Xtended Reality – Applying Immersive Technologies
(When it's right, when it's wrong, which is right, which is wrong)
Knowledge Extraction

Craig Clark
Chief Technology Officer (CTO)
Director of Engineering
Kratos Defense
Orlando, Florida
craig.clark@kratosdefense.com

ABSTRACT

The age of fully immersive reality is upon us. We have technology that provides us a completely 'Artificial' experience, and technology that can blend our current reality with digital reality, also referred to XR technologies. The Reality-Virtuality Continuum (ref 1) is real and is part of the modeling, simulation, and training (MS&T) landscape. The reality of immersive technology is that it is not a one size fits all solution. In fact, the use of anything but full reality in some training systems could be considered negative training. This document is designed to help define when and which technology should be considered across the spectrum of training systems which take out some of the mysteries of what is meant by the Reality-Virtuality Continuum and what it means to providing best solutions for the MS&T community.

To fully realize the training potential of XR technology, several aspects must be considered. The technologies that enable virtual and real-world interaction across the continuum spectrum will be discussed in this study. It will also outline the technology that offers the best fidelity, cost, environmental impact, and user-friendliness trade-offs.

The available technologies that enable physical and virtual immersion differ in their approaches as well as the number of sensors required, latency, logistics, cost, and other factors. Edge detection, object recognition, object tracking, depth sensing, and chroma key technologies are the main means through which the virtual and real worlds can communicate with one another. Each of these technologies has advantages and disadvantages that could influence the final choice to satisfy the demands of warfighters.

This study will go over all the current developments in MR technology as well as business and governmental strategies for maximizing its value for warfighters. There are a few limitations to be aware of when it comes to edge detection, object recognition, or depth sensing. Edge detection locates numerous edges to match a known physical position, whereas object recognition detects an object and its orientation. Choosing how to display mixed reality (MR) has a higher overhead in both of the prior solutions. To select how to show MR, depth sensing can make use of cameras or sensors that are similar to lidar; nevertheless, it might not be effective for larger solutions where depth turns into a very dynamic variable. Chroma key technology offers extremely quick rendering capabilities but also has more space requirements and illumination issues.

ABOUT THE AUTHOR

Mr. Craig Clark is the Chief Technology Officer (CTO) and Director of Engineering at Kratos Defense. He is the head of Kratos' Research and Development (R&D) for all of Kratos training and simulation (KTS). Over the past 7 years, Craig has directed business unit technologies from product operation improvements to research and development. Prior to being the CTO, Craig was the senior software architect for Kratos and designed the training application framework for the modeling and simulation baseline for KTS. He holds a BS in computer science and an MSEM from the University of Central Florida. Craig started his career in the Naval Nuclear Propulsion Program, U.S. Navy. He was a naval instructor and has spent plenty of time underway running propulsion systems. He worked in telecommunications until 9/11 and subsequently changed his career to defense and training. He became a principal architect for General Dynamics and designed urban operation systems which has trained over a million U.S. Army soldiers.

Xtended Reality (XR) – Applying Immersive Technologies (When it's right, when it's wrong, which is right, which is wrong)

Knowledge Extraction

Craig Clark
Chief Technology Officer (CTO)
Director of Engineering
Kratos Defense
Orlando, Florida
craig.clark@kratosdefense.com

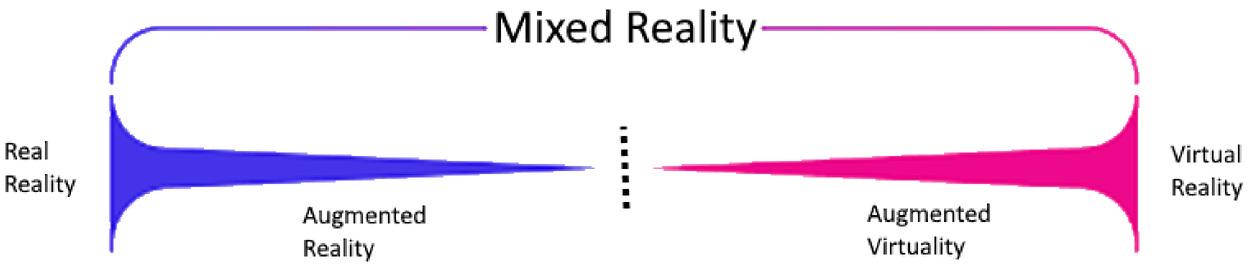
INTRODUCTION

Finding the best Extended Reality (XR) application might be difficult when trying to meet the needs of the warfighter. Fully immersive virtual reality (VR), also known as augmented reality (AR), and augmented reality (AR), where content is painted over direct view optics, are frequently not the best solutions for meeting training demands. MR, which combines virtual and real-world content, can be a potent tool for helping warfighters acquire, experience, and understand the precise objectives they need to be proficient. It can be difficult for both technology professionals and clients to grasp the underlying technologies required to create MR solutions to know how to implement MR in the many use cases.

Using the right MR technology depends entirely on the requirements. Based on an accurate grasp of the current technology, these requirements will undoubtedly alter. The movement of real-world and virtual items in the scene, as well as fidelity, cost, footprint, portability, and human usability issues, are all important considerations. The optimum usage of MR today will depend on knowing which technologies are available and which one to adopt. There is no one size fits all approach for MR implementations.

VIRTUALITY CONTINUUM, IN A NUTSHELL

In the Virtuality Continuum spectrum, the scale goes between two realities: the physical real world and the virtual world. The physical real-world training needs little explanation. This would be typically called on the job training or using physical training aids. The pure virtual training is a pure computer simulation and more recently fully immersive VR head mounted display (HMD). There is training technologies that provide solutions across the entire spectrum.



Pros/cons of real-world training

Pro:
With few exceptions for most things real world training is the most accurate and best for muscle memory development. It is the reality component that may be the only thing that can produce the tactile feedback or experience of getting an unpredictable result of doing something never done before. There is no more accurate way to portray the real world than with the real world (obvious). The real thing is usually considered the best way to teach and

usually gives the trainee several of the intangible factors that the real thing provides. Intangible factors including several unconscious factors, gut feelings, and even biological chemical factors that are not duplicatable with technology. These factors, based on the training, can be the most important factors for getting the most accurate training.

Cons:

Some of the biggest challenges with real world, also known as grounded world, training is that it can be very expensive, highly dangerous, or nearly impossible for a vast number of training needs. Costs in fuel, bullets, and human life can be limiting to the training task. If the job is to educate on how to clean up a biological or chemical weapon it would make sense that it is better to abstract out the actual poison either by a placebo or producing a computer-generated model of such a disaster. Usually sending untrained personnel into any high-risk situation is not good “On the Job” training.

Pros/cons of Pure VR training

Pro:

Without leaving the safety of a schoolhouse, virtual reality enables soldiers to train across the globe and even into space. Military schools all across the world are already utilizing this power to enhance lesson ideas. Additionally, it enables individuals to have worldly experiences in ways they never could previously. You can fight in the streets of Paris, swim at test depth with a submarine, and fly on the wing of an F-35. It has unlocked a universe of limitless opportunities and possibilities. Virtual reality gives us the ability for rewards without risk.

Cons:

A major drawback to VR is that it lessens human contact and may detract from value in human interaction. Virtual reality for extended periods of time can cause people to lose sight of their real life, the real world, true real-world behavior. There are consequences for putting your hand into a fire or shooting an unarmed civilian. It can, especially if not done with high fidelity, be negative training. This could cause improper confidence when caution and attention to detail could mean life or death.

Additionally, it may have certain detrimental bodily side effects. A major benefit of virtual reality is that it lets you engage in exciting, dangerous activities while being safe. On the other side, several users have complained of sickness as a result of VR headgear fooling their senses. People have also complained of eyestrain as a result of the numerous visual stimuli you frequently encounter in virtual reality. Computer Vision Syndrome can be a byproduct of the pervasive use of computers.

Mixing of Real and Virtual reality or Mixed Reality

The space between real reality and virtual reality is identified as mixed reality. It extends from Augmented reality to Augmented virtuality. Augmented reality can be thought of as mostly real but some computer interaction with the student. A passive lens (eyeglasses) that lets you see reality as it is with the naked eye and putting computer labels or images onto the lens can give the student some feedback to what they are seeing or basic interaction with the outside world. On the other side of the spectrum the training may be mostly virtual but gives some tactile feedback like vibration or olfactory (Smell) to allow your real senses pick up on training queues.

MIXED REALITY, THE BEST OF BOTH WOLRDS

Numerous war fighters and end users have found that the usage of MR in a training setting is the most valuable experience offered outside of genuine "on-the-job" training. It enables the end user to hold actual tactile feedback-producing physical devices in their hands. In many cases, it is hard to reproduce this input in a virtual or simulated environment. The use of a complex weapon is a prime illustration. Only using the original device or an exact duplicate will allow you to develop muscle memory for the precise location of safety features, pinch points and how to "clear" a weapon in complete darkness. The next time the device is picked up, it should be easy to remember if your hand is close to the weapon's pinch point because it gives you tactile input (feeling) to remind you of this.

BENEFITS TO MR

Being able to observe both the real world and the virtual world simultaneously can greatly enhance training. The ability to physically touch genuine items that provide tactile feedback to the user is one of MR's major advantages. Because many people do not suffer motion sickness or may have latent (delayed) motion sickness due to having real-world correlation, staying longer during the training is another advantage. The sense of mixing the virtual and real worlds can provide a more natural experience and behavior benefit beyond full immersion technology, even though it is not considered to be a fully immersive system.

The trainee's ability to interpret MR accurately is heavily influenced by latency. The adoption and usability of an MR system depend heavily on the quality of the experience (QoE). The latency, measured in milliseconds for the eye and brain, must be below 33ms to have the least negative impact on the goals of minimized motion sickness and greatest immersive experience. This means that each component, including the tracking system, camera system, and game engine or image generator (IG), must operate inside the 33ms time limit.

APPROACHES TO THE MR EXPERIENCE

Through the use of both real-world and computer-generated virtual aspects, mixed reality can improve an experience (in this example, training). Although primarily referring to the visual aspect, the other senses—tactile, auditory, and olfactory—also contribute to the realization of the whole experience. Two elements, tracking technology and visual blending technology, and how they interact, are essential for successfully fusing the two realities. The majority of MR experiences rely on a head-mounted device (HMD). Visual blending technology enables elements from the two realities to selectively participate in the mix and be presented to the user's eyes. Tracking technology enables the two realities (physical and virtual) to be brought in the same coordinate space.

Video Blending technologies:

MR Video Blending technologies can be broken into 2 primary technologies

- • Optical See Through (passive) - During the passive experience, your natural eye can look through a specific lens or lenses into the outside world. The lens(es) is equipped with technologies that offer a reflective surface for displaying virtual visual data (example: Hololens). It's vital to remember that the real world is always present in passive MR scenes; it can only be partially obscured.
- • Video See Through (VST) - VST shows a camera view integrated into the HMD by using high-resolution video cameras to capture the real environment, which is then digitally combined with computer generated information in the GPU. One illustration is the Varjo XR-3.

Tracking Technologies:

The technology integrator might select either of the two main tracking approaches or a combination of the two. The two methods are as follows:

- **Outside-in tracking:** this type of tracking places tracking components on the operator's head-mounted display or the tracked items while the tracking system is located outside the operator, or the thing being tracked. The HMD has a complementary sensor that enables an electronic system to calculate the position and orientation of the HMD with respect to the fixed references in the physical world. The technologies in this category include Magnetic Tracking, Global Position System (GPS), Infrared (IR) Tracking, Ultrasonic Tracking, and Camera Tracking.
- **Inside Out tracking:** Without the use of active external devices, the operator's HMD can determine its location and orientation in the real world. Additionally, depending on the position and direction of the HMD, the HMD may be able to detect additional physical things (controllers or user objects).

These two MR kinds certainly have different latency profiles. Real-world eye and mind latency in passive visual MR is what we refer to as a baseline or zero latency. The virtual items that appear in passive visualizations exhibit delay and can seem to be floating or swimming virtual elements. Depending on the fidelity or items viewed in the HMD, the influence on latency with passive HMD experience can change. By using artificial intelligence in the HMD to predictably show virtual items while considering the user and the system's use case, perceived latency can be reduced. Passive HMDs may cause the least amount of motion sickness because the movements of the human body and the real environment are most similar.

TECHNOLOGIES AND MIXED REALITY

The creation of an MR application is challenging and depends on several technologies for different use cases. The relationship between a number of these technologies and properly developed MR systems is discussed below. Although there are many articles on MR technologies, MS&T community is working on the correct application for the technologies The following technologies will aid in the practical adoption of MR applications by business and warfighters.

TECHNICS IN MR Technologies:

Being able to combine the real world with the physical world is based on sensors identifying the physical environment that allows some type of correlation with the virtual environment. Each technic uses a variation on sensors that allows the following:

- Object recognition
- Tracking entities
- Edge detection
- Depth sensing
- Color Keying

- **Object Recognition: Marker MR**

When the system recognizes a physical object in the surroundings and follows the object to determine how to paint the virtual picture, this is called object recognition tracking for MR. This type of mixing can be used as inclusive or exclusive to the virtual scene. If the object needs to be seen in the virtual it can allow the image of the physical object to be visible through filtering the object to be seen. If it is object does not want to be seen (fiducial marking) it is hidden behind the virtual or painted over. Typically, object recognition is carried out during video capture (OpenCV) and is determined by video recognition algorithms.

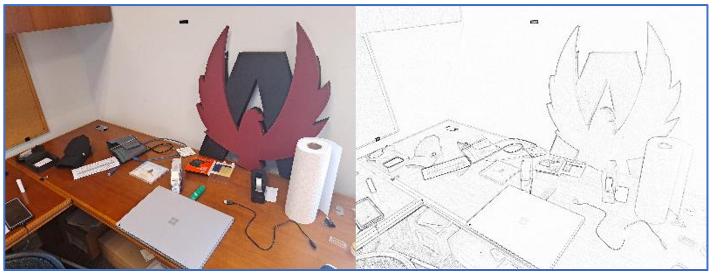
A functional use case for this type of MR technology is if a student is in a classroom environment or maintenance trainer when specific objects can be in the view of the student. Usually, the system is calibrated to accept very specific objects including QR codes, fiducials, or other deterministic objects that can be easily identified in video capture algorithms. The more unique an object is, the better the results of visual detection, latency, and system accuracy.
In a maintenance trainer with several items of interest, each component can have a QR sticker or visual marker (i.e.; aircrew compartment of aircraft) as a visual cue to discuss components or where to paint the virtual world or technical manual. This technique works well in large spaces that require the operator to move around in a dynamic work environment.

- **Tracking Entities: Marker-less MR**

When an MR is used by the system based on reference of static objects, this is known as marker-less tracking for MR. Typically, a beacon or puck that provides the location and position of the actual thing is used as relative location for the simulation. An object's orientation and location can be known, which enables a device to have properties based on the data. These properties can be used to determine where to position the virtualized scene based on the location of an object or to allow placement of occlusion models that either include or exclude the object in a virtual environment. A visual at the end of the device, such as a firehose spraying simulated water on a fire or a bullet cartridge emerging from a rifle, would be one use case. The real rifle appears passive or digital to the eye, but when the trigger is pressed, only the virtual version of the bullet is shown. The use of cameras, LEDs, IMUs, magnetics, or IR reflecting balls can all be used for marker-less tracking. This uses other sensor systems that do not mix based on object shape or marker.
A functional use case for this type of MR solution is if a student is in a large room environment or operational trainer where head tracking and a device (i.e., weapon) both need to be tracked for visual orientation and device simulation (flashlight or weapon). Usually, the system is calibrated at the beginning of an exercise by placing an object with a known location for a datum point. This type of technology can be difficult to configure but calibration is relatively easy with a proper process for the user. This type of tracking is very fast and accurate and can be found

in several visual systems preconfigured for location. It works well in large spaces that require the operator to have room to move from place to place in a dynamic work environment. The use of inertia motion tracking (IMU), optical tracking, and magnetic tracking are common technologies used for this MR technical approach.

- **Edge Detection:**



The use of edge detection tracking is described as finding contrasts in a visual system to identify edges of objects for location and spatial constraints of the surroundings. This provides a visual system to map the area for algorithms to determine distance and placement of virtual objects. This is primarily useful when all the physical world components stay relatively static, multiple physical areas are to be used, and the spaces cannot be altered by logistic tracking devices (markers, color key, objects).

This kind of MR technology can be used effectively in a classroom or multi-station setting, such as a maintenance trainer. When space and environmental adjustment are not required, it can be employed in a genuine tactical device like a helicopter or ground vehicle. Several MR HMDs, like the Hololens, utilize it as one of their main techniques. It is highly helpful for device location, overlay, and component recognition solutions since it transitions very effectively from different positions within a space.

Without substantial training or complex documentation, this kind of equipment is quite simple to calibrate. It does take time to incorporate desired scenarios, depending on the setting. Although configuration can be time-consuming, once finished, it is static. Although calibration is simple, it needs to be done in a static setting. The drawback of this type of technology is that it functions best in static settings. The system's visuals can be either low-fidelity or moderate-fidelity. If the environment's lighting conditions vary, it could be very difficult. As a result, the tracking system's quality may alter according to how the contrasts and shadows affect edge detection sensing. Sometimes the images can appear to move away from a fixed point; this is usually caused by latency in the rendering of the images. When there is a lot of movement or the speed of the movement increases, this issue is doubled. Errors in the edge detection computations in the physical world are the main cause of this problem.

- **Depth (Spatial-Based Tracking):**

Where the real world ends and the virtual one begins can be determined by using depth or spatial-based tracking. This is most helpful when every element of the physical world is nearby the system's user and static. The optimum application of this technology typically occurs in immediate regions that are just out of the user's grasp. To precisely position and put the virtual world, the depth tracking system can make use of a variety of sensors, such as lidar.

If a student simply needs to see real, tactile objects at arm's length in a classroom setting, this type of MR technology could be useful. For instance, when a user wants to be able to see their hands on the gearshift and steering wheel. The user will be able to see exactly where their hands are on the steering wheel, and the required tactile cues will be provided for the pupil. The user may look down and see the shifter and put their hands on it to control the device. The gear shift is at the exact same location as the real automobile. The driver would be able to operate the car in a virtual environment because everything past the steering wheel and gear shift would be virtual. Without considerable training or complex documentation, this type of technology is quite simple to adjust for the user. Setting it up doesn't involve a lot of logistics, and most beginners to advanced system users can handle it. With clear instructions, it should be straightforward to set up in minutes.

This kind of technology presents a hurdle in that it prevents complete mental involvement. Even though they are of excellent quality, images may have problems, including tearing when the virtual environment meets the real world. This is especially true of digital viewing HMDs. Depth tracking errors and accurate contour data in the processing of virtual images are additional determinants of accuracy in the placement of the virtual scene. When there is a lot of movement or the speed of the movement increases, this issue is doubled. Errors in the depth estimations near the edge of the physical world are the main cause of this difficulty. Due to the latency of a whole visual picture or frame, the speed of rendering the virtual content is also somewhat slow.

- **Color keying: (chromakey)**

The real-world objects and virtual objects can be distinguished clearly with the use of color keying technology for MR. A secondary tracking system is always going to be necessary with this kind of MR. This implies that in order to complete the MR experience, independent additional tracking systems like those mentioned above will need to be used. Typically, this would only include watching how the trainer's participants' heads moved. Color keying technology has been available for a while and is frequently utilized in a variety of ways to achieve live or post-production effects. Hollywood productions frequently provide the idea that their most important cast members are placed in danger while nevertheless running no risk of harm. It can offer a high-quality hack or cheat to placing students in a virtual environment in training systems. Only digital MR, not passive MR, is capable of using color-key. This means that color-keying technologies won't be able to be employed with passive HMDs like the Hololens.

Color-key works by placing the virtual environment according to a particular spectrum of light (color). A precise algorithm that can produce a seamless visual transition between the real and virtual scene is the replacement of color pixels with virtual pixels. This makes it possible for a clear visual environment, increasing a system's capacity for mental immersion. Often, the user is unable to distinguish between the real world and the virtual world because their minds start to combine the two into a single image. This technology also offers quick and straightforward algorithms, which is a bonus. The algorithm's processing speed greatly minimizes latency. It is even possible to implement color keying algorithms in hardware chips rather than software. This can also significantly lower latency to the point where the only bottleneck is the speed at which the monitor renders images.

It is possible to use the color key materials in both small classroom settings and much bigger spaces. The fundamentals of how the system immerses the participant include the direction the student is looking in and the use of color key technology. When a device (such a weapon) needs to be tracked for visual orientation and device simulation, it can be employed in highly dynamic operational training. For the mind to believe that the physical and virtual worlds are in the same space, the sharp edge of the physical world against the background of color creates a highly realistic, reasonable divide.

Depending on how color keying is used, the drawbacks of a color key MR solution can be very different. One of the biggest obstacles for this kind of MR is the amount of work required to integrate a color key background into an environment. The area required for color keying may require a substantially greater footprint for large-dynamic operating trainers than for any other MR system.

The system's lighting has a significant impact on how the technology works. The effectiveness of the color keying algorithm might be significantly impacted by erratic or subpar illumination. Grainy visual effects may result from hot (bright) or cold (dark) lighting spots, which can leave holes in the virtual scene. The technology's operation may be affected by shadows. A blockage between the light source and the color screen will result in shadows, cold patches, and a reduction in the MR quality. The room's materials should be matte finished to prevent reflections of the light or other objects in the space. This might result in lighting hotspots that also leave holes in the virtual world. Additionally, lighting reduces the MR's viewing area and heats the environment in which it is located. To allow for color reflection, the lights must be present in the room but cannot be a part of the MR scene. Depending on the illumination used, the amount of setup can change significantly.

The best lighting for color key backgrounds is even lighting. This uniform illumination requires thoughtful placement in the training space. Lighting fixtures take up room, produce heat, and need the training space to be cooled. This hardware might be difficult to install and quite heavy for the best lighting. The camera of the MR HMD receives color reflections from a color screen, carpet, or sheet.

Utilized color key material reflects source light back to the HMD cameras. A wide spectrum of light is produced by the colors and saturation of reflected light. Some of the white light is reflected as various hues and saturations of the surface it is reflecting on when utilized with a color drop background. The spectrum of light reflected broadens as a result of this feature. The spectrum expands with increasing light intensity. The difficulty with a larger spectrum of light is that it also widens the algorithm for color keying of the particular light spectrum.

A larger spectrum of light presents difficulties since it enables objects to be keyed out of the scene when they are not intended to be. If green is utilized as the keying color and the trainees are dressed in green, this is clear. The students start to resemble half-erased people or people who are both virtual and real. More objects with similar colors can be added to the room by narrowing the color range used in the algorithm.

Electroluminescent (EL) illumination can be used to solve most lighting problems in MR. EL lighting produces light internally by passing an electric current through phosphorus media. The non-thermal transformation of electrical energy into light energy is known as EL. Cool, continuous, low-power fixtures that emit a gentle light without any glare are referred as EL lighting. When compared to other illumination sources, it produces a relatively narrow light spectrum, which creates the ideal condition for MR color isolation.

Since EL lighting is built into the trainer's walls, it practically occupies no space. It is water resistant, as thin as wallpaper, and can be produced in a wide range of forms and sizes. Since EL lighting produces little to no heat, the area doesn't need as much additional cooling. Low-wattage and long-lasting EL devices are available. Other advantages of EL lighting include adaptability, excellent visibility in the dark, and ease of viewing without eye strain. Since EL lighting is created without harmful ingredients, it is also landfill-friendly.

- **Compositing Technologies**

To select what elements the user sees the blending technologies generate a mask that occlude one object or parts of it to reveal elements from the layer behind. These masks can be fixed relative to the physical world or attached to physical objects and move and rotate with them. The masks can be generated using three methodologies:

- Static Mask – this mask is defined and calibrated in the physical world and provides a portal to blend the physical. Think about sitting into a cockpit that is physical, but everything outside is virtual.
- Depth Mask – To distinguish between the real and the physical, depth information on the physical world is collected. Typically, a threshold value is used, and anything past that point turns virtual.
- Chroma Key Mask – a specific color in the real environment (usually green or blue) is replaced with pixels from the virtual environment.
- Dynamic Mask – This mask is used to paint a virtual item over a physical object that is monitored in 3D space, for as showing a virtual image over an articulate display that is actually a cardboard placeholder or exposing a genuine display over a virtual background.

CONCLUSION

In all use cases of the reality-virtuality continuum, the developer and the end user want the best of all the solutions The best training technology is high fidelity, simple and easy to use, inexpensive, and overall provides training that adds value to the warfighter’s abilities. These needs cannot be all be met by a single solution. There is currently no solution that provides a higher level of fidelity than real world experience. A balanced combination of real and virtual technology (MR) including high quality displays, realistic visual game engines, and highly accurate physics models provide the best alternative to the risks and costs of real-world training.

For many training solutions high fidelity MR becomes impractical when solutions do not have to meet near real experiences. Pure VR works well if a user only needs to learn basic location familiarity for a cockpit or device locations. Passive visual systems like the Hololens are much better suited to perform in areas that need simple tutorial walk through of tactical systems. The Hololens is a great option if the main visualization area must remain in the physical world or if only straightforward icon overlays are required. Using depth sensing technology to render the virtual world when providing a window of real-world visuals is a great use of depth MR technology. Depth is typically integrated directly into the HMD, like the Varjo XR-3 system, that allows for simple set up and very good visual acuity and speed where needed. Compared to a color key solution, there is a significant reduction in cost, logistics, and simplicity

All prospective system users should have a basic understanding of their goals, as well as what constitutes an acceptable system's pricing and capabilities. It will be simpler for both industry and procurement teams to choose solutions that deliver the best value for our warfighter if they are aware of these crucial variables when choosing a system or building a system.

REFERENCES

1. Milgram, Paul; H. Takemura; A. Utsumi; F. Kishino (1994). "Augmented Reality: A class of displays on the reality-virtuality continuum". *Proceedings of SPIE - The International Society for Optical Engineering Vol. 2351*. Retrieved 2021-06-01.
2. CAICT. 2017. Virtual Reality/Augmented Reality White Paper (2017). Retrieved from https://www.eng.it/en/white-papers/ar-mr-vr
3. MDPI. 2019. A Review on Mixed Reality: Current Trends, Challenges and Prospects, Retrieved from https://www.mdpi.com/2076-3417/11/5/2417