

PRINCIPLES AND PRACTICE



Augmented **REALITY**



Dieter **SCHMALSTIEG**
Tobias **HÖLLERER**

FREE SAMPLE CHAPTER

SHARE WITH OTHERS



Augmented Reality

Augmented Reality

Principles and Practice

Dieter Schmalstieg
Tobias Höllerer

◆ Addison-Wesley

Boston • Columbus • Indianapolis • New York • San Francisco • Amsterdam • Cape Town
Dubai • London • Madrid • Milan • Munich • Paris • Montreal • Toronto • Delhi • Mexico City
São Paulo • Sydney • Hong Kong • Seoul • Singapore • Taipei • Tokyo

Many of the designations used by manufacturers and sellers to distinguish their products are claimed as trademarks. Where those designations appear in this book, and the publisher was aware of a trademark claim, the designations have been printed with initial capital letters or in all capitals.

The authors and publisher have taken care in the preparation of this book, but make no expressed or implied warranty of any kind and assume no responsibility for errors or omissions. No liability is assumed for incidental or consequential damages in connection with or arising out of the use of the information or programs contained herein.

For information about buying this title in bulk quantities, or for special sales opportunities (which may include electronic versions; custom cover designs; and content particular to your business, training goals, marketing focus, or branding interests), please contact our corporate sales department at corpsales@pearsoned.com or (800) 382-3419.

For government sales inquiries, please contact governmentsales@pearsoned.com.

For questions about sales outside the U.S., please contact intlcs@pearson.com.

Visit us on the Web: informati.com/aw

Library of Congress Cataloging-in-Publication Data

Names: Schmalstieg, D. (Dieter), author. | Höllerer, Tobias, 1970– author.

Title: Augmented reality : principles and practice / Dieter Schmalstieg,
Tobias Höllerer.

Description: Boston : Addison-Wesley, 2016. | Includes bibliographical
references and index.

Identifiers: LCCN 2016009049 | ISBN 9780321883575 (pbk. : alk. paper)

Subject: LCSH: Augmented reality.

Classification: LCC QA76.9.A94 S36 2016 | DDC 006.8—dc23

LC record available at <https://lccn.loc.gov/2016009049>

Copyright © 2016 Pearson Education, Inc.

All rights reserved. Printed in the United States of America. This publication is protected by copyright, and permission must be obtained from the publisher prior to any prohibited reproduction, storage in a retrieval system, or transmission in any form or by any means, electronic, mechanical, photocopying, recording, or likewise. For information regarding permissions, request forms and the appropriate contacts within the Pearson Education Global Rights & Permissions Department, please visit www.pearsoned.com/permissions/.

ISBN-13: 978-0-321-88357-5

ISBN-10: 0-321-88357-8

Text printed in the United States on recycled paper at RR Donnelley in Crawfordsville, Indiana.

First printing, June 2016

Publisher

Mark L. Taub

Executive Editor

Laura Lewin

Development Editor

Susan Brown Zahn

Managing Editor

Sandra Schroeder

Full-Service Production Manager

Julie B. Nahil

Project Editor

Thistle Hill Publishing Services

Copy Editor

Jill Hobbs

Indexer

John S. Lewis

Proofreader

Anna Popick

Technical Reviewers

Reinhold Behringer

Doug Bowman

Kiyoshi Kiyokawa

Editorial Assistant

Olivia Basegio

Cover Designer

Chuti Prasertsith

Compositor

Shepherd, Inc.

To Ursula, Katharina, and Florian
—Dieter

To Julie, Clara, and Luisa
—Tobias

This page intentionally left blank

Contents

Preface	xix
Acknowledgments	xxv
About the Authors	xxvii
1 Introduction to Augmented Reality	1
Definition and Scope	3
A Brief History of Augmented Reality	4
Examples	13
Industry and Construction	13
Maintenance and Training	17
Medical	18
Personal Information Display	20
Navigation	21
Television	22
Advertising and Commerce	25
Games	27
Related Fields	28
Mixed Reality Continuum	28
Virtual Reality	29
Ubiquitous Computing	30
Summary	31
2 Displays	33
Multimodal Displays	34
Audio Displays	34
Haptic, Tactile, and Tangible Displays	35
Olfactory and Gustatory Displays	37
Visual Perception	39
Requirements and Characteristics	40
Method of Augmentation	40
Ocularity and Stereoscopy	42

Focus	45
Occlusion	47
Resolution and Refresh Rate	48
Field of View	50
Viewpoint Offset	51
Brightness and Contrast	53
Distortions and Aberrations	55
Latency	55
Ergonomics	55
Social Acceptance	55
Spatial Display Model	56
Visual Displays	58
Near-Eye Displays	59
Handheld Displays	69
Stationary Displays	72
Projected Displays	78
Summary	84
 3 Tracking	 85
Tracking, Calibration, and Registration	86
Coordinate Systems	87
Model Transformation	88
View Transformation	88
Projective Transformation	89
Frames of Reference	89
Characteristics of Tracking Technology	90
Physical Phenomena	90
Measurement Principle	91
Measured Geometric Property	91
Sensor Arrangement	91
Signal Sources	92
Degrees of Freedom	92
Measurement Coordinates	92
Spatial Sensor Arrangement	93
Workspace Coverage	94

Measurement Error	94
Temporal Characteristics	95
Stationary Tracking Systems	96
Mechanical Tracking	96
Electromagnetic Tracking	97
Ultrasonic Tracking	98
Mobile Sensors	99
Global Positioning System	99
Wireless Networks	101
Magnetometer	102
Gyroscope	102
Linear Accelerometer	103
Odometer	104
Optical Tracking	105
Model-Based versus Model-Free Tracking	106
Illumination	106
Markers versus Natural Features	109
Target Identification	113
Sensor Fusion	117
Complementary Sensor Fusion	117
Competitive Sensor Fusion	117
Cooperative Sensor Fusion	118
Summary	120
4 Computer Vision for Augmented Reality	121
Marker Tracking	123
Camera Representation	124
Marker Detection	126
Pose Estimation from Homography	128
Pose Refinement	132
Multiple-Camera Infrared Tracking	132
Blob Detection	133
Establishing Point Correspondences	133
Triangulation from Two Cameras	135
Triangulation from More Than Two Cameras	137

Matching Targets Consisting of Spherical Markers	137
Absolute Orientation	137
Natural Feature Tracking by Detection	138
Interest Point Detection	140
Descriptor Creation	144
Descriptor Matching	145
Perspective-n-Point Pose	146
Robust Pose Estimation	148
Incremental Tracking	149
Active Search	150
Kanade-Lucas-Tomasi Tracking	151
Zero-Normalized Cross-Correlation	152
Hierarchical Search	154
Combined Detection and Tracking	155
Simultaneous Localization and Mapping	156
Five-Point Algorithm for Essential Matrix	157
Bundle Adjustment	158
Parallel Tracking and Mapping	159
Relocalization and Loop Closure	160
Dense Mapping	161
Outdoor Tracking	164
Scalable Visual Matching	165
Prior Information from Sensors	167
Prior Information from Geometry	169
Simultaneous Tracking, Mapping, and Localization	170
Summary	176
5 Calibration and Registration	179
Camera Calibration	180
Internal Camera Parameters	180
Correcting Lens Distortion	182
Display Calibration	183
Single Point Active Alignment Method	185
Head-Mounted Display Calibration Using a Pointing Device	186
Hand-Eye Calibration	188

Registration	190
Geometric Measurement Distortions	190
Error Propagation	191
Latency	192
Filtering and Prediction	192
Summary	194
6 Visual Coherence	195
Registration	196
Occlusion	199
Occlusion Refinement	201
Probabilistic Occlusion	202
Model-Free Occlusion	202
Photometric Registration	205
Image-Based Lighting	207
Light Probes	208
Offline Light Capturing	210
Photometric Registration from Static Images	210
Photometric Registration from Specular Reflections	211
Photometric Registration from Diffuse Reflections	212
Photometric Registration from Shadows	214
Outdoor Photometric Registration	214
Reconstructing Explicit Light Sources	215
Common Illumination	216
Differential Rendering	216
Real-Time Global Illumination	218
Shadows	220
Diffuse Global Illumination	223
Specular Global Illumination	225
Diminished Reality	227
Determination of the Region of Interest	228
Observation and Modeling of the Hidden Area	228
Removal of the Region of Interest	229
Projector-Based Diminished Reality	230

Camera Simulation	231
Lens Distortion	231
Blur	232
Noise	234
Vignetting	234
Chromatic Aberrations	234
Bayer Pattern Artifacts	235
Tone Mapping Artifacts	235
Stylized Augmented Reality	236
Summary	237
 7 Situated Visualization	 239
Challenges	241
Data Overload	242
User Interaction	242
Registration Errors	243
Visual Interference	243
Temporal Coherence	244
Visualization Registration	245
Locally Registered Situated Visualization	245
Globally Registered Situated Visualization	246
Registration Uncertainty	247
Annotations and Labeling	248
Labeling Fundamentals	248
Optimization Techniques	249
Temporal Coherence	250
Image-Guided Placement	252
Legibility	253
X-Ray Visualization	254
Ghostings from Object Space	255
Ghostings from Image Space	256
Implementation with G-Buffers	258
Spatial Manipulation	260
Explosion Diagrams	260
Space Distortion	262

Information Filtering	265
Knowledge-Based Filter	265
Spatial Filter	265
Combined Knowledge-Based and Spatial Filter	267
Summary	270
8 Interaction	271
Output Modalities	272
Augmentation Placement	272
Agile Displays	274
Magic Lenses	276
Input Modalities	279
Tracking and Manipulation of Rigid Objects	279
Body Tracking	281
Gestures	282
Touch	283
Physically Based Interfaces	286
Tangible Interfaces	286
Tangibles on Surfaces	287
Tangibles with Generic Shape	287
Tangibles with Distinct Shapes	289
Transparent Tangibles	293
Virtual User Interfaces on Real Surfaces	294
Augmented Paper	295
Multi-view Interfaces	297
Multi-display Focus+Context	297
Shared Space	298
Multiple Locales	300
Cross-View Interaction	303
Haptic Interaction	304
Multimodal Interaction	304
Conversational Agents	306
Summary	309

9 Modeling and Annotation	311
Specifying Geometry	312
Points	313
Planes	314
Volumes	315
Specifying Appearance	317
Semi-automatic Reconstruction	319
Free-Form Modeling	322
Annotation	325
Summary	328
 10 Authoring	329
Requirements of AR Authoring	330
Real-World Interfaces	330
Hardware Abstraction	332
Authoring Workflow	332
Elements of Authoring	333
Actors	333
Story	334
Stages	334
Interactions	334
Setup	335
Stand-Alone Authoring Solutions	335
Desktop Authoring	335
Authoring by Performance	337
Plug-In Approaches	339
Web Technology	341
Summary	342
 11 Navigation	345
Foundations of Human Navigation	346
Exploration and Discovery	347
Route Visualization	347

Viewpoint Guidance	350
Guidance Toward a Target Object	350
Guidance Toward a Target Viewpoint	354
Multiple Perspectives	354
Simultaneous Multiple Perspectives	355
Transitional Interfaces	357
Summary	360
12 Collaboration	361
Properties of Collaboration Systems	362
Co-located Collaboration	364
Individual Displays and Views	366
Gaze Awareness	368
Agile Collaboration in Shared Space	368
Remote Collaboration	370
Video Sharing	371
Video Sharing with Virtual Objects	372
Video Sharing with Geometric Reconstruction	374
Pointing and Gestures	375
Remote Collaboration with Agile Users	376
Summary	377
13 Software Architectures	379
AR Application Requirements	380
Environment Control and Scene Dynamics	380
Display Space	381
Real–Virtual Consistency	381
Semantic Knowledge	381
Physical Space	382
Software Engineering Requirements	382
Platform Abstraction	382
User Interface Abstraction	383
Reusability and Extensibility	383

Distributed Computing	384
Decoupled Simulation	384
Distributed Object Systems	385
Object Management	386
Case Study: <i>SHEEP</i>	387
Dataflow	389
Dataflow Graphs	390
Multimodal Interaction	390
Threads and Scheduling	391
Case Study: Wearable Augmented Reality Setup	393
Scene Graphs	395
Fundamentals of Scene Graphs	395
Dependency Graph	397
Scene Graph Integration	397
Distributed Shared Scene Graph	399
Developer Support	400
Parameter Configuration	401
Declarative Scripting	401
Case Study: Augmented Reality Tour Guide	403
Procedural Scripting	405
Mixed-Language Programming	405
Runtime Reconfiguration	405
Choosing an AR Platform	407
Summary	407
14 The Future	409
What May Drive Business Cases	410
Professional Users	410
Consumers	411
An AR Developer’s Wish List	411
Low-Level Camera API	412
Multiple Cameras	412
Wide-Field-of-View Cameras	413
Sensors	413
Unified Memory	413

Parallel Programming on the Mobile GPU	414
Better Displays	414
Taking AR Outdoors	415
Uncooperative Users	415
Limited Device Capabilities	416
Localization Success Rate	416
Interfacing with Smart Objects	417
Confluence of Virtual Reality and Augmented Reality	418
Augmented Humans	419
AR as a Dramatic Medium	420
AR as a Social Computing Platform	421
Summary	422
References	423
Index	473

This page intentionally left blank

Preface

Over the past 20 years, the use of information technology has undergone a clear transition from stationary office and desktop computing—first to the web, then to social media, and then to mobile computing. Sales of smartphones and tablet computers have far outpaced the sales of conventional desktop PCs for years now, even if one places laptop or notebook computers within the desktop category.

While the predominant user interface style of today has not radically departed from the *desktop computing* of the 1990s (or the 1981 Xerox Star, for that matter), the way members of the young generation today attain computer literacy has changed: Apps and cloud computing are replacing the computer desktop in many cases. Computing has shifted from office or home office work to an anywhere-and-anytime activity.

Enter: Augmented Reality

As users move away from the desktop, it increasingly makes sense to include the physical world in our computing experience. Given that the physical world is not flat and is not composed of written documents, a new user interface metaphor becomes necessary. **Augmented reality** (AR) has the potential to become the leading user interface metaphor for *situated computing*. Augmented reality has the unique quality of providing a direct link between the physical reality and virtual information about that reality. The world becomes the user interface, leading to the familiar proclamation:

Back to the real world!

Virtual reality, the vision of immersing ourselves in artificial worlds, has propelled the development of game consoles with amazing 3D graphics and led to consumer devices such as head-mounted displays and gesture-tracking devices. Even so, a user interface metaphor such as virtual reality, which by definition monopolizes our attention, is not necessarily a good fit for everyday and spontaneous use of computing.

Instead, we increasingly rely on computer interfaces that make possible casual use and provide information in small, easily understood portions. We feel a need for **ubiquitous computing**. This can take the form of *calm* computing, which operates behind the scenes without the user intervening or even consciously noticing. If ubiquitous *interaction* is required, though, augmented reality excels as an appropriate user interface technology.

Why a Book on Augmented Reality?

Multiple overlapping research fields are contributing to the development of augmented reality, and the associated body of knowledge is growing fast. The authors of this book have been contributing to this body of knowledge as researchers since the 1990s. However, the main motivation for this book came from teaching classes on augmented reality at the authors' home institutions at Graz University of Technology and the University of California, Santa Barbara. In the preparation for these classes, it became evident that no single text is available that covers both the breadth and the depth of this rapidly evolving field. Some notes were available from tutorials at various conferences, several of which the authors were involved in, starting as early as SIGGRAPH 2001. A lot of ground has been covered since then, and the authors were motivated to assemble all this knowledge in a systematic way with an eye toward both novel concepts and practical information. Hence, this book was born.

What's in the Book?

As its title suggests, the book strives for a compromise between principles and practice. Our objective was to make it interesting and usable for both academic researchers and practitioners, especially engineers, who are interested in augmented reality applications. The book, therefore, is intended to be usable both as a textbook and as a reference. To get the most out of it, readers should have a basic understanding of computer science in general, and some knowledge of, and interest in, computer graphics and computer vision is helpful. We don't hesitate to refer to existing literature that explains specific aspects of the necessary background in more detail than we could within the constraints of a single volume. At the same time, we were careful to introduce and clearly explain any specific augmented reality concepts that go beyond basic knowledge, to make the book self-contained. Using the following structure, we present the technical and methodological foundations of AR.

Chapter 1, "Introduction to Augmented Reality," sets the stage by presenting a working definition of augmented reality, providing a brief history of the field, and then walking the reader through various application examples of this powerful real-world user interface technology. We conclude the chapter with a contextualization within the spectrum of related technologies and research fields.

Chapter 2, "Displays," deals with displays, a fundamental enabling technology for augmented reality. Based on some foundations of visual perception, various display technologies that are suitable for augmented reality are discussed—in particular, head-mounted displays, handheld displays, and projective displays. The chapter also discusses nonvisual displays, such as auditory and haptic devices.

Chapter 3, "Tracking," gives an introduction to tracking, one of the core technologies underlying augmented reality. We first discuss the characteristics that are necessary to understand how

tracking—and measurement systems in general—work. We then discuss traditional stationary tracking systems and compare them to mobile sensors. Optical tracking as the most prominent tracking technology is given extensive treatment. The chapter concludes by sketching the principles of sensor fusion.

Chapter 4, “Computer Vision for Augmented Reality,” picks up the issue of optical tracking from Chapter 3 and gives a detailed account of computer vision algorithms for real-time pose estimation, i.e. for determining a camera’s viewing position and orientation from the observed imagery. To make this topic manageable and address readers with a wide variety of backgrounds, the chapter is structured along a suite of case studies. Every case study introduces only the knowledge necessary for it to be self-contained, so the reader does not have to accumulate in-depth knowledge of computer vision first. Moreover, advanced mathematical topics, which in practice are often used as a black box by relying on a software library such as OpenCV, are marked so that the reader can safely skip over them.

Chapter 5, “Calibration and Registration,” deals with methods for calibration and registration of the devices used in augmented reality. Calibration of the digital cameras used for the optical tracking described in Chapter 3 is a necessary prerequisite to deliver repeatable, accurate behavior in augmented reality applications. Registration is the process that aligns the physical and virtual parts of the augmented reality experience geometrically, thereby giving rise to the illusion of a coherent mixed environment.

Chapter 6, “Visual Coherence,” focuses on a family of computer graphics techniques that together produce a seamlessly blended view of real and virtual objects. It includes phenomena such as correct occlusion between virtual and real objects, or correct shadowing between virtual and real objects. We also explain diminished reality, or the removal of real objects from a scene, and examine the simulation of physical cameras.

Chapter 7, “Situated Visualization,” is dedicated to visualization techniques. Visualization has the objective of making information comprehensible. In the context of augmented reality, this means that the computer-generated information that is geometrically registered to objects in the physical scene must be positioned and styled in such a way that it can be easily understood by its users. We deal with both two-dimensional augmentations (such as textual labels) and three-dimensional augmentations (such as synthesized views of the interior of objects, so-called “ghostings”).

Chapter 8, “Interaction,” examines the various interaction techniques and interaction styles that are relevant for augmented reality applications. The topics range from simple situated information browsing to full three-dimensional interaction. We specifically discuss props, widgets, and hand-based interaction, and the connection of augmented reality to tangible user interfaces of various forms. We also take a look at multimodal and agent-based interfaces for augmented reality.

Chapter 9, “Modeling and Annotation,” is concerned with the topic of interactive modeling—that is, the creation of new geometric content through augmented reality. User interfaces that are embedded in a three-dimensional environment provide a powerful approach for re-creating a digital version of this environment. This capability is invaluable for all applications that deal with visual computing.

Chapter 10, “Authoring,” discusses authoring approaches for augmented reality. The content of augmented reality presentations and information databases needs to be designed and created the same way that web content is authored today. Augmented reality content can be authored with conventional tools or in augmented reality itself. Authoring is concerned with aspects of the application that go beyond geometric and visual properties—in particular, establishing the semantics and the behavior of the application. Preferably, authoring should be content-driven and require no or only minimal traditional programming effort. We discuss various approaches to address this need, and also examine recent efforts to combine augmented reality authoring with emerging open web standards.

Chapter 11, “Navigation,” deals with navigational guidance—a particularly relevant aspect of augmented reality as a user interface. Orientation in unfamiliar environments is an important application challenge involving mobile information systems. We present an overview of navigational guidance techniques implemented using augmented reality, and compare them to digital maps.

Chapter 12, “Collaboration,” investigates collaboration. Augmented reality has strong potential as a medium that can be used for communication among individuals. This encompasses both co-located collaboration, which is enriched by the additional cues afforded by a shared augmented reality system, and remote collaboration, which can be significantly supported by augmented reality technology and, in the process, provide new forms of remote presence.

Chapter 13, “Software Architectures,” analyzes the underlying architectures of augmented reality systems. Augmented reality has complex requirements, as it must combine aspects of real-time systems, multimedia systems, and often also distributed systems. Combining these requirements in a flexible way that can be mastered by an application programmer is a difficult endeavor. We discuss various architectural patterns such as distributed objects, dataflow systems, and scene graphs, and present a number of case studies.

Chapter 14, “The Future,” reviews possible trajectories of augmented reality as it moves from a research field with demonstrated usefulness in prototype applications to potentially universal consumer adoption. As part of this effort, the chapter considers which roadblocks and unresolved issues remain to be overcome. It also summarizes trends and insights from all of the material presented in this book and sketches a future research agenda.

How to Use the Book and the Related Material

How you use this book will depend on your relationship to the field of augmented reality and the degree and focus of your interest. We discuss three types of roles that this relationship or interest might take.

If you are a developer: Professional developers can use the book for inspiration and guidance in the design, implementation, and evaluation of augmented reality applications. Readers with such backgrounds will find useful information on hardware setups in the display, tracking, and interaction chapters. They will benefit from the chapters on visual coherence, visualization, and authoring for the development of application content, and learn about appropriate registration technologies in the tracking, computer vision, and calibration chapters. User interface design is informed by the chapter on interaction and following chapters. Finally, the chapter on software architectures provides important information for actual implementation work.

If you are a teacher: The book is useful as a text for several different types of university-level courses. A graduate course on augmented reality can use it as the primary textbook. A course on computer graphics or visual computing could use the chapters on visual coherence and visualization as an introduction to graphical aspects of augmented reality. A course on computer vision can use the chapters on tracking and registration for teaching important real-time computer vision techniques. A human–computer interaction course can utilize the chapters on interaction, modeling, authoring, navigation, and collaboration to provide detailed coverage of augmented reality concepts.

If you are a researcher: This book can serve as a comprehensive reference guide for researchers interested in the development or evaluation of experimental augmented reality applications. The research agenda in the concluding chapter also provides researchers and students with a list of important questions to be addressed in the field.

Companion Website

The companion website to the book can be found at the following address:

<http://www.augmentedrealitybook.org>

Augmented reality is rapidly evolving. To make this book a dynamic working document, this companion website provides additional information, including teaching materials. This site contains information and links related to the latest augmented reality research and applications. This is an open effort, so readers are invited to contribute to this collection. Your comments will help us to update the website, as well as future editions of this book.

Register your copy of *Augmented Reality* at informit.com for convenient access to downloads, updates, and corrections as they become available. To start the registration process, go to informit.com/register and log in or create an account. Enter the product ISBN (9780321883575) and click Submit. Once the process is complete, you will find any available bonus content under "Registered Products."

Acknowledgments

This book would not have been possible without the encouragement and expertise of many friends and colleagues. First, we offer our gratitude to the reviewers who provided invaluable insights and suggestions for improvements: Reinhold Behringer, Doug Bowman, André Ferko, Steffen Gauglitz, Kiyoshi Kiyokawa, Tobias Langlotz, Vincent Lepetit, Gerhard Reitmayr, Chris Sweeney, and Daniel Wagner.

Second, we want to thank our editors at Addison-Wesley: Peter Gordon, who believed in the idea of this book and helped us establish the publishing contract, and Laura Lewin and Olivia Basegio, who continuously encouraged us and provided us with great advice.

Third, we want to thank all colleagues that provided us with additional image materials: Aaron Stafford, Alessandro Mulloni, Alexander Plopski, Andreas Butz, Andreas Geiger, Andreas Hartl, Andrei State, Andrew Maimone, Andy Gstoll, Ann Morrison, Anton Fuhrmann, Anton van den Hengel, Arindam Day, Blair MacIntyre, Brigitte Ludwig, Bruce Thomas, Christian Pirchheim, Christian Reinbacher, Christian Sandor, Claudio Pinhanez, Clemens Arth, Daniel Wagner, David Mizell, Denis Kalkofen, Domagoj Baričević, Doreé Seligmann, Eduardo Veas, Erick Mendez, Ernst Kruijff, Ethan Eade, Florian Ledermann, Gerd Hesina, Gerhard Reitmayr, Gerhard Schall, Greg Welch, Gudrun Klinker, Hannes Kaufmann, Henry Fuchs, Hiroyuki Yamamoto, Hrvoje Benko, István Barakonyi, Ivan Sutherland, Jan Herling, Jens Grubert, Jonathan Ventura, Joseph Newman, Julien Pilet, Kiyoshi Kiyokawa, Lukas Gruber, Mark Billinghurst, Markus Oberweger, Markus Tatzgern, Martin Hirzer, Matt Swoboda, Matthias Straka, Michael Gervautz, Michael Kenzel, Michael Marner, Morten Fjeld, Nassir Navab, Oliver Bimber, Pascal Fua, Pascal Lagger, Peter Kán, Peter Mohr, Peter Weir, Philipp Descovic, Qi Pan, Ralph Schönfelder, Raphael Grasset, Remo Ziegler, Simon Julier, Stefan Hauswiesner, Stefanie Zollmann, Steffen Gauglitz, Steve Feiner, Taehhee Lee, Takuji Narumi, Thanh Nguyen, Thomas Richter-Trummer, Tom Drummond, Ulrich Eck, Vincent Lepetit, Wayne Piekarski, William Steptoe, Wolfgang Broll, and Zsolt Szalavári.

Fourth, we want to thank all colleagues and students at Graz University of Technology and University of California, Santa Barbara, for the countless discussions on and off the topics in this book and for providing great working environments.

Finally, we would like to thank our families for supporting us and being patient during the not-so-short time of creating this book!

Dieter Schmalstieg
Graz, Austria, April 2016

Tobias Höllerer
Santa Barbara, California, April 2016

This page intentionally left blank

About the Authors

Dieter Schmalstieg is full professor and head of the Institute of Computer Graphics and Vision at Graz University of Technology (TUG), Austria. His current research interests are augmented reality, virtual reality, real-time graphics, user interfaces, and visualization. He received Dipl.-Ing. (1993), Dr.techn. (1997), and Habilitation (2001) degrees from Vienna University of Technology. Dr. Schmalstieg is author or coauthor of more than 300 peer-reviewed scientific publications. His organizational roles include associate editor in chief of *IEEE Transactions on Visualization and Computer Graphics*, member of the editorial advisory board of *Computers & Graphics* and of Springer's *Virtual Reality* journal, member of the steering committee of the IEEE International Symposium on Mixed and Augmented Reality, chair of the EUROGRAPHICS working group on Virtual Environments (1999–2010), key researcher of the K-Plus Competence Center for Virtual Reality and Visualization in Vienna, and key researcher of the Know-Center in Graz. In 2002, Dr. Schmalstieg received the START Career Award presented by the Austrian Science Fund. In 2012, he received the IEEE Virtual Reality Technical Achievement Award for seminal contributions to the field of augmented reality. He was elected as a senior member of IEEE, as a member of the Austrian Academy of Sciences, and as a member of the Academia Europaea. Since 2008, he has also been director of the Christian Doppler Laboratory for Handheld Augmented Reality.

Tobias Höllerer is professor of computer science at the University of California, Santa Barbara, where he co-directs the Four Eyes Laboratory, conducting research in the “four I’s” of Imaging, Interaction, and Innovative Interfaces. Dr. Höllerer holds a Diploma in Informatics from the Technical University of Berlin, as well as an M.S., an M.Phil., and a Ph.D. in computer science from Columbia University. He is a recipient of the U.S. National Science Foundation’s CAREER Award for his work on “Anywhere Augmentation,” which enables mobile computer users to place annotations in 3D space wherever they go. He is an IEEE senior member and was named an ACM Distinguished Scientist in 2013. Dr. Höllerer is author or coauthor of more than 150 peer-reviewed journal and conference publications in the areas of augmented and virtual reality, information visualization, 3D displays and interaction, mobile and wearable computing, and social computing. Several of these publications have been selected for Best Paper or Honorable Mention Awards at such venues as the IEEE International Symposium on Mixed and Augmented Reality (ISMAR), IEEE Virtual Reality (VR), ACM Virtual Reality Software and Technology, ACM User Interface Software and Technology, ACM MobileHCI, IEEE SocialCom, and IEEE CogSIMA. Dr. Höllerer is an associate editor of *IEEE Transactions on Visualization and Computer Graphics*. Among his many organizational roles for scientific conferences, he served as program chair for IEEE VR 2015, ICAT 2013, and IEEE ISMAR 2009 and 2010; as general chair of IEEE ISMAR 2006; and as member of the steering committee of IEEE ISMAR.

This page intentionally left blank

CHAPTER 1

INTRODUCTION TO AUGMENTED REALITY

Virtual reality is becoming increasingly popular, as computer graphics have progressed to a point where the images are often indistinguishable from the real world. However, the computer-generated images presented in games, movies, and other media are detached from our physical surroundings. This is both a virtue—everything becomes possible—and a limitation.

The limitation comes from the main interest we have in our daily life, which is not directed toward some virtual world, but rather toward the *real world* surrounding us. Smartphones and other mobile

devices provide access to a vast amount of information, anytime and anywhere. However, this information is generally disconnected from the real world. Consumers with an interest in retrieving online information from and about the real world, or linking up online information with the real world, must do so individually and indirectly, which, in turn, requires constant cognitive effort.

In many ways, enhancing mobile computing so that the association with the real world happens automatically seems an attractive proposition. A few examples readily illustrate this idea's appeal. Location-based services can provide personal navigation based on the Global Positioning System (GPS), while barcode scanners can help identify books in a library or products in a supermarket. These approaches require explicit actions by the user, however, and are rather coarse grained. Barcodes are useful for identifying books, but not for naming mountain peaks during a hiking trip; likewise, they cannot help in identifying tiny parts of a watch being repaired, let alone anatomic structures during surgery.

Augmented reality holds the promise of creating direct, automatic, and actionable links between the physical world and electronic information. It provides a simple and immediate user interface to an electronically enhanced physical world. The immense potential of augmented reality as a paradigm-shifting user interface metaphor becomes apparent when we review the most recent few milestones in human-computer interaction: the emergence of the World Wide Web, the social web, and the mobile device revolution.

The trajectory of this series of milestones is clear: First, there was an immense increase in access to online information, leading to a massive audience of information consumers. These consumers were subsequently enabled to also act as information producers and communicate with one another, and finally were given the means to manage their communications from anywhere, in any situation. Yet, the physical world, in which all this information retrieval, authoring, and communication takes place, was not readily linked to the users' electronic activity. That is, the model was stuck in a world of abstract web pages and services without directly involving the physical world. A lot of technological advancement has occurred in the field of location-based computing and services, which is sometimes referred to as *situated computing*. Even so, the user interfaces to location-based services remain predominantly rooted in desktop-, app-, and web-based usage paradigms.

Augmented reality can change this situation, and, in doing so, redefine information browsing and authoring. This user interface metaphor and its enabling technologies form one of today's most fascinating and future-oriented areas of computer science and application development. Augmented reality can overlay computer-generated information on views of the real world, amplifying human perception and cognition in remarkable new ways.

After providing a working definition of augmented reality, we will briefly review important developments in the history of the research field, and then present examples from various application areas, showcasing the power of this physical user interface metaphor.

Definition and Scope

Whereas virtual reality (VR) places a user inside a completely computer-generated environment, augmented reality (AR) aims to present information that is directly registered to the physical environment. AR goes beyond mobile computing in that it bridges the gap between virtual world and real world, both spatially and cognitively. With AR, the digital information appears to become part of the real world, at least in the user's perception.

Achieving this connection is a grand goal—one that draws upon knowledge from many areas of computer science, yet can lead to misconceptions about what AR really is. For example, many people associate the visual combination of virtual and real elements with the special effects in movies such as *Jurassic Park* and *Avatar*. While the computer graphics techniques used in movies may be applicable to AR as well, movies lack one crucial aspect of AR—interactivity. To avoid such confusion, we need to set a scope for the topics discussed in this book. In other words, we need to answer a key question: What is AR?

The most widely accepted definition of AR was proposed by Azuma in his 1997 survey paper. According to Azuma [1997], AR must have the following three characteristics:

- Combines real and virtual
- Interactive in real time
- Registered in 3D

This definition *does not require* a specific output device, such as a head-mounted display (HMD), nor does it limit AR to visual media. Audio, haptics, and even olfactory or gustatory AR are included in its scope, even though they may be difficult to realize. Note that the definition does require real-time *control* and spatial *registration*, meaning precise real-time alignment of corresponding virtual and real information. This mandate implies that the user of an AR display can at least exercise some sort of interactive viewpoint control, and the computer-generated augmentations in the display will remain registered to the referenced objects in the environment.

While opinions on what qualifies as real-time performance may vary depending on the individual and on the task or application, interactivity implies that the human–computer interface operates in a tightly coupled feedback loop. The user continuously navigates the AR scene and controls the AR experience. The system, in turn, picks up the user's input by tracking the user's viewpoint or pose. It registers the pose in the real world with the virtual content, and then presents to the user a *situated visualization* (a visualization that is registered to objects in the real world).

We can see that a complete AR system requires at least three components: a tracking component, a registration component, and a visualization component. A fourth component—a spatial model (i.e., a database)—stores information about the real world and about the virtual world (Figure 1.1). The real-world model is required to serve as a reference for the tracking

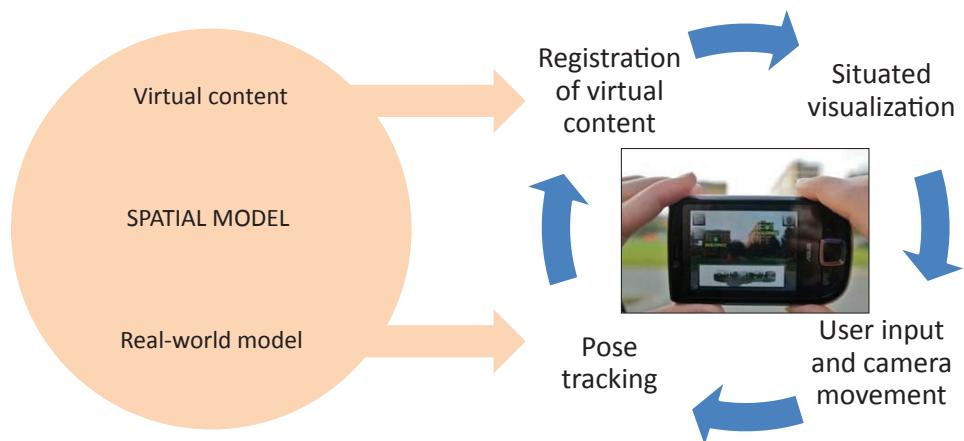


Figure 1.1 AR uses a feedback loop between human user and computer system. The user observes the AR display and controls the viewpoint. The system tracks the user's viewpoint, registers the pose in the real world with the virtual content, and presents situated visualizations.

component, which must determine the user's location in the real world. The virtual-world model consists of the content used for the augmentation. Both parts of the spatial model must be registered in the same coordinate system.

A Brief History of Augmented Reality

While one could easily go further back in time to find examples in which informational overlays were layered on top of the physical world, suffice it to say that the first annotations of the physical world with *computer-generated* information occurred in the 1960s. Ivan Sutherland can be credited with starting the field that would eventually turn into both VR and AR. In 1965, he postulated the *ultimate display* in an essay that contains the following famous quote:

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's [1965] essay includes more than just an early description of immersive displays, however. It also contains a quote that is less often discussed, but that clearly anticipates AR:

The user of one of today's visual displays can easily make solid objects transparent—he can "see through matter!"



Figure 1.2 The Sword of Damocles was the nickname of the world's first head-mounted display, built in 1968. Courtesy of Ivan Sutherland.

Shortly thereafter, Sutherland constructed the first VR system. In 1968, he finished the first head-mounted display [Sutherland 1968]. Because of its weight, it had to be suspended from the ceiling and was appropriately nicknamed "Sword of Damocles" (Figure 1.2). This display already included head tracking and used see-through optics.

Advances in computing performance of the 1980s and early 1990s were ultimately required for AR to emerge as an independent field of research. Throughout the 1970s and 1980s, Myron Krueger, Dan Sandin, Scott Fisher, and others had experimented with many concepts of mixing human interaction with computer-generated overlays on video for interactive art experiences. Krueger [1991], in particular, demonstrated collaborative interactive overlays of graphical annotations among participant silhouettes in his Videoplace installations around 1974.

The year 1992 marked the birth of the term "augmented reality." This term first appeared in the work of Caudell and Mizell [1992] at Boeing, which sought to assist workers in an airplane factory by displaying wire bundle assembly schematics in a see-through HMD (Figure 1.3).



Figure 1.3 Researchers at Boing used a see-through HMD to guide the assembly of wire bundles for aircraft. Courtesy of David Mizell.

In 1993, Feiner et al. [1993a] introduced KARMA, a system that incorporated knowledge-based AR. This system was capable of automatically inferring appropriate instruction sequences for repair and maintenance procedures (Figure 1.4).

Also in 1993, Fitzmaurice created the first handheld spatially aware display, which served as a precursor to handheld AR. The Chameleon consisted of a tethered handheld liquid-crystal display (LCD) screen. The screen showed the video output of an SGI graphics workstation of the time and was spatially tracked using a magnetic tracking device. This system was capable of showing contextual information as the user moved the device around—for example, giving detailed information about a location on a wall-mounted map.

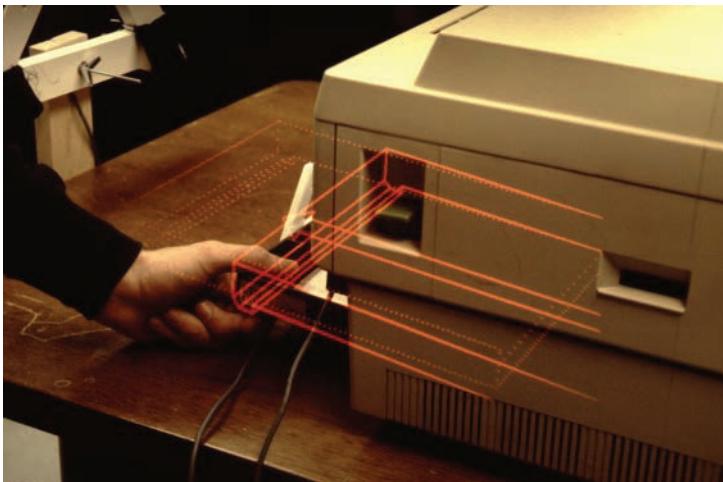


Figure 1.4 (top) KARMA was the first knowledge-driven AR application. (bottom) A user with an HMD could see instructions on printer maintenance. Courtesy of Steve Feiner, Blair MacIntyre, and Doreé Seligmann, Columbia University.

In 1994, State et al. at the University of North Carolina at Chapel Hill presented a compelling medical AR application, capable of letting a physician observe a fetus directly within a pregnant patient (Figure 1.5). Even though the accurate registration of computer graphics on top of a deformable object such as a human body remains a challenge today, this seminal work hints at the power of AR for medicine and other delicate tasks.



Figure 1.5 View inside the womb of an expecting mother. Courtesy of Andrei State, UNC Chapel Hill.

Around the mid-1990s, Steve Mann at the MIT Media Lab implemented, and experimented with, a “reality mediator”—a waist-bag computer with a video see-through HMD (a modified VR4 by Virtual Research Systems) that enabled the user to augment, alter, or diminish visual reality. Through the WearCam project, Mann [1997] explored wearable computing and mediated reality. His work ultimately helped establish the academic field of wearable computing, which, in those early days, had a lot of synergy with AR [Starner et al. 1997].

In 1995, Rekimoto and Nagao created the first true—albeit tethered—handheld AR display. Their NaviCam was connected to a workstation, but was outfitted with a forward-facing camera. From the video feed, it could detect color-coded markers in the camera image and display information on a video see-through view.



Figure 1.6 One of the applications of the Studierstube system was teaching geometry in AR to high school students. Courtesy of Hannes Kaufmann.

In 1996, Schmalstieg et al. developed Studierstube, the first collaborative AR system. With this system, multiple users could experience virtual objects in the same shared space. Each user had a tracked HMD and could see perspectively correct stereoscopic images from an individual viewpoint. Unlike in multi-user VR, natural communication cues, such as voice, body posture, and gestures, were not affected in Studierstube, because the virtual content was added to a conventional collaborative situation in a minimally obtrusive way. One of the showcase applications was a geometry course [Kaufmann and Schmalstieg 2003], which was successfully tested with actual high school students (Figure 1.6).

From 1997 to 2001, the Japanese government and Canon Inc. jointly funded the Mixed Reality Systems Laboratory as a temporary research company. This joint venture was the largest industrial research facility for mixed reality (MR) research up to that point [Tamura 2000] [Tamura et al. 2001]. Among its most notable achievements was the design of the first coaxial stereo video see-through HMD, the COASTAR. Many of the activities undertaken in the lab were also directed toward the digital entertainment market (Figure 1.7), which plays a very prominent role in Japan.

In 1997, Feiner et al. developed the first outdoor AR system, the Touring Machine (Figure 1.8), at Columbia University. The Touring Machine uses a see-through HMD with GPS and orientation tracking. Delivering mobile 3D graphics via this system required a backpack holding a



Figure 1.7 *RV-Border Guards* was a multiuser shooting game developed in Canon's Mixed Reality Systems Laboratory. Courtesy of Hiroyuki Yamamoto.

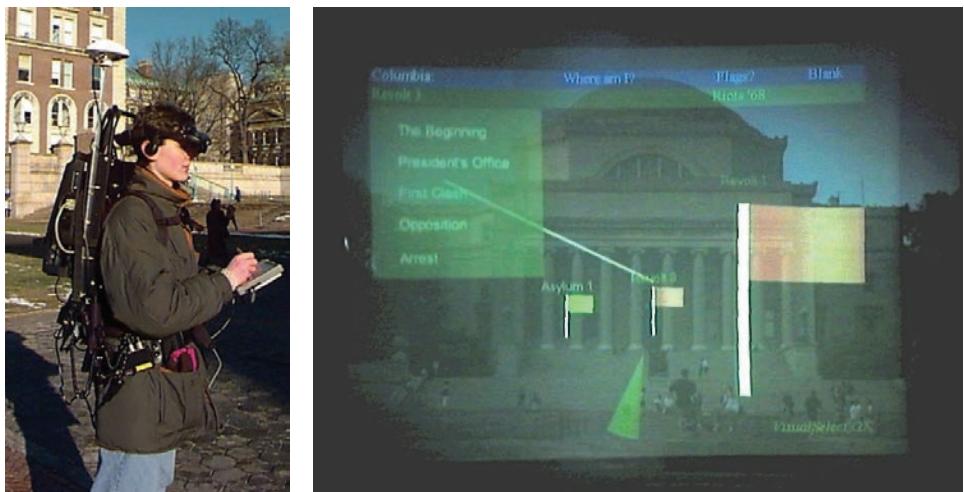


Figure 1.8 The Touring Machine was the first outdoor AR system (left). Image of the *Situated Documentaries* AR campus tour guide running on a 1999 version of the Touring Machine (right). Courtesy of Columbia University.

computer, various sensors, and an early tablet computer for input [Feiner et al. 1997] [Höllerer et al. 1999b].

Just one year later, in 1998, Thomas et al. published their work on the construction of an outdoor AR navigation system, Map-in-the-Hat. Its successor, Tinmith (few people know that this name is actually an acronym for “This is not map in the hat”), evolved into a well-known experimental platform for outdoor AR. This platform was used for advanced applications, such as 3D surveying, but is most famous for delivering the first outdoor AR game, *ARQuake* (Figure 1.9). This game, which is a port of the popular first-person shooter application *Quake* to Tinmith, places the user in the midst of a zombie attack in a real parking lot.

In the same year, Raskar et al. [1998] at the University of North Carolina at Chapel Hill presented the Office of the Future, a telepresence system built around the idea of structured light-scanning and projector-camera systems. Although the required hardware was not truly practical for everyday use at the time, related technologies, such as depth sensors and camera-projection coupling, play a prominent role in AR and other fields today.

Until 1999, no AR software was available outside specialized research labs. This situation changed when Kato and Billinghurst [1999] released ARToolKit, the first open-source software



Figure 1.9 Screenshot of ARQuake, the first outdoor AR game. Courtesy of Bruce Thomas and Wayne Piekarski.

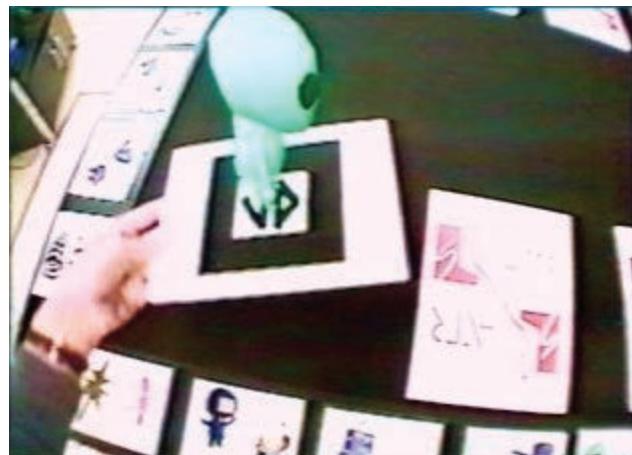


Figure 1.10 A person holding a square marker of ARToolKit, the popular open-source software framework for AR. Courtesy of Mark Billinghurst.

platform for AR. It featured a 3D tracking library using black-and-white fiducials, which could easily be manufactured on a laser printer (Figure 1.10). The clever software design, in combination with the increased availability of webcams, made ARToolKit widely popular.

In the same year, Germany’s Federal Ministry for Education and Research initiated a €21 million program for industrial AR, called ARVIKA (Augmented Reality for Development, Production, and Servicing). More than 20 research groups from industry and academia worked on developing advanced AR systems for industrial application, in particular in the German automotive industry. This program raised the worldwide awareness of AR in professional communities and was followed by several similar programs designed to enhance industrial application of the technology.

Another noteworthy idea also appeared in the late 1990s: IBM researcher Spohrer [1999] published an essay on Worldboard, a scalable networked infrastructure for hyperlinked spatially registered information, which Spohrer had first proposed while he was working with Apple’s Advanced Technology Group. This work can be seen as the first concept for an AR browser.

After 2000, cellular phones and mobile computing began evolving rapidly. In 2003, Wagner and Schmalstieg presented the first handheld AR system running autonomously on a “personal digital assistant”—a precursor to today’s smartphones. One year later, the *Invisible Train* [Pintaric et al. 2005], a multiplayer handheld AR game (Figure 1.11), was experienced by thousands of visitors at the SIGGRAPH Emerging Technologies show floor.

It took several years, until 2008, for the first truly usable natural feature tracking system for smartphones to be introduced [Wagner et al. 2008b]. This work became the ancestor of the

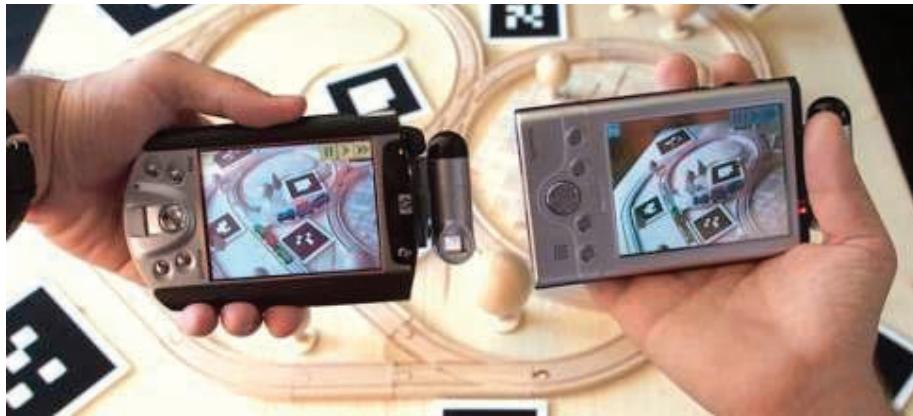


Figure 1.11 The Invisible Train was a handheld AR game featuring virtual trains on real wooden tracks. Courtesy of Daniel Wagner.

popular Vuforia toolkit for AR developers. Other noteworthy achievements in recent years in the area of tracking include the parallel tracking and mapping (PTAM) system of Klein and Murray [2007], which can track without preparation in unknown environments, and the KinectFusion system developed by Newcombe et al. [2011a], which builds detailed 3D models from an inexpensive depth sensor. Today, AR developers can choose among many software platforms, but these model systems continue to represent important directions for researchers.

Examples

In this section, we continue our exploration of AR by examining a set of examples, which showcase both AR technology and applications of that technology. We begin with application domains in which AR technologies demonstrated early success—namely, industry and construction. These examples are followed by applications in maintenance and training, and in the medical domain. We then discuss examples that focus on individuals on the move: personal information display and navigational support. Finally, we present examples illustrating how large audiences can be supported by AR using enhanced media channels in, for example, television, online commerce, and gaming.

Industry and Construction

As mentioned in our brief historic overview of AR, some of the first actual applications motivating the use of AR were industrial in nature, such as Boeing's wire bundle assembly needs and early maintenance and repair examples.

Industrial facilities are becoming increasingly complex, which profoundly affects their planning and operation. Architectural structures, infrastructure, and machines are planned using computer-aided design (CAD) software, but typically many alterations are made during actual construction and installation. These alterations usually do not find their way back into the CAD models. In addition, there may be a large body of legacy structures predating the introduction of CAD for planning as well as the need for frequent changes of the installations—for example, when a factory is adapted for the manufacturing of a new product. Planners would like to compare the “as planned” to the “as is” state of a facility and identify any critical deviations. They would also like to obtain a current model of the facility, which can be used for planning, refurbishing or logistics procedures.

Traditionally, this is done with 3D scanners and off-site data integration and comparison. This process is lengthy and tedious, however, and it results in low-level models consisting of point clouds. AR offers the opportunity to perform on-site inspection, bringing the CAD model to the facility rather than the reverse. Georgel et al. [2007], for example, have developed a technique for still-frame AR that extracts the camera pose from perspective cues in a single image and overlays registered, transparently rendered CAD models (Figure 1.12).

Schönfelder and Schmalstieg [2008] have proposed a system based on the Planar (Figure 1.13), an AR display on wheels with external tracking. It provides fully interactive, real-time discrepancy checking for industrial facilities.



Figure 1.12 AR can be used for discrepancy analysis in industrial facilities. These images show still frames overlaid with CAD information. Note how the valve on the right-hand side was mounted on the left side rather than on the right side as in the model. Courtesy of Nassir Navab.



Figure 1.13 The Planar is a touchscreen display on wheels (left), which can be used for discrepancy analysis directly on the factory floor (right). Courtesy of Ralph Schönenfelder.

Utility companies rely on geographic information systems (GIS) for managing underground infrastructure, such as telecommunication lines or gas pipes. The precise locations of the underground assets are required in a variety of situations. For example, construction managers are legally obliged to obtain information on underground infrastructure, so that they can avoid any damage to these structures during excavations. Likewise, locating the reason for outages or updating outdated GIS information frequently requires on-site inspection. In all these cases, presenting an AR view that is derived from the GIS and directly registered on the target site can significantly improve the precision and speed of outdoor work [Schall et al. 2008]. Figure 1.14 shows Vidente, one such outdoor AR visualization system.

Camera-bearing micro-aerial vehicles (drones) are increasingly being used for airborne inspection and reconstructions of construction sites. These drones may have some degree of autonomous flight control, but always require a human operator. AR can be extremely useful in locating the drone (Figure 1.15), monitoring its flight parameters such as position over ground, height, or speed, and alerting the operator to potential collisions [Zollmann et al. 2014].



Figure 1.14 Tablet computer with differential GPS system for outdoor AR (left). Geo-registered view of a virtual excavation revealing a gas pipe (right). Courtesy of Gerhard Schall.



Figure 1.15 While the drone has flown far away and is barely visible, its position can be visualized using a spherical AR overlay. Courtesy of Stefanie Zollmann.

Maintenance and Training

Understanding how things work, and learning how to assemble, disassemble, or repair them, is an important challenge in many professions. Maintenance engineers often devote a large amount of time to studying manuals and documentation, since it is often impossible to memorize all procedures in detail. AR, however, can present instructions directly superimposed in the field of view of the worker. This can provide more effective training, but, more importantly, allows personnel with less training to correctly perform the work. Figure 1.16 reveals how AR can assist with the removal of the brewing unit of an automatic coffee maker, and Figure 1.17 shows the disassembly sequence for a valve [Mohr et al. 2015].

If human support is sought, AR can provide a shared visual space for live mobile remote collaboration on physical tasks [Gauglitz et al. 2014a]. With this approach, a remote expert can explore the scene independently of the local user's current camera position and can communicate via

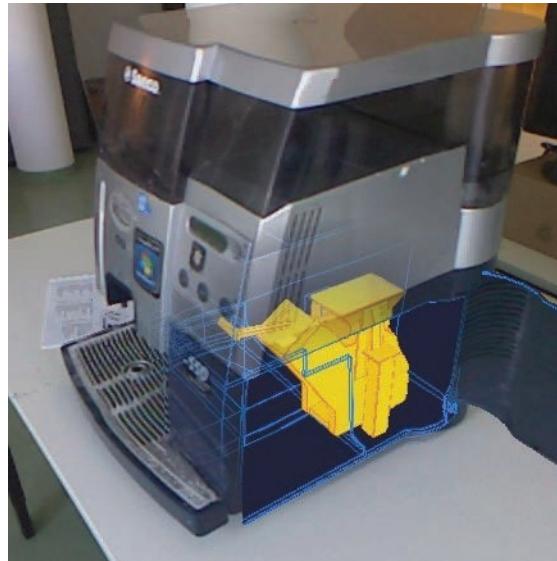


Figure 1.16 Ghost visualization revealing the interior of a coffee machine to guide end-user maintenance. Courtesy of Peter Mohr.

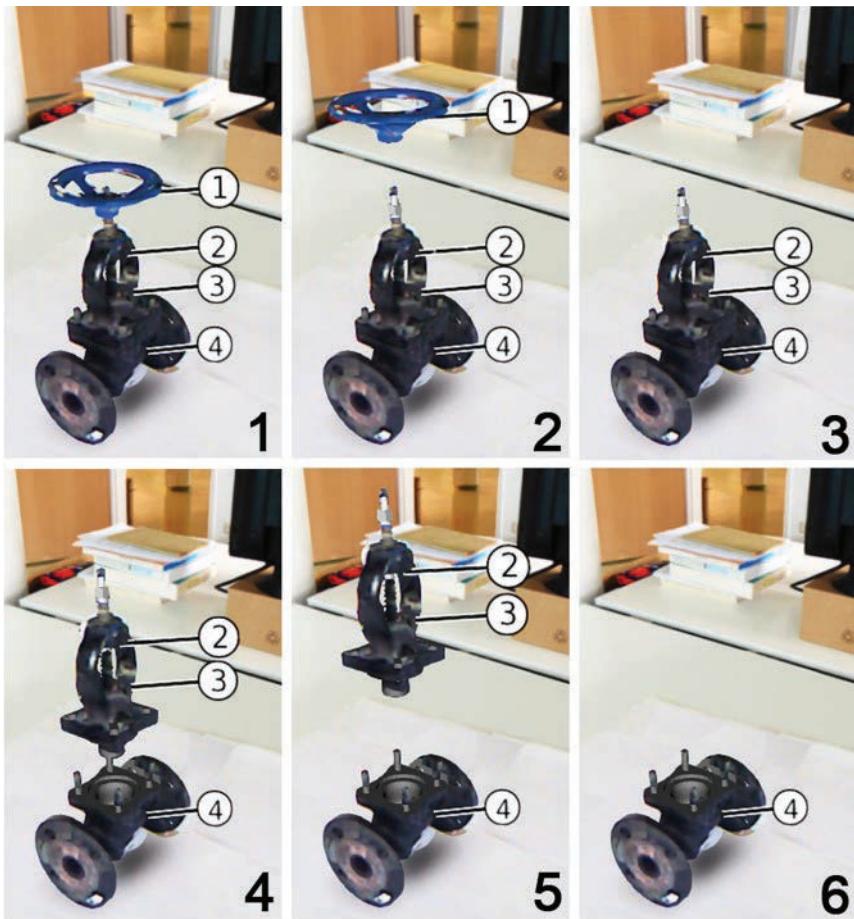


Figure 1.17 Automatically generated disassembly sequence of a valve. Courtesy of Peter Mohr.

spatial annotations that are immediately visible to the local user in the AR view (Figure 1.18). This can be achieved with real-time visual tracking and reconstruction, eliminating the need for preparation or instrumentation of the environment. AR telepresence combines the benefits of live video conferencing and remote scene exploration into a natural collaborative interface.

Medical

The use of X-ray imaging revolutionized diagnostics by allowing physicians to see inside a patient without performing surgery. However, conventional X-ray and computed tomography devices separate the interior view from the exterior view of the patient. AR integrates these



Figure 1.18 A car repair scenario assisted by a remote expert via AR telepresence on a tablet computer (top). The remote expert can draw hints directly on the 3D model of the car that is incrementally transmitted from the repair site (bottom). Courtesy of Steffen Gauglitz.

views, enabling the physician to see directly inside the patient. One such example, which is now commercially available, is the Camera Augmented Mobile C-arm, or CamC (Figure 1.19). A mobile C-arm is used to provide X-ray views in the operating theater. CamC extends these views with a conventional video camera, which is arranged coaxially with the X-ray optics to deliver precisely registered image pairs [Navab et al. 2010]. The physician can transition and blend

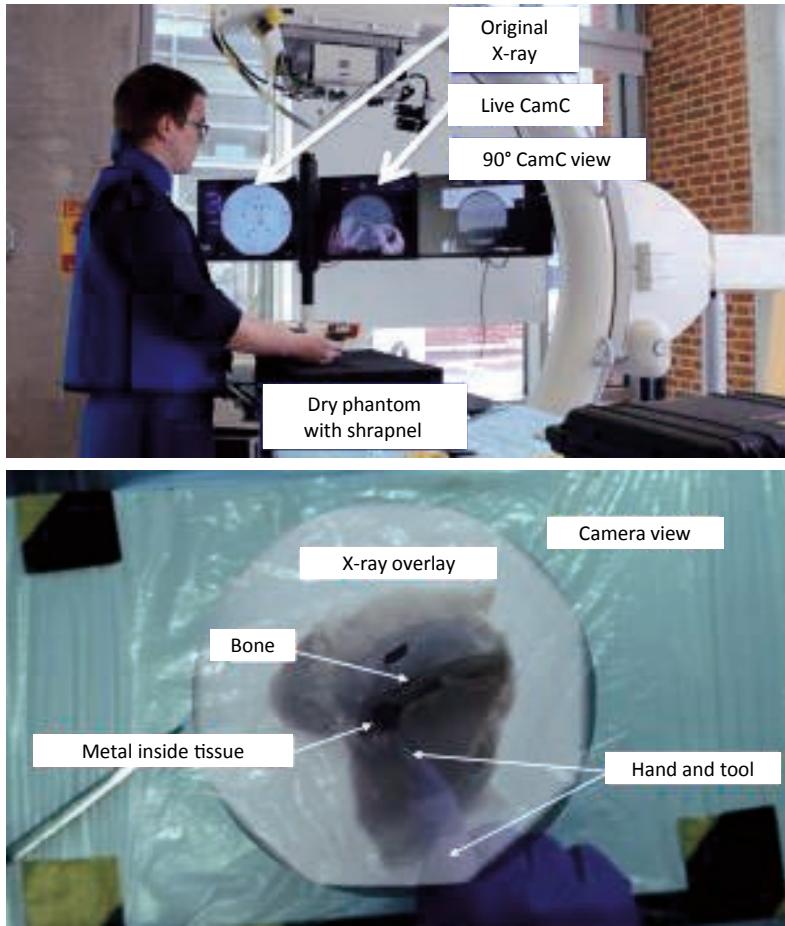


Figure 1.19 The CamC is a mobile C-arm, which allows a physician to seamlessly blend between a conventional camera view and X-ray images. Courtesy of Nassir Navab.

between the inside and outside views as desired. CamC has many clinical applications, including guiding needle biopsies and facilitating orthopedic screw placement.

Personal Information Display

As we have seen, several specific application domains can profit from the use of AR technology. But can this technology be applied more broadly to support larger audiences in completing everyday tasks? Today, this question is being answered with a resounding “yes.” A large variety of AR browser apps are already available on smartphones (e.g., Layar, Wikitude, Junaio, and others). These apps are intended to deliver information related to *places of interest* in the user’s environment, superimposed over the live video from the device’s camera. The places of interest

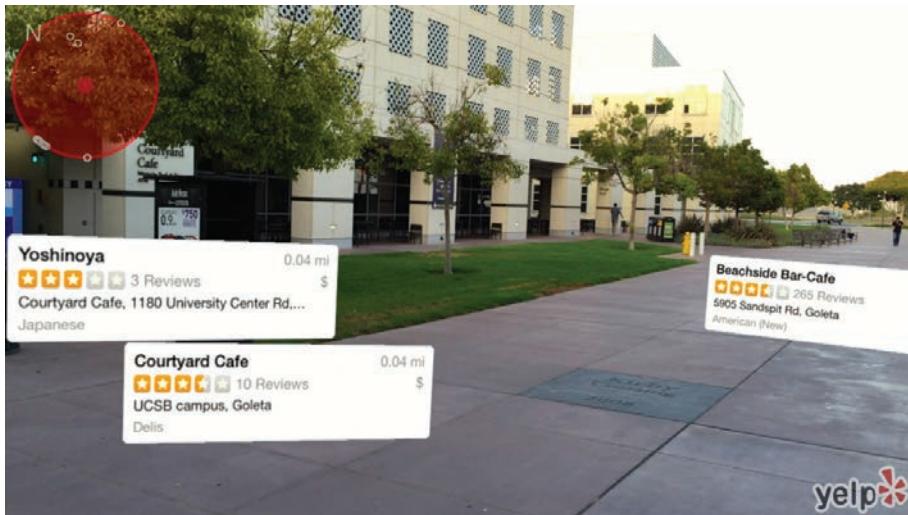


Figure 1.20 AR browsers such as Yelp Monocle superimpose points of interest on a live video feed.

are either given in geo-coordinates and identified via the phone's sensors (GPS, compass readings) or identified by image recognition. AR browsers have obvious limitations, such as potentially poor GPS accuracy and augmentation capabilities only for individual points rather than full objects. Nevertheless, thanks to the proliferation of smartphones, these apps are universally available, and their use is growing, owing to the social networking capabilities built into the AR browsers. Figure 1.20 shows the AR browser *Yelp Monocle*, which is integrated into the social business review app *Yelp*.

Another compelling use case for AR browsing is simultaneous translation of foreign languages. This utility is now widely available in the *Google Translate* app (Figure 1.21). The user just has to select the target language and point the device camera toward the printed text; the translation then appears superimposed over the image.

Navigation

The idea of heads-up navigation, which does not distract the operator of a vehicle moving at high speeds from the environment ahead, was first considered in the context of military aircraft [Furness 1986]. A variety of see-through displays, which can be mounted to the visor of a pilot's helmet, have been developed since the 1970s. These devices, which are usually called heads-up displays, are mostly intended to show nonregistered information, such as the current speed or torque, but can also be used to show a form of AR. Military technology, however, is usually not directly applicable to the consumer market, which demands different ergonomics and pricing structures.



Figure 1.21 Google Translate superimposes spontaneous translations of text, recognized in real time, over the camera image.

With improved geo-information, it has become possible to overlay larger structures on in-car navigation systems, such as road networks. Figure 1.22 shows *Wikitude Drive*, a first-person car navigation system. The driving instructions are overlaid on top of the live video feed rather than being presented in a map-like view. The registration quality in this system is acceptable despite being based on smartphone sensors such as GPS, as the inertia of a car allows the system to predict the geography ahead with relative accuracy.

Figure 1.23 shows a parking assistant, which overlays a graphical visualization of the car trajectory onto the view of a rear-mounted camera.

Television

Many people likely first encountered AR as annotations to live camera footage brought to their homes via broadcast TV. The first and most prominent example of this concept is the virtual 1st & 10 line in American football, indicating the yardage needed for a first down, which is superimposed directly on the TV screencast of a game. While the idea and first patents for creating such on-field markers for football broadcasts date back to the late 1970s, it took until 1998 for the concept to be realized. The same concept of annotating TV footage with virtual overlays has successfully been applied to many other sports, including baseball, ice hockey, car racing,



Figure 1.22 Wikitude Drive superimposes a perspective view of the road ahead. Courtesy of Wikitude GmbH.



Figure 1.23 The parking assistant is a commercially available AR feature in many contemporary cars. Courtesy of Brigitte Ludwig.



Figure 1.24 Augmented TV broadcast of a soccer game. Courtesy of Teleclub and Vizrt, Switzerland (LiberoVision AG).

and sailing. Figure 1.24 shows a televised soccer game with augmentations. The audience in this incarnation of AR has no ability to vary the viewpoint individually. Given that the live action on the playing field is captured by tracked cameras, interactive viewpoint changes are still possible, albeit not under the end-viewer's control.

Several competing companies provide augmentation solutions for various broadcast events, creating convincing and informative live annotations. The annotation possibilities have long since moved beyond just sports information or simple line graphics, and now include sophisticated 3D graphics renderings of branding logos or product advertisements.

Using similar technology, it is possible—and, in fact, common in today's TV broadcasts—to present a moderator and other TV personalities in virtual studio settings. In this application, the moderator is filmed by tracked cameras in front of a green screen and inserted into a virtual rendering of the studio. The system even allows for interactive manipulation of virtual props.

Similar technologies are being used in the film industry, such as for providing a movie director and actors with live previews of what a film scene might look like after special effects or other compositing has been applied to the camera footage of a live set environment. This application of AR is sometimes referred to as Pre-Viz.



Figure 1.25 The lifestyle magazine *Red Bulletin* was the first print publication to feature dynamic content using AR. Courtesy of Daniel Wagner.

Advertising and Commerce

The ability of AR to instantaneously present arbitrary 3D views of a product to a potential buyer is already being welcomed in advertising and commerce. This technology can lead to truly interactive experiences for the customer. For example, customers in Lego stores can hold a toy box up to an AR kiosk, which then displays a 3D image of the assembled Lego model. Customers can turn the box to view the model from any vantage point.

An obvious target for AR is the augmentation of printed material, such as flyers or magazines. Readers of the *Harry Potter* novels know how pictures in the *Daily Prophet* newspaper come alive. This idea can be realized with AR by superimposing digital movies and animations on top of specific portions of a printed template. When the magazine is viewed on a computer or smartphone, the static pictures are replaced by animated sequences or movies (Figure 1.25).

AR can also be helpful for a sales person who is trying to demonstrate the virtues of a product (Figure 1.26). Especially for complex devices, it may be difficult to convey the internal operation with words alone. Letting a potential customer observe the animated interior allows for much more compelling presentations at trade shows and in show rooms alike.

Pictofit is a virtual dressing room application that lets users preview garments from online fashion stores on their own body (Figure 1.27). The garments are automatically adjusted to match



Figure 1.26 Marketing presentation of a Waeco air-conditioning service unit. Courtesy of magicleansapp.com.

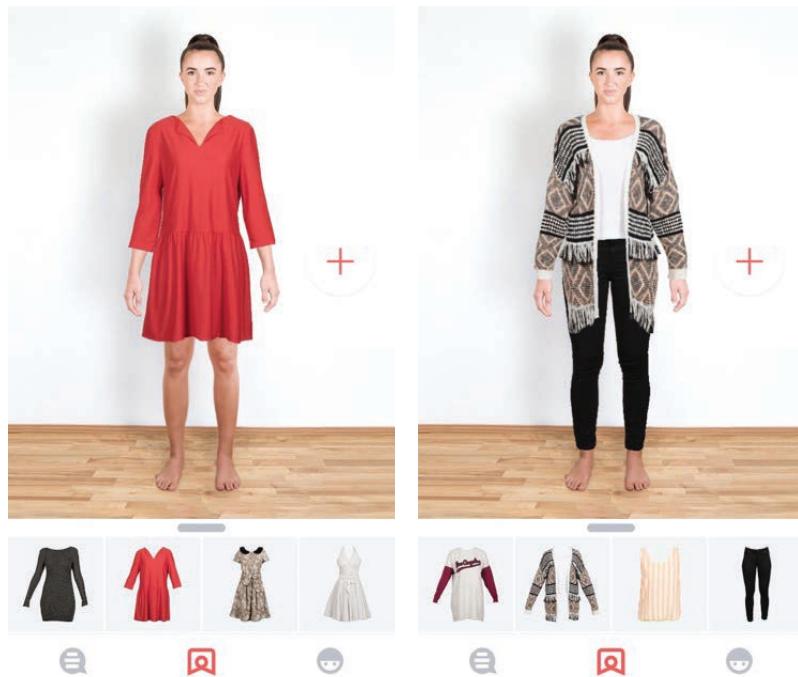


Figure 1.27 Pictofit can extract garment images from online shopping sites and render them to match an image of the customer. Courtesy of Stefan Hauswiesner, ReactiveReality.

the wearer's size. In addition, body measurements are estimated and made available to assist in the entry of purchase data.

Games

One of the first commercial AR games was *The Eye of Judgment*, an interactive trading card game for the Sony PlayStation 3. The game is delivered with an overhead camera, which picks up game cards and summons corresponding creatures to fight matches.

An important quality of traditional games is their tangible nature. Kids can turn their entire room into a playground, with pieces of furniture being converted into a landscape that supports physical activities such as jumping and hiding. In contrast, video games are usually confined to a purely virtual realm. AR can bring digital games together with the real environment. For example, Vuforia SmartTerrain (Figure 1.28) delivers a 3D scan of a real scene and turns it into a playing field for a "tower defense" game.



Figure 1.28 Vuforia SmartTerrain scans the environment and turns it into a game landscape.
© 2013 Qualcomm Connected Experiences, Inc. Used with permission.



Figure 1.29 Using a TV-plus-projector setup, the IllumiRoom extends the game world beyond the boundaries of the screen. Courtesy of Microsoft Research.

Microsoft's *IllumiRoom* [Jones et al. 2013] is a prototype of a projector-based AR game experience. It combines a regular TV set with a home-theater projector to extend the game world beyond the confines of the TV (Figure 1.29). The 3D game scene shown in the projection is registered with the one on the TV, but the projection covers a much wider field of view. While the player concentrates on the center screen, the peripheral field of view is also filled with dynamic images, leading to a greatly enhanced game experience.

Related Fields

In the previous section, we have highlighted a few AR applications. Other compelling examples of applications only tangentially match the definition we have given of AR. These applications often come from the related fields of mixed reality, ubiquitous computing, and virtual reality, which we briefly discuss here.

Mixed Reality Continuum

A user immersed in virtual reality experiences only virtual stimuli, for example, inside a CAVE (a room with walls consisting of stereoscopic back-projections) or when wearing a closed HMD. The space between reality and virtual reality, which allows real and virtual elements to be combined to varying degrees, is called **mixed reality**. In fact, some people prefer the term "mixed

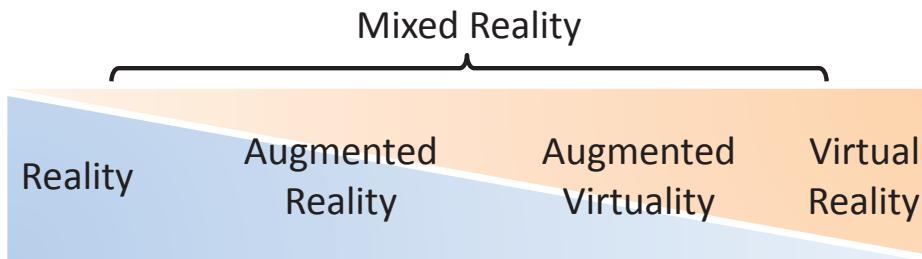


Figure 1.30 The mixed reality continuum captures all possible combinations of the real and virtual worlds.

“reality” over “augmented reality,” because they appreciate the broader and more encompassing notion of MR.

This view can be attributed to Milgram and Kishino [1994], who proposed a continuum (Figure 1.30) spanning from reality to virtual reality. They characterized MR as follows:

[MR involves the] merging of real and virtual worlds somewhere along the “virtuality continuum” which connects completely real environments to completely virtual ones.

Benford et al. [1998] go one step further, arguing that a complex environment will often be composed of multiple displays and adjacent spaces, which constitute “mixed realities” (note the plural). These multiple spaces meet at “mixed reality boundaries.”

According to this perspective, augmented reality contains primarily real elements and, therefore, is closer to reality. For example, a user with an AR app on a smartphone will continue perceiving the real world in the normal way, but with some additional elements presented on the smartphone. The real-world experience clearly dominates in such a case. The opposite concept, **augmented virtuality**, prevails when there are primarily virtual elements present. As an example, imagine an online role-playing game, where the avatars’ faces are textured in real time with a video acquired from the player’s face. Everything in this virtual game world, except the faces, is virtual.

Virtual Reality

At the far right end of the MR continuum, virtual reality immerses a user in a completely computer-generated environment. This removes any restrictions as to what a user can do or experience in VR. VR is now becoming increasingly popular for enhanced computer games. New designs for HMD gaming devices, such as the Oculus Rift or HTC Vive, are receiving a great deal of public attention. Such devices are also suitable for augmented virtuality applications. Consequently, AR and VR can easily coexist within the MR continuum. As we will see later, transitional interfaces can be designed to harness the combined advantages of both concepts.

Ubiquitous Computing

Mark Weiser proposed the concept of **ubiquitous computing** (ubicomp) in his seminal 1991 essay. His work anticipates the massive introduction of digital technology into everyday life. Contrasting ubicomp with virtual reality, he advocates bringing the “virtuality” of computer-readable data into the physical world via a variety of computer form factors, which should sound familiar to today’s technology users: inch-scale “tabs,” foot-scale “pads,” and yard-scale “boards.”

Depending on the room, you may see more than 100 tabs, 10 or 20 pads, and one or two boards. This leads to our goal for initially deploying the hardware of embodied virtuality: hundreds of computers per room. [Weiser 1991]

This description includes the idea of mobile computing, which allows users to access digital information anytime and anywhere. However, it also predicts the “Internet of Things,” in which all elements of our everyday environment are instrumented. Mackay [1998] has argued that augmented things should also be considered as a form of AR. Consider, for example, home automation, driver assistance systems in cars, and smart factories capable of mass customization. If such technology works well, it essentially disappears from our perception. The first two sentences of Weiser’s 1991 article succinctly express this model:

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

UbiComp is primarily intended as “calm computing”; that is, human attention or control is neither required nor intended. However, at some point, control will still be necessary. A human operator away from a desktop computer, for example, may need to steer complex equipment. In such a situation, an AR interface can directly present status updates, telemetry information, and control widgets in a view of the real environment. In this sense, AR and ubicomp fit extremely well: AR is the ideal *user interface* for ubicomp systems.

According to Weiser, VR is the opposite of ubicomp. Weiser notes the monolithic nature of VR environments, such as a CAVE, which isolate a user from the real world. However, Newman et al. [2007] suggest that ubicomp actually combines two important characteristics: **virtuality** and **ubiquity**. Virtuality, as described by the MR continuum, expresses the degree to which virtual and reality are mixed. Weiser considers location and place as computational inputs. Thus, ubiquity describes the degree to which information access is independent from being in a fixed place (a terminal). Based on these understandings, we can arrange a family of technologies in a “Milgram–Weiser” chart as shown in Figure 1.31.

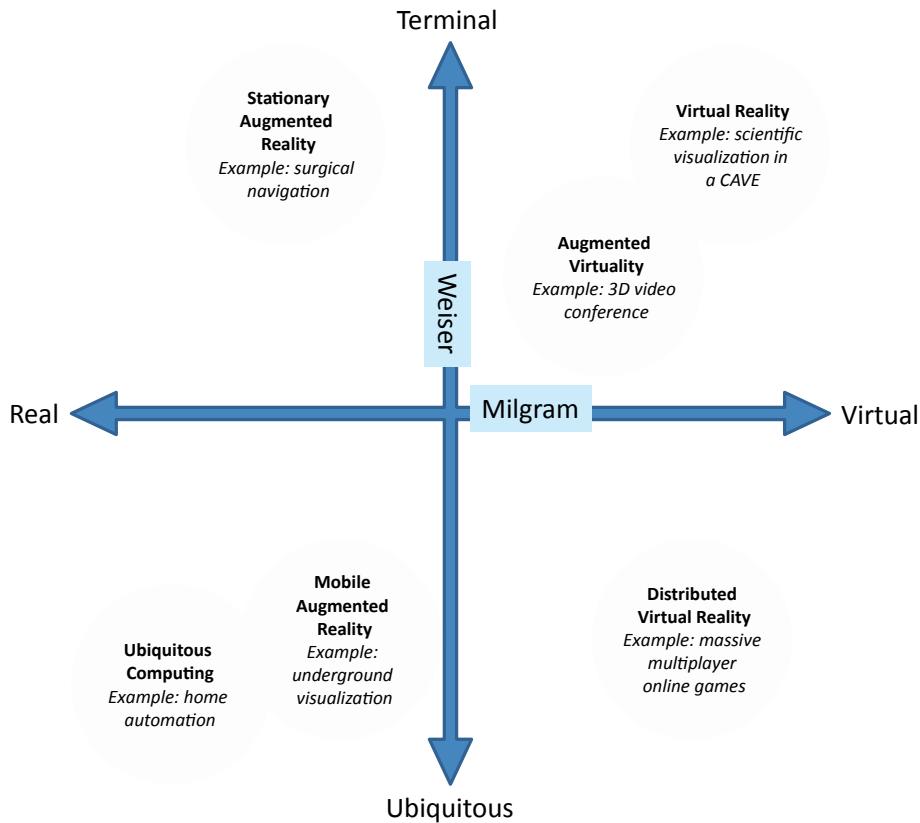


Figure 1.31 The Milgram–Weiser chart visualizes the relationships of various user interface paradigms.

Summary

In this chapter, we provided an introduction to the research field and practical occurrences of augmented reality. For a working definition, augmented reality relies on three key components: (1) the combination of virtual and real information, with the real world as the primary place of action and (2) interactive, real-time updates of (3) virtual information registered in 3D with the physical environment. Different technologies can be used to realize such a concept. The first part of this book provides an overview of technologies for displays (Chapter 2), tracking technologies (Chapters 3, 4, and 5) and graphics (Chapters 6 and 7). The second part of the book (Chapters 8 through 14) deals with interactive techniques.

We also presented a brief history of the field and then went on a whirlwind tour of AR application examples, with the goal of suggesting the enormous potential that AR holds as an interface metaphor to computing in the physical world (sometimes referred to as situated computing). While many specific application possibilities exist, such as AR for equipment maintenance or AR for surgery, one can also envision AR turning into a more general interface paradigm, redefining the overall browsing experience for computing in the physical world. Application examples from the domains of personal information display and navigation hint at that potential.

We concluded this chapter with a discussion of related fields. In doing so, we placed AR within the scope of Milgram's mixed reality continuum and contrasted AR with Weiser's concept of ubiquitous computing.

INDEX

Numbers

2D

- annotations, 325
- establishing 3D point from 2D observations, 136–137
- radar maps in navigation guidance, 350

3D

- ability of AR to produce 3D views, 25
 - annotations, 325
 - in AR games, 27
 - combining 3D tracking with 3D scanning, 106
 - establishing 3D point from 2D observations, 136–137
 - tracking, 87
 - using 3D scanners in construction and industry, 14
 - volumetric displays and, 58
- 3D Puppetry, authoring by performance, 337
- 3DOF. *See also* Degrees of freedom (DOF)
- in gyroscope, 102
 - in linear accelerometers, 104
 - in measurement systems, 92
- 3DS Max plug-in, for authoring, 340
- 6DOF. *See also* Degrees of freedom (DOF)
- complementary sensor fusion, 117
 - in measurement systems, 92
- Perspective-n-Point (PnP) problem, 146–147
- tracking/manipulating rigid objects, 280

A

- Aberrations. *See also* Distortions, 55
- Absolute measurements, versus relative, 93
- Absolute orientation
- alignment of two tracking systems, 188
 - in multiple-camera infrared tracking, 137–138
- Abstraction
- hardware abstraction, 332
 - platform abstraction, 382–383
 - UI abstraction, 383
- Accelerometers, mobile sensors for tracking, 103–104
- Accommodation. *See also* Focus
- in multifocal display, 46
 - at various distances, 45
- Accommodation-vergence conflict, 45–46, 64, 69
- Accuracy, of measurements, 95
- Acquisition stage, of VST pipeline, 197–198
- Active illumination, in optical tracking, 107

Active light probes, 208

Active searches

- incremental tracking technique, 150–151
- zero-normalized cross-correlation and, 152–153

Active sources, signals, 92

Actors, authoring elements, 333–334

Advertising, examples of use of AR, 25–27

Agile displays, as output modality, 274–276

Agile projectors, 275

A-GPS. *See* Assisted GPS

Airbrush, in specifying appearance, 318–319

AIREAL prototype, extrinsic haptic displays and, 36

AiRScouter (Brother), retinal scanning displays, 65

AJAX, producing/consuming multimedia information, 341

Albedo constant

- in diffuse reflectance, 214
- double shadowing problem and, 222–223

ALIVE system, conversational agents, 307

alVRed, for declarative scripting, 403

Amazon Fire Phone, user-tracking systems with, 184

AMIRE. *See* Authoring Mixed Reality

Angles, in determining location, 91

Annotations

- collaboration and, 325–327
- image-guided placement, 252–253
- labeling, 248–249
- legibility, 253–254
- optimization techniques, 249–250
- overview of, 248
- temporal coherence, 250–252
- use in navigation guidance, 350

Antiradiance, 225

Appearance

- free-form modeling, 322
- specifying, 317–319

Applications

- AR examples, 13, 23, 24
- AR software requirements, 380–382
- benefits of modularization, 333
- hardware abstraction and, 332

APRIL. *See* Augmented Reality Presentation and Interaction Language

AR²Hockey, agile collaboration in shared space, 369

Architecture. *See* Software architectures

Argon

- procedural scripting, 405
- web channels in Argon browser, 342

AR-Jig, for free-form modeling, 323–324

- AR Karoke, as conversational agents, 308
 ARML. *See* Augmented Reality Modeling Language
 Arrows
 guidance provided by, 350
 guidance toward target viewpoint, 354
 pointing and gestures in remote collaboration, 376
 ARTHUR, for co-located collaboration, 366
 Artifacts
 camera simulation and, 235
 exposing hidden, 254
 ARTToolKit
 marker tracking and, 123
 open-source software platform for AR, 11–12
 as programming framework, 330
 in wearable AR setup as case study, 394
 ARVIKA. *See* Augmented Reality for Development, Production and Servicing
 Assisted GPS (A-GPS)
 cooperative sensor fusion and, 118
 wireless networks in tracking, 101
 Audio Aura system, 35
 Audio displays, 34–35
 “Augment-able reality,” 325
 Augmentation
 augmented browsing, 276–279
 augmented humans, 419–420
 augmented paper, 295–297
 augmented surfaces, 298
 methods, 40
 optical see-through (OST), 40–41
 placement of, 272–274
 surfaces with multi-view interfaces, 298
 transparent tangibles and, 293
 video see-through (VST), 41–42
 Augmented reality, introduction
 advertising and commercial uses, 25–27
 benefits of, 2
 Caudell and Mizell in history of, 5–6
 definition and scope, 3–4
 example applications, 13
 Feiner in history of, 6–7, 9–10
 Fisher in history of, 5
 Fitzmaurice in history of, 6
 game uses, 27–28
 industry and construction uses, 13–16
 information display uses, 20–21
 Kato and Billinghurst in history of, 11–12
 Klein and Murray in history of, 13
 Krueger in history of, 5
 limitations of virtual reality and, 1–2
 maintenance and training uses, 17–18
 Mann in history of, 8
 medical uses, 18–20
 mixed reality related to, 28–29
 navigation uses, 21–22
 Raskar in history of, 11
 Rekimoto and Nagao in history of, 8
 related fields, 28
 Sandin in history of, 5
 Schmalstieg in history of, 9, 12
 Spoehr in history of, 12
 summary, 31–32
 Sutherland in history of, 4–5
 Tamura in history of, 9
 television uses, 22–25
 Thomas in history of, 11
 ubiquitous computing related to, 30–31
 virtual reality related to, 29
 Wagner and Schmalstieg in history of, 12
 Augmented Reality for Development, Production and Servicing (ARVIKA), 12
 Augmented Reality Modeling Language (ARML), 342
 Augmented Reality Presentation and Interaction Language (APRIL)
 AR tour guide as case study, 403–404
 declarative scripting, 402
 for desktop authoring, 336–337
 Authoring
 actors as element of, 333–334
 desktop authoring, 335–337
 hardware abstraction and, 332
 interactive behavior and, 334
 overview of, 329–331
 by performance, 337–339
 plug-in approaches, 339–340
 real-world interfaces and, 331
 setup for, 335
 stages of, 334
 story element of, 334
 summary, 342–343
 web technologies for, 341–342
 workflow for, 332–333
 Authoring Mixed Reality (AMIRE)
 desktop authoring, 336
 object management, 386
 Auto-correlation matrix, corner detectors based on, 141–142
 Avalon
 distributed shared scene graphs, 399
 object management, 387
 runtime reconfiguration, 405–406
 scene graph integration, 399
 Avango
 declarative scripting, 402–403
 distributed shared scene graphs, 399
 object management, 386
 scene graph integration, 399
 Aveyant Glyph, retinal scanning displays, 65

B

- Bag-of-words model, 166
 Barcodes
 benefits of augmented reality services, 2
 marker target identification, 114

- Bat system, ultrasonic tracking, 98
Bayer masks
 color filters, 235
 distortions and aberrations in OST and VST
 HMDs, 55
Bayer pattern artifacts, camera simulation
 and, 235
Billboard, rendering blur effects, 233
Binocular depth cues, 39, 196–197
Binocular HMD, 42–43
Blindness, uses of haptic displays, 36
Blob detection, in multiple-camera infrared tracking, 133
Bluetooth, 101–102
Blur effects, camera simulation and, 232–233
Body
 augmentation and, 273
 body tracking techniques, 281–282
Body-stabilized, frames of reference, 89–90
“Bore-sight” approach, to display calibration, 183–184
Bounding boxes, in label placement, 250
BRIEF, in descriptor creation, 144
Brightness, requirements for visual displays, 53–54
Brother AiRScouter, retinal scanning displays, 65
Browsers
 AR apps on smartphones, 20–21
 first concept for AR browser, 12
 social computing platforms and, 421
 web channels in Argon browser, 342
Bundle adjustment, for drift correction in mapping, 158–159
Business cases, drivers
 consumers, 411
 overview of, 410
 professional users, 410–411
Buttons, defining interactive behaviors, 334
- C**
- C, mixed language programming, 405
C++
 mixed language programming, 405
 object management, 386
C#
 object management, 386
 procedural scripting, 405
CAD. *See* Computer-aided design
Calibration
 of authoring hardware, 335
 of cameras, 180
 comparing with tracking and registration, 86–87
 correcting lens distortion, 182–183, 232
 of displays, 183–185
 geometric measurement distortions and, 190–191
 hand-eye calibration, 188–190
 of HMD using pointing device, 186–187
 internal camera parameters and, 180–181
 of measurements, 86–87
 overview of, 180
 single point active alignment method (SPAAM), 185–186
 summary, 194
 of virtual camera, 51
 in visual coherence, 196–198
Calibration target, 180
Camera Augmented Mobile C-arm (CamC), medical uses of AR, 19–20
Camera simulation
 Bayer pattern artifacts, 235
 blur effects, 232–233
 chromatic aberrations, 234–235
 compensating for lens distortion, 231–232
 compositing and, 199
 matching noise effects, 234
 overview of, 231
 tone mapping artifacts, 235
 vignetting, 234
Camera tracking, 57
Cameras. *See also* Multiple-camera infrared tracking
 calibration of, 180
 correcting lens distortion, 182–183
 depth cues, 196–197
 dynamic lighting, 225
 future features, 412–413
 internal and external parameters, 125–126, 180–181
 optical tracking and, 105–106
 PointGrey Ladybug camera, 119
 representing in marker tracking, 124–125
 triangulation from more than two cameras, 136–137
 triangulation from two cameras, 135–136
 virtual mirror using front-facing, 72–74
 YUV format (luminance and chrominance), 235
Cascading Style Sheets (CSS), producing/consuming multimedia information, 341
Cathode ray tubes (CRTs), resolution and refresh rate, 49
CAVE, object management with, 386
CCD. *See* Charge-coupled device
Ceiling lights, for passive illumination, 106
Cell phones, in history of augmented reality, 12
Chameleon display, in history of augmented reality, 6
Channels, user access to information channels via web, 341–342
Charge-coupled device (CCD), 105
Chromatic aberrations, camera simulation and, 234–235
Client-side scripting, producing/consuming multimedia information, 341

- CMOS. *See* Complementary metal oxide semiconductor
- Co-located collaboration
- agility in shared space, 368–369
 - gaze awareness, 368
 - individual displays and views, 366–368
 - overview of, 364–366
- Co-Optical Axis See-Through Augmented Reality (COASTAR)
- coaxial stereo VST HMD, 9
 - parallax-free VST HMD, 67
 - viewpoint offset, 52–53
- CoCube, use of tangible interfaces, 292
- Collaboration
- agility in shared space, 368–369
 - annotations supporting, 325–327
 - co-located, 364–366
 - gaze awareness, 368
 - individual displays and views, 366–368
 - modularized applications supporting, 333
 - overview of, 361–362
 - pointing and gestures, 375–376
 - properties of collaboration systems, 362–363
 - remote collaboration, 370–371
 - remote collaboration with agile users, 376–377
 - summary, 377
 - video sharing, 371–372
 - video sharing with geometric reconstruction, 374–375
 - video sharing with virtual objects, 372–373
- Color
- Bayer masks for filtering, 235
 - chromatic aberrations, 234–235
 - tone mapping artifacts, 235
- Columbia MARS, OST display designs, 61
- Common illumination. *See also* Illumination
- compositing and, 198
 - differential rendering, 216–217
 - diffuse global illumination, 223–225
 - overview of, 216
 - real-time global illumination, 218–220
 - shadows and, 220–223
 - specular global illumination, 225–227
- Communication space, versus task space in collaboration, 362–363
- Compatibility issues
- hardware abstraction and, 332
 - platform abstraction and, 382–383
- Competitive sensor fusion, 117–118
- Complementary metal oxide semiconductor (CMOS), 105
- Complementary sensor fusion, 117
- Compositing
- camera simulation. *See* Camera simulation
 - common illumination. *See* Common illumination
 - diminished reality. *See* Diminished reality
- occlusion. *See* Occlusion
- scene compositing, 259–260
- stages of VST pipeline, 197–199
- stylization. *See* Stylized AR
- Computed tomography (CT), medical uses of AR, 18–20
- Computer-aided design (CAD), examples of use of AR, 14
- Computer mouse, as 2D odometry, 104
- Computer-supported cooperative work (CSCW), 362
- Computer vision
- absolute orientation, 137–138
 - active search tracking technique, 150–151
 - blob detection, 133
 - bundle adjustment for drift correction, 158–159
 - camera representation, 124–126
 - combining detection and tracking, 155–156
 - dense mapping, 161–164
 - descriptor creation, 144–145
 - descriptor matching, 145–146
 - determining corners using Harris detector, 141–142
 - differences of Gaussian (DOG) detection method, 142
 - establishing point correspondences, 133–135
 - features from accelerated segment test (FAST) detection method, 143–144
 - hierarchical searches, 154–155
 - incremental tracking, 149–150
 - interest point detection, 140–141
 - Kanade-Lucas-Tomasi (KLT) tracker, 151–152
 - marker detection, 126–128
 - marker tracking, 123–124
 - matching targets consisting of spherical markers, 137
 - multiple-camera infrared tracking, 132–133
 - natural feature tracking by detection, 138–140
 - Nister's five-point algorithm, 157–158
 - outdoor tracking, 164–165
 - overview of, 121–122
 - parallel tracking and mapping, 159–160
 - Perspective-n-Point (PnP) problem, 146–147
 - pose estimation from homography, 128–131
 - pose refinement, 132
 - prior information from geometry, 169–170
 - prior information from sensors, 167–169
 - relocalization and loop closure, 160–161
 - robust pose estimation, 148–149
 - scalable visual matching, 165–167
 - simultaneous localization and mapping (SLAM), 156–157
 - simultaneous tracking, mapping, and localization, 170–176
 - summary, 176–177
 - triangulation from more than two cameras, 136–137

- triangulation from two cameras, 135–136
zero-normalized cross-correlation (ZNCC), 152–153
- Configuration**
parameter configuration, 401
runtime reconfiguration, 405–406
- Construct3D**
co-located collaboration, 365–366
free-form modeling, 324
- Construction**, examples of use of AR, 13–16
- Constructive solid geometry**
modeling volumes, 316
in semi-automatic reconstruction, 321
- Consumers**, as business case driver, 411
- Contact lens**
for HMD focus optics, 62
future of optical see-through displays, 415
social acceptance of HMDs and, 56
- Content**, 330
- Content actors**, as objects, 333–334
- Content design frameworks**, 330
- Context Compass**, use in navigation guidance, 351–352
- Context-sensitive (situated) computing**, 72. *See also* Situated visualization
- Contrast**, requirements for visual displays, 53–54
- Conversational agents**, 306–309
- Cooperative sensor fusion**, 118–120
- Coordinate systems**
frames of reference, 89–90
global versus local, 92–93
globally registered situated visualization, 246–247
locally registered situated visualization, 245–246
model, view, and projective transformation and, 88–89
overview of, 87–88
registration of, 179
- CORBA**
general-use middleware for AR, 385
object management, 386–387
scene graph integration, 397
- Corner detectors**, based on auto-correlation, 141–142
- COTERIE**, distributed shared scene graphs, 400
- Cows vs. Aliens game**, 369–370
- CPU/GPU processors**, future of, 413
- Cross-view interactions**, 303–304
- CRTs**. *See* Cathode ray tubes
- CSCW**. *See* Computer-supported cooperative work
- CSS**. *See* Cascading Style Sheets
- CT**. *See* Computed tomography
- CUIML**, for declarative scripting, 402
- Cultural education**, design frameworks for, 331
- Cybernetic organisms (Cyborgs)**, 419–420
- Cyber-physical systems**, interfacing with smart objects, 417
- D**
- DART**. *See* Designer's Augmented Reality Toolkit
- Dataflow**
dataflow graphs, 390, 394
in distributed object systems, 380
multimodal interactions, 390–391
overview of, 389–390
threads and scheduling, 391–393
wearable AR setup as case study, 393–395
- Data overload**, challenges in situated visualization, 242
- DataTiles**, use of transparent tangibles, 293
- Data transformation**, dealing with data overload, 242
- Data types**, multimodal interactions, 391
- Declarative scripting**, developer support for, 401–403
- Decoupled simulation**, AR engineering requirements, 384
- Degrees of freedom (DOF)**
complementary sensor fusion, 117
in gyroscope, 102
in linear accelerometers, 104
in measuring systems, 92
Perspective-n-Point problem and, 146–147
tracking/manipulating rigid objects, 280
- DEM**. *See* Digital elevation models
- Dense mapping**, SLAM, 161–164
- Dense matching**, tracking by detection, 139
- Dense SLAM**, 164
- Dependency graphs**, 397
- Depth cues**
camera calibration and, 196–197
occlusion and, 199
shadows as, 220
visual perception and, 39
- Depth image integration**, dense mapping, 163–164
- Depth maps**, indexing radiance cache in projective space, 219
- Depth of field**
focus and, 45
in multifocal display, 46
rendering depth-of-field effect using ray-tracing, 233
- Depth sensors**
developer wish list, 413
use in semi-automatic reconstruction, 321
use with hand gestures in remote collaboration, 376
- Descriptors**, natural features
creating, 144–145
matching, 145–146
tracking, 139–140
- Design frameworks**, versus programming frameworks, 330
- Designer's Augmented Reality Toolkit (DART)**, plug-in approaches to authoring, 340

- Desktop authoring, 335–337
 Desktop displays, 72
Detection. *See* Natural features, tracking by detection
 Developer support
 AR tour guide as case study, 403–404
 declarative scripting, 401–403
 mixed language programming, 405
 overview of, 400–401
 parameter configuration, 401
 platform selection, 407
 procedural scripting, 405
 runtime reconfiguration, 405–406
 Device abstraction software frameworks, 335
 DGPS. *See* Differential GPS
 Differences of Gaussian (DOG), natural feature detection method, 142
 Differential GPS (DGPS), 100
 Differential instant radiosity
 diffuse global illumination and, 224
 specular global illumination and, 225
 Differential rendering, for real-world illumination, 216–217
 Diffuse global illumination, 223–225
 Diffuse reflections
 albedo constant in, 214
 local illumination compared with global illumination, 206
 photometric registration from, 212–214
 Digital cameras. *See* Cameras
 Digital elevation models (DEM), prior information from sensors, 170
 Digital projection mapping. *See also* Projectors, 78
 DigitalDesk, augmented paper, 295
 Diminished reality
 compositing and, 198–199
 determining region of interest, 228
 observing and modeling hidden areas, 228–229
 overview of, 227
 projector-based diminished reality, 230–231
 removing region of interest, 229–230
 Direct illumination, 216
 Direct linear transformation (DLT)
 homography estimation, 128–131
 triangulation from more than two cameras, 136–137
 using with SPAAM calibration, 187
 Discovery, in navigation, 347
 Display tracking, spatial display model, 57
 Displays
 agile, 274–276
 audio, 34–35
 augmentation, 273
 brightness and contrast, 53–54
 calibration, 183–185
 calibration using pointing device, 186–187
 calibration using SPAAM, 186–187
 co-located collaboration, 366–368
 COASTAR, 9
 developer wish list, 414
 distortions and aberrations, 55
 dynamic shader lamps, 82
 ergonomics, 55
 everywhere projectors, 83
 field of view, 50–51
 focus, 45–47
 handheld, 8, 69–72
 haptic, tactile, and tangible, 35–37
 head-mounted projector displays, 80–81
 heads-up displays, 21–22
 history of augmented reality, 5–6
 multi-view interfaces, 297–298
 near-eye, 59–61
 occlusion, 47–48
 ocular and stereoscopy, 42–45
 olfactory and gustatory, 37–39
 optical see-through, 40–41
 optical see-through HMDs, 61–66
 overview of, 33–34
 projected, 78
 requirements, 40, 381
 resolution and refresh rate, 48–49
 social acceptance of, 55–56
 spatial AR projectors, 78–80
 spatial display model, 56–57
 stationary and desktop, 72
 summary, 84
 Sutherland on immersive, 4–5
 video see-through (VST), 41–42
 video see-through (VST) HMDs, 66–69
 viewpoint offset, 51–53
 virtual mirror, 72–74
 virtual showcase, 74–76
 visual, 58–59
 visual perception and, 39
 window and portal, 75–77
 Distance, odometers measuring, 104
 Distinct shapes, tangible interfaces for, 289–292
 Distortions
 correcting lens distortion, 182–183, 231–232
 geometric measurement, 190–191
 in OST and VST HMDs, 55
 requirements for visual displays, 55
 space distortion, 262–265
 Distributed computing, AR engineering requirements, 384
 Distributed object systems
 AR building on, 380
 overview of, 385–386
 SHEEP game, 387–389
 software for managing objects, 386–387
 Distributed Open Inventor, distributed shared scene graphs, 400
 Distributed user interface, 297

Districts, spatial knowledge for navigation, 346
 DLT. *See* Direct linear transformation
 DOF. *See* Degrees of freedom
 DOG. *See* Differences of Gaussian
 Dominant planes, estimating in reconstruction, 319–320
 DoubleDigitalDesk, pointing and gestures in remote collaboration, 375
 Double shadowing problem, 222–223
 Dramatic uses of AR, 420–421
 Dual phantom rendering algorithm, 262
 DWARF
 object management, 387
 as research framework, 330
 runtime reconfiguration, 405–406
 scene graph integration, 397
 SHEEP game case study, 387–389
 Dynamic error, registration and, 190
 Dynamic lighting, from RGB-D camera, 225
 Dynamic registration, 87
 Dynamic shader lamps, 82

E

Edges
 ghostings from image space and, 256–257
 natural feature tracking, 112
 spatial knowledge for navigation, 346
 Ego-centric
 map perspective, 354–357
 navigation tasks, 346
 transitional interfaces for multiple views, 358–359
 Electromagnetic radiation, sensors for physical phenomena, 90–91
 Electromagnetic tracking, types of stationary tracking systems, 97
 Electronic compass, mobile sensors for tracking, 102
 ELMO HMD, 48
 Embodied conversational agent, 306
 EMMIE
 co-located collaboration, 366
 uses of cross-view interaction, 303
 Engineering requirements, for AR software, 382–384
 Entities, defining entity behavior in authoring workflow, 332
 Environment
 control and scene dynamics, 380–381
 outdoor tracking and, 164
 Environment maps
 capturing offline light, 210
 for image-based lighting, 207–208
 indexing radiance cache by orientation, 219
 modeling incoming light, 206–207
 spherical and cubic, 209

Epipolar line, establishing point correspondences, 134–135
 Epson Moverio, display mounting options, 60
 Ergonomics
 of near-eye displays, 60
 requirements for visual displays, 55
 Errors
 compensating for jitter, 192–194
 registration and, 191–192, 243
 static and dynamic, 190
 Essential matrix, establishing point correspondences, 134–135
 Events
 dataflow, 389
 handling aggregate, 391
 Everywhere Displays Projector, 83
 Exo-centric
 map perspective, 354–357
 navigation tasks, 347
 Exploration
 advantages of AR for, 347
 aspects of navigation, 346
 Explosion diagrams, in spatial manipulation, 260–262
 Extended Kalman filter, compensating for jitter, 193
 Extensibility, AR engineering requirements, 383–384
 eXtensible Markup Language (XML)
 ARML as XML dialect, 342
 declarative scripting, 402
 Extrinsic haptic displays, 36
 Extrusion, creating volumes, 315
 Eye. *See* Visual perception
The Eye of Judgment game, 27
 Eye-through HMD, gaze awareness in collaboration, 368
 Eye tracking
 for display calibration, 185
 spatial display model, 57

F

Facebook, using AR as social computing platform, 422
 FAST. *See* Features from accelerated segment test
 “Fat finger problem,” touch interaction and, 284
 Features from accelerated segment test (FAST)
 detecting interest points, 157
 natural feature detection method, 143–144
 Fiber-optic (laser) gyroscopes, 103
 Fiducials. *See* Markers
 Field of view (FOV)
 cameras with wide, 413
 of near-eye displays, 60
 OST displays, 61
 requirements for visual displays, 50–51
 VST displays, 67
 wide-FOV displays, 67–68, 414

- Field-sequential color displays, 49
 Film industry, Pre-Viz AR application, 24
 Filter nodes, dataflow graphs, 390
 Filters
 applying stylization filters, 199
 information filtering. *See* Information filtering
 for reducing data overload, 242
 registration and, 192–194
 Fingerprinting, wireless networks in tracking, 101
 Fire Phone (Amazon), user-tracking systems
 with, 184
 FLARToolKit, plug-in approaches to authoring, 340
 Flash, plug-in approaches to authoring, 340
 Flashlight (handheld projector), browsing with, 277
 Foam cutting, free-form modeling, 323
 Focal depth, of displays, 45
 Focus
 blur effects, 232–233
 displays with variable, 415
 multifocal displays, 46–47
 near-eye focus issue, 62–63
 overview of, 45–46
 Focus+Context relationship
 multi-display interface, 297–298
 tangibles and surfaces, 293
 FogScreen, transparent immaterial display, 76–77
 Food Simulator, uses of olfactory and gustatory
 displays, 37–38
 Force fields, treating label placement as
 optimization problem, 250
 Form factor, display designs, 61
 FOV. *See* Field of view
 FOV2GO, 68
 Frame of reference, coordinate systems, 89–90
 Frameworks
 device abstraction software framework, 335
 hardware abstraction in, 332
 platform abstraction, 382–383
 platform selection, 407
 programming versus content design
 frameworks, 330–331
 scene graph integration, 397–399
 Free-form modeling, 322–324
 Frustum, guidance toward target viewpoint, 354
 Fundamental matrix, establishing point
 correspondences, 134–135
 Future of
 augmented humans, 419–420
 camera-related features, 412–413
 depth and pose sensors, 413
 displays, 414–415
 dramatic uses of AR, 420–421
 memory and processors, 413–414
 outdoor AR, 415–417
 overview of, 409–410
 parallel programming, 414
 smart object interfaces, 417–418
 social computing uses of AR, 421–422
 summary, 422
 VR/AR integration, 418–419
- ## G
- Games
 AR as dramatic medium, 420–421
 collaboration in, 369–370
 content-driven nature of, 330
 examples of use of AR, 27–28
 history of augmented reality and, 12
 Gaussian filters, rendering blur effects, 233
 Gauss–Newton method, nonlinear optimization, 132
 Gaze awareness, in co-located collaboration, 368
 Gaze objects, recording specular reflections, 209–210
 G-buffers. *See* Geometric buffers
 GearVR (Samsung), 68
 General-purpose graphics processing unit
 (GPGPU), 413–414
 Generic shapes, tangible interfaces for, 287–289
 Geographic information systems (GIS)
 audio display use and, 35
 construction and industry examples of use of
 AR, 15–16
 outdoor AR visualization system, 15–16
 prior information from sensors, 169
 Geometric buffers (G-buffers)
 buffer rendering, 258–259
 implementing ghosting, 258–260
 processing, 259
 scene compositing, 259–260
 Geometry
 constructive solid geometry, 316
 determining location, 91
 measurement distortions in sensors, 190–191
 prior information from, 169–170
 scene understanding based on geometric
 information, 322
 sharing video with geometric reconstruction,
 374–375
 specifying in general, 312
 specifying planes, 314–315
 specifying points, 313–314
 specifying volumes, 315–317
 Gestures
 body tracking techniques, 281–283
 enhancement of annotations, 325
 as input modality, 282–283
 manipulating rigid objects, 280–281
 PaperWindows and, 296–297
 in remote collaboration, 375–376
 Ghosting
 from image space, 256–258
 implementing with G-buffers, 258–260

- from object space, 255–256
overview of, 254
- GIS.** See *Geographic information systems*
- Glasstron OST display (SONY), 61–62
- Global coordinates, versus local, 92–93
- Global illumination
common illumination and, 216
diffuse, 223–225
local illumination compared with, 205–206
real-time, 218–220
reconstructing explicit light sources, 215–216
specular, 225–227
- Global Positioning System (GPS)
AR browser apps on smartphones, 21
audio display use and, 35
location-based services in mobile computing, 2
mobile sensors for tracking, 99–101
in outdoor AR, 16
prior information from sensors, 167–169
sensor fusion and, 117–118
wireless networks in tracking, 101–102
- Global registration, visualization registration, 246–247
- Glyphs, guidance provided by, 350
- Google Cardboard, 68
- Google Earth, storing geo-referenced interest points, 341
- Google Glass
augmented humans, 420
social acceptance of HMDs and, 56
spatial arrangement, 60
visual display mounting options, 60
as VST display, 67
- Google Maps, prior information from sensors, 169
- Google Project Tango, 407, 413
- Google Sketchup, plug-in approaches to authoring, 340
- GP GPU (general-purpose graphics processing unit), 414
- GPS.** See *Global Positioning System*
- GPUs. See *Graphical processing units*
- Graph traversal, processing scene graphs, 396
- Graphical processing units (GPUs)
mobile, 414
treating label placement as optimization problem, 250
- Graphs
dataflow graphs, 390
dependency graphs, 397
distributed shared scene graphs, 399–400
scene graphs, 395–397
- Gravity
measuring with linear accelerometer, 104
sensors for physical phenomena, 90–91
- Gravity sensor, defining vertical planes, 314
- Gravity-Space, touch interaction and, 285
- Gustatory displays, 37–39
- Gyroscopes
linear accelerometers, 103–104
mobile sensors for tracking, 102–103
- ## H
- Halos, use in navigation guidance, 351
- Hand-eye calibration, alignment of tracking systems, 188–190
- Handheld devices, Indoor Signpost for, 348
- Handheld displays
agile displays, 274, 276
browsing with, 277
overview of, 69–72
personal displays in collaboration, 366–368
social acceptance of, 56
user-tracking systems with, 184
- Handheld projectors, browsing with, 277
- "Hand of God" (HOG), pointing and gestures in remote collaboration, 376
- Hand-referenced display, 273
- Hand tracking, body tracking techniques, 281–283
- Hand video, pointing and gestures in remote collaboration, 376
- HandyAR, gestures and, 283
- Haptic displays, 35–37
- Haptic interaction, 304
- Haptic reality, 37
- Haptic virtuality, 37
- Hardware
abstraction in frameworks, 332
setup of authoring hardware, 335
- Harris detector
detecting interest points, 157
determining corners, 141–142
- HDR. See *High dynamic range*
- Head-mounted displays (HMDs)
agile displays, 274–276
brightness and contrast, 53–54
browsing with, 277
calibration generally, 183–185
calibration using pointing device, 186–187
calibration using SPAAM, 185–186
Caudel and Mizells' see-through HMD, 5–6
in co-located collaboration, 366
COASTAR, 9
in collaboration, 368
distortions and aberrations, 55
ELMO HMD, 48
ergonomics, 55
field of view, 51
gaming devices and, 29
head-mounted projector displays (HMPD), 80–81
Mann's "reality mediator," 8
near-eye displays, 59–61
not essential to definition of AR, 3

- Head-mounted displays (HMDs) (*continued*)
 ocularity and stereoscopy, 42–43
 optical see-through (OST), 61–66
 personal displays in collaboration, 366–368
 social acceptance of, 55–56
 Sword of Damocles, 5
 Touring Machine, 9–11
 video see-through (VST), 66–69
 viewpoint offset, 51–53
- Head-mounted projector displays (HMPD), 80–81
- Head-referenced display, augmentation and, 273
- Head-related transfer function (HRTF), 35
- Heads-up display (HUD), in AR system, 90
- Heads-up navigation, examples of use of AR, 21–22
- Head tracking, spatial display model, 57
- Hidden areas
 observing and modeling, 228–229
 problem areas addressed in diminished reality, 227
- Hierarchical searches
 descriptor matching, 145
 incremental tracking technique, 154–155
- High dynamic range (HDR)
 obtaining HDR images using light probe, 209
 using environment maps for image-based lighting, 207
- History of augmented reality
 Caudell and Mizell in, 5–6
 Feiner in, 6–7, 9–10
 Fisher in history of, 5
 Fitzmaurice in, 6
 Kato and Billinghurst in, 11–12
 Klein and Murray in, 13
 Krueger in history of, 5
 Mann in, 8
 Raskar in, 11
 Rekimoto and Nagao in, 8
 Sandin in history of, 5
 Schmalstieg in, 9, 12
 Spoher in, 12
 Sutherland in, 4–5
 Tamura in, 9
 Thomas in, 11
 Wagner and Schmalstieg in, 12
- HMDs. *See* Head-mounted displays
- HMPD. *See* Head-mounted projector displays
- HOG. *See* “Hand of God”
- HoloDesk, touch interaction and, 286
- Holograms
 as display technology, 58–59
 verifying authenticity of, 353
- Holographic displays, types of visual displays, 58–59
- HoloLens (Microsoft)
 near-eye displays, 64
 optical see-through display, 44
 platform selection, 407
 spatial audio support, 35
 wide-field-of-view cameras, 413
- Homography
 estimation, 128–130
 pose estimation, 130–131
- Horn algorithm, absolute orientation, 138
- HRTF. *See* Head-related transfer function
- HSV color space, image-guided placement of labels, 252
- HTC/Valve Steam VR, use of HMDs in gaming devices, 29
- HTML5, producing/consuming multimedia information, 341
- HUD. *See* Heads-up display
- Human body. *See* Body
- Human vision. *See* Visual perception
- I**
- Ice, general use middleware for AR, 385
- ICP. *See* Iterative closest point algorithm
- ID (Identifier), wireless network, 101
- Illumination
 active illumination, 107
 common illumination, 216
 differential rendering for real-world illumination, 216–217
 diffuse global illumination, 223–225
 local illumination compared with global illumination, 205–206
 in optical tracking, 106
 passive illumination, 106–107
 real-time global illumination, 218–220
 reconstructing explicit light sources, 215–216
 shadows and, 220–223
 specular global illumination, 225–227
 storing directional illumination, 212–213
 structured light, 107–109
- IllumiRoom* projector, use in AR games, 28
- Image-based lighting
 environment maps for, 207–208
 radiance maps for, 219
- Image-guided placement, of labels, 252–253
- Image space, ghostings from, 256–258
- ImageTcLAR, procedural scripting, 405
- Imitation Game, 418–419
- Immersion factors, AR displays and, 42
- IMU. *See* Inertial measurement unit
- Incremental tracking
 combining detection and tracking, 155–156
 hierarchical searches, 154–155
 Kanade-Lucas-Tomasi (KLT) tracker, 151–152
 overview of, 149–150
 zero-normalized cross-correlation (ZNCC), 152–153
- Indirect illumination, 216
- Indoor Signpost, for handheld devices, 348
- Industry, examples of use of AR, 13–16
- Inertia, sensors for physical phenomena, 90–91

- Inertial measurement unit (IMU), sensor fusion and, 118–119
- Information channels, web access, 341–342
- Information display, examples of use of AR, 20–21
- Information filtering
- combined knowledge-based and spatial filter, 267–269
 - knowledge-based filtering, 265
 - overview of, 265
 - spatial filtering, 265–267
- "Information seeking mantra," 242–243
- Infrared light
- combining magnetic and infrared sensing, 119
 - sensors for physical phenomena, 90–91
- Infrared tracking. *See* Multiple-camera infrared tracking
- Innovega, iOptik platform, 62
- Inpainting, observing and modeling hidden areas, 228
- Input modalities
- gestures, 282–283
 - overview of, 276–279
 - physically based interfaces, 286
 - touch, 283–286
 - tracking bodies, 281–282
 - tracking/manipulating rigid objects, 279–281
- Input peripherals, hardware abstraction and, 332
- Inside-out tracking, spatial sensor arrangement, 93–94
- Instant radiosity
- diffuse global illumination and, 224
 - shadow mapping and, 220
 - specular global illumination and, 225
- Interaction
- agile displays, 274–276
 - augmentation placement, 272–274
 - augmented paper, 295–297
 - challenges in situated visualization, 242–243
 - conversational agents, 306–309
 - cross-view interactions, 303–304
 - defining interactive behaviors, 334
 - gestures, 282–283
 - haptic interactions, 304
 - input modalities, 276–279
 - magic lens, 276–279
 - multi-display Focus+Context, 297–298
 - multimodal interactions, 304–306, 390–391
 - multiple locales, 300–302
 - multi-view interfaces, 297
 - output modalities, 272
 - overview of, 271
 - physically based interfaces, 286
 - shared space, 298–300
 - summary, 309–310
 - tangible interfaces, 286–287
 - tangibles on surfaces, 287
 - tangibles with distinct shapes, 289–292
 - tangibles with generic shapes, 287–289
- touch, 283–286
- tracking bodies, 281–282
- tracking/manipulating rigid objects, 279–281
- transparent tangibles, 292–294
- virtual user interfaces on real surfaces, 294–295
- Interest points
- detecting, 140–141
 - natural feature tracking, 112
 - storing geo-referenced, 341
- Interfaces. *See also* User interfaces
- authoring and, 331
 - multi-view. *See* Multi-view interfaces
 - real/virtual consistency, 381
 - smart object interfaces, 417–418
 - tangible. *See* Tangible interfaces
 - transitional interfaces for multiple views in navigation, 357–359
 - virtual interfaces, on real surfaces, 294–295
- Interference, as challenge in situated visualization, 243–244
- Internet of Things (IoT), interfacing with smart objects, 417–418
- Intersense IS-600/IS-900, ultrasonic tracking, 98
- Intrinsic haptic displays, 36
- Inverted file structure, scalable visual matching, 166
- Invisible Train, agile collaboration in shared space, 369
- IoT. *See* Internet of Things
- Irradiance, 207
- Irradiance cache, 219, 226–227
- Irradiance maps, 207
- Irradiance volumes
- shadow mapping and, 220
 - storing indirect illumination, 223–224
- Iterative closest point (ICP) algorithm, dense mapping, 162
- J**
- Java, object management, 386
- Java RMI, general use middleware for AR, 385
- JavaScript
- procedural scripting, 405
 - producing/consuming multimedia information, 341
- JESS, declarative scripting, 402
- JIMM, use in semi-automatic reconstruction, 320
- Jitter, compensating for, 192–193
- K**
- Kalman filters, compensating for jitter, 193
- Kanade-Lucas-Tomasi (KLT) tracker, incremental tracking technique, 151–152
- KARMA application
- history of augmented reality, 7
 - knowledge-based filtering, 265

- k-d trees
 descriptor matching, 145
 photon maps used with sparse k-d trees, 219
- Keyframes, natural feature tracking, 112
- Kinect (Microsoft)
 dense mapping, 162–163
 tracking human motions, 108–109
- KinectFusion system, history of augmented reality and, 13
- Kinesthetic feedback, 36
- KinÈtre, authoring by performance, 337
- KML format, Google Earth, 341
- Knowledge-based filtering
 combined knowledge-based and spatial filter, 267–269
 overview of, 265
- L**
- Labels/labeling
 anchor point selection, 249
 area representation, 249
 image-guided placement of, 252–253
 internal and external, 248
 legibility of, 253–254
 optimization techniques, 249–250
 placement objectives, 248–249
 temporal coherence, 250–252
- Landmarks, spatial knowledge for navigation, 346
- Lasers
 laser (fiber-optic) gyroscopes, 103
 laser range finder in semi-automatic reconstruction, 321
 as light source, 108
 sensors for physical phenomena, 90–91
- Latency
 distortions and aberrations in OST and VST HMDs, 55
 registration and, 192
 temporal characteristics of tracking systems, 95
- Layout, for reducing data overload, 242
- LCDs. *See* Liquid crystal displays
- LED spotlights, source of active illumination, 107
- LEGO robot, as conversational agent, 308–309
- Lens. *See also* Magic lens
 correcting lens distortion, 182–183, 231–232
 vignetting, 234
- LIDAR. *See* Light-radar
- Light factorization algorithm, 224
- Light-field displays
 near-eye displays, 64
 types of visual displays, 58–59
- Light fields
 offline light capturing, 210
 photometric registration from specular reflections, 211
- Light probes
 for acquisition of radiance map, 208–210
 reflection in human eye as, 211
- Light propagation volumes, 220
- Light-radar (LIDAR), laser sensing, 108
- Light sources
 in optical tracking, 106–109
 reconstructing explicit, 215–216
- LightSpace
 multiple depth sensors for touch interaction, 284–285
 touch interaction and, 285
 uses of cross-view interaction, 304
- Light transport, in real-time global illumination, 218–219
- Linear accelerometers, in tracking, 103–104
- Liquid crystal displays (LCDs)
 history of augmented reality, 6
 resolution and refresh rate, 49
- LiveSphere, panoramic video with, 372
- Local coordinates, versus global, 92–93
- Locales, multi-view interfaces, 300–302
- Local illumination, compared with global, 205–206
- Localization
 device capabilities and, 416–417
 outdoor tracking and, 164
 relocalization and loop closure, 160–161
 simultaneous localization and mapping. *See* Simultaneous localization and mapping success rate of, 416
- Local registration, visualization registration, 245–246
- Location, geometric methods for determining, 91
- Loops, relocalization and loop closure, 160–161
- Lukas-Kanade algorithm, 151–152
- M**
- Machine learning, FAST detection method and, 143–144
- Magic book
 multiple displays in collaboration, 367
 tangible interfaces for distinct shapes, 290
- MagicLeap, 64
- Magic lens
 focus area of, 278
 as output modality, 277–279
 spatial filtering, 265–267
 use of tangible interfaces, 292
- MagicMeeting
 use in co-located collaboration, 366
 use of tangible interfaces, 290, 292
- "Magic mirror" environment, ALIVE system providing, 307
- Magnetic flux, sensors for physical phenomena, 90–91

- Magnetic tracking
combining magnetic and infrared sensing, 119
magnetometers in, 102
types of stationary tracking systems, 97
- Magnetometers, 102
- Maintenance, examples of use of AR, 17–18
- Map-in-the-Hat, outdoor navigation system, 11
- Mano-a-Mano, use in co-located collaboration, 366
- MapLens, agile collaboration in shared space, 369–370
- Maps/mapping
exo-centric and ego-centric perspectives, 354–357
simultaneous localization and. *See* Simultaneous localization and mapping
simultaneous multiple perspectives in maps, 356–357
use for outdoor navigation, 349
uses of augmented paper, 296
- MARA, as conversational agent, 308
- Markers
authoring process defining interactive behaviors, 334
black-and-white shapes, 111
gaze objects as, 209
matching targets consisting of spherical markers, 137
retro-reflective foil, 111–112
square and circular designs, 109–110
as tangible objects, 287–289
target identification, 114–115
- Marker tracking
camera representation, 124–126
detecting, 126–128
optical tracking, 109
overview of, 123–124
pose estimation from homography, 128–131
pose refinement, 132
- Markup languages, social computing and, 421
- MARS. *See* Mobile augmented reality system
- Mathematic notation, 123
- Matrices, establishing point correspondences, 134–135
- Maya plug-in, for authoring, 340
- Measurement systems. *See also* Sensors, tracking
absolute and relative measurements, 93
accuracy of GPS systems, 100
calibration of, 86–87
coordinates, 92–93
degree of freedom, 92
errors and, 94–95
geometric distortions, 190–191
sensor arrangement, 91
sensors for physical phenomena, 90–91
spatial sensor arrangement, 93–94
temporal characteristics, 95
workspace coverage, 94
- Mechanical tracking, types of stationary systems, 96–97
- Medical, examples of use of AR, 18–20
- MEMS. *See* Micro-electromechanical system
- M-estimator, robust pose estimation, 148
Meta 2, spatial audio support, 35
- Meta 2 Development Kit
optical see-through display, 44
spatial audio support, 35
- MetaCookie, uses of olfactory and gustatory displays, 38
- Metaio Mobile SDK
platform selection, 407
plug-in approaches to authoring, 340
- Micro-electromechanical system (MEMS)
linear accelerometers, 103–104
mobile sensors for tracking, 102–103
- Microlens, in near-eye displays, 58
- Microsoft
HoloLens. *See* HoloLens
Kinect, 108–109, 162–163
- MicroVision Nomad, retinal scanning displays, 65
- MirageTable, touch interaction and, 286
- Mixed language programming, 405
- Mixed reality (MR)
continuum, 29
haptic mixed reality, 37
related fields to AR, 28–29
research facility for, 9–10
- Mixed Reality System Laboratory, 9–10
- Mobile AR, agile collaboration, 369
- Mobile augmented reality system (MARS), 336
- Mobile displays, 274
- Mobile sensors
GPS, 99–101
gyroscopes, 102–103
linear accelerometers, 103–104
magnetometers, 102
odometers, 104
in tracking, 99
wireless networks in, 101–102
- Mobility, outdoor tracking and, 164
- Model-free occlusion, 202–205
- Model transformation
coordinate system and, 88–89
spatial display model, 56
- Models/modeling
annotation, 325–327
free-form modeling, 322–324
model-based versus model-free tracking, 106
overview of, 311–312
in rendering phantoms, 200
semi-automatic reconstruction, 319–322
specifying appearance, 317–319
specifying geometry, 312
specifying planes, 314–315
specifying points, 313–314

- M**
- Models/modeling (*continued*)
 - specifying volumes, 315–317
 - summary, 328
 - Modularity, supporting collaboration and reuse, 333
 - Monocular depth cues
 - camera calibration and, 196–197
 - visual perception and, 39
 - Monocular HMDs, 42–43
 - MORGAN
 - declarative scripting, 402
 - object management, 387
 - scene graph integration, 397
 - Motion blur effects, 232–233
 - Mounting options, visual displays, 60
 - Movies, AR as dramatic medium, 420
 - MR. *See* Mixed reality
 - Multi-display Focus+Context, 297
 - MultiFi, variant of virtual ether, 299
 - Multifocal displays, 46–47
 - Multimedia, creating multimedia assets in authoring workflow, 332
 - Multimodal displays
 - audio, 34–35
 - haptic, tactile, and tangible, 35–37
 - olfactory and gustatory, 37–39
 - overview of, 34
 - Multimodal interactions
 - dataflow and, 390–391
 - overview of, 304–306
 - Multiple-camera infrared tracking
 - absolute orientation, 137–138
 - blob detection, 133
 - establishing point correspondences, 133–135
 - matching targets consisting of spherical markers, 137
 - overview of, 132–133
 - triangulation from more than two cameras, 136–137
 - triangulation from two cameras, 135–136
 - Multiple cameras, 412–413
 - Multiple-locale interface, 300–302
 - Multi-touch detection, 283–284
 - Multi-view interfaces
 - cross-view interactions, 303–304
 - interaction, 297–298
 - multi-display Focus+Context, 297
 - multiple locales, 300–302
 - shared space, 298–300
- N**
- National Tele-Immersion Initiative, 374
 - Natural features
 - target identification, 115–116
 - in tracking, 109, 112–113
 - Natural features, tracking by detection
 - descriptor creation, 144–145
 - descriptor matching, 145–146
 - determining corners using Harris detector, 141–142
 - DOG detection method, 142
 - FAST detection method, 143–144
 - interest point detection, 140–141
 - overview of, 138–140
 - PnP problem, 146–147
 - robust pose estimation, 148–149
- N**
- Navigation
 - examples of use of AR, 21–22
 - exploration and discovery, 347
 - foundations of human navigation, 346–347
 - guidance toward target object, 350–354
 - guidance toward target viewpoint, 354
 - multiple perspectives, 354–355
 - overview of, 345
 - route visualization, 347–350
 - simultaneous multiple perspectives, 355–357
 - summary, 360
 - transitional interfaces for multiple views, 357–359
 - viewpoint guidance, 350
 - Near-eye displays
 - microlens use in, 58
 - OST. *See* Optical see-through
 - overview of, 59–61
 - VST. *See* Video see-through
 - Networked RTK (NRTK), 101
 - Nister's five-point algorithm, 157–158
 - Nodes, dataflow
 - dataflow graphs, 390
 - scene graphs, 395–397
 - threads and scheduling and, 391
 - Nodes, navigation
 - spatial knowledge for, 346
 - use in outdoor and indoor navigation, 349–350
 - Noise effects, matching in camera simulation, 234
 - Nonlinear optimization, pose refinement and, 132
 - Non-photorealistic rendering (NPR), 236
 - NPR. *See* Non-photorealistic rendering
 - NRTK. *See* Networked RTK

- tracking/manipulating rigid objects, 279–281
virtual representation of real-world objects, 260–262
- Obliq, declarative scripting, 402
- Occlusion
depth cues and, 197
model-free, 202–205
overview of, 199
phantom use for, 199–200
probabilistic, 202
refinement, 201–202
removing via X-ray visualization, 254
requirements for visual displays, 47–48
- Ocularity, requirements for visual displays, 42–45
- Oculus Rift
binocular HMDs, 43
field of view, 51, 67–68
high-resolution displays, 68
use of HMDs in gaming devices, 29
- Odometer, mobile sensors for tracking, 104
- Odometry, visual. *See* Simultaneous localization and mapping
- Office of the Future, telepresence system, 11
- Offline light, capturing with environment map, 210
- Offset, in specifying planes, 315
- Olfactory displays, 37–39
- OmniTouch, use of virtual interfaces, 294–295
- OpenSceneGraph, scene graph integration, 399
- OpenSG, scene graph integration, 399
- OpenStreetMap, 169
- OpenTracker
declarative scripting, 402
as device abstraction software framework, 335
multimodal interactions, 391
object management, 386
scene graph integration, 399
wearable AR setup as case study, 393–395
- “Optical camouflage,” projector-based diminished reality and, 230
- Optical see-through (OST)
brightness and contrast, 53–54
comparing resolution with VST, 49
conversational agents, 308
developer wish list, 414–415
distortions and aberrations, 55
HMDs, 61–66
methods of augmentation, 40–41
stereo capabilities, 43–44
user-tracking systems with, 184
- Optical tracking
illumination in optical tracking, 106–109
markers in, 109–112
model-based versus model-free, 106
natural features in, 112–113
overview of, 105–106
target identification, 113–116
- Optics. *See* Computer vision
- Optimization techniques, for labels, 249–250
- OptiX ray tracer, 226
- Outdoor AR, future of, 415–417
- Outdoor photometric registration, 214–215
- Outdoor tracking
overview of, 164–165
prior information from geometry, 169–170
prior information from sensors, 167–169
scalable visual matching, 165–167
simultaneous tracking, mapping, and localization, 170–176
- OutlinAR, use in semi-automatic reconstruction, 320
- Output modalities
agile displays, 274–276
augmentation placement, 272–274
magic lens, 276–279
overview of, 272
- Output peripherals, hardware abstraction and, 332
- Outside-in tracking, spatial sensor arrangement, 93–94
- P**
- P3P algorithm, applying to PnP problem, 146–147
- Paddles, use of tangible interfaces, 290
- Panoramas, extending video sharing, 372
- Paper, augmented, 295–297
- PaperWindows, 296–297
- Parallax-free
viewpoint offset and, 52–53
VST displays, 67
- Parallel programming, 413–414
- Parallel tracking and mapping (PTAM) system
history of augmented reality and, 13
overview of, 159–160
- Parameter configuration, developer support, 401
- Parking assistant, AR applications, 23
- Particle filters, compensating for jitter, 193
- Passive illumination, in optical tracking, 106–107
- Passive light probes, 208
- Passive sources, signal sources, 92
- Pedometers, linear accelerometer use by, 104
- Performance, authoring by, 337–339
- Personal information display, examples of use of AR, 20–21
- Personal interaction panel, use of tangible interfaces, 290–291
- Perspective-n-Point (PnP) problem, 146–147
- Perspectives, navigation
multiple perspectives, 354–355
simultaneous multiple perspectives, 355–357
transitional interfaces for multiple views, 357–359
- Phantom Omni (Geomagic Touch) haptic device, 36
- Phantoms
dual phantom rendering algorithm, 262
model-free occlusion, 202
occlusion refinement, 201–202

- Phantoms (continued)**
- phantom geometry for physical coherence, 286
 - probabilistic occlusion, 202
 - use with occlusion, 199–200
 - virtual representation of real-world objects, 260–262
- Photometric registration**
- compositing and, 198–199
 - from diffuse reflections, 212–214
 - global illumination used for, 225
 - image-based lighting, 207–208
 - light probe use, 208–210
 - offline light capturing, 210
 - outdoors, 214–215
 - overview of, 205–207
 - reconstructing explicit light sources, 215–216
 - from shadows, 214
 - from specular reflections, 211–212
 - from static images, 210–211
- Photon maps**
- combining with ray tracing, 226
 - use in sparse k-d trees, 219
- Physically based interfaces**, as input modality, 286
- Physical phenomena**, sensors for, 90–91
- Physical world**
- comparing AR and VR, 382
 - real/virtual consistency, 381
- Pick-and-drop**, cross-view interaction, 303
- Pictofit** virtual dressing room, in fashion industry, 25–27
- Pinch gestures**, manipulating rigid objects, 280–281
- Pinlight**, OST platform, 62–63
- Pipes-and-filters architecture**, 389
- Pixels**
- accuracy of registration, 51
 - noise and, 234
- Pixel shaders**, rendering blur effects, 233
- PixMix** method, for removing region of interest, 229
- Planar system**, 14
- Planes**
- estimating dominant planes in reconstruction, 319–320
 - specifying geometry, 314–315
 - use in specifying volumes, 316–317
- Platforms.** *See* **Frameworks**
- Plug-in approaches**, to authoring, 339–340
- PnP.** *See* **Perspective-n-Point problem**
- Point correspondences**, establishing in multiple-camera infrared tracking, 133–135
- Pointers**, defining interactive behaviors, 334
- PointGrey Ladybug camera**, 119
- Pointing**, in remote collaboration, 375–376
- Pointing device**, calibration of HMD, 186–187
- Points**
- specifying geometry, 313–314
 - use in specifying planes, 314
 - use in specifying volumes, 315
- Polarization-based stereo**, 58, 367
- Polygonal models**, in rendering phantoms, 200
- Polygons**, specifying volumes and, 315
- Portability**, platform abstraction and, 383
- Portal displays**, 75–77
- Pose**
- estimation from homography, 128–131
 - pose sensors, future, 413
 - refinement, 132
 - robust pose estimation, 148–149
- Pose graph optimization**, 159
- PowerSpace**, for desktop authoring, 335
- Pre-Viz AR application**, film industry, 24
- Precision**, of measurements, 95
- Prediction**, compensating for jitter, 193–194
- Printed media** (magazines/flyers), augmentation, 25
- Prior information**
- from geometry, 169–170
 - natural feature tracking and, 139
 - from sensors, 167–169
- Privacy**
- in collaborative environment, 367
 - social acceptance of HMDs and, 56
- Privacy lamp**, 367
- Probabilistic occlusion**, 202
- Procedural scripting**, 405
- ProFORMA**, use in semi-automatic reconstruction, 320
- Programming languages**
- declarative scripting, 401–403
 - mixed-language programming, 405
 - object management, 386
 - procedural scripting, 405
 - versus content design frameworks, 330
- Projected displays**
- dynamic shader lamps, 82
 - everywhere projectors, 83
 - head-mounted projector displays (HMPD), 80–81
 - overview of, 78
 - spatial AR projectors, 78–80
- Projective transformation**
- coordinate system and, 88–89
 - spatial display model, 56–57
- Projector-based diminished reality**, 230–231
- Projector-camera systems**
- agile displays, 274–275
 - touch interaction and, 284
- Projectors**
- agile projectors, 275
 - IllumiRoom* projector in AR games, 28
 - Touch Projector, 304
 - transparent tangibles and, 293
- Properties**, of collaboration systems, 362–363
- Proprioception**, 273
- PTAM.** *See* **Parallel tracking and mapping system**
- PTI Vuforia**, 27, 340, 407

Pull strategy, for thread management, 392–393
Pupil
 focus and, 45
 response to light, 39
Push strategy, for thread management, 392

Q

Qualcomm Vuforia. *See also* PTI Vuforia, 340

R

Radial distortion, correcting, 182–183
Radiance
 defined, 207
 storing on surface points or surface patches, 219
Radiance cache, 219
Radiance maps
 for image-based lighting, 219
 light probes for acquiring, 208–210
 using environment maps for image-based lighting, 207
Radio signals, sensors for physical phenomena, 90–91
Radiosity methods
 diffuse global illumination and, 223–224
 storing radiance, 219–220
Random measurement error, 94–95
Random sampling consensus (RANSAC)
 robust pose estimation, 148
 scalable visual matching, 166
Range finders, 321
Range or workspace coverage of sensor, tracking, 94
Raycasting
 performing operations at a distance, 313
 selection of object in 3D environment, 278–279
Ray direction, in raycasting, 313
Rays, triangulation from two cameras, 135–136
Ray-tracing
 rendering depth-of-field effect, 233
 specular global illumination and, 225–227
Razer Hydra magnetic tracker, 97
Real entities, linking to virtual entities, 325, 332
Real-time global illumination, 218–220
Real-time kinematics (RTK) GPS, 100–101
Real world
 as part of AR definition, 1–4, 29
 real/virtual consistency, 381
 real-world interfaces for authoring, 331
Red, green, blue. *See* RGB
Reflections
 local illumination compared with global illumination, 206
 photometric registration from diffuse reflections, 212–214

photometric registration from specular reflections, 211–212
reflection in human eye as light probe, 211
Refresh rate, requirements for visual displays, 48–49
Region of interest (ROI)
 determining when working with diminished reality, 228
problem areas addressed in diminished reality, 227
removing, 229–230
Registration. *See also* Visualization registration
alignment of spatial properties, 86
alignment of virtual and real objects, 87
error propagation and, 191–192
errors as challenge in situated visualization, 243
essential components in definition of AR, 3
filtering and prediction, 192–194
geometric measurement distortions and, 190–191
latency and, 192
overview of, 179–180, 190
stages of VST pipeline, 197–198
summary, 194
in visual coherence, 196–198
Relative measurements, versus absolute, 93
Relighting, differential rendering and, 217
Remote collaboration
 with agile users, 376–377
 overview of, 370–371
 pointing and gestures, 375–376
 video sharing, 371–372
 video sharing with geometric reconstruction, 374–375
 video sharing with virtual objects, 372–373
Rendering
 blur effects, 233
 buffer rendering, 258–260
 choosing rendering method, 219
 dealing with depth cues, 197
 differential rendering for real-world illumination, 216–217
 non-photorealistic rendering (NPR), 236
 shadow mapping in, 222
Reprojection error, pose refinement and, 132
Research framework, collaborative AR system as, 330
Resolution
 field of view and, 50–51
 requirements for visual displays, 48–49
 of sensors, 95
Retinal scanning displays, 65
Retro-reflective material, 80–81
Reuse, as engineering requirement, 383–384
RGB (red, green, blue)
 converting to YUV, 235
 image-guided placement of labels, 252
 tone mapping artifacts, 235

- RGB-D camera, dynamic lighting from, 225
 Rigid objects, tracking/manipulating, 279–281
 Rockwell Collins SimEye, 60
 ROI. *See* Region of interest
 Room in Miniature, simultaneous multiple perspectives in maps, 356
 Rotary plate, use of tangible interfaces, 290, 292
 Rotational velocity, measuring with gyroscope, 102
 Routes, spatial knowledge for navigation, 346
 Route visualization, in navigation, 347–350
 RQ factorization, use in camera calibration, 181
 RTK GPS. *See* Real-time kinematics GPS
 Runtime engine, in authoring workflow, 332
 Runtime reconfiguration, developer support, 405–406
- S**
- Salience, ghostings from image space and, 257–258
 Samsung
 GearVR, 68
 Smart Window, 75–77
 Satellites, in GPS system, 99
 Scalable visual matching, outdoor tracking, 165–167
 Scale invariant feature transform (SIFT), descriptor creation, 144
 Scale space search, 142
 Scene graphs
 declarative scripting, 402
 dependency graphs, 397
 distributed shared scene graphs, 399–400
 function of, 380
 high-level programming frameworks, 330
 integration of, 397–399
 overview of, 395–397
 Scenes
 compositing, 259–260
 confluence of VR and AR, 418
 in real-time global illumination, 218
 in stories, 334
 understanding based on geometric information, 322
 Scheduling dataflow, 391–393
 Screen-stabilized frames of reference, 89
 Scripting languages
 declarative scripting, 401–403
 function of, 380
 procedural scripting, 405
 use with advanced interactions, 334
 Search
 active search tracking technique, 150–151
 descriptor matching, 145
 zero-normalized cross-correlation (ZNCC), 152–153
 Search trees, 166
- See-through displays
 optical. *See* Optical see-through video.
 See Video see-through
 Semantic knowledge, real/virtual consistency, 381–382
 Semi-automatic reconstruction, 319–322
 Semi-dense mapping, 164
 SensaBubble, uses of olfactory and gustatory displays, 37
 SenseShapes, multimodal interactions, 305
 Sensorama simulator, uses of olfactory and gustatory displays, 37
 Sensor fusion
 competitive, 117–118
 complementary, 117
 cooperative, 118–120
 overview of, 117
 synchronization, 91
 Sensors
 augmented humans, 419
 future of, 413
 prior information from, 167–169
 synchronization, 91
 use of depth sensors in semi-automatic reconstruction, 321
 use with hand gestures in remote collaboration, 376
 Sensors, tracking
 arrangement of, 91
 combining magnetic and infrared sensing, 119
 competitive sensor fusion, 117–118
 complementary sensor fusion, 117
 cooperative sensor fusion, 118–120
 filtering compensating for jitter, 192–193
 geometric measurement distortions, 190–191
 Global Positioning System (GPS), 99–101
 gyroscopes for, 102–103
 linear accelerometer in, 103–104
 magnetometers for, 102
 measurement errors and, 94–95
 mobile sensors, 99
 multiple depth sensors for touch interaction, 284–285
 odometer in, 104
 optical tracking and, 105–106
 for physical phenomena, 90–91
 range or workspace coverage of, 94
 resolution of, 95
 spatial sensor arrangement, 93–94
 wireless networks in, 101–102
 Service localization protocol (SLP), 386
 Setup, of authoring hardware, 335
 SFM. *See* Structure from motion
 SH. *See* Spherical harmonics
 Shading, depth cues and, 197

- Shadows**
 common illumination and, 216
 computation of, 220–223
 depth cues and, 197
 instant radiosity for mapping, 220
 passes of shadow volume technique, 220–221
 photometric registration from, 214
 shadow mapping, 222
- Shape hints**, 256
- Shapes**
 free-form modeling, 322
 tangible interfaces for distinct, 289–292
 tangible interfaces for generic, 287–289
- Shared space**
 co-located collaboration, 364
 multi-view interfaces, 298–300
 properties of collaboration systems, 362
- Sharing.** *See also* Collaboration
 agility in shared space, 368–369
 annotations supporting, 325–327
 distributed shared scene graphs, 399–400
 video, 371–372
 video with geometric reconstruction, 374–375
 video with virtual objects, 372–373
- SHEEP game**, 387–389
- “Shooting gallery” approach, to display calibration, 183–184
- Shutter glasses**
 for stereoscopic effect, 74–75
 view-dependent spatial AR, 79–80
- SIFT.** *See* Scale invariant feature transform
- Signals**
 measuring, 91
 sources, 92
- Signpost, route visualization and**, 348
- Simulation, engineering requirement**, 384
- Simultaneous localization and mapping (SLAM)**
 building system of outdoor annotations, 327
 bundle adjustment for drift correction, 158–159
 combining 3D tracking with 3D scanning, 106
 dense mapping, 161–164
 live assembly method of scene, 46
 Nister’s five-point algorithm, 157–158
 outdoor tracking, 170–176
 overview of, 156–157
 parallel tracking and mapping, 159–160
 relocalization and loop closure, 160–161
 semi-dense mapping, 164
 sparse map used for registration of
 annotations, 325
 use in semi-automatic reconstruction,
 319–320, 322
 video sharing with geometric reconstruction,
 374–375
- Single point active alignment method (SPAAM)**
 calibration of display, 185–186
 calibration of HMD using pointing device,
 186–187
- Sink nodes, dataflow graphs**, 390
- Situated computing**
 location-based services and augmented
 reality, 2
 smartphones and tablets and, 72
- Situated Documentaries**, 336
- Situated modeling**
 reconstruction method providing data
 for, 319
 specifying geometry, 312
- Situated visualization**
 annotations, 248
 challenges in, 241–242
 combined knowledge-based and spatial filter,
 267–269
 data overload and, 242
 explosion diagrams, 260–262
 ghostings from image space, 256–258
 ghostings from object space, 255–256
 globally registered, 246–247
 image-guided placement, 252–253
 implementing ghostings with G-buffers,
 258–260
 information filtering, 265
 knowledge-based filtering, 265
 labeling, 248–249
 legibility, 253–254
 locally registered, 245–246
 optimization techniques, 249–250
 overview of, 239–241
 registration errors, 243
 space distortion, 262–265
 spatial filtering, 265–267
 spatial manipulation, 260
 summary, 267–269
 temporal coherence, 244, 250–252
 uncertainty of registration, 247–248
 user interaction, 242–243
 visual interference, 243–244
 visualization registration, 245
 X-ray visualization, 254–255
- SixthSense**, use of virtual interfaces, 294
- Skeleton tracking**
 body tracking techniques, 281
 in KinÈtre, 337
- SLAM.** *See* Simultaneous localization and
 mapping
- SLP.** *See* Service localization protocol
- Smart object interfaces**, 417–418
- Smart Window (Samsung)**, 75–77
- Smartphones**
 AR browser apps on, 20–21
 example applications for AR, 12–13
 handheld displays and, 69
 mobile sensors, 99
 situated (context-sensitive) computing, 72
- Smelling Screen**, uses of olfactory and gustatory
 displays, 37

- Social acceptance, requirements for visual displays, 55–56
- Social computing, 421–422
- Software architectures
- application requirements, 380–382
 - AR tour guide as case study, 403–404
 - dataflow, 389–390
 - declarative scripting, 401–403
 - dependency graphs, 397
 - developer support, 400–401
 - distributed object systems, 385–386
 - distributed shared scene graphs, 399–400
 - engineering requirements, 382–384
 - integration of scene graphs, 397–399
 - mixed-language programming, 405
 - multimodal interactions, 390–391
 - object management, 386–387
 - overview of, 379
 - parameter configuration, 401
 - platform selection, 407
 - procedural scripting, 405
 - runtime reconfiguration, 405–406
 - scene graphs, 395–397
 - SHEEP game case study, 387–389
 - summary, 407
 - threads and scheduling, 391–393
 - wearable AR setup as case study, 393–395
- SONY Glasstron OST display, 61–62
- Sound sensors, for physical phenomena, 90–91
- Source nodes, dataflow graphs, 390
- SPAAM. *See* Single point active alignment method
- Space-carving
- creating volumes, 316
 - use in semi-automatic reconstruction, 320
- Space distortion
- overview of, 262
 - variable perspective view, 263–265
- Sparse maps, use with annotation, 325
- Sparse matching, natural feature tracking by detection, 139
- Spatial AR
- AR displays and, 42
 - dynamic shader lamps and, 82
 - everywhere projectors, 83
 - multifocal displays, 46
 - projected displays, 78–79
 - view-dependent, 79–80
- Spatial audio, 35
- Spatial filtering
- combined knowledge-based and spatial filter, 267–269
 - overview of, 265–267
- Spatial knowledge, for navigation, 346
- Spatial location. *See* Tracking
- Spatial manipulation
- explosion diagrams, 260–262
 - interfacing with smart objects, 417–418
- overview of, 260
- space distortion, 262–265
- Spatial models
- essential in definition of AR, 3–4
 - spatial display model, 56–57
- Spatial resolution, video displays and, 49
- Spatial sensor arrangement, 93–94
- Specular global illumination, 225–227
- Specular reflections
- light transport and, 218
 - local illumination compared with global illumination, 206
 - photometric registration from, 211–212
 - recording, 209
- Spherical harmonics (SH)
- indexing radiance cache, 220
 - storing directional illumination, 212–213
- Spill forest, descriptor matching, 145
- Sports, augmented TV broadcasts, 24
- Spotlights, source of active illumination, 107
- SpotScents, uses of olfactory and gustatory displays, 37
- Static error, registration and, 190
- Static images, photometric registration from, 210–211
- Static registration, 87
- Stationary displays, 72
- Stationary tracking systems
- electromagnetic tracking, 97
 - mechanical tracking, 96–97
 - overview of, 96
 - ultrasonic tracking, 98
- Stencil buffer, shadow volume technique based on, 220–221
- Stereoscopic displays, 58
- Stereoscopy, requirements for visual displays, 42–45
- Stories
- AR as dramatic medium, 420–421
 - chronology in authoring, 334
- Structural modeling system, 322
- Structure from motion (SFM)
- scalable visual matching, 166–167
 - simultaneous localization and mapping, 157
- Structured light, in optical tracking, 107–109
- Studierstube collaborative AR system
- APRIL use with, 336–337
 - declarative scripting, 402
 - history of augmented reality, 9
 - as multiple-locale interface, 300
 - as research framework, 330
- Stylized AR
- applying stylization filters to composites, 199
 - overview of, 236
- Stylus-and-tablet interface, 293–294
- Sun, source of passive illumination, 106
- SURF, descriptor creation, 144

- Surfaces**
- augmented, 298
 - tangible interfaces for, 287
 - virtual user interfaces on real surfaces, 294–295
- Surface textures**, specifying appearance, 317
- Surveillance**, social acceptance of HMDs and, 56
- Surveys**, spatial knowledge for navigation, 346
- Sword of Damocles**
- head-mounted display (HMD), 5
 - near-eye displays, 59–60
- Synchronicity**, properties of collaboration systems, 362
- Synchronized dual phantom rendering algorithm**, 262
- Systematic measurement error**, 94–95
- T**
- Tablet computers**
- handheld displays and, 69
 - mobile sensors, 99
 - situated (context-sensitive) computing, 72
 - uses of multi-view interfaces, 297–298
 - in wearable AR setup as case study, 393
- Tactile displays**, 35–37
- Tactile feedback**, displays and, 36
- Tangible AR**, 273, 287
- Tangible displays**, 35–37
- Tangible interfaces**
- augmentation and, 273
 - with distinct shapes, 289–292
 - with generic shapes, 287–289
 - interfacing with smart objects, 417
 - overview of, 286–287
 - on surfaces, 287
 - transparent tangibles, 292–294
- Tangible replicas**, sharing video with virtual objects, 373
- Target identification**
- feature target identification, 115–116
 - marker target identification, 114–115
 - overview of, 113
- Targets**
- guidance toward target object, 350–354
 - guidance toward target viewpoint, 354
 - matching targets consisting of spherical markers, 137
- Task space**, versus communication space in collaboration, 362–363
- Tcl libraries**, 405
- Tele-collaboration**, 371
- Telepresence**
- collaborative AR interface and, 18–19
 - experiencing live remote location, 362
 - in Office of the Future, 11
 - video sharing with geometric reconstruction, 375
- Television (TV)**
- AR as dramatic medium, 420–421
 - examples of use of AR, 22–25
 - resolution and refresh rate of TV cameras, 49
- Templates**, template-based approach to authoring, 336
- Temporal characteristics**
- of collaboration, 362
 - tracking, 95
- Temporal coherence**
- challenges in situated visualization, 244
 - of labels, 250–252
- Temporal resolution**, video displays and, 49
- Texture maps**, compensating for lens distortion, 232
- Textures**, specifying surface appearance, 317
- THAW**, uses of multi-view interfaces, 297
- Thermal feedback**, in displays, 36
- Threads**
- dataflow, 391–393
 - distributed object systems and, 385
- Time-multiplexed stereo**, 49, 58, 367
- Tinmith**
- declarative scripting, 402
 - outdoor AR navigation system, 11
 - route visualization and, 347–348
 - runtime reconfiguration, 405–406
 - scene graph integration, 399
- Tone mapping artifacts**, camera simulation and, 235
- Torso-referenced display**, augmentation and, 273
- Total Immersion D'Fusion**
- platform selection, 407
 - plug-in approaches to authoring, 340
- Touch**
- as input modality, 283–286
 - virtual touch widgets, 294
- Touch Projector**, uses of cross-view interaction, 304
- Touchscreens**, 284
- Tour Guide** case study, 403–404
- Touring Machine**, history of augmented reality, 9–11
- TOWNWEAR system**, laser (fiber-optic) gyroscopes, 103
- Tracking**
- absolute and relative measurements, 93
 - bodies, 281–282
 - comparing with registration and calibration, 86–87
 - coordinate systems, 87–88, 92–93
 - degree of freedom (DOF) and, 92
 - by detection. *See* Natural features, tracking by detection
 - electromagnetic tracking, 97
 - essential component in definition of AR, 3
 - frames of reference, 89–90
 - gestures, 281–283

- T**
- Tracking (*continued*)
 - Global Positioning System (GPS), 99–101
 - gyroscopes for, 102–103
 - handheld displays and, 69
 - illumination in optical tracking, 106–109
 - incremental. *See* Incremental tracking
 - infrared. *See* Multiple-camera infrared tracking
 - interfacing with smart objects, 418
 - linear accelerometer in, 103–104
 - magnetometers for, 102
 - marker tracking. *See* Marker tracking
 - measurement and, 91
 - measurement errors and, 94–95
 - mechanical tracking, 96–97
 - mobile sensors, 99
 - model-based versus model-free, 106
 - model, view, and projective transformation and, 88–89
 - odometer in, 104
 - optical tracking, 105–106
 - outdoor tracking. *See* Outdoor tracking
 - overview of, 85
 - range or workspace coverage of sensor, 94
 - rigid objects, 279–281
 - sensor arrangement, 91
 - sensor fusion, 117–120
 - sensors for physical phenomena, 90–91
 - signal sources, 92
 - simultaneous localization and mapping. *See* Simultaneous localization and mapping
 - spatial display model, 57
 - spatial sensor arrangement, 93–94
 - stationary systems, 96
 - summary, 120
 - by synthesis, 151
 - target identification, 113–116
 - temporal characteristics, 95
 - ultrasonic tracking, 98
 - wireless networks in, 101–102
 - Training, examples of use of AR, 17–18
 - Transitional interfaces, for multiple views in navigation, 357–359
 - Transparent tangibles, 292–294
 - Travel, navigation for, 346
 - Triangulation
 - geometric methods for determining location, 91
 - from more than two cameras, 136–137
 - from two cameras, 135–136
 - Trilateration, geometric methods for determining location, 91
 - Truncated signed distance function, dense mapping, 163–164
 - Tsai algorithm, for camera calibration, 180–181
 - Tsai-Lenz algorithm, applying to hand-eye calibration, 189–190
 - Tukey estimator, robust pose estimation, 148–149
 - Tunnels, use in navigation guidance, 351–353
 - Turing Test, 418
 - TV. *See* Television
 - Twitter, AR use in social computing, 421
- U**
- Ubiquitous computing (ubicomp)
 - future of AR and, 410
 - related fields to AR, 30–31
 - UI. *See* User interface
 - Ultrasonic tracking, 98
 - UML charts, in desktop authoring, 336–337
 - Unit, declarative scripting, 402
 - Unity3D plug-in, for authoring, 340
 - Unscented transform, in error correction, 193
 - Update rate
 - choosing, 219
 - temporal characteristics of tracking systems, 95
 - User interaction. *See* Interaction
 - User interface (UI)
 - multi-view. *See* Multi-view interfaces
 - physically based, 286
 - real/virtual consistency, 381
 - tangible. *See* Tangible interfaces
 - UI abstraction, 383
 - virtual interfaces, on real surfaces, 294–295
 - Users
 - actors interacting with, 333–334
 - consumers as business case driver, 411
 - professional users as business case driver, 410–411
 - remote collaboration with agile users, 376–377
 - taking AR outdoors, 415–417
- V**
- Vampire mirror, for privacy in collaborative environment, 367–368
 - Variable perspective view, space distortion and, 263–265
 - Vergence
 - focus and, 45–46, 64, 69
 - in multifocal display, 46
 - Vidente, outdoor visualization system, 15
 - Video
 - hand video in remote collaboration, 376
 - sharing, 371–372
 - sharing with geometric reconstruction, 374–375
 - sharing with virtual objects, 372–373
 - Video conferencing, remote AR compared with, 371–372
 - Videoplace Responsive Environment, remote collaboration, 370–371

- Video see-through (VST)
 brightness and contrast, 53–54
 camera optics, 46
 comparing resolution with OST, 49
 compensating for lens distortion, 231–232
 distortions and aberrations, 55
 gaze awareness in collaboration, 368
 HMDs, 66–69
 methods of augmentation, 41–42
 pipeline stages, 197–198
 stereo capabilities, 43–45
Video-textured phantoms, virtual representation of real-world objects, 260–262
View-dependent spatial AR, 79–80
Viewpoint guidance, navigation
 overview of, 350
 toward target object, 350–354
 toward target viewpoint, 354
Viewpoint offset, requirements for visual displays, 51–53
Viewpoints, 354
Views
 co-located collaboration, 366–368
 virtual views for scalable visual matching, 167
View transformation
 coordinate system and, 88–89
 dealing with data overload, 242
 spatial display model, 56
Vignetting, camera simulation and, 234
Virtual entities, linking to real, 332
Virtual ether, shared space and, 298–299
Virtual interfaces, on real surfaces, 294–295
Virtual light points (VPL), diffuse global illumination and, 224
Virtual mirror
 overview of, 72–74
 virtual showcase compared with, 74–75
Virtual objects, sharing video with, 372–373
Virtual-physical tool, use of virtual interfaces, 294
Virtual reality (VR)
 comparing with AR, 1–3
 related fields to AR, 29
 VR/AR integration, 418
Virtual showcase
 compared with virtual mirror, 74–76
 user-tracking systems with, 184
Virtual studios, segmentation of foreground and background, 204–205
Virtual touchscreens, 284
Virtual views, scalable visual matching, 167
Virtual world, real/virtual consistency, 381
Visual coherence. *See also* Computer vision
 Bayer pattern artifacts, 235
 blur effects, 232–233
 camera simulation, 231
 chromatic aberrations, 234–235
 common illumination, 216
 compensating for lens distortion, 231–232
 determining region of interest, 228
 differential rendering for real-world illumination, 216–217
 diffuse global illumination, 223–225
 diminished reality, 227
 image-based lighting, 207–208
 light probe use, 208–210
 matching noise effects, 234
 model-free occlusion, 202–205
 observing and modeling hidden areas, 28–229
 occlusion, 199
 occlusion refinement, 201–202
 offline light capturing, 210
 overview of, 195
 phantom use for occlusion, 199–200
 photometric registration, 198–199, 205–207
 photometric registration from diffuse reflections, 212–214
 photometric registration from shadows, 214
 photometric registration from specular reflections, 211–212
 photometric registration from static images, 210–211
 photometric registration outdoors, 214–215
 probabilistic occlusion, 202
 projector-based diminished reality, 230–231
 real-time global illumination, 218–220
 reconstructing explicit light sources, 215–216
 registration in, 196–198
 removing region of interest, 229–230
 shadows and, 220–223
 specular global illumination, 225–227
 stylized augmented reality, 236
 summary, 237
 tone mapping artifacts, 235
 vignetting, 234
Visual displays. *See also* Computer vision
 handheld, 69–72
 mounting options, 60
 near-eye, 59–61
 OST HMDs, 61–66
 overview of, 58–59
 requirements for, 40
 stationary and desktop, 72
 virtual mirror, 72–74
 virtual showcase, 74–76
 VST HMDs, 66–69
 window and portal, 75–77
Visual Interaction for Archeology (VITA), 367
Visual interference, challenges in situated visualization, 243–244
Visualization, *See also* Situated visualization, 3
Visualization registration
 globally registered, 246–247
 locally registered, 245–246

- V**
- Visualization registration (*continued*)
 - overview of, 245
 - uncertainty of registration, 247–248
 - Visual mapping, dealing with data overload, 242
 - Visual odometry. *See* Simultaneous localization and mapping (SLAM)
 - Visual perception
 - displays and, 39
 - requirements for visual displays, 40
 - Visual Turing Test, 419
 - VITA. *See* Visual Interaction for Archeology
 - VjControl, runtime reconfiguration, 405–406
 - Vocabulary trees, 166
 - Volumes
 - free-form modeling, 323
 - specifying geometry, 315–317
 - Volumetric displays, 58
 - VPL. *See* Virtual light points
 - VR Juggler, runtime reconfiguration, 405–406
 - VST. *See* Video see-through
 - Vuforia
 - platform selection, 407
 - plug-in approaches to authoring, 340
 - SmartTerrain for 3D use in AR games, 27
- W**
- Wacom tablet, in wearable AR setup as case study, 393
 - Wayfinding. *See also* Navigation
 - aspects of navigation, 346
 - route visualization and, 347–350
 - Waypoints, guidance toward target object, 350–354
 - Wearable technologies
 - augmented humans, 419–420
 - near-eye displays, 64
 - uses of haptic displays, 36
 - wearable AR setup as case study, 393–395
 - Web browsers. *See* Browsers
 - WebGL, producing/consuming multimedia information, 341
 - WebKit
 - Argon browser built on, 342
 - procedural scripting, 405
 - Web technologies, 341–342
 - Welbo animated agent, uses of conversational agents, 307–308
 - WIDE5, 68
- WiFi**, mobile sensors for tracking, 101–102
- WIM. *See* World-in-miniature**
- WIMP** (windows, icons, menus, pointers), 382–383
- Window displays**, 75–77
- Windowed bundle adjustment**, for drift correction, 159
- Windows**, icons, menus, pointers (WIMP), 382–383
- Wireless networks**, mobile sensors for tracking, 101–102
- Workflow**, for authoring, 332–333
- Worldboard**, first concept for AR browser, 12
- World-in-miniature (WIM)**
 - as multiple-locale interface, 302
 - simultaneous multiple perspectives in maps, 356–357
 - transitional interfaces for multiple views, 358
- World-stabilized**, frames of reference, 89
- X**
- X-rays**, medical uses of AR, 18–20
 - X-ray visualization**
 - ghostings from image space, 256–258
 - ghostings from object space, 255–256
 - implementing ghostings with G-buffers, 258–260
 - overview of, 254–255
 - simultaneous multiple perspectives in maps, 356
 - X3D scene graph**, 402
 - XML. *See* eXtensible Markup Language**
- Y**
- YCbCr color space**, 252
 - YUV format (luminance and chrominance)**, 235
- Z**
- Z-buffers**
 - ghostings from object space and, 255
 - occlusion, 47
 - in rendering phantoms, 199–200
 - segmentation of foreground and background, 203–205
 - Zero-normalized cross-correlation (ZNCC)**, incremental tracking technique, 152–153