

Evaluating the Benefits of Augmented Reality for Task Localization in Maintenance of an Armored Personnel Carrier Turret

Steven J. Henderson

Columbia University*

Steven Feiner



Figure 1: (Left) A mechanic wearing a tracked head-worn display performs a maintenance task inside an LAV-25A1 armored personnel carrier. (Right) The AR condition in the study: A view through the head-worn display captured in a similar domain depicts information provided using augmented reality to assist the mechanic. (The view through the head-worn display for the LAV-25A1 domain was not cleared for publication due to security restrictions, necessitating the substitution of images from an alternative domain throughout this paper.)

ABSTRACT

We present the design, implementation, and user testing of a prototype augmented reality application to support military mechanics conducting routine maintenance tasks inside an armored vehicle turret. Our prototype uses a tracked head-worn display to augment a mechanic's natural view with text, labels, arrows, and animated sequences designed to facilitate task comprehension, location, and execution. A within-subject controlled user study examined professional military mechanics using our system to complete 18 common tasks under field conditions. These tasks included installing and removing fasteners and indicator lights, and connecting cables, all within the cramped interior of an armored personnel carrier turret. An augmented reality condition was tested against two baseline conditions: an untracked head-worn display with text and graphics and a fixed flat panel display representing an improved version of the laptop-based documentation currently employed in practice. The augmented reality condition allowed mechanics to locate tasks more quickly than when using either baseline, and in some instances, resulted in less overall head movement. A qualitative survey showed mechanics found the augmented reality condition intuitive and satisfying for the tested sequence of tasks.

KEYWORDS: maintenance, service, repair, attention, localization augmented reality

*{henderson,feiner}@cs.columbia.edu

INDEX TERMS: H.5.1 [Information Interfaces and Presentation]: Multimedia Information Systems—*Artificial, augmented and virtual realities*; H.5.2 [Information Interfaces and Presentation]: User Interfaces—*Training, help, and documentation*

1 INTRODUCTION

Maintenance and repair operations represent an interesting and opportunity-filled problem domain for the application of augmented reality. The majority of activities in this domain are conducted by trained maintenance personnel applying established procedures to documented designs in relatively static and predictable environments. These procedures are typically organized into sequences of quantifiable tasks targeting a particular item in a specific location. These characteristics and others form a well-defined design space, conducive to a variety of systems and technologies that could assist a mechanic in performing maintenance.

Physically navigating these tasks, however, can be extremely time consuming and requires significant head and neck movement when transitioning between tasks. Maintenance sequences can also be difficult to traverse cognitively because they require mechanics to first position a given task in a presumed model of the environment and then correctly identify this location in the physical world. This problem is particularly acute when maintaining complex systems, such as those found in industrial, military, and aerospace domains. Maintenance sequences in such systems typically span dozens of tasks involving potentially unfamiliar objects randomly distributed across a given area. Moreover, movement in and around these systems can be complicated by their sheer size or by structural characteristics that restrict a mechanic's view and freedom of movement.

In this paper, we examine and document how augmented reality (AR) can assist in reducing the time and effort in navigating larger

sequences of maintenance tasks in such complex systems. We describe the design and user testing of a prototype AR application (Figure 1) for assisting mechanics in navigating realistic and challenging repair sequences inside the cramped interior of an armored vehicle turret. Our application uses AR to enhance localization in standard maintenance sequences with on-screen instructions, attention-directing symbols, overlaid labels, context-setting 2D and 3D graphics, and animated sequences. This information is combined on a mechanic’s natural view of the maintenance task using a tracked head-worn display (HWD).

Our contributions include a domain-specific user study examining professional mechanics using our system to maintain actual equipment in a field setting. Our user study demonstrates how mechanics performing maintenance sequences under an AR condition were able to locate tasks more quickly than when using several baseline conditions. We also document specific instances when the AR condition allowed mechanics to perform tasks with less overall head movement than when using these baselines. Finally, we convey the qualitative insights of these professional mechanics with regard to the intuitiveness, ease of use, and acceptability of our approach.

2 RELATED WORK

There has been much interest in applying AR to maintenance tasks. This interest is reflected in the formation of several collaborative research consortiums specifically dedicated to the topic—ARVIKA [11], Services and Training through Augmented Reality (STAR) [20], and ARTESAS [2]. These and other efforts have resulted in a sizable body of work, much of which is surveyed by Ong, Yuan, and Nee [18].

The majority of related work focuses on specific subsets of the domain, which we categorize here as activities involving the inspection, testing, servicing, alignment, installation, removal, assembly, repair, overhaul, or rebuilding of human-made systems [30]. Within this categorization, assembly tasks have received the most attention. Caudell and Mizell [6] proposed a seminal AR prototype to assist in assembling aircraft wire bundles. Subsequent field testing of this system by Curtis and colleagues [8] found the prototype performed as well as baseline techniques, but faced several practical and acceptance challenges. Reiners and colleagues [21] demonstrated a prototype AR system that featured a tracked monocular optical see-through (OST) HWD presenting instructions for assembling a car door. Baird and Barfield [4] showed that users presented with screen-fixed instructions on untracked monocular OST and opaque HWDs completed a computer motherboard assembly task more quickly than when using fixed displays or paper manuals.

Tang and colleagues [27] studied the effectiveness of AR in assembling toy blocks and found users made fewer dependent errors when aided by registered instructions displayed with a tracked stereoscopic OST HWD, compared to traditional media. An experiment by Robertson, MacIntyre and Walker [22] discovered that subjects assembled toy blocks more quickly while viewing registered instructions on a tracked binocular video see-through (VST) HWD than when using non-registered variants. Zauner and colleagues [33] demonstrated a prototype system for employing AR in a furniture assembly task. Qualitative studies by Nilsson and Johansson involving a medical assembly task [16] and by Salonen and Säski involving 3D puzzle assembly [23] suggest strong user support for AR.

Additional work outside the assembly domain includes Feiner, MacIntyre, and Seligmann’s knowledge-based AR maintenance prototype [10], which used a tracked monocular OST HWD to present instructions for servicing a laser printer. Ockermann and Pritchett [17] studied pilots performing preflight aircraft inspections while following instructions presented on an untracked OST

HWD. The results of this study demonstrated an undesired over-reliance on computer-generated instructions. Smailagic and Siemorek [26] designed a collaborative wearable computer system that displayed maintenance instructions on an untracked opaque HWD. Schwald and Laval [24] proposed a prototype hardware and software framework for supporting a wide range of maintenance categories with AR. Knöpfle and colleagues [13] developed a prototype AR application and corresponding authoring tool to assist mechanics in removing and installing components, plugs, and fasteners. Platonov and colleagues [19] developed a similar proof-of-concept system featuring markerless tracking.

There is also notable work on the general task of localizing a user’s attention in AR. Feiner, MacIntyre, and Seligmann [10] used a 3D rubberband line drawn from a screen-fixed label to a possibly offscreen target object or location. Biocca and colleagues developed the “Attention Funnel” [5], a vector tunnel drawn to a target, similar to “tunnel-in-the-sky” aviation cockpit head-up displays, and showed that it reduced search time compared to world-fixed labels or audible cues. Tönnis and Klinker [28] demonstrated that an egocentrically aligned screen-fixed 3D arrow projected in AR was faster at directing a car driver’s attention than an exocentric alternative. Wither, DiVerdi, and Höllerer [32] compared the performance of various displays to support visual search for text in AR (a task supported by localization), but did not detect any significant differences between display conditions. Schwerdtfeger and Klinker [25] studied AR attention-directing techniques to help users find and pick objects from stockroom storage bins. Their frame-based technique outperformed static 3D arrows and variants of the Attention Funnel.

Two aspects of our contributions distinguish them from this previous work. First, other than the wire bundle assembly research conducted by Curtis and colleagues [6], our research is the only project we know of to include a quantitative study of professional users employing AR for maintenance tasks under field conditions. Our work differs from the wire bundle assembly research by examining a more diverse set of maintenance tasks (e.g., inspection, alignment, removal, and installation) in a more restrictive environment using different comparison conditions. Second, our work is the first within the maintenance domain that articulates the potential benefits of AR for reducing head movement.

3 PROTOTYPE

We developed a hardware and software architecture for studying AR applications for maintenance. This architecture allowed us to create an implemented prototype that we evaluated through the user study described in Section 4. We note that our prototype is a laboratory proof-of-concept system to explore the potential benefits of AR for supporting maintenance procedures under field conditions and is not a production-ready implementation. Therefore, our software and hardware choices did not have to reflect the needs of a production environment.

Our prototype supports United States Marine Corps (USMC) mechanics operating inside the turret of an LAV-25A1 armored personnel carrier. The LAV-25 (of which the LAV-25A1 is a variant) is a light-wheeled military vehicle, and the turret portion is a revolving two-person enclosed, cockpit-like station in the middle of the vehicle. The entire turret volume is approximately 1 cubic meter, but a large amount of electrical, pneumatic, hydraulic, and mechanical infrastructure encroaches from random directions and in close proximity to the crew’s operating space. A mechanic servicing the turret works while sitting in one of two seats that are each fixed along the longitudinal axis of the turret. The resulting work area is approximately 0.34 cubic meters and spans the entire area surrounding the mechanic.



Figure 2: A typical localization sequence in our prototype. (a) A screen-fixed arrow indicates the shortest rotation distance to target. (b) As the user orients on the target, a semi-transparent 3D arrow points to the target. (c) When the user orients on the target, the 3D arrow begins a gradual fade to full transparency. (d) When the arrow has completely faded, a brief highlighting effect marks the precise target location.

Because we did not have regular access to the vehicle, we used an extensive set of 3D laser scans to create a mostly virtual mock-up of the turret, which we used in our lab during development. We then finalized our design in an actual turret in two separate pilot tests prior to the user study in the real turret. The first pilot test involved prototype testing with users at the Marine Corps Logistics Base in Albany, Georgia. We used this pilot test to refine our tracking configuration and gather user feedback about our interface and attention-directing techniques. The second pilot test involved four mechanics from the population recruited for our user study described in Section 4. These mechanics experienced nearly the same test procedure as other participants, but their data was excluded after we modified two tasks to reduce the overall execution time of our experiment.

3.1 Software

Our AR application software was developed as a game engine “mod” using the Valve Source Engine Software Development Kit. The engine “player” serves as a virtual proxy for the user and is positioned by location information from the tracking hardware. All virtual content in the AR scene is provided by custom game engine models, GUI elements, and other components. Full resolution stereo video from two Point Grey Firefly MV cameras is stretched to the scene back buffer via an external DLL that hooks the game engine’s instance of the DirectX graphics interface via the Windows Detours library. The entire scene is rendered in stereo at 800×600 resolution with an average frame rate of 75 fps. (Note: the effective video frame rate is approximately 25 fps, due to software upscaling from 2×640×480 to 2×800×600.)

At any given point in time, the application assumes a state representing a particular task (e.g., toggling a switch) within a larger maintenance sequence (e.g., remove an assembly). For each task, the application provides five forms of augmented content to assist the mechanic:

- *Attention-directing information* in the form of 3D and 2D arrows.
- *Text instructions* describing the task and accompanying notes and warnings.
- *Registered labels* showing the location of the target component and surrounding context.
- *A close-up view* depicting a 3D virtual scene centered on the target at close range and rendered on a 2D screen-fixed panel.
- *3D models* of tools (e.g., a screwdriver) and turret components (e.g., fasteners or larger components), if applicable, registered at their current or projected locations in the environment.

Attention-directing graphics follow a general sequence (Figure 2) that depends on 6DOF user head pose. If the target component is behind the mechanic, a screen-fixed green arrow points the user in the shortest rotational direction to the target. Once the target is within 180° (yaw) of the user’s line of sight, a tapered red semi-

transparent 3D arrow appears directing the user toward the target. The tail of the arrow is smoothly adjusted and placed along the far edge of the display at each frame, based on the vector between the target and the user’s projected line of sight on the near clipping plane. This ensures the arrow provides a sufficient cross-section for discernment. As the user approaches the target, the arrow increases in transparency and eventually disappears and spawns a highlighting effect for five seconds at the precise position of the target. Depending on task preferences and settings, the 3D arrow will reengage if the angle between the user’s head azimuth and the direction to target exceeds 30°.

For more complex or potentially ambiguous tasks, animated 3D models are added to the user’s view. These animations show the correct movement of tools or components required to accomplish a particular task. For example, when mechanics are instructed to remove or install fasteners, animated tools demonstrate the correct tool motion to accomplish the task, as shown in Figure 1 (right). Animation sequences are controlled and synchronized so they begin when the 3D arrow disappears, and play for a finite period of time (five seconds).

If a mechanic wishes to replay an animated sequence or control its speed, they can use a wireless wrist-worn controller, shown in Figure 3, which serves as the primary means of interaction in our prototype. The controller uses a custom 2D interface application written using the Android SDK, and provides forward and back buttons that allow the mechanic to navigate between maintenance tasks. When viewing tasks with supporting animation, additional buttons and a slider are provided to start, stop, and control the speed of animated sequences (Figure 3 inset). These animation buttons are hidden for non-animated tasks.

3.2 Hardware

We experimented with two HWDs while developing our prototype. The display we eventually used for user trials (Figure 1) is a custom-built stereo VST HWD constructed from a Headplay 800×600 resolution color stereo gaming display with a 34° diagonal field of view (FOV). We mounted two Point Grey Firefly MV 640×480 resolution cameras to the front of the HWD, which were connected to a shared IEEE 1394a bus on the PC. The cameras are equipped with 5mm micro lenses and capture at 30 fps. This application executes on a PC running Windows XP Pro, with an NVIDIA Quadro 4500 graphics card. We also experimented with, and initially intended to use, an NVIS nVisor ST color stereo OST HWD. We selected this display because of its bright 1280×1024 resolution graphics, 60° diagonal FOV, and high transmissivity. However, during pilot testing, we discovered that vehicle assemblies directly in front of and behind the seats prevented users from moving their head freely while wearing the relatively large nVisor HWD. This necessitated use of our custom-built HWD.



Figure 3: A mechanic uses the wrist-worn controller to cue the next task in a repair sequence. The inset view shows additional features that appear during applicable tasks for controlling animations.

Tracking is provided by a NaturalPoint OptiTrack tracking system. The turret's restricted tracking volume and corresponding occluding structures created a non-convex and limited stand-off tracking volume, which led us to employ 10 tracking cameras to achieve ample coverage. Because we were focused on research, rather than practical deployment, we were not concerned with the disadvantages of adding a large number of cameras to the existing turret. In contrast, future production-ready AR maintenance systems might instead use cameras and other sensors built into the task environment or worn by the maintainer, possibly in conjunction with detailed environment models.

The OptiTrack system typically uses passive retroreflective markers illuminated by IR sources in each camera. During pilot testing, we discovered that numerous metallic surfaces inside the turret created spurious reflections. Although we were able to control for all of these with camera exposure settings or by establishing masked regions in each camera, these efforts greatly reduced tracking performance. Therefore, we adopted an active marker setup, using 3 IR LEDs arrayed in an asymmetric triangle on the HWD. Given the confined space inside the turret, we were concerned that a worker's head position could potentially move closer than the 0.6 meter minimum operating range of the OptiTrack. However, experimentation revealed that, for any point inside our work area, at least four cameras could view a user's head from beyond this minimum operating range. Moreover, the active marker setup prevented the possibility of IR light from cameras reflecting off the user's head or hair at close range. The tracking software utilizes the NaturalPoint Tracking Tools 2.0 application, which streams tracking data at 60 Hz to the PC running the AR application over a dedicated gigabit ethernet connection. The tracking application runs on an Alienware M17 notebook running Windows Vista, with an additional enhanced USB controller PC Card.

We implemented our wrist-worn controller using an Android G1 phone (Figure 3). The device displays a simple set of 2D con-

trols and detects user gestures made on a touch screen. These gestures are streamed to the PC running the AR application over an 802.11g link. The mechanic attaches the device to either wrist using a set of Velcro bracelets.

4 USER STUDY

We designed a user study to compare the performance and general acceptance of our prototype (the AR condition) to that of an enhanced version of the system currently used by USMC mechanics. We also included an untracked version of our prototype in the study as a control for HWD confounds. Six participants (all male), ages 18–28, were recruited from a recent class of graduates attending the USMC Light Wheeled Mechanic Course in Aberdeen Proving Ground, Maryland. Participant computer experience ranged from minimal to extensive, with the majority rating their experience with mechanical systems as above a basic level of exposure. Two participants identified themselves as requiring contact lenses or glasses, and both determined that the separate left and right eye focus adjustments on the Headplay display provided adequate correction. All participants were right-handed.

4.1 Baseline Comparison Technique

In our experiment, we wanted to compare our prototype against current techniques used by USMC mechanics while performing maintenance task sequences. These techniques principally involve the use of an Interactive Electronic Technical Manual (IETM) [1], a 2D software application deployed on a portable notebook computer carried and referenced by mechanics while completing tasks. IETM users browse electronic documents in portable document format (PDF) using a specialized reader, an example of which is shown in Figure 4. We felt that a comparison against this system would not be compelling for several reasons. First, the extra time required to navigate this software, which affords less user control than common PDF readers, is significant. Second, the perspective views featured in the software are drawn from arbitrary locations and contain minimal context, which requires users to browse multiple pages with the suboptimal interface. As a result, any task completion or localization metrics would be heavily influenced by the time required to negotiate the IETM interface.

Therefore, we designed and adopted an improved version of the IETM interface to use as a baseline in the study. This baseline (the LCD condition) features static 3D scenes presented on a 19" LCD monitor. The monitor was fixed to the right of the mechanic (who sat in the left seat of the turret during our experiment), on an azimuth of roughly 90° to the mechanic's forward-facing seated direction. The LCD was positioned and oriented to reflect how

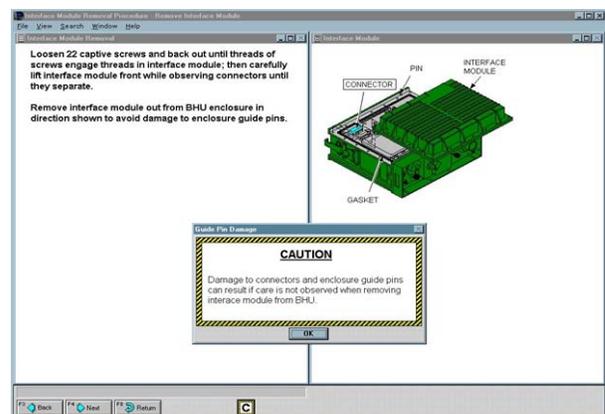
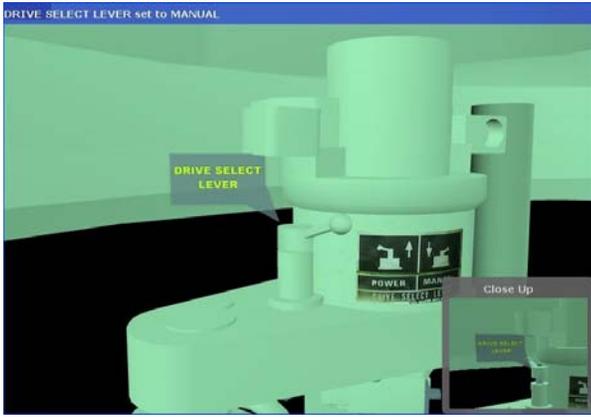
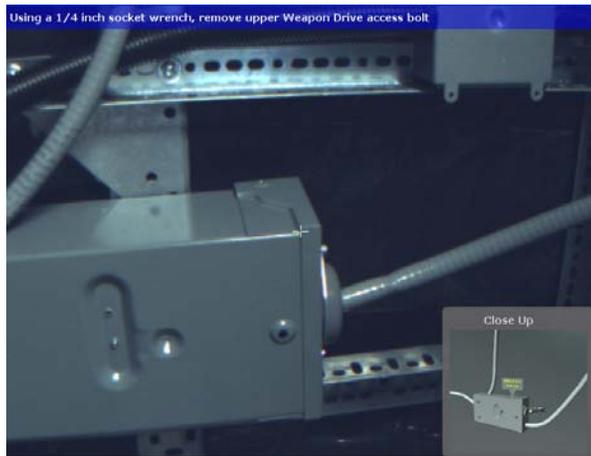


Figure 4: Example screen shot from the currently used IETM interface.



(a) LCD condition.



(b) HUD condition.

Fig 5: Baseline comparison techniques.

mechanics naturally arrange IETM notebook computers while working from the left seat. During each task, the LCD presents a single static 3D scene rendered in VR. Each static scene, such as the example shown in Figure 5(a), is rendered using the same engine that generates virtual content for the AR condition and depicts identical text instructions, 3D labels, close-up graphics, and animated sequences (if applicable). Additional 3D models are added to the scene to depict the central component of interest as well as important surrounding context. For each task, static perspective views were chosen that generally correspond to how each scene would naturally appear to a user sitting in the left seat. The FOV for each scene in the LCD condition was widened to 50° to approximate the perspectives used in IETMs. When experiencing the LCD condition during the user study, mechanics control the displayed scene by manipulating the system state with the wrist-worn controller.

To control for the general effects of wearing a HWD, we added a third condition featuring an untracked version of our AR prototype. This HUD (head-up display) condition uses screen-fixed graphics that depict text instructions and close-up views identical to those in the AR condition. An example screen shot is shown in Figure 5(b). While experiencing the HUD condition, participants wear the same HWD worn in the AR condition, and interact with the application using the same wrist-worn controller used in both the AR and LCD conditions.

Task	Description	Pitch	Azimuth
T1	Switch A OFF	31.3	40.2
T2	Remove Bulb X	26.5	35.9
T3	Switch B ON	28.2	33.0
T4	Remove Bolt #1	25.3	15.9
T5	Switch C OFF	38.5	42.1
T6	Inspect Assembly #1	30.9	58.0
T7	Inspect Assembly #2	19.4	-34.4
T8	Drive Lock to LOCK	-10.1	132.2
T9	Install Bulb Y	20.2	36.5
T10	Switch D OFF	23.2	39.5
T11	Switch E ON	25.2	45.1
T12	Lever 23 to Manual	11.9	-42.0
T13	Remove Bolt #2	35.9	17.8
T14	Remove Screw K	19.0	37.0
T15	Install Bolt # 1	25.3	15.9
T16	Connect Cable	29.3	19.1
T17	Install Screw S	19.0	43.9
T18	Install Bolt #2	35.9	17.8

Table 1: Selected tasks (with descriptions expurgated for publication) and corresponding pitch and azimuth measured from 0.7 meters above the center of the left turret seat.

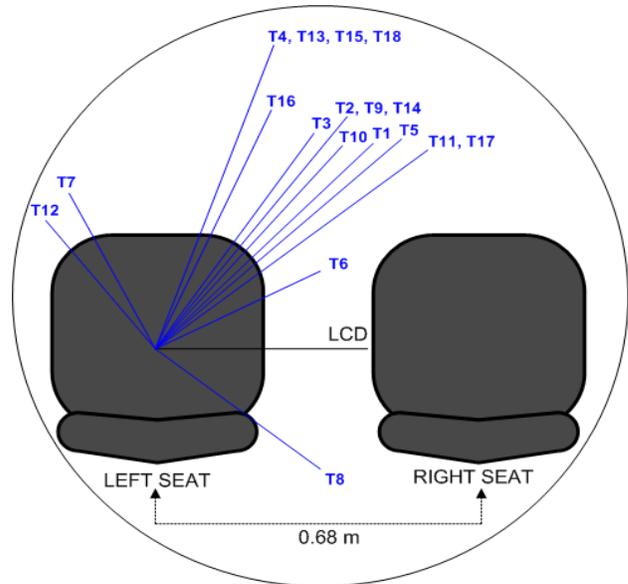


Figure 6: Approximate task azimuths and distances as viewed from above the turret and looking down. Neighboring task identifiers are separated by commas.

4.2 Tasks

We selected 18 representative maintenance tasks for inclusion in the user study from among candidates listed in the operator's manual for the vehicle [31]. Table 1 summarizes the selected set of tasks, and Figure 6 shows their approximate arrangement inside the turret. These tasks serve as individual steps (e.g., removing a bolt) performed as part of a larger maintenance sequence (e.g., replacing a pump). We specifically avoided adopting correct contiguous sequences of tasks to mitigate experiential influences in the experiment. We selected tasks that a trained mechanic could perform while sitting in the left seat and reasonably complete in under five minutes. We also sought to include a diversity of tasks

representing various strata within the larger spectrum of maintenance operations defined in [30].

4.3 Procedure

A within-subject, repeated measures design was used consisting of three conditions (AR, LCD, and HUD) and 18 maintenance tasks. The experiment lasted approximately 75 minutes and was divided into three blocks, with a short break between blocks. Each block consisted of all 18 tasks for each condition. Block order was counterbalanced across participants using a Latin square approach. The task order within blocks was fixed to facilitate the comparison of display conditions. We addressed potential learning effects by selecting 18 different task locations, which effectively prevented memorization of task order. At the start of the experiment, participants were asked to sign a consent form and then sit in the left seat of the turret. Before each block, the participant was given an orientation on how to wear the equipment used in each particular condition. In the AR and HUD conditions, this consisted of fitting and focusing the HWD, with an additional brief calibration step for the AR condition. For the LCD condition, participants donned a light-weight head band affixed with IR LEDs to facilitate the collection of tracking data. No portion of this apparatus entered the participant's field of view during the experiment.

Before each block, each participant was afforded an opportunity to rehearse the condition using five practice trials until they felt comfortable. Tools and fasteners required for tasks within the block were arrayed on a flat waist-high structure to the right of the seat and their locations highlighted to the participant.

The timed portion of the block consisted of 18 trial tasks distributed throughout the mechanic's work area. Each trial began when the mechanic pressed the "next" button on the wrist-worn controller. This started the overall task completion timer, and triggered the presentation of instructional text, close-up views, and labels associated with the trial task. In the AR condition, cueing information (i.e., the red or green arrow) was simultaneously activated, prompting the user to locate the target. The localization time was recorded when the user positioned their head such that the target location entered and remained within a 200 pixel radius of the center of the display for more than one second. In the AR and HUD conditions, a cross-hair was displayed to the participant to remind them to center each target. In the LCD condition, which presented static VR scenes for each task during the experiment, collected tracking data was replayed in a discrete event simulation after the experiment to calculate the localization time. Following target localization, overall task completion timing continued until the mechanic gestured on the wrist-worn controller for the next task. The block then proceeded to the next task until the participant experienced all 18 tasks.

5 QUANTITATIVE RESULTS

5.1 Data Preparation

We performed several preprocessing steps prior to analyzing our results. First, because the tracker coordinate system was centered above the left camera of our VST HWD, we translated tracking data points to a position coincident with the center of rotation for the participant's head. This was accomplished by adding a small offset vector v to each reading, where v was estimated by combining HWD measurements with population-specific anthropometric data from Donelson and Gordon [9] and supplemented by Paquette and colleagues [12].

We then removed spurious points in the recorded tracking and completion time datasets. For tracking data, we applied a moving average filter as defined by Law and Kelton [14]. After some

experimenting, we selected a window size of 0.25 seconds, which was applied to all six degrees of freedom. For completion time data, we manually inspected the task completion timestamps that were triggered when the subject gestured for the next task using the wrist-worn controller. In several instances, subjects made accidental double gestures, then immediately (usually within two seconds) gestured on the "back" button to reload the appropriate task. We identified and removed eight of these instances.

Our final data preparation step involved normalizing position and orientation data for each subject. Because the HWD was worn differently by each user, the relative position and orientation of the tracker to tasks in the experiment varies by subject. To standardize all subjects to a common reference frame, we individually normalized each subject's position and orientation data, as suggested by Axholt, Peterson, and Ellis [3].

5.2 Completion Time Analysis

We applied a 3 (Display Condition) \times 18 (Task) repeated measure ANOVA to task completion time with our participants as the random variable. Using $\alpha=0.05$ as our criterion for significance, the Display condition had a significant main effect on completion time ($F_{(2,34)}=5.252$, $p < 0.05$). The mean task completion times for each condition were 42.0 seconds (AR), 55.2 seconds (HUD), and 34.5 seconds (LCD) and are shown as asterisks in the Tukey plots in Figure 7 [7]. A subsequent pairwise comparison of these means revealed that the AR condition was 23% faster than the HUD condition, which was not significant. The LCD condition was 18% faster than the AR condition (not significant), and 37% faster than the HUD condition, which was statistically significant ($p < 0.04$). The set of maintenance tasks used in the study had a significant main effect on completion time ($F_{(17,34)}=8.063$, $p < 0.001$), which we expected, given the varying levels of effort required to perform each task.

5.3 Localization Time Analysis

We applied a 3 (Display Condition) \times 18 (Task) repeated measure ANOVA on task localization time with our participants as the random variable. Display condition exhibited a significant main effect on localization time ($F_{(2,34)}=42.444$, $p < 0.05$). The mean task localization times were 4.9 seconds (AR), 11.1 seconds (HUD), and 9.2 seconds (LCD), as shown in Figure 8. A subsequent pairwise comparison of mean task localization time revealed that AR was 56% faster than HUD, which was statistically significant ($p < 0.001$) and 47% faster than LCD ($p < 0.004$), which was also statistically significant. LCD was 17% faster than HUD, which was statistically significant ($p < 0.02$). As expected, the par-

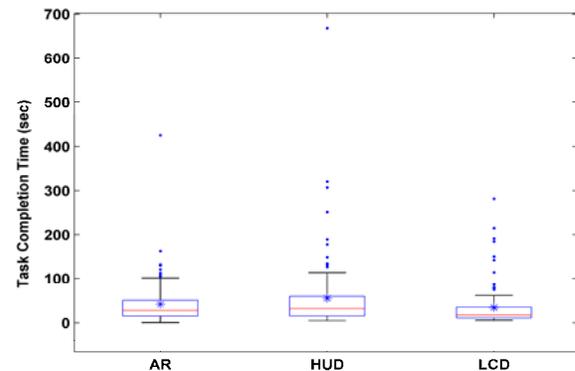


Figure 7: Task completion times (seconds) for AR, HUD, and LCD. An asterisk marks the mean task completion time for each condition.

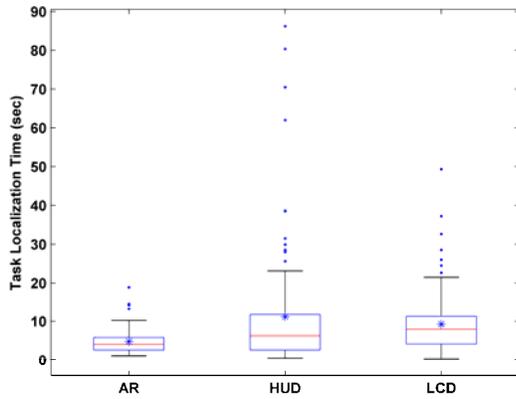


Figure 8: Task localization times (seconds) for AR, HUD, and LCD. An asterisk marks the mean task localization time for each condition.

particular set of selected maintenance tasks used in the study did not have a significant main effect on localization time ($F_{(2,34)}=1.533$, $p < 0.103$).

5.4 Error Analysis

Errors in our experiment were defined as instances when a subject performed a task to completion on the wrong item, and were logged by the observer during the experiment. Examples of errors included toggling an incorrect switch, removing an incorrect bolt, or inspecting the wrong item. In general, we found mechanics made few errors, and confirmed this with a 3 (Display Condition) \times 18 (Task) repeated measure ANOVA on task errors with our participants as the random variable. Display condition did not exhibit a significant main effect on total errors ($F_{(2,34)}=1.00$, $p=0.410$). This corroborates earlier findings by Robinson and colleagues [22].

5.5 Head Movement Analysis

Our analysis of head movement focused on the range of head rotation, rotational exertion and velocity, translational exertion and velocity, and the distribution of view direction. This analysis was confined to only the localization portion of each task because it was difficult to isolate head movements from overall body motion during the hands-on portion of some tasks. In these tasks, the user remained relatively static during localization, but adopted many different body poses once they began the physical portion of the task.

5.5.1 Range of head rotation

Table 2 depicts the descriptive statistics for overall ranges in head rotation about each axis and across all tasks. Left and right head rotation about the neck (azimuth or yaw) was the greatest, and generally conforms to the relative task azimuths shown in Table 1. The ranges of head yaw are fairly consistent across display condi-

Axis		Min	Max	Range
Pitch	AR	-4.0	74.1	78.0
	HUD	-6.3	64.7	71.0
	LCD	1.9	85.9	84.0
Roll	AR	-8.3	34.0	42.3
	HUD	-21.8	34.1	55.9
	LCD	-45.3	53.9	99.2
Yaw	AR	-56.4	143.2	199.6
	HUD	-67.0	134.2	201.2
	LCD	-39.3	143.3	182.6

Table 2: Ranges (in degrees) in head rotation across all tasks.

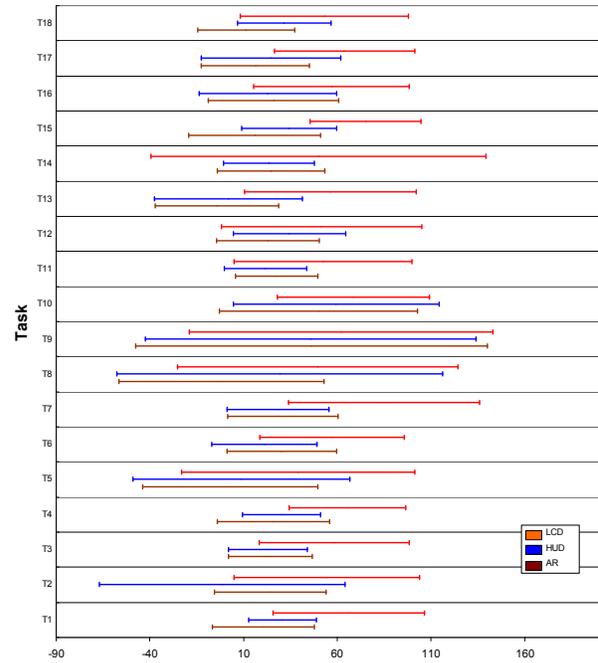


Figure 9: Ranges of head rotation (degrees yaw) for all participants across each task. Tasks are stacked in layers. Each task layer shows ranges for AR (bottom), HUD (middle), and LCD (top).

tions as the LCD sits within the range of absolute task azimuths.

A comparison of ranges by task, shown in Figure 9, provides a more revealing picture of the effect of display condition on head yaw. It should be noted that the range information includes transient head movements between tasks and thus range intervals are not necessarily centered on the target task. Despite this fact, the ranges exhibited by AR are lower than those exhibited by LCD in all but three tasks.

5.5.2 Rotational exertion and velocity

Rotational head exertion during each task was estimated for each participant by summing the change in head pitch, yaw, and roll Euler angles at each interval of the recorded data. Rotational velocity during each task was calculated for each participant by dividing this total rotational exertion in each axis by the time required to locate the task. Table 3 summarizes these statistics. A 3 (Display Condition) \times 18 (Task) repeated measure ANOVA was performed separately for each statistic along each axis, with participants as the random variable. In this analysis, display condition had a significant effect on pitch exertion ($F_{(2,34)}=12.206$, $p < 0.05$), roll exertion ($F_{(2,34)}=34.496$, $p < 0.05$), and yaw exertion ($F_{(2,34)}=32.529$, $p < 0.05$). Pairwise comparisons of mean rotational exertion about each axis are summarized in Table 3. For rotational velocity, display condition had a significant main effect on mean pitch velocity ($F_{(2,34)}=12.205$, $p < 0.05$), mean roll velocity ($F_{(2,34)}=48.875$, $p < 0.05$), and mean yaw velocity ($F_{(2,34)}=44.191$, $p < 0.05$). Table 3 captures the paired comparison of means.

5.5.3 Translational head exertion and velocity

Translational head exertion during each task was estimated for each participant by individually summing the change in Euclidean distance exhibited between each interval of the recorded data. The result represents the total Euclidean distances the head traveled during localization. A 3 (Display Condition) \times 18 (Task) repeated measure ANOVA test revealed a significant main effect of display condition on translational exertion ($F_{(2,34)}=17.467$, $p < 0.05$). The mean translational head exertions were 0.25 meters (AR), 0.36

Statistic	AR	HUD	LCD	Comparisons
<i>Mean Rotational Exertion (°)</i>				
Pitch	21.8	38.8	56.6	AR 43% less travel than HUD AR 62% less travel than LCD* HUD 32% less travel than LCD*
Roll	12.7	21.2	47.2	AR 40% less travel than HUD AR 75% less travel than LCD* HUD 55% less travel than LCD*
Yaw	42.9	50.2	122.3	AR 15% less exertion than HUD AR 65% less exertion than LCD* HUD 59% less exertion than LCD*
<i>Mean Rotational Velocity (°/s)</i>				
Pitch	4.2	2.8	4.4	AR 53% faster than HUD LCD 5% faster than AR LCD 61% faster than HUD
Roll	2.4	1.7	4.0	AR 48% faster than HUD* LCD 62% faster than AR LCD 140% faster than HUD*
Yaw	7.7	4.0	13.5	AR 95% faster than HUD* LCD 74% faster than AR* LCD 240% faster than HUD*

*statistically significant difference ($p < 0.05$)

Table 3: Rotational and Translation Exertions and Velocities

meters (HUD), and 0.68 meters (LCD) and are shown in Figure 10. A subsequent pairwise comparison of mean translational exertion revealed that AR required 29% less translational movement than HUD (not significantly significant), and 62% less translational head movement than LCD, which was statistically significant ($p=0.007$). HUD was 46% faster than LCD, which was statistically significant ($p=0.003$).

Translational head velocity was estimated for each participant by dividing total translational head exertion by the time required to locate the task. A 3 (Display Condition) \times 18 (Task) repeated measure ANOVA test revealed a significant main effect of display condition on translational velocity ($F_{(2,34)}=19.907$, $p < 0.05$). The mean translational head velocities were 0.05 meters/second (AR), 0.03 meters/second (HUD), and 0.08 meters/second (LCD) and are shown in Figure 11. A subsequent pairwise comparison of means revealed that AR was 60% faster than HUD, which was

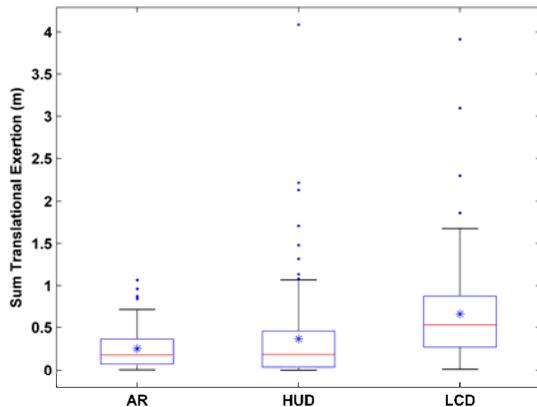


Figure 10: Translational head exertion (meters) during localization in AR, HUD, and LCD conditions. An asterisk marks the mean translational head velocity.

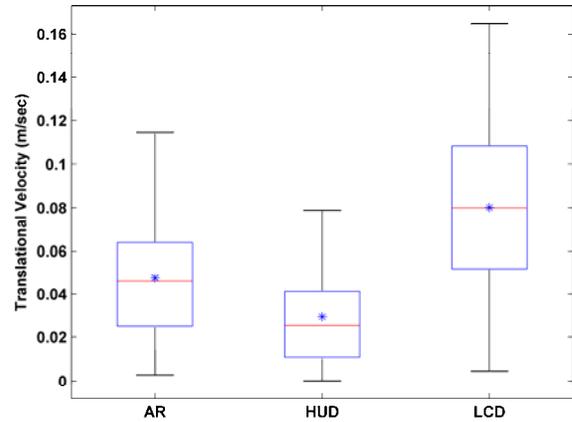


Figure 11: Translational head velocity (m/sec) during localization in AR, HUD, and LCD conditions. An asterisk marks the mean translational head velocity.

statistically significant ($p=0.019$). LCD was 40% faster than AR, which was statistically significant ($p=0.001$), and 62% faster than HUD, which was also statistically significant ($p=0.03$).

5.5.4 View Direction Distribution

We also examined the distribution of a user's view direction to understand the overall directionality of head exertion. Although proven metrics exist for this purpose, such as Distance from Center Point (DFCP), demonstrated by Axholt, Peterson, and Ellis [3], we confined our analysis to descriptive statistics. For each task in the experiment, we examined the distribution of normalized pitch and yaw angles exhibited by all participants. For pitch, these distributions were similar across display conditions and reflected a normal distribution. For yaw, we discovered more interesting differences. For AR and HUD, the distributions followed a normal distribution centered about the task azimuth. For LCD, this analysis revealed a distinctive bimodal distribution for most tasks. A comparative set of distributions is shown in Figure 12.

The presence of a second peak in the LCD condition was clearly a result of the mechanic spending time looking at the LCD. While this was not a surprise, the distributions help us quantify how and where participants accumulated head movements.

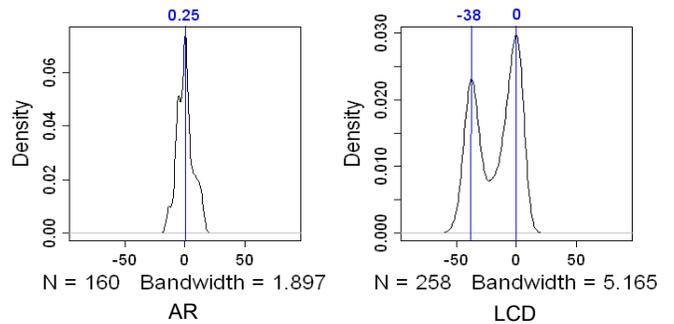


Figure 12: Distribution of normalized yaw angles for AR and LCD for Tasks T4. In each plot, the value $x=0$ indicates the population's mean yaw orientation.

6 QUALITATIVE RESULTS

We asked each participant to complete a post-experiment questionnaire. This questionnaire featured five-point Likert scale questions (1 = most negative, 5 = most positive) to evaluate ease of use, satisfaction level, and intuitiveness for each display condition. The summary results from these ratings, shown in Figure 13, are difficult to generalize given our small population size and individual rating systems. In terms of ease of use, median response for LCD (5.0) was highest, followed by AR (4.5) and HUD (3.5). For rating satisfaction, median response to AR (5.0) was highest, followed by LCD (4.0) and HUD (4.0). For rating intuitiveness, median response to AR (4.5) tied with LCD (4.5), followed by HUD (4.0).

When we asked participants to rank the techniques as to which they would rather use to perform a task, 4 of the 6 participants ranked the LCD condition first. A Friedman test indicated this was a significant ranking ($\chi^2_{(6,2)}=7.0$, $p=0.03$). Subsequent pairwise Wilcoxon test revealed that LCD was ranked significantly better than HUD ($p<0.02$). When asked to rank the techniques as to how intuitive they were, 4 of the 6 participants ranked the AR condition first. However, a Friedman test indicated this was not a significant ranking ($\chi^2_{(6,2)}=4.33$, $p=0.12$).

7 DISCUSSION

The quantitative and qualitative results reveal several encouraging findings. A statistically significant result showed the AR display condition allowed mechanics to locate tasks more quickly than when using the LCD display condition (representing an improved version of the IETMs currently employed in practice). Because AR was also faster at localization than HUD, we can conclude that the 3D registered localization information overlaid on the mechanic's view of the task (e.g., arrows, labels, and models) contributed to this result.

We were not surprised by the lack of statistical separation between the mean completion times for the AR and LCD display conditions. This can be explained, in part, by examining post-localization information requirements in what Neumann and Majoros define as the workpiece portion of a task [15]. This portion of a task is primarily kinesthetic and psychomotoric in nature and requires visual information supporting such activities as adjustment, alignment, and detection. While we did provide some alignment and routing information (e.g., content showing the position and movements of select tools and turret components), this was common to both AR and LCD. Moreover, this information was based on an object's static position in the turret, as we did not dynamically track tools or components in response to user actions.

Therefore, once a mechanic began physically manipulating objects in a task, they tended to not require information provided by the display. This was especially noticeable in more demanding tasks, such as installing a fastener that was partially occluded from view and requiring very precise alignment of movable surfaces before tightening. These tasks typically required mechanics to explore and manipulate the task from several different vantage points. During these more extensive physical manipulations, the LCD condition offered the clearest and widest view, which we suspect allowed the mechanics to complete this portion more quickly. More research with various types of HWDs, potentially including the tracking of task components, user hands, and tools, is required to articulate the advantages AR can afford during this portion of a maintenance task vis-a-vis display capabilities.

An additional statistically significant result of the study showed that the AR display condition allowed mechanics to incur fewer translational and rotational head movements at lower velocities than the LCD display condition during task localization. Descriptive statistics show that, in general, subjects experiencing the AR

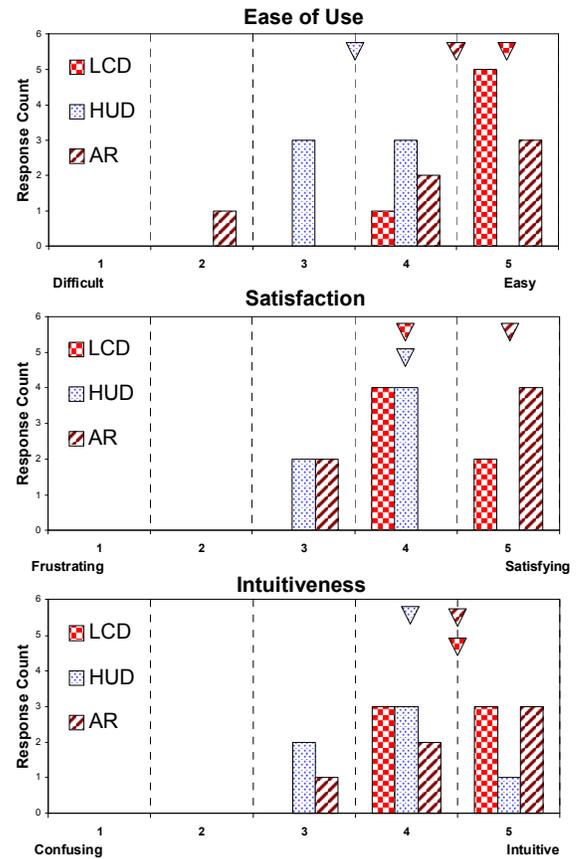


Figure 13: Survey response histograms by condition for ease of use (top), satisfaction (middle), and intuitiveness (bottom). Median values for each condition are shown as triangles.

condition also required smaller ranges of head movement. This result highlights the ability of AR to potentially reduce overall musculoskeletal workloads and strain related to head movement during maintenance tasks. However, more work is required to reconcile strain reductions afforded by reduced head and neck movement with the added strain of wearing a HWD. A technique proposed by Tümler and colleagues [29], which uses heart rate variability to measure strain, could be useful for this analysis.

Our qualitative results provide additional encouragement for the application of AR to maintenance tasks. Despite the disadvantage of wearing a bulky, relatively low-resolution prototype VST HWD with fixed focus cameras, and a narrow FOV, participants rated the AR condition at least as favorably as LCD in terms of satisfaction and intuitiveness. Future AR systems employing lighter, more comfortable displays with wider FOVs and higher resolutions could improve these results. Mechanics also provided some interesting written remarks when responding to our survey. While many participants acknowledged the visibility constraints experienced while using AR, they dismissed this limitation with profuse appreciation for the 3D arrows, labels, and animated sequences. Several participants mentioned the potential utility of the tool in maintaining hydraulic and electrical systems in particular.

8 CONCLUSIONS AND FUTURE WORK

We were pleased that our prototype AR application proved more effective than an enhanced baseline system at reducing time and head movement during the localization portion of maintenance tasks. We were especially encouraged to achieve these results with a population of professionally-trained mechanics working in a field setting, who expressed support for our approach.

In the future, we plan to repeat a version of the experiment in a less cramped maintenance environment that would allow us to use the NVIS nVisor HWD described in Section 3. This would provide participants experiencing the AR display condition a clearer and wider view during the workpiece portion of the task, which we hypothesize will reduce the overall completion time for AR. We plan to incorporate tracked physical objects under repair to provide additional information to assist mechanics as they physically manipulate the task environment. We are also interested in studying the benefits of AR for localizing users who are performing tasks that require substantial locomotion between tasks.

In closing, we have presented the results of a user study exploring the use of a prototype AR application supporting professional mechanics in a field setting. The prototype AR application allowed a population of mechanics to locate individual tasks in a maintenance sequence more quickly than when using an improved version of currently employed methods. The prototype AR application also resulted in mechanics making fewer overall head movements during localization and was favorably received by users in terms of intuitiveness and satisfaction.

ACKNOWLEDGEMENTS

This research was funded in part by ONR Grant N00014-04-1-0005 and generous gifts from NVIDIA and Google. We thank Bengt-Olaf Schneider, who provided the StereoBLT SDK used to support the display of stereo camera imagery, and Kyle Johnsen, who advised on use of the OptiTrack system. We are grateful for the assistance of cadre and students at Aberdeen Proving Ground, as well as the support of engineers at the Marine Corps Logistics Base including Mike Shellem, Curtis Williams, Andrew Mitchell, and Alan Butterworth. We also thank David Madigan and Magnus Axholt for insights shared during design and analysis of our experiment. Permission to use the LAV-25 photos shown in this paper was granted by the USMC LAV Program Manager.

REFERENCES

- [1] *Interactive Electronic Technical Manuals - General Content, Style, Format and User-Interaction Requirements*, Department of Defense, 2007.
- [2] "ARTESAS - Advanced Augmented Reality Technologies for Industrial Service Applications," <http://www.artesas.de>, May, 2009.
- [3] M. Axholt, S. Peterson, and S.R. Ellis, "User Bovesight Calibration Precision for Large-format Head-up Displays," *Proc. Virtual Reality Software and Technology (VRST '08)*, 2008, pp. 141-148.
- [4] K.M. Baird and W. Barfield, "Evaluating the Effectiveness of Augmented Reality Displays for a Manual Assembly Task," *Virtual Reality*, vol. 4, 1999, pp. 250-259.
- [5] F. Biocca, A. Tang, C. Owen, and F. Xiao, "Attention Funnel: Omnidirectional 3D Cursor for Mobile Augmented Reality Platforms," *Proc. CHI (CHI '06)*, 2006, pp. 1115-1122.
- [6] T.P. Caudell and D.W. Mizell, "Augmented Reality: An Application of Heads-up Display Technology to Manual Manufacturing Processes," *Proc. Int. Conf. on System Sciences*, Hawaii: 1992, pp. 659-669 vol.2.
- [7] J.M. Chambers, *Graphical Methods for Data Analysis*, Duxbury Resource Center, 1983.
- [8] D. Curtis, D. Mizell, P. Gruenbaum, and A. Janin, "Several Devils in the Details: Making an AR Application Work in the Airplane Factory," *Proc. Int. Workshop on Augmented Reality (IWAR '98)*, 1999, pp. 47-60.
- [9] S.M. Donelson and C.C. Gordon, "1995 Matched Anthropometric Database of US Marine Corps Personnel: Summary Statistics," 1996.
- [10] S. Feiner, B. Macintyre, and D. Seligmann, "Knowledge-based Augmented Reality," *Communic. ACM*, vol. 36, 1993, pp. 53-62.
- [11] W. Friedrich, "ARVIKA-Augmented Reality for Development, Production and Service," *Proc. Int. Symp. on Mixed and Augmented Reality (ISMAR '02)*, 2002, pp. 3-4.
- [12] C.C. Gordon, S.P. Paquette, J.D. Brantley, H.W. Case, and D.J. Gaeta, *A Supplement to the 1995 Matched Anthropometric Database of US Marine Corps Personnel: Summary Statistics*, U.S. Army, 1997.
- [13] C. Knöpfle, J. Weidenhausen, L. Chauvigné, and I. Stock, "Template Based Authoring for AR Based Service Scenarios," *Proc. IEEE Virtual Reality (VR '05)*, 2005, pp. 249-252.
- [14] A.M. Law and W.D. Kelton, *Simulation Modeling and Analysis*, McGraw-Hill Higher Education, 1997.
- [15] U. Neumann and A. Majoros, "Cognitive, Performance, and Systems Issues for Augmented Reality Applications in Manufacturing and Maintenance," *Proc. IEEE Virtual Reality (VR '98)*, 1998, pp. 4-11.
- [16] S. Nilsson and B. Johansson, "Fun and Usable: Augmented Reality Instructions in a Hospital Setting," *Proc. of the 19th Australasian Conf. on Computer-Human Interaction*, 2007, pp. 123-130.
- [17] J.J. Ockerman and A.R. Pritchett, "Preliminary Investigation of Wearable Computers for Task Guidance in Aircraft Inspection," *Proc. Int. Symp. on Wearable Computers (ISWC '98)*, 1998, pp. 33-40.
- [18] S.K. Ong, M.L. Yuan, and A.Y.C. Nee, "Augmented Reality Applications in Manufacturing: A Survey," *Int. Journal of Production Research*, vol. 46, 2008, pp. 2707-2742.
- [19] J. Platonov, H. Heibel, P. Meier, and B. Grollmann, "A Mobile Markerless AR System for Maintenance and Repair," *Proc. Int. Symp. on Mixed and Augmented Reality (ISMAR'06)*, 2006, pp. 105-108.
- [20] A. Raczynski and P. Gussmann, "Services and Training Through Augmented Reality," *Proc. European Conf. on Visual Media Production (CVMP '04)*, 2004, pp. 263-271.
- [21] D. Reiners, D. Stricker, G. Klinker, and S. Müller, "Augmented Reality for Construction Tasks: Doorlock Assembly," *Proc. Int. Workshop on Augmented Reality (IWAR '98)*, 1999.
- [22] C.M. Robertson, B. MacIntyre, and B.N. Walker, "An Evaluation of Graphical Context when the Graphics are Outside of the Task Area," *Proc. Int. Symp. on Mixed and Augmented Reality (ISMAR '08)*, 2008, pp. 73-76.
- [23] T. Salonen and J. Sääski, "Dynamic and Visual Assembly Instruction for Configurable Products Using Augmented Reality Techniques," *Advanced Design and Manufacture to Gain a Competitive Edge*, 2008, p. 23.
- [24] B. Schwald and B.D. Laval, "An Augmented Reality System for Training and Assistance to Maintenance in the Industrial Context," *Journal of WSCG*, vol. 11, 2003, pp. 425-432.
- [25] B. Schwerdtfeger and G. Klinker, "Supporting Order Picking with Augmented Reality," *Proc. Int. Symp. on Mixed and Augmented Reality (ISMAR '08)*, 2008, pp. 91-94.
- [26] A. Smailagic and D. Siewiorek, "User-centered Interdisciplinary Design of Wearable Computers," *SIGMOBILE Mob. Comput. Commun. Rev.*, vol. 3, 1999, pp. 43-52.
- [27] A. Tang, C. Owen, F. Biocca, and W. Mou, "Comparative Effectiveness of Augmented Reality in Object Assembly," *Proc. CHI (CHI '03)*, 2003, pp. 73-80.
- [28] M. Tönnis and G. Klinker, "Effective Control of a Car Driver's Attention for Visual and Acoustic Guidance Towards the Direction of Imminent Dangers," *Proc. Int. Symp. on Mixed and Augmented Reality (ISMAR '06)*, 2006, pp. 13-22.
- [29] J. Tümler, R. Mecke, M. Schenk, A. Huckauf, F. Doil, G. Paul, E. Pfister, I. Böckelmann, and A. Roggentin, "Mobile Augmented Reality in Industrial Applications: Approaches for Solution of User-related Issues," *Proc. Int. Symp. on Mixed and Augmented Reality (ISMAR '08)*, 2008, pp. 87-90.
- [30] U.S. Army, *Maintenance Operations and Procedures (Field Manual 4-30.3)*, 2007.
- [31] U.S. Marine Corps, *Light Armored Vehicle Operator's Manual (LAV-25A1)*, 2003.
- [32] J. Wither, S. DiVerdi, and T. Höllerer, "Evaluating Display Types for AR Selection and Annotation," *Proc. Int. Symp. on Mixed and Augmented Reality (ISMAR '07)*, 2007, pp. 1-4.
- [33] J. Zauner, M. Haller, A. Brandl, and W. Hartman, "Authoring of a Mixed Reality Assembly Instructor for Hierarchical Structures," *Proc. Int. Symp. on Mixed and Augmented Reality (ISMAR '03)*, 2003, pp. 237-246.