Augmented Reality

Ashutosh Maleti, Kalpit Bhanawat and Naveen Kumar Kedia

Department of IT, JECRC, Jaipur

1. Introduction

The field of Augmented Reality (AR) has existed for just over one decade, but the growth and progress in the past few years has been remarkable. In 1997, the first author published a survey (based on a 1995 SIGGRAPH course lecture) that defined the field, described many problems, and summarized the developments up to that point. Since then, the field has grown rapidly. In the late 1990s, several conferences specializing in this area were started, including the International Workshop and Symposium on Augmented Reality, the International Symposium on Mixed Reality, and the Designing Augmented Reality Environments workshop. Some well-funded interdisciplinary consortia were formed that focused on AR, notably the Mixed Reality Systems Laboratory in Japan and Project ARVIKA in Germany. A freely-available software toolkit (the ARToolkit) for rapidly building AR applications is now available. Because of this wealth of new developments, an updated survey is needed to guide and encourage further research in this exciting area.

The goal of this new survey is to cover the recent advances in Augmented Reality that are not covered by the original survey. This survey will not attempt to reference every new paper that has appeared since the original survey; there are far too many new papers. Instead, we reference representative examples of the new advances.

What is Augmented Reality? The basic goal of an AR system is to enhance the user’s perception of and interaction with the real world through supplementing the real world with 3D virtual objects that appear to coexist in the same space as the real world. Many recent papers broaden the definition of AR beyond this vision, but in the spirit of the original survey we define AR systems to share the following properties:

1) Blends real and virtual, in a real environment
2) Real-time interactive
3) Registered in 3D

Registration refers to the accurate alignment of real and virtual objects. Without accurate registration, the illusion that the virtual objects exist in the real environment is severely compromised. Registration is a difficult problem and a topic of continuing research.

Note that this definition of AR is not restricted to particular display technologies, such as a Head-Mounted Display (HMD). Nor is it limited to the visual sense. AR can potentially apply to all senses, including touch, hearing, etc. Certain AR applications also require removing real objects from the environment, in addition to adding virtual objects. For example, an AR visualization of a building that used to stand at a certain location would first have to remove the current building that exists there today. Some researchers call the task of removing real objects Mediated or Diminished Reality, but this survey considers it a subset of Augmented Reality.

Milgram defined a continuum of Real to Virtual environments, where Augmented Reality is one part of the general area of Mixed Reality. In both Augmented Virtuality and Virtual Environments (a.k.a Virtual Reality), the surrounding environment is virtual, while in AR the surrounding environment is real. This survey focuses on Augmented Reality and does not cover Augmented Virtuality or Virtual Environments.

This new survey will not duplicate the content of the 1997 survey. That paper described potential applications such as medical visualization, maintenance and repair of complex equipment, annotation and path planning. It summarized the characteristics of AR systems, such as the advantages and disadvantages of optical and video approaches to blend virtual and real, and problems in the focus and contrast of displays and the portability of AR systems. Registration was highlighted as a basic problem. The survey analyzed the sources of registration error and described strategies for reducing the errors. Please refer to the original survey for details on these topics.

The remainder of this survey organizes the new developments into the following categories: Enabling Technologies, Interfaces and Visualization, and New Applications. Enabling Technologies are advances in the basic technologies required to build a compelling AR environment: displays, tracking, registration, and calibration. The Interfaces and Visualization section describes new research in how users interact with AR systems and what they see displayed. This covers new
user interface metaphors, data density and occlusion problems, more realistic rendering and human factors studies. New Applications include outdoor and mobile systems, collaborative AR, and commercial developments. This survey concludes by describing several areas requiring further research.

2. Enabling Technologies

2.1. See-Through Displays

Display technology continues to be a limiting factor in the development of AR systems. There are still no see-through displays that have sufficient brightness, resolution, field of view, and contrast to seamlessly blend a wide range of real and virtual imagery. Furthermore, many technologies that begin to approach these goals are not yet sufficiently small, lightweight, and low-cost.

Nevertheless, the past few years have seen a number of advances in see-through display technology.

Presence of well-known companies: Established electronics and optical companies, such as Sony and Olympus, now produce opaque, color, LCD-based consumer head-worn displays intended for watching videos and playing video games. While these systems have relatively low resolution (180K-240K pixels), small fields of view (ca. 30º horizontal), and do not support stereo, they are relatively lightweight (under 120 grams) and offer an inexpensive option for video see-through research. Sony introduced true SVGA resolution optical see-through displays, including stereo models (later discontinued), which have been used extensively in AR research.

Parallax-free video see-through displays: One of the challenges of video see-through display design is to ensure that the user’s eyes and the cameras effectively share the same optical path, eliminating parallax errors that can affect the performance of close-range tasks. The Mixed Reality Systems Laboratory developed a relatively lightweight (340 gram) VGA resolution video see-through display, with 51º horizontal field of view, in which the imaging system and display system optical axes are aligned for each eye.

Support for occlusion in optical see-through displays: In conventional optical see-through displays, virtual objects cannot completely occlude real ones. Instead, they appear as ghost images through which real objects can be seen. One experimental display addresses this by interposing an LCD panel between the optical combiner and the real world, making it possible to opacify selected pixels. To avoid having the LCD appear out of focus, it is sandwiched between a pair of convex lenses and preceded by an erecting prism to invert the image of the real world.

Support for varying accommodation: Accommodation is the process of focusing the eyes on objects at a particular distance. In conventional optical see-through displays there is a conflict between the real world, viewed with correctly varying accommodation, and the virtual world, viewed on a single screen with fixed accommodation. In contrast, while conventional video see-through displays provide the same fixed accommodation distance for both real and virtual worlds, the effect is wrong except for those objects that are at the display’s fixed apparent distance. Both cases can result in eyestrain and visual artifacts. Prototype video and optical see-through displays have been developed that can selectively set accommodation to correspond to vergence, by moving the display screen or a lens through which it is imaged. One version can cover a range of .25 m to infinity in .3 sec.

Eyeglass displays: Ideally, head-worn AR displays would be no larger than a pair of sunglasses. Several companies are developing displays that literally embed display optics within conventional eyeglasses. MicroOptical has produced a family of eyeglass displays in which the image of a small color display, mounted facing forward on an eyeglass temple piece, is reflected by a right angle prism embedded in a regular prescription eyeglass lens. Minolta’s prototype forgettable display is intended to be light and inconspicuous enough that the user forgets that it is being worn. Others see only a transparent lens, with no indication that the display is on, and the display adds less than 6 grams to the weight of the eyeglasses.

Virtual retinal displays: In contrast to the virtual images produced by the displays discussed above, virtual retinal displays form their images directly on the retina. These displays, which are being developed commercially by MicroVision, literally draw on the retina with lowpower lasers whose modulated beams are scanned by microelectromechanical mirror assemblies that sweep the beam horizontally and vertically. Potential advantages include high brightness and contrast, low power consumption, and large depth of field.

2.2. Projection Displays

An alternate approach to AR is to project the desired virtual information directly on those objects in the physical world that are to be augmented. In the simplest case, the augmentations are intended to be coplanar with the surface on which they are projected and can be projected monoscopically from a room-mounted projector, with no need for special eyewear. Examples include a projection of optical paths taken through simulated elements on a virtual optical bench, and an application
where a remote user controls a laser pointer worn by
another user to point out objects of interest.

Generalizing on the concept of a multi-walled CAVE
environment, Raskar and colleagues show how large
irregular surfaces can be covered by multiple overlapping
projectors, using an automated calibration procedure that
takes into account surface geometry and image overlap.
They use stereo projection and liquid crystal shutter
eyewear to visualize 3D objects. This process can also be
applied to true 3D objects as the target, by surrounding
them with projectors.

Another approach for projective AR relies on head-worn
projectors, whose images are projected along the viewer’s
line of sight at objects in the world. The target objects are
clothed with a retroreflective material that reflects light
back along the angle of incidence. Multiple users can see
different images on the same target projected by their own
head-worn systems, since the projected images cannot be
seen except along the line of projection. By using
relatively low output projectors, non-retroreflective real
objects can obscure virtual objects.

While these are strong advantages, the use of projectors
poses a challenge for the design of lightweight systems
and optics. Figure 4 shows a new prototype that weighs
under 700 grams. One interesting application of projection
systems is in Mediated Reality. Coating a haptic input
device with retroreflective material and projecting a model
of the scene without the device camouflages the device by
making it appear semi-transparent.

2.4. Calibration and Autocalibration

AR systems generally require extensive calibration to
produce accurate registration. Measurements may include
camera parameters, field of view, sensor offsets, object
locations, distortions, etc. The basic principles of camera
 calibration are well established, and many manual AR
calibration techniques have been developed. One
approach to avoiding a calibration step is the development
of calibration-free renderers. Since Kutulakos and Vallino
introduced their approach of calibration-free AR based on
a weak perspective projection model, Seo and Hong
extended it to cover perspective projection, supporting
traditional illumination techniques. Another example
obtained camera focal length without an explicit metric
calibration step. The other approach to reducing
 calibration requirements is autocalibration. Such
algorithms use redundant sensor information to
automatically measure and compensate for changing
calibration parameters.

3. Interfaces and Visualization

In the last five years, AR research has become broader in
scope. Besides work on the basic enabling technologies,
researchers are considering problems of how users will
interact and control AR applications, and how AR
displays should present information.

3.1. User Interface and Interaction

Until recently, most AR interfaces were based on the
desktop metaphor or used designs from Virtual
Environments research. One main trend in interaction
research specifically for AR systems is the use of
heterogeneous designs and tangible interfaces.
Heterogeneous approaches blur the boundaries between
real and virtual, taking parts from both worlds. Tangible
interfaces emphasize the use of real, physical objects and
tools. Since in AR systems the user sees the real world
and often desires to interact with real objects, it is
appropriate for the AR interface to have a real component
instead of remaining entirely virtual.

In one example of such an interface, the user wields a real
paddle to manipulate furniture models in a prototype
interior design application. Through pushing, tilting,
swatting and other motions, the user can select pieces of
furniture, drop them into a room, push them to the desired
locations, and smash them out of existence to eliminate
them.

Other examples include the Studierstube Personal
Interaction Panel (PIP), several game applications, and
Sony’s Augmented Surfaces system. The Studierstube PIP
is a blank physical board that the user holds, upon which
virtual controls are drawn. The tangible nature of the
interface aids interaction with the controls. The Mixed
Reality Systems Lab created several AR gaming systems.
In the AR Hockey system, two users played an air hockey
game by moving a real object that represents the user’s
paddle. In the RVBorder Guards game, users combat
virtual monsters by using gestures to control their weapons
and shields. In Sony’s Augmented Surfaces system, users
manipulate data through a variety of real and virtual
mechanisms. Users see data through both projective and
handheld displays. A real model of a camera, placed upon
the projection of a top-down view of a virtual room,
generates a 3D rendering of the room from the viewpoint
of that camera.

Similarly, the EMMIE system mixes several display and
device types and enables transferring data across devices
through various operations. EMMIE supports colocated
and remote collaboration amongst several simultaneous
users. The development of collaborative AR interfaces is
the other major trend in interaction research; these are discussed later in the Applications section. Researchers have started exploring collaboration in heterogeneous environments. For example, the Studiers tube and MARS systems support collaboration between co-located and remote users interacting with AR, VR and desktop displays. Another application of such cross-paradigm collaboration is the integration of mobile war fighters (engaged with virtual enemies via AR displays) collaborating with units in a VR military simulation. Alternately, the Magic Book interface allows one or more AR users to enter a VR environment depicted on the pages of the book; when they descend into the immersive VR world, the AR users see an avatar appear in the environment on the book page. The Magic Book requires the display to be able to completely block the user’s view of the world when they descend into the VR environment.

Maximizing performance for a particular application may requiring tuning an interface specifically for that application. The needed modifications may not be initially obvious to the designers, requiring iterative design and user feedback.

3.2. Visualization Problems

Researchers have begun to address problems in displaying information in AR displays, caused by the nature of AR technology or displays. Work has been done in visualizing the registration errors and avoiding hiding critical data due to density problems.

Error visualization: In some AR systems, registration errors are significant and unavoidable. For example, the measured location of an object in the environment may not be known accurately enough to avoid visible registration error. Under such conditions, one approach to rendering an object is to visually display the area in screen space where the object could reside, based upon expected tracking and measurement errors. This guarantees that the virtual representation always contains the real counterpart. Another approach when rendering virtual objects that should be occluded by real objects is to use a probabilistic function that gradually fades out the hidden virtual object along the edges of the occluded region, making registration errors less objectionable.

Data density: If the real world is augmented with large amounts of virtual information, the display may become cluttered and unreadable. The distribution of data in screen space varies depending on the user’s viewpoint in the real world. Julie uses a filtering technique based on a model of spatial interaction to reduce the amount of information displayed to a minimum while keeping important information in view. The framework takes into account the goal of the user, the relevance of each object with respect to the goal and the position of the user to determine whether or not each object should be shown. The EMMIE system models the environment and tracks certain real entities, using this knowledge to ensure that virtual information is not placed on top of important parts of the environment or on top of other information.

3.3. Advanced Rendering

Ideally, virtual augmentations would be indistinguishable from real objects. Such high-quality renderings and compositions are not currently feasible in real time. However, researchers have begun studying the problems of removing real objects from the environment (a.k.a. Mediated Reality) and more photorealistic rendering (although not yet in real time).

Mediated Reality: The problem of removing real objects is more than simply extracting depth information from a scene, as discussed previously in the section on tracking; the system must also be able to segment individual objects in that environment. Lepetit discusses a semiautomatic method for identifying objects and their locations in the scene through silhouettes. This enables the insertion of virtual objects and deletion of real objects without an explicit 3D reconstruction of the environment.

Photorealistic rendering: A key requirement for improving the rendering quality of virtual objects in AR applications is the ability to automatically capture the environmental illumination information. Two examples of work in this area are an approach that uses ellipsoidal models to estimate illumination parameters and Photometric Image-Based Rendering.

4. Future Work

Despite the many recent advances in AR, much remains to be done. Here are nine areas requiring further research if AR is to become commonly deployed.

Ubiquitous tracking and system portability: Several impressive AR demonstrations have generated compelling environments with nearly pixel-accurate registration. However, such demonstrations work only inside restricted, carefully prepared environments. The ultimate goal is a tracking system that supports accurate registration in any arbitrary unprepared environment, indoors or outdoors. Allowing AR systems to go anywhere also requires portable and wearable systems that are comfortable and unobtrusive.

Ease of setup and use: Most existing AR systems require expert users (generally the system designers) to calibrate
and operate them. If AR applications are to become commonplace, then the systems must be deployable and operable by non-expert users. This requires more robust systems that avoid or minimize calibration and setup requirements. Some research trends supporting this need include calibration-free and auto-calibration algorithms for both sensor processing and registration.

Broader sensing capabilities: Since an AR system modifies the user’s perception of the state of the real environment, ideally the system needs to know the state of everything in the environment at all times. Instead of just tracking a user’s head and hands, an AR system should track everything: all other body parts and all objects and people in the environment. Systems that acquire real-time depth information of the surrounding environment, through vision-based and scanning light approaches, represent progress in this direction.

Interface and visualization paradigms: Researchers must continue developing new interface techniques to replace the WIMP standard, which is inappropriate for wearable AR systems. New visualization algorithms are needed to handle density, occlusion, and general situational awareness issues. The creation and presentation of narrative performances and structures may lead to more realistic and richer AR experiences.

Proven applications: Many concepts and prototypes of AR applications have been built but what is lacking is experimental validation and demonstration of quantified performance improvements in an AR application. Such evidence is required to justify the expense and effort of adopting this new technology.

User studies and perception issues: Few user studies have been performed with AR systems, perhaps because few experimenters have access to such systems. Basic visual conflicts and optical illusions caused by combining real and virtual require more study. Experimental results must guide and validate the interfaces and visualization approaches developed for AR systems.

Photorealistic and advanced rendering: Although many AR applications only need simple graphics such as wireframe outlines and text labels, the ultimate goal is to render the virtual objects to be indistinguishable from the real. This must be done in real time, without the manual intervention of artists or programmers. Some steps have been taken in this direction, although typically not in real time. Since removing real objects from the environment is a critical capability, developments of such Mediated Reality approaches are needed.

AR in all senses: Researchers have focused primarily on augmenting the visual sense. Eventually, compelling AR environments may require engaging other senses as well (touch, hearing, etc.) For example, recent systems have demonstrated auditory [52] and haptic AR environments [88].

Social acceptance: Technical issues are not the only barrier to the acceptance of AR applications. Users must find the technology socially acceptable as well. The tracking required for information display can also be used for monitoring and recording. How will non-augmented users interact with AR-equipped individuals? Even fashion is an issue: will people willingly wear the equipment if they feel it detracts from their appearance?

6. Acknowledgements

Part of this work was funded Office of Naval Research contracts N00014-99-3-0018, N00014-00-1-0361, N00014-99-1-0249, N00014-99-1-0394, N00014-990683, and NSF Grant IIS-00-82961.

7. References

[1] Y. Akatsuka, G. Bekey, Compensation for end to end delays in a VR system, Proc. Virtual Reality Ann. Int'l Symp. 98. (VRAIS '98). Atlanta, 14-18 Mar. 1998, pp. 156-159.

[2] ARToolkit. http://www.hitl.washington.edu/research/shared_space/

[3] R. Azuma, A Survey of Augmented Reality, Presence: Teleoperators and Virtual Environments vol. 6, no. 4, Aug. 1997, pp. 355-385.

[4] R. Azuma, et. al., A Motion-Stabilized Outdoor Augmented Reality System, Proc. IEEE Virtual Reality 99. Houston, TX, 13-17 Mar. 1999, pp. 252-259.

[5] R. Azuma, et. al., Tracking in unprepared environments for augmented reality systems, Computers & Graphics vol. 23, no. 6, Dec. 1999, pp. 787-793.

[6] R. Behringer, Registration for Outdoor Augmented Reality Applications Using Computer Vision Techniques and Hybrid Sensors, Proc. IEEE Virtual Reality 99. Houston, TX, 13-17 Mar. 1999, pp. 244-251.

[7] R. Behringer, et. al., A Wearable Augmented Reality Testbed for Navigation and Control, Built Solely with Commercial-Off-The-Shelf (COTS) Hardware, Proc. Int'l Symp. Augmented Reality 2000 (ISAR 00). Munich, 5-6 Oct. 2000, pp. 12-19.
[8] M. Billinghurst, H. Kato, Collaborative Mixed Reality, *Proc. Int l Symp. Mixed Reality (ISM R '99). Mixed Reality - Merging Real and Virtual Worlds*, Yokohama, Japan, 9-11 Mar. 1999, pp. 261-284.

[9] F.A. Biocca, J.P. Rolland, Virtual Eyes Can Rearrange Your Body: Adaptation to Visual Displacement in See-Through, Head-Mounted Displays, *Presence: Teleoperators and Virtual Environments.* vol. 7, no. 3, June 1998, pp. 262-277.

[10] A. Butz, et. al., Enveloping Users and Computers in a Collaborative 3D Augmented Reality, *Proc. 2nd Int l Workshop Augmented Reality. (IWAR ’99).* San Francisco, 20-21 Oct. 1999, pp. 35-44.

[11] R. Cavallaro, The FoxTrax Hockey Puck Tracking System, *IEEE CG&A.* vol. 17, no. 2, Mar./April 1997, pp. 6-12.

[12] G. Sahoo, R. K. Tiwari, “Designing an Embedded Algorithm for Data Hiding using Steganographic Technique by File Hybridization”, IJCSNS, Vol. 8, No. 1, pp. 228-233, January 2008.

[13] Kumar Gunjan, R. K. Tiwari and G. Sahoo, “Towards Securing APIs in Cloud Computing” International Journal of Computer Engineering and Applications, Volume 2, Issue 2, 2014.

[14] Abu Salim, R. K. Tiwari and S. Tripathi “Secure Cloud Environment a Novel Approach”, International Journal of Computer Engineering and Applications Volume XI, Issue IX.

[15] Rajesh Kumar Tiwari “Hybrid database A Steganographic Approach”, International Journal of Computer Engineering and Applications, Volume II, Issue I, 2013.