Augmented Reality Coaching in an Extended Training Environment

Richard St. Augustine

Paratus Associates

Beavercreek, Ohio

richard.staugustine@gmail.com

Richard Goad

Paratus Associates

Beavercreek, Ohio

rcgoad5@gmail.com

ABSTRACT

This paper presents a methodology for maximizing the outcome(s) of training using augmented reality (AR) coaching techniques that exploit interpersonal coordination and an enriched environment. Many commercial and defense training programs involve multiple trainees in synchronous operation. Personnel availability and scheduling reduce the volume and quality of on-demand training availability. Current virtual/augmented-reality training programs focus on asynchronous tasks and roles, which restricts the improvement possibilities of those programs.

Our AR Coaching focuses on simulating the appearance and movement of synchronous human navigation and action in accordance with career-specific training guidelines. We leverage cutting-edge augmented-reality software-development frameworks to construct virtual images of personnel involved in collaborative and competitive training sessions. These objects maneuver and engage in pre-programmed paths and tasks for which the user must account when determining where and how to execute his mission requirements. For example, he establishes the muscle memory of following a teammate while remaining out of sight from an enemy combatant navigating a shared environment.

Paratus will provide a mobile application that presents a simulation environment complete with virtual humans involved in the training operation. The trainee will be able to physically navigate an extended plane while viewing his training environment on the mobile display. This display's visual depiction of the training simulation will change in coordination with his position and orientation.

ABOUT THE AUTHORS

Richard St. Augustine is a software engineer with a degree in Mechanical Engineering from Wright State University and a diverse programming background that includes robotics, additive manufacturing, mobile applications, and data processing. He is the technical lead for software development and data science.

Richard Goad is a program manager with several decades of experience managing and leading major acquisition programs principally in the FMS environment and research and development projects for various DoD laboratories.

Interactive Coaching of Highly Independent Participants in AR

Richard St. Augustine
Paratus Associates
Beavercreek, Ohio
richard.staugustine@gmail.com

Richard Goad
Paratus Associates
Beavercreek, Ohio
rcgoad5@gmail.com

INTRODUCTION

Trainers in any collaborative or competitive environment utilize many mental, physical, tangible, and intangible techniques for visualizing and anticipating the interactions of teammates and competitors. Helpful visualizations include the locations to (and means by) which the individuals and groups will move to effectively complete their assignments. For example, after watching all of the opponent's game film, an American football coach may create a booklet of the opponent's top 30 plays. Then, he could use scout players, screenshots, videos, and drawings to present the formations and plays to his own players. This visualization provides players with understanding of the location vectors, eye progressions, and footwork required to complete their assignments and outmaneuver the competition.

Today's audiovisual technology and techniques make this visualization incredibly accurate and well-informed, but two primary capabilities remain absent:

1. Creating synchronous progressions of anticipatory visualizations and
2. Viewing the same circumstances from multiple potential perspectives.

We combine these issues into a central problem statement: trainers will benefit from the ability to create, adjust, and impose accurate visualizations of what the trainees will likely see during the event for which they're training. Virtual Reality (VR) and Augmented Reality (AR) provide this capability, but their development presently remains limited to a predetermined (mostly) stagnant position vector. By focusing on the visibility of surroundings provided with AR, we will leverage this technology to expand the operating space of training programs while providing virtual objects and sequences in accordance with relevant job-specific training protocols.

Related Innovation

Multiple industries currently use VR and/or AR to enhance training. The Gaming Research Integration for Learning Lab at Wright Patterson Air Force Base in Dayton, Ohio (the Headquarters for the Air Force Research Laboratory) develops augmented-reality programs for aircraft maintenance training (Clement et al, 2020). The since-disbanded Local Games Lab ABQ at the University of New Mexico used AR to improve classroom learning by superimposing augmented-reality graphics originally developed for the growing gaming industry (Holden, 2014). Strivr provided teams in the National Football League with the capability of watching game film using virtual reality (VR) glasses/goggles to allow quarterbacks to prepare their decision-making with a more life-like perspective (Tanier, 2016). VAR Systems developed three-dimensional playbooks for any position that can be viewed with a 360-degree perspective (*3D Playbook*, VAR). ARHT developed holographic technology that would allow instructors to conduct realistic training via "in-person" briefings at multiple sites without leaving the headquarters (*ARHT Suite*, ARHT).

Technological Overview

In the Association for Computing Machinery (ACM) Computing Surveys, Cao et al. provide an in-depth summary of the current technology and methodology used in creating programs such as those described in the previous section. Because our program will be deliverable via phone application, it would fall under the category of Mobile Augmented Reality (MAR). Figure 1 illustrates the systems commonly involved in this process.

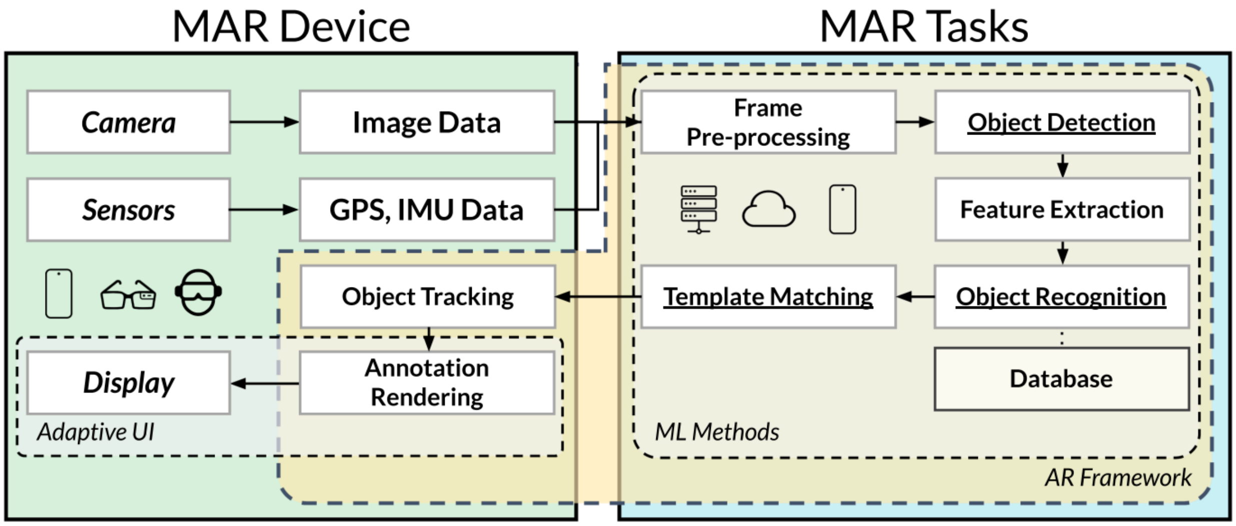


Figure 1. Typical Mobile Augmented Reality (MAR) Pipeline, Split into MAR Device and Tasks.

Our implementation of this technology will become relevant through a phased process. In Phase One, the trainee will use a mobile phone or tablet and its camera to view an augmented training environment while physically navigating that space and viewing it from multiple perspectives. To properly view the scenario from multiple perspectives, the application must detect the present environment's objects and features and overlay them with those that are absent yet required by the relevant training doctrine. Then, the virtual objects need to be anchored to a path required by the training simulation in relation to the physical environment so that the trainee can view them from multiple perspectives as he would those physically present. Finally, the program must analyze the trainee's movements in relation to the recognized and programmed environment so that the trainer can track and edit performance metrics and sequences in the same manner that he would execute repetitions of a training event. In Phase Two, the application will utilize head-mounted displays to maintain a fixed perspective relative to the trainee's body while freeing his hands to execute manual tasks on the move relevant to the training objectives. In Phase Three, the program will adopt a virtual "coach" that can analyze performance and provide feedback with direct input from the sensor data of the display device that learns from repeat use and can improve data-backed, performance-based feedback over time. Finally, in Phase Four, the application will respond to and accommodate multiple use cases in a shared training environment to allow multiple trainees to use the program and collectively improve team performance through a shared, real-time feedback system that accounts for and adopts to multi-personnel participation in a shared training mission.

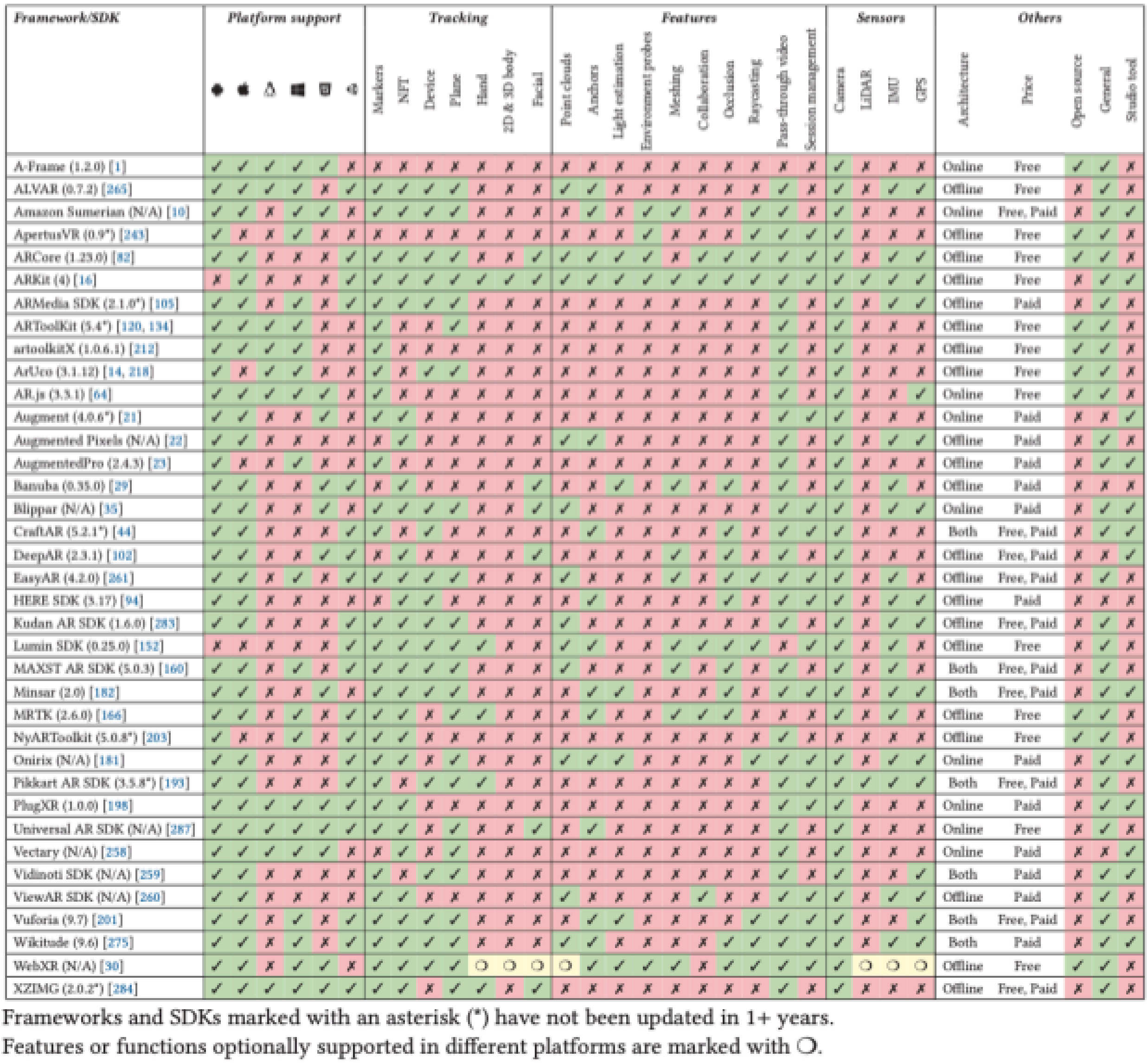
Leveraging Frameworks

Devices that provide the systems detailed in Figure 1 rely on frameworks developed by and for the companies that provide such devices to execute any function. These software development kits (SDKs) optimize programs for specific devices while providing nuanced capability, and thus dozens of them are created for any functionality. Table 1 details the summary of the capabilities and applications of the frameworks and SDKs in development and use (as of January 2023) for reality augmentation provided by Cao et al. We will focus on Apple products due to our technical lead's experience in mobile development with Swift, Apple's primary programming language (those frameworks with check marks in the second sub-column of the second column). We aim to use free services whenever possible and as described in the previous section, utilize frameworks that provide functionality with only the sensors germane to the present phase of development.

Consumer Market

Much of the recent innovation (as mentioned in the "RELATED INNOVATION" section) focuses on commercial and defense industries. While we will tap into the technologies available from this market, we plan to scale to the individual consumer. As a small company with an ear to the ground, we understand that successful training starts with trainee engagement. Individual users must view the technology as effective, engaging, and entertaining. By utilizing gaming technology to renovate implementations of longstanding training protocols, we can revive interest in the required work without sacrificing (and in fact, increasing) the benefits of training. Furthermore, a focus on consumer-based training programs significantly expands the available market while lowering the threshold of adoption. Everyone with a smartphone can download an app, and the computational requirements of current phones allow such demanding tasks as using your camera to visualize furniture directly in your space. As of 2021, the consumer market made up 85% of the purchases of AR/VR equipment (Alsop, 2022, *Augmented…2026*).

Table 1. Comparisons of Several Available Features in Mobile Augmented Reality (MAR) Frameworks and Software Development Kits (SDKs).



Framework/SDK	*Platform support*						*Tracking*							*Features*										*Sensors*				*Others*				
							Markers	NFT	Device	Plane	Hand	2D & 3D body	Facial	Point clouds	Anchors	Light estimation	Environment probes	Meshing	Collaboration	Occlusion	Raycasting	Pass-through video	Session management	Camera	LiDAR	IMU	GPS	Architecture	Price	Open source	General	Studio tool
A-Frame (1.2.0) [1]	✓	✓	✓	✓	✓	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✓	✗	✗	✗	Online	Free	✓	✓	✗
ALVAR (0.7.2) [265]	✓	✓	✓	✓	✗	✓	✓	✓	✓	✓	✗	✗	✗	✓	✓	✗	✗	✗	✗	✗	✗	✓	✗	✓	✗	✓	✓	Offline	Free	✗	✓	✗
Amazon Sumerian (N/A) [10]	✓	✓	✗	✓	✓	✗	✓	✓	✓	✓	✗	✗	✗	✗	✓	✗	✓	✓	✗	✗	✓	✓	✗	✓	✗	✗	✗	Online	Free, Paid	✗	✓	✓
ApertusVR (0.9*) [243]	✓	✗	✗	✓	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✓	✗	✗	✗	✓	✓	✓	✓	✗	✗	✗	Offline	Free	✓	✓	✗
ARCore (1.23.0) [82]	✓	✓	✗	✗	✗	✓	✓	✓	✓	✓	✗	✗	✓	✓	✓	✓	✓	✗	✓	✓	✓	✓	✓	✓	✗	✓	✓	Offline	Free	✓	✓	✗
ARKit (4) [16]	✗	✓	✗	✗	✗	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	Offline	Free	✗	✓	✓
ARMedia SDK (2.1.0*) [105]	✓	✓	✗	✓	✗	✓	✓	✓	✓	✓	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✓	✗	✓	✗	✓	✓	Offline	Paid	✗	✓	✗
ARToolKit (5.4*) [120, 134]	✓	✓	✓	✓	✗	✗	✓	✗	✗	✓	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✓	✗	✓	✗	✗	✗	Offline	Free	✓	✓	✗
artoolkitX (1.0.6.1) [212]	✓	✓	✓	✓	✗	✗	✓	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✓	✗	✓	✗	✗	✗	Offline	Free	✓	✓	✗
ArUco (3.1.12) [14, 218]	✓	✗	✓	✓	✗	✗	✓	✗	✓	✓	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✓	✗	✓	✗	✗	✗	Offline	Free	✓	✓	✗
AR.js (3.3.1) [64]	✓	✓	✓	✓	✓	✗	✓	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✓	✗	✓	✗	✗	✓	Online	Free	✓	✓	✗
Augment (4.0.6*) [21]	✓	✓	✗	✗	✓	✗	✓	✓	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✓	✓	✓	✗	✗	✗	Online	Paid	✗	✗	✓
Augmented Pixels (N/A) [22]	✓	✓	✗	✗	✗	✗	✗	✓	✗	✗	✗	✗	✗	✓	✓	✗	✗	✗	✗	✗	✗	✓	✗	✓	✗	✓	✓	Offline	Paid	✗	✓	✗
AugmentedPro (2.4.3) [23]	✓	✗	✗	✓	✗	✗	✓	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✓	✗	✓	✗	✗	✗	Offline	Paid	✗	✓	✓
Banuba (0.35.0) [29]	✓	✓	✗	✗	✓	✓	✗	✓	✗	✗	✗	✗	✓	✗	✗	✓	✗	✓	✗	✓	✗	✓	✗	✓	✗	✓	✗	Offline	Paid	✗	✗	✗
Blippar (N/A) [35]	✓	✓	✗	✗	✓	✗	✓	✓	✓	✓	✗	✗	✓	✓	✗	✗	✗	✗	✗	✗	✗	✓	✓	✓	✗	✓	✓	Online	Paid	✗	✓	✓
CraftAR (5.2.1*) [44]	✓	✓	✗	✗	✗	✓	✓	✗	✓	✗	✗	✗	✗	✗	✓	✗	✗	✗	✗	✓	✗	✓	✓	✓	✗	✗	✗	Both	Free, Paid	✗	✓	✓
DeepAR (2.3.1) [102]	✓	✓	✗	✗	✓	✓	✗	✓	✗	✗	✗	✗	✓	✗	✗	✗	✗	✓	✗	✓	✗	✓	✗	✓	✗	✗	✗	Offline	Free, Paid	✗	✗	✓
EasyAR (4.2.0) [261]	✓	✓	✗	✓	✗	✓	✓	✓	✓	✓	✗	✗	✗	✓	✗	✗	✗	✓	✗	✓	✓	✓	✓	✓	✗	✓	✗	Offline	Free, Paid	✗	✓	✗
HERE SDK (3.17) [94]	✓	✓	✗	✗	✗	✗	✗	✓	✓	✗	✗	✗	✗	✗	✓	✗	✗	✗	✗	✓	✗	✓	✓	✓	✗	✓	✓	Offline	Paid	✗	✗	✗
Kudan AR SDK (1.6.0) [283]	✓	✓	✗	✗	✗	✓	✓	✓	✓	✓	✗	✗	✗	✓	✗	✗	✗	✗	✗	✗	✗	✓	✗	✓	✗	✓	✓	Offline	Free, Paid	✗	✓	✗
Lumin SDK (0.25.0) [152]	✗	✗	✗	✗	✗	✓	✓	✓	✓	✓	✓	✗	✗	✓	✓	✗	✗	✓	✓	✓	✓	✗	✓	✓	✗	✓	✗	Offline	Free	✗	✓	✗
MAXST AR SDK (5.0.3) [160]	✓	✓	✗	✓	✗	✓	✓	✓	✓	✓	✗	✗	✗	✓	✗	✗	✗	✓	✗	✓	✗	✓	✗	✓	✗	✓	✗	Both	Free, Paid	✗	✓	✗
Minsar (2.0) [182]	✓	✓	✗	✗	✓	✗	✓	✓	✓	✓	✗	✗	✗	✗	✓	✓	✗	✗	✗	✓	✗	✓	✓	✓	✗	✓	✓	Both	Free, Paid	✗	✓	✓
MRTK (2.6.0) [166]	✓	✓	✗	✓	✗	✓	✓	✓	✗	✓	✓	✗	✗	✗	✓	✗	✗	✓	✓	✓	✗	✗	✗	✓	✗	✓	✗	Offline	Free	✓	✓	✗
NyARToolkit (5.0.8*) [203]	✓	✗	✗	✓	✗	✓	✓	✓	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✓	✗	✗	✗	✗	✗	Offline	Free	✓	✓	✗
Onirix (N/A) [181]	✓	✓	✗	✗	✗	✓	✓	✓	✗	✓	✗	✗	✗	✓	✓	✓	✗	✗	✗	✗	✓	✓	✓	✓	✗	✓	✓	Online	Paid	✗	✓	✓
Pikkart AR SDK (3.5.8*) [193]	✓	✓	✗	✗	✗	✓	✓	✗	✓	✓	✓	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✓	✓	✓	✓	✓	✓	Both	Free, Paid	✗	✓	✗
PlugXR (1.0.0) [198]	✓	✓	✓	✓	✓	✓	✓	✓	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✓	✓	✓	✗	✗	✗	Online	Paid	✗	✓	✓
Universal AR SDK (N/A) [287]	✓	✓	✓	✓	✓	✓	✓	✓	✗	✓	✗	✗	✓	✗	✓	✗	✗	✗	✗	✗	✗	✓	✗	✓	✗	✗	✗	Online	Free	✗	✓	✗
Vectary (N/A) [258]	✓	✓	✓	✓	✓	✗	✗	✓	✗	✓	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✓	✗	✓	✗	✗	✗	Online	Paid	✗	✗	✓
Vidinoti SDK (N/A) [259]	✓	✓	✗	✗	✗	✗	✓	✗	✓	✓	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✓	✗	✓	✗	✗	✓	Both	Paid	✗	✓	✓
ViewAR SDK (N/A) [260]	✓	✓	✗	✗	✗	✗	✓	✓	✗	✗	✗	✗	✗	✓	✗	✗	✗	✗	✓	✗	✗	✓	✓	✓	✗	✓	✓	Offline	Paid	✗	✓	✗
Vuforia (9.7) [201]	✓	✓	✗	✓	✗	✓	✓	✓	✓	✓	✗	✗	✗	✗	✓	✓	✗	✗	✗	✗	✓	✓	✓	✓	✗	✗	✓	Both	Free, Paid	✗	✓	✗
Wikitude (9.6) [275]	✓	✓	✗	✓	✗	✓	✓	✓	✓	✓	✗	✗	✗	✓	✓	✗	✗	✗	✗	✓	✓	✓	✓	✓	✗	✓	✓	Both	Paid	✗	✓	✓
WebXR (N/A) [30]	✓	✓	✗	✓	✓	✗	✓	✓	✓	✓	❍	❍	❍	❍	✓	✓	✓	✓	✗	✓	✓	✓	✓	✓	❍	❍	❍	Offline	Free	✓	✓	✗
XZIMG (2.0.2*) [284]	✓	✓	✓	✓	✓	✓	✓	✓	✗	✓	✓	✗	✓	✗	✗	✗	✗	✗	✗	✗	✗	✓	✗	✓	✗	✗	✗	Offline	Free, Paid	✗	✓	✗

Frameworks and SDKs marked with an asterisk (*) have not been updated in 1+ years.
Features or functions optionally supported in different platforms are marked with ❍.

METHODOLOGY

Preparation/Prerequisites

Before virtual technology, trainers needed to either physically provide all of the requirements for training for a mission or simply find the minimum requirements to reach the threshold of preparedness. Modernized training reduces the struggles of optimization by quickly providing a plethora of physical and virtual equipment and scenarios. AR focuses on combining the present and absent physical requirements of training by two primary means:

1. Anchoring virtual objects to the scanned environment or
2. Presenting virtual objects on an arbitrary virtual plane so that they always remain within the user's vision.

In general, if virtual objects move within an AR environment, they will do so via option 2 (presumably, for convenience and to keep the user from putting himself in harm's way trying to track the object by turning and moving without sufficient regard to his surroundings). In our case, we want to establish a requirement for an open environment where we both track and provide moving objects with which the trainee must interact (i.e., avoid) in real-time and in sequence using training protocol.

Looking back at Figure 1, we can break down the challenges of MAR training into two categories: Device and Task. Device refers to the independent aspects of the device in use that we cannot control but with which we must interact. Task refers to the ways in which the Device responds to the environment-dependent input that influences the outputs with which the user must interact to achieve the training outcomes. We access the Device via platform-specific SDKs, while we define and influence the Task via platform-compatible frameworks. In combination, the application provides an environmentally interactive training program that follows pre-programmed, job-specific protocols delivered to the trainee via nearly ubiquitous mobile devices as the only true equipment requirement. In this way, training becomes available nearly on-demand without delays in factors such as personnel availability that correspond to increased cost and decreased performance.

Expanding the training environment to include the benefits of movement (i.e., muscle memory, or habit training, for mission tasks) requires little-to-no technological innovation. We hypothesize that lack of present adoption depends on career perspective and job requirement variability. For example, when the NFL began working with virtual reality, companies began with the universally accepted most important position: quarterback. His most recognizable struggle is in reading a defense before and after the snap from within the pocket (a small area with a radius of no more than a few yards) and making decisions based upon what he sees. Without the *need* to recognize and input tangible surroundings, it makes more sense to begin with an entirely virtual simulation. With that said, the short-term timeline gains of ignoring the physical environment restricts its application to positions, assignments, or careers with little required mobility. User movement requires the program to respond and adapt to synchronous events (two inputs occurring at the same time through different sensors). This added complexity is proved possible by user's being able to make hand gestures that impact game play (rather than watching a movie independent of user input), but it introduces more variables.

Most directly, the variable perspective requires the program to attach virtual objects to the environment along a path independent of the user's unpredictable location vector. This path must fit within the environment and begin and end with predictability. To accomplish this, we reference training doctrine and past mission-relevant data. The training event occurs because the program expects that the user will experience several events that occurred in previous or similar situations for which the trainee needs to prepare to perform his best when the results of that interaction will become permanently impactful to his and his peers' careers. Throughout the conception of the training program, creators documented the relative, general placement of obstacles and events, which the trainer follows to set up the course and environment. In developing an augmented-reality version, the programmer must define these (and eventually, create a library of them from which the trainee may choose to practice encountering). To define them in simulation, the career in question must have clearly defined protocol that an outsider can initiate via code. Our company benefits from experience in engineering, combat, education, and athletics that can help interpret less-than-easily translatable doctrine, which reduces the barriers to this new method of training.

Visualization

According to Zollman et al, after interpreting the doctrine into required objects and their location and velocity vectors, three steps follow in what they call the visualization pipeline (Cao et al, 2023):

1. Filtering,

2. Mapping, and
3. Rendering.

Many algorithms work to complete these steps, and they're found within many of the frameworks shown in Table 1. The selection of these depends upon the platform, use of mounted displays, complexity of doctrine and participation, degree of accuracy required, and quality of graphics desired. Additionally, with the added complexity that we discussed in utilizing a multi-perspective, motion-based AR, several factors become crucial to maintaining a simple, effective user experience. Cao et al. list the following crucial factors in delivering an effective user experience:

- Information Filtering and Clustering,
- Depth Cues and Ordering of Objects,
- Illumination Estimation,
- Sensor Error Estimation and Adaptation, and
- Dynamic Content Placement.

Information Filtering and Clustering

When adding virtual objects to the physical plane, the user's view can quickly become cluttered. While volume in optical stimuli can create an enhanced training environment, the ability to execute this with deliberate accuracy ensures a pathway to predefined training objectives. Two methods of regulation include filtration algorithms and clustering methodology. Machine learning regularly uses clustering as a means of categorization and prediction, so those algorithms can help in the use of virtual presentation despite the absence of traditional machine-learning outcomes in the early phases of development (like improved calculations in feedback displayed to a user).

Filtration algorithms utilize one or many spatial or knowledge- or location-based filters. Spatial filters categorize the relevancy of displayable data by distinguishing between what the user's current perspective allows them to see versus what remains momentarily out of sight. In our application, this filtration must account for two sets of location and velocity vectors: those of the user and those of the virtual, mobile objects. Knowledge-based filters utilize subject-matter-expert background information to program the placement and movement of objects. These filters translate the training doctrine into code. Location-based filters accept spatial information from device sensors to provide a user-centric feedback loop that allows virtual training visuals to adapt to training performance in real time. This will become relevant and required in Phase Three of development with the implementation of a virtual "coach" to assist the trainee before, during, and after the completion of the sessions.

Depth Cues and Ordering of Objects

Physical objects naturally fall behind in view other objects between the viewer and themselves. This must happen programmatically for the overlay of virtual objects in physical space—as well as virtual objects in relation to one another depending on the location vector of the viewer and the independent location vectors of the virtual objects. The latter objective depends solely on the knowledge-based filtering, while the former requires at least one or a combination of reading contours of surrounding objects, strict calculations of position and timing, and/or overlying invisible encodings to the view that serve as additional nodes by which the location vectors of objects and plotted and determined.

Geometric features combine with and assist depth mapping of all pixels in an image. Apple's Portrait camera mode automatically calculates depth for pixels to enable its Vision framework to identify three-dimensional poses of people for recognition and categorization. Similar visual cues and calculations allow frameworks to recognize contours to rebuild them on top of the view to superimpose those contours with varying transparency on top of virtual objects to simulate and convey depth. This information added on top of the image creates a feedback loop of positional tracking and object ordering to maintain a realistic dynamic visual that the program requires for effective mobile training. Recent frameworks utilize "X-ray vision" to track clusters of data as they pass "overtop" or "in front of" other imposed data in a manner similar to the spatial filtering that tracks the dynamic positioning of physical and virtual objects while they remain outside of the user's view.

Illumination Estimation

Unlike VR programs, AR programs must anticipate and account for the lighting of the area in which the application will overlay. Virtual objects should match the illumination and contours of their environment, and that illumination affects the depth perception of the trainee and the seamlessness of the order of objects. Many models have been developed for various careers and applications in computer image rendering. In general, these use RGB-D (Red, Blue, Green—Distance) camera data that use either a predetermined, arbitrary illumination reference object or an object in the image with the clearest view that the image processing uses as a reference by which to adjust the illumination of virtual objects.

Sensor Error Estimation and Adaptation

Device sensor tracking continues to improve in accuracy, but variation and error will always remain present for on-board positional data estimation in contrast with fixed, external data tracking. For example, St. Augustine previously worked on a project determining the accuracy of an autonomous navigation quadcopter using an early generation iPhone's positional tracking and an RGB-D camera tracking the changes in location of contrast points such as corners and edges of visually distinct objects. To measure and calculate accuracy, the on-board data was contrasted with a fixed, objective, external motion capture system with multiple camera vantages that used combined depth data to provide "actual" location data. Sensor tracking accuracy noticeably and significantly deteriorated in correlation with a less distinctive visual environment in combination with larger velocity and acceleration vectors. This remains true for a mobile device "mounted" to a person with variable position, velocity, and acceleration. One can confirm this fact by looking at location-tracking data used to find devices (which can only reduce the resolution to an area on the order of square feet) and by comparing the movement data (i.e., steps taken or distance traveled) registered by a mobile device before and after a smart accessory (i.e., smart watch) pairs with the device and tracks the user's movement from direct attachment to the body. Our program's sensor error will depend upon the advancement of the device in use, the contrast of the visual environment, and the later-phase addition of mounted devices used for visual display and positional and motion tracking.

Errors from sensors in registering user, physical, and virtual objects pose a variety of issues ranging from an inaccurate user experience to a dangerous behavioral impact when training for military missions. In anything that humans use that could fail, the severity of the outcome influences the magnitude of the factor of safety required in design. This means that the application must incorporate the risk involved in the training based on the career field and its training doctrine and estimate an allowable error. Additionally, error tracking must occur in real time of use to adapt the progress of the training event to the potential escalation of error (and this risk). The key to estimating and adapting to errors lies in clearly defining the locations of targets for training objectives and recursively monitoring the progress of the trainee in relation to those targets. Then, the program should automatically conclude when the divergence from training objectives or safety levels reaches unacceptable magnitude.

Dynamic Content Placement

Here, we reference both the virtual objects themselves as well as the means of identifying them for the user. This could include a combination of easily recognizable features, floating labels, or visual tracking markers. These will become increasingly involved and valuable with the progress through phased development as a virtual coach joins activity and as multiple individuals begin collaborative or competitive simultaneous use. Before this can take place visually for the user, it must programmatically occur. The application must clearly label pre-programmed objects and data with the foresight to automatically label a large volume of metadata collected once the training begins. Programmatic labels should correspond to visual labels with the understanding that quick, concise legibility will significantly impact the ease of use and safety of the user's intuitive experience. Multiple algorithms exist both within Table 1's frameworks and adoptable into those frameworks that instantaneously process the objects and open space within the field of view to allow and accept labeling that never impedes, and always enhances, the training experience. The major keys to using these algorithms are anticipating the settings in which the training will take place and the perspective changes (i.e., head turning, eye movement) in which the trainee will engage throughout the duration of the training event.

EVALUATION

Beyond using algorithms for real-time evaluation of sensor accuracy and training progression, a new design requires adapting evaluation metrics to determine if the application works well and enjoyably for users and if it enhances the achievement of training goals. Some of the methods for evaluating new protocols will easily apply (i.e., how fast a user completes a course). Other methods become difficult to apply in the enhanced reality world (i.e., other participants in the training evaluating each other). Enhancing training requires an optimization of multiple evaluation methods; otherwise, we're adding technology for the sake of innovation and at best, breaking even in terms of added cost and complexity versus improved outcomes. On top of the evaluation of the relevance of the training delivery system, we need to evaluate the effectiveness of the specific version of implementation (i.e., the seamlessness of the user interaction with this iteration of the product). Cao et al. focus on four evaluation categories:

1. Information Perception,
2. Program Manipulation,
3. Task-Oriented Outcomes, and
4. Other Subjective Metrics.

Information Perception

How comprehensive would an expert observer consider the information presented and topics addressed throughout the training session? Would he consider the information easy to understand—in its volume, order, and delivery? How many repetitions or explanations must take place before the trainee will successfully complete the training session?

Program Manipulation

What physical, mental, time, and space requirements need to be met for someone to use this application? How easily can the least-equipped/lowest-performing trainee use it? What type and level of effort will a trainee need to exert to operate the program in its intended/instructed capacity?

Task-Oriented Outcomes

What efficiency would quantify this training session (time to accomplish goals relative to time to execute tasks)? What efficiency gain or loss comes from using this training platform compared to an entirely physical or entirely virtual program? Would an expert consider the training tasks performed "better" (i.e., better form)? Would he consider them more effectively learned (i.e., better accuracy or precision)? Could one say that trainees perform more of the desired tasks correctly as compared to other training programs?

Other Subjective Metrics

How much would potential trainees like to use this program based on what they know and/or saw about it? Would previous users like to use it again? Will other companies display interest in implementing a repurposed program for themselves after watching the program in action for a different career field? Can the program retain the users complete focus throughout the entirety of the session? Do trainees feel better about themselves or more excited about their jobs after using the application? Is the application generating discussion in environments outside of the training environment? How much control do the user's feel they have over the progression of the session as compared to their traditional training environment?

CHALLENGES

Market

A large volume of companies works (or claims to work) in/with VR. This varies from displaying objects in space to view them three-dimensionally to playing immersive VR games. A growing number of industries continue to adopt the technology. As of 2021, the Global AR, VR, and Mixed Reality (MR) market was valued at $28 billion; by 2028, that number is expected to reach $250 billion (The Insight Partners, 2023). In 2021, the MAR market was valued at $12.5 billion; by 2026, that value is expected to exceed $36 billion (Alsop, 2022, *Mobile…2026*). Small companies struggle to penetrate such an oversaturated market, but utilizing a new approach in mobile training in combination with a focus on the consumer versus the commercial industry provide a unique selling point. For 2024, the projected commercial use cases are estimated to value at $8.2 billion, while the consumer use cases are projected to value at $21.1 billion (Howarth, 2024). Even with an ideal adoption of the concept by a commercial company, our use case fits within the market share with the most space because it only requires a consumer's phone to download and use.

Indirect/Hardware

As a software-focused company, we can remain largely at the whim of the pace and direction of hardware developers. Processing capabilities, graphics, mobility, and communication all impact the consumer experience of consumer and commercial applications, but these remain out of our control. Luckily, we benefit from the rapidly growing interest in the development of this technology by companies like Meta and Apple with interest in constructing micro-universes within an interactive virtual landscape. The pace and quality of their graphics development and seamless device interaction will affect the quality of the experience, but with reliable growth prospects, this becomes an aspect on which we don't need to spend time, energy, and money. Instead, we can focus on identifying new perspectives, expanding training landscapes, and designing intuitively for the consumer and user.

Direct

The Stephenson Technologies Corporation recently submitted a proposal to AFWERX for a concept of a Mixed Augmented Reality Utility training range that would provide enhanced training for careers such as Explosive Ordnance Disposal operators (Battle, 2022). This project's description most closely matches our concept, but the vagueness of the project description merely includes the possibility of aligning with our application idea. Its emphasis on a "training range" lends to the concept by not the implementation. Beyond that, any of the previously mentioned companies working in AR or VR could branch out to include mobile virtual objects onto an expanded training space with a mobile trainee. The challenge for all of these companies remains the same challenges that face ours; although their experience in programming a virtual environment benefits them, their focus on fixed user locations or complete immersion inhibits their perspective on the applications of mobility. Many careers need only a restricted area of operation (i.e., a pilot in a cockpit), and those need significant development in graphics and expert procedural input to reach maximum engagement and efficiency. The main reasons for divergence from that path lie with the needs of the consumer—in contrast to the needs of the commercial industry. Traditionally, training for a career requires taking expensive courses or joining a corporation with resources to spare trainers in addition to workers. Supplying these resources independently empowers the consumer and employee and improves the value of work completed. This further lowers the threshold for training requirements and thus the financial and time burden on employers.

REFERENCES

Alsop, T (2022). *Augmented Reality (AR) and Virtual Reality (VR) Headset Shipments Worldwide from 2021 to 2026, by Market*. Statista. https://www.statista.com/statistics/1301629/ar-vr-headset-shipments-by-market/#:~:text=In%202021%2C%209.54%20million%20augmented,recorded%20in%20the%20commercial%20market

Alsop, T (2022). *Mobile Augmented Reality (AR) Market Revenue Worldwide from 2021 to 2026*. Statista. https://www.statista.com/statistics/282453/mobile-augmented-reality-market-size/

ARHT Suite. ARHT. https://www.arht.tech/products/

Battle, J (2022). *Mixed Augmented Reality Utility (MARU)*. AFWERX Challenge. https://afwerxchallenge.com/operationaleffectiveness/advanced-training-concepts-challenge/solution-submission-phase-478/suggestion/13710

Cao et al (2023). “Mobile Augmented Reality: User Interfaces, Frameworks, and Intelligence”. *ACM Computing Surveys, Vol. 55, No. 9, Article 189,* pgs 1-36.

Clement et al (2020). “Live-Virtual-Constructive Demonstration Proves to Be a Game Changer”. *AFRL Fight’s On!*, *Vol. 37*, pgs 1-4.

Holden, C (2014). “The Local Games Lab ABQ: Homegrown Augmented Reality”. *TechTrends Vol. 58*, pgs 42-48.

Howarth, J (2024). *24+ Augmented Reality Stats (2024-2028)*. Exploding Topics. https://explodingtopics.com/blog/augmented-reality-stats

The Insight Partners (2023). *Augmented Reality and Virtual Reality Market Growth to Hit 36.9% CAGR, Globally, by 2028 – Exclusive Report by the Insight Partners*. GlobeNewswire by notified. https://www.globenewswire.com/en/news-release/2023/11/01/2771367/0/en/Augmented-Reality-and-Virtual-Reality-Market-Growth-to-Hit-36-9-CAGR-Globally-by-2028-Exclusive-Report-by-The-Insight-Partners.html

Products. VAR. https://virtualcoachfootball.com/products/

Tanier, M (2016). *Future of the NFL: The Virtual, Augmented, 3D, 360-Degree Football Experience*. Bleacher Report. https://bleacherreport.com/articles/2659861-future-of-the-nfl-the-virtual-augmented-3d-360-degree-football-experience