U.S. Department
of Transportation
Federal Railroad
Administration

Office of Research,
Development and Technology
Washington, DC 20590

Augmented Reality for Railroad Operations Using Head-up Displays

DOT/FRA/ORD-23/20

Final Report | August 2023

REPORT DOCUMENTATION PAGE	*Form Approved* *OMB No. 0704-0188*

1. REPORT DATE *(DD-MM-YYYY)* **8/11/23**	**2. REPORT TYPE** Technical Report	**3. DATES COVERED** *(From - To)* April 2019–October 2021

4. TITLE AND SUBTITLE
Augmented Reality for Railroad Operations Using Head-up Displays

5a. CONTRACT NUMBER
693JJ6-18-C-000010

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)
Andrew M. Liu: 0000-0001-9184-6527, Charles M. Oman: 0000-0002-5576-3510
Allison Zhang: 0000-0002-1136-1683, Amelia Gagnon: 0000-0003-2326-3260, ,
Kirk H. Mathews*:0000-0002-3068-3194, James Brooks*: 0000-0002-8395-7494,
Bradford Miller *: 0000-0001-6985-5769, Saikat Ray Majumder*: 0000-0002-5845-0472, Sarah Felix*: 0000-0002-9048-0014, Kaivalya Bakshi: 0000-0001-6980-0311,*

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Human Systems Laboratory
Massachusetts Institute of Technology
Cambridge, MA 02139
*GE Global Research, Niskayuna, NY 12309

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
U.S. Department of Transportation
Federal Railroad Administration
Office of Railroad Policy and Development
Office of Research, Development and Technology
Washington, DC 20590

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSOR/MONITOR'S REPORT NUMBER(S)
DOT/FRA/ORD-23/20

12. DISTRIBUTION/AVAILABILITY STATEMENT
This document is available to the public through the FRA website.

13. SUPPLEMENTARY NOTES
COR: Michael Jones

14. ABSTRACT
A team from MIT’s Human Systems Laboratory designed the locomotive HUD as a wide field of view augmented reality head-up display (AR-HUD). The technical feasibility of an AR-HUD was assessed through literature review and hardware tests. To study human factors issues, an AR-HUD prototype was designed, reviewed by experienced engineers, then implemented in the FRA Cab Technology Integration Laboratory simulator. The engineers’ behavior was not significantly altered and using the AR-HUD reduced the time spent looking away from the forward view. Subjective feedback from the engineers confirmed the acceptability and potential benefit of using HUDs.

15. SUBJECT TERMS
Augmented reality, head-up displays, eye tracking, head tracking, locomotive, cognitive evaluation, train simulator, human factors

16. SECURITY CLASSIFICATION OF:			**17. LIMITATION OF ABSTRACT**	**18. NUMBER OF PAGES**	**19a. NAME OF RESPONSIBLE PERSON**
a. REPORT	**b. ABSTRACT**	**c. THIS PAGE**		55	**19b. TELEPHONE NUMBER** *(Include area code)*

Standard Form 298 (Rev. 8/98)
Prescribed by ANSI Std. Z39.18

METRIC/ENGLISH CONVERSION FACTORS

ENGLISH TO METRIC

LENGTH (APPROXIMATE)

1 inch (in) = 2.5 centimeters (cm)
1 foot (ft) = 30 centimeters (cm)
1 yard (yd) = 0.9 meter (m)
1 mile (mi) = 1.6 kilometers (km)

AREA (APPROXIMATE)

1 square inch (sq in, in^2) = 6.5 square centimeters (cm^2)
1 square foot (sq ft, ft^2) = 0.09 square meter (m^2)
1 square yard (sq yd, yd^2) = 0.8 square meter (m^2)
1 square mile (sq mi, mi^2) = 2.6 square kilometers (km^2)
1 acre = 0.4 hectare (he) = 4,000 square meters (m^2)

MASS - WEIGHT (APPROXIMATE)

1 ounce (oz) = 28 grams (gm)
1 pound (lb) = 0.45 kilogram (kg)
1 short ton = 2,000 pounds (lb) = 0.9 tonne (t)

VOLUME (APPROXIMATE)

1 teaspoon (tsp) = 5 milliliters (ml)
1 tablespoon (tbsp) = 15 milliliters (ml)
1 fluid ounce (fl oz) = 30 milliliters (ml)
1 cup (c) = 0.24 liter (l)
1 pint (pt) = 0.47 liter (l)
1 quart (qt) = 0.96 liter (l)
1 gallon (gal) = 3.8 liters (l)
1 cubic foot (cu ft, ft^3) = 0.03 cubic meter (m^3)
1 cubic yard (cu yd, yd^3) = 0.76 cubic meter (m^3)

TEMPERATURE (EXACT)

[(x-32)(5/9)] °F = y °C

METRIC TO ENGLISH

LENGTH (APPROXIMATE)

1 millimeter (mm) = 0.04 inch (in)
1 centimeter (cm) = 0.4 inch (in)
1 meter (m) = 3.3 feet (ft)
1 meter (m) = 1.1 yards (yd)
1 kilometer (km) = 0.6 mile (mi)

AREA (APPROXIMATE)

1 square centimeter (cm^2) = 0.16 square inch (sq in, in^2)
1 square meter (m^2) = 1.2 square yards (sq yd, yd^2)
1 square kilometer (km^2) = 0.4 square mile (sq mi, mi^2)
10,000 square meters (m^2) = 1 hectare (ha) = 2.5 acres

MASS - WEIGHT (APPROXIMATE)

1 gram (gm) = 0.036 ounce (oz)
1 kilogram (kg) = 2.2 pounds (lb)
1 tonne (t) = 1,000 kilograms (kg)
= 1.1 short tons

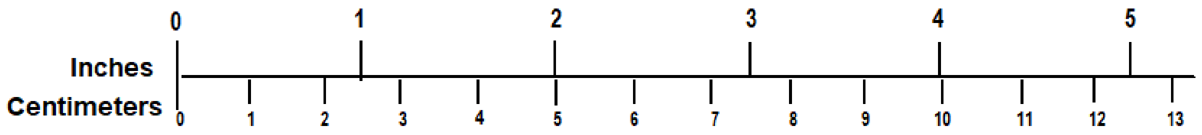
VOLUME (APPROXIMATE)

1 milliliter (ml) = 0.03 fluid ounce (fl oz)
1 liter (l) = 2.1 pints (pt)
1 liter (l) = 1.06 quarts (qt)
1 liter (l) = 0.26 gallon (gal)
1 cubic meter (m^3) = 36 cubic feet (cu ft, ft^3)
1 cubic meter (m^3) = 1.3 cubic yards (cu yd, yd^3)

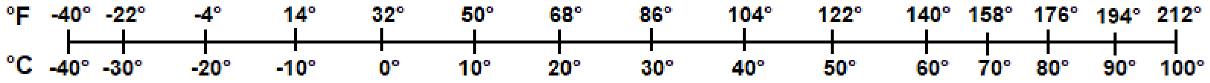
TEMPERATURE (EXACT)

[(9/5) y + 32] °C = x °F

QUICK INCH - CENTIMETER LENGTH CONVERSION



QUICK FAHRENHEIT - CELSIUS TEMPERATURE CONVERSION



For more exact and or other conversion factors, see NIST Miscellaneous Publication 286, Units of Weights and Measures. Price $2.50 SD Catalog No. C13 10286

Updated 6/17/98

Acknowledgements

The authors wish to thank Matt Isaacs, Jared Young, and Lafcadio Flint at the Volpe National Transportation Systems Center for their help in developing the testbed at the Cab Technology Integration Laboratory and for conducting the simulator studies. The authors also thank George Newman for his assistance in reviewing augmented reality head-up display designs and recruiting engineers for the simulator study. Finally, the authors thank the eight engineers who participated in the cognitive evaluation and simulator studies for providing their expertise.

Contents

Illustrations

Tables

Executive Summary

The topic of locomotive head-up displays (HUDs) has not been comprehensively reviewed since 2007, when Davies et al. examined the benefits of adapting aviation-style narrow field of view (FOV) HUDs for locomotives. The authors concluded that, though there were many safety benefits, there were also significant technical obstacles to integrating the HUDs into a locomotive. Since that time, new research and development activities on locomotive and automotive HUDs have continued, including the announcement of an augmented reality head-up display (AR-HUD) in a production vehicle which projects virtual images into the external scene. However, these implementations still retain the basic limited FOV format. Recent releases of commercially available, large transparent displays may provide a platform for a display that encompasses the entire front windscreen of a locomotive. From September 2018 to September 2021, the Federal Railroad Administration (FRA) sponsored the Massachusetts Institute of Technology (MIT) and GE to design a wide FOV locomotive AR-HUD and complete two research activities that assess the human factors and necessary hardware to realize this concept.

Engineers designed a prototype, then implemented it in the Cab Technology Integration Laboratory (CTIL) locomotive simulator as a testbed to investigate whether the AR-HUD concept would provide performance and/or safety benefits to the engineer in normal railroad operations. Five experienced passenger engineers participated in a user study, drove two versions of a locomotive HUD, and provided feedback on their design and use. Researchers recorded the participants' gaze, and analyzed how gaze behavior changed with the introduction of the HUD. The results showed that participants spent less time looking away from the forward view. The conformal information presented in the AR-HUD might be useful to direct engineer attention to important track-side objects, such as upcoming signals. Keeping gaze in the forward view, even if it is focused on HUD information, is crucial for the engineer to detect and identify potential hazards. Engineers will continue to analyze trip data to provide additional insight into the effects of HUD use. Future work should investigate whether safety benefits are realized when using a HUD. Analysis of train handling performance also suggests that there is no detrimental effect of using a HUD while operating a train. Subjective feedback from the engineers confirms the acceptability and potential benefit of using a HUD.

A survey of the current technology literature and product announcements shows that there are numerous promising developments in display technology that could support AR-HUD development, but they do not yet meet technical requirements for an AR-HUD rail application. Transparent organic light-emitting diode (OLED) displays are now available in sizes that match freight locomotive window widths and are available for commercial purchase but are relatively expensive and have limited light transmissibility. Display modules for automotive AR-HUDs are inexpensive but project a limited FOV. Further technical developments are needed for both technologies. Commercially available tracking systems, which are needed to accurately register HUD images with the external scene, are inexpensive and widely available. One widely available tracker system, the Intel RealSense D435 depth tracking camera, was integrated into a hardware testbed to test system accuracy and the measured performance was deemed acceptable for AR-HUDs.

The results from the human factors study and technical assessment do not present any immediate concerns that would suggest halting AR-HUD research and development. However, additional technology and human factors questions must be addressed in a more realistic testbed (e.g., in an

operating locomotive) before the beginning of commercial product development. The CTIL test represents an ideal case where AR-HUD symbology is always presented at the proper location. A real-world testbed is needed to understand the role of virtual image distances on engineer perception and acceptance. This knowledge will determine the type of display best suited for the AR-HUD and the need for additional hardware or optics. Additionally, the display technology should be tested under typical operating conditions (e.g., night, bright sunlight, etc.) to determine whether existing hardware can meet the brightness and contrast requirements.

1. Introduction

From September 2018 to September 2021, the Federal Railroad Administration (FRA) sponsored a research team from the Massachusetts Institute of Technology's (MIT) Human Systems Laboratory, the General Electric (GE) Global Research Center, and the FRA Cab Technology Integration Laboratory (CTIL) at the Volpe National Transportation Systems Center. This research team developed and tested a proof-of-concept augmented reality head-up display (AR-HUD) for railroad operations within a simulation environment and examined whether it enhanced the engineer's safe and efficient operation of the train. A subsequent human-in-the-loop experiment measured how the display affected the operator's mental workload and situational awareness, as well as changes to the safe handling of the train.

1.1 Background

The proposed AR-HUD is a large-field display that uses the entire engineer-side front windshield of the locomotive to display train information to the engineer during a trip. Possible methods to display the computer-generated symbology include large transparent organic light-emitting diode (OLED) displays or projection onto the windshield through combining optics. As shown in Figure 1, the AR-HUD has two types of symbology: (1) *Fixed* symbology remains in the same location on the display and, in general, shows train state information, such as the current speed, track speed limit, and location by milepost; and (2) *Conformal* symbology provides information that is connected to a physical object or location in the external environment, such as an upcoming signal or the beginning of a speed restriction, and changes location as the train approaches the object or location. Because locomotive engineers can freely move about the cab while the train is in motion, the AR-HUD also tracks the engineer's head position to draw the conformal symbology at the appropriate location to appear stable in the external environment. This feature differentiates the AR-HUD from current aviation and automotive head-up-displays (HUDs) which require the operator's head to remain within a small area to properly render the conformal symbology.

Figure 1. The AR-HUD concept. Conformal symbology highlights an approaching signal while static elements provide speed and location information

The presentation of critical information in the vehicle operator's forward view of the external environment is presumed to enhance safety because the operator's gaze is not diverted by looking at the physical displays and controls. However, careful consideration of the information presented in a "head-up" manner is required to avoid scene clutter, occlusion of the external environment, and excessive mental workload. Developing a proof-of-concept enables exploration of general scientific, technical, and design issues, such as testing the validity of the safety-enhancement assumption or defining the requirements for content and format of the information to be displayed.

1.2 Objectives

The project objectives were to (1) create a HUD interface that incorporates augmented reality (AR) symbology using feedback from experienced engineers, (2) develop a prototype system that could be demonstrated in the CTIL simulator, (3) evaluate the utility of the AR-HUD prototype with experienced engineers in CTIL and determine how driving performance and gaze patterns may change, and (4) assess the technical feasibility of producing an AR-HUD with current technology.

1.3 Overall Approach

The project team took a clean-sheet design approach to the AR-HUD interface by examining typical operating tasks through previously completed task analyses (Groshong, 2016) and determining the required information for those tasks. A cognitive walk-through of the preliminary design with an experienced locomotive engineer was performed to refine the design. The second design was evaluated through an on-line survey with two engineers, resulting in a third design that was integrated into CTIL. Two variants of the AR-HUD interface were implemented in CTIL for the user evaluations, and compared to the users driving with no HUD present. Subjective evaluation of the AR-HUD complemented the quantitative performance measures and gaze tracking data that were recorded. In parallel to the prototype development, the team conducted a literature review and technical feasibility assessment of producing a locomotive AR-HUD from currently available technology.

1.4 Scope

The project examined the relative utility of a wide field of view (FOV) HUD and AR-HUD in a representative operational situation compared to the existing head-down operating display (HDD). The demonstration system implemented in CTIL was meant to evaluate the usefulness of the interface design as well as understand how gaze and vigilance may change with use of the AR-HUD. The design was intended primarily for use in freight locomotives, but the displayed information would also support most aspects of passenger train operations with appropriate format changes. Researchers did not implement in CTIL all the hardware capabilities of the system needed to support an AR-HUD in an actual operating locomotive, such as a head tracking system or the optics needed to project the symbology onto the external environment. Researchers also did not examine the effects of various environmental conditions, such as nighttime or inclement weather operations. Thus, the evaluation represents only the use of the AR-HUD under ideal viewing conditions. As part of the feasibility assessment, a laboratory testbed incorporating head tracking and a simple display was created as a proof-of-concept and to characterize tracking hardware performance.

1.5 Organization of the Report

Section 2 presents a brief literature review of efforts to create HUDs for rail applications as well as recent developments in the automotive and maritime industries using emerging technology for relevant applications. Section 3 details the methodology to create an initial “clean sheet” design of the AR-HUD. Section 4 describes the initial design in detail. Section 5 interprets two reviews with experienced railroad engineers to evaluate and refine the initial design. Section 6 describes the current implementation of the AR-HUD in CTIL. Section 7 reviews the human-in-the-loop experiment design and discusses the results from the study, including recommendations for design changes. Section 8 analyzes the technical feasibility of implementing an AR-HUD system into current freight locomotives using current and near future technology. Appendix A assesses the technical feasibility of implementing an AR-HUD using currently available technologies.

2. AR-HUD Literature Review

Davies et al. (2007) provided the first detailed examination of HUDs for rail application in the 2007 UK Rail Safety and Standards Board report, which examined the technical, safety, and business cases for aviation-style narrow FOV HUDs in locomotives. The authors performed a small user study to examine possible HUD benefits in speed maintenance and workload, including conformal symbology highlighting the next signal. The results showed no difference in speed maintenance and a slight decrease in subjective workload. Since then, two patent applications have been filed in the US for similar AR-HUD systems: one in 2014 by Porsch and Schaeper (2017) of Siemens AG (awarded in 2017) and a second in 2018 by Miglianico et al. (2018). The Siemens patent describes a ceiling-mounted projector that uses a portion of the front windshield to display the combined output from two projectors, one showing the forward scene captured by a separate camera and a second that projects the HUD overlay after the scene from the first camera is parsed by an "evaluation device." This implementation would eliminate the need to track engineer head position to maintain the display if the head is not at the assumed eye point.

The Alstom patent describes a system similar to current automotive HUDs mounted on the forward engineer's console with additional hardware to permit a greater range of adjustment for different driver heights. In addition, Agarwal, supported by Bombardier, developed a user interface concept for a European Train Control System HUD in a 2019 master's thesis at KTH Stockholm. The concept applied a design process using Scenario Development to quickly reiterate through scenarios, producing a useful mapping between mental activities of the engineer, their actions, and the technical function required of them to guide interface design (Agarwal, 2018). The proposed AR-HUD solution goes beyond these systems by using AR hardware for head tracking which will enable the engineer to move about the cab without compromising the accuracy of conformal symbology placement. This makes better use of the wide FOV display so that external objects further to the sides of the track can be highlighted.

Background research in other applications of AR, particularly automotive, provide insight to applications in locomotion AR. Kim et al. (2018) showed a significant improvement in the response time to decelerate the vehicle in response to a pedestrian collision warning displayed on an AR display when compared to a control condition. In addition, the authors identified improved decision making while using conformal information (i.e., pedestrian and projected location) compared to a text alert stating "BRAKE." While the response time for braking was slower with the conformal display compared to the text message, there were also fewer false positives and a reduced risk for accidents due to emergency braking. This was attributed to the conformal information providing more information about the immediacy and magnitude of braking required. Locomotive engineers may have increased situational awareness with the use of conformal symbology. Interestingly, in a different automotive HUD experiment, Tonnis et al. (2005) found that implementation of conformal symbology more ambiguously indicated the location of a threat, and that drivers were better able to locate the threat with a static top-down view of the car. Drawing both possible conclusions together suggests that the association between conformal symbology and a specific object or location must be unambiguous to realize the benefits of a conformal AR display over traditional HUDs.

AR interfaces have also been tested in maritime applications, such as entering a harbor. On the ship's bridge, the watch officers often move around the deck while commanding the vessel, so keeping conformal symbology stable in their view also requires head tracking. Holder and Motz (2014) examined one possible AR implementation using transparent wide FOV displays and identified similar technical issues described later in this report. Their conclusions from a user study in a simulation environment agreed with Davies et al. (2007), showing the benefits of reduced head down time and localization of information, especially in situations with reduced visibility, confined waters, and high-speed operations. Current commercially available AR systems such as the Raymarine ClearCruise AR[1] are only implemented on HDDs and use synthetic imagery superimposed on a video stream from ship mounted cameras (i.e., "synthetic vision systems"), comparable to what is available in aviation. Morgère et al. (2014) conducted another study for maritime applications to examine head-worn AR displays for merging course information, hazard or other sea marker information, and real time data together to reduce cognitive workload and increase safety.

Whereas maritime and automotive AR-HUDs are highly focused on navigation issues as well as collision avoidance, railroad engineers are primarily focused on vehicle control in the longitudinal dimension. Thus, the primary goal for the rail AR-HUD should be to provide situation awareness to the engineer to appropriately control the speed of the train at the present location and under the present operating rules. There are many system design decisions to consider for an AR application for locomotive engineers, which may further constrain the choices and implementation of AR-HUD symbology (e.g., should the display be monoscopic or stereoscopic display, should the symbology be fixed on the screen or conformal with the external world). Section 3 describes the approach used to identify the type and location of the information that should be present in the AR-HUD design.

---

[1] Raymarine, Clear Cruise [AR] Augmented Reality Display for Raymarine Axiom MFDs.

3. AR-HUD Design Methodology

The design goal for the AR-HUD was to create a secondary information display supporting the most common events, or scenarios, that engineers encounter during a typical trip. In a previous FRA-funded study, typical freight railroad operations were separated into key operational scenarios to develop automation systems that drive with the engineer (Brooks et al., 2022). The team developed a hybrid model combining aspects of cognitive work analysis, hierarchical task analysis, and an abstraction hierarchy to understand the information requirements of the engineer in these scenarios (Brooks et al., 2022). Table 1 lists these key scenarios and the underlying information requirements that form the initial operating contexts for the AR-HUD design process.

3.1 Scenarios

Table 1 classifies the graphic scenarios, represented by computer-generated imagery (CGI), that are used in the AR-HUD development.

Table 1. Key scenarios to design AR-HUD

Scenario #	**Description**
0	Train Operation in Absence of Other Scenarios
1	*Signal Response (non-Clear)* – Changing speed or stopping in response to a signal indication
2	*Temporary Speed Restriction* – Complying with a speed limit within the designated zone
3	*Maintenance-of-Way* – Approaching and passing through a work area as directed by the foreman
4	*Passing a grade crossing* – Approaching and passing through a grade crossing
5	*Stopping on a grade* – Speed management to stop at a designated location on a grade
6	*Meet and Pass* – Meeting a schedule to allow opposing traffic to pass on a single track

The four “key operational scenarios” identified in Brooks et al. (2022) were (1) responding to signal, (2) complying with temporary speed restrictions (TSRs), (3) complying with a maintenance-of-way (MOW) zone (i.e., work area), and (4) providing a warning at a grade crossing. In each of these cases, the engineer needs to know the location of the event and the current operating constraints (e.g., signal indication, speed limit, and rules for entering a MOW zone). The engineer could also benefit from the increased situational awareness provided by the HUD when stopping at a specific location on a grade or incline, or participating in a meet and pass that requires schedule information (i.e., managing traffic traveling in opposite directions on a single track). Finally, a baseline scenario was identified as the absence of the named scenarios, such as when the engineer is operating the train at a given speed without a change in acceleration or a need to respond to signals.

3.2 Determining Necessary Information to Display

Table 2 presents a list of the types of information engineers use when making operational decisions (e.g., notch setting, current speed, or mile marker). This includes information that is physically available from HDDs, including track charts, bulletins, or hand-written notes. The table also lists information that can be obtained from the external environment or synthesized entirely by the engineer based on previously acquired information. However, the information sources inside the locomotive require the engineer to look away from the forward scene. The AR-HUD may more readily provide the necessary information to engineers and thus reduce their workload and minimize time looking away from the external environment. A key design question is determining what information to show without cluttering the display or increasing the engineer's workload.

Table 2. Information required by engineer to safely operate a train

1. Acceleration	2. Acceleration Trend	3. Air Brake Setting	4. Brake Line Pressure - End
5. Brake Line Pressure - Head	6. Current Speed	7. Desired Speed	8. Display Grade
9. Does this brake configuration fit with the geography	10. Throttle/Dynamic Brake setting	11. Has the consist end passed a certain landmark/marker	12. Identify Broken Gate/Grade Crossing
13. Identify Grade/Speed Mismatch	14. Identify Intruders	15. Identify Landmarks	16. Identify Places Where Mode Change May Occur
17. Identify User-Inputted Markers	18. Location Limit Delta	19. On/Off Optimized Profile Indication	20. Other Train Locations
21. Potential Obstacle Locations	22. Previous Signal	23. Projected collision warning	24. Projected ETA for several upcoming markers
25. Signal Change	26. Signal Information	27. Signal Speed Limit	28. Time Until Pressurized
29. Trip Advisor Information	30. Weather	31. Where is the consist end	

3.3 Determining Operational Relevance of Information

To date, there is no formal design methodology for determining the minimum information to display. Thus, the authors developed an ad-hoc, quantitative process to rank the information by operational relevance. For each piece of information in Table 2, one of the authors assigned two values: (1) Information Acquisition Cost, which represented the engineer's mental effort required to acquire the information from its location (e.g., display gauge and environmental cue), and (2) Information Salience, which represented the importance of the information to the safe operation of the train in each scenario listed in Table 1. By calculating the product of the weight of the information source and the salience to a given scenario, the authors determined a utility value for each piece of information that could be included on the AR-HUD design. There is a physical constraint on how much information can be displayed on the AR-HUD, and this ranking provides a criterion for selecting the most important information to be displayed.

3.3.1 Information Acquisition Cost

In Table 2, the information sources were clustered into three sections: HDDs within the cab, information outside the cab in the external environment, such as signal aspects, and information synthesized from these sources. The acquisition cost reflects the effort spent retrieving the information and storing it until needed. Synthesized information, such as the expected stopping location, is computed from other cues (e.g., current terrain and brake settings), so the engineer must locate the underlying information before synthesizing and committing it to short-term memory. If the underlying variables change then the engineer must also recalculate the synthesized information. The researchers assigned a cost equal to 1, since obtaining this information requires a high mental workload. Any information obtained from an in-cab display (e.g., the brake line pressure) is always present in the same location and can be readily retrieved with a single glance. These information sources were assigned a value of 0.25, the lowest weight, since they are easily obtained and require no storage in memory. While information from these sources is important, it can easily be viewed by the engineer and therefore may be redundant if included again on the AR-HUD. Information that is obtained outside the cab, such as a mile marker or speed limit sign, is episodic and must be kept in the engineer's working memory or written down. The increase in workload to obtain and store this information resulted in an assigned weight of 0.5. These initial weights were selected to represent relative differences based on subjective estimates of workload, and could be further refined with data from workload experiments, such as eye movement studies.

3.3.2 Temporal Importance to a Task/Scenario

A second, separate weight was assigned to the information based on when it influences engineer actions during the scenario. The weights correspond to four separate categories of information: primary, secondary, tertiary and background. Primary information, given the highest weight of 1, is critical to determining how to proceed in a situation. If not available, the situation resolution cannot proceed. Using Scenario #1: Signal Response as an example, primary information would include both the past and upcoming signal aspect, in addition to whether the signal has changed. The engineer cannot act to increase, maintain, or decrease speed without this knowledge. Secondary information, which the engineer will use immediately after obtaining the primary information, was assigned the next highest weight of 0.5; it has a lower weighting since its interpretation is dependent on the context provided by the primary information. In Scenario #1, current train speed is one type of secondary information used immediately after determining the upcoming signal aspect as the engineer determines how to comply with the primary information. Tertiary information is directly relevant to the actions taken by the engineer in the scenario but not necessarily essential, therefore assigned a lower weight of 0.25. In Scenario #1, this would include information such as the landmark the engineer uses to identify when braking should commence. Finally, background information could alter the timing or magnitude of the engineer's actions but is not directly used to optimally complete the scenario at hand. In Scenario #1, the current weather (e.g., rain or snow) was considered background data and assigned a weight of 0.1. As in the case of the information acquisition cost, these weights were chosen to reflect the subjective importance of the different categories with an arbitrarily chosen interval difference. They should be refined in the future through additional studies of engineer behavior.

3.3.3 Results of Weighting

Table 3 lists the weights for acquisition cost and temporal importance to the individual scenarios listed in Table 1. The importance of an information source across several scenarios is evident when there are high weights assigned to multiple scenarios. Table 4 show the aggregated weight for each information source, obtained by taking the product of all the individual scenario and acquisition cost weights and scaling the values so they are all greater than or equal to 1. The sources are then listed from highest to lowest aggregate weight with several tiers of "equivalent" weight within the ranking.

Table 3. Intermediate results in weighting information sources

	Info Found	Scenario 1	Scenario 2	Scenario 3	Scenario 4	Scenario 5	Scenario 6
Acceleration	1	0.5	1	0.5	0.5	1	0.5
Acceleration Trend	0.25	0.5	1	0.5	0.5	1	0.5
Air Brake Setting	0.25	0.5	1	0.25	1	1	0.25
Brake Line Pressure - End	0.25	0.5	1	0.25	1	1	0.25
Brake Line Pressure - Head	0.25	0.5	1	0.25	1	1	0.25
Current Speed	0.25	0.5	1	1	0.5	1	1
Desired Speed Location	1	0.5	1	1	0.5	0.25	1
Display Grade	0.5	0.5	0.5	0.1	0.1	1	0.1
Does this brake configuration fit with the geography	1	0.5	1	0.25	1	1	0.25
Throttle/Dynamic Brake Setting	1	0.5	1	0.25	1	1	0.25
Has the consist end passed a certain landmark/marker	1	0.1	0.1	1	1	0.1	1
Identify Broken Gate/Grade Crossing Signal	0.5	0.1	0.1	0.1	1	0.1	0.1
Identify Grade/Speed Mismatch	1	0.1	0.25	0.25	0.25	0.5	0.25
Identify Intruders	0.5	0.25	0.25	1	1	0.25	0.25
Identify Landmarks	0.5	0.25	0.25	0.25	0.25	0.5	1
Identify Places Where Mode Change May Occur	1	0.1	0.1	0.25	0.1	0.1	0.1
Identify User-Inputted Markers	1	0.1	0.25	0.25	0.25	0.5	0.5
Location Limit Delta	1	0.5	1	1	0.5	0.25	0.5
On/Off Optimized Profile Indication	1	0.5	0.5	0.25	0.25	0.25	0.25
Other Train Locations	0.5	0.25	0.25	1	0.25	0.25	1
Potential Obstacle Locations	1	0.25	0.25	1	1	0.25	1
Previous Signal	1	1	0.25	0.25	0.5	0.25	0.25
Projected collision warning	1	0.25	0.25	1	1	0.25	1
Projected estimated time of arrival for several upcoming markers	1	0.1	0.25	0.25	0.25	0.25	1
Signal Change	1	1	0.5	0.25	0.5	0.25	0.25
Signal Information	0.5	1	0.25	0.25	0.5	0.25	0.25
Speed Limit	0.5	0.5	1	1	0.5	0.25	0.5
Time Until Pressurized	1	0.5	1	0.25	1	1	0.25
Trip Advisor Information	0.25	0.5	0.5	0.25	0.25	0.25	0.25
Weather	0.5	0.1	0.1	0.1	0.1	0.1	0.1
Where is the consist end	1	0.1	0.1	1	1	1	1

Table 4. Aggregate weights of information sources ranked from highest to lowest

	Weight *10^6	Rank
Current Speed	62,500	1
Desired Speed Location	62,500	1
Acceleration	62,500	1
Location Limit Delta	31,250	2
Throttle/Dynamic Brake Setting	31,250	2
Does this brake configuration fit with the geography	31,250	2
Time Until Pressurized	31,250	2
Speed Limit	15,625	3
Acceleration Trend	15,625	3
Projected collision warning	15,625	3
Potential Obstacle Locations	15,625	3
Where is the consist end	10,000	4
Air Brake Setting	7,813	5
Brake Line Pressure - Head	7,813	5
Brake Line Pressure - End	7,813	5
Signal Change	3,906	6
Previous Signal	1,953	7
Other Train Locations	1,953	7
Identify Intruders	1,953	7
Has the consist end passed a certain landmark/marker	1,000	8
Signal Information	977	9
On/Off Optimized Profile Indication	977	9
Identify Landmarks	977	9
Projected ETA for several upcoming markers	391	10
Identify User-Inputted Markers	391	10
Trip Advisor Information	244	11
Identify Grade/Speed Mismatch	195	12
Display Grade	125	13
Identify Broken Gate/Grade Crossing Signal	5	14
Identify Places Where Mode Change May Occur	3	15
Weather	1	16

4. AR-HUD Design

In the initial AR-HUD design, researchers placed only a few limitations on information representation, such as avoiding obstruction of the central view. In some instances, the conceptual design was representative rather than precise.

Five classes of information were defined to sort the placement of information display elements on the AR-HUD: (1) train system status information, (2) speed/acceleration information, (3) map information, (4) warning information, and (5) target speed information. A template was created for the AR-HUD (Figure 2) by combining these categories, the weighted importance of various pieces of information, and the team's past experience designing the moving map display (Voelbel, 2017). Train system status information was placed on the left side of the windscreen to provide spatial consistency with the layout of the AAR-105 control stand. Speed, acceleration, and warning information were placed centrally above the area where the engineer's gaze is most often directed. Map information was placed on the right edge of the display, and target speed information was overlaid on the track ahead of the train.

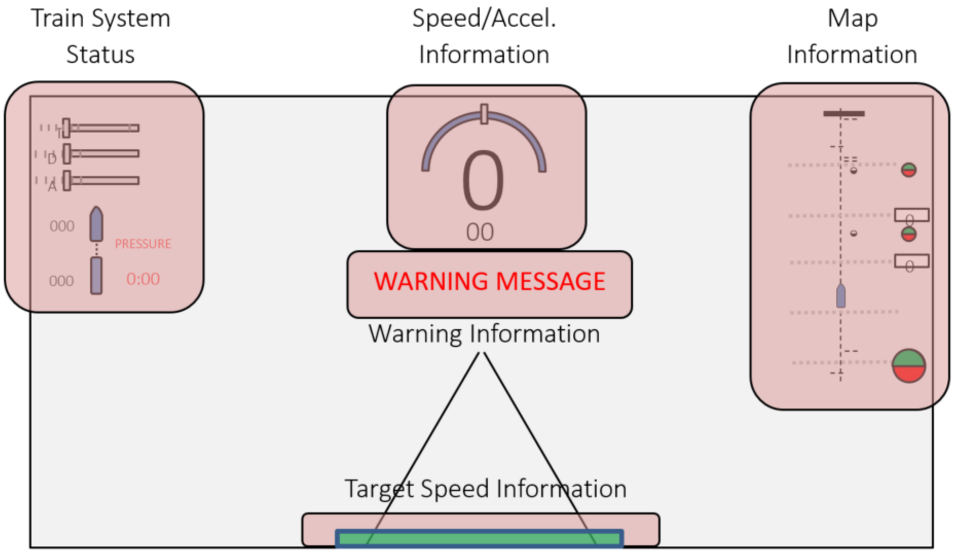


Figure 2. AR-HUD information display

4.1 Display Information Classes

The following information is included in the HUD display.

4.1.1 Train System Information

Primary train control status is displayed in a pictorial format that corresponds to the physical controls of the AAR-105 control stand. Brake information is displayed in a numeric format.

- *Notch Settings*

  Three sliders represent the position of the throttle, dynamic brake, and air brake handles (Figure 2). The left side of the icon is the lowest possible setting, while the right side represents the highest. This design choice corresponds with the physical controls for the two brake levers on the familiar AAR 105 console. They are arranged top to bottom in the order they are most often used, throttle position on the top, dynamic brake position in the middle, and the air brake position on the bottom. This icon is gray to be unobtrusive yet readily available (Figure 3).

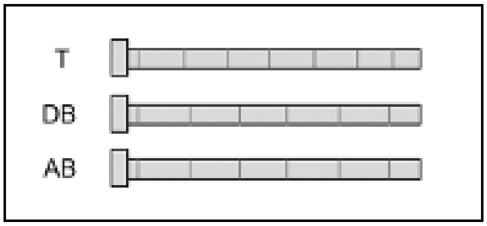


Figure 3. Display symbology for the throttle, dynamic brake, and air brake

- *Pressure Information*

  To convey brake line pressure information, a train icon representing the front and rear of the train is surrounded by pressure information (Figure 4). The upper car, which is pointed, signifies the front of the train while the box-shaped car at the bottom represents the end of the train. The consist head brake pipe (BP) pressure in psi is displayed next to the icon of the head of the train, and end-of-train (EOT) pressure in psi is displayed next to the icon at the rear of the train. Three additional pressures are displayed between BP and EOT: brake cylinder (BC), main reservoir (MR), and the equalizing reservoir (ER) pressure, needed if the consist uses traditional pneumatic brakes. Of these five pressures, the BP and EOT values are always visible. Since the BC, MR, and ER pressures only change when the air brake is being used, they will only appear on the AR-HUD when the pressures are not within the normal range of non-utilization. This helps keep a clean workspace during at-speed operations and alerts the engineer to any unexpected or abnormal conditions if the pressure information appears. Further investigation is needed to determine whether this is convenient or if engineers prefer to always see all five numbers on the AR-HUD, when this information is already found on the HDD. Additionally, if the line pressure is not at the maximum, a red PRESSURE warning will appear next to the consist icon. While this icon is redundant when the engineer actively applies the air brake, it will also trigger a warning to the engineer if there is a leak in the air line. When the air brake is released and the line is recharging, a countdown based off current recharging tendencies will appear to let the engineer know when they can expect full braking capabilities to return as the line recharges.

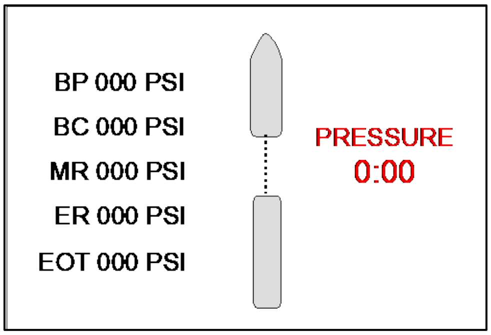


Figure 4. Display symbology for brake pressure information

4.1.2 Speed/Acceleration Information

Figure 5 shows the speed and acceleration information displayed at the top of the windscreen above the vanishing point of the view directly ahead of the train. This central location minimizes the distance for the engineer to change their gaze between the track ahead and the speed and acceleration information. It also copies the location of the speed information in most primary operating display screens and other secondary displays, such as the one used for GE Trip Optimizer.

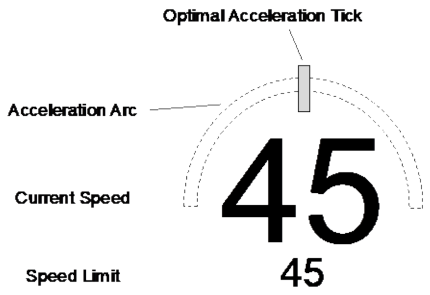


Figure 5. Speed, speed limit, and acceleration information

- *Current Speed*

  Current speed is critical for the engineer to monitor. During nominal operation, it is white in color. The speed numbers will change from white to yellow at 73-75 mph and from yellow to red at speeds greater than 75 mph to correspond with the GE Trip Optimizer display.

- *Speed Limit*

  The speed limit is shown below the current speed for situational awareness. It is smaller than the current speed indicator to avoid confusion with the train speed, and is displayed in white.

- *Acceleration Curves*

  When the train is at a constant speed, a short vertical line at the 12 o'clock position is displayed above the speed display. When the train is accelerating, an arc sweeps out clockwise from the center. When the train decelerates, the curve sweeps counterclockwise from the top. The maximum arc extent is 90 degrees. This corresponds to the movement of an analog speedometer, in which the needle moves clockwise when accelerating and counterclockwise when slowing down. The arc length of the curve in the AR-HUD design is proportional to the current, instantaneous acceleration of the locomotive up to a prescribed minimum or maximum. Some locomotive operator displays have a digital acceleration field that shows the projected increase in speed per time unit (typically mph/min) at the current acceleration. Anecdotally, engineers have mixed opinions on the usefulness of the acceleration display, which is possibly more useful during training or when driving an unfamiliar consist.

- *Desired Acceleration Tick Mark*

  This mark shows a "desired" acceleration defined by an acceleration or braking curve that could, minimize the use of air brakes or in-train forces. It could also be used as a training tool to help novice engineers learn the appropriate control changes to achieve the desired profile. When the train acceleration is close to the desired level, the tick mark will be near the end of the acceleration curve, and both symbols will be shown in green for ease of interpretation. If the acceleration curve is different from the tick mark, indicating an acceleration other than the directed one, the curve and tick mark will be red. In instances where neither acceleration nor deceleration is needed, (e.g., when the train is moving at the target speed) the indicator will be gray in color.

4.1.3 Map Information

The information pertaining to the train's surroundings is found on the right side of the windscreen. This design decision was made because when an engineer wants to know more information about their surroundings, they already look forward through the front windscreen and to the right through the right window. A top-down moving map view was designed to display this information, based on engineer feedback from researchers' previous work developing a tablet-based moving map display (Voelbel, 2017),. The format is familiar due to wide use in navigation displays (e.g., hand-held devices and automotive displays) and can be shown along the vertical side of the display. Typical train displays (e.g., i-ETMS and Trip Optimizer) use a profile map instead, which would need to be placed along the bottom or top. A description of the features of the moving map are below.

- *Scrolling Map Concept*

  As shown in Figure 6, the map design is a train-centric, scrolling map showing icons of upcoming landmarks approximately 5 miles in front of and 2 miles behind the current train location. The train icon stays in the same visual location while the landmark icons scroll down as the train travels along the track, similar to how a runner on a treadmill stays in one relative location while the treadmill scrolls beneath them. The scrolling map does not reflect curves or grade since this information is required knowledge for an engineer qualified on a route.

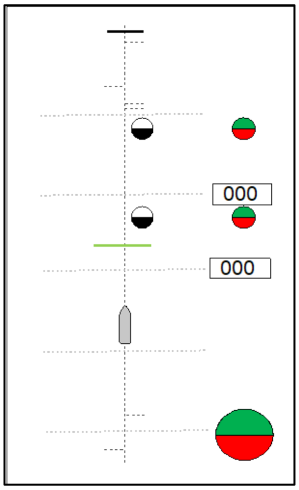


Figure 6. Example of the scrolling map

- *Selection of Data on the Scrolling Map*

  The criteria for displaying landmarks comes from two requirements: first, the landmark must be pertinent to the direct tasking of engineers; second, the landmark must be found throughout the rail system. This resulted in the display of the following list of landmarks: mile markers, interlock locations, grade crossing locations, and surrounding signal locations. Other landmarks that qualified for inclusion were curve lubricators and dragging equipment detectors, but they are not displayed because they do not affect train control (as verified by discussions with a retired engineer). Additionally, geographical features such as rivers or mountains were not included because they could most likely be seen from the cab and because the driver would not need to act on or be alerted to their presence.

- *Mile Marker Location*

  The iconography for each of these landmarks is designed to be a pictorial representation of the landmark itself (Figure 7). To indicate mile marker locations, a horizontal, gray dotted line is drawn across the dashed black line symbolizing the track at the location of the marker. The line extends to a mile marker icon which is represented by a black number over a white rectangular background, the same as displayed in the CTIL rail simulator. As the mile markers and mile marker location bars scroll through the map, only the upcoming 2-mile markers are visible, and once they scroll down the map past the train icon, they disappear.

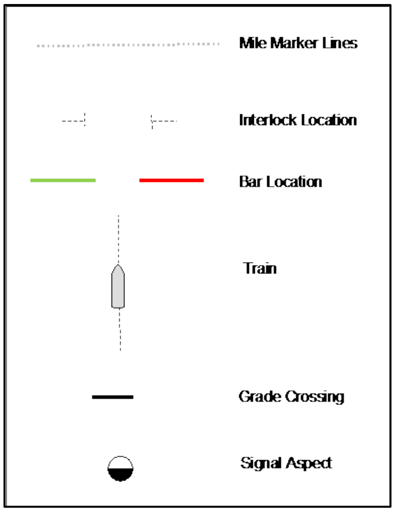


Figure 7. Iconography for the scrolling map

- *Interlock Locations*

  For interlock locations, the short arm extending to one side of the track indicates the side from which the interlock is entering or diverging. They are displayed as dashed lines, to remain consistent with the main track's dashed line display.

- *Grade Crossing Location Icon*

  Grade crossings are denoted by a bold, black line crossing the track. The thickness of the line provides a greater contrast with the track and other symbols crossing the track to enhance their detection since engineers are required to sound their horn when approaching most crossings.

- *Target Speed Predictor Icon*

  A green or red bar crossing the track marks the location where the train is predicted to reach a desired target speed based on current train speed, location and acceleration. This information is potentially useful when approaching a restricted speed area or stopping at a desired location (e.g., a stop signal). As the train approaches the target speed, the bar will be shown in green and move closer to the train icon, while it will be shown in red and move away from the train icon if train speed diverges from the target.

- *Signal Location Icon*

  An icon is placed next to the track to indicate a signal location. The icon, a circle with a white upper field and black lower field, was selected to provide a high contrast marker that would not be confused with the actual colors of the current signal indication.

- *Current and Future Signal Indication Icons*

  The actual indications displayed are in a separate, larger colorized icon to the right of the black and white signal aspect location icons. These are offset so they don't conflict with the scrolling map but are visible to give the engineer increased situational awareness about upcoming sections of track. The previous signal indication governing speed for the current signal block is also displayed in the lower right corner of the scrolling map. This alleviates the need for the engineer to remember current speed limits and is automatically updated as the train passes a signal.

4.1.4 Warning Information

Any warning messages that need to be conveyed to the engineer (e.g., a detected overspeed or an intruder/obstacle) are clearly indicated by a flashing red text message that appears below the speed and acceleration display. This position, as shown in Figure 8, is central in the field and near the area of the track where an engineer will often look for obstructions, switches, or other signals. The message text will succinctly state the reason for the warning using a known list of terms such as "overspeed" or "intruder."

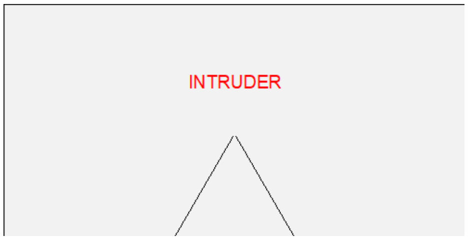


Figure 8. Warning example

4.1.5 Target Speed Information

The AR-HUD provides the benefit of showing conformal symbology that conveys information in relation to objects and their real-world location. This may be most useful with synthesized information which does not correspond to a physical object but is connected to a physical location (e.g., the projected location where the train will attain a desired speed). Depicting the predicted location of reaching this desired speed could help the engineer slow down to the appropriate speed in advance of a temporary speed restriction or even stopping at a red signal (target speed = 0). The system incorporates two symbols to indicate this synthesized information.

- *Target Speed Predictor*

  As mentioned in the description of the moving map display, the target speed predictor indicates the predicted location where the train will attain the desired speed based on the current acceleration. The icon, shown as a green bar in Figure 9, enables engineers to see precisely where they can expect to reach the desired speed in relation to signal aspects, grade crossings, or intruders. The color reflects whether the engineer will reach a desired speed before (green color) or after (red color) a mile marker or physical object along the track. The reference object could be determined by the engineer or via a machine vision system. In cases where the distance it takes to reach a desired speed is greater than 1 mile, the bar will be absent.

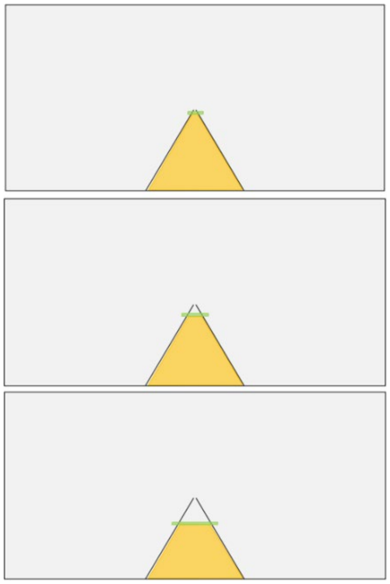


Figure 9. Example of Target Speed Predictor (green bar) and Headway (yellow) symbols as the locomotive approaches the target

- *Target Speed Headway*

  Target Speed Headway, or the distance between the locomotive and Target Speed Predictor location, is represented by a virtual flat surface overlaid on the track, extending from the front of the locomotive to the Target Speed Predictor (Figure 9) or terminating at the vanishing point of the track. It follows the contours of the track and serves to anchor the Predictor to the physical environment and to depict the perceived distance to the Predictor. As the overlay spans the width of the track, it is also used to indicate to the engineer that they are operating in a reduced speed area (yellow) or that they are clear to accelerate to the maximum speed limit. This is redundant to the digital speed limit below the current speed, however, this articulation uses both color and shape to make safe operating speeds clear to the operator.

4.1.6 Other Display Characteristics

Since the AR-HUD has information overlaid onto the windscreen, the information may sometimes be visually lost against the background of trees, buildings, or other objects in the environment. Likewise, the information displayed on the AR-HUD may overlap with objects in the environment that must always remain visible. The actual ranges of transparency (e.g., brightness and relative contrast to the background) will need to be determined through experiments in real operating conditions or a more detailed analysis. For this design effort, representative renderings of the AR-HUD were created using Microsoft PowerPoint to adjust the transparency of the symbology. These renderings, included in Appendix A, were used during the research team's review with a subject matter expert (see Section 5.1).

4.1.6.1 AR-HUD Transparency

- Information Transparency

  Any level of transparency of the displayed information will reduce the engineer's ability to clearly see other objects of interest, (e.g., signals or obstructions on the track). For example, even if Target Speed Predictor and Headway symbols are displayed at a low transparency, they may still obscure relatively dark objects likes switches and prevent the engineer from correctly assessing their setting. Alphanumeric symbols may need to be displayed at higher levels of brightness to be clearly visible and interpretable to the engineer. A future solution may be to use a machine vision system to determine the depth location of certain objects and render the AR-HUD symbology in a manner that avoids occlusion with the desired objects.

- Information Background Transparency

  AR-HUD-exterior environment conflicts are common in areas of dense foliage and tall surrounding landmarks, since it can be difficult to read information on the AR-HUD, even at 100 percent opacity. To increase the relative contrast, the team compared backgrounds of different colors and opacities using the PowerPoint renderings. The background must be large enough to depict the major groups of information (excluding the Target Speed display), with an opacity that allows for signals and other important objects and landmarks to be seen even when covered by the background. For the CTIL implementation, the researchers chose a sky blue color for the background to provide additional contrast with the AR-HUD information and minimal visual conflict with the exterior environment. A neutral gray background may be more suitable in practice as ambient lighting conditions change throughout the trip.

4.1.6.2 Declutter Mode and Go-Away Mode

The researchers created two ways for engineers to control what they see on the AR-HUD, since the display of the AR-HUD is automated. One method is Declutter Mode, which clears certain information from the AR-HUD to engineer preference. The minimal set could be determined by their weighted ranking, their location on the display (e.g., to remove information in the central field), or other criteria. Go-Away Mode clears all information from the display, and the windscreen again becomes a regular window. This is desirable if the AR-HUD is malfunctioning or obscuring visibility. These two modes could be controlled by switches on the main control console display.

4.1.6.3 Context-dependent Information Display

Some AR-HUD symbols could be shown only when the train is not in/at a desired state . When the engineer is driving the train at approximately the desired state, less information will be displayed, to potentially reduce distractions from monitoring the external environment. Examples and potential implementations of this property for warning messages, certain pressure information when braking, and the Target Speed indicators are discussed below.

- Warning Information Operating Rules

  The warning messages will only appear in emergency or unusual circumstances, such as when the train is speeding, and disappear when the train returns to a normal operating state.

- Brake Pressure Information Operating Rules

  The pressure information for various components of the brake system is not always required by the engineer. In instances when the braking system is not used, the pressures can be expected to remain within a non-braking operating range, so it may be acceptable to only show the brake pipe pressure. Other pressure information can be found on the HDD, if needed. Additional information, such as the ER pressure or brake recharge time, will appear during braking to assist the engineer with slowing the train and will remain visible until the brakes are fully recharged. This supports engineer situational awareness and helps the engineer maintain their gaze on the track ahead.

- Target Speed Information Operating Rules

  The inherent behavior of the Target Speed displays, particularly the Headway Indicator, provides a clear indication of whether the train is at the desired operating state. As the train diverges from the desired target speed, the Headway Indicator extends from the bottom portion of the screen to clearly indicate that engineer action is needed. As the train approaches the desired state, the indicator will shrink until disappearing as the Target Speed Predictor moves to a point just in front of the locomotive (e.g., the bottom of the display). This allows for a greater FOV when headway information is not necessary.

Figure 11 shows two renderings of the AR-HUD overlaid on representative scenes from CTIL simulation, illustrating what engineers see during a typical trip. Figure 10 shows the AR-HUD symbology overlaid on a sample screen capture from CTIL simulation runs and illustrates what engineers see during their approach to a stop signal. The train speed is 10 mph and decreasing at the prescribed deceleration shown by the green arc to the left. The Target Speed Predictor indicates that the stopping point on the track ahead (speed = 0 mph) will be reached before the signal. The Train Status Display shows that the throttle is at notch zero, the dynamic brake is fully engaged, and the air brake is also in use. Figure 11 shows the train is traveling at 55 mph. The Target Speed Predictor is placed at the front of the train on the bottom of the screen, indicating it has attained the target speed. The Train Status Display shows that the throttle is at notch three and no braking is in effect.

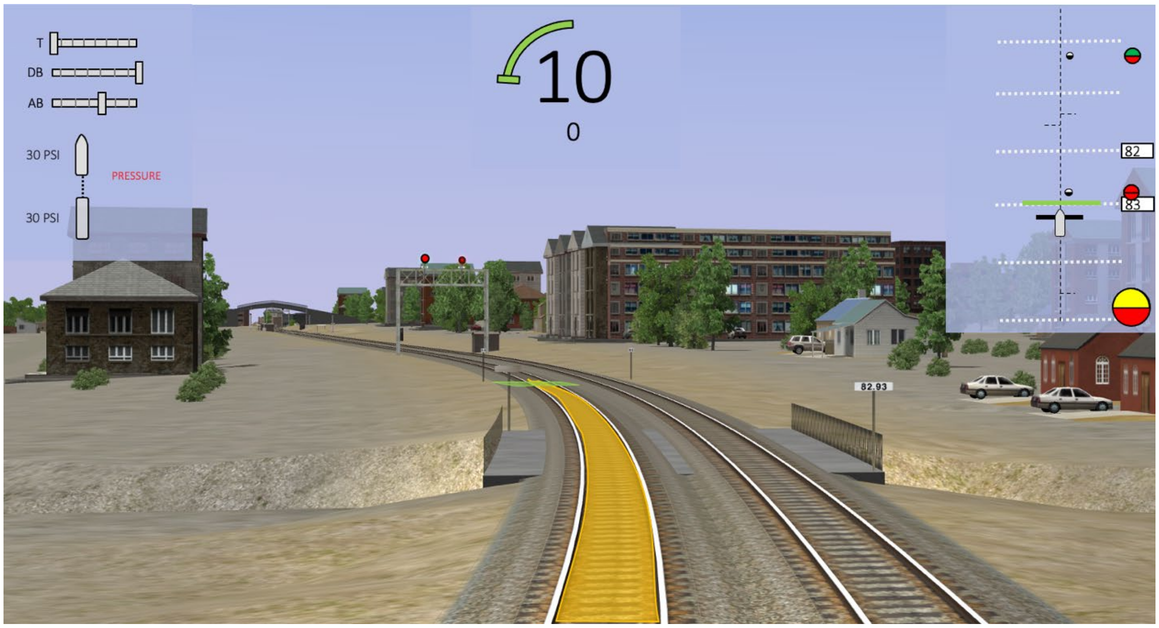


Figure 10. AR-HUD of train approaching stop signal

Figure 11. AR-HUD of train at the target speed

5. Evaluation of the Initial AR-HUD Design

5.1 Cognitive Walkthrough for Design Refinement

Following the completion of the initial design, a cognitive walkthrough was conducted with an experienced retired locomotive engineer. The cognitive walkthrough method is often employed in usability studies to evaluate the ease of learning to operate a system within the scope of specific user tasks (Wharton, J., Lewis, & Polson, 1994). Rather than providing explicit training before the walkthrough, the reviewers are encouraged to explore the interface to complete the task, revealing the elements that are easily understandable or usable. The typical walkthrough session begins with a detailed introduction to the system followed by one or more exercises in which a task scenario is presented to the reviewer, who must perform a specified series of actions to complete the task.

Since the AR-HUD is not an interface on which control actions can be performed, the research team's walkthrough focused on the operator's ease in obtaining the appropriate information to manage specific operating scenarios. The AR-HUD walkthrough began with an introduction to the larger project, followed by a detailed explanation of the individual display components. The hardware and underlying implementation of the AR-HUD was not provided as background information since the review was focused on information content and display. The behavior of the dynamic elements (e.g., the Target Speed Predictor) was explained through a series of PowerPoint mockup slides showing the evolution of the display over time. Unlike the process described in Wharton et al. (1994), feedback about each element was solicited after the description. Following the system description, the reviewer was presented with a generic scenario of transiting a section of mainline track with one interlocking and stop signal. A series of slides depicting the progress of the train was showed in succession with feedback from the engineer about the display and system behavior. Following a short break, a second operating scenario was presented to the engineer in the same manner, except the engineer was asked to answer three to four questions about the state of the train based on the images. Additional feedback about the scenario and display behavior was discussed after the exercise. The third and final session consisted of showing a series of single slides depicting a moment during a trip. The current train state (e.g., speed and last signal) was described to the engineer followed by a series of questions to probe the engineer's knowledge about the present and future states and actions that could be obtained from the display. A final discussion period allowed the experimenter to follow-up on any feedback and the engineer to provide any other feedback. Although not part of the cognitive walkthrough process, the engineer also reviewed the information weighting scheme used to determine the priority of information for the AR-HUD and was satisfied with their assignment, aside from the grade, which they identified as more important.

5.1.1 Summary of Results

Overall, the engineer's feedback suggested that the design did not omit any crucial information for train control, although the grade of the track and status of the reverser were two suggested additions. The reverser state is only necessary during the start-up process, and qualified engineers memorize grade information, though it could be a useful memory aid if it did not clutter the moving map display. The engineer seemed to readily understand the format and location of the displayed information, as they correctly answered all questions that required retrieval of specific information from the display. Some novel information that is currently unavailable (e.g., the air brake re-charging time or the current speed limits) was perceived to be beneficial. Similarly, the moving map display was mentioned on multiple occasions as a beneficial addition, despite some concerns over the readability of the small symbols. Other novel elements, such as the acceleration arc display, may be less useful since acceleration can also be estimated from the electric motor ammeters. However, the values displayed by the arc also incorporate the effects of the train consist and grade, so the arc may still provide a benefit. The engineer also mentioned the Target Speed Predictor display several times. Primarily, the engineer thought the maximum useful display range for the icon was only about 1/2 to 1 mile ahead of the train, presumably since the preview time at these distances would be 30 seconds or longer, which allows adequate time to adjust their actions.

The issue of information transparency and the visibility of objects in the environment was a notable topic in the post-session debrief. Despite the team's initial concerns, the engineer did not consider the potential occlusion of equipment on the track (e.g., hot box detector and dragging equipment detector) by the Target Speed Predictor to be a major problem since they are passive devices and not actively monitored like a signal. On the other hand, the occlusion of signals by the speed and moving map displays could be problematic, since engineers are required to monitor signals until the locomotive has passed them. However, because the moving map is on the right side, the duration of obscuration will be short, minimizing the chance of a signal change during that time. The engineer will also be able to quickly confirm the signal state because signals are displayed on the moving map. The location of the speed display is more problematic for observing signals mounted above the track and the location may need to be shifted to avoid this issue.

The sole use of paper mockups limited the ability of the engineer to fully appreciate the dynamic nature of two prominent elements: the acceleration arc and the Target Speed Predictor. For example, during the presentation of the slide series for a specific task scenario, the engineer failed to notice or report changes in the location of the Target Speed Predictor. The discrete nature of the presentation could have contributed to this inability to detect movement of the predictor, but this could easily have resulted from the engineer's attention being focused elsewhere, or the lack of familiarity in monitoring the novel display. Future cognitive walkthroughs should be conducted with an operating prototype or videoclips captured from a working system to ensure that the dynamic nature of the display is faithfully reproduced.

5.2 Engineer Assessment of AR-HUD Information Elements Via Online Survey

A Qualtrics survey was developed and used to interview a group of nine US freight railroad engineers. The goal was to define the information requirements and possible benefits of a new type of AR-HUD. A locomotive windscreen mounted transparent display provides a wide FOV while maintaining AR conformality using image-based head tracking. Based on previous experimental studies of volunteer subjects in a locomotive simulator (Groshong, 2016; Price, 2021), Human Systems Laboratory and GE used hierarchical cognitive task analyses and other analytic methods to identify the tasks, goals, and display information requirements for a reference freight rail operational driving "scenario." When selecting information to display, it is critical to factor in expert opinions at an early stage using cognitive walkthrough techniques or locomotive simulator evaluations using prototype displays. Stakeholders have different backgrounds and expectations, so solicitation and evaluation of opinions is inevitably challenging. Due to the COVID-19 pandemic, in-person cognitive walkthrough evaluations were not feasible. Instead, the research team decided to recruit engineers who had freight and/or passenger experience and administer an online survey. Each participant was first individually briefed on details of the AR-HUD concept in a 30-minute online (Zoom) session. Participants then completed a self-administered survey developed using the Qualtrics online platform.[2] During the 30-minute participant briefing, storyboard photos and diagrams of the prototype display's elements were described in the context of a normal mainline freight operational scenario, involving accelerating from a stop, maintaining speed, coming to a planned stop, or responding to a sudden change. Display elements were grouped as they related to status, operational, or predictive information categories. The online survey required 25-30 minutes to complete. All responses were coded to maintain participant anonymity. The survey included a review of the main briefing points, and asked participants about operational background (e.g., years of freight and/or passenger rail experience in various roles, as well as any prior experience with HUDs and augmented reality systems).

Then, several dozen multiple choice questions, grouped by information category, asked participants to identify the most useful AR-HUD display elements for each phase. Subjects were also asked to indicate the relative importance of each category of display information for each phase, using a virtual pointer on an analog sliding scale. Participants could qualify, clarify, or amplify their answers using 18 free-text, open-ended question boxes. In a final section, participants were asked about specific non-normal operational phases that might change their relative rankings and suggestions for future AR-HUD enhancements. Details on the briefing and questionnaire are available online (Zhang, 2021). Unfortunately, in part due to the pandemic, recruiting had to be conducted by email only, and went more slowly than anticipated. By the time the AR-HUD was finalized, responses had been obtained from only two of the nine planned participants. Only a qualitative analysis was conducted. Although their background differed (e.g., one operator had primarily freight experience and the other passenger), both were highly experienced (>10 years) and many of their responses were congruent. As expected, speed was the most important single status element, with signal state and mile marker close behind. Operational data seemed more important than predictive data, at least for these two experienced engineers. Avoiding information overload is important, and the ability to customize and declutter may be useful. They disagreed on the importance of display element attention flashing and predictive information when stopping. The most highly ranked feature enhancement was adding other train locations to the HUD's rolling map. When asked what percentage of the time they thought engineers currently spent looking down at displays rather than out the windshield, their responses were 5 and 25 percent, and motivated the effort to add eye tracking to the simulator study.

---

[2] Qualtrics. (2022).

6. AR-HUD Implementation in CTIL

The following is a description of the AR-HUD configuration in CTIL (Figure 12).

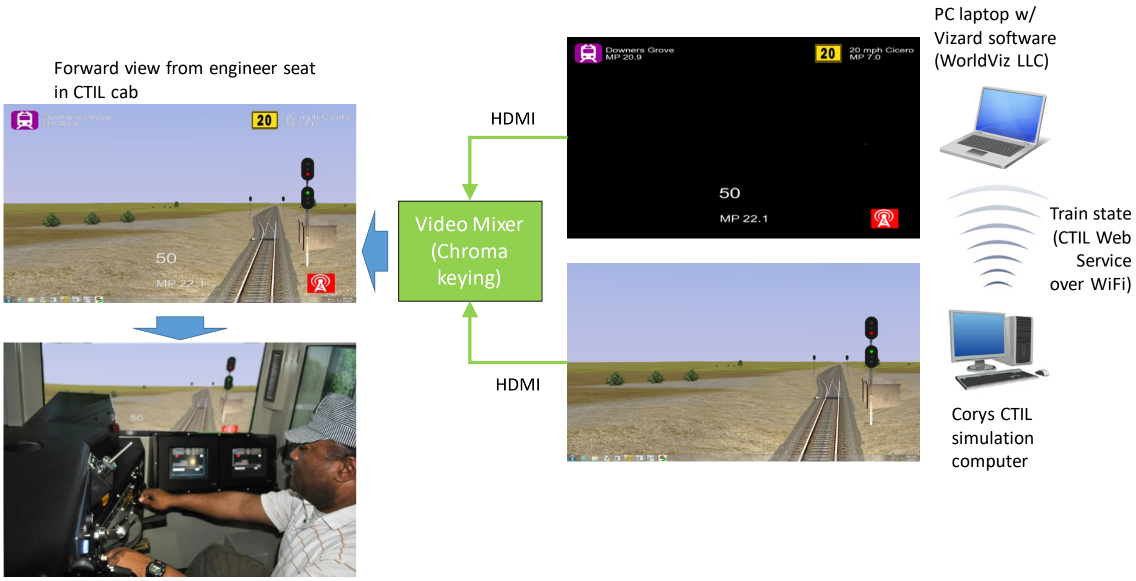


Figure 12. Schematic diagram of the CTIL HUD hardware configuration

6.1 AR-HUD CTIL System Architecture

The current implementation of the AR-HUD is based on the HUD prototype developed under FRA contract #DTFR53-16-P-00052 “Cab Technology Integration Laboratory Head-Up Display upgrade”. The HUD software runs in the Vizard v7.2 Integrated Development Environment (WorldViz, Santa Barbara, CA) within the Windows 10 operating system. Vizard is a widely used Python-based programming tool to develop augmented and virtual reality applications. Figure 13 illustrates the system architecture. The HUD laptop computer (top right) outputs a High-Definition Multimedia Interface (HDMI) signal, routed to the CTIL video mixer, which combines the HUD signal with HDMI video output for the engineer’s forward view from the CTIL simulation computer (bottom right). The HUD symbology is presented on a unique color background, (shown in black in the figure) which is used as the chroma-key for blending the imagery with the CTIL video signal (top left). This technique is commonly known as “green-screen” technology that is widely used in CGI for television and movies. The host laptop also connects to a local wireless network to communicate with the CTIL network via the CTIL web service, which is a simple Hypertext Transfer Protocol server that pushes the current train parameters (e.g., speed, location, and control lever state) in response to a client request. The current HUD computer is an Asus ROG Zephyrus G14 gaming laptop with AMD Ryzen processor and nVidia GTX 1650 graphics, which updates the HUD symbology at up to 10 Hz depending on the HUD features being displayed.

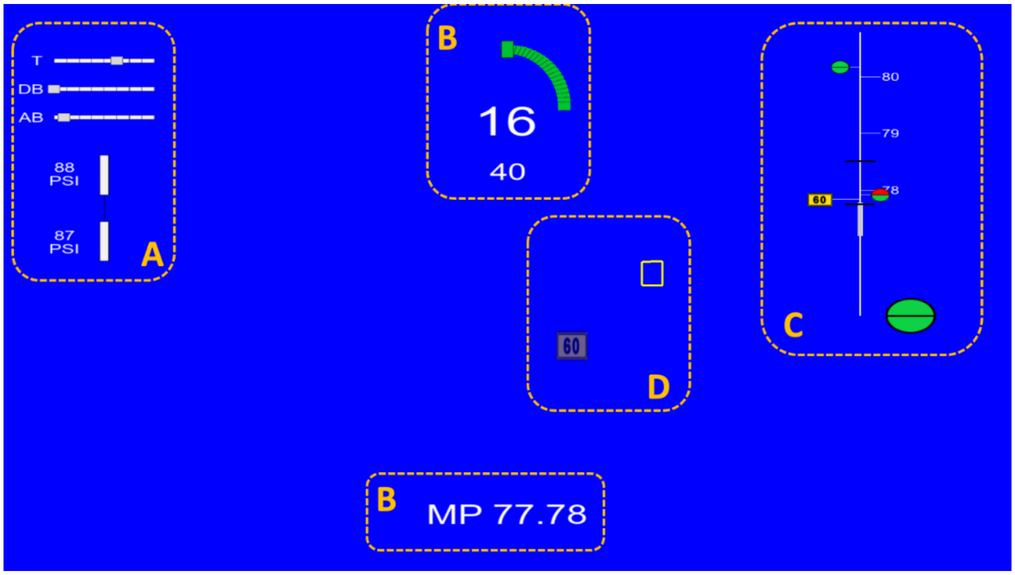


Figure 13. Screen capture of AR-HUD as currently implemented

6.2 Current CTIL Implementation

The current AR-HUD code extends the base HUD application described above to include several of the paper design elements described above. Figure 14 depicts the display elements in the AR-HUD that was used for the user evaluation study. The blue screen background is the key-color for blending the AR-HUD symbology with the CTIL scene using "green-screen" technology and would not be used in a production AR-HUD. The red lines represent the train tracks. Four of the five information classes are illustrated in the figure and identified by the orange dotted regions and letter (A-D). The orange elements are not part of the AR-HUD.

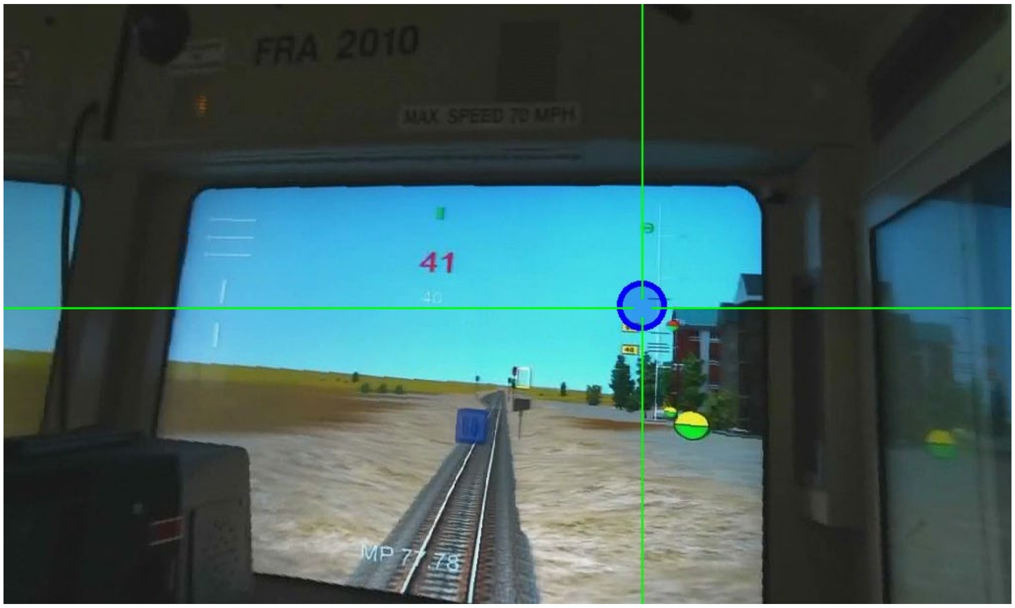


Figure 14. CTIL implementation of AR-HUD viewed by the engineer. The blue circle indicates the engineer's current gaze point

Symbol Group A graphically depicts the current settings for the throttle (T), dynamic brake (DB), and the air brake (AB). The setting of the physical handles, (e.g., throttle notch) is obtained from the CTIL simulation model via its web service so the display indications correspond to the physical setting of the handles. Front and end of train locomotive air brake pressures (showing "88 PSI" and "87 PSI," respectively) are also directly available from the CTIL simulation model.

Symbol Group B contains train speed (in mph), the current speed limit (in mph, if not track speed), the acceleration arc display (in mph/min), and the train milepost location (at the bottom). At the maximum extent of the arc (as shown in Figure 13), the train's current acceleration is 10 mph/min or greater.

Symbol Group C is the moving map display, which currently can display upcoming mileposts and signals, including signal aspect, if that information is available from CTIL. The preview distance ahead of the train (MP 77.78 to 80.78 in this case) was set at 3 miles. The map also displays the 2 miles behind the front end of the train. Both display distances are customizable through a configuration file. Mileposts are shown in white numerals to the right of the track. Wayside signals are shown along the track, on the side corresponding to their physical location. The grade crossings are depicted by the horizontal black lines. Speed restrictions will be implemented once the information is extracted from the CTIL track database. The location of speed changes, (e.g., the start or end of speed restrictions) are displayed at the milepost location to the left of the track. The large signal at the bottom right shows the signal indication for the current block being traversed.

Symbol Group D shows the conformal HUD symbology. Figure 13 shows the yellow conformal signal overlay directing attention to an upcoming signal, which appears 1 mile before the train reaches it. The size of the signal overlay scales with the distance to enhance the sensation of approaching the signal. Also shown in the figure is an overlay indication of the location of a change in track speed (e.g., the start or end of a speed restriction).

Figure 14 shows an example of the engineer's view of the AR-HUD while driving the locomotive. The slight offset of the signal overlay is due to an approximation of the CTIL perspective transformation as well as discrepancies in the database of signal locations. For clarity in CTIL, the symbology was not offset from the scene by partially transparent backgrounds, although conformal objects (e.g., the signal overlay and track speed indications) were presented at 50 percent transparency to avoid complete occlusion of the external environment.

6.2.1 AR-HUD Development Status

The frameworks of several other AR-HUD features have been integrated into CTIL. However, due to scheduling constraints and underlying discrepancies between track models and the CTIL simulation, their performance was not sufficiently reliable to be included in the final implementation. In Group A, an algorithm for brake recharging time was developed based on the GE's quantitative train model, however, the CTIL simulation displayed faster system dynamics. The CTIL model is proprietary, so further progress to reconcile the model performance would have required additional help from CORYS, the manufacturer. For the Moving Map (Group C), the locations of interlockings have not been included in the moving map since the present experiment scenarios did not include any track changes. The information is known and could be added to the Moving Map when needed. For the AR conformal symbology set (Group D), the

Target Speed Predictor, based on current acceleration, was integrated into AR-HUD. However, proper alignment of the Predictor and Target Speed Headway overlays with the track could not be attained due to unresolved discrepancies between the GE and CTIL track databases. Thus, the predictor was also not included in the Moving Map, although the information could be readily added.

7. Human-in-the-loop Evaluation

7.1 Introduction

The goal of human-in-the-loop evaluation was to collect data about engineer gaze patterns and train handling performance during typical locomotive operations to study how the use of HUDs alters engineer behavior. Using a HUD is expected to decrease the amount of time spent looking at the HDD, since critical information (e.g., speed, location) is replicated in the forward view. The team used six consulting engineers' subjective opinions about the CTIL implementation and its ease of use to further refine the current HUD design.

7.1.1 CTIL Simulation and Data Collection

In the experiments carried out in CTIL, the engineers drove a simulated freight consist comprised of 3 head-end locomotives and 45 cars (25 empty cars followed by 20 loaded cars). The consist was 3,144 ft in length and weighed 4,296 tons. Engineers drove the approximately 30-mile section of track profile computer graphic imagery (Figure 15). There were no additional speed restrictions along the route, although there were groups of maintenance workers distributed along the track. Engineers were instructed to proceed past the work parties safely, without needing to communicate with the foreman.

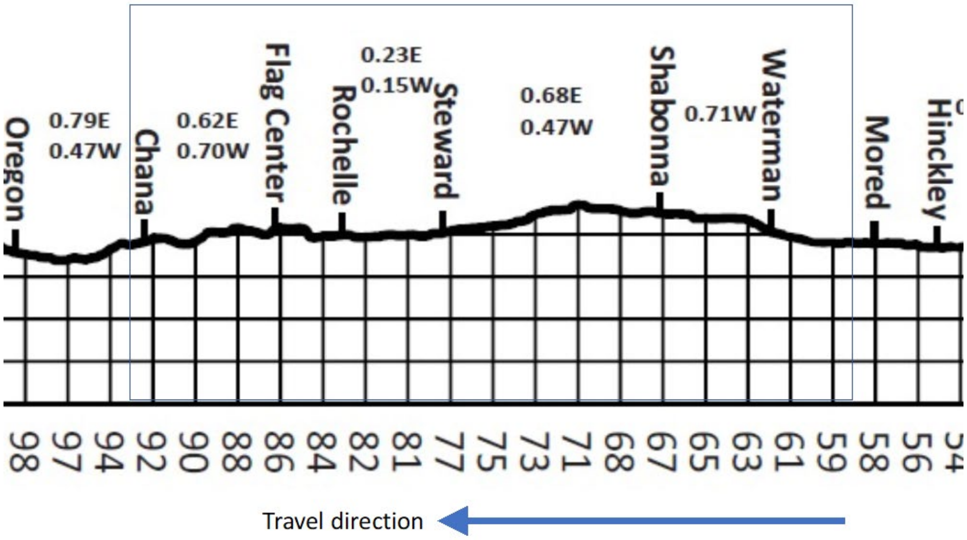


Figure 15. Track profile of the research subdivision from MP 58.8 to MP 93.3 westbound direction

The standard HUD configuration consisted of the symbology in display groups A, B, and C. The AR-HUD configuration also included the symbol group D, the signal overlay, and the speed change symbol.

Engineer gaze data was collected with the Argus Technology (Tyngsboro, MA) ETVision system. This consisted of a lightweight, head-worn frame with binocular, 180 Hz eye cameras and a 720p scene camera. These cameras were connected to a battery powered controller that wirelessly communicated data to the laptop system (Dell G7). The laptop ran the ETVision software, which calculated the fixation measurements (Figure 16). The seven areas of interest (AOIs) were:

1. Forward (Fwd): Captures glances through the engineer windshield to the outside scene, except when a HUD AOI overlaps the location.
2. HDD: Captures glances to the primary train display located beneath the front windshield.
3. Control (Ctl): Captures glances at the train control handle display or the brake pressure display.
4. Speed (Spd): Captures glances at the HUD speed display located near the top center of the windshield.
5. Location (Loc): Captures glances at the HUD location display located near the bottom center of the windshield.
6. Map: Captures glances at the HUD moving map located at the right edge of the windshield.
7. Signal (Sig): Captures glances at the AR-HUD signal overlay. The overlay appears when the train is 1 mile away from the signal along the track, even if the signal itself is occluded by other objects in the scene. Fixations in this AOI were classified manually because the ETVision software could not automatically perform the classification.

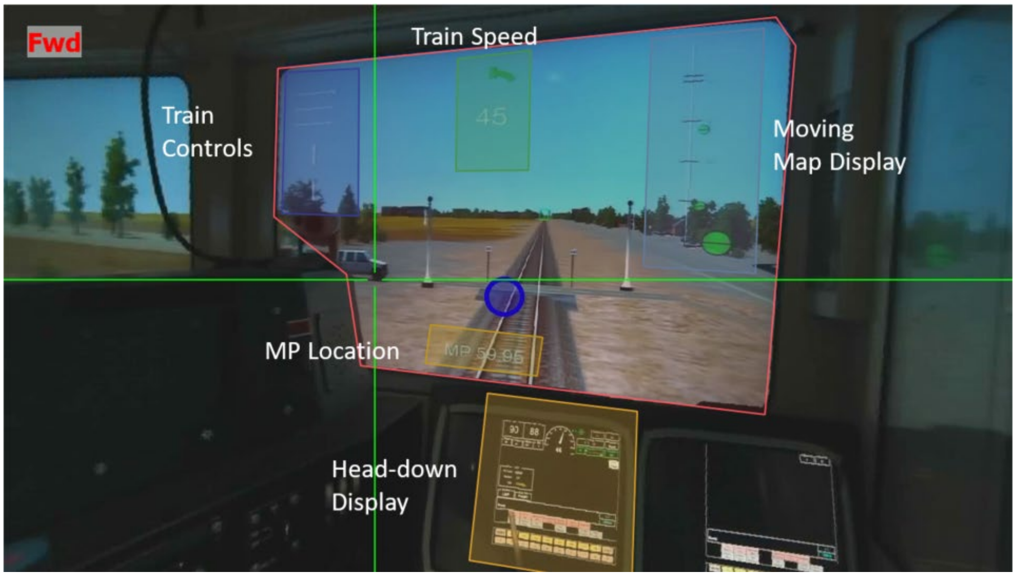


Figure 16. The AOIs used in the gaze analysis

Any fixations outside the AOIs listed above (e.g., when the engineer referred to their paperwork) were classified in the ‘OUTSIDE’ AOI. Once the classification was finished, the dwell time and fixation frequency for each AOI were computed. For this report, only data taken from MP 58.8 to MP 72 was analyzed. This section of track provided a generic scenario within which to compare gaze patterns for each display condition.

7.1.2 Subjects

The six experienced locomotive engineers who participated in the study were recruited from the local population of freight and passenger engineers. The participants were between 27 and 53 years of age (mean: 38.0, SD: 9.6) with 2-6 years of experience in passenger rail service. The average experience was 4.6 years, excluding one engineer with 25 years of experience. Two engineers also had experience on freight rail service (2 years and 25 years). Each engineer's familiarity with the term "Augmented Reality" was assessed on the following 5-point scale:

1. "I have never heard of the term before."
2. "I have heard of the term but cannot define it."
3. "I can give you a definition but am not sure that it is correct."
4. "I know the meaning of the term but have not used it (the technology)."
5. "I know the meaning of the term and have used it before."

The scores were distributed as follows: 2, 2, 2, 4, 4, and 5 (mean: 3, SD: 1.4).

7.1.3 Experiment Protocol

Two days before the scheduled experiment, the engineers completed the informed consent form online, including a COVID-19 disclosure showing no recent exposures or current symptoms. On the morning of the experiment, engineers arrived at the FRA CTIL simulator for the experiment session. Engineers reviewed and signed the informed consent form, the COVID-19 addendum, and subject payment forms. The protocol was approved by the MIT Committee on the Use of Humans as Experimental Subjects (COUHES) (Protocol #2003000117, C. Oman, PI) prior to experiment sessions.

The first experiment session consisted of training and practice. Participants received a brief introduction to CTIL, including how to operate the simulator and a description of the train route and nominal scenario. After the introduction, they drove the simulated train over the entire section of route used in the experiment to become familiar with the territory and train handling qualities. Following a short break, they drove the route a second time using the AR-HUD to become familiar with the display and the behavior of the conformal symbology. Participants also wore the eye tracking system during the latter portion of the second trip to experience how it felt while driving the train. Each trip took approximately 45-55 minutes to complete.

After another short break, each participant began a series of four experimental sessions starting with a trip using the normal locomotive displays (No-HUD). Each trip began with a short briefing on the trip details, such as the train cargo and temporary speed restrictions. The participant donned the eye tracking system and completed a brief calibration protocol. As the trip proceeded along the route, the engineer performed normal activities, including obeying trackside signals and looking for malfunctioning wayside equipment and potential hazards. After a lunch break, the participant drove two trips using the HUD, with the order of HUD trials balanced across the participants. Finally, each engineer completed a fourth trip which repeated the conditions of the first trip. The subject's physical actions and voice communications were recorded for further analysis using audiovisual systems built into CTIL . The eye tracking system recorded the subject's eyes and view of the environment but did not capture any other part of the subject's face. After the fourth trip, subjects were asked to provide feedback about using the HUDs, including subjective ratings about the individual display elements.

7.2 Results

In this section the research team describes the results from the portion of the trip between MP 58.8 (start) and MP 72. This portion of track does not include any speed restrictions but includes six clearly indicated signals and several grade-crossings with and without gates and/or vehicles. This provides a generic scenario for comparing train handling performance and gaze behavior among the display conditions. This segment of the trip was typically completed in 40-45 minutes by the engineers. The first engineer participant was used as a pilot subject, and several changes suggested by the engineer were made to the HUDs for subsequent engineers (this engineer's data was not included in the analysis described below). With a limited data set, the differences in observed gaze and train handling behavior were not statistically tested for significance.

7.2.1 Gaze Results

Figure 17 shows the group averages for the percentage of total gaze time spent in each of the individual AOIs as well as outside of any AOI, such as when the engineer consults paperwork. In the typical operating condition with no HUD, the engineers spent around 77 percent of their gaze duration looking out the front windshield (Fwd AOI), 15 percent of their time glancing at the primary train display (HDD AOI), and the remaining time looking elsewhere (e.g., looking at the paperwork or control stand). The percentages do not sum to 100 percent since this is averaging over the five engineers. The time spent in the forward view agrees with the subjective assessment provided by the one engineer who participated in the second cognitive walkthrough described in Section 5.2.

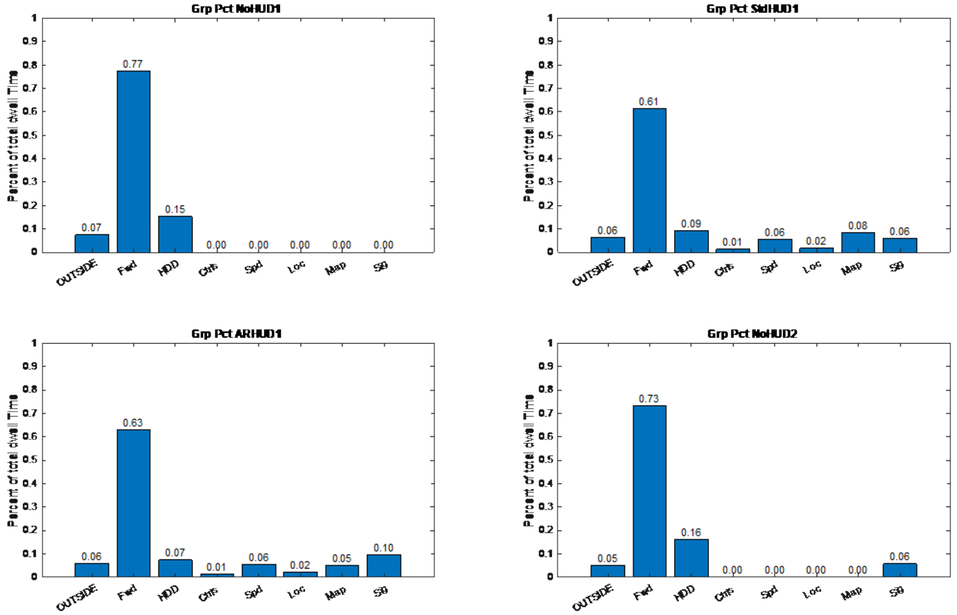


Figure 17. Percentage of total gaze time for each AOI and each display condition. Top left: First No-HUD condition, top right: Std-HUD condition, bottom left: AR-HUD conditions, bottom right: Second No-HUD condition

In the Std-HUD condition, the average amount of time the engineer spends looking out the forward view is reduced to 61 percent of the total time, while in the AR-HUD condition, the engineer spends 63 percent of their total time looking out the forward view. The time spent looking at the main train display (HDD) also decreased from 16 to 9 percent with the Std-HUD and to 7 percent with the AR-HUD. Since there is very little change in the amount of time spent outside the AOIs, engineer gaze is primarily re-distributed over the HUD elements.

Within the Std-HUD condition, the highest proportion of gaze time was spent looking at the map (8 percent), speed (6 percent), and upcoming signal or signal location (6 percent) AOIs. The engineers viewed the location and train control elements the least, at 2 percent and 1 percent, respectively. In the AR-HUD condition, the proportion of time spent looking at the upcoming signal or signal location increased to 10 percent with a concomitant reduction in the time spent looking at the map overlay, down to 6 percent.

Though the trends in gaze behavior of the individual engineers (Figure 18) were generally similar over all the display conditions, there were notable individual differences that could alter the interpretation of the average trends previously described. For example, Subject CH spent the largest proportion of time looking at the HDD, especially in the two HUD conditions, suggesting the AR-HUD might have a slightly larger effect on reducing "head-down time" during a trip than observed. Similarly, the observed increase in total time spent looking at signals in the AR-HUD condition resulted from just two of the five subjects, so it is unclear whether the effect of signal overlay is an artifact of a small sample size or a true effect of the HUD. In the AR-HUD condition, Subject MC spent a much larger proportion of time looking in the Fwd AOI compared to the other subjects. This was an artifact of not manually re-classifying gaze into the signal overlay AOI because of an offset in the eye tracker calibration. Assuming the time spent looking at the signal was near the group average, the proportion of time spent looking at the HDD would be similar to the other engineers.

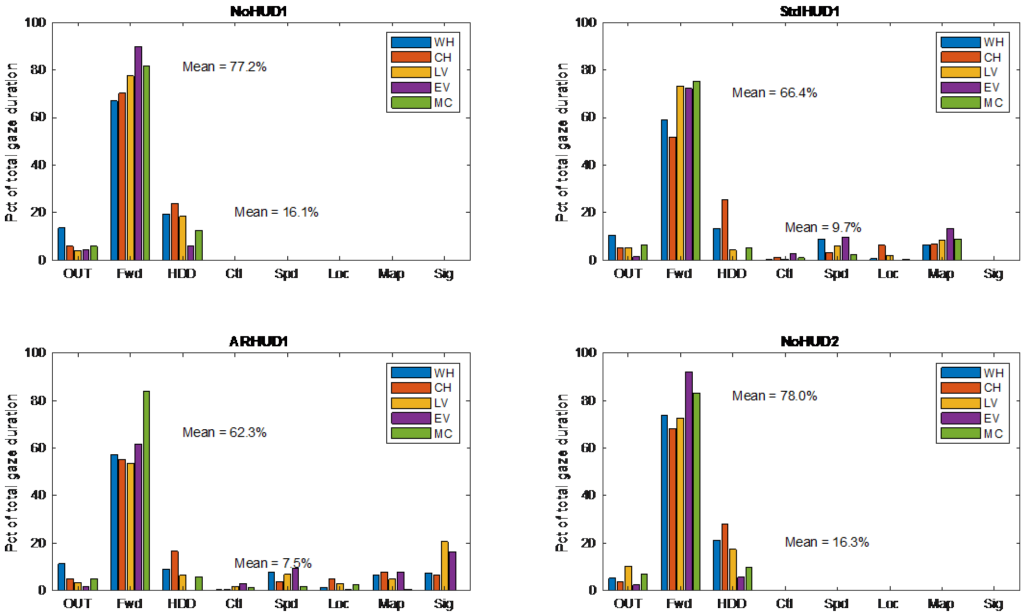


Figure 18. Gaze duration per AOI by subject and display condition

7.2.2 Train Handling

Train handling performance was assessed qualitatively by comparing the train speed, fuel usage, and in-train forces during a trip under each of the display conditions. Figure 19 shows the elevation profile of the segment with the sequence of notch settings from one engineer superimposed. From the trip start, all engineers quickly increased the throttle to notch eight through mile five, after which they had to actively control their speed through the undulating terrain. Train handling performance was expected to differ among the engineers, but display conditions were not expected to affect overall train handling performance.

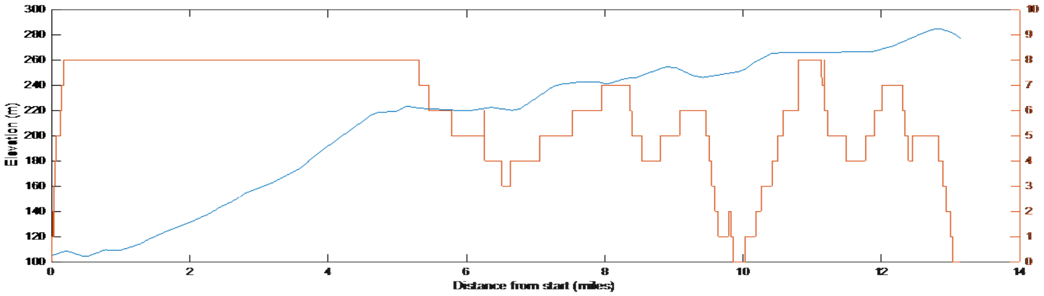


Figure 19. Elevation profile of trip segment shown in blue with the notch settings for one engineer (No-HUD condition) shown in red

The train speed profiles for each subject and display condition are shown in Figure 20. As expected, the profile was largely the same across engineers and display conditions. The largest outliers on the exponential speed increase (miles 0-6) occurred on the first trip (No-HUD condition), but the engineers' performance became consistent for the last three trips. The speed profiles began to diverge in the second half of the trip (miles 7-13), and speed differences of up to 3 mph were observed in most trips. There were no cases of engineers exceeding the 60-mph speed limit during any trips over this segment. The average time to traverse this segment was the longest for the first No-HUD trip (15.9 min) and similar for the other three trips (Std-HUD: 15.4 min, AR-HUD: 15.6 min, NoHUD2: 15.5 min).

Figure 21 shows each engineer's fuel usage for each trip grouped by the display condition. The group means are indicated by the star symbol in each display condition. For all engineers, the fuel usage was the lowest in their first trip but was consistently higher for each of the subsequent trips. The group mean of the last trip (NoHUD2) may be skewed higher by Subject LV who used the most fuel of any engineer across all trips. These results seem to correlate with the lower observed speeds and longer duration of the first trip compared to the remaining three trips.

The average peak draft forces during the trips (Figure 21) showed a similar trend to the fuel use. Draft forces occur when the train is "stretched" due to lead locomotive acceleration. The peaks were identified as the local maxima compared to their neighbors in the observed sequence of maximum draft forces, and the magnitudes were averaged over the trip segment. For four engineers, the lowest average forces were observed in the first trip and were consistently higher for the last three trips. The exception was Subject WH whose highest average maximum draft force occurred on the first trip. The group mean peak buff forces exhibited a different trend, in that the highest forces were observed in the two trips without a HUD. Buff forces are the compressive forces that occur when the train becomes "bunched" together under lead locomotive deceleration or braking. This trend was consistent for four engineers, the exception being Subject

LV. It is possible the HUD provided some preview information from its moving map display that helped the engineer minimize braking, but given the limited number of subjects and inter-subject variability, it seems more likely to be an experimental artifact.

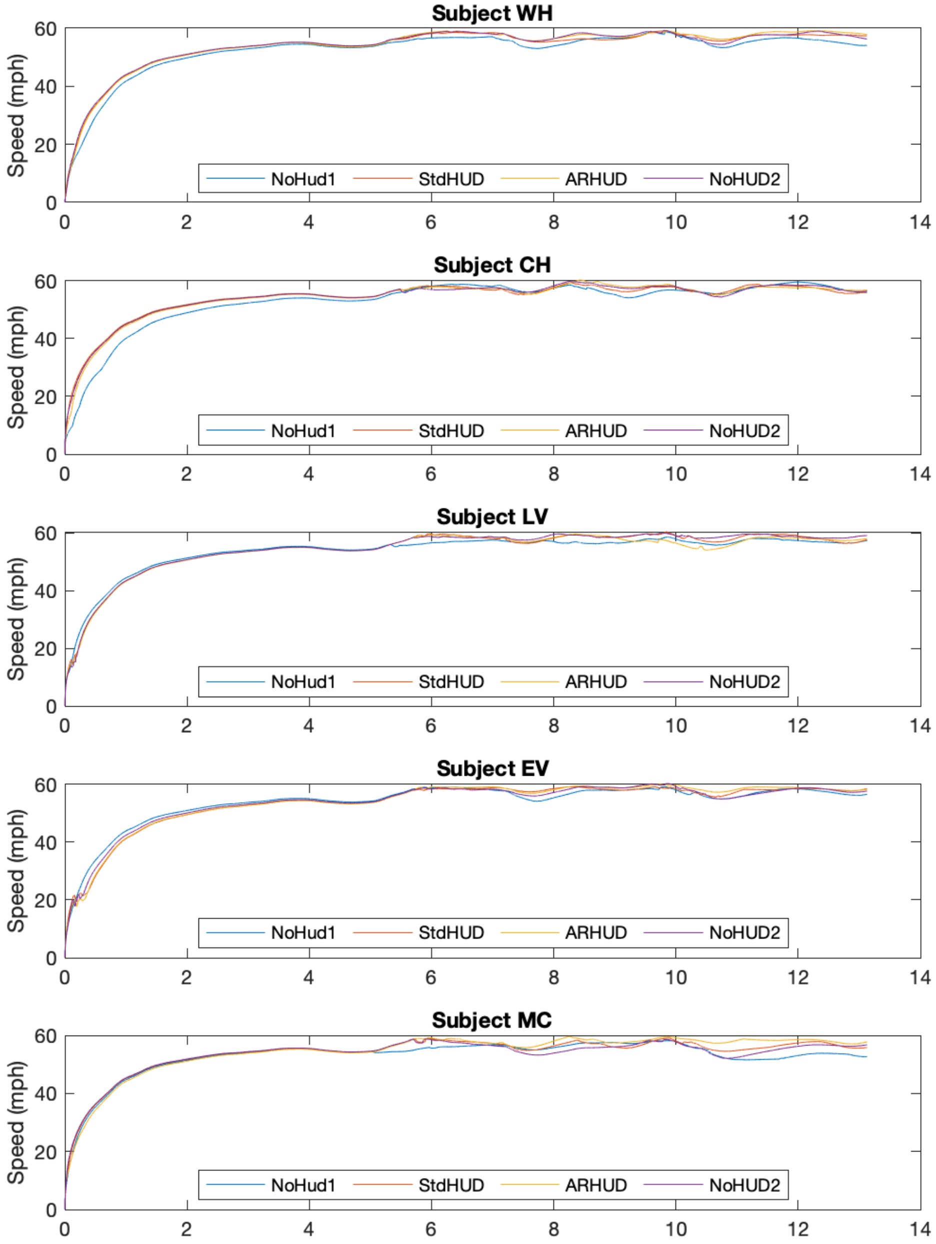


Figure 20. Train speed over the trip segment. Each plot shows one engineer's data for each display condition

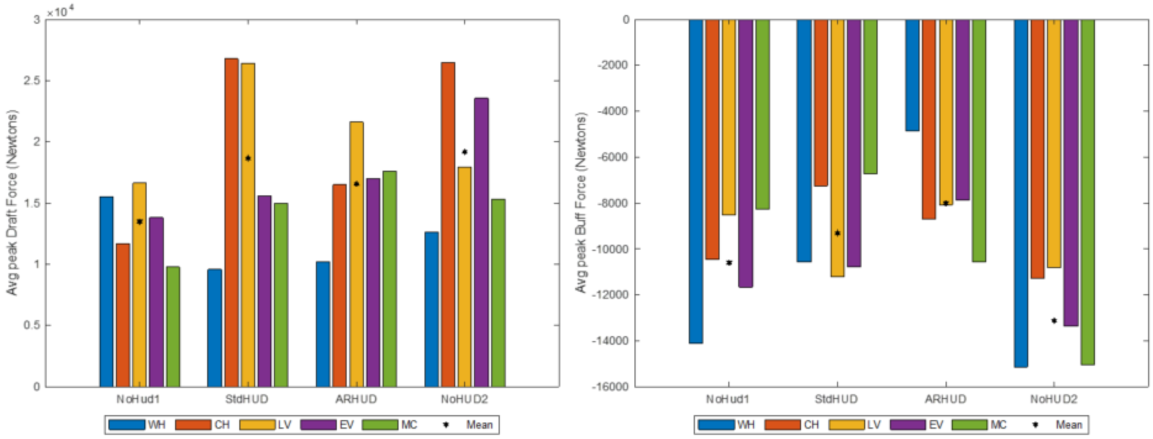


Figure 21. (Left) average peak Draft forces for each engineer grouped by display condition. (Right) average peak Buff forces for each engineer grouped by display condition. The group mean for each display condition is indicated by the star symbol

7.2.3 Subjective responses

After the driving sessions were completed, subjects were asked to rate the usefulness of the individual HUD elements on a 5-point Likert scale (Table 5). Speed and location were the highest rated elements, both assigned a rating of 5. These were followed by the moving map and acceleration display, which were given ratings of 4 or 5. The AR elements indicating the location of speed changes and signals were considered moderately useful, with ratings between 2 and 4, while the remaining HUD elements providing brake and throttle information were deemed the least useful.

Table 5. Subjective ratings of usefulness for individual HUD elements (1 = not useful; 5 = very useful)

	Speed	Location	Moving Map	Accel Display	AR – Speed Changes	AR – Signal	Brake Pressure	Brake Handle	Throttle Handle
Mean	5	5	4.6	4.4	3.2	2.8	2.8	2.6	2.4
SD	0	0	0.6	0.6	0.8	0.8	1.1	0.6	0.9

As expected, displaying the train speed and location were unanimously rated as the most useful information for the HUD. All engineers thought the display elements were well placed, although two engineers suggested grouping the information together to reduce gaze shifts. Placing information high in the visual field (e.g., speed in the HUD) could obstruct the engineer's view of bridge-mounted signals or catenary wires (for passenger service). However, this seems to be preferable to presenting the information low in the visual field where it might obscure the track.

The moving map display was also highly rated by the engineers. The symbology was simple and clear, and presenting it in the forward view simplified the confirmation of information with the external scene. None of the engineers thought the position along the right edge of the window would obstruct their view. Two engineers liked the top-down presentation as it simplified the task of understanding when the end of the train passed a signal or speed restriction. The acceleration display was similarly rated highly useful as a HUD element. The graphical representation provided the information in an intuitive manner, although one engineer preferred a digital representation (as on the HDD) and another engineer suggested adding numbers to the graphical display. Overall, the engineers thought it would be most useful in adverse weather conditions or if an engineer was less familiar with a territory. Both the moving map and acceleration displays have HDD equivalents. Positive Train Control (PTC) and energy management systems like GE Trip Optimizer presently include a moving map with similar information, shown in profile, and the train acceleration is provided in a digital format on the main train display. In a future study, it would be useful to compare how engineer gaze behavior and workload differ between the HDD and HUD implementations.

The two AR display elements, the speed limit indicators and signal overlays, were given mid-level ratings of usefulness with additional comments that sometimes were contradictory. One engineer thought that the AR speed change indicator provided better location precision, but another engineer thought the moving map provided better overall context for understanding how to act on the speed change. Three engineers noted that the information depicted by AR speed indicators was also present on the moving map, so some simplification was needed. For the signal overlays, one engineer found them to be useful for locating signals early, while another thought that engineers qualified on a route would find them redundant.

The brake pressure display was potentially useful in freight service, but the engineers, all of whom were in passenger service, thought it would be unnecessary for the shorter and lighter passenger trains. The displays of the throttle and brake handle position were rated as the least useful elements. Engineers commented that the information was readily available by touch or a quick glance, so the HUD did not provide any real benefit.

7.3 Discussion

The results from this user study suggest that the use of a HUD does not introduce significant changes to engineer behavior and performance for typical mainline operations. The addition of the HUD led to a shift of the engineer's gaze away from the main train display to the information presented on the HUD. The eye tracking data supports the general belief that HUD usage keeps the engineer's gaze in the forward scene for a greater proportion of time. The performance changes that were observed in the train handling measures were more likely the result of improved familiarity with the train and territory after driving the same trip multiple times. In this case, engineers seemingly settled on a consistent driving strategy after the first trip as evidenced by similar trip time, fuel usage, and maximum draft forces observed during the final three trips. The familiarity allowed them to traverse the initial segment of a longer trip in a shorter time but at the expense of slightly higher fuel use and higher in-train forces.

The engineer comments regarding the AR-HUD conformal elements correspond with observations by Tonnis et al. (2005) that any ambiguity in the location of conformal elements renders them much less useful. In this study, some engineers commented that speed information from the moving map provided better situation awareness than the conformal virtual sign. Conformal information may be better suited to automobile applications since the drivers are able to react to events much closer to the vehicle. Engineers must plan their actions on a longer time scale, so the greater preview of the moving map provides better contextual support. Future research should examine whether the predictive conformal information (e.g.,the speed target indicator that was included in the AR-HUD design but not implemented in CTIL) would provide some benefit to train handling.

The initial segment of the trip provided a baseline scenario in which the engineer was primarily concerned with speed maintenance in the absence of external influences such as speed restrictions. Data was also collected for the remaining 21 miles of the trip which included several speed restrictions and rolling terrain that present a more challenging driving task for the engineer. Future analysis will examine whether HUD use provides any performance benefit in higher workload scenarios.

8. Hardware Technical Feasibility Assessment

This section discusses the feasibility of implementing AR-HUDs for freight train locomotive cabs given the current state of the enabling technologies. The broad motivation is to understand which required supporting system capabilities are relatively mature, and to determine in which cases further development is necessary to provide adequate performance. The assessment focuses on the two critical components of the AR-HUD: the wide-FOV display and a system to track the engineer's head in a three-dimensional (3-D) space for proper rendering of the HUD symbology over the view of the external environment. A detailed report on the technical feasibility of implementing a locomotive AR-HUD is attached as Appendix A.

8.1 Displays

The proposed AR-HUD requires the display module to offer a large FOV, a long virtual image distance, and high specifications (e.g., brightness, saturation, and clarity). Ideally, the module can be integrated with or replace the forward window(s) of the locomotive. Recently, wide screen OLEDs have become commercially available, such as the 55-inch display panel produced by LG electronics[3] or the Xiaomi Mi TV Lux Transparent Edition, announced in 2020[4]. The dimensions of these displays cover almost 90 percent of the width and 100 percent of the height of the engineer's window in a typical GE locomotive. There are several advantages to using a transparent display including (1) the larger display area, (2) that no equipment is required within the existing interior volume of the locomotive cab, and (3) that there are no issues of occluding the optical path of the display system, as occurs with current HUD designs explored by Davies et al. (2007). The two disadvantages of this approach are the inability to easily project AR-HUD symbology to other virtual image locations off the display plane and the relatively low transmissivity of current panels, which would affect the engineer's view of the external environment in low-light conditions.

AR-HUDs are now being introduced into commercially available automobiles, such as the Mercedes S-class MBUX AR-HUD. Systems from Texas Instruments, Continental/DigiLens, Envisics and WayRay project the display imagery onto a combiner integrated onto the front windshield of the vehicle. This allows the HUD designer to place the virtual images anywhere behind the display plane from a few meters to optical infinity, which reinforces the illusion of the virtual object being part of the external environment and minimizes the need for the observer to constantly change the focal point of their eyes. Additional benefits include their relatively low unit cost and small packaged volume to fit in a vehicle. The main disadvantage is that their FOV is limited to approximately 10 degrees horizontally whereas an engineer might have a horizontal FOV of 65 degrees when seated.

The automotive manufacturer Geely and Prism Entertainment developed the Aerial Intelligent Display (AID) system that realizes omni-directional 3-D images without media. This system, with a FOV of 16 degrees (h) by 8 degrees (v) and a size of 2-3.8L[5], reconstructs the light field in the fixed space in the car and forms a physical image without relying on any medium (e.g.,

---

[3] LG, See Beyond, LG Transparent OLED Signage, report No. 55EW5F-A.

[4] Press release can be found at 54.6" Xiaomi Mi TV Lux OLED Transparent Edition - Specifications.

[5] Prism Holography, Aerial Holo Intelligent Display, China.

windshield). No details about the system appear to be publicly available but released images show a similarly sized display to automotive AR-HUDs.

Beyond display technology development, there are several display specifications that need to be determined before a locomotive AR-HUD can be fielded, including the maximum or minimum luminance or brightness of the symbology, and how the relative contrast with external objects will affect the operator's perception of virtual and real objects in many possible ways. Some specifications like contrast can be analytically determined; others, such as the preferred virtual image distance, will require additional human subject testing.

8.2 Head Tracking

Commercial off-the-shelf (COTS) tracking systems (e.g., the Intel RealSense D435 depth tracking camera[6]) are powerful, compact, low-cost systems capable of operating within the confines of a locomotive. These systems will probably require additional engineering changes robust enough for the rugged environmental conditions in which locomotives typically operate, including extreme temperatures or vibrations. The error in tracking head position is a critical specification since it will impact the location of conformal symbology viewed by the engineer. A prototype AR-HUD was created at GE to perform the initial laboratory tests on D435 tracker performance (Figure 19). The measured errors in head depth tracking were determined to be around 5 percent of the distance between camera and target and 6-8 percent of the distance when moving laterally. To illustrate how this affects the sizing of conformal symbology, consider the case of drawing a box on the AR-HUD to highlight a distant signal head, essentially a point object to the engineer. For this measured error, a conformal box roughly subtending a 6-7-degree visual angle would be needed to ensure the target object stays within the box. At a typical head distance of approximately 48 inches from the windscreen, the width of the window would subtend an angle of approximately 60 degrees. Additional theoretical analysis is underway to better understand how these errors impact the perception of typical scenes, such as approaching a switch to multiple tracks when several signals are present in the view.

8.3 Overall Assessment of Technical Feasibility

Overall, the current state-of-the-art technology, although promising, is not ready to support product development of a wide FOV AR-HUD for locomotives. The primary limitation is developing appropriate display capabilities. Commercially available large transparent OLED displays have the appropriate display area but are still expensive and have possible limitations like low light transmissibility. Projection display AR-HUD units included in current production vehicles are compact and affordable but have a very limited FOV. Future research should explore whether multiple small automotive AR-HUD units could be combined into a virtual large display, although this adds cost and complexity to the system. Automobile AR-HUD component manufacturers may not be aware of the potential market for rail applications, so developing technological partnerships may help spur further development for the rail industry.

COTS head-tracking systems are mature and capable of supporting the proposed AR-HUD concept. These COTS systems are low cost and have demonstrated sufficient accuracy to track the head in real-time in a lab setting. Current technology for supporting autonomous vehicle

---

[6] Intel RealSense, Intel® RealSense™ Depth Camera D435.

operation also meets computational needs. This report describes development of a demonstration AR-HUD system based on a small, low-cost stereo camera which was used to quantify head tracking accuracy and its effect on symbology size requirements. This information, whether assessed offline using image data or in real-time, can aid in the selection of HUD technology or determination of improvements for enhanced performance of driving in train cabs.

Finally, additional human factors testing is needed to determine the design requirements for parameters such as virtual image distance, brightness, or contrast under all possible operating conditions. Meeting these requirements may necessitate additional hardware development or indicate whether further display and tracker improvements are necessary.

9. Conclusion

The topic of locomotive HUDs has not been comprehensively reviewed since Davies et al. (2007), although some research and development activities have been reported since 2014. In this research, a team from MIT's Human Systems Laboratory designed a wide FOV locomotive AR-HUD. To assess the technical feasibility of the concept, researchers reviewed the literature and the technologies needed to develop a locomotive AR-HUD. Current large area display technologies, (e.g., transparent OLED displays) are still expensive and have limited light transmissibility, whereas the technology used in automotive AR-HUDs have a very limited FOV, so further technical development for rail application will be required. Commercially available head tracking systems are inexpensive and tests of their accuracy in a hardware testbed indicate their performance should be suitable for an operating rail AR-HUD.

An AR-HUD prototype was designed, reviewed by experienced engineers, then implemented in FRA's CTIL locomotive simulator as a testbed to investigate whether the AR-HUD concept would provide performance and/or safety benefits to the engineer in normal railroad operations. A user study was conducted with five experienced passenger engineers, who drove two versions of a locomotive HUD and provided feedback on its design and use (a sixth subject was used as a pilot subject to finalize the course, and their test data was discarded). Engineer gaze was recorded during the simulated trips and one section of the trip was analyzed to determine how gaze behavior changed with the introduction of the HUD. The results showed that less time was spent looking away from the forward view and that the conformal information presented in the AR-HUD might be useful to direct engineer attention to important track-side objects such as upcoming signals. Keeping gaze in the forward view, even if it is focused on HUD information, is crucial for the engineer to detect and identify potential hazards. Continuing analysis of trip data will provide additional insight into the effects of HUD use. Future work should investigate whether this safety benefit is realized when using a HUD. Analysis of train handling performance also suggests that there is no detrimental effect from using HUDs while operating a train. Subjective feedback from the engineers confirms the acceptability and potential benefit of using HUDs.

Further study is necessary to answer the additional human factors questions raised in this report, including the necessary virtual image distances for conformal objects and other display parameters (e.g., brightness and contrast) under all operating conditions. Remaining technology questions include whether large transparent OLED light transmissibility is acceptable or if current automotive AR-HUDs can be combined into a larger display. These additional research questions will require developing a hardware testbed, preferably in an operating locomotive.

10. References

Agarwal, R. (2018). *User Interface Design of Head-Up Display Using Scenarios: An Early Stage Innovation Project at Bombarier Transportation*. Degree Project in Mechanical Engineering Second Cycle. Report No. TRITA-ITM-EX 2018:697. Stockholm, Sweden: KTH Royal Institute of Technology, School of Industrial Engineering and Management.

Brooks, J., Miller, B., Subrahmaniyan, N., Liu, A., Groshong, H., Oman, C., & Houpt, P. (2022). *Function Allocation and New Roles for Humans in Rail Automation*. Research Results No. RR 22-18. Washington, DC: U.S. Department of Transportation, Federal Railroad Administration.

Davies, K., Hinton, J., Thorley, P., & Buse, D. (2007). *Feasibility of Head-Up Displays in Driving Cabs* (T513). Rail Safety and Standards Board.

Groshong, H. L. (2016). *Task Modeling and Assessment for Human-System Interaction in Freight Rail Operations*. S.M. Master of Science. Cambridge, MA: Massachusetts Institute of Technology, Institute for Data, Systems, and Society in Partial Fulfillment of the Requirements for the Degree of Master of Science in Technology and Policy.

Holder, E., & Motz, F. (2014). *Display options for head-up information: Getting the right information in the right place*. *Proceedings of the e-Navigation Underway North America 2014 Conference*, Vallejo, CA.

Kim, H., Gabbard, J. L., Anon, A. M., & Misu, T. (2018). Driver Behavior and Performance with Augmented Reality Pedestrian Collision Warning: An Outdoor User Study. *IEEE Transactions on Visualization and Computer Graphics, 24*(4), 1515–1524. doi:10.1109/TVCG.2018.2793680.

Miglianico, D., Moyart, L., Houbre, F., Nagtegaele, P., & Blanche, J. (2018). USA Patent No.: Washington, DC: U.S. Patent and Trademark Office.

Morgère, J. C., Diguet, J. P., & Laurent, J. (2014, 26–28 Aug. 2014). *Mobile Augmented Reality System for Marine Navigation Assistance. Proceedings of the 2014 12th IEEE International Conference on Embedded and Ubiquitous Computing*.

Porsch, R., & Schaeper, W. (2017). USA Patent No. 9592844 B2. Washington, DC: U.S. Patent and Trademark Office.

Price, R. (2021). *The Expert Locomotive Engineer's Mental Model*. Research Results No. RR 21-15. Washington, DC: U.S. Department of Transportation, Federal Railroad Administration.

Tonnis, M., Sandor, C., Klinker, G., Lange, C., & Bubb, H. (2005, 5–8 Oct. 2005). *Experimental evaluation of an augmented reality visualization for directing a car driver's attention. Proceedings of the Fourth IEEE and ACM International Symposium on Mixed and Augmented Reality (ISMAR'05)*.

Voelbel, K. (2017). *Development and Evaluation of a Moving Map Display for Rail Applications*. Technical Report No DOT/FRA/ORD-17/24. Washington, DC: U.S. Department of Transportation, Federal Railroad Administration.

Wharton, C., J., R., Lewis, C., & Polson, P. (1994). The Cognitive Walkthrough Method: A Practitioner's Guide. In J. Nielsen & R. L. Mack (Eds.), *Usability Inspection Methods* (pp. 105–140.). New York: John Wiley & Sons, Inc.

Zhang, A. T. (2021). *An Approach for Conducting Stakeholder Interviews for a Novel Technology, as Applied to the Locomotive Industry.* S.M. Master of Science. Cambridge, MA: Massachusetts Institute of Technology.

Appendix A.
Technical Feasibility Report: Implementing AR-HUDs with Current Technologies

Technical Feasibility Assessment of Augmented Reality Head-up Displays in Railroad Locomotives

Guangliang Zhao[1], Kaivalya Bakshi[1], H. Kirk Mathews[1], and Andrew M Liu[2]

[1] General Electric Research Center, Niskayuna, NY

[2] Massachusetts Institute of Technology, Cambridge, MA

7 February 2022

Final report for GE subcontract to FRA Contract 693JJ6-18-C-000010 (C. Oman, PI)

Abstract

This technical report examines the feasibility of implementing an Augmented Reality Head-up Displays (AR-HUDs) for train locomotive cabs using currently available technology. The broad motivation is to understand the required supporting system capabilities and provide specifications of the AR-HUD technology, in relation to technical issues arising from the large FOV specific to train cabs. In this context, the selection and specification of hardware for such a system will depend on issues associated with human visual perception besides required supporting infrastructure.

Davies et al. [1] concluded that the addition of HUDs including AR in locomotive cabs could significantly reduce the impact of accidents caused due to driver error. The study considered the adoption of a HUD form-factor commonly used in aviation that has a limited field-of-view and small viewing volume from where the observer could view the symbology. With the development and deployment of low-cost trackers and displays supporting Augmented Reality, there is an opportunity to create a new class of wide field-of-view HUDs that support a locomotive engineer's typical activities. This report first surveys the literature, including publications and technical manuals and primarily focused on AR-HUDs for automobiles, then describes the challenges involved in the system design and implementation of this technology for train cabs are distinct from those in the case of automobiles.

The report assesses the state of the art for the two critical components of the AR-HUD: (1) the wide-field-of view display and (2) a system to track the engineer's head in 3-D space for proper rendering of the HUD symbology over the view of the external environment. The AR-HUD requires the display module to offer a large field of view, a long virtual image distance, and high specifications (e.g., in brightness, saturation, clarity) that, ideally, can be integrated or can replace the forward window(s) of the locomotive. Current display technology for the wide field-of-view AR-HUD is not quite ready for product development in a locomotive. Large transparent commercially available OLED displays have the appropriate display area but are still expensive and have possible limitations such as low light transmissibility. Projection display AR-HUD units included in current production vehicles are compact and affordable but have a very limited field-of-view. Future research should explore whether multiple small automotive AR-HUD units could be "stitched" together into a virtual large display although this adds cost and complexity to the system. Automobile AR-HUD component manufacturers may not be aware of the potential market for rail applications, so developing relationships may help spur further development.

Commercial off-the-shelf (COTS) tracking systems are mature and capable of supporting the proposed AR-HUD concept. These COTS systems are low cost and have demonstrated sufficient accuracy to track the head in real-time as demonstrated in the lab. Computational needs can also be met with current technology used for supporting autonomous vehicle operation. The report describes development of a demonstration AR-HUD system based on a small low-cost stereo camera which was used to quantify head tracking accuracy and its effect on symbology size requirements. This information, whether assessed offline using image data or in real-time, can aid in the selection or determining improvements in the selected HUD technology for enhanced performance of driving in train cabs.

Furthermore, additional human factors testing is required to determine the design requirements for parameters such as virtual image distance, brightness, or contrast under all possible operating conditions, as meeting these requirements may necessitate additional hardware development or indicate whether further display improvements are necessary.

Introduction

Head-Up Displays (HUDs), originally developed for pilots to effectively view information while keep the head positioned "up", have been commercialized in the automotive industry since 1988. The head-up display (HUD) for conveying real-time information in cars has not only been popular among drivers [2] but has also been shown to enhance driving efficiency and safety [3-5]. The integrated approach of using a HUD toward a general safety concept for car drivers has now been implemented in drivers assist technologies in several higher end automotive vehicles. Since most car accidents occur due to human errors in longitudinal and lateral car control, such as late braking or lane departure, such systems typically focus on displaying pathway information to the driver. The benefit of HUDs in comparison with head-down displays (HDDs) or other secondary displays is that the driver is kept *in the loop* of the task, as opposed to being taken out of the loop when looking at and reacting to a secondary display.

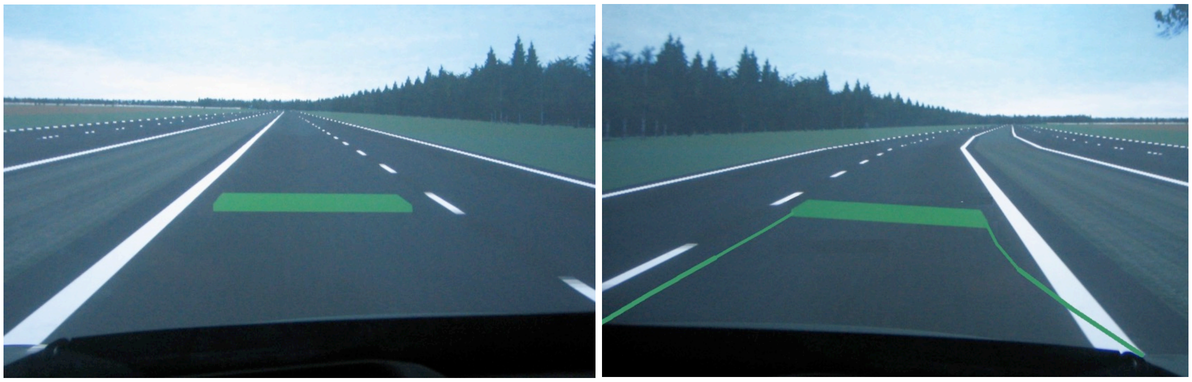


Figure 1 - Screen shot of the simulation environment in [5] that shows the predicted braking point on a straight road (left) and for a curved road, including the anticipated path (right).

Davies et al. [1] concluded that the addition of HUDs in locomotive cabs could significantly reduce the impact of accidents caused due to driver error. However, the study cites the limitations of using the small HUDs, which are appropriate in automobiles, due to data legibility and other issues caused by the relatively larger volume of the *head-box* (i.e., the 3D bounding box within which the driver's head is typically located) in trains.

An Augmented Reality (AR) HUD, unlike the traditional HUD, will overlay perceptual information on the real-world objects over a much wider field-of view and without restriction of head movements. This concept has recently been explored as a technology for driver-assistance of road vehicles, although it has previously been explored for many other applications such as medical visualization, maintenance and repair, and military aircraft navigation and targeting[1]. The automotive industry's investment in AR-HUD research has led to the first production vehicle equipped with AR-HUD in 2021[2]. It is expected that AR-HUDs will be equipped in more motor vehicles in the next few years, which will, in turn, improve the technology, from materials, sensors, semiconductors, to software and interface design. With the

---

[1] Available: https://innoviz.tech/innovizone

[2] Available: https://arstechnica.com/cars/2020/07/augmented-reality-heads-up-displays-for-cars-are-finally-a-real-thing/

maturity of the supply chain in automotive industry, the locomotive industry will be more likely use AR-HUDs along with other driving assistance technologies to improve safety and performance.

The broad motivation of this report is to understand the required supporting system capabilities and provide specifications of the AR-HUD technology, in relation to the technical issues of integrating the display into train cabs. This report includes a literature review towards determining the feasibility of implementing various types of AR-HUDs for train cabs. Since there is a body of literature, including publications and technical manuals on the topic of AR-HUDs for automobiles, a large proportion of the reviewed material belongs to this category. However, the challenges involved in the system design and implementation of this technology for train cabs are distinct from those in the case of automobiles. The core issues in the synthesis of an AR-HUD framework for train cabs are related to the need for a wide field-of-view (FOV) and movement of the driver within the cab. These core issues can be briefly described as:

1. What is the best focal length or image location to display information, and what technology can implement this requirement?
2. Can the use of head-tracking and gaze-detection ensure accurate registration of conformal symbols as the engineer moves about the locomotive cabin?
3. What supporting technology (e.g., machine vision) is needed to perceive the external environment while displaying information conformal with it?

This report explores the potential of transparent organic LED (OLED) and other display technologies to meet the wide field-of-view display requirement and it discusses several human factors display requirements that must be satisfied in any implementation. The report also presents a preliminary study on the viability of a commercial off-the-shelf stereo camera system for performing the necessary head tracking to support conformal display of symbology. The study attempted to quantitatively determining the head tracking accuracy of the system and its effect on symbology size requirements. Finally, the report examines the requirements for any supporting technology that would add additional information that could be shown on the AR-HUD. This information, whether assessed offline using image data or in real-time, can aid in the selection or determining improvements in the selected HUD technology for enhanced performance of driving in train cabs.

System Components

Many companies have been working on AR-HUD systems, by developing the hardware components or integrating end-to-end AR-HUD solutions. For example, Panasonic, as a solution supplier, announced its AR-HUD system with a demo[3] showing an AR rendering featuring large depth and field of view, using the Envisics display hardware module. It has accurate markings for road curvature, lane line, bridge with height, cyclist, road obstacles, etc. For automotive applications, the AR-HUD system is typically a display module of advanced driver assistance systems (ADAS) or autonomous vehicles (AV). The highly integrated ADAS or AV uses all the sensors and inferencing engines to feed to AR-HUD for rendering. Aftermarket automotive AR HUD system may not achieve the same capability compare to that in the car

---

[3] Available: https://na.panasonic.com/us/news/panasonic-automotive-brings-expansive-artificial-intelligence-enhanced-situational-awareness-driver

manufactures' integrated ADAS or AV, and thus is not a focusing market for the major AR HUD manufactures.

For locomotives, the major market will be the upgrade of existing locomotive units. Depending on existing locomotive configuration, an AR-HUD system may use existing sensors and inferencing engines on board, if any, or may need to build in all the sensor and computing hardware and corresponding software components. Considering the trend of modularization in locomotives, a plug-and-play feature may be required to upgrade an AR-HUD system, which requires the system to be self-contained, meaning the system should not just include the display module, but also include all the sensors and computing power as well as software.

In general, an end-to-end AR-HUD system may contain the following hardware components: multi-modality sensors, edge computing, and display module. The software components may include sensor fusion, scene intelligence, and AR rendering. A basic locomotive AR-HUD system may display limited information such as train operation data, basic railway asset detection, which requires fewer sensors, and today's edge computing power can be sufficient. An advanced locomotive AR-HUD system may require a full suite of sensors like that required for ADAS or AV and need large computing power. The inferencing engine in advanced AR-HUD system may include additional situation awareness capabilities like obstacle detection, pedestrian detection, yard assets detection, track condition detection, collision prediction and more, in order to cover the point protection use cases for yard, railway, and grade crossing, etc. Due to the similarity of AR-HUDs for locomotives and automotive, many hardware and software technologies could be shared. Thus, the discussion considers the AR-HUD technologies originally developed for automotive applications but also have potential application in locomotives.

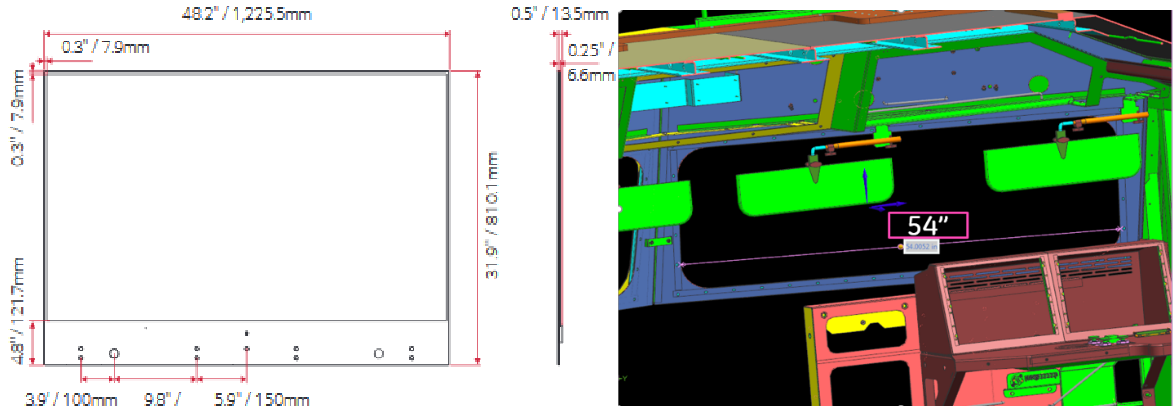


Figure 2 - Dimensions of the 55" OLED LG 55EW5F-A (left) and the width of the engineer side window for comparison.

Large transparent Organic LED Displays

To utilize the entire width of the engineer's windscreen in a typical freight locomotive, a display will need to be as wide as 54" (Figure 2). Large 55" (diagonal) OLED displays are presently sold in the consumer market by companies such as LG, Panasonic and Samsung[4] and have a width of over 48". We explore the benefits and limitations of using such displays in comparison to tube and LCD displays.

---

[4] LG 55" OLED panel - https://www.lg-informationdisplay.com/oled-signage/brand

Display Comparison

Traditional CRT HUDs are typically used with a *combiner*, which is a non-opaque and partially transparent surface on which the light from the tube is projected to form the image viewed by the driver. The primary advantage is that the image location can be manipulated by appropriately designing the combiner's focal length. Combiners are typically small with a narrow visual field of view and may not easily scale to the size needed for an AR-HUD. To achieve a true augmented reality, the projection requires a field of view greater than 10˚ and a long virtual image distance (VID) of at least 7m.[5] Long VID will help reduce the parallax effect, which is the misalignment of the virtual object and the real-world object, especially when the human eye position changes. A study shows the optimal VID between 12m and 15m will result in a better user experience and lower cost, and the elimination of wedge film and eye-tracking for automotive use case [6]. Locomotive cabs, while relatively spacious compared to automobiles, would still be limited to VIDs around two meters without any optics, so the parallax effect ,which could lead to eye fatigue as observers shift their focus between the virtual and real objects, must still be resolved.

A consistent issue with combiner-based HUDs is the appearance of a double image due to the two surfaces of the combiner device, especially when the windshield is used as the combiner. This issue can be resolved by using wedge angled lenses [7] or holographic lenses [8] with the design depending on the desired virtual image location [9]. Additionally, the image location maybe changed easily by small changes to the diffusor lens used in combiner based systems [7]. Note, that the discussion of methods and techniques for mitigating the double image and controlling the image location in combiner systems are compatible with various light sources including LEDs and lasers. Finally, increasing the desired FOV does not cause a major change in the size of the mirror used in the projection-based HUD system [7, 10].

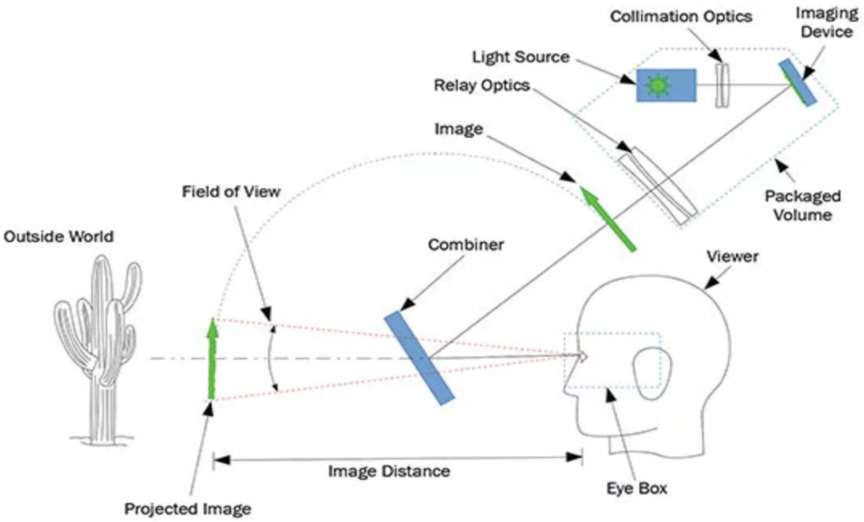


Figure 3 - Combiner-based HUD. Image from Blanche [9].

---

[5] https://www.ti.com/lit/wp/dlpy009/dlpy009.pdf?ts=1627078553620

Although the presentations in [7, 11] are related to automobiles, the discussion on the optics technology translates directly to any HUD system and can thus be applied to installation in locomotives.

Automotive OLED HUDs

The idea of using the entire span of the windshield as the combiner for wide FOV conformal HUD is being pursued in the car industry. 2 shows a cursory comparison of the dimensions of a commercially available OLED display and the train windshield. LG was awarded patents for using flexible transparent (medium sized) OLEDs for AR-HUD in cars [12] (Figure 4), a method to reduce location and angular error in indicating conformal information in the AR-HUD [10]. Transparent OLEDs with dimensions comparable to the windshield have been proposed for AR-HUD concepts in cars [13], although actual application is not commercially available.

The advantages of using such large sized displays attached to the windshield as the HUD compared to combiner-based HUDs (whether LED compatible or not) are:

1. The greater flexibility in positioning information and minimizing occlusion of the external environment, which can be particularly important when displaying conformal information about, say, the location of a signal.
2. Availability of displays of sizes comparable to the cab windshield as indicated in Figure 2,

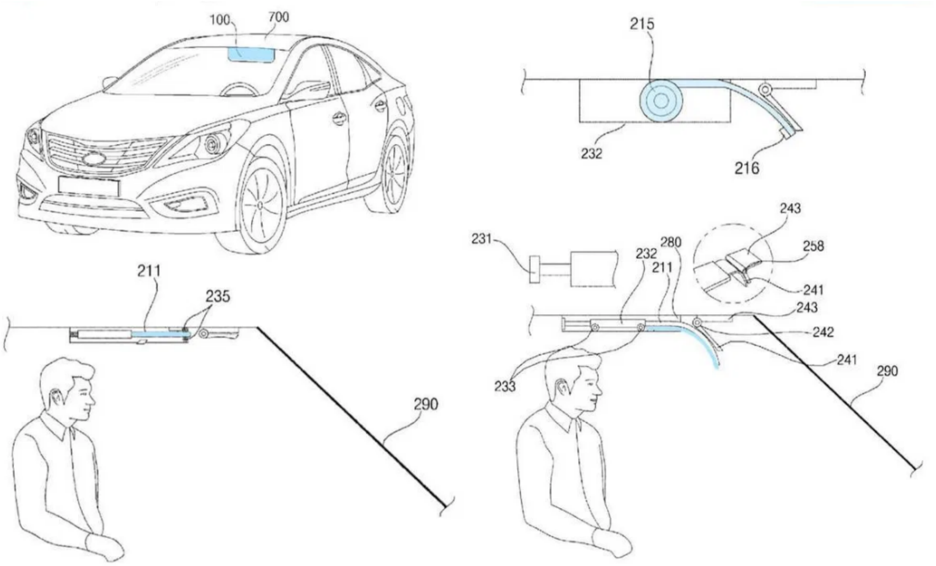


Figure 4 - Flexible transparent OLED for an automotive AR-HUD [12]

3. The flexible configuration as in Figure 4 requires minimal requirement of supporting infrastructure for retrofitting train cabs in comparison to the combiner based HUDs which require significant space in the console or roof mountings to house the optics.
4. Compared to all other see-through AR-HUD displays, including all combiner-based systems, LED based displays have been shown to overcome the difficulty in using occlusion as a visual cue for depth perception [11]. This difficulty is caused in combiner-based systems since they are spatial light modulators which additively combine images and lack the capability to block physical light. Indeed, this difficulty could be overcome by using two spatial light modulators, although this is

not practical for most AR systems since this requires significantly more power and increasing the device form factor through increased complexity of the optical-electrical system. Further, this strategy will lead to the difficulty in maintaining the calibration of two light modulation systems, which is difficult in long-term use. Finally, true hard-edge occlusion would require a pixel-based modulation of the intensity of the display image, which is an important leveraging point when using LED displays. In [11] this is achieved by employing a digital micro-mirror device to combine the LED display AR image with the physical image.

AR-HUDs in Automotive Applications

Several newly developed technologies that purportedly provide a large projection volume and field of view, high level of image clarity, brightness, depth, and saturation, fewer image distortions, and provide the user with an immersive augmented reality experience. In all the examples below, the largest field-of-view is still less than 15˚

Figure 5 – Example of the AR-HUD for the 2021 Mercedes S-class vehicle using TI DLP technology. Source: Mercedes

Texas Instruments uses its Digital Light Processing (DLP) projector, a digital micromirror device (DMD) [14], to perform spatial light modulation with a resolution of 1.3 megapixels and projects onto a 10˚ horizontal, 5˚ vertical aperture HUD, with the image appearing 33 feet (10m) away. A DMD is an optical micro-electronic mechanical system (MEMS) consisting of hundreds of thousands of moving micromirrors that are controlled by underlying CMOS electronics. This display technology is used in the 2021 Mercedes S-Class AR-HUD.[6] (Figure 5)

---

[6] Available https://www.motor1.com/news/436339/mercedes-s-class-ar-hud/

Continental has partnered with DigiLens to develop their Advanced Augmented Reality HUD using waveguide technology that will have a packaged volume < 14 liters.[7] (**Error! Reference source not found.**) The HUD is based on DigiLens' proprietary photopolymer material to enable a thinner multi-layer waveguide with a light engine to direct the light from underneath with maximum efficiency for a full HUD display solution. The published specifications list brightness at ~12,000 cd/m^2, a contrast of ~1500:1, resolution of ~80px/degree. AR content can be displayed over an area 10˚ x 3˚ at a VID = ~8m. The smaller "status" content is shown over an area of 6˚ x 1˚ at a VID = ~2m.

Figure 6- Example of the Continental Advanced Augmented Reality HUD. Source: Continental

Envisics and WayRay are start-up companies who are focusing on developing new laser holographic technologies from optics and electronics components using system engineering and interface design, primarily targeting the automotive industry. Envisics is expecting the mass production of vehicles using its technology in 2023, and as an investor and partner of Envisics, Hyundai Mobis is targeting mass production by 2025[8] [7]. The upcoming 2023 Cadillac Lyriq will be their first vehicle equipped with a HUD and has confirmed will use an Envisics AR-HUD[9]. WayRay has partnered with Karma Automotive to provide AR-HUD units (Figure 7) for their electric vehicles.[10]

---

[7] Available: https://www.continental-automotive.com/en-gl/Passenger-Cars/Information-Management/Head-Up-Displays-(3)/Holographic-HUD.

[8] Available: https://techcrunch.com/2020/10/07/envisics-nabs-50m-for-its-in-car-holographic-display-tech/

[9] Available: https://cadillacsociety.com/2020/10/08/cadillac-lyriq-to-feature-augmented-reality-hud/

[10] Available: https://wayray.com/deep-reality-display/#about

Figure 7 – Left: WayRay picture generating unit (PGU). Right: PGU location in a vehicle showing the relative size of the unit. Source: Wayray

The automotive manufacturer Geely and Prism Entertainment have developed their Aerial Intelligent Display (AID) system which realizes omni-directional 3D images without media. This system, with a FOV of 16˚(h) by 8˚(v) and a size of 2-3.8L[11], reconstructs the light field in the fixed space in the car and forms a physical real image without relying on any medium (e.g., windshield). It will be mass-produced on the Geely ICON and will be launched in the second half of 2021[12].

HUD Perceptual Issues

Since the main purpose of using a HUD instead of a HDD is to present information to the driver in the line-of-sight (LOS), the perceptual issues in their implementation revolve around the visibility of the displayed information.

1. Luminance: The extrema of the luminance of the display must be manipulatable in order to adapt to the surrounding luminance. Although human sight has a high tolerance to high intensity, the display intensity is required to be variable to account for the luminance of the background scene. Additionally, transparent displays have a low transmissivity (e.g., 38 % for the LG 55” OLED product [15]) which could become a critical issue for night-time drivability. This may not be an issue for projective displays, however, depending on the transmissivity of the required coatings. Please see [1] for specifications of the lower and upper bounds of required luminance.
2. Contrast: The relative luminance between the display and background scene must be greater than one for the display information to be visible. Additionally, if the relative contrast is very high, then it will occlude anything behind the display and may capture the focus of attention unintentionally. A contrast of 0.2 to 5 is desired in order to assure effectiveness of the displayed

---

[11] Available: https://www.mholos.com/czqxcpjjjfa

[12] Available: https://inf.news/en/auto/17f8727dc2e73310821eb7076840eac7.html

imagery [1]. The range of luminescence observed during the daytime ranges from 1000 nits to 10000 nits and while LED displays with up to 10000 nits are available, the luminescence on the LG OLED display is only 400 nits, implying a minimum contrast of 0.04 which is very low and a maximum contrast of 0.4. This issue will become more important in the case of displaying conformal imagery.

3. Image location: As previously discussed, the location of the HUD virtual image relative to actual objects is a critical factor in the workload reduction capability of a chosen HUD. The HUD should minimize the focal distance between the objects in the view ahead and the display information image to minimize the effort to refocus the eyes from the HDD to the LOS scenery. Further, the presence of two images in the same LOS, each demanding attention will contribute to the attention capture. In cars, the desired focal length (display image location) can range from ~ 10 to 20 m depending on the mode of driving (highway versus city driving). The desired FOV is ~ 10° to 15°. It is therefore essential to determine the desired focal length of the HUD in a train cab, to select or design the AR-HUD. Such studies are not yet available in the general literature to the knowledge of the authors.
4. Viewing zone: In a train cab, a driver has a significantly large volume within which their head moves, in contrast to the analogous situation in cars or aircraft. This is because the driver can stand up and move about the cab while the cab is in operation. Additionally, operation of the air brake requires the driver to turn while sitting in several cab designs. Since all HUDs have a fundamental limitation in terms of the visibility of information from a location larger than a certain volume or wider than a certain gaze volume, specification of those parameters is also critical to the design or selection of the HUD system.
5. Occlusion: All optical see-through AR-HUD displays, which includes all combiner-based systems, have a major disadvantage when leveraging occlusion as a visual cue for depth perception. This is because the key idea used in such systems is to additively display images on a single display without the capability to block physical light. LED based displays have been shown to overcome this critical issue [11].

Supporting Additional Sensing Capabilities

Beyond overlaying virtual objects over the scene under known operating conditions, an AR-HUD could display additional information such as potential obstacles or enhanced vision in poor weather through a machine vision system like what is used in advanced safety systems for cars. Cognitive Pilot, a Russia startup, announced their Cognitive Rail Pilot system with video demonstrating the scene intelligence capabilities including several rail assets detection and segmentation[13] and is testing the system on Russian Railways[14]. Cognitive Rail Pilot automatically detects signal states, switch position and potential obstacles which could be presented on the AR-HUD overlaying the corresponding objects. The following sections briefly describe the additional sensing and computing requirements needed to support these additional features.

---

[13] Available: https://en.cognitivepilot.com/products/cognitive-rail-pilot/

[14] Available: https://tadviser.com/index.php/Project:The_first_trains_with_artificial_intelligence_from_Cognitive_Technologies_began_to_test_the_Russian_Railway

Multi-modality Sensors

Depending on the capability of an AR-HUD system, different sensor modalities might be used for environment perception. A basic AR-HUD system may only need one forward facing camera for bare-minimum scene intelligence. To have full weather and 24/7 coverage for different use cases like point protection and operation assistance, the system may include some or all sensors like color camera, thermal camera, Lidar, Radar, along with inertial navigation system (INS) and GPS, etc.

A forward-facing color camera could be the primary sensor for AR-HUD. The sensor is widely available and inexpensive and can provide a high-resolution image for perception. The advancement of AI and Deep Learning (DL) technology has improved image analysis significantly, and the development of AI hardware has made edge computing with high-accuracy and real-time inferencing possible. While a color camera is best for perception tasks like object detection and classification, and semantic understanding, it is also highly dependent on visible light and weather conditions. Therefore, it is usually combined with other sensors to have a full day and all-weather coverage. For in-cabin driver monitoring, a color camera and IR camera are used together, and a stereo camera may also be used to infer the depth for accurate inferencing, e.g., for driver's eye localization and tracking.

A thermal camera passively senses the infrared energy emitted from objects in the field of view and generate the heat signature as images. It provides similar perception capability as a color camera and has the advantage in detecting pedestrian and warm objects. It is more robust to weather conditions, dusty or smoke environments, and does not depend on visible light. It can capture objects at a distance. However, it is typically more expensive at the desired specifications compared to a color camera, thus is not widely used in AV development. However, considering its capability of detecting objects at large distance and at night, it might be a good option for locomotives.

Automotive Radar is optimized to detect large metallic moving objects or pedestrians, get the objects' distance, speed, and direction, and are mature with low cost. It could be a good fit for locomotive AR-HUD applications compared to other types of radar. Compared to Lidar or cameras, Radar is more robust to weather condition like rain, snow, and fog and it also does not depend on visible light. Long range automotive Radar could detect objects to a distance of about 300 meters. Short range Radar has a wide field of view which is good for blind-spot detection, obstacle detection, collision avoidance, and is suitable for scenarios like point protection during rail yard operation.

Edge Computing

To achieve real-time environment perception and situation inferencing, powerful onboard computing is required, especially when the analytic models include state-of-the-art (SOTA) artificial intelligence (AI) and deep neural networks (DNN). Depending on the capability and level of intelligence, the end-to-end AR-HUD system including scene intelligence may need computing power equivalent to SAE driving automation level 2 (hands off partial automation) to level 4 (mind off high automation) [19]. Currently there are few products that offer level 2+ computing power. For example, Nvidia Drive AGX Orin system-on-chip (SoC)[15] offers 200 TOPs (trillions of operations per second) processing speed to deliver level 2+ AI-assisted driving, at a price point under $1000 which would be a negligible cost in a locomotive. Again,

---

[15] Available: https://blogs.nvidia.com/blog/2019/12/17/orin-soc/

to our knowledge, there's no AI-accelerator edge device that pass the railway electronic system compliance standards.

Sensor Fusion

Intelligent scene awareness for AR-HUD is achieved by combining data from multi-modality sensors. The same sensor synchronization and fusion are also required in ADAS or AV. Each type of sensor has its advantages and disadvantages, and the fused data from multiple sensors could be more robust and reliable. For example, an object may not be detected by the color camera due to sun glare but could be detected by Radar. An advanced system may have many sensors that generate huge amounts of data which require high computing power for real-time processing. For example, NIO Aquila Super Sensing features 33 sensing units, including 11 8MP cameras, 1 ultralong-range LiDAR, 5 millimeter wave radars, 12 ultrasonic sensors, 2 positioning units, V2X and ADMS. Aquila can generate 8GB data per second [21]. A few sensors for rear sensing tasks in automotive may not be needed in locomotives, however, locomotives may need more sensors to cover a much longer distance in front for emergency stopping. So, depending on the level of intelligence, the number of sensors for locomotive may vary. A basic AR-HUD system comparable to SAE level 1 automation (hands on driver assistance) may only need as few as one forward facing camera. For advanced AR-HUD system comparable to SAE level 4 automation's perception capability, considering the coverage of use cases from yard point protection to railway operation to 24/7 reliable assistance, the total number of perception sensors may be similar to what is used in automotive applications.

Scene Intelligence

The AR-HUD itself is only a display module, and the scene intelligence involves what intelligent features can be inferred. It includes awareness and perception modules, which determines the capability of the AR-HUD system as well as the capability of AV, or in other words, it's the "brain". For locomotive, the scene awareness may include a semantic segmentation model which classifies the scene at the pixel level to categories like ground, rail track, track ballast, building, bridge, etc., an object detection model that detect objects of interest like pedestrians, railcars, motor vehicles, signals, signs, switches, or other railroad assets, scene text recognition model to recognize the signs, mileposts, or speed limits, etc., localization model that localize the detected objects using algorithms like simultaneous localization and mapping, object tracking model that generates spatial-temporal tracklets based on their velocity and direction for those moving objects. The scene perception module will organize the information from the scene awareness module, interpret the meaning and impact, and provide perception results for the display module. Examples could be predicting the possibility of a collision with a motor vehicle that is passing the crossing gate or determining if a pedestrian on adjacent track is maintenance crew or a trespasser.

Currently, data-driven deep learning models are widely used for scene intelligence, and some model accuracies have surpassed human in AI challenges. For AV related challenges, the SOTA pedestrian detection model on the CityPersons [16] test has reached 7.5 Reasonable MR^{-2} (log miss-rate). [17] However, to achieve reliable scene intelligence, a large amount of real-world vehicle operating data is needed to train and test the AV system. A study shows that for autonomous vehicles to achieve 95% of the safety of human drivers, 275 million failure-free autonomous miles are needed, and billions of miles are needed to make AV 10% or 20% safer than human. [18] For the locomotive industry, there are only a few related public datasets such as the real-world data collected for development of PTC or autonomous locomotives.[16] However, It is likely that the Class 1 railroads could quickly acquire the necessary train data given that US freight trains covered roughly 445 million miles in 2019 (reported by the US Bureau of Transportation Statistics).

Head tracking requirements

As previously explained, the bounding head-box obtained from the driver's motion in the cab is a critical specification for designing the HUD. The largest dimensions of the head-box will prescribe the largest viewing angle and affect the virtual image location design, in turn affecting the design of the display. Additionally, this information must be used in real-time to change the location of the virtual image so that it conforms with the environmental reality. Such systems have been proposed [13] and use a viewing angle calculation unit for estimating a LOS direction of a driver using a face direction, detected based on face images of the driver, and center positions of pupils and calculating a viewing angle. This feedback is then used by the matching unit, which matches a location of real-world information located in front of a driver's seat with a location of corresponding virtual object information located in front of the driver's seat, based on the LOS direction and viewing angle. To determine the head-tracking volume, accuracy, and refresh rate requirements, we developed a laboratory mock-up of an AR-HUD with head tracking using commercial off-the-shelf components and open-source software. The goal was to implement a prototype that would demonstrate the feasibility of this approach and provide estimates for the necessary speed and accuracy of the system.

Head tracking using stereo camera: preliminary results

This section describes preliminary experiments using the Intel D435 stereo camera for head tracking, towards determining the head-box volume. This camera uses the Intel RealSense Vision Processor D4 featuring high depth resolution of up to 1280 × 720 at 30 frames per second, long-range capabilities, global shutter technology and a wide FOV (91.2° × 65.5°). The depth camera offers accurate depth perception when the object is moving or the device is in motion, and it covers a wider FOV in that case, minimizing blind spots. The camera is easily portable owing to the small size (90 mm × 25 mm × 25 mm), making it ideal for performing head tracking in train cab simulators or in the real world. Please see the data sheet for detailed mechanical, data processing specifications including humidity, maximum sustained (0° C to 40° C) and short exposure (non-operational) temperature (-40° C to 70° C) [19].

---

[16] Available: https://fortune.com/2019/07/29/autonomous-trains-challenges/ or https://www.post-gazette.com/business/tech-news/2021/02/21/North-Shore-Wabtec-full-automation-railroads-safety-positive-train-control-human-error/stories/202102180186

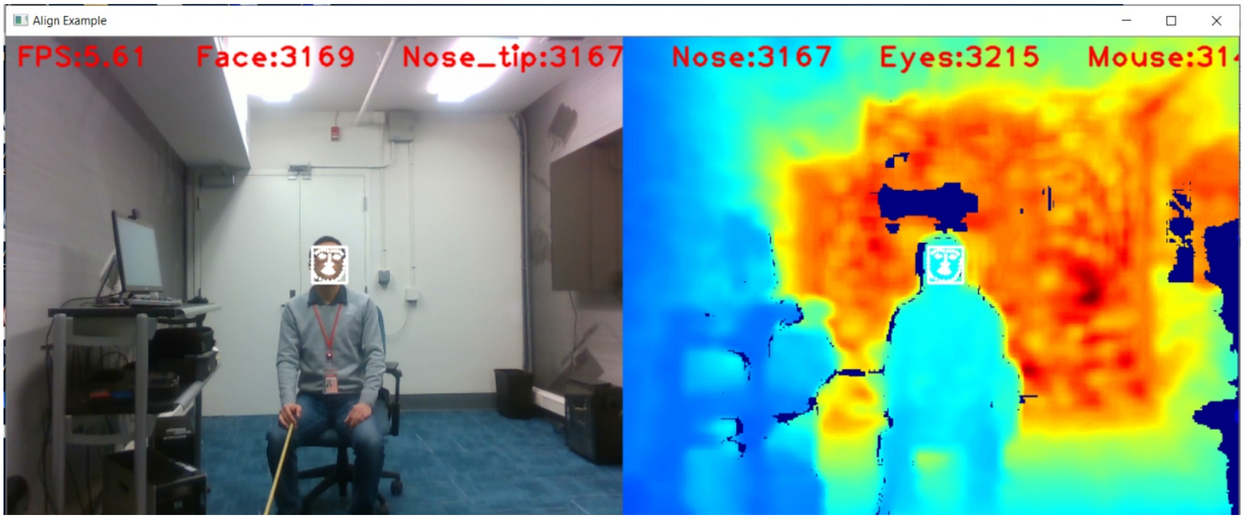


Figure 8 – Left: Testing the depth perception APIs at GE Research. Right: Image acquired by the stero deph camera.

The intelrealsensepy2 Python 3 API hosted by Intel[17] was used to obtain and manipulate the video stream. Dlib is a modern C++ toolkit containing machine learning algorithms and tools for creating complex software in C++ to solve real world problems. We use a pre-trained model of the facial landmarks tool that takes in an image region (i.e., the facial image) and the dlib.shape_predictor() method in Python attempts to identify the locations of important facial landmarks such as the corners of the mouth and eyes and the tip of the nose that define the pose of the head. A preliminary validation of the depth detection APIs was conducted in the Transportation Lab at the GE Research Center. (Figure 8)

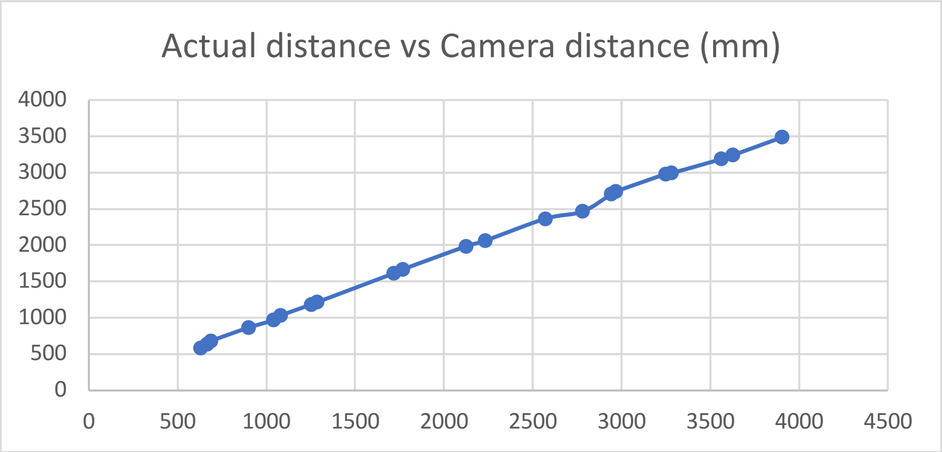


Figure 9 - Comparison of the camera depth measurements with physical measurements (x-axis).

Readings of Z axis measurement or the depth distance of the face from the camera were taken intermittently over ~ 12 ft (in ~ 0.5 ft increments) for two different subjects (two different faces) with the camera and compared with the distance measured with a tape measure ("ground truth"). The face detection algorithms could distinguish the face from the surroundings at distances ~ 11.5 ft. The average mean error over all measurements was 7.48 % although the datasheet claims this value as 2 %. Most

---

[17] Available at https://github.com/IntelRealSense/librealsense

error values were close to 5% possibly indicating a bias in the camera measurement. Furthermore, the error increases rapidly over the range ~ 10 ft to 12 ft from the camera and that the readings become increasingly unstable at larger distances (Figure 9).

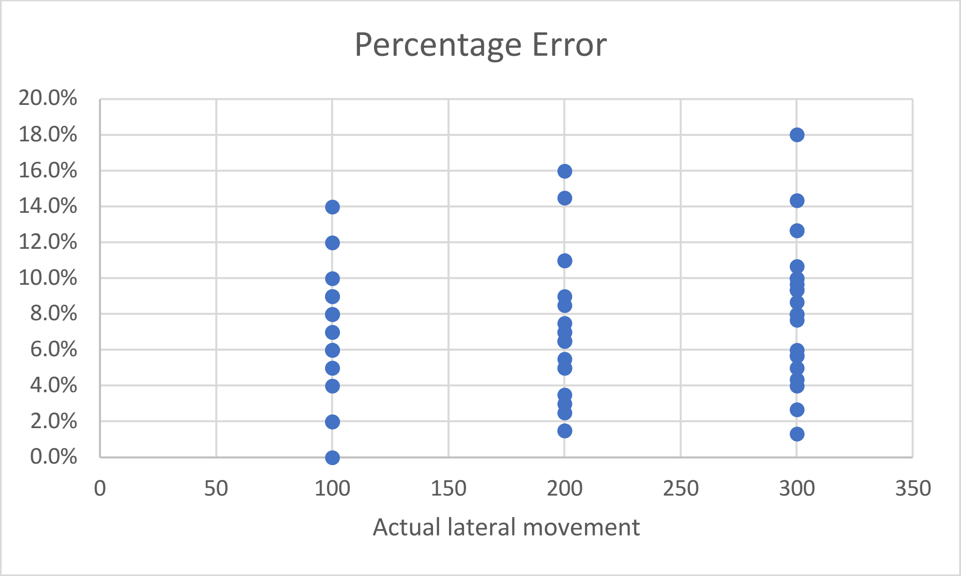


Figure 10 - Percentage error in the measurement of lateral motion of the head and face at a fixed distance from the camera.

Measurements to estimate the error in lateral position are shown in Figure 10. For this experiment, the distance of the face from the camera was assumed to be constant at 1000 mm while the head was laterally displaced by 100-300 mm. The mean lateral error over all displacements was calculated as 6.5% compared to the RMS error of 2% claimed in the datasheet [19].

Error Analysis

Consider an object which we want to conformally annotate in the AR view of the driver presented at a display screen. Achieving this task would require real-time estimation of the positions of the head as well as the object. Clearly, the accuracy of this estimation will affect the precision with which a conformal image (say a point) can be placed at the appropriate location on the screen in which it appears in the vision of the driver. We would then like to know the impact of the depth and lateral eye position estimation provided by a tracker configuration similar to the previous section, on the intended position (S in Figure 11) of the displayed conformal image (which should ideally coincide with the position of the object as seen by the driver) corresponding to an object in the environment.

Using the co-ordinates defined in the figure, we further denote

$$r_e := \sqrt{z_e^2 + x_e^2},\ L_{eo} := z_e - z_o \text{ and } L_{so} := z_s - z_o.$$

Since $z_e = r_e \cos\theta_e$, $x_e = r_e \sin\theta_e$ and $x_s = x_e + \frac{z_e - z_s}{z_e - z_o}(x_o - x_e)$ it can be shown that

$$dz_e = dr_e \cos\theta_e - r_e d\theta_e sin, dx_e = dr_e \sin\theta_e + r_e d\theta_e \cos\theta_e$$

and the error sensitivity is given by

$$dx_s = dx_e + \frac{z_s - z_o}{(z_e - z_o)^2}\, dz_e(x_o - x_e) - \frac{z_e - z_s}{z_e - z_o}\, dx_e.$$

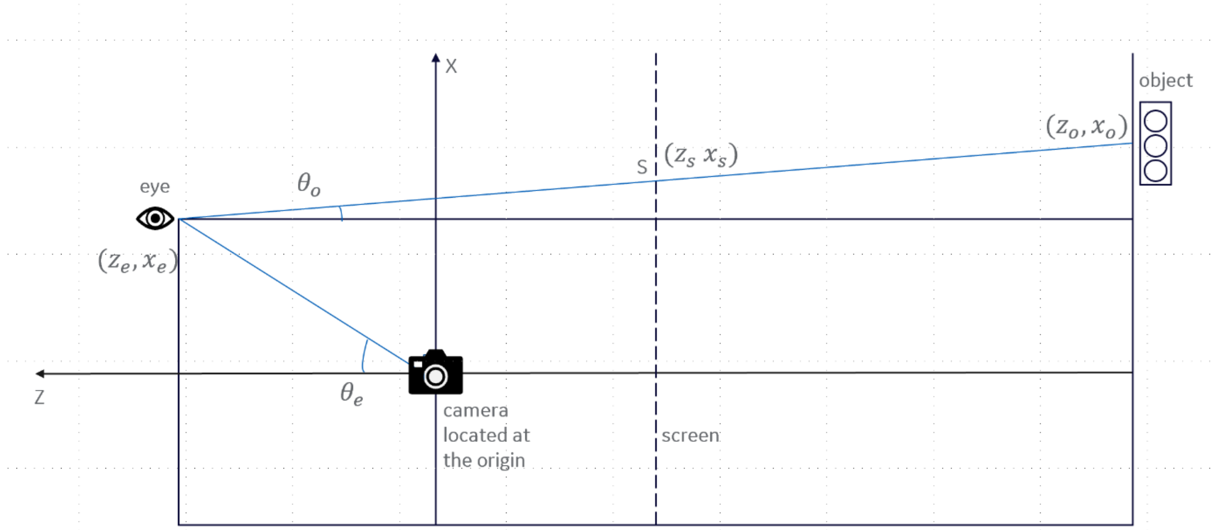


Figure 11 - Schematic depicting the coordinates of the location of the driver's eyes, the screen, and an object external to the driving cab.

Rewriting this equation, we have

$$dx_s = \frac{z_s - z_0}{(z_e - z_0)^2}\, dz_e(x_o - x_e) + \frac{z_s - z_0}{z_e - z_0}\, dx_e = \frac{x_o - x_e}{z_e - z_o}\frac{L_{so}}{L_{eo}}\, dz_e + \frac{L_{so}}{L_{eo}} dx_e.$$

From our preliminary experiments we have estimates of the percentage error in the measurement of z_e and x_e for various distances z_e and x_e. By doing further algebra on the previous equation, we can obtain

$$dx_s = \frac{L_{so}}{L_{eo}}\,(dz_e \tan\theta_o + dx_e).$$

Further, denoting $dx_e' := r_e d\theta_e$, using the trigonometric representation of dz_e and dx_e and the sum of angles trigonometric formulae we have

$$dx_s = \frac{L_{so}}{L_{eo}}\frac{1}{\cos\theta_o}(dr_e \sin(\theta_e + \theta_o) + dx_e' \cos(\theta_e + \theta_o)).$$

The absolute error dr_e corresponds to the depth estimation absolute error and the lateral position estimation absolute error corresponding to the experiments described in the previous section. We note that this implies a linear relationship between the error in the intended location of a conformal image and the position and lateral eye position estimates. However, for a given location of the object (given θ_o), this relationship is nonlinear in relation to different eye positions. The implication is that the design of the location of the camera position is an important factor in mitigating the absolute error in displaying conformal imagery. In what follows, we first provide some data on locomotives and from earlier studies which is used here to outline a method to estimate this error displaying conformal imagery. In what follows, we first provide some data on locomotives and from earlier studies which is used here to outline a method to estimate this error.

The dimensions of the operating cabin of a GE locomotive are depicted in **Error! Reference source not found.** with the following numbered key:

6. Height from the Raised floor to Operator Desk= 28.59”
7. Height from the floor to the Operator Desk= 36.59”
8. Chair to Screen Distance to Centre position of Side Rail = 30.95+3.25 = 34.20”
9. Chair to Screen Distance to #2 End position of Side Rail (Max.) = 30.95 + 3.25 + 14.75 = 48.95”
10. Chair to Screen Distance to #1 End position of Side Rail (Min.) = 30.95 + 3.25 - 14.75 = 19.45”

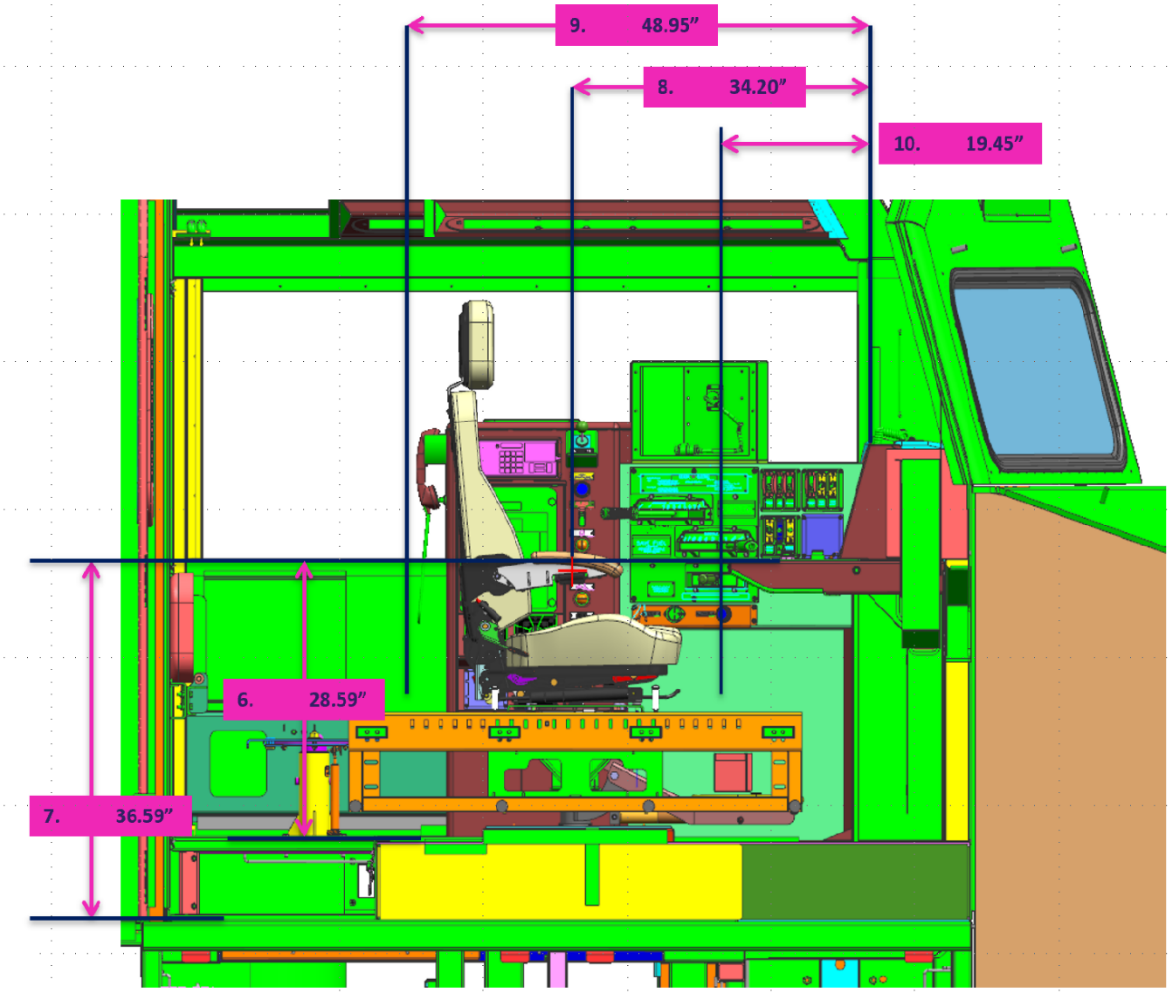


Figure 12 - Schematic depicting internal dimensions of the GE locomotive cabin.

The relationship between the head-box in which the driver's head stays during most of the time of travel, the cab console which is the likely candidate for attaching the camera to determine in real-time the location of the driver's head and the windshield, which is a candidate for the display screen, is shown below. We can now combine the information we have about a typical locomotive cabin in the above figure and knowledge of the head-box from the prior literature to provide a quantification of the error in the intended position of the conformal image on the screen.

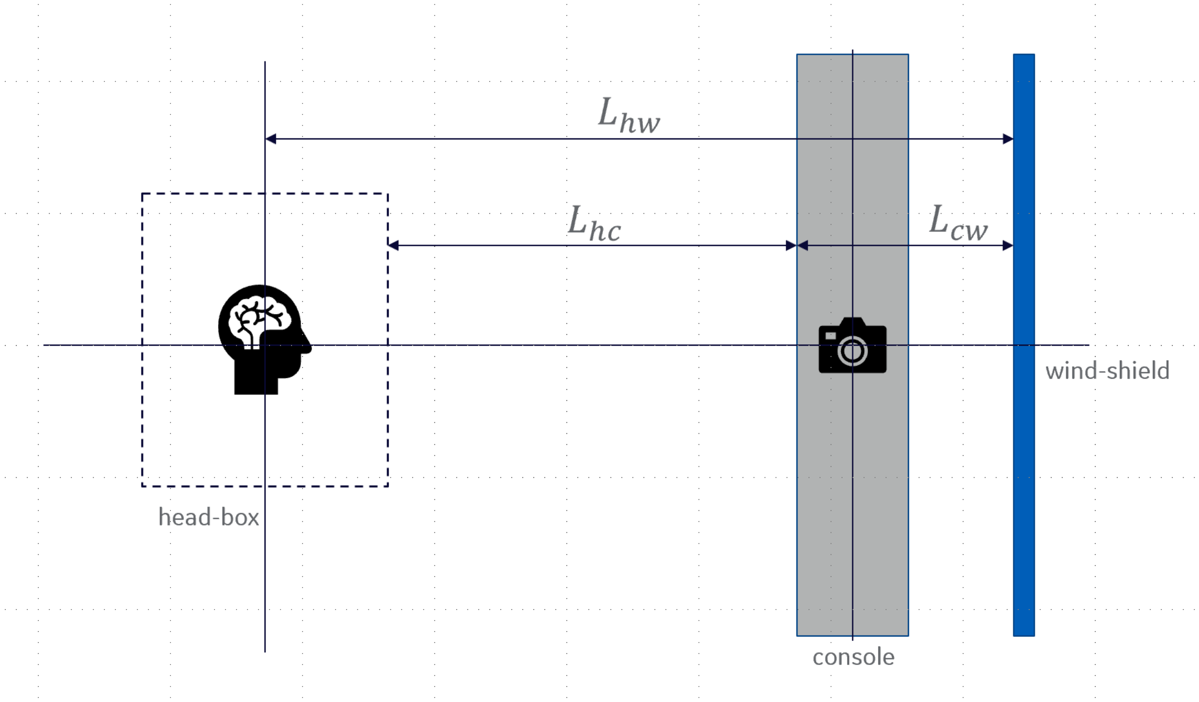


Figure 13 - Schematic depicting the head-box of the driver [16].

An estimate of the head-box containing the driver's head location in a GE locomotive appears in [20] with dimension of width (w) × height (h) × depth (d) = 9.4" × 4.6" × 8.4". The driver's head is found in this volume box appropriately 90% of the time, well within view of a 40° field-of-view console-mounted camera [21]. However, about 6% of the time, the head is turned too far for the eyes to be seen, and 2% of the time the engineer works elsewhere in the cab.

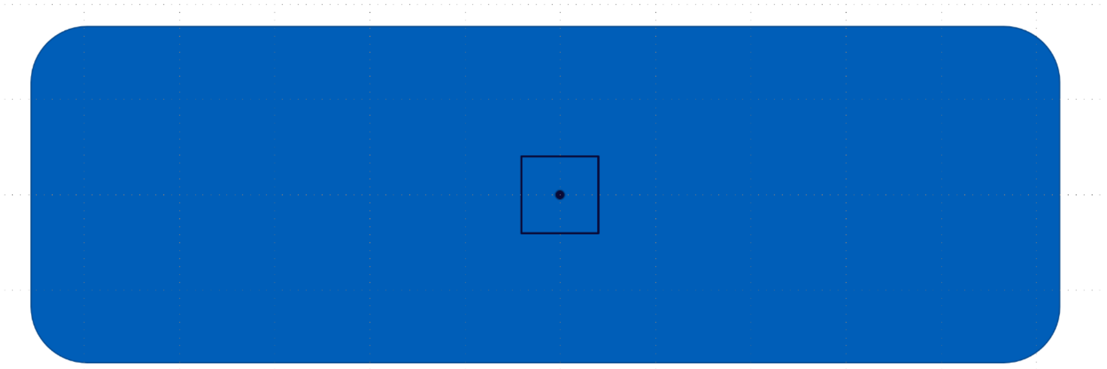


Figure 14 - Schematic showing 1:5 scaled worst-case one standard deviation error box in indicating a point of zero dimensions on a single pane of the windshield which is conformal with a point in the external environment

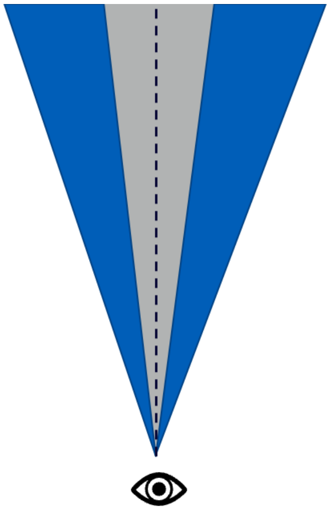


Figure 15 - Schematic showing the worst-case angular error overlaid with the maximum estimated FOV of the driver in indicating a point in the FOV which is located on the dashed line.

The highest spatial standard deviation in conformally displaying a point at the windshield is approximately 4.7“/2 while the windshield dimensions for the cabin considered above are 17.65” × 54”. The schematic in Figure 14 is a 1:5 scaled depiction of the size of a box conformally displayed around the intended target point in the environment needed to ensure would lie within the area of the box assuming the largest tracking error. The visual angle error is approximately 6.4° when two cameras are used for tracking the eyes of the driver while it is 6.9° if only one camera is used. Figure 15 shows the worst-case visual angle error overlaid with the total field of visual range of approximately 20°.

Note that the error in the conformal display of a point which is analyzed in this section is a function of the sum $\theta_e + \theta_o$ and that the spatial and angular errors noted here as well as depicted in the schematics above assume the value of this sum to be such that the error is maximal.

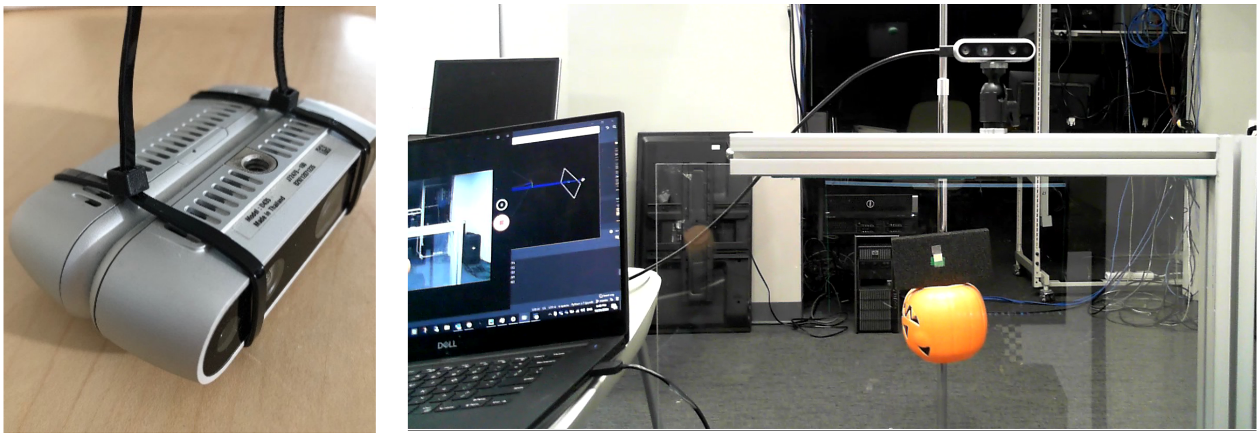


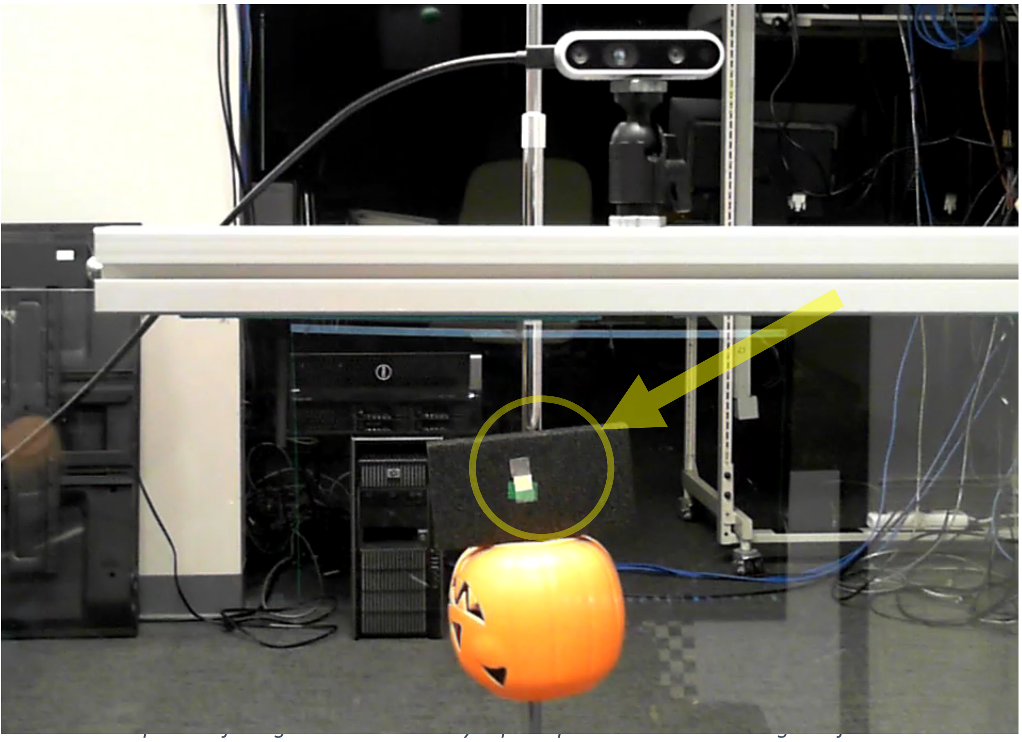
Figure 16 - Left: dual Intel RealSense cameras for inward eye tracking and outward object detection, localization, and tracking. Right: experiment setup and a white square target on black Styrofoam board attached to a pole, overlaid in real time with green virtual symbol on a screen between observer and the target

Demonstration System Implementation

A demonstration AR-HUD system with head tracking using dual Intel D435 cameras in a minimalist setup mimicking in-cab AR-HUD placement was conducted at GE Research for evaluating the performance of the commodity depth camera for in-cab eye tracking. Two cameras were bundled together back-to-back, with one for inward head and eye tracking and another for outward object detection, localization, and tracking (**Error! Reference source not found.**, Left). The cameras were attached to the top of a frame with a clear Lexan screen fixed underneath. A LED monitor was placed between observer and the screen, with LED display facing the screen. The virtual symbols displayed on the monitor were reflected by the screen and clearly visible to the observer while also appearing to be overlaid on objects behind the screen. For the demonstration a pole was placed on the other side of screen with a white square target affixed to a black Styrofoam board mounted on the pole for better visualization (**Error! Reference source not found.**, Right).

For the demonstration, the observer moves on one side of the screen while fixating on the white square. The face keypoint detection model analyzes the inward camera's color image to detect the location of the observer's eyes, and a camera-to-world module localized the eyes' 3D position in world coordinate system from the camera's depth image. The white square is captured and localized from the outward camera within the same world coordinate system. The intersection location on the screen of the eye ray from the observer to the target is calculated, and a green virtual symbol is generated on the LED monitor which is reflected at that exact intersection location on the screen. The observer will see a green mark on the screen overlaying on the white square. (**Error! Reference source not found.**)

Additional experiments with white square replaced by moving target indicate that the overall alignment of the virtual image to the target does not cause discomfort, despite the error of eye tracking.



Figu

Conclusions and Recommendations

The key technologies that will enable wide field-of-view Augmented Reality-HUD for rail operations, namely a wide field-of-view display, accurate head tracking, and sufficient computing power, are all approaching technological readiness levels that could enable test system development. Early limited implementations of AR-HUDs are now being launched in commercially available luxury cars in although the initial feature sets will be limited to basic functions like navigation and display vehicle information and displayed in a limited field of view. Continuing commercialization should lead to further technological improvements which will enable the full suite of locomotive AR-HUD capability.

Display technology for the wide field-of-view AR-HUD is not quite ready for product development in a locomotive. Large transparent OLED displays that have the appropriate display area are commercially available but are still expensive and have possible limitations such as low light transmissibility. On the other hand, projection display AR-HUD units that are now being introduced into production vehicles are compact and affordable but with a limited field-of-view. Future research should explore whether multiple small automotive AR-HUD units could be “stitched” together into a virtual large display although this adds cost and complexity to the system. Additionally, several human factors questions relating to the virtual image distance, image brightness and contrast in a wide variety of operating conditions should also be studied to determine the suitability of any type of display and display requirements for a future system. AR-HUD component manufacturers may not be aware of the potential market for rail applications, so developing a relationship may help spur further development.

Commercial off-the-shelf (COTS) tracking systems are mature and capable of supporting the proposed AR-HUD concept. These COTS systems are low cost and have demonstrated sufficient accuracy to track the head in real-time as demonstrated in the lab. Computational needs can also be met with current technology used for supporting autonomous vehicle operation.

References

[1] K. Davies, J. Hinton, P. Thorley, and D. Buse, "Feasibility of Head-Up Displays in Driving Cabs," Rail Safety and Standards Board, Final Report T513, 2007. [Online]. Available: http://www.rssb.co.uk/pdf/reports/research/T513_rpt_final.pdf

[2] S. Sechrist, "The Expanding Vision of Head-Up Displays: HUDs for Cars at Display Week 2017," *Information Display,* vol. 5/17, pp. 18-23, 2017.

[3] Y. Hwang, B.-J. Park, and K.-H. Kim, "Effects of Augmented-Reality Head-up Display System Use on Risk Perception and Psychological Changes of Drivers," *ETRI Journal,* vol. 38, no. 4, pp. 757-766, 2016/08/01 2016, doi: 10.4218/etrij.16.0115.0770.

[4] M. Tonnis, C. Sandor, G. Klinker, C. Lange, and H. Bubb, "Experimental evaluation of an augmented reality visualization for directing a car driver's attention," in *Fourth IEEE and ACM International Symposium on Mixed and Augmented Reality (ISMAR'05)*, 5-8 Oct. 2005 2005, pp. 56-59, doi: 10.1109/ISMAR.2005.31.

[5] M. Tonnis, C. Lange, and G. Klinker, "Visual Longitudinal and Lateral Driving Assistance in the Head-Up Display of Cars," in *2007 6th IEEE and ACM International Symposium on Mixed and Augmented Reality*, 13-16 Nov. 2007 2007, pp. 91-94, doi: 10.1109/ISMAR.2007.4538831.

[6] Texas Instruments, "The Importance of a Longer VID for AR HUDs," Texas Instruments, Technical White Paper, 2021.

[7] M. Marcus, "Simultaneous head-up display windshield wedge-angle and layer-thickness measurements," in *SPIE News*, ed: SPIE, 2016.

[8] P.-A. Blanche, "Holographic Combiners Improve Head-Up Displays," in *Photonics Media*, ed, 2019.

[9] D. Kessler and M. Bablani, "Head-Mounted Optical Apparatus Using an oLED Display. ," USA Patent US 8,094,377 B2, 2012. [Online]. Available: https://patentimages.storage.googleapis.com/6b/3e/16/c022494d32e292/US8094377.pdf

[10] Texas Instruments, "DLP® Technology: Solving design challenges in next generation of automotive head-up display systems," Texas Instruments, White Paper 2017.

[11] B. Krajancich, N. Padmanaban, and G. Wetzstein, "Factored Occlusion: Single Spatial Light Modulator Occlusion-capable Optical See-through Augmented Reality Display," *IEEE Transactions on Visualization and Computer Graphics,* vol. 26, no. 5, pp. 1871-1879, 2020, doi: 10.1109/TVCG.2020.2973443.

[12] J. Uhm, "Vehicle Display Device and Vehicle Comprising Same," USA Patent US 10086762 B2 Patent Appl. 2017/0349098 A1, 18 October 2018, 2018. [Online]. Available: https://patents.google.com/patent/US20170349098A1/en

[13] K.-H. Kim, H.-S. Park, D.-S. Yoon, J.-W. Choi, Y.-S. Hwang, and J.-H. Park, "Automotive Augmented Reality Head-Up Display Apparatus and Method," USA Patent 8994558 B2, 2015. [Online]. Available: https://patents.google.com/patent/US8994558B2/en

[14] D. Dudley, W. Duncan, and J. Slaughter, "Emerging Digital Micromirror Device (DMD) Applications," *MOEMS display and imaging systems,* vol. 4985, 2003.

[15] LG Electronics Inc., "LG Transparent Signage: 55EW5F-A," ed, 2019.

[16] S. Zhang, R. Benenson, and B. Schiele, "CityPersons: A Diverse Dataset for Pedestrian Detection," in *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 21-26 July 2017 2017, pp. 4457-4465, doi: 10.1109/CVPR.2017.474.

[17] I. Hasan, S. Liao, J. Li, S. U. Akram, and L. Shao, "Generalizable Pedestrian Detection: The Elephant In The Room," in *2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 20-25 June 2021 2021, pp. 11323-11332, doi: 10.1109/CVPR46437.2021.01117.

[18] N. Kalra and S. M. Paddock, *Driving to Safety: How Many Miles of Driving Would It Take to Demonstrate Autonomous Vehicle Reliability?* Santa Monica, CA: RAND Corporation, 2016.

[19] Intel, "Intel® RealSense Product Family D400 Series Datasheet," ed, 2020.

[20] A. Aboukhalil, "Systems Aproach to the Design of Locomotive Fatigue Management Technologies," SM Masters, Aeronautics & Astronautics, Massachusetts Institute of Technology, Cambridge, MA, 2006.

[21] A. Aboukhalil, C. M. Oman, S. Popkin, J. K. Pollard, and H. Howarth, "Quantitative assessment of locomotive cab environment and engineer head movement for development of an alertness monitor," presented at the International Conference on Fatigue Management in Transportation Operations, Seattle, WA, 2005.

Abbreviations and Acronyms

ACRONYM	DEFINITION
AB	Air Brake
AAR	Association of American Railroads
AOI	Area of Interest
AR	Augmented Reality
AR-HUD	Augmented Reality Head-Up Display
BC	Brake Cylinder
CTIL	Cab Technology Integration Laboratory
COTS	Commercial off-the-shelf
CGI	Computer-Generated Imagery
DB	Dynamic Brake
EOT	End-of-Train
ER	Equalizing Reservoir
FRA	Federal Railroad Administration
FOV	Field of View
GE	General Electric
HDD	Head-down Display
HUD	Head-up Display
HDMI	High-Definition Multimedia Interface
MR	Main Reservoir
MOW	Maintenance-of-Way
MIT	Massachusetts Institute of Technology
MP	Milepost
OLED	Organic Light-Emitting Diode
TSR	Temporary Speed Restrictions