

The Evolution and Future Scope of Augmented Reality

Charvi Agarwal¹ and Narina Thakur²

^{1,2}Department of Computer Science & Engineering, Bharati Vidyapeeth's College of Engineering
New Delhi, India

1

Abstract

This paper evaluates current practices and research being done in the field of augmented reality, reviews best practices and suggests new ways of incorporating this technology in our everyday lives. In doing so, it discusses the problems faced while implementing this technology and some ways to overcome them. It describes enabling technologies such as displays and trackers that are essential in building any AR system. This paper describes various fields in which AR is already being used (such as medical, military, entertainment, manufacturing, education etc.) and also presents new areas of applications for it. It can be used as a starting point for anyone who is new to this area of technology.

Keywords: augmented reality, enabling technologies, applications, limitations

1. Introduction

Imagine a technology with which you could see more than others see and hear more than others hear. Augmented Reality (AR) enables computer generated virtual imagery to exactly overlay physical real world objects in real time. The main goal of an AR system is to improve the user's perception as well as interaction with the real world by supplementing, or *augmenting*, the real world with 3D virtual objects that appear to coexist in the same space as the real world. As opposed to virtual reality (VR), where the real world is entirely replaced by a virtual one, AR allows the user to interact with virtual images using real objects in a seamless way. A widely accepted definition of AR is any system which

1. Combines real and virtual world,
2. Is interactive in real time, and
3. Is registered in 3D [1].

Here, the term registration means the accurate alignment of real and virtual objects with respect to each other. Without accurate registration, the appearance of the coexistence of virtual elements in the real environment with physical objects would be severely compromised. Registration has proved to be a difficult problem and remains a topic of continuing research.

Here, three main aspects of the given definition must be mentioned. Firstly, it is not restricted to specific display technologies such as a head-mounted display (HMD), which was usually always associated with

AR in the past. It is also not limited to the sense of sight, as AR can potentially apply to all senses, including hearing, touch, and smell. Further, mediated or diminished reality, which refers to overlaying virtual objects to obscure real ones, is also considered AR. For example, we would have to remove the current building that exists at a location in order to visualize a new building to take its place.

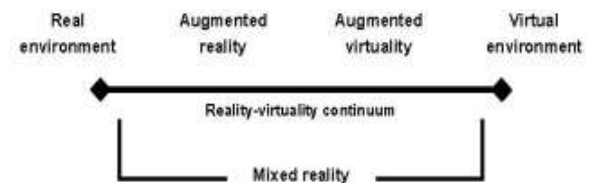


Fig. 1: Reality-virtuality continuum [10]

For anyone who is new to this field and wants to get familiarized with it, this paper gives an overview of important technologies, applications and limitations of AR systems, along with its future scope.

2. Brief History and Trends

- The first AR prototypes were made in the 1960s and used a see-through display to present 3D graphics. They were created by Ivan Sutherland and his students at Harvard University and the University of Utah.
- A group of researchers at U.S. Air Force's Armstrong Laboratory, the NASA Ames Research Center, the University of North Carolina at Chapel Hill and the Massachusetts Institute of Technology continued research during the 1970s and 1980s. During this time, we saw the advent of mobile devices such as the Sony Walkman, personal digital organizers and digital watches. This expedited wearable computing in the 1990s as personal computers became small enough in size to be able to be worn at all times.
- The term "augmented reality" was coined by Caudell and Mizell in the early 1990s, who were developing an AR system to facilitate workers put together wiring harnesses at Boeing Corporation. Also, early palmtop

computers such as the Psion I were introduced.

- AR became a distinct field of research by the late 1990s and several conferences such as the International Workshop and Symposium on Augmented Reality, the International Symposium on Mixed Reality, and the Designing Augmented Reality Environments workshop began.
- Nowadays, several mobile platforms such as tablet PCs, smartphones and personal digital assistants (PDAs) exist which are capable of supporting AR technologies. It is now possible to build AR applications using software development toolkits such as the ARToolKit, which are available for free. Over the years, many surveys have also appeared that provide an overview of various advancements in the field of AR.

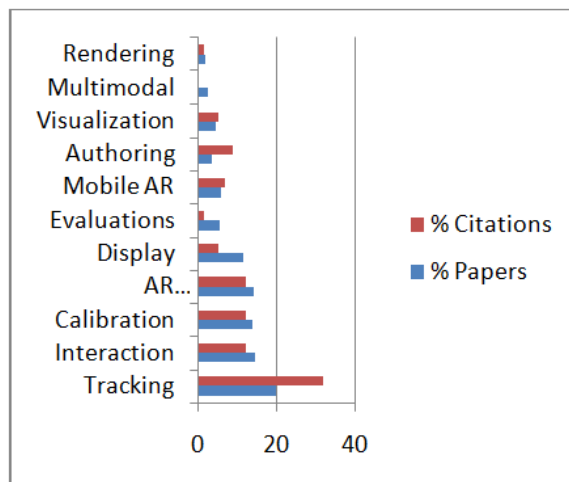


Fig. 2: Proportion of Highly Cited Papers

Figure 2 shows the proportion of papers that have been published over the years in the given categories. These papers were published in conferences such as ISMAR, ISMR, ISAR and IWAR [7]. It is clear that tracking techniques dominate the research done in the field of AR, which is expected as achieving accurate tracking remains one of the major challenges that are faced in implementing AR systems. Interaction techniques, calibration, display and applications are some of the other categories which have been extensively explored over time, as enabling technologies of AR systems continue to be a limiting factor and more and more applications of AR are being discovered.

3. Enabling Technologies

Enabling technologies are the advances in basic technologies required to build an AR system. The technological demands for AR are much higher than for virtual environments or VR, which is why the field of AR took longer to mature than that of VR. However, the key components needed to build an AR system have remained the same since Ivan Sutherland's pioneering work of the 1960s. Displays, trackers, and graphics computers and software remain essential in many AR experiences. These are discussed as follows.

3.1 Display Technologies

Display technology continues to be a limiting factor in the development of AR systems. There are still no see-through displays that have adequate brightness, contrast, resolution and field of view to blend a wide range of real and virtual imagery in a seamless manner. Moreover, many technologies that attempt to approach these goals are not yet sufficiently small, lightweight, and low-cost. The display technologies mainly focus on three types: see-through head-mounted displays, projection-based displays and handheld displays.

See-through Displays: See-through HMDs are mostly used to allow the user to view the real world with virtual objects superimposed on it by optical or video technologies. They can be broadly divided into two categories: optical see-through and video see-through HMDs.

- **Optical see-through displays** are those that allow the user to see the real world with their natural eyes, and graphics are overlaid onto the users view by using a holographic optical element, half silvered mirror or similar technology. The main advantage of these displays is that they offer a superior view of the real world including a natural, instantaneous view of the real scene, seamlessness between aided and periphery views as well as simple and lightweight structures. They are also considered the safer alternative as the user can still see the real world even if the power is turned off.
- **Video see through displays** are those in which the user has a video view of the real world with virtual graphics overlaid on it. There are several advantages to this approach, such as consistency between real and synthetic views, and availability of a variety of image processing techniques like

the correction and control of intensity, tint and blending ratio. Hence, video see-through displays can handle occlusion problems more easily compared to optical see-through displays. They are also the cheapest and easiest to implement.

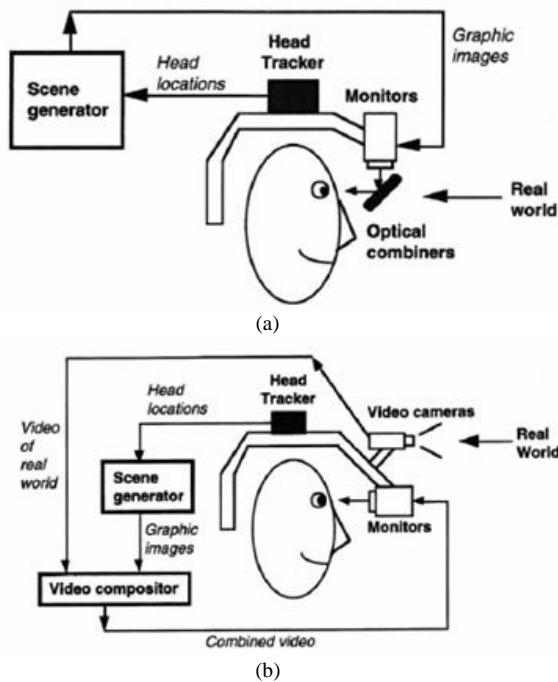


Fig. 3: Conceptual diagrams of (a) Optical see-through and (b) Video see-through HMDs [1].

Projection-based Displays: Projection-based displays are a good option for applications that do not require several users to wear anything. They can be used to provide a wide field of view. A variety of projection-based display techniques have been proposed for displaying graphical information directly on real objects and surfaces. Bimber and Fröhlich presented projector illumination techniques for creating correct occlusion effects with view-dependent optical see-through displays [7]. In this way, they combined the desirable features of projectors and optical displays.

Handheld Displays: These include handheld see-through displays as well as handheld projectors. Although these displays are bulkier than head-mounted displays, they are presently the best workaround to introduce AR to a mass market owing to low production costs and ease of use. They are also highly mobile and provide minimal intrusion. Devices such as mobile phones and tablets can be used to support AR applications. For instance, several

navigational and gaming applications can be efficiently implemented on handheld platforms.

Spatial Displays: This type of display technology includes screen-based video see-through displays, spatial optical see-through displays, and projective displays. These techniques prove to be useful for large presentations and exhibitions that do not require a lot of interaction. This technique is now extensively being applied in the world of sports television where environments such as swimming pools and race tracks are well defined and easy to augment. Head-up displays (HUDs) are a form of spatial optical see-through which are becoming a standard extension for production cars to project navigational directions in the windshield. Spatial see-through displays may however appear misaligned when users move around in open spaces, for example when AR overlay is presented on a transparent screen.

Aural and Haptic Displays: Aural display applications in AR are mostly limited to self-explanatory mono (0-dimensional), stereo (1-dimensional) or surround (2-dimensional) headphones and loudspeakers. Today, 3D aural display can be found in highly immersive simulations of virtual environments and augmented virtuality, or technologies still in experimental stages. 'Haptic audio' refers to sound that is *felt* rather than heard, and is already being applied in consumer devices such as Turtle Beach's Ear Force5 headphones to increase the sense of realism and impact, as well as to enhance user interfaces of mobile phones.

3.2 Tracking Techniques

It is crucial to accurately track the user's viewing orientation and location for AR registration. Several techniques have been employed to build tracking systems and many of these have shown great accuracies in indoor environments. These successful trackers often use hybrid techniques that combine various techniques to overcome their individual limitations and provide better results. These are discussed below.

Sensor-Based Tracking Techniques: Before displaying virtual objects in a real environment, an AR system must be capable of sensing the environment and tracking the viewer's (relative) movement preferably with six degrees of freedom, i.e. three variables (x, y, and z) for position and three angles (yaw, pitch, and roll) for orientation. Sensor-based tracking techniques are based on various kinds of sensors such as magnetic, inertial, acoustic, optical and mechanical sensors. All of these have their own

advantages and disadvantages. For example, magnetic sensors are light and have a high update rate, but they can be distorted by any nearby metallic substance or the presence of another device with its own magnetic field that disturbs the magnetic field of the sensor.

Vision-Based Tracking Techniques: Vision-based tracking techniques can use image processing techniques to calculate the camera pose and angle relative to real world objects. Thus, they are analogous to closed loop systems which can dynamically correct errors. In the field of computer vision, tracking techniques are classified into two main categories:

- **Feature-based:** In these methods, the underlying objective is to find a correspondence between 2D image features and their 3D world frame coordinates. When this is done, the camera pose can be found from projecting the 3D coordinates of the feature into the observed 2D image coordinates and minimizing the distance to their corresponding 2D features.
- **Model-based:** Model-based tracking techniques are the ones that explicitly use a model of the features of tracked objects such as a CAD model or 2D template of the object based on the distinguishable features. These kinds of trackers usually construct their models based on lines or edges in the model.

Hybrid Tracking Techniques: For some AR applications, computer vision alone cannot sufficiently provide robust tracking solutions. Thus, many hybrid methods have been developed that combine several sensing technologies. Several researchers aimed to combine inertial and computer vision technologies to provide closed-loop-type tracking. Vision-based tracking has low jitter and no drift, but it is slow, which is a major drawback. Compared to these tracking methods, inertial tracking offers desirable complementary features. It is fast, robust and can be used for motion prediction when changes occur rapidly. Furthermore, we can recover object pose using measurements of acceleration and rotational speed, but inertial trackers can drift due to accumulation of noise.

4. User Interface and Interaction

For AR systems, one main trend in interaction research is the use of heterogeneous designs and tangible interfaces. Heterogeneous approaches aim to blur the boundaries between the physical and virtual worlds by taking elements from both worlds. Tangible interfaces highlight the use of real, physical objects and tools. In AR systems, since the user sees the real world and often wants to interact with real objects, it is fitting for the AR interface to have a real component instead of remaining an entirely virtual environment.

4.1 Tangible AR

Previously, Ishii developed the concept of tangible user interfaces (TUIs), which are interfaces where users can manipulate digital information with physical objects. Tangible interfaces are a powerful concept because the physical objects used have properties which are familiar, physical constraints, and affordances which make them easy to use. We can apply the same concept in AR interfaces where we can combine the intuitiveness of the physical input devices with the enhanced display possibilities provided by virtual image overlays.

4.2 Collaborative AR

Researchers have started exploring collaboration in heterogeneous environments. For instance, the Studierstube and MARS systems support collaboration between co-located and remote users interacting with AR, VR and desktop displays [8]. Another application of such a cross-paradigm collaboration is the integration of mobile warfighters (engaged with virtual enemies via AR displays) collaborating with units in a VR military simulation.

4.3 Gesture Recognition

Instead of using hand-worn trackers, hand movement can also be tracked visually which would be more convenient for the user. A head-worn or collar-mounted camera that points at the user's hands can be used for gesture recognition. Using gesture recognition, an AR system could automatically draw up reports of activities. Cameras can also be useful to record and document the user's view, e.g. for providing a live video feed for teleconferencing, for informing a remote colleague about the findings of AR field-workers, or just for documenting everything that is taking place in front of the user of the mobile AR system.

4.4 Hybrid UI

Since each modality has its own set of drawbacks and benefits, AR systems are likely to use a multimodal UI. As an example, a synchronized combination of gestures, sound, speech, vision and haptics may provide users with a more natural, predictable and robust UI.

5. Existing Applications

There has been considerable amount of research and development in the following fields of applications for AR.

5.1 Medical Applications

- Doctors could use augmented reality as a visualization and training aid for surgery. It may be possible to collect 3D datasets of a patient in real time using sensors like magnetic resonance imaging (MRI), ultrasound imaging, or computed tomography scans (CT). These datasets could then be rendered and combined in real time with a view of the real patient. This would effectively provide a doctor "X-ray vision" inside a patient. Since surgeons can detect some features with the naked eye that they cannot see in computerized scans, and vice versa, AR would provide surgeons access to both types of data simultaneously, thereby enhancing their field of view with more information and increasing efficiency.
- AR might also be useful for training purposes. Virtual instructions could remind a novice surgeon of the required steps, without the need to look away from a patient to consult a manual.
- Roaming nurses and doctors could benefit from important information being delivered directly to their glasses, as it would also save time.

5.2 Manufacturing and Repair

- Instructions may be easier to understand if they were available as 3D drawings superimposed upon the actual equipment showing step-by-step the tasks to be done, instead of appearing as manuals with text and pictures which would be more cumbersome to follow. These superimposed 3D drawings can be animated as well, which

would make the directions even easier to follow.

- AR can also be extremely helpful when interacting with complex machinery and structures. For instance, it can provide "x-ray vision" to look into machines or automatically probe the environment using extra sensors to bring the user's attention to problem sites.

5.3 Annotation and Visualization

- AR could be used to annotate objects and environments with public or private information. Applications using public information would assume the availability of public databases to draw upon. For example, a handheld display such as a smartphone could provide information about the contents of library shelves as the user walks around the library in real time. A user could point to the parts of an engine model and the AR system would display the name of the part that is being pointed at. Similarly, AR could provide the list of items available at a grocery store. Alternately, these annotations might be private notes attached to specific objects. For example, a person might make a list of reminders and attach a window containing this list in her field of vision.
- AR can aid general visualization tasks as well. Architects with a see-through HMD might be able to look out a window and see how a proposed new skyscraper would change their view. AR might also give architects "X-ray vision" inside a building, showing where the pipes, electric lines, and structural supports are inside the walls.

5.4 Entertainment and Gaming

- In cases where a large public is reached, AR can also serve advertisers to show virtual ads and product placements. Well-known environments such as football fields, swimming pools, race tracks etc. are easily prepared which video see-through augmentation through tracked camera feeds easy.
- Augmented reality allows gamers to experience virtual game play in a real world environment. In the last few years there have been a lot of improvements of technology, resulting in enhanced movement detection

and the possibility for the Wii to exist, but also direct detection and tracking of the player's movements. Piekarski and Thomas created a game called "ARQuake" where mobile users fight virtual enemies in a real environment [4]. A number of other games have been developed for prepared indoor environments, such as the alien-battling "Aqua-Gauntlet", dolphin-juggling "ContactWater", "ARHockey", and "2001 AR Odyssey".

5.5 Military Aircraft

- For many years, military aircraft and helicopters have used head-up displays (HUDs) and helmet mounted sights (HMS) to superimpose vector graphics upon the pilot's view of the real world, thereby augmenting it. Besides providing basic navigation and flight information, this technology has the capability of registering targets in the pilot's field of view, providing a way to aim the aircraft's weapons. For example, if the weaponry in a fighter jet is made slave to the pilot's HMS, the pilot can aim the weapon simply by looking at the target.
- Training in large-scale combat scenarios and simulating real-time enemy action, as in the Battlefield Augmented Reality System (BARS), can be useful to military users.

5.6 Education and Training

- Augmented reality applications can supplement a standard curriculum. Elements such as text, images, video and audio can be superimposed into a student's real time view of the environment. Educational material such as textbooks, flashcards and other reading material can contain embedded "markers" that produce supplementary information to the student in a multimedia format, when they are scanned by an AR device.
- AR technology also allows learning via remote collaboration, in which students and instructors who are not at the same physical location can share a common virtual learning environment augmented by virtual objects and learning materials, and interact with another within that environment.
- Kaufmann et al. (2006) introduced the Construct3D tool for math and geometry

education, based on the StudierStube framework [8].

- MIT Education Arcade introduced game-based learning in "Mystery at the Museum" and "Environmental Detectives" which promote engagement as well as learning owing to their rich backstories, distinct characters, synthetic activities and embedded replay [10].
- In art education, Caarls et al. (2009) present multiple examples where AR is used to create new forms of visual art [10].

6. New Applications

The above mentioned applications have been explored but many of them are still not widely used due to technological limitations such as accuracy and portability, and social reluctance. At the same time, some new areas have proven to have a tremendous potential in use of AR technology. These are discussed as follows.

6.1 Collaboration

A major potential benefit of AR is having multiple people view, discuss, and interact with 3D models simultaneously. Collaborative environments allow integration with existing tools and practices and enhance practice by supporting distributed, remote and collocated activities that would otherwise be impossible. Employees of the same company located in offices that are continents apart could interact with each other in the same virtual environment. Collaborative AR systems can use both see-through handheld displays (as seen in MagicBook) and see-through head-worn displays (such as in StudierStube).

6.2 Navigation

- Rekimoto (2000) presented NaviCam for indoor use that augmented a video stream from a hand held camera using fiducial markers for position tracking [10].
- Starner et al. (1998) considered applications and limitations of AR for wearable computers, including problems of finger tracking and facial recognition [10].
- Narzt et al. discussed navigation paradigms for pedestrians and cars that overlay routes, highway exits, follow-me cars, dangers, fuel prices, etc. They prototyped video see-through PDAs and mobile phones and

envision eventual use in car windshield heads-up displays.

6.3 Tourism

- AR applications can augment a user's experience when traveling by providing real time information on displays regarding a location and its attributes, including reviews and comments made by previous visitors of the site which might be helpful to tourists.
- AR applications allow tourists to experience simulations of historical events, places and objects by overlaying them onto their current view of a site. Vlahakis et al. present the ArcheoGuide project that reconstructs a cultural heritage site in Olympia, Greece. Using this system, visitors can view as well as learn ancient architecture and customs.
- AR applications may also present location information using audio, for example by announcing features of interest at a particular site as they come into the user's field of view.

7. Limitations

Today, AR faces several technical challenges regarding stereo view, color depth, luminance, high resolution, contrast, focus depth and field of view. Researchers have begun to address problems in displaying information in AR displays that are caused by the nature of AR technology or displays. Work has been done in visualizing the registration errors, avoiding hiding critical data due to density problems and at the same time not cluttering the screen with excessive information. However, before AR becomes accepted as part of the user's everyday life, issues regarding intuitive interfaces, cost, weight, power usage, ergonomics, and appearance must also be addressed. Some of the major problems are discussed below.

1. **Portability and Outdoor Use:** Most mobile AR systems are bulky and cumbersome, requiring a heavy backpack to carry the PC, sensors, display, batteries, and other components. Connections between all the devices must be able to withstand outdoor use, including weather and shock. Optical and video see-through displays are usually not suited for outdoor use due to low brightness, contrast, resolution, and field of view. However, laser-powered displays offer a new alternative to overcome this problem.
2. **Tracking and Calibration:** Tracking in unprepared/outdoor environments remains a challenge but hybrid approaches are becoming small enough to be added to mobile devices. Calibration of these devices is still complicated and extensive, but it may be solved through calibration-free or auto-calibrating approaches that minimize set-up requirements.
3. **Latency:** A major source of dynamic registration errors are system delays. Techniques like precalculation, temporal stream matching and prediction of future viewpoints may solve some delay. Through careful system design, system latency can be scheduled to reduce errors and pre-rendered images can be shifted at the last instant to compensate for pan-tilt motions. Likewise, image warping may correct delays in 6DOF motion (both translation and rotation).
4. **Depth Perception:** Problems such as accommodation-vergence conflicts or low resolution and dim displays cause object to appear further away than they really are. Correct occlusion ameliorates some depth problems, as does consistent registration for different eye point locations. In an experiment by Biocca and Rolland (1998), subjects exhibit a large overshoot in a depth-pointing task after removing the HMD.
5. **Data Density:** If the real world is augmented with a large amount of virtual information, the display may become cluttered and overpopulated with unnecessary data. The distribution of data in screen space varies depending on the user's viewpoint in the real world. The user interface must follow some guidelines as not to overload the user with information and at the same time must prevent the user from overly relying on the AR system such that important cues from the environment are missed.
6. **Social Acceptance:** Making AR a part of everyday life may be more challenging than expected, as many factors play a role in social acceptance of AR ranging from unobtrusive fashionable appearance (gloves, helmets, etc.) to privacy concerns. For example, Accenture's Assistant blinks a light when it records for the sole purpose of alerting the person who is being recorded. These issues must be addressed before AR is widely accepted.
7. **Adaptation and Long-Term Use:** User adaptation to AR equipment can negatively

impact performance. AR displays that are uncomfortable may not be suitable for long-term use. One study found that binocular displays, where the same image is shown on, caused significantly more discomfort than monocular displays, in both eye strain as well as fatigue.

8. Conclusion

AR has come a long way but still has some distance to go before industries, the military and the general public will accept it as a familiar user interface. With social acceptance, it would be possible for widespread use of AR systems in everyday life.

Schools could employ AR technology to facilitate education at all levels, be it toddlers learning about shapes or college students visualizing and learning about different components of a machine. Navigation would become even easier and more accurate. Personal reminders annotated through AR would be extremely useful and would expedite communication. With effective utilization, AR will revolutionize the way people live and the way industries work.

With mediated reality and efficient visualization, AR has a tremendous potential in fields such as art, architecture, construction, manufacturing and repair. It also has the potential to dominate the areas of education, training, and simulation as well as medical and military applications if exploited fully. In order to achieve this, all the aforementioned limitations must be overcome.

Registration accuracy, latency and tracking accuracy are the main areas of concern that require more improvements. 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. Also, 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. Portability is also still an issue especially with HMDs.

Current research has been focusing on virtual retinal displays and contact lenses that might contain the elements for display embedded into the lens including LEDs, integrated circuitry, and an antenna to provide wireless communication. With the advent of such complementary technologies as tactile networks, artificial intelligence, cybernetics, and (non-invasive) brain-computer interfaces, AR might soon pave the way for ubiquitous (anytime-anywhere) computing of a more natural kind.

References

- [1] Azuma, Ronald T. "A survey of augmented reality." *Presence* 6.4 (1997): 355-385.
- [2] Azuma, R., Baillet, Y., Behringer, R., Feiner, S., Julier, S., & MacIntyre, B. (2001). Recent advances in augmented reality. *Computer Graphics and Applications, IEEE*, 21(6), 34-47.
- [3] Feiner, S., Macintyre, B., & Seligmann, D. (1993). Knowledge-based augmented reality. *Communications of the ACM*, 36(7), 53-62.
- [4] Piekarski, Wayne, and Bruce Thomas. "ARQuake: the outdoor augmented reality gaming system." *Communications of the ACM* 45.1 (2002): 36-38.
- [5] Wagner, Daniel, et al. *Towards massively multi-user augmented reality on handheld devices*. Springer Berlin Heidelberg, 2005.
- [6] Reitmayr, Gerhard, and Tom W. Drummond. "Going out: robust model-based tracking for outdoor augmented reality." *Mixed and Augmented Reality, 2006. ISMAR 2006. IEEE/ACM International Symposium on*. IEEE, 2006.
- [7] Zhou, Feng, Henry Been-Lirn Duh, and Mark Billinghurst. "Trends in augmented reality tracking, interaction and display: A review of ten years of ISMAR." *Proceedings of the 7th IEEE/ACM International Symposium on Mixed and Augmented Reality*. IEEE Computer Society, 2008.
- [8] Szalavári, Zsolt, et al. "'Studierstube': An environment for collaboration in augmented reality." *Virtual Reality* 3.1 (1998): 37-48.
- [9] Schmalstieg, Dieter, Tobias Langlotz, and Mark Billinghurst. *Augmented Reality 2.0*. Springer Vienna, 2011.
- [10] Van Krevelen, D. W. F., and R. Poelman. "A survey of augmented reality technologies, applications and limitations." *International Journal of Virtual Reality* 9.2 (2010): 1.
- [11] ARToolkit, www.hitl.washington.edu/artoolkit
- [12] Augmented Reality Wikipedia, http://en.wikipedia.org/wiki/Augmented_reality