Operationalizing Persistent Augmented and Virtual Environments in Naval Aviation Maintenance

Maj Michael Ashmore	**Jake Ramirez, Kalvin Lam**	**Dr. Amela Sadagic**
PMTRASYS, U.S. Marine Corps	**NIWC Pacific**	**Naval Postgraduate School**
Orlando, FL	**San Diego, CA**	**Monterey, CA**
michael.ashmore1@usmc.mil	**jake.p.ramirez.civ@us.navy.mil,**	**asadagic@nps.edu**
	kalvin.k.lam.civ@us.navy.mil	

ABSTRACT

The demand for more efficient approaches in aviation maintenance, repair, and overhaul (MRO) for mission readiness and automation of life-cycle sustainment requires the introduction of novel approaches. Those solutions will be even more significant if they address the operational needs of individuals and provide tangible enterprise-level benefits. Corrosion detection is just one example of the maintenance tasks that, if not attended to in a systematic and timely manner, can result in heightened repair costs and, ultimately, lead to diminished operational readiness. We present a novel framework for the maintenance domain that utilizes augmented reality technology, the persistent augmented and virtual environment - PAVE, and digital twin concepts. As a proof of concept, we designed and developed a prototype system, the F/A-18 Tracking Corrosion and Assessment Tool (FATCAT), which integrated highlighted technological approaches. The resulting system was tested in a formal user study involving 32 active-duty USN F/A-18 Super Hornet aircraft maintainers who represented day and night crew work shifts. The participants used the FATCAT tool in the context of visual inspections and aircraft corrosion control, performing multiple tasks typical of their job roles on real aircraft. The results demonstrated that a strategic combination of novel technological approaches, together with intuitive user interaction metaphors, superior data visualization, and data capture, can support multiple stakeholders across the aviation MRO domain. The objective and subjective data sets reflected participants' ability to use the tool and complete the tasks after a very short familiarization with the tool's functionalities. The use of the tool resulted in high satisfaction levels with the system and a desire to use it in their daily practice. A set of concepts and technologies used in the FATCAT is applicable to preventive and corrective maintenance tasks in aviation and any other domains that require regular inspections and repairs.

ABOUT THE AUTHORS

Major Michael Ashmore is a Project Officer for PMTRASYS in Orlando, Florida within Marine Corps Systems Command. Before joining Team Orlando, Major Ashmore earned a Masters of Science in Modeling and Simulation from the Naval Postgraduate School MOVES institute and served in billets ashore and afloat as an aircraft maintenance officer with MALS-11, the Snakes of VMFA-323, and Team Quicksand aboard the USS Nimitz with CVW-17.

Jake Ramirez is a software engineer and data scientist out of San Diego, California. He earned a Bachelor of Science in Computer Science from Point Loma Nazarene University in 2020. Jake began his career with the Naval Information Warfare Center Pacific in 2018 and continues to support NIWC Pacific in the Reverse Engineering, Science, and Technology for Obsolescence, Restoration, and Evaluation (RESTORE) Laboratory.

Kalvin Lam is a software developer currently supporting the RESTORE Laboratory. He graduated from the University of California San Diego in 2020 with a Bachelor of Science in Computer Science and immediately began supporting the U.S Navy at NIWC Pacific located in San Diego, California. Passionate about bringing modern, user-friendly design to the fleet, Kalvin lends his experience in agile methodologies to develop applications with an emphasis on streamlining the user experience without compromising functionality.

Dr. Amela Sadagic is a Research Associate Professor and a Co-director of the Center for Advanced Manufacturing and the Consortium for Advanced Manufacturing Research and Education (CAMRE) at the Naval Postgraduate School (NPS) in Monterey, CA. Her research interests include virtual and augmented reality, human factors, training systems, diffusion and large-scale adoption of technology. Dr. Sadagic holds a Ph.D. degree in Computer Science from University College London, UK.

Operationalizing Persistent Augmented and Virtual Environments in Naval Aviation Maintenance

Maj Michael Ashmore
PMTRASYS, U.S. Marine Corps
Orlando, FL
michael.ashmore1@usmc.mil

Jake Ramirez, Kalvin Lam
NIWC Pacific
San Diego, CA
jake.p.ramirez.civ@us.navy.mil,
kalvin.k.lam.civ@us.navy.mil

Dr. Amela Sadagic
Naval Postgraduate School
Monterey, CA
asadagic@nps.edu

INTRODUCTION

Naval aviation is a critical component of power projection; it enables maritime services to operate effectively within the air domain from naval ships within the contested battle space. The sustainment of these systems requires extensive resources for mission readiness, life cycle sustainment, and consistent system upgrades. Those operational needs generate a demand for more efficient approaches in aviation maintenance, repair, and overhaul (MRO), as well as ensuring mission readiness and automation of the life-cycle sustainment. While current solutions can be further perfected, to achieve substantial improvements, gain novel capabilities, and realize tangible enterprise-level benefits, it will be necessary to introduce innovative solutions. We present a novel framework for the maintenance domain that utilizes augmented reality technology, persistent augmented and virtual environment (PAVE), and digital twin concepts. As a proof-of-concept, we designed and developed a prototype system, the F/A-18 Tracking Corrosion and Assessment Tool (FATCAT), that integrated highlighted technological approaches. This research effort addresses two research questions: (1) What is the feasibility of implementing a persistent augmented reality application that enables aircraft maintainers to execute corrosion control tasks more efficiently?, and (2) What is the usability of the FATCAT augmented reality tool that supports aircraft corrosion control?

The main contributions of this paper are: (1) The work represents a novel solution derived from studying the challenges of aircraft maintenance; it addresses the needs identified in task analysis and known to the community, (2) Proposal of a novel framework for the aircraft maintenance that employs PAVE concepts, (3) The design and implementation of a proof-of-concept prototype persistent augmented reality application that supports the visual inspection phase of aircraft corrosion maintenance, and (4) Formal user study that examined usability of the FATCAT tool.

BACKGROUND

Persistent Virtual Environments

The concept of persistent environments, the environments that are 'always up and running', has existed in the Virtual Reality (VR) domain for a long time. A typical example is a virtual environment (VE) that enables individuals worldwide to connect in a cohesive, shared virtual environment, allowing them to meet, interact, and communicate in real-time. Even when a particular user is not active (logged out), the persistent environment will continue to be inhabited and used by other users.

The work on persistent VEs has been inspired by the multi-user dungeons (MUDS) that originated in the late 1970s (Kelly & Rheingold, 1993). These text-based, real-time environments allow their users to enter their virtual world, interact with other users, and cease the virtual session by leaving the environment at will. A crucial building block for persistent VEs was a specialized network solution—multi-user networked virtual worlds—capable of connecting multiple individuals in a coherent 3D environment. The availability of high-bandwidth, wide-area communications and the emergence of World Wide Web applications provided a crucial impetus for VR researchers to expand their reach beyond local area networks, leverage these new technologies, and develop systems that would enable the networking of multiple virtual worlds, allowing users from all over the world to interact with environments and with each other in real time.

An early effort in that direction was undertaken by Zyda and his team in the Graphics and Video Laboratory of the Department of Computer Science at the Naval Postgraduate School, who created NPSNET (Zyda et al., 1992;

Macedonia et al., 1995). The NPSNET utilized the Multicast Backbone (MBONE) (Eriksson, 1994) and was defined as a multiplayer 3D virtual environment over the internet, designed to serve a medium to large number of simultaneous users (in the authors' words, "more than 1,000") (Macedonia et al., 1995). Another good example of an early version of a persistent virtual environment was the Collaborative Architecture Via Immersive Navigation (CALVIN) developed by the University of Illinois at Chicago's Electronic Visualization Laboratory (Leigh et al., 1996; Leigh & Johnson, 2002). The system supported collaborative design sessions in VE attended by multiple individuals.

The representative examples of early persistent virtual environments created and maintained by industry include Second Life (SL) by Linden Lab, launched in 2003 (Second Life, 2025), with an array of other gaming and social media solutions that has followed, like the AltSpaceVR (no longer operational since March 2023) (AltSpaceVR, 2025), Fortnite (Fortnite, 2025), Decentraland (Decentraland, 2025), Roblox (Roblox, 2025), The Sandbox (The Sandbox, 2025), VRChat (VRChat , 2025), and Meta's Horizons worlds (Horizon World, 2025).

Computer-Supported Aircraft Maintenance

Any organization that operates and maintains a large number of systems critical for its operational readiness requires employing some form of computer-supported solutions. In the case of the Department of the Navy, maintaining and operating thousands of air, sea, and land assets would not be possible without a plethora of such system solutions. The primary types of system solutions for aircraft maintenance processes include Computerized Maintenance Management Systems (CMMS), electronic publications, and Aircraft Health Monitoring Systems (AHMS). The examples of specific instantiations of those solutions include the Naval Aviation Logistics Command Management Information System (NALCOMIS), Relational Supply (RSUPPLY), Decision Knowledge Programming for Logistics Analysis and Technical Evaluation system (DECKPLATE), Vector, One Touch Support, Interactive Electronic Technical Manuals (IETMS), the F-18 Automated Maintenance Environment (FAME), and Hornet Health Assessment and Readiness Tool (HhART). Nevertheless, the demands and needs of the maintenance domain are significantly larger than what the current versions of those system solutions can provide.

PERSISTENT AUGMENTED AND VIRTUAL ENVIRONMENT - PAVE

Recent emergence of the low-cost commercial off the shelf (COTS) virtual reality (VR), augmented reality (AR), smartphones, sensor technologies, and the advancements made in the domain of artificial intelligence, including the industry-based system solutions and infrastructure, raised the potential of addressing the daily needs of a large number of prospective users across many domain in a manner, scale, and with the efficiency that was not possible to realize before. The systemic and persistent application of these systems in daily operations and training onboard Navy ships and Marine Corps installations has a significant impact on DON operations, providing a strategic advantage for augmenting human performance and increasing the level of operational readiness.

The concept of persistent augmented environments and their use to address the needs of the DoD was proposed by Dr. Sadagic in 2019 (Sadagic, 2019). The persistent mode of PAVE's operation is reflected in the real-time collection and manipulation of two major types of data, which, once collected, contribute to the growing number of historical data sets in a particular domain. They include system-generated datasets (e.g., sensor data) and human-generated datasets (e.g., data resulting from human interactions with virtual and real environments, as well as collaborations). Actions of individual operators would also be added to the global historical data sets, which are continuously updated. The system, in its implementation, could be envisioned as a database with integrated understandings of the applications and users who would be provided with different access privileges to specific data (a user, for example, could only insert, view, or edit the information).

This aggregate set of diverse information would enable enhanced situation awareness among operators, allowing not only technology-supported decision-making based on past data but also the simulation of 'what if' scenarios and the forecasting of future events. The same data sets and display platforms can also be leveraged to support a variety of training needs. Personal display solutions, such as AR and VR glasses, can be worn by personnel who require their functionality at all times, providing immediate access to new system capabilities anywhere and at any time. Future decision-makers will be able to utilize immersive and non-immersive AR and VR display solutions as *windows* into complex data sets that are continuously updated to support their needs in real-time.

We envision the use of PAVE environments to support collaboration, communication, and training needs at the enterprise level. A key novel aspect of this approach is the concept of systematically utilizing technological solutions by individuals to support specific tasks at all times, and persistently collecting data from platform systems to derive organization-wide capabilities and gain a strategic advantage in decision-making. The recipients of the benefits of such a system will be individual trainees and operators, as well as decision-makers at various levels of an organization's operations.

FATCAT SYSTEM DESIGN AND IMPLEMENTATION

We designed a proof-of-concept AR application to demonstrate a practical use case of the PAVE environment, specifically serving the visual inspection phase of aircraft corrosion maintenance. We titled it the F/A-18 Tracking Corrosion and Assessment Tool (FATCAT). The goal is to study the feasibility of implementing this application and enable aircraft maintainers to execute corrosion control tasks more efficiently.

Design Criteria

The research team conducted a thorough task and gap analysis before designing the FATCAT system. Additionally, the team engaged with active duty U.S. Navy aircraft maintainers who shared their experiences with the current approach to conducting corrosion detection and provided suggestions for new (desired) functionalities. That information was analyzed, and a subset of features was selected as most compelling in terms of novel capabilities that would address identified gaps for implementation. A few examples of this include the following: identify specific areas of interest to ensure those areas are inspected, identify the types of corrosion during inspections so that the maintainer can appropriately assign action items to each type, and record data in a standard manner to ensure consistent documentation. The team also added a set of system features and requirements; examples include: the application should have a guided step-by-step walk-through on how to perform a corrosion inspection, there should be a standardized work order form (existing or new) that the user should be able to fill out within the application with work order generation automated based on location specific authoritative data encoded within the digital twin, and the application should be able to access the database of known maintenance data from previous inspections. The underlying infrastructure should set conditions for live maintenance data generation on the aircraft within the interface, as well as lay the groundwork for incorporating computer vision features to aid maintainers with the visual inspection process. A complete set of design criteria is presented in [Ashmore, 2024].

New Functionalities Supported by the PAVE and FATCAT: A Novel Framework for Aircraft Maintenance

The PAVE and FATCAT system solutions are viewed as fundamental building blocks for future Maintenance, Repair, and Operations on Deck - a framework that incorporates visualization and display technologies to support operators, along with a set of underlying principles for persistent environments with wireless connectivity between different parts of the overall system. The data collection conducted each time an operator performs any preventive or corrective maintenance on an aircraft provides not only the insights needed by individual maintainers (who have access to information generated by past maintenance jobs on a specific aircraft), but it also supports enterprise-level decision-making. With historical data on all maintenance jobs performed by all maintainers and on all aircraft, the USN can deploy analytical tools on that dataset and obtain insights into each aircraft's performance in various environments, its susceptibility to corrosion, or any other issue that may be reported during the maintenance procedure. The same type of answers could be provided for the entire family of aircraft, and, if needed, requests could be made to address frequently identified issues that were detected this way. In the fleet, the majority of maintenance training occurs on the job. A PAVE system brings a mechanism for conducting aided maintenance training on the aircraft, with access to the most up-to-date maintenance data. Training mode vs active maintenance interactions is possible through a common interface with training tailored to maintenance trends and specific maintainer's performance as the system connects directly to authoritative maintenance management systems and maintainer's log sets.

The introduction of the PAVE and FATCAT, much like the introduction of any new system, would also involve the design and acquisition of infrastructure to support their operation (for example, wireless networking, a database that stores historical data sets, and AR headsets), training of the workforce, as well as continuous monitoring and assessment of system and operators' efficiency including the effect that new solutions have on workflows (both positive and negative effects). While currently available AR headsets are predominantly used for indoor operation,

future AR headsets will likely support work in outdoor conditions (daylight readability). The operators could also use tablets as portable AR visual display solutions; many of them perform very well outdoors. However, their use prevents hands-free operation, which is characteristic of VR and AR headsets.

Programming and Development Environment

Hardware environment: The team used the following hardware environment: (a) Microsoft HoloLens 2 headset, and (b) MSI Raider GE76 12UE, 12th Generation Intel® Core™ i9-12900H, 16GB RAM, NVIDIA® GeForce RTX™ 3080 TI GPU Laptop. The choice of the AR headset was led by the user's needs: the device needed to enable full visibility of the real world (optical see-through enables that, while video pass-through still suffers from field of view (FOV) limitations inherent to the camera and display FOV)), the device had to provide reasonable FOV for the virtual (synthetic) elements for the environment, it had to be light and allow easy ventilation and not overheat while worn for an extended period of time.

Software environment: The software environment included: (a) Blender for 3D model development, (b) Unity, Visual Studio Code (VS), Mixed Reality Toolkit (MRTK), and Git Lab repository for software development, (c) Figma collaborative interface design tool and Microsoft PowerPoint for storyboard design, and (d) Slack for team communication and coordination.

System architecture: The system architecture consisted of a laptop equipped with a set of applications (Unity, MRTK), a Wi-Fi hotspot for wireless local area network (WLAN), and a HoloLens 2 headset. The wireless network between devices enabled easy transfer of code and data from the laptop to the headset and vice versa.

3D Model and User Interface

The research team employed a digital twin (DT) model of the aircraft instead of the actual aircraft (Figure 1); the DT was very close to the size of the actual F/A-18 Super Hornet (Figure 1). This decision allowed for much-needed flexibility when conducting the user study—a virtual replica of the aircraft was always available for experimentation and sessions with the study participants, which could not be the case with the actual aircraft. We were also able to instrument that 3D model with the number and variety of corrosion (Figure 2), dirt textures (Figure 3), and the bare metal (Figure 4), which guaranteed a mix most suitable for the study.). When fully employed, the presented concept would align a high-fidelity digital twin with the real aircraft and only visually render appropriate location-specific information within the AR interface. For the user study, the entire aircraft and augmented data were virtualized.

In addition to desired textures, we constructed an annotation collision system that allowed us to identify user interaction in segments of the model where additional data had to be made available to the user (for example, information about the existing work orders for specific areas, and images of the affected corrosion areas taken by previous maintainers).

Figure 1. Prototype Digital Twin in an Empty Hangar Parking Spot at an Operational Airfield

Figure 2. Galvanic corrosion texture on the trailing edge flap of the DT

Figure 3. Dirt texture on the outer wing of the DT

Figure 4. Bare metal texture on the aileron of the DT

The primary component of the user interface consisted of visual and auditory information presented to the user, while a gesture recognition system and speech recognition system for auditory commands facilitated the user's interactions with the system. A set of users' gestures that were supported and required to be mastered to interact with the tool included single-hand gestures (tap, hold, drag) and double-hand gestures (hold, resize, rotate). The users were also able to request 'take a picture' from the system; this functionality allowed users to save images of the affected areas, substantiate their decisions, and complete documentation about the work that was done.

A simple home menu allowed users to either conduct a calibration (adjust the size of the 3D model and place it anywhere in the physical space) or start the main simulation and the maintenance task menu (Figure 5). The latter option initiated the menu that provided necessary information about the steps of the task (informing the user about the current step), textual instructions informing the user about the expected actions, an orange arrow that cues the user towards the area of the aircraft that needs to be inspected, a photo of what the inspection area looks like without the corrosion (the baseline information), photos of example corrosion effects for easier recognition and identification of correct type of corrosion, annotation mode On/Off toggle, and a Next/Back toggle for easy navigation to the next or previous step in the main inspection task. If corrosion is detected, the user can activate the annotation tool (Figure 6) and create a new annotation (Figures 7 and 8) unless the system informs the user about the current outstanding work order for the same area.

Figure 5. Maintenance Task Menu

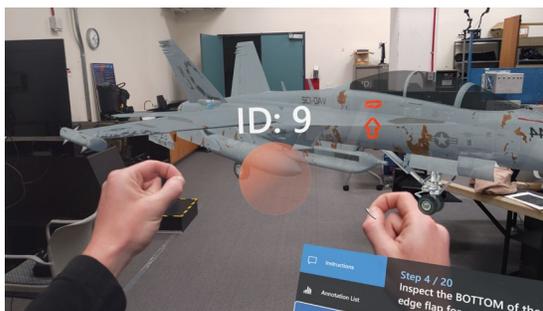**Figure 6. User Interface for Creating an Annotation**

Figure 7. User Interface for Expanding (Enlarging) an Annotation Area
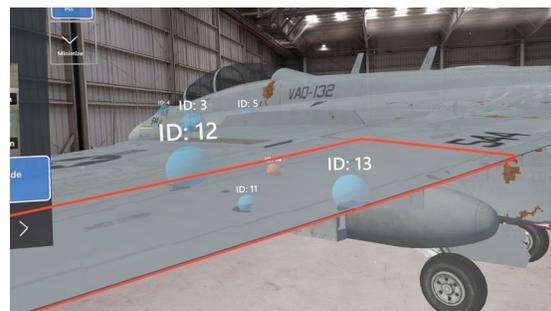

Figure 8. Corrosion Annotation

USER STUDY

The execution of the user study was approved by the NPS Institutional Review Board (IRB), with reference number NPS.2023.0023-IR-EP6-7-A. The protocol stipulates that participation in the study is voluntary, anonymous, and adheres to the professional research ethics, applicable laws, and other requirements of IRB-approved studies.

The goal of the user study was to collect data that would provide a firm foundation for addressing our second research question, 'What is the usability of the FATCAT augmented reality tool that supports aircraft corrosion control?' The methodology used for the user study followed standard approaches in usability studies, employing mixed methods to collect a comprehensive set of objective and subjective data that provide insights into user performance and their perception of the system. As mentioned earlier, the study used a digital twin of an actual F/A-18G Super Hornet aircraft (Figure 1).

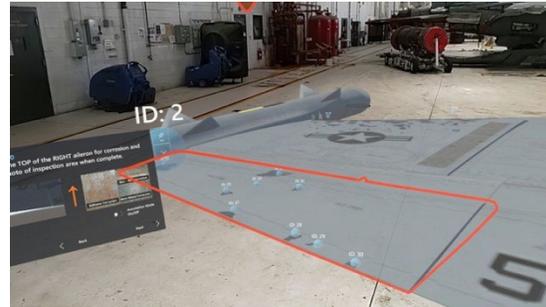**Figure 9. Third-person View of User Performing the Corrosion Inspection**

Figure 10. Corrosion Cases Annotated on Simulated Aircraft Inside a Hangar Bay

Apparatus

The apparatus used for the study included (a) Microsoft HoloLens 2 headset with FATCAT AR tool and digital model of the F/A-18G Super Hornet, (b) a laptop (MSI Raider GE76 12UE, 12th Generation Intel® Core™ i9-12900H, 16GB RAM, NVIDIA® GeForce RTX™ 3080 TI GPU), and (c) a hotspot for the Wi-Fi connectivity.

The study was conducted over three workdays and two nights of aircraft maintenance shifts, inside a hangar bay with doors closed to prevent direct sunlight and ensure the optimal light conditions necessary for the HoloLens 2 display. The virtual sessions with all subjects took place in the same aircraft hangar space aboard the flightline of an active naval air station. The existence of the physical aircraft and live maintenance tasks next to the experimental area added a level of realism that could not be achieved in the laboratory. This environment, along with the deployment of private 5G networks across the U.S. Department of Defense, demonstrates further progress in the adoption of the presented technology in relevant operational settings.

Participants

A total of 32 participants took part in the user study; they were active-duty USN F/A-18 Super Hornet aircraft maintainers (both day and night crew work shifts were represented). The age range was between 19 and 53 years (average = 28, standard deviation = 8.5 years; one person did not provide information). Their military service ranged from 11 months to 30 years (average = 7.4 years, st. dev. = 7.3 years). Thirty-one (31) participants completed the task, and only one participant was unable to do so due to an unexpected HoloLens 2 system crash; the objective data for that participant were excluded from the study. Therefore, the analysis of the objective data set focused on the data collected from 30 participants.

Task and Procedure

Each subject was asked to execute a simulated visual corrosion inspection on a modified version of the F/A-18G Super Hornet 182 DSI using the FATCAT tool. The task included 20 inspections (steps) in specified areas; an example step was: 'Inspect the top of the left trailing-edge flap for corrosion and take a photo of the inspection area when complete.'

Each study participant followed the same procedure. The steps included: (1) study brief and IRB consent form, (2) baseline simulator sickness questionnaire (baseline SSQ), (3) familiarization with HoloLens 2 headset, practice of hand gestures in an unrelated training environment (examples: object selection, pressing and pinching gestures), and the techniques for reduction of cybersickness symptoms, (4) pre-task simulator sickness questionnaire (pre-task SSQ), (5) study task execution, (6) post-task simulator sickness questionnaire (post-task SSQ), (7) system usability scale (SUS) questionnaire, (8) post-task custom survey, (9) demographics questionnaire, and (10) a short explanation of the study with opportunity to ask questions related to their participation in the study.

Data Collection and Results

A participant's ID, unique for each individual, allowed the research team to link objective and subjective data sets for each individual. The objective data set was designed to gather precise data on user performance. In contrast, subjective data are strategically used to gather user opinions, reflections, and preferences that extend beyond the performance data logged by the system. For example, if a participant experienced issues with the tool and did not perform well, the custom-made questions enabled the team to identify the most probable cause of such poor performance.

Objective Data Sets

The system-logged data sets, reflecting all user actions with the FATCAT tool, were collected for each participant using dedicated print statements in the application. Each action was timestamped with parameters that determined the type of action and values significant for a particular user action. For example, the system collected information about each element of the user interface that was selected (such as menu items), annotations being created and manipulated in the space (including their scale and coordinates in the local 3D space), the pictures of affected areas taken, and voice commands issued. In addition to such discrete user actions, the system logged a participant's local coordinates (X, Y, Z) and head orientation (yaw, pitch, roll) every second during the session, effectively describing the route that a participant navigated during the interactive session. Having such a rich data set not only allowed the research team to gain a precise understanding of each action and derive the elements of the user performance, but also to treat it as a form of playback, enabling the research team to view and analyze data and participants' actions from any viewpoint in the 3D space. The eye gaze was not collected and analyzed.

Once the data set was recorded on the headset, it was transferred to an external storage medium wirelessly – this simulated the way a future AR maintenance system would automatically save the session data, automate the work order generation process, and perform computer-aided quality assurance actions. For this study, the system generated two log files (one in CSV format and one in .log file format), which were then analyzed by the research team to derive necessary insights into user performance. The HoloLens 2 system crashed for participant #4, preventing that person from completing the session, and the system log data for participant #32 could not be recovered (the subjective data set was collected from that participant). The recorded data suggest that all participants executed all inspection steps.

Visual inspection time (task completion time): This value was measured between the participant pressing "Start demo maintenance" and the "Done" menu item after step 20; Figure 11 shows the visual inspection time used by each participant (data for 30 participants). The average time was 24 minutes 12 seconds, Min = 11 minutes and 24 seconds, Max = 40 minutes and 52 seconds, and St. dev. = 8 minutes and 46 seconds.

Number of pictures taken: The data logging system lost the images captured by two participants, and one participant's session had to be stopped because the HoloLens kept crashing. As a result, we obtained information about captured pictures for 29 participants. All participants took pictures; on average, they took 19 pictures during the main task, which consisted of 20 inspection steps. One participant took 31 pictures (Max), and another participant took only six pictures in 20 inspection steps (Min). The data also suggests that 14 participants took more than one picture during individual inspection sets, one participant took 20 pictures, and 14 participants missed to take pictures during certain inspection steps. To ensure that complete documentation is obtained about all inspections conducted by the maintainers, future versions of the system can remind them about this element of their task.

Use of the *View Active Annotations* function: Seventeen (17) participants did not turn on (activate) this function at any time during their inspection, and three participants did not open the annotations list menu. As a result, those individuals could not know if other maintainers had already reported specific corrosion annotations. All participants used the menu button to save and cancel annotations (they could have done it with the voice commands as well).

Subjective Data Sets

The collection of the subjective data sets was conducted using the Qualtrics tool, a web-based survey and report generation platform; the tool's server has been operated by the NPS to ensure the privacy of the collected data. A wireless cellphone hotspot provided a mobile internet connection to Qualtrics at the study location. The study collected the following sets of subjective data: (a) simulator sickness questionnaire (SSQ), (b) system usability scale (SUS) questionnaire, (c) post-task custom survey, and (d) demographics questionnaire. The number of analyzed subjective data sets was 31.
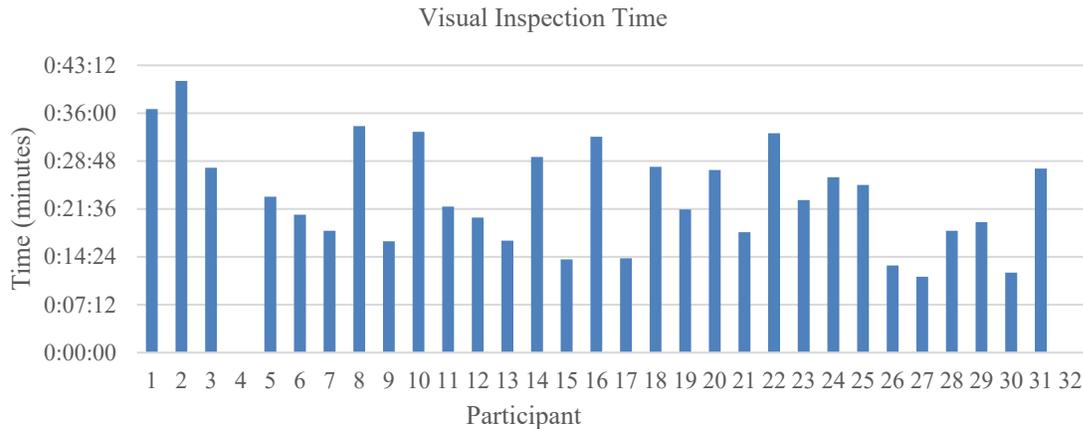

Figure 11. Time participants dedicated to the simulated inspection task

SSQ questionnaire: This is a standardized tool that enables the assessment of the severity of cybersickness symptoms after exposure to simulated environments (Kennedy et al., 1993). The tool allows for collecting values for 16 symptoms classified into three groups (nausea, oculomotor disturbance, disorientation); each symptom could be reported at one of the following four levels: None (=0), Slight (=1), Moderate (=2), or Severe (=3). Once the SSQ data were calculated as per Kennedy et al. (1993), a score of 0 was declared as having no symptoms, less than five was classified as negligible symptoms, 5-10 being minimal, 10-15 significant, 15-20 meant that the symptoms were a concern, and scores greater than 20 were considered harmful and simulator was qualified as bad (Stanney et al., 1997; Bimberg et al., 2020). The team collected three sets of SSQ values:

1. *Baseline SSQ* values that reflect the state the participants were in before any exposure to the augmented environment: Average: 0.84 (negligible level), Min: 0.00, Max: 11.22, St. dev.: 2.46.
2. *Pre-task SSQ* values after the training session in AR, just before the main task in FATCAT tool: Average: 0.60, (negligible level) Min: 0.00, Max: 7.48, St. dev.: 1.92
3. *Post-task SSQ* values right after the main task in FATCAT tool: Average: 2.41 (negligible level), Min: 0.00, Max: 22.44, St. dev.: 5.18. One participant reported a harmful post-task SSQ score (22.44), with slight symptoms of eye strain, difficulty focusing, blurred vision, and dizziness with eyes closed. Our past studies, which used AR display solutions, also reported slightly elevated values for symptoms in the oculomotor group (Timmerman, 2018; Turner & Cruz, 2022).

No SSQ symptom response was reported at the moderate or severe level at any time during the study.

System usability scale (SUS): The SUS is frequently used to assess the perceived usability and "appropriateness to a purpose" of a system (Brooke, 1996). The questionnaire consists of 10 statements, and participants rate each statement using the 5-point Likert scale (1 = strongly disagree, …, 5 = strongly agree). The final SUS score for each individual was calculated using recommendations by Bangor et al. (2009) for adjective ratings. The Average SUS score was 74.92 (*Good*), Min = 37.50, Max = 97.50, and St. dev. = 11.82. Four participants reported a SUS score of 90 and above, and only one participant scored the system with a score of less than 50 (37.50).

Post-task questionnaire: A small set of questions, customized for the FATCAT tool, was presented to participants after they completed the main inspection task. The participants gave comments on the level of realism presented in the simulation, the navigation and interaction in the system, the corrosion annotation tool, their performance, the

perceived value of the tool, the potential value of implementing the application concept, the overall usability, and desired (new) features that could be added to the tool.

When asked about the perceived value of the tool, 68% of participants were "very satisfied" and 29% were "satisfied" (Figure 12). No negative responses were reported in association with the overall experience. One participant stated the tool would be a "great asset and a high potential to affect the speed of aircraft maintenance and accuracy of documentation." Different elements of the system were also assessed (Figure 13). Eighty-one percent of participants found the overall experience with the user interface to be *very easy*, *easy*, or *somewhat easy*, and only 13% found it *somewhat difficult*.
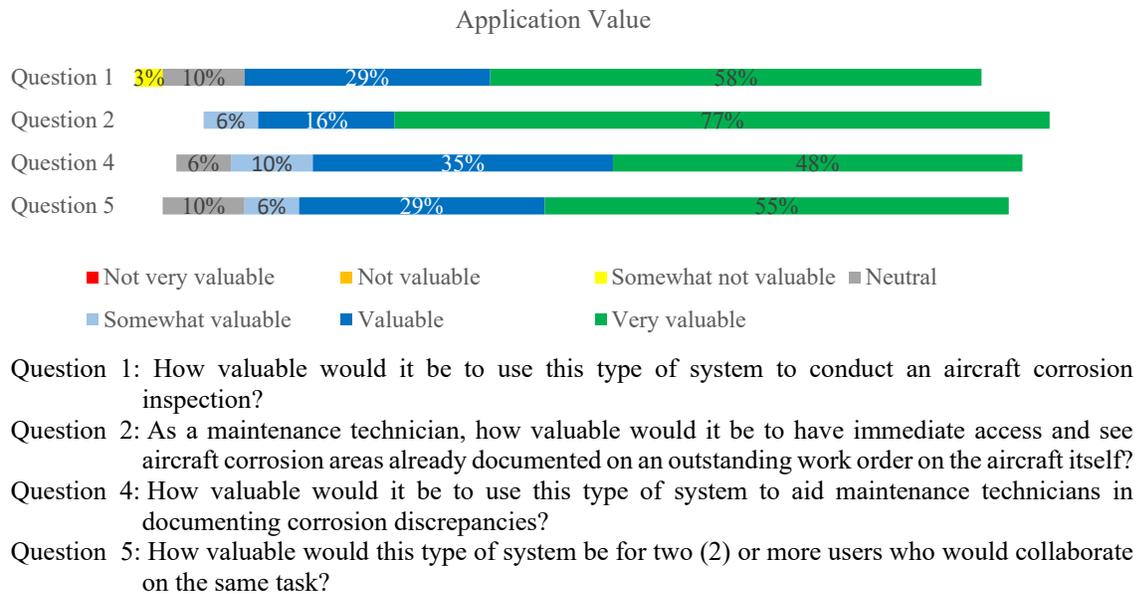
Application Value

Question 1: How valuable would it be to use this type of system to conduct an aircraft corrosion inspection?

Question 2: As a maintenance technician, how valuable would it be to have immediate access and see aircraft corrosion areas already documented on an outstanding work order on the aircraft itself?

Question 4: How valuable would it be to use this type of system to aid maintenance technicians in documenting corrosion discrepancies?

Question 5: How valuable would this type of system be for two (2) or more users who would collaborate on the same task?

Figure 12. Participants' view of the application value

User Interface Usability

Question 7: How was your overall experience with the user interface portion of the prototype system?
Question 8: How easy or difficult was it to navigate through the scene?
Question 9: How easy or difficult was it to interact with the menus?
Question 10: How easy or difficult was it to take a photo of each area inspected?
Question 11: How easy or difficult was it to interact with the settings brightness slider?
Question 12: How easy or difficult was it to make annotations for the 3D model in the areas where corrosion was identified?
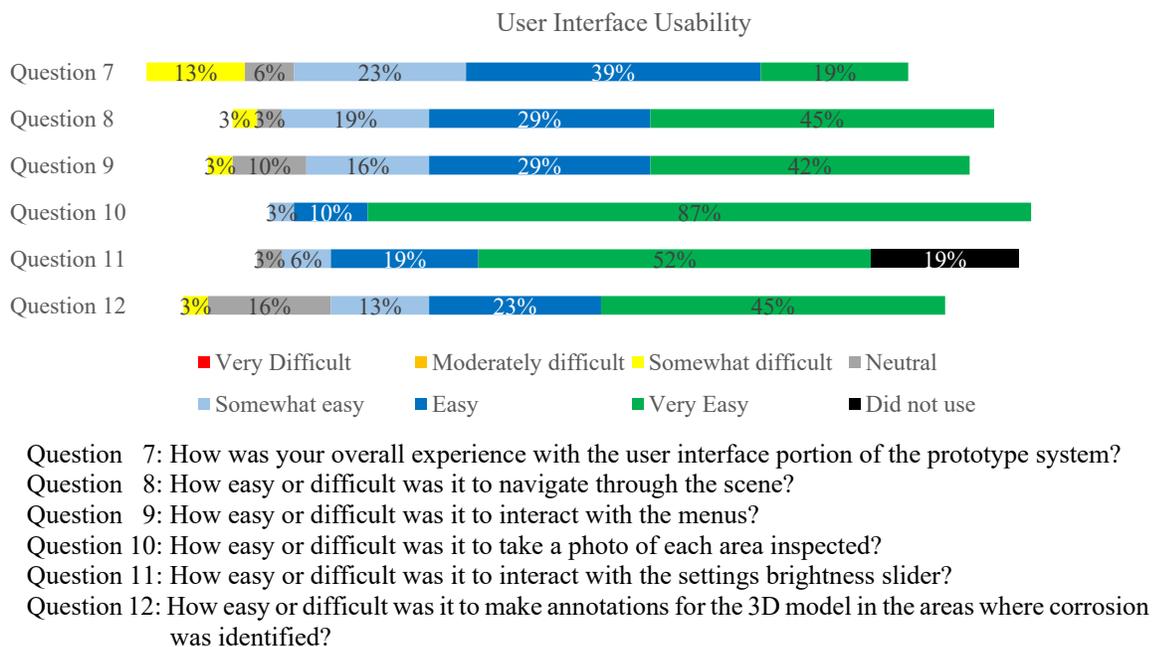
Figure 13. Participants' view of the application interface usability

When asked to rate their performance, 32% of participants were *very satisfied*, 32% were *satisfied*, 16% were *somewhat satisfied*, 13% were *neutral*, and 6% were *somewhat dissatisfied*. No participant reported being *moderately dissatisfied* or *strongly dissatisfied*.

DISCUSSION, RECOMMENDATIONS, AND FUTURE WORK

Several limitations were identified in the study. The FATCAT tool used in this study focused solely on visual inspections of the aircraft exterior. Most maintenance jobs, however, also involve removing panels and disassembling various components. The maintainers are also required to climb on top of the aircraft; our study avoided this requirement due to safety concerns, the protective gear they would need to wear under the HoloLens, and the need to keep the execution simpler (otherwise, a ladder would have to be carried and repositioned each time a top view of the aircraft was needed). The study also used only three corrosion-related textures in the model, which was simplistic. The job of distinguishing corrosion types and severity on the real aircraft is not as clear-cut in its nature. That is why adding an even larger variety of example pictures than what we had in the FATCAT interface could be a valuable aid for maintainers. Finally, the maintenance task simulated in our study was executed only indoors, while the actual maintenance tasks happen both indoors and outdoors. Adoption to outdoor tasks would require exploring alternate hardware solutions and expanding network connectivity to the ramp of aircraft MRO facilities.

The results obtained in the study helped us answer the research questions. Both objective and subjective data sets suggested that it was feasible to implement and utilize a persistent augmented environment to facilitate corrosion detection and maintenance tasks. All participants successfully completed their tasks and were satisfied with their performance. The application's frame rate remained consistently at 60 frames per second, with minor drops that did not impact users' performance. Regarding the usability of the system, the FATCAT tool's SUS score was rated as *good*. The level of realism was rated very high, and participants were very positive about the tool's functionalities. Sixty-six percent of study participants reported that they would be *very likely* to use FATCAT for daily and turnaround aircraft inspections, and 19% reported that they would *likely* use it. Users' readiness to employ the new tool in their daily practice constitutes a significant element in the overall receptiveness for the adoption of the novel technology.

The average SSQ score was rated as *negligible,* and the tool did not cause any SSQ symptoms at *moderate* or *severe* levels. We can only speculate that more extended use of the AR headset, which would simulate the length of the actual maintenance task, would not cause a dramatic change in the SSQ symptoms; however, only a new study with more prolonged AR exposure would be able to confirm or refute this.

The team identified two elements of the overall effort that could be improved: one concerned the need for more extensive stress testing of the system, and the other element was related to the pinch gesture, which was not sufficiently mastered by the study participants. Possible remedies would involve extended training sessions and extensive practice of the pinch gesture or the use of a different interface, such as a virtual button, instead of the pinch gesture.

The avenues of future work include utilizing real aircraft and object recognition capabilities through computer vision and artificial intelligence, which would enable the recognition of both the maintainer's position and the identification of the segments of the aircraft being worked on, as well as monitoring of the correct procedures in maintenance tasks. Any change identified on the aircraft would also be reflected in changes to the corresponding 3D model, with appropriate documentation being automatically generated. We also envision the development of a photo database and tail number-specific linkage to the database within maintenance management system tools. Being able to view historical examples of particular corrosion types in specific aircraft areas would be an invaluable tool for all maintainers, regardless of their level of expertise. Finally, the annotations do not need to be limited to corrosion incidents – they can be used as a communication language between technicians and maintainers to signal safety concerns or areas of the aircraft that require specific attention or repair.

CONCLUSION

The primary objective of this research was to examine the viability of using the concept of persistent augmented and virtual environments to assist in the aircraft maintenance task; a specific type of this task, investigated in the user

study, was the corrosion detection task. That objective was achieved by creating the FATCAT AR tool and conducting a user study to test its usability. A rich set of objective and subjective data collected in the user study helped us identify the elements of the tool and user experience that were rated as successful, as well as the elements that require additional work. Additionally, we were able to determine the future likely avenues and opportunities for integrating the AR technology into contemporary operations. While our study involved the military community, the aircraft maintenance task also occurs in the civilian domain, and the results of our study are applicable to that domain as well.

ACKNOWLEDGEMENTS

REFERENCES

AltSpaceVR, (2025). Retrieved from https://en.wikipedia.org/wiki/AltspaceVR

Ashmore, M. (2024). *Corrosion Control on The LVC Grid — Persistent Augmented and Virtual Environments In Support of Aircraft Maintenance.* Master's Thesis, Naval Postgraduate School, Monterey, CA.

Bangor, A., Kortum, P., & Miller, J. (2009). Determining what individual SUS scores mean: Adding an adjective rating scale. *Journal of usability studies, 4(3),* 114-123. https://uxpajournal.org/determining-what-individual-sus-scores-mean-adding-an-adjective-rating-scale/

Bimberg, P., Weissker, T., & Kulik, A. (2020, March). On the usage of the simulator sickness questionnaire for virtual reality research. *In 2020 IEEE conference on virtual reality and 3D user interfaces abstracts and workshops (VRW)*, 464-467. IEEE. https://doi.org/10.1109/VRW50115.2020.00098

Brooke, J. (1996). SUS-A quick and dirty usability scale. *Usability evaluation in industry*, *189*(194), 4-7.

Decentraland (2025). Retrieved from https://decentraland.org

Eriksson, H. (1994). Mbone: The multicast backbone. *Communications of the ACM, 37*(8), 54-60.

Fortnite (2025). Retrieved from https://www.fortnite.com

Kelly, K., & Rheingold, H. (1993). The dragon ate my homework. *Wired*, July, 69.

Kennedy, R. S., Lane, N. E., Berbaum, K. S., & Lilienthal, M. G. (1993). Simulator sickness questionnaire: An enhanced method for quantifying simulator sickness. *The international journal of aviation psychology*, 3(3), 203-220. https://doi.org/10.1207/s15327108ijap0303_3

Leigh, J., & Johnson, A. E. (2002). Supporting transcontinental collaborative work in persistent virtual environments. *IEEE Computer Graphics and Applications*, *16*(4), 47-51. https://doi.org/10.1109/38.511853

Leigh, J., Johnson, A. E., Vasilakis, C. A., & DeFanti, T. A. (1996). Multi-perspective collaborative design in persistent networked virtual environments. *In Proceedings of the IEEE 1996 Virtual Reality Annual International Symposium*, 253-260. IEEE. https://doi.org/10.1109/VRAIS.1996.490535

Macedonia, M. R., Brutzman, D. P., Zyda, M. J., Pratt, D. R., Barham, P. T., Falby, J., & Locke, J. (1995). NPSNET: A multi-player 3D virtual environment over the internet. *In Proceedings of the 1995 symposium on Interactive 3D graphics*, 93-ff.

Meta Horizon (2025). Retrieved from https://horizon.meta.com

Roblox (2025). Retrieved from https://www.roblox.com

Sadagic, A. (2019). *Persistent Augmented Environment in Support of Naval Operations.* Project proposal submitted to the NPS Naval Research Program, Monterey, CA, USA, 2019.

Stanney, K. M., Kennedy, R. S., & Drexler, J. M. (1997, October). Cybersickness is not simulator sickness. In *Proceedings of the Human Factors and Ergonomics Society annual meeting*, *41*(2), 1138-1142. Sage CA: Los Angeles, CA: SAGE Publications. https://doi.org/10.1177/107118139704100292

Second Life (2025), Retrieved from https://secondlife.com

The Sandbox (2025). Retrieved from https://www.sandbox.game

Timmerman, M. (2018). *Enabling Collaboration and Visualization of Complex Operational Technology Computer Networks with Augmented Reality Technology*. Master's Thesis, Naval Postgraduate School, Monterey, CA, Dec 2018.

Turner, J. & Cruz, J. (2022). Supporting Mission Planning With a Persistent Augmented Environment (PAE). Master's Thesis, Naval Postgraduate School, Monterey, CA, Jun 2022.

VRChat (2025). Retrieved from https://hello.vrchat.com

Zyda, M. J., Pratt, D. R., Monahan, J. G., & Wilson, K. P. (1992). NPSNET: Constructing a 3D virtual world. *In Proceedings of the 1992 symposium on Interactive 3D graphics*, 147-156. https://doi.org/10.1145/147156.147186