second refraction theorem.) *If we compute the cross-product with* n *and the vectorized form of Snell's law, we obtain* $\eta_2(n \times r') = \eta_1(n \times r)$*. If we define the* angle of incidence α_i *and the* angle of refraction α_t*, we can substitute the cross-products and obtain Snell's law of refraction:* $\eta_1 \sin \alpha_i = \eta_2 \sin \alpha_t$*.*

Laws of reflection. Since the vector relation applies in general, we can assume r and r' to be located within the same medium with a refraction index η_1. Consequently, r is reflected at i into r' (Figure 2.6).

We can derive the following two theorems of reflection from the vectorized form of Snell's law.

Theorem 2.2.3 (First reflection theorem.) *Since* $r' = r + (a/\eta_1)n$*, the reflected ray* r' *lies on the plane of incidence.*

Theorem 2.2.4 (Second reflection theorem.) *If we compute the cross-product with* n *and the vectorized form of Snell's law, we obtain* $n \times r' = n \times r$*. If we reassign* α_t *to be the* angle of reflection*, we can substitute the cross-products and obtain Snell's law of reflection:* $-\sin \alpha_t = \sin \alpha_i$ *or* $-\alpha_t = \alpha_i$ *for* $-\pi/2 \leq \alpha_i \leq \pi/2$ *and* $-\pi/2 \leq \alpha_t \leq \pi/2$*.*

Note that the law of reflection is formally based on the assumption that $\eta_2 = -\eta_1$.

Critical angle and total internal reflection. Since $-1 \leq \sin \alpha_i \leq 1$, we can derive $-(\eta_1/\eta_2) \leq \sin \alpha_t \leq (\eta_1/\eta_2)$ from the second refraction theorem. It therefore holds that $-(\pi/2) \leq \alpha_i \leq (\pi/2)$ and $-\gamma \leq \alpha_t \leq \gamma$, whereby γ is called the *critical angle* and is defined by $\gamma = \sin^{-1}(\eta_1/\eta_2)$.

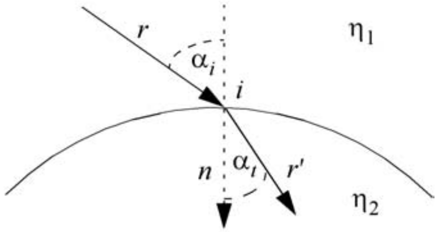


Figure 2.5. Snell's law of refraction.

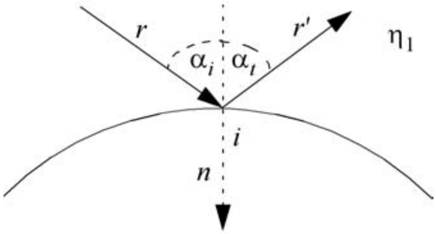


Figure 2.6. Snell's law of reflection.

If α_i becomes sufficiently large when entering an optically sparser medium, then α_t exceeds 90° and r is reflected from the interfacing surface, rather than being transmitted. This phenomenon is known as *total internal reflection*.

We can differentiate between two cases:

1. r enters an optically denser medium ($\eta_1 < \eta_2$): r is refracted for all angles of incidence α_i. If $\alpha_i = \pi/2$, then $\alpha_t = \sin^{-1}(\eta_1/\eta_2) = \gamma$.

2. r enters an optically sparser medium ($\eta_1 > \eta_2$): If $\alpha_i < \gamma = sin^{-1}(\eta_1/\eta_2)$, then r is refracted. Otherwise, r is reflected, due to total internal reflection.

2.2.2 The Formation of Point Images

Optical instruments can form *images* from a number of point-like light sources (so-called *objects*). Light rays that are emitted from an object can be reflected and refracted within the optical instrument and are finally perceived by a detector (e.g., the human eye or a photographic film). If all light rays that are emitted from the same object p_o travel through the optical system which bundles them within the same image p_i, then the points p_o and p_i are called a *stigmatic pair*. Consequently, this image-formation property is called *stigmatism*, and the optical system that supports stigmatism between all object-image pairs is called an *absolute optical system*.

The basic precondition for stigmatism can also be derived from Fermat's principle. It states that the optical path length for every light ray travelling from p_o to p_i is constant:

$$L(p_o \rightarrow p_i) = \eta_1(i_x - p_o) + L(i_x \rightarrow j_x) + \eta_2(p_i - j_x) = \text{const}$$

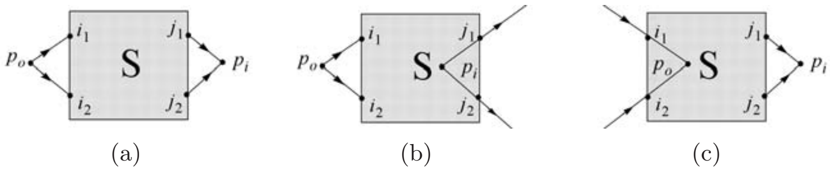


Figure 2.7. Stigmatic image formation. (a) real object, real image, (b) real object, virtual image, and (c) virtual object, real image.

where η_1 and η_2 are the refraction indices at the entrance and the exit of the optical system.

If points (objects or images) are formed by a direct intersection of light rays, then these points are called *real*. Figure 2.7(a) illustrates the *real object* p_o whose emitted light rays pass through an optical system (the filled square) and intersect at the *real image* p_i.

If light rays do not directly intersect at a point (e.g., if they diverge after exiting the optical instrument), they can form *virtual points*. Since human observers are only able to detect the directions of light rays, rays diverging from an optical system can appear to intersect within the system. These images are called *virtual images* (Figure 2.7(b)).

The location of virtual points can be determined by extending the exiting light rays in the negative direction. Consequently, this portion of the optical path is negative and must be subtracted from the total path length.

As illustrated in Figure 2.7(c), objects can also be virtual. In this case, the entering light rays have to be extended to find the location of the corresponding virtual object. Similar to the relationship of the optical path to a virtual image, the sub-path to a virtual object also has to be subtracted from the total path length.

The production of absolute optical systems is difficult, since the only surfaces that are easy to build and support stigmatism (some only for a single object-image pair) are planar or spherical surfaces. Therefore, most optical instruments only approximate stigmatic image formation. The introduced deviation from the ideal image is called *aberration*. Some examples of reflective and refractive optical systems are given in the following sections.

2.2.3 Reflective Optics

In the case of exclusively reflective optical systems (*mirrors*), the medium that light rays travel through is homogeneous, thus $\eta_1 = \eta_2 = \eta$ and

$i_x = j_x$. Consequently, the optical path length equation can be simplified:

$$L(p_o \rightarrow p_i) = \eta((i_x - p_o) + (p_i - i_x)) = \text{const.}$$

It can be further idealized that a mirror is surrounded by air, and that the medium air is approximately equivalent to the medium of a vacuum ($\eta = 1$), then two stigmatic points which are formed within air are defined by

$$L(p_o \rightarrow p_i) = (i_x - p_o) + (p_i - i_x) = \text{const.}$$

Planar mirrors. In the case of planar mirrors p_o is real while p_i is virtual (Figure 2.8 (a)), and all points i_x of the simplified optical path equation describe the surface of a rotation-hyperboloid with its two focal points in p_o and p_i. Planes represent a special variant of a rotation-hyperboloid, where $L(p_o \rightarrow p_i) = 0$. Planar mirrors are absolute optical systems that map each object p_o to exactly one image p_i. Since this mapping is bijective, invertible, and symmetrical for all points, it provides stigmatism between all objects and images. This means that images which are generated from multiple image points preserve the geometric properties of the reflected objects that are represented by the corresponding object points.

If we represent the mirror plane by its normalized plane equation within the three-dimensional space $f(x, y, z) = ax+by+cz+d = 0$, then the image for a corresponding object can be computed as follows: With respect to Figure 2.8(a), it can be seen that the distance from p_o to the mirror plane equals the distance from the mirror plane to p_i (i.e., $a = a'$). This can be derived from the simplified optical path equation with simple triangulation.

If we now define the ray $r_p = p_o + \lambda n$, where $n = (a, b, c)$ ($f(x, y, z)$ is normalized) is the normal vector perpendicular to the mirror plane and λ an arbitrary extension factor of n, we can insert the components of r_p into f and solve for λ:

$$f(r_p) = nr_p + d = n(p_o + \lambda n) + d = 0 \rightarrow \lambda = -\frac{1}{nn}(np_o + d).$$

Since $|n| = 1$, we can set $nn = 1$ and solve $\lambda = -(np_o + d) = a = a'$. Consequently the intersection of r_p with f is given by

$$i_p = p_o - (np_o + d)n.$$

Since $a = a'$, the image point p_i results from

$$p_i = p_o - 2(np_o + d)n. \tag{2.1}$$

In Chapter 6, we show how Equation (2.1) can be expressed as a homogeneous 4×4 transformation matrix that can be integrated into fixed function rendering pipelines to support optical combination with mirror beam combiners.

With respect to Snell's first reflection theorem, we can determine the reflection ray r' of the original ray r as

$$r' = r - 2n(nr). \tag{2.2}$$

In the case of planar mirrors, n is constant for all surface points i. However, this equation is also valid for non-planar mirrors with individual normal vectors at each surface point. In this case, the intersection i of r with the mirror surface and the normal n at i has to be inserted in Equations (2.1) and (2.2). Note that Equation (2.2) is a common equation used by ray-tracers to compute specular reflection rays. The curvilinear behavior of reflections at non-planar mirrors can be well expressed with ray-tracing techniques, since they are based on the optical foundations of light rays.

Non-planar mirrors. In contrast to planar mirrors, non-planar mirrors do not provide stigmatism between all objects and images. In fact, only a few surface types generate just one true stigmatic pair. For all other objects (or for objects reflected by other surfaces), the corresponding images have to be approximated, since the reflected light rays do not bundle exactly within a single point.

Like planar mirrors, *convex mirrors* generate virtual images from real objects. This is because light rays always diverge after they are reflected. Rotation-paraboloids (parabolic mirrors), for instance, can generate just one true stigmatic pair (Figure 2.8(b)).

The extended light rays bundle in one virtual point p_i only if p_o is located at infinity. This point is the *focal point* f of the paraboloid. The distance between the focal point and the surface is called the focal distance or *focal length* f. For example, the focal length of a convex mirror is defined as $f = -r/2$, the focal length of a concave mirror is given by $f = r/2$, and the focal length of a planar mirror is $f = 0$, where r is the surface radius.

If p_o is not located at infinity, the extended light rays do not bundle exactly within a single image. Thus, p_i has to be approximated (Figure 2.8(c)). Note, that in this case, images formed by multiple image points appear to be a reduced and deformed version of the reflected object that is represented by the corresponding object points.

In addition to rotation-paraboloids, other mirror surfaces (such as rotation-hyperboloids and prolate ellipsoids) can generate a single true

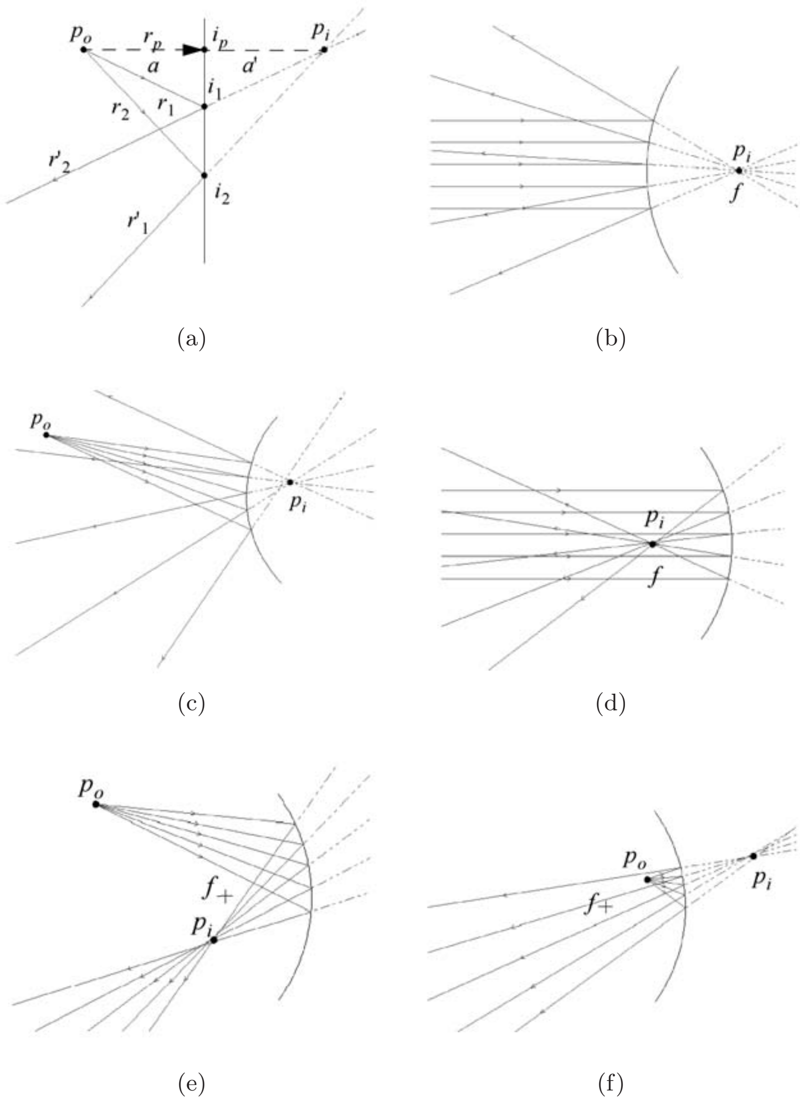


Figure 2.8. (a) Planar mirror; (b) convex parabolic mirror with object at infinity; (c) convex parabolic mirror with an object at a finite distance away from the mirror surface; (d) concave parabolic mirror with object at infinity; (e) concave parabolic mirror with an object at a finite distance behind its focal point; (f) concave parabolic mirror with an object at a finite distance in front of its focal point.

stigmatic pair. In general we can say that the true stigmatic pair, generated by such surfaces, is always their two focal points. Mirror surfaces other than the ones mentioned above do not generate true stigmatic pairs at all.

Concave mirrors can generate both virtual and real images from real objects because the reflected light rays converge or diverge, depending on the location of the object with respect to the focal point. As in the convex case, only the above mentioned surface types can generate just one true stigmatic pair which consists of their two focal points. For other surface types (or for objects that do not match the focal points), images can only be approximated.

Figure 2.8(d) illustrates an example of a concave parabolic mirror, where p_o is located at infinity and p_i is generated at f.

If p_o is not located at infinity, p_i has to be approximated. However, depending on the position of the object with respect to the focal point, the image can be either real or virtual. If, on the one hand, the object p_o is located behind the focal point f (as illustrated in Figure 2.8(e)) the reflected light rays converge and approximately bundle within the real image p_i (also located behind the focal point). Note, that in this case, images formed by multiple image points appear to be an enlarged, flipped, and deformed version of the reflected object that is represented by the corresponding object points.

If, on the other hand, p_o is located between the surface and f (as illustrated in Figure 2.8(f)) the reflected light rays diverge and their extensions approximately bundle within the virtual image p_i. Note, that in this case, images formed by multiple image points appear to be an enlarged and deformed version of the reflected object that is represented by the corresponding object points—yet, it is not flipped.

Note that if $p_o = f$, then p_i is located at infinity (i.e., the reflected light rays are parallel). In this case, p_i is neither real nor virtual.

2.2.4 Refractive Optics

In the case of refractive optical systems (*lenses*), the medium that light rays travel through is inhomogeneous. This means that, with respect to the simplified optical path equation, light rays pass through two different media with different densities and refraction indices ($\eta_1 \neq \eta_2$), where η_1 denotes the refraction index of the medium that surrounds the lens, and η_2 is the refraction index of the lens material. Since the rays are redirected when they change into another medium, their entrance and exit points