A Standard Test Method for Evaluating Navigation and Obstacle Avoidance Capabilities of AGVs and AMRs

Adam Norton[1], Peter Gavriel[1], and Holly Yanco[1]

ABSTRACT

Automatic guided vehicles (AGVs) and autonomous mobile robots (AMRs) are now ubiquitous in industrial manufacturing environments. These systems all must possess a similar set of core capabilities, including navigation, obstacle avoidance, and localization. However, there are few standard methods to evaluate the capabilities and limitations of these systems in a way that is comparable. In this paper, a standard test method is presented that can be used to evaluate these capabilities and can be easily scaled and augmented according to the characteristics of the system under test. The test method can be configured in a variety of ways to exercise different capabilities, all using a common test apparatus to ease test set up and increase versatility. For each test configuration, conditions are specified with respect to the a priori knowledge provided to the system (e.g., boundary and/or obstacle locations) and the obstacles in the environment. Robustness of system capabilities is evaluated by purposefully introducing misalignment between the characteristics of the physical and virtual environment (e.g., providing representations of obstacles in the system's map when they are not physically present). Example test performance data from an AMR is provided. The goal of this work is to provide a common method to characterize the performance of mobile systems in industrial environments that is easily comparable and communicated for both commercial and developmental purposes. This work is driven by existing

[1] New England Robotics Validation and Experimentation (NERVE) Center, University of Massachusetts Lowell, 110 Canal St., Lowell, MA, 01852, USA

Corresponding author: Adam Norton, adam_norton@uml.edu

standards and those in development by the ASTM F45 Committee on Driverless Automatic Guided Industrial Vehicles and will influence the development of new standards within the committee.

Keywords

Automatic guided vehicles, autonomous mobile robots, navigation, obstacle avoidance, localization, standard test methods, evaluation, benchmarking

Introduction

Automatic guided vehicles (AGVs) and autonomous mobile robots (AMRs) are used very commonly in industrial environments to pick up, transport, and deliver goods from one place to another. These types of systems require the ability to intelligently navigate through their environment in order to accurately and reliably perform these tasks on their own, utilizing their own internal models of the environment (i.e., a map) and making updates to those models based on local sensing. While traversing through a space, they must also avoid walls, structures, and obstacles within the environment so as to not cause damage. In addition to physical boundaries and obstacles, virtual boundaries may also be set in order to designate work zones or restrict access to hazardous or otherwise occupied areas.

For more advanced AGVs or AMRs, their navigation capabilities include making decisions based on a combination of their global understanding of the space (e.g., which route to take based on path planning rules or optimization factors), local sensing of the environment (e.g., path planning around obstacles, updating routes in real-time as required), and how their local sensing of the environment can update their global understanding of the space. However, there exists no unified, standard way to holistically evaluate the performance of these capabilities and allow for comparison across multiple systems. Such measurement tools would be beneficial to manufacturers of these systems and potential customers for procurement. The ASTM F45 Committee on Driverless Automatic Guided Industrial Vehicles [1] develops standards that apply to this type of evaluation, as well as other standards bodies such as ANSI/ITSDF (e.g., B56.5 - 2012 [2]) and ISO (e.g., 13482 [3]). As of the writing of this article, there is no method for evaluating more advanced capabilities, such as the ability to deviate from initial navigation plans based on changes in the environment. To that end, this article presents the design of a standard test method for evaluating navigation and obstacle avoidance capabilities, leveraging some existing constructs from other standards, with the intent of being used as the basis for potential future standards through the ASTM F45 committee.

Background and Related Work

There are a wide range of systems available for use in industrial environments that could be classified as either AGVs or AMRs. While no commonly accepted definitions of each term exist, AGVs typically navigate using more fixed solutions (e.g., following a magnetic tape line on the ground) while AMRs can navigate using more flexible solutions (e.g., reacting to local obstacles detected by 2D and 3D sensors on a high-level path, planned using global map data generated with the same or similar sensors via simultaneous localization and mapping (SLAM) techniques). AMRs can also be found in environments outside of warehouses, including offices, hospitals, and hotels. The capabilities of some systems may fall in between these descriptions (see [4] for a literature review of common AMR capabilities). Regardless, both types of systems are used for similar applications which all require navigation and obstacle avoidance capabilities in order to operate effectively. One of the first products of the ASTM F45 committee was ASTM

F3200-18a Standard Terminology for Driverless Automatic Guided Industrial Vehicles [5]. In this standard, a new term was coined that is used to represent both AGVs and AMRs, recognizing the spread of capabilities between systems that fall into each category:

- <u>Autonomous-Unmanned Ground Vehicle, A-UGV</u>: automatic, automated, or autonomous vehicle that operates while in contact with the ground without a human operator.

Throughout this article, the term "A-UGV" is used to represent the applicable systems that possess navigation and obstacle avoidance capabilities, of which can be measured using the proposed test method. The physical characteristics of A-UGVs can vary in terms of size (e.g., low profile systems like the Omron LD Series [6], large automated fork lift vehicles from American In Motion [7]), locomotion methods (e.g., omniwheels for holonomic movements like the Kinova MOVO [8], wheels for Ackermann steering like the Seegrid vision guided vehicles [9]), and onboard sensors (e.g., 2D lidar is present on almost all A-UGVs, 3D sensors for perception like the Fetch Mobile Manipulator [10], cameras for visual odometry like Seegrid [9]). The software characteristics of A-UGVs will also vary greatly, not just between manufacturers and models of A-UGVs but also on a single deployed A-UGV. An A-UGV may utilize different autonomous behaviors or adjust its control settings (e.g., movement speed) when operating in different parts of an environment, such as adjusting its obstacle avoidance parameters when moving through a narrow passage. Overviews of control and navigation methods for A-UGVs are presented in [11], [12], and [13]. It is necessary to consider each possible characteristic when developing A-UGV performance test methods such that these capabilities can be accurately exercised and highlighted.

Metrics used to evaluate A-UGV navigation capabilities typically refer to measuring the characteristics of the trajectory and path used by the system to move within an environment; these can be grouped into metrics that consider (a) the proximity to obstacles, (b) dimensions of the trajectory towards the goal, or (c) the smoothness of the trajectory [14]. The metrics in groups (a) and (c) are concerned with comparing the accuracy of the trajectory performed by the A-UGV to that which was commanded, which can be analyzed using external optical tracking systems [15]. While movement accuracy is an element of navigation, evaluating at this level assumes the A-UGV is attempting to follow a commanded trajectory. This type of behavior may be true for more traditional AGVs, but not the case for AMRs which may continuously update their navigation plans based on their live sensor data with the main intent of finding the best method to reach the goal. Metrics in group (b) can also be used to measure the accuracy of a performance trajectory if relevant, but can be more generically expressed in terms of the distance and time between the A-UGV and its desired goal. These metrics are directly applicable to a manufacturing environment wherein the throughput rate of performing a task is very important. Additionally, more advanced navigation capabilities – such as those that deal with dynamic environments – may only be able to be measured using more generic metrics as they relate to the performance of a navigation task. The test method presented in this article relies on these metrics, but does not prohibit others (e.g., accuracy of trajectories) from being measured if desired/accessible.

Existing relevant standards do not speak to the measurement of more advanced navigation and obstacle avoidance capabilities. The ANSI/ITSDF B56.5-2012 standard requires certain specifications with regards to a vehicle's safety measures when obstacles are detected via a 2D sensor in order for that vehicle to be compliant with the standard [2]. However, no measures of performance are recorded as part of that standard. So, ASTM F3265-17 Standard Test Method for Grid-Video Obstacle Measurement [16] was developed in order to enable such performance to be measured. That test method involves measuring the velocity of an A-UGV before and after it encounters an obstacle that dynamically enters its path. Additionally, the ISO 13842:2014 standard specifies distances around an A-UGV in relation to its environment and other agents, dictating when emergency stop behaviors should be engaged if an object/agent enters those distance thresholds [3]. These standards are useful for determining the safety characteristics of an A-UGV's behaviors, but they are not concerned with more intelligent capabilities used for decision-making which are the driving capabilities for navigation and obstacle avoidance. A test method for obstacle avoidance was preliminarily developed by Norton and Yanco [17], focusing on building a taxonomy of unique obstacle characteristics and rendering them as standard test props. That work has since evolved and follows in the development lineage of the test method presented in this article. More advanced navigation and obstacle avoidance testing of A-UGVs has been executed towards the generation of performance standards, such as in Yoon and Bostelman [18], utilizing several types of environmental conditions (e.g., grades, gaps), obstacles (e.g., overhanging, moving), and combinations thereof.

The ASTM F45 committee is developing several standards to be used towards performance measurement of A-UGVs. As of the writing of this article, two standard test methods have been published: ASTM F3244-17 Standard Test Method for Navigation: Defined Area [19] and the aforementioned ASTM F3265-17. The ASTM F3244-17 test method evaluates an A-UGV's performance to navigate through a straight aisle or turns between two aisles defined by physical or virtual boundaries from a start location to an end location. The width of the aisles and type of turn between the two adjoining aisles is variable and set by the test requestor (a term from the ASTM F3200-18a standard, meaning the person or organization selecting the tests and defining the conditions under which they are performed). This level of variability allows for the test method to be scaled based on the type of test looking to be run; e.g., narrowing the aisle widths to test for navigation through confined space. This test method is useful for measuring more elemental characteristics of an A-UGV's ability to traverse through an area without colliding into boundaries, but the A-UGV is not required to make more complex decisions as part of its navigation, such as choosing which route to take to reach a goal or replanning its movements around an encountered obstacle.

Additional standards that are currently in development by the ASTM F45 committee are centered around defining larger environments wherein obstacles and other features (such as network impairments) could be rendered in order to exercise these capabilities, but they are not yet standardized. Another test method in development is used to confirm the docking performance of an A-UGV, meaning its ability to position itself repeatably in the environment after navigating to a goal. The test method proposed in this article describes an optional measurement method using grids on the floor to measure the position of the A-UGV at the goal (see the **Apparatus** section). It is intentionally presented from a high-level, intending to align with this in-development test method in the future. There are also standard practices related to recording environmental conditions (ASTM F3218-17 [20]) and an A-UGV's hardware and software configurations (ASTM F3327-18 [21]). The test method presented in this article can also leverage these standards.

Several frameworks for evaluating and defining the capabilities of intelligent systems have been developed including the Complexity Levels of Environment and Obstacles (CLEO) [22], Autonomy Levels for Unmanned Systems (ALFUS) [23], and Performance Measures Framework for Unmanned Systems (PerMFUS) [24]. The CLEO framework was aimed at characterizing an autonomous, mobile system's capabilities for navigation and obstacle avoidance through increasingly complex and dynamic environments. Performance metrics including geometric correctness (between generated maps and the real world) and time to update (the generated maps) were used to evaluate each system in the framework. The ALFUS framework seeks to characterize systems according to their capabilities across three axes: mission complexity, environmental complexity, and human/operator independence (i.e., autonomy level). The PerMFUS framework is an extension of ALFUS, implementing metrics to use in such characterization. While the CLEO framework was aimed specifically at navigation and obstacle avoidance capabilities, ALFUS/PerMFUS were structured such that any autonomous functionality could be characterized.

The previously described standard test method developments from ASTM F45 may eventually connect to some of the more advanced capabilities described in these frameworks, but as of the writing of this article they do not. The test method proposed in this article is aimed at exercising more advanced navigation and obstacle avoidance capabilities of A-UGVs, which can be used to characterize performance under either of these frameworks. A more dedicated effort to characterize A-UGVs specifically is currently under development within the ASTM F45 committee by defining the set of distinct capabilities that an A-UGV may possess. The intention is to define capabilities that could be demonstrated in a standard test method, allowing manufacturers to claim these capabilities according to the standard. Versions of this concept have been previously published as A-UGV capability levels [25], and will eventually be proposed as a standard. The test method presented in this article is designed such that it could be used to demonstrate some of these capabilities.

Scope

The test method described in this article is concerned with measuring autonomous navigation and obstacle avoidance capabilities of A-UGVs. While these concepts are generally understood throughout the AGV and AMR communities, there are several terms used in this article whose definitions could be misunderstood. To this end, the following definitions from ASTM

F3200-18a [5] apply to this article:

- Localization: Ability of the A-UGV to determine its pose within an environment map.
- Navigation: Deciding on and controlling the direction of travel derived from localization and the environment map.
- Obstacle: Static or moving object or feature not present in the map, that obstructs the intended movement.
- Obstacle avoidance: Autonomously avoiding impact with obstacles (for example, stopping, driving around).

There are two additional terms that are used to describe A-UGV navigation and obstacle avoidance capabilities covered in this article that are not covered by the ASTM F3200 terminology standard. For the scope of this article, they are as follows:

- Route: An area defined by boundaries which can be traversed by a A-UGV from a start location in order to reach a goal location; multiple routes may exist in a space if there are multiple unique areas each of which require different minimum distances to be traversed in order to reach a goal location from a start location (e.g., aisles in between racking in a warehouse; hallways throughout a building).
- Path: The trajectory taken by the A-UGV when traversing through a route.

Using the test method described in this article, the A-UGV will be faced with various challenges that may require it to alter which route it traverses and/or make updates to its planned path in order to complete a task. An important factor to these behaviors is the A-UGV's knowledge about its environment before executing a task that can be used for planning (*a priori*) and updates to its knowledge while performing a task (*in situ*). When an A-UGV begins performing a navigation task, it likely contains a priori knowledge such as a map of the environment that contains the position of boundaries (defining the routes that can be taken) and waypoints that define start and goal locations. Using just this information, the A-UGV can plan to traverse a specific route along a particular path in order to reach the goal. If elements of the environment are encountered in situ that are misaligned with the A-UGV's map – such as encountering obstacles – this may influence the route and/or path it takes to reach the goal. If the goal is no longer reachable by the A-UGV, it may require human intervention. The A-UGV may also update its map to include the obstacles it encountered such that it can plan around those obstacles in the future should it traverse that area again. The updates to the map may also decay after a certain amount of time (which could be instantaneously) or be overwritten if the environment changes again causing the A-UGV to update its map, which is common in the lifecycle of A-UGVs with this functionality. Gill et al. [26] utilized the ASTM F3244-17 [19] test method to evaluate the performance of their A-UGV, expanding its usage to include similar incongruities between the A-UGV's internal model and the physical reality of the environment, simulating the types of issues that may arise when long-term autonomy is deployed. In order to similarly exercise the robustness of an A-UGV's navigation capabilities like these, the test method presented in this article specifies several knowledge conditions which will vary misalignments between the A-UGV's internal model of the environment and the true physical nature of the environment. This aspect is referred to as "model misalignment" throughout the remainder of this article.

The software configuration of the A-UGV will determine what strategy is used. Some strategies are better suited for certain deployment scenarios, but the test method described in this article do not dictate which should be used (although the test requestor may impose specific conditions to be used). Rather, performance in the test method will elicit the A-UGV's behaviors and allow them to be demonstrated in a standardized manner. As such, the test method are designed to allow for A-UGVs to demonstrate capabilities such as utilizing the shortest route and/or path to reach a goal, updating their navigation plans while avoiding obstacles, and maintaining localization so they don't get lost while navigating, among others.

The characteristics of the A-UGV must also be considered; the test method described in this article are designed to be scalable such that any A-UGV can be tested. This includes the physical dimensions of the A-UGV and if it is deployed with additional components as part of an A-UGV System (or A-UGVS, defined in ASTM F3200 as an A-UGV and all associated components, equipment, software, and communications necessary to make a fully functional system), such as an enterprise controller that commands traffic throughout a facility and uses additional external sensors to influence behavior. The ASTM F3327-18 Standard Practice for Recording the A-UGV Test Configuration [21] can also be used to capture each of these characteristics prior to testing. The type of sensors used by the A-UGV in order to maintain localization and/or detect and avoid obstacles may also influence the characteristics of the boundaries and the obstacles used in the test method.

The intention is for the development of this test method to coalesce with the standards generated by the ASTM F45 committee and borrows many of its testing conventions (e.g., fault conditions, success criteria, etc.) from existing standards or work items in development. To that end, the test method proposed in this article holistically evaluates the navigation and obstacle avoidance capabilities of an A-UGV, rather than evaluating the performance of the A-UGV's components that enable these types of capabilities. To follow with some of the basic tenets of standard test method development, the materials used to build the apparatus and measurement methods used for evaluation should not be cost prohibitive. The test method described in this article is intended for usage across the A-UGV community, either as a tool used during research and development to measure capabilities, or during a procurement process to validate an A-UGV's performance. In either case, expensive testing would limit broader proliferation of the test method. More detailed and complex analysis methods may be applied to the test method if desired (see the **Performance Measures** section), but are not required.

Requirements

All of the previously described factors of the test method's scope form a set of requirements for the development of the test method. They are distilled as follows:

- **R1**: All systems qualified as A-UGVs (i.e., both AGVs and AMRs) must be able to utilize their navigation capabilities when performing in the test method.
- **R2**: Due to the variety of possible A-UGV characteristics – including dimensions, sensors, control algorithms, and locomotion methods – the test method must be malleable and scalable to accommodate the characteristics of the A-UGV being tested.
- **R3**: Due to the intertwined nature and variety therein of the components of A-UGV navigation (e.g., path planning, locomotion, localization, obstacle avoidance, motion control, etc.), performance must be evaluated holistically.
- **R4**: Due to the variety of operational requirements for different A-UGV scenarios and user needs, the success criteria for the test method must also be malleable.
- **R5**: To coalesce with ASTM F45 standards, components and structures of existing and in-development standards should be leveraged.
- **R6**: Test apparatus fabrication and measurement techniques must be performed using cheap, readily available materials and methods that are inexpensive.

Test Method

The underlying structure of the test method allows for a multitude of test configurations, each of which exercise different A-UGV capabilities related to navigation and obstacle avoidance. Regardless of each individual testing configuration, the bounding dimensions of the apparatus are maintained (based upon a standard unit, *u*, which is influenced by the A-UGV being tested; see the **Apparatus** section), the same procedure is used, and the commands given to the A-UGV to perform the task are always the same: navigate from the start location to the goal location, which is either from point A to B or B to A. For all test configurations, the A-UGV traverses through the apparatus attempting to reach its goal while negotiating any obstructions placed in the apparatus as well as any model misalignments it encounters. The task is performed multiple times in order to reach statistical significance of the performance measures. Rules for the test method dictate fault conditions that will result in the failure of a task repetition (and, subsequently, of the entire test), but the success criteria for each test can vary and is defined prior to performing the test as different scenarios will call for different expected performance. Each component is further detailed in the following subsections.

ASTM Standards		Summary	Relevance to Proposed Test Method
Definitions	ASTM F3200-18a Standard Terminology for Driverless Automatic Guided Industrial Vehicles [5]	Defines many terms related to A-UGV performance testing	Terms referenced include A-UGV, localization, navigation, obstacle, and obstacle avoidance
	ASTM WK65139 Standard Guide for A-UGV Capabilities (in development)	Defines various possible A-UGV capabilities that one can possess	By varying test configuration parameters, A-UGV capabilities as defined in this standard including those in the goal navigation (pre-programmed and in situ), localization, and obstacle avoidance categories can be demonstrated
Testing Capabilities	ASTM F3244-17 Standard Test Method for Navigation: Defined Area [19]	Evaluates ability of an A-UGV to navigate within boundaries to reach a goal	Apparatus boundary definition and fault conditions from this standard are leveraged
	ASTM WK57000 Standard Test Method for Docking Driverless Automatic Guided Industrial Vehicles (in development)	Evaluates ability of an A-UGV to position itself at a goal position	Position measurement at a goal via paper grids in the environment from this standard can be used (optional)
	ASTM F3265-17 Standard Test Method for Grid-Video Obstacle Measurement [16]	Evaluates ability of an A-UGV to reduce velocity when an obstacle is detected	A-UGV velocity measurement via paper grids in the environment from this standard can be used (optional)
	ASTM WK65141 Standard Guide for Combining A-UGV Standards (in development)	Defines how to combine multiple standards together to conduct more complex testing of an A-UGV (e.g., navigating through multiple defined areas while avoiding obstacles and with communication impairments)	The proposed test method is essentially a combination of multiple standards (in addition to some unique elements) and will be definable by utilizing this standard
Describing Conditions	ASTM F3218-17 Standard Practice for Recording Environmental Effects for Utilization with A-UGV Test Methods [20]	Describing environmental conditions (e.g., lighting, temperature, ground surface) when conducting a test with an A-UGV	Utilize this standard to describe the environment when conducting a test
	ASTM F3327-18 Standard Practice for Recording the A-UGV Test Configuration [21]	Describing characteristics of an A-UGV that is to be tested	Utilize this standard to describe the A-UGV when conducting a test
	ASTM WK60390 Standard Practice for Describing Stationary Obstacles Utilized within A-UGV Test Methods (in development)	Describing the characteristics of stationary obstacles used when conducting a test with an A-UGV	Utilize this standard to describe any stationary obstacles present when conducting a test
	ASTM WK68031 Standard Practice for Describing Moving Obstacles Utilized within A-UGV Test Methods (in development)	Describing the characteristics of moving obstacles used when conducting a test with an A-UGV	Utilize this standard to describe any moving obstacles present when conducting a test (out of scope for this article)
	ASTM WK54431 Standard Practice for Performance Testing of an A-UGV Under Varied Communication Conditions (in development)	Describing and implementing impairments to the communication channels between an A-UGV, its controller, and/or other infrastructure when conducting a test	Utilize this standard to describe and implement any communication impairments present when conducting a test

Table 1. Relationship between the proposed test method and ASTM F45 standards that are published or are in development.

This test method leverages some of the testing [15] components utilized in ASTM F3244-17 [19]. In that test method, an A-UGV traverses through a straight aisle or turns between two aisles from a start location to an end location. This method also leverages several concepts in other standard test methods currently in development by the ASTM F45 committee that are not fully standardized as of the writing of this article; see Table 1 for an overview of the proposed test method's relevance to existing standards and those in development.

Apparatus

The test apparatus is comprised of an area defined by boundaries containing several possible routes for the A-UGV to traverse in order to navigate between two locations: A and B. Each route is made up of a series of squares that connect together to form a contiguous segment of space that the A-UGV can traverse through (see Figure 1). A standard unit, u, is used to dictate the dimensions of the squares, in turn setting the width and length of each continuous section within a route. The value of u should be larger than the A-UGV's width (w) such that if the width of an aisle is set to u, the A-UGV could fit through it. From here, the value of u can increase to allow for room on either side of the A-UGV as it traverses, to allow room for obstacles to be placed into the environment that the A-UGV can navigate around, or can be sized to match a target implementation environment. The value of u is intended to be flexible so that a variety of environments can be represented in the test method, allowing for any A-UGV to be tested. It is recommended that the value of u be no smaller than $3w$ for exploratory testing of an A-UGV's navigation capabilities to allow for some obstacles to also be used (see the **Obstacle Configurations** section).

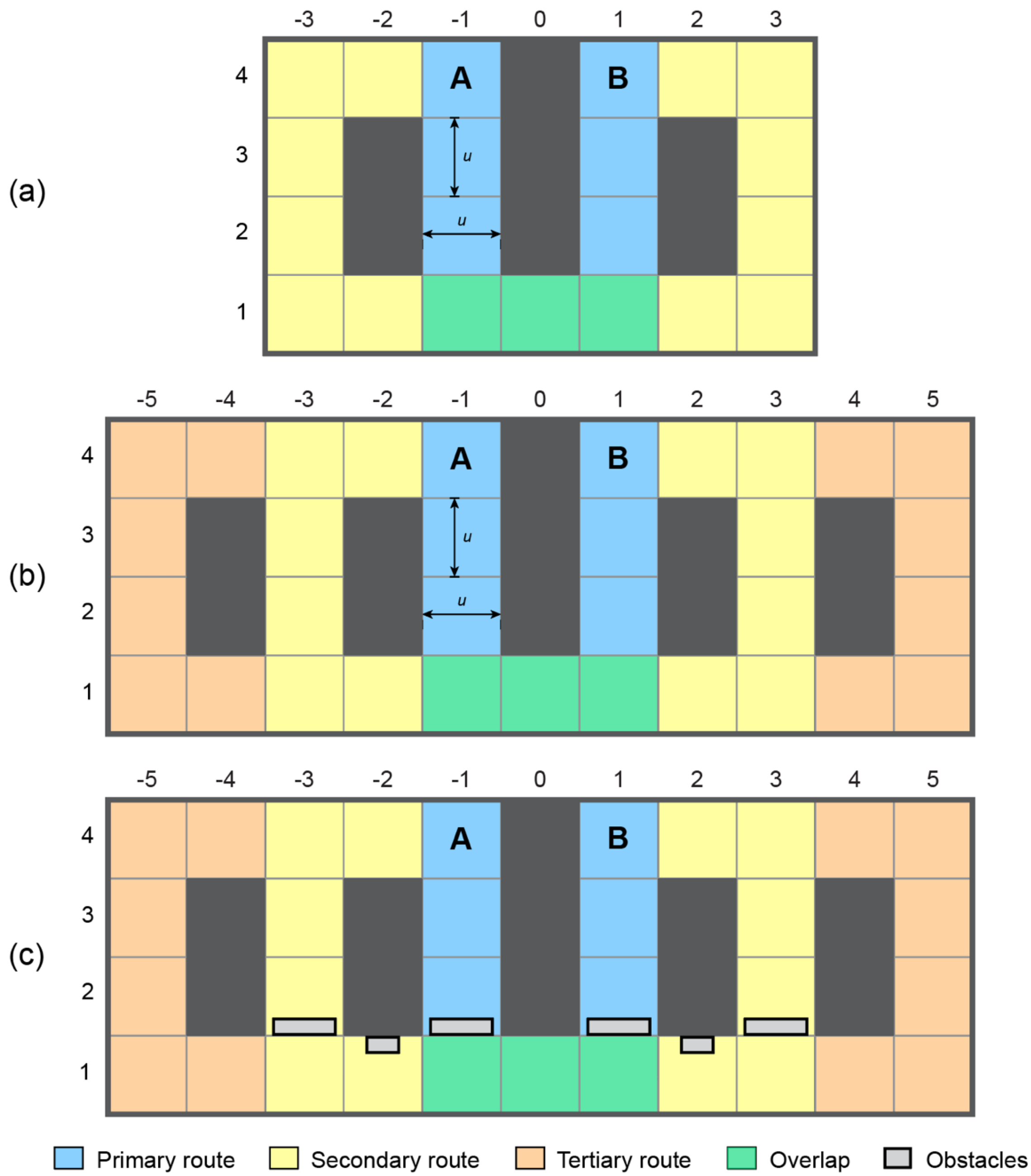

Figure 1. Two variations of the test apparatus: (a) simplified version without tertiary route, (b) larger version with tertiary route. An example layout of symmetrically placed obstacles is shown in (c). Each location block in the apparatus can be identified by using a pair of X,Y coordinates.

Two versions of the test apparatus are shown in Figure 1: (a) with primary and secondary routes, (b) with primary, secondary, and tertiary routes. Depending on the value of u, the size of the test apparatus may be rather large and difficult to implement, so the smaller version (a) is provided as an alternative, however the lack of a tertiary route will limit the types of capabilities that can be evaluated (see the **Test Configurations** section). The boundaries of the apparatus can be physical (e.g., wood panels, steel panels, fencing) or virtual (e.g., positions defined in the A-UGV's map, possibly marked by tape on the floor); this is consistent with ASTM F3244-17 [19].

Most A-UGVs are designed to be symmetrical in terms of their dimensions, shape, weight, locomotion methods, and sensor coverage, but their behaviors may not be consistently symmetrical. The test apparatus is designed to be symmetrical in order to exercise this factor; i.e., when the A-UGV is commanded to traverse from A to B it is required to turn left, and when commanded from B to A it is required to turn right. This symmetry should be maintained in that if any obstructions are added (see the **Obstacle Configurations** section), an identical obstacle should be added to the opposite side of the apparatus such that their positions are symmetrical; see Figure 1 for an example. Each side of the apparatus can be referred to as the A side or the B side based on which location is found on that side.

By maintaining symmetry, the primary route requires the least amount of distance to be traversed (8*u*) when navigating A to B or B to A, and therefore likely the least amount of time to navigate making it the optimal route to take. However, if the primary route is obstructed in any way, the A-UGV may instead opt to take the secondary route in order to reach the goal, which requires more distance to be traveled (16*u*) and most likely will take longer to navigate than the primary route. If both the primary route and the secondary route are obstructed (as is the case in Figure 1(c)), the tertiary route requires even more distance to be traveled (24*u*) and more time to navigate than the secondary route.

Each location – A and B – is notionally defined as the center of the square it occupies, shown in Figure 1. The A-UGV's understanding of this location may be defined virtually in its map (e.g., x,y coordinates), physically (e.g., QR codes on the floor, the end of a magnetic tape line), or as a combination of both. Tape lines are added to the edges of the A and B location squares for ground truth measurement of when the A-UGV reaches each location: when the full body of the A-UGV crosses the line, it has reached the location (see Figure 2). This measurement method is consistent with ASTM F3244-17. For more precise measurements regarding the A-UGV's end position once it has finished navigating, additional ground truth measurements can be implemented via grids on the floor centered around the intended end position of fiducials on the A-UGV once it reaches its goal (see Figure 2). This measurement method is currently under development by the ASTM F45 committee as a standard test method.

For the scope of this article, the test method is limited to the straight boundaries and 90 degree turns shown in Figure 1. However, it is possible that the test apparatus could be configured such that the boundaries were angled more or less than 90 degrees to form each turn, or the boundaries could be curved. Doing so may introduce complexities in maintaining a constant width of u throughout the apparatus. Those implementing the test method can configure the apparatus as such in order to fit a specific application if desired, but the standard measurement of u must be maintained throughout and each side of the apparatus must be symmetrical to the other.

Obstacle Configurations

Obstacles can be positioned within the apparatus in order to influence the behavior of the A-UGV under test. The obstacles will either partially or fully obstruct one of the routes that can be taken, causing path and/or route deviations. Obstacles can be positioned anywhere within the apparatus – except within locations A and B – so long as a copy of that obstacle is positioned in a symmetrical location (i.e., it's location on the A side of the apparatus is mirrored on the B side).

Obstacles can be positive (i.e., making contact with or elevated above the ground plane; e.g., boxes, chairs) or negative (i.e., a void from the ground plane down; e.g., holes, cracks). Obstacle designs are not specified as part of the test method as any obstacle that is considered relevant for the test being conducted can be utilized, whether it be a real object (a "genuine obstacle") or a fabricated prop used solely for testing (an "artifact obstacle"). Rather, several unique configurations of the obstacles' dimensional and positional relationship to the apparatus boundaries are specified, each of which can be used to exercise different A-UGV capabilities.

Figure 2. Counter-clockwise from top left: Photos of an A-UGV in the test apparatus with obstacles; the A-UGV leaving location A by crossing the yellow line; the A-UGV reaching location B by crossing the yellow line; measuring the accuracy of the A-UGV's end position at a location by using grids on the floor and fiducials on the A-UGV (pencils) that correspond with the center of each grid; detail of positioning accuracy measurement method.

When selecting the value of u, the available space left for the A-UGV to traverse when an obstacle is present within the apparatus must be considered. The A-UGV's expected behaviors will change depending on the dimensions of the A-UGV, the obstacles present, the value of u, and the A-UGV's navigation configuration (e.g., how much distance it will attempt to leave between it and the obstacle). Regardless of the specifics of these relationships, three obstacle configuration categories are defined based on the minimum available traversal space between an obstacle and the apparatus boundaries and/or other obstacles (referred to as "available traversal space"): minor, moderate, and severe. See Table 2 for definitions of each configuration. The obstacle configuration categories are intended to be used as a high-level description of obstacle placement within a test; the specific locations of the obstacles, their characteristics (e.g., dimensions, shapes, materials), available traversal space, etc., should always be reported alongside the configuration category. See Figure 3 for example obstacle layouts of each obstacle configuration category. Note that in these layouts any obstacle can be used so long as the width of the volume the obstacle occupies complies with all the necessary conditions of an obstacle configuration category.

Obstacle Configuration Category	**Available Traversal Space**	**A-UGV Capability Focus**
Minor	>200% of the A-UGV width	Path deviation to avoid obstacles without deviating from route
Moderate	≤200% and >100% of the A-UGV width	Path and/or route deviation to avoid obstacles
Severe	≤100% of the A-UGV width	Route deviation to avoid obstacles

Table 2. Definitions for the three obstacle configuration categories and the affiliated A-UGV capabilities they are intended to exercise.

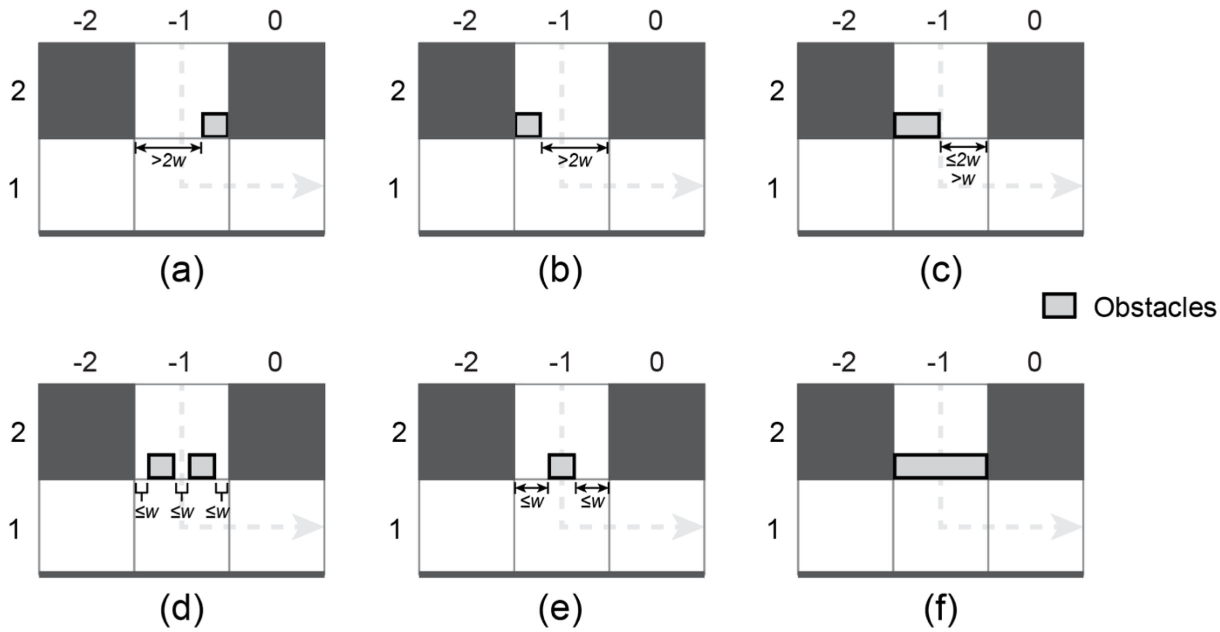

Figure 3. Layouts of notional obstacle configurations: minor (a, b), moderate (c), and severe (d-f). The grey dotted line arrow shows the intended direction of travel for the A-UGV to traverse to its goal.

The characteristics of the obstacles to be used for each configuration category can be determined empirically per each A-UGV width (*w*) and the chosen *u* value. A set of obstacles may be selected before the value of *u* is set and be used to determine what its value should be. The determination of these variables is intentionally flexible.

Knowledge Conditions

Three conditions for model misalignments are specified in order to evaluate the robustness of an A-UGV's navigation capabilities. Each condition varies what information is given to the A-UGV a priori and what information must be gained in situ during task execution. The definitions of each knowledge condition are as follows:

- Condition 1: Obstacles in the physical space are not present on the map the A-UGV uses to navigate, meaning the a priori knowledge given to the A-UGV is inaccurate. The A-UGV must detect the obstacles in situ or else it could collide with the obstacles.
- Condition 2: Obstacles in the physical space are present on the map given to the A-UGV, meaning the a priori knowledge given to the A-UGV is accurate. The A-UGV can treat the edges of obstacles as boundaries and can use them for localization (if relevant).
- Condition 3: Obstacles are present on the map given to the A-UGV, but they are not in the physical space, meaning the a priori knowledge given to the A-UGV is inaccurate. The A-UGV may rely solely on its internal map to plan its movements, potentially causing it to take a longer route than necessary or causing it to believe it is stuck.

Each of these conditions could conceivably occur naturally to an A-UGV during day-to-day operations wherein an obstacle is encountered that is not part of its map (condition 1), which gets added to the map such that it can be avoided in the future (condition 2), but then the obstacle is moved (condition 3). As such, each obstacle added to the test apparatus is assigned a knowledge condition as part of specifying a test configuration. By explicitly implementing these knowledge conditions for a given test, each scenario can be objectively evaluated to gain a better understanding of the A-UGV's capabilities in the presence of such uncertainty. It is also possible to conduct testing wherein the knowledge condition updates throughout performance of the test (paired with adjusting the presence/absence of obstacles), more like that of a real scenario, but the implementation of dynamic conditions is not in scope of this article.

Another factor that will affect an A-UGV's navigation capabilities is its treatment of newly learned knowledge. Depending on the A-UGV's software configuration, detecting model misalignment may spur an update to its knowledge, such as updating its map of the environment by either adding elements that are newly detected (condition 1) or removing those that are no longer found (condition 3). This aspect is specifically targeted when using the second variation of the test apparatus with the tertiary route (see Figure 1(b)) and there are severe obstacles in the primary and secondary route, blocking them from being fully traversed. For example, if testing using condition 1, the expected order of operations is:

1. The A-UGV attempts to take the primary route to reach the goal, but detects an obstacle blocking the way.
2. The A-UGV replans to take the secondary route to reach the goal.
3. The A-UGV attempts to take the secondary route to reach the goal, but detects an obstacle blocking the way.
4. The next expected action depends on how the A-UGV's configuration is set:
    a. If the A-UGV maintains the newly gained knowledge of the obstacle blocking the primary route, then the A-UGV should replan to take the tertiary route to reach the goal.
    b. If the A-UGV does not maintain the knowledge of the obstacle blocking the primary route, then it may replan to take the primary route to reach the goal.
5. Depending on what happened during step 4:
    a. The A-UGV takes the tertiary route and goes through the same route replanning on the opposite side of the apparatus, likely reaching the goal.
    b. The next action is the same as step 1 and the A-UGV becomes stuck in a loop or determines that there is no valid path, likely never reaching the goal.

Regardless of the A-UGV's configuration for maintaining knowledge while traversing to a goal location, the A-UGV must lose all gained knowledge at the end of a task repetition. If the

A-UGV were allowed to maintain its knowledge in between repetitions of navigating through the same environment, then testing performed wherein in situ knowledge must be gained (i.e., conditions 1 and 3) would only be exercised during the first repetition. Such a test would then effectively become condition 2 after the first repetition. More information regarding the number of task repetitions is provided in the **Procedure** section.

Test Configurations

The test method can be configured by selecting a combination of obstacle configuration categories, obstacle characteristics, obstacle positions, and knowledge conditions. The unique scenarios that each combination specifies will require different A-UGV capabilities in order for the A-UGV to successfully navigate to the goal. When specifying a test configuration, a simple abbreviated format is used:

- <u>Routes</u>: Primary = P, Secondary = S, Tertiary = T
- <u>Obstacle Configurations (OC)</u>: Minor = $_{mi}$, Moderate = $_{mo}$, Severe = $_{se}$, Empty = $_{e}$
- <u>Knowledge Conditions (KC)</u>: 1 = $_{1}$, 2 = $_{2}$, 3 = $_{3}$
- <u>Test configuration identifier format</u>: $P_{OC,KC}S_{OC,KC}T_{OC,KC}$

Some example test configurations are defined in Table 3 and can be seen in Figure 4. Each explicitly exercises different A-UGV navigation capabilities or a combination thereof. While the example test configurations shown only utilize minor and severe obstacles, moderate obstacles could be used in place of either. Depending on the specific dimensions of the moderate obstacle and the A-UGV's navigation configuration, the expected behavior of the A-UGV will vary, whereas the expected A-UGV behavior should be more consistent across minor and severe obstacles (path deviations and route deviations, respectively). Similar A-UGV capabilities can be tested in each variation of the test apparatus. For example, $P_{se,1}S_{se,3}$ in version (a) and $P_{se,1}S_{se,1}T_{se,3}$ in version (b), both require route updates based on in situ knowledge and the a priori knowledge given at the start must get updated based on gained in situ knowledge (i.e., recognizing that the secondary or tertiary routes, respectively, are actually free of obstacles) in order to reach the goal. However, test configurations that are intended to exercise an A-UGV's capability at maintaining knowledge can only be performed in version (b) of the test apparatus as it requires the tertiary route.

For all test configurations, an A-UGV's obstacle avoidance capabilities can be further exercised by varying the characteristics of the obstacles used (e.g., positive, negative, basic shapes that only require 2D sensing, complex shapes that require 3D sensing, etc.), so long as their overall dimensional characteristics as described in the **Obstacle Configurations** section are maintained. More than one obstacle can be applied to a single route if more cluttered environments are to be simulated, such as using multiple minor obstacles to further exercise path planning.

It is recommended that a $P_eS_eT_e$ test configuration be used as a baseline (i.e., all routes free of obstacles) in order to compare A-UGV performance in other configurations.

Obstacle Configurations and Knowledge Conditions per Route			A-UGV Capability Focus	Test Configuration Identifier
Primary	**Secondary**	**Tertiary**		
Minor, 1	Empty	Empty	Path updates based on in situ knowledge	$P_{mi,1}S_eT_e$
Minor, 2	Empty	Empty	Path planning based on a priori knowledge	$P_{mi,2}S_eT_e$
Minor, 3	Empty	Empty	Path planning based on a priori knowledge which could get updated based on in situ knowledge	$P_{m,3}S_eT_e$
Severe, 1	Empty	Empty	Route updates based on in situ knowledge	$P_{se,1}S_eT_e$
Severe, 2	Empty	Empty	Route planning based on a priori knowledge	$P_{se,2}S_eT_e$
Severe, 3	Empty	Empty	Route planning based on a priori knowledge which could get updated based on in situ knowledge; updating based on in situ knowledge will result in less distance traveled	$P_{se,3}S_eT_e$
Severe, K	Minor, K	Empty	Same as $P_{se,k}S_eT_e$ plus path updates based on in situ knowledge	$P_{se,k}S_{mi,k}T_e$
Severe, 1	Severe, 1	Empty	Route updates based on in situ knowledge, but knowledge must be maintained in order to reach the goal	$P_{se,1}S_{se,1}T_e$
Severe, 2	Severe, 2	Empty	Route planning based on a priori knowledge	$P_{se,2}S_{se,2}T_e$
Severe, 3	Severe, 3	Empty	Route planning based on a priori knowledge which could get updated based on in situ knowledge; updating based on in situ knowledge will result in less distance traveled	$P_{se,3}S_{se,3}T_e$
Severe, K	Severe, K	Minor, K	Same as $P_{se,k}S_{se,k}T_e$ plus path updates based on in situ knowledge	$P_{se,k}S_{se,k}T_{mi,k}$
Severe, 1	Severe, 1	Severe, 1	Route updates based on in situ knowledge, but regardless of whether or not the knowledge is maintained, it should recognize that it is stuck	$P_{se,1}S_{se,1}T_{se,1}$
Severe, 2	Severe, 2	Severe, 2	Route planning based on a priori knowledge, which should recognize that it is stuck	$P_{se,2}S_{se,2}T_{se,2}$
Severe, 3	Severe, 3	Severe, 3	Route planning based on a priori knowledge which could get updated based on in situ knowledge; updating based on in situ knowledge will result in it being able to reach the goal	$P_{se,3}S_{se,3}T_{se,3}$
Severe, 1	Severe, 1	Severe, 3	Route updates based on in situ knowledge, but knowledge must be maintained so as to not get stuck, and a priori knowledge must get updated based on in situ knowledge in order to reach the goal	$P_{se,1}S_{se,1}T_{se,3}$

Table 3. Example test configurations using a combination of obstacle configurations and knowledge conditions. Visualizations of these configurations can be seen in Figure 4. A knowledge condition value of K is shown for test configurations when 1, 2, or 3 could be used.

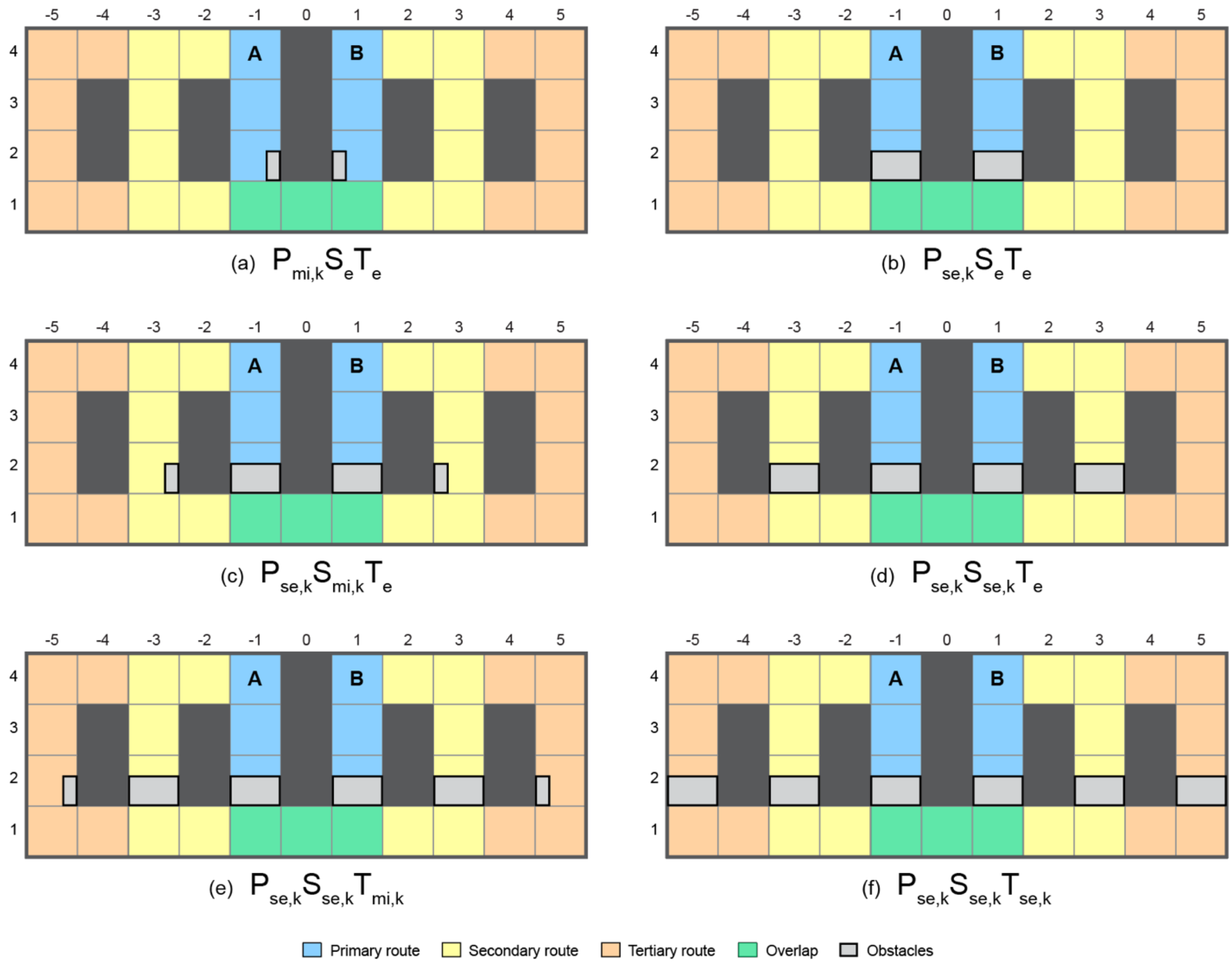

Figure 4. Example test configurations with minor and severe obstacles placed in varying routes. A knowledge condition value of K is shown for all configurations as conditions 1, 2, or 3 could be used.

Success Criteria

Success criteria is set by the test requestor to dictate the type of performance expected to be demonstrated by the A-UGV in order to deem the entire test successful and individual task repetitions successful. The test success criteria is based on achieving the desired sample size of successful task repetitions. To be consistent with ASTM F3244-17 [19], the minimum required test success criteria is 30 successful task repetitions (meaning 30 A-B or 30 B-A). This is based on the confidence and probability of success threshold values associated with achieving 29 successful repetitions with zero failures (see Table 4), which is then rounded up to 30 for simplicity. Task success criteria is set based on the expectations of the A-UGV for a given test configuration. It can be set as either "goal reaching" or "stuck state" task success criteria, meaning a task repetition is deemed successful if the A-UGV reaches its goal or detects that it is stuck and cannot reach its goal, respectively. For example, in the $P_{se,1}S_{se,1}T_{se,1}$ test configuration, the A-UGV cannot physically reach the goal, so stuck state task success criteria would be set as it is expected to recognize that it is stuck. Maximum task time (i.e., when the A-UGV crosses the tape lines) and positioning accuracy (i.e., offset distance of the fiducials on the A-UGV from the center of the grids on the floor) can also optionally be set as task success criteria.

		Probability of Success Threshold		
		0.99	**0.95**	**0.90**
Confidence	**0.99**	459	90	44
	0.95	299	59	29
	0.90	230	45	22
	0.85	189	37	19
	0.80	161	32	16

Table 4. Number of repetitions required to achieve different confidence measured against the probability of success threshold with zero failures. This table is referenced from ASTM F3244-17 [19].

Regardless of test configuration, several fault conditions can occur that will result in failing a task repetition and subsequently the entire test (as zero failures are allowed). The fault conditions are as follows:

- If the A-UGV makes contact with and/or crosses the boundaries of the apparatus or the obstacles.
- If human intervention (e.g., hitting the emergency stop) occurs during task performance (i.e., while the A-UGV is navigating).

- If the A-UGV exceeds the maximum task time (if set).

Procedure

After determining the value of *u*, fabricating the apparatus, laying ground truth measurement markers (i.e., tape lines and/or positioning grids on the floor), and selecting obstacles to match all desired obstacle configurations to be tested, the following steps are taken to conduct a test in any test configuration:

1. Select the desired test configuration and define the success criteria for the test.
2. Position all obstacle configurations with knowledge condition 2 or 3 in the apparatus.
3. Command the A-UGV to build its a priori knowledge of the environment (i.e., make a map of the apparatus and any obstacles present) including its understanding of the A and B locations.
4. Position the A-UGV at location A.
5. Position all obstacle configurations with knowledge condition 1 in the apparatus. Remove all obstacle configurations with knowledge condition 3 from the apparatus.
6. Command the A-UGV to navigate to location B.
7. The next step depends on how the task success criteria is set:
    a. If the goal reaching task success criteria is met, then the A-B task repetition is successful. Clear all knowledge gained by the A-UGV during task performance to return its a priori knowledge state to that of step 3.
    b. If the stuck state task success criteria is met, then the A-B task repetition is successful. Clear all knowledge gained by the A-UGV during task performance to

return its a priori knowledge state to that of step 3. Manually reposition the A-UGV at location A and skip to step 10.

8. Command the A-UGV to navigate to location A.
9. If the goal reaching task success criteria is met, then the B-A task repetition is successful. Clear all knowledge gained by the A-UGV during task performance to return its a priori knowledge state to that of step 3.
10. Repeat steps 6-9 for the desired sample size.
11. Record timing data (and positioning data if desired) throughout the test.
12. Calculate the performance measures.

Performance Measures

Task time is recorded based on when the full body of the A-UGV has crossed the tape line at the start location and once the full body of the A-UGV has crossed the tape line at the goal location (if goal reaching task success criteria is used) or when the A-UGV stops navigating (if stuck state task success criteria is used). Internal measures provided by the A-UGV of its location within the test apparatus should not be used for timing data as it can be prone to localization inaccuracies and latency issues. Rather, this measurement can be recorded via external measurement through direct observation of the A-UGV as it crosses the tape lines. Timing data should be reported for all task repetitions as well as an average with standard deviation. Positioning accuracy can be reported if desired; a standard test method is currently under development by the ASTM F45 committee in order to measure this capability.

If the task success criteria is met and no fault conditions are incurred during a task repetition, then that task repetition is successful. If the test success criteria is met meaning the required number of successful task repetitions was achieved, then the entire test is successful. All reports of performance measures should be coupled with the test configuration (similar to Table 3) and specific details regarding the obstacle characteristics and positions. Video footage should be taken of all test performance from multiple angles in order to verify navigation timing data and to determine if any collisions with the apparatus and/or obstacles occurred. If desired/feasible, more expensive measurement methods could be used to gain higher-resolution ground truth data of the A-UGV's movements, such as through the use of a motion capture system. This detailed level of measurement is not necessary, but optional.

Example Performance Data

In order to exercise this test method, validation testing was performed using a commercially available research A-UGV platform, the Fetch Robotics Mobile Manipulator [10], a commercially available research robot platform. The A-UGV was configured using a rudimentary Robot Operating System (ROS) [27] navigation stack, utilizing the *AMCL* [28] package for localization and the *move_base* [29] package for path planning. It should be noted that all performance measures presented in this section are not wholly representative of the A-UGV's capabilities, and are presented only as a means to demonstrate use of the test method. Under a different software configuration the A-UGV may have performed differently.

To perform this testing, the width of the A-UGV (w) was measured (~500mm) and the value of u was set ($3w$ = 1500mm). Using these measurements, a test apparatus was fabricated. Due to available space in the laboratory, test apparatus version (a) was used (i.e., without the tertiary route). 27 unique test configurations in addition to a baseline were performed. Test success criteria was set to 20 total repetitions in both directions (i.e., 10 A-B and 10 B-A) for all tests with goal reaching task success criteria, and 5 repetitions in only the A-B direction for all tests with stuck state task success criteria. Given limited resources for conducting testing, a variety of test configurations with less than the required minimum task repetitions (i.e., 30) was favored over conducting very few test configurations with the necessary repetitions. A summary of the performance data can be found in Table 5. See Figure 5 for screenshots from videos of test performance showing each unique obstacle configuration.

During testing, the A-UGV was logging its position within its internal map of the environment. These positions reported by the A-UGV were not used to calculate performance data, but rather provide insight into some of the issues encountered by the A-UGV. For example, during test 25, the A-UGV collided with one of the apparatus boundaries, resulting in a failed test. By plotting the positions reported by the A-UGV (i.e., where the A-UGV thinks it is), it can clearly be seen that a localization error occurred, likely causing the collision; see Figure 6 for plots of the A-UGV's position according to its localization data of some example tests. As shown in Figure 6(d), the A-UGV believed it was actually on the B side of the apparatus (which was incorrect) just prior to the collision.

Some inconsistent behavior was exhibited by the A-UGV, particularly in test configurations utilizing moderate obstacle configurations. For example, during test 4 ($P_{mo,1}S_e$), the A-UGV took the secondary route for 3 repetitions, sometimes when traversing A-B and sometimes B-A, resulting in a high standard deviation (12.8 seconds). Similarly, during test 9 ($P_{mo,3}S_e$), the A-UGV took the secondary route 5 times, but only when traversing on the A side when traversing B-A, demonstrating a lack of symmetry in its performance. For all tests using obstacles with knowledge condition 3, the A-UGV navigated as if its internal model of the environment was accurate (i.e., avoiding the supposed locations of the obstacles even though they were not physically present). In test 18, the A-UGV took the secondary route for all repetitions, even though it technically could have used the primary route to minimize task time. However, depending on the intended application environment for the A-UGV, prioritizing its internal model of the environment over the physical reality of the environment in this way may be preferred. For example, if the A-UGV's map contained the positions of safety barriers blocking the top of a stairway, but someone unknowingly moves the safety barriers now leaving an opening; it would likely be desirable for the A-UGV to continue to behave as if the safety barriers were in place.

Discussion and Future Work

This test method has been exercised with one A-UGV thus far as an initial demonstration of how it can be used to evaluate navigation and obstacle avoidance capabilities. Additional A-UGVs of varying characteristics (e.g., different sizes, locomotion methods, sensors, etc.) will be exercised in the test method, including platforms intended for research applications and those aimed for use in real world deployments in industry. Those falling into the latter category should be designed with more hardened, consistent behaviors to demonstrate more repeatable performance measures. For future testing, the minimum task repetitions (30) will be used in order to produce more accurate testing results. The apparatus set up will also be expanded to include the tertiary route such that in situ knowledge maintenance can be evaluated. More advanced A-UGV capabilities than those covered by this article can be exercised using the test method, such as avoidance of dynamic obstacles or coordination of movements throughout an environment across a fleet of A-UGVs. Both of those scenarios could be tested simultaneously, in that a fleet of A-UGVs will have to avoid collisions with one another. If the A-UGVs are sharing information via an enterprise controller, then their movements should be coordinated to limit traffic issues. The value of u may need to increase when testing with multiple moving agents in order to make room for more than one A-UGV. These capabilities and others will be considered for the next iteration of the test method, and researchers are encouraged to expand the test method to fit their evaluation needs. While only a single A-UGV (more specifically, an AMR) has been evaluated using the test method, it is expected that other A-UGVs (i.e., an AGV) will be able to be tested given the flexibility of apparatus dimensions. Also, the test method utilizes the same boundary definitions and fault conditions from ASTM F3244-17 [19], which has been used with both AMRs and AGVs. With these points in mind, R1 and R2 are assumed to have been met, which will be validated with continued testing of more A-UGVs.

Each test configuration possible within the test method is aimed at exercising various A-UGV capabilities (see Table 3), which are enabled by a core set of competencies that an A-UGV most possess in order to demonstrate that capability. The parameters that make up a test configuration each require these competencies in order for the test to be successful: the obstacle configurations require the A-UGV to detect and react in order to not collide with the obstacle and continue attempting to reach the goal, while the varied knowledge conditions can cause issues with localization and navigation due to model misalignment. However, it is difficult to compare more detailed metrics for these competencies, such as the accuracy of an A-UGV's detection of the size of an obstacle compared to its actual size, or the accuracy of localization by comparing the A-UGV's local map to the global map. Doing so would assume that all A-UGV's software representations of these characteristics used a standard format, which will most likely not be the case. The proposed test method instead relies on the use of directly observable metrics (e.g., crossing lines on the ground, physical contact with the environment) for evaluating performance according to the set success criteria, which is malleable according to the test requestor. With these factors in place, both R3 and R4 are met. Additionally, the use of directly observable ground truth measurements via tape on the ground, paper grids, and video cameras satisfies R6.

A successful test result means that the capabilities associated with the test configuration used can be assumed to perform reliably (according to the confidence and probability thresholds per the number of repetitions performed) for that A-UGV when in similar context (i.e., same physical/software configuration, environment boundaries, obstacle configurations, and knowledge conditions). The presentation of a test result should always be coupled with all of these associated parameters, as performance in one context is not necessarily transferrable to another. The parameters of the test method are delineated such that successful performance under each demonstrates sufficiently different capabilities, so caution should be exercised when extending the interpretation of a test result to a scenario with different parameters. However, the underlying fact that a capability was or was not demonstrated can be interpreted more broadly, as the possession of these capabilities already can set apart A-UGV's from one another. For example, an A-UGV that is able to adapt to new scenarios (i.e., replan their path and route when encountering obstacles) may be better suited for a deployment environment where the layout changes very frequently, such as a smaller, leaner, more dynamic manufacturing environment that adjusts based on new or rotating jobs. Conversely, an A-UGV that cannot replan its movements and instead waits for an obstacle to be removed from its environment may be better suited for a larger, more concrete manufacturing operation that operates on a very strict layout. Between these two A-UGVs, there may be a tradeoff between flexibility and speed, capabilities that the test method can be used to demonstrate and measure.

The results of performance in this test method can be presented similarly to the data found in Table 5, but more detail will be needed, including the specific characteristics of the obstacles used, their positions within the apparatus, environmental conditions, etc. Other standards developed through the ASTM F45 committee, such as ASTM F3218-17 [20] for recording environmental conditions and ASTM WK60390 (in development) for describing the characteristics of obstacles which could be used, can be leveraged to provide more of this information (see Table 1 for a breakdown of all ASTM F45 standards and those in development). Each standard generated through the ASTM F45 committee includes an example reporting form that can be used to record and present the results of a test; a similar type of report form will need to be generated for this test method. ASTM WK65141, another standard currently in development, is a practice for combining multiple standards together to form more complex testing scenarios, similar to the test method described in this article. The proposed test method is intended to serve as an example implementation of similar testing concepts to ASTM WK65141 and can be used to inform the development of that standard. The proposed test method leverages sufficient elements and structure of the referenced standards, while also new testing components such as using explicit misalignments between the A-UGV's internal model of the environment and the true physical nature of the environment (i.e., knowledge conditions), that R5 is considered to have been met.

Test	Test Config	Success Criteria		% of successful reps	Avg task time (s)	Stdev task time (s)	Notes
		Test: # of reps, direction	Task: type				
Baseline	P_eS_e	20, both	Goal reaching	100%	25.1	12.6	
1	$P_{mi,1}S_e$	20, both	Goal reaching	100%	21.5	5.9	
2	$P_{mi,2}S_e$	20, both	Goal reaching	100%	20.0	6.6	
3	$P_{mi,3}S_e$	20, both	Goal reaching	95%	23.2	13.2	Aborted navigation
4	$P_{mo,1}S_e$	20, both	Goal reaching	100%	25.6	12.8	
5	$P_{mo,2}S_e$	20, both	Goal reaching	100%	24.8	11.4	
6	$P_{mo,3}S_e$	20, both	Goal reaching	0%	n/a	n/a	Aborted navigation
7	$P_{mo,1}S_e$	20, both	Goal reaching	30%	18.5	5.9	Collided with obstacle
8	$P_{mo,2}S_e$	20, both	Goal reaching	30%	21.6	8.9	Aborted navigation
9	$P_{mo,3}S_e$	20, both	Goal reaching	100%	22.4	8.1	
10	$P_{mo,1}S_e$	20, both	Goal reaching	100%	39.3	11.5	
11	$P_{mo,2}S_e$	20, both	Goal reaching	100%	37.2	8.8	
12	$P_{mo,3}S_e$	20, both	Goal reaching	100%	42.8	13.4	
13	$P_{se,1}S_e$	20, both	Goal reaching	100%	36.9	7.6	
14	$P_{se,2}S_e$	20, both	Goal reaching	100%	36.8	9.3	
15	$P_{se,3}S_e$	20, both	Goal reaching	100%	39.4	12.8	
16	$P_{se,1}S_e$	20, both	Goal reaching	100%	34.9	6.0	
17	$P_{se,2}S_e$	20, both	Goal reaching	100%	39.1	10.5	
18	$P_{se,3}S_e$	20, both	Goal reaching	100%	33.8	4.9	
19	$P_{mo,1}S_{mo,1}$	5, A-B	Stuck state	40%	n/a	n/a	Collided with boundary
20	$P_{mo,2}S_{mo,2}$	5, A-B	Stuck state	100%	n/a	n/a	
21	$P_{mo,3}S_{mo,3}$	5, A-B	Stuck state	100%	n/a	n/a	
22	$P_{se,1}S_{se,1}$	5, A-B	Stuck state	100%	n/a	n/a	
23	$P_{se,2}S_{se,2}$	5, A-B	Stuck state	100%	n/a	n/a	
24	$P_{se,3}S_{se,3}$	5, A-B	Stuck state	100%	n/a	n/a	
25	$P_{se,1}S_{se,1}$	5, A-B	Stuck state	60%	n/a	n/a	Collided with boundary
26	$P_{se,2}S_{se,2}$	5, A-B	Stuck state	100%	n/a	n/a	
27	$P_{se,3}S_{se,3}$	5, A-B	Stuck state	100%	n/a	n/a	

Table 5. Summary of example performance data from validation testing. Failed tests are shaded in gray. Note that tests with identical test configuration identifiers were unique to each another due to the different obstacle positions being used, even if they are considered to be in the same obstacle configuration category.

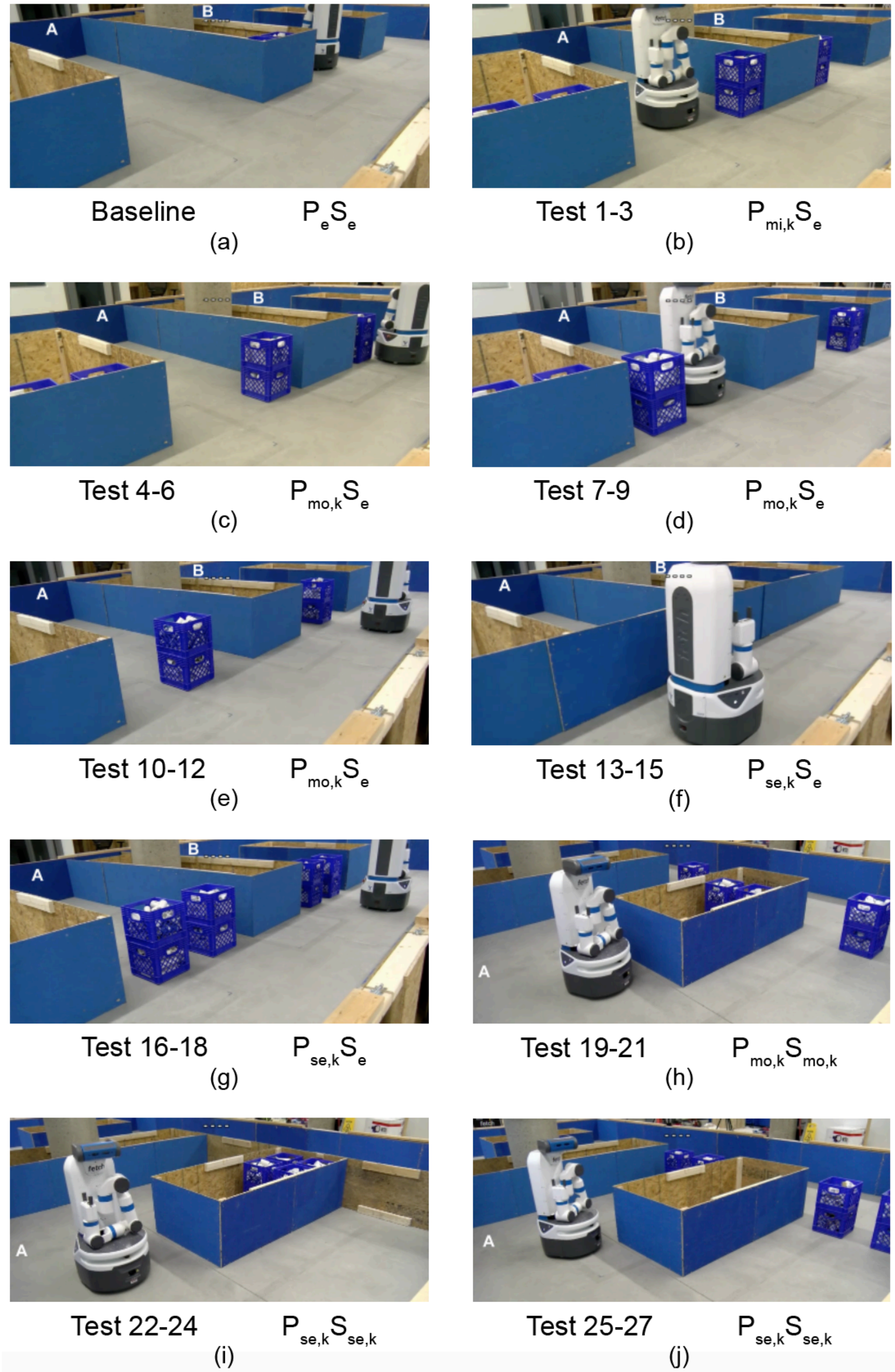

Figure 5. Screenshots from videos of test performance showing each unique obstacle configuration. A knowledge condition value of K is shown for test configurations when 1, 2, or 3 could be used.

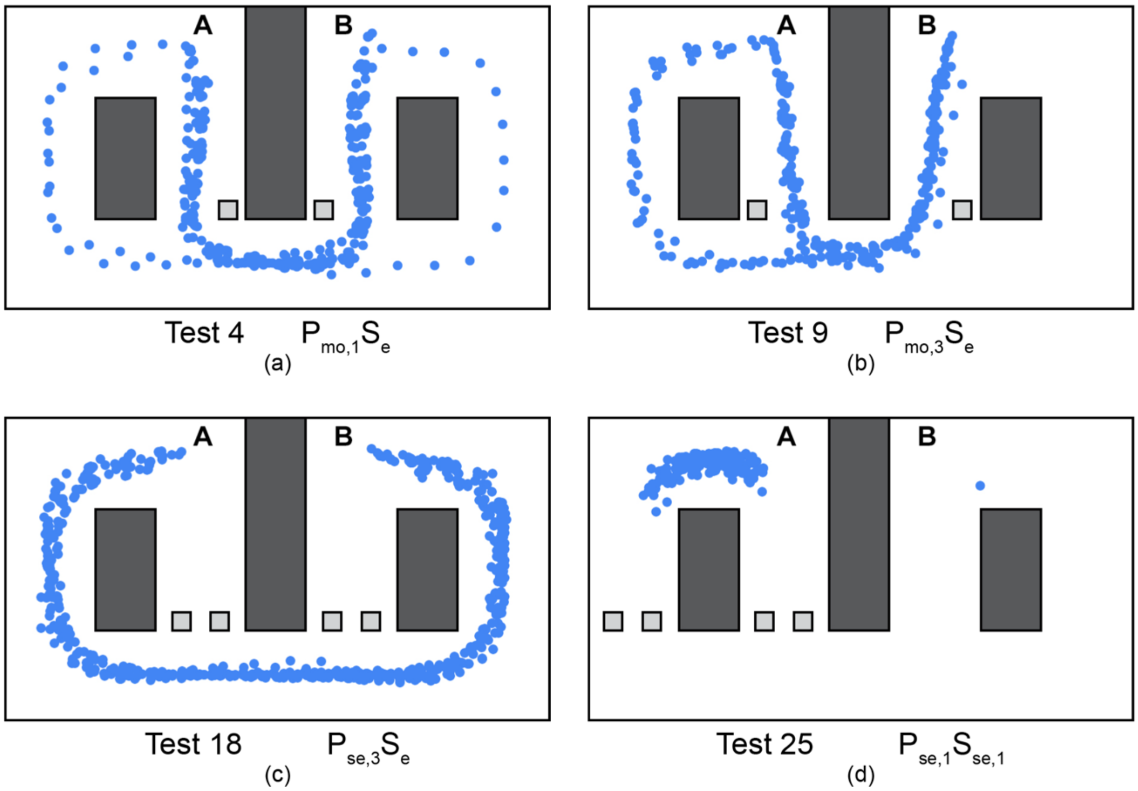

Figure 6. **Plots of the A-UGV's position according to its localization data during some of the example performance data test configurations. Note: test 25 (d) only used obstacles placed on the A side of the apparatus because it was physically impossible or the A-UGV to reach the B side so was therefore unnecessary.**

Conclusions

This article presents a test method for evaluating the navigation and obstacle avoidance capabilities of A-UGVs; the design of the apparatus, obstacle configurations, knowledge conditions, test configurations, procedure, and performance measures are specified. The test method is designed using similar constructs of existing or in-development standards from the ASTM F45 Committee on Driverless Automatic Guided Industrial Vehicles [1] such that if it were to become a standard its development would logically coalesce with the other standards. Example performance data was collected using a commercially available A-UGV in order to validate the test method using a variety of test configurations. Further testing will be conducted using other A-

UGVs in order to produce a spread of performance data. This work will be used to form the basis of a proposed standard practice or usage guide for the ASTM F45 committee to develop.

ACKNOWLEDGMENTS

This research has been supported in part by the National Institute of Standards and Technology (NIST) under award 70NANB17H256. The authors would like to thank the ASTM F45 committee for their input and support in developing this test method.

References

[1] ASTM International, "Committee F45 on Driverless Automatic Guided Industrial Vehicles," 2019. [Online]. Available: https://www.astm.org/COMMITTEE/F45.htm. [Accessed August 2019].

[2] Industrial Truck Standards Development Foundation, *ANSI/ITSDF B56.5-2012, Safety Standard for Driverless, Automatic Guided Industrial Vehicles and Automated Functions of Manned Industrial Vehicles,* 2012.

[3] International Organization for Standardization, *ISO 13482:2014 Robots and robotic devices–Safety requirements–Non-medical personal care robot,* Geneva, Switzerland: International Organization for Standardization, 2014.

[4] M. Schneier and R. Bostelman, "Literature Review of Mobile Robots for Manufacturing," U.S. Department of Commerce, National Institute of Standards and Technology, 2015.

[5] ASTM International, *F3200-18a Standard Terminology for Driverless Automatic Guided Industrial Vehicles,* West Conshohocken, PA: ASTM International, 2018.

[6] Omron Corporation, "LD Series Mobile Robots/Lineup | OMRON Industrial Automation," [Online]. Available: http://www.ia.omron.com/products/family/3664/lineup.html. [Accessed 17 October 2019].

[7] America In Motion, "America In Motion - Our Products," [Online]. Available: https://www.weareaim.com/Products.aspx. [Accessed 17 October 2019].

[8] Kinova Inc., "MOVO Mobile Manipulator," [Online]. Available: https://www.kinovarobotics.com/en/knowledge-hub/movo-mobile-manipulator. [Accessed 17 October 2019].

[9] Seegrid, "Platform - Seegrid," [Online]. Available: https://seegrid.com/platform/. [Accessed 17 October 2019].

[10] Fetch Robotics, Inc., "Fetch Mobile Manipulator," [Online]. Available: https://fetchrobotics.com/robotics-platforms/fetch-mobile-manipulator/. [Accessed 17 October 2019].

[11] S. G. Tzafestas, "Mobile robot control and navigation: A global overview," *Journal of Intelligent & Robotic Systems,* vol. 91, no. 1, pp. 35-58, 2018.

[12] M. Hoy, A. S. Matveev and A. V. Savkin, "Algorithms for collision-free navigation of mobile robots in complex cluttered environments: a survey," *Robotica,* vol. 33, no. 3, pp. 463-497, 2015.

[13] A. Pandey, D. R. Parhi and S. Pandey, "Mobile robot navigation and obstacle avoidance techniques," *International Robotics and Automation Journal,* vol. 2, no. 3, 2017.

[14] N. D. M. Ceballos, J. A. Valencia and N. L. Ospina, "Quantitative Performance Metrics for Mobile Robots Navigation," *Mobile Robots Navigation,* 2010.

[15] R. Bostelman, T. Hong and G. Cheok, "Navigation performance evaluation for automatic guided vehicles," in *IEEE International Conference on Technologies for Practical Robot Applications (TePRA)*, 2015.

[16] ASTM International, *F3265-17 Standard Test Method for Grid-Video Obstacle Measurement,* West Conshohocken, PA: ASTM International, 2017.

[17] A. Norton and H. Yanco, "Preliminary Development of a Test Method for Obstacle Detection and Avoidance in Industrial Environments," *ASTM STP1594 Autonomous Industrial Vehicles: From the Laboratory to the Factory Floor,* pp. 23-40, 2016.

[18] S. Yoon and R. Bostelman, "Analysis of Automatic through Autonomous – Ummanned Ground Vehicles (A-UGVs) Towards Performance Standards," in *Proceedings of IEEE International Symposium on RObotic and SEnsors Environments (ROSE)*, Ottawa, Ontario, 2019.

[19] ASTM International, *F3244-17 Standard Test Method for Navigation: Defined Area,* West Conshohocken, PA: ASTM International, 2017.

[20] ASTM International, *F3218-17 Standard Practice for Recording Environmental Effects for Utilization with A-UGV Test Methods,* West Conshohocken, PA: ASTM International, 2017.

[21] ASTM International, *F3327-18 Standard Practice for Recording the A-UGV Test Configuration,* West Conshohocken, PA: ASTM International, 2018.

[22] R. Madhavan, R. Bostelman, Z. Kootbally, R. Lakaemper, S. K. Gupta and S. Balakirsky, "Smart, Flexible, and Safety Industrial Mobile Robots: Performance Evaluation & Standardization Efforts," in *ITEA Annual Technology Review Conference*, 2011.

[23] H.-M. Huang, K. Pavek, B. Novak, J. Albus and E. Messina, "A Framework for Autonomy Levels for Unmanned Systems (ALFUS)," in *Proceedings of the AUVSI's Unmanned Systems North America Conference*, 2005.

[24] H.-M. Huang, E. Messina, A. Jacoff, R. Wade and M. McNair, "Performance Measures Framework for Unmanned Systems (PerMFUS): Models for Contextual Metrics," in *Proceedings of the 10th Performance Metrics for Intelligent Systems Workshop*, 2010.

[25] R. Bostelman and E. Messina, "A-UGV Capabilities," in *Proceedings of the Third IEEE International Conference on Robotic Computing (IRC)*, Naples, Italy, 2019.

[26] J. S. Gill, M. Tomaszewski, Y. Jia, P. Pisu and V. N. Krovi, "Evaluation of Navigation in Mobile Robots for Long-Term Autonomy in Automotive Manufacturing Environments," *SAE Technical Paper,* 2019.

[27] Open Source Robotics Foundation, "ROS.org | Powering the world's robots," [Online]. Available: https://www.ros.org/. [Accessed 17 October 2019].

[28] Open Source Robotics Foundation, "amcl - ROS Wiki," [Online]. Available: http://wiki.ros.org/amcl. [Accessed 17 October 2019].

[29] Open Source Robotics Foundation, "move_base | ROS Wiki," [Online]. Available: http://wiki.ros.org/move_base. [Accessed 17 October 2019].

[30] H. Durrant-Whyte and T. Bailey, "Simultaneous Localization and Mapping: Part I," *IEEE Robotics & Automation Magazine,* pp. 99-110, June 2006.