CLAIMS

What is claimed is:

1. A method comprising:

receiving, at a perception module configured in operational logic of a robotic control system, image data from a robot's sensors,

wherein the perception module comprises a detection module, a scene segmentation module, and a mapping module,

wherein the robot's sensors include at least one of cameras, lidar sensors, inertial measurement unit (IMU) sensors, wheel encoders, and other sensors, and

wherein the operational logic of the robotic control system comprises executable instructions stored by a storage device of the robotic control system and executed by a processor of the robotic control system;

detecting, by the detection module, objects from the image data, as two-dimensional (2D) bounding boxes with object classes;

generating, by the detection module:

predicted three-dimensional (3D) object locations, using the 2D bounding boxes and a ground plane; and

two-dimensional - three-dimensional (2D-3D) bounding boxes with class labels, the 2D-3D bounding boxes based on the 2D bounding boxes and the predicted 3D object locations;

generating, by the scene segmentation module:

a multi-class segmentation map, using a segmentation model to segregate a floor boundary and other relevant regions in the image data;

an edge map including the floor boundary and other relevant boundaries, with semantic information; and

semantic boundary masks, from the multi-class segmentation map, wherein the semantic boundary masks identify relevant boundaries and their semantic information;

generating, by the mapping module, using the 2D-3D bounding boxes and the semantic boundary masks:

a scene layout map, wherein the scene layout map includes global elements relevant for global mapping; and

a local occupancy map, wherein the local occupancy map includes local elements useful for local path planning and local obstacle avoidance;

receiving, at a navigation module configured in the operational logic of the robotic control system, the scene layout map, and the local occupancy map, wherein the navigation module includes:

a simultaneous localization and mapping module (SLAM);

a global mapper module; and

a fusion and inflation module;

generating, by the SLAM, using lidar data, IMU data, and wheel encoding data:

a SLAM global map, which comprises a 2D occupancy grid representation of an environment with obstacle information at lidar height and real-time location information of the robot; and

a robot pose;

generating, by the global mapper module, using the SLAM global map, the lidar data, and the scene layout map:

a navigation global map, which represents an improved 2D occupancy grid representation of the environment over the SLAM global map;

generating, by the fusion and inflation module using the lidar data, the navigation global map, and the local occupancy map:

a fused local occupancy map, with the lidar data and information from the navigation global map and the local occupancy map, making the fused local occupancy map useful for obstacle avoidance;

an inflated global map, which includes buffer regions for the robot around obstacles; and

an inflated local map based on the fused local occupancy map, which includes the buffer regions for the robot around the obstacles;

receiving, at a robot brain configured in the operational logic of the robotic control system, the inflated global map, the robot pose, and the inflated local map;

generating, by the robot brain, robot action commands based on at least one of the inflated global map, the robot pose, and the inflated local map;

receiving, at an action module configured in the operational logic of the robotic control system, the robot action commands; and

controlling robot actuators in response to the robot action commands.

2. The method of claim 1, further comprising:

receiving, at the robot brain, an interface signal from a robot user interface; and

generating, by the robot brain, the robot action commands based on at least one of the inflated global map, the robot pose, the inflated local map, and the interface signal.

3. The method of claim 1, further comprising:

detection filtering, by the detection module, to remove the 2D-3D bounding boxes that are not on the ground or are inside of a shovel on the robot.

4. The method of claim 1, further comprising:

running an edge detection algorithm, by the scene segmentation module, on the multi-class segmentation map, resulting in the semantic boundary masks with the semantic information.

5. The method of claim 1, further comprising:

adding, using a labeling module included in the mapping module, additional semantic labels to the segmentation boundary masks by:

determining the points in the edge map that are inside a corresponding 2D bounding box or a corresponding 2D-3D bounding box; and

assigning all of the points inside of the corresponding 2D bounding box or the corresponding 2D-3D bounding box a same label as that of the corresponding 2D bounding box or the corresponding 2D-3D bounding box, thereby resulting in a semantically rich boundary map;

warping, using a top view transformation module included in the mapping module, the semantically rich boundary map into a point cloud with real-world coordinates and semantic label classes by using at least one of:

a lookup table that stores pixel mappings from an image space to real-world coordinates; and

a homography matrix that maps pixels from the image space to the real-world coordinates in real time;

filtering, by a scene layout module included in the mapping module, the point cloud with the real-world coordinates and the semantic label classes, keeping the semantically relevant points needed for the global mapping and discarding points not needed for the global mapping; and

filtering, by a local occupancy module included in the mapping module, the point cloud with real-world coordinates and semantic label classes, keeping the semantically relevant points needed for the local path planning and the obstacle avoidance.

6. The method of claim 1, further comprising:

processing, using a filter and fusion module included in the SLAM, the lidar data, the IMU data, and the wheel encoding data by:

removing noise and potentially unreliable data from the lidar data and the IMU data resulting in filtered lidar data and filtered IMU data;

removing the noise and potentially unreliable data from an angular velocity output of the wheel encoding data; and

fusing non-angular velocity output data of the wheel encoding data with the filtered IMU data to generate filtered and fused odometry data.

7. The method of claim 6, further comprising:

generating, using a main pipeline included in the SLAM, the 2D occupancy grid representation of the environment with the obstacle information at lidar height and the real-time location information of the robot by:

receiving the filtered lidar data and the filtered and fused odometry data;

creating a new 2D point registration for each new laser measurement at a given odometry reading;

estimating and correcting odometry slippages between each odometry reading by scan-to-scan matching the new 2D point registrations, thereby resulting in odometry slippage data points;

adding the odometry slippage data points to a pose-graph, resulting in an optimized pose-graph;

looking for loop closure in a chain of ‘N’ of the odometry slippage data points, wherein the loop closure represents a process of determining if a current location observed by the robot's sensors has been previously visited by the robot;

on condition the loop closure is detected:

correcting odometry poses for each new 2D point registration based on the optimized pose-graph, thereby resulting in loop closure pose corrections; and

forming a common 2D occupancy grid using the 2D point registrations and probabilistically updated 2D point registrations, wherein the common 2D occupancy grid is the 2D occupancy grid representation of the environment with the obstacle information at lidar height and the real-time location information of the robot.

8. The method of claim 7, further comprising:

processing, using a sensor data filter, the lidar data, and the scene layout map by:

removing the noise and potentially unreliable data from the lidar data; and

passing the scene layout map through a semantic filter that caters to the filtering of points from the scene layout map based on the semantic information provided, resulting in filtered lidar data and a filtered scene layout map where unreliable and irrelevant semantic labels have been removed during mapping.

9. The method of claim 8, further comprising:

generating, using a multi-sensor data registration, an enhanced 2D occupancy grid representation of the environment with the obstacle information at lidar height, from the filtered lidar data and the filtered scene layout map, by:

receiving the filtered lidar data and the filtered scene layout map;

creating a new enhanced 2D point registration for each new laser measurement and the filtered scene layout map using the real-time location information of the robot from the SLAM; and

updating all registrations from all of the robot's sensors, probabilistically, together in an enhanced common 2D occupancy grid based on predetermined confidence values of the robot's sensors.

10. The method of claim 9, further comprising:

receiving, by loop closure integration, the new enhanced 2D point registrations and the SLAM global map including the loop closure pose corrections;

reiterating, temporally, over the new 2D point registrations that are near the loop closure pose corrections for all of the updated registrations of the robot's sensors from the multi-sensor data registration to provide reiterated loop closure pose corrections;

updating map pose data of each of the robot's sensors using the reiterated loop closure pose corrections;

re-updating the updated registrations of each of the robot's sensors with the updated map pose data, thereby resulting in enhanced registrations; and

re-updating, probabilistically with the enhanced registrations, respective cells in the enhanced common 2D occupancy grid, thereby resulting in the navigation global map.

11. An apparatus comprising:

a robot;

a processor; and

a memory storing instructions that, when executed by the processor, configure the apparatus to:

receive, at a perception module configured in operational logic of a robotic control system, image data from the robot's sensors,

wherein the perception module comprises a detection module, a scene segmentation module, and a mapping module,

wherein the robot's sensors include at least one of cameras, lidar sensor, inertial measurement unit (IMU) sensors, wheel encoder, and other sensors, and

wherein the operational logic of the robotic control system comprises executable instructions stored by a storage device of the robotic control system and executed by a processor of the robotic control system;

detect, by the detection module, objects from the image data, as two-dimensional (2D) bounding boxes with object classes;

generate, by the detection module:

predicted three-dimensional (3D) object locations, using the 2D bounding boxes and a ground plane; and

two-dimensional - three-dimensional (2D-3D) bounding boxes with class labels, the 2D-3D bounding boxes based on the 2D bounding boxes and the predicted 3D object locations;

generate, by the scene segmentation module:

a multi-class segmentation map, using a segmentation model to segregate a floor boundary and other relevant regions in the image data;

an edge map including the floor boundary and other relevant boundaries, with semantic information; and

semantic boundary masks, from the multi-class segmentation map, wherein the semantic boundary masks identify relevant boundaries and their semantic information;

generate, by the mapping module, using the 2D-3D bounding boxes and the semantic boundary masks:

a scene layout map, wherein the scene layout map includes global elements relevant for global mapping; and

a local occupancy map, wherein the local occupancy map includes local elements useful for local path planning and local obstacle avoidance;

receive, at a navigation module configured in the operational logic of the robotic control system, the scene layout map, and the local occupancy map, wherein the navigation module includes:

a simultaneous localization and mapping module (SLAM);

a global mapper module; and

a fusion and inflation module;

generate, by the SLAM, using lidar data, (IMU) data, and wheel encoding data:

a SLAM global map, which comprises a 2D occupancy grid representation of an environment with obstacle information at lidar height and real-time location information of the robot; and

a robot pose;

generate, by the global mapper module, using the SLAM global map, the lidar data, and the scene layout map:

a navigation global map, which represents an improved 2D occupancy grid representation of the environment over the SLAM global map;

generate, by the fusion and inflation module using the lidar data, the navigation global map, and the local occupancy map:

a fused local occupancy map, with the lidar data and information from the navigation global map and the local occupancy map, making the fused local occupancy map useful for obstacle avoidance;

an inflated global map, which includes buffer regions for the robot around obstacles; and

an inflated local map based on the fused local occupancy map, which includes the buffer regions for the robot around the obstacles;

receive, at a robot brain, the inflated global map, the robot pose, and the inflated local map;

generate, by the robot brain configured in the operational logic of the robotic control system, robot action commands based on at least one of the inflated global map, the robot pose, and the inflated local map;

receive, at an action module configured in the operational logic of the robotic control system, the robot action commands; and

control robot actuators in response to the robot action commands.

12. The apparatus of claim 11, wherein the instructions further configure the apparatus to:

receive, at the robot brain, an interface signal from a robot user interface; and

generate, by the robot brain, the robot action commands based on at least one of the inflated global map, the robot pose, the inflated local map, and the interface signal.

13. The apparatus of claim 11, wherein the instructions further configure the apparatus to:

detection filter, by the detection module, to remove the 2D-3D bounding boxes that are not on the ground or are inside of a shovel on the robot.

14. The apparatus of claim 11, wherein the instructions further configure the apparatus to:

run an edge detection algorithm, by the scene segmentation module, on the multi-class segmentation map, resulting in the semantic boundary masks with the semantic information.

15. The apparatus of claim 11, wherein the instructions further configure the apparatus to:

add, using a labeling module included in the mapping module, additional semantic labels to the segmentation boundary masks by:

determine the points in the edge map that are inside a corresponding 2D bounding box or a corresponding 2D-3D bounding box; and

assign all of the points inside of the corresponding 2D bounding box or the corresponding 2D-3D bounding box a same label as that of the corresponding 2D bounding box or the corresponding 2D-3D bounding box, thereby resulting in a semantically rich boundary map;

warp, using a top view transformation module included in the mapping module, the semantically rich boundary map into a point cloud with real-world coordinates and semantic label classes by using at least one of:

a lookup table that stores pixel mappings from an image space to real-world coordinates; and

a homography matrix that maps pixels from the image space to the real-world coordinates in real time;

filter, by a scene layout module included in the mapping module, the point cloud with the real-world coordinates and the semantic label classes, keeping the semantically relevant points needed for the global mapping and discarding points not needed for the global mapping; and

filter, by a local occupancy module included in the mapping module, the point cloud with real-world coordinates and semantic label classes, keeping the semantically relevant points needed for the local path planning and the obstacle avoidance.

16. The apparatus of claim 11, wherein the instructions further configure the apparatus to:

process, using a filter and fusion module included in the SLAM, the lidar data, the IMU data, and the wheel encoding data by:

remove noise and potentially unreliable data from the lidar data and the IMU data resulting in filtered lidar data and filtered IMU data;

remove the noise and potentially unreliable data from an angular velocity output of the wheel encoding data; and

fuse non-angular velocity output data of the wheel encoding data with the filtered IMU data to generate filtered and fused odometry data.

17. The apparatus of claim 16, wherein the instructions further configure the apparatus to:

generate, using a main pipeline included in the SLAM, the 2D occupancy grid representation of the environment with the obstacle information at lidar height and the real-time location information of the robot by:

receive the filtered lidar data and the filtered and fused odometry data;

create a new 2D point registration for each new laser measurement at a given odometry reading;

estimate and correct odometry slippages between each odometry reading by scan-to-scan matching the new 2D point registrations, thereby resulting in odometry slippage data points;

add the odometry slippage data points to a pose-graph, resulting in an optimized pose-graph;

look for loop closure in a chain of 'N' of the odometry slippage data points, wherein the loop closure represents a process of determining if a current location observed by the robot's sensors has been previously visited by the robot;

on condition the loop closure is detected:

correct the odometry poses for each new 2D point registration based on the optimized pose-graph, thereby resulting in loop closure pose corrections;

form a common 2D occupancy grid using the 2D point registrations and probabilistically updated 2D point registrations, wherein the common 2D occupancy grid is the 2D occupancy grid representation of the environment with the obstacle information at lidar height and the real-time location information of the robot.

18. The apparatus of claim 17, wherein the instructions further configure the apparatus to:

process, using a sensor data filter, the lidar data, and the scene layout map by:

remove the noise and potentially unreliable data from the lidar data; and

pass the scene layout map through a semantic filter that caters to the filtering of points from the scene layout map based on the semantic information provided, resulting in filtered lidar data and a filtered scene layout map where unreliable and irrelevant semantic labels have been removed during mapping.

19. The apparatus of claim 18, wherein the instructions further configure the apparatus to:

generate, using a multi-sensor data registration, an enhanced 2D occupancy grid representation of the environment with the obstacle information at lidar height, from the filtered lidar data and the filtered scene layout map by:

receive filtered lidar data and the filtered scene layout map;

create a new enhanced 2D point registration for each new laser measurement and the filtered scene layout map using the real-time location information of the robot from the SLAM; and

update all registrations from all of the robot's sensors, probabilistically, together in an enhanced common 2D occupancy grid based on predetermined confidence values of the robot's sensors.

20. The apparatus of claim 19, wherein the instructions further configure the apparatus to:

receive, by loop closure integration, the new enhanced 2D point registrations and the SLAM global map including the loop closure pose corrections;

reiterate, temporally, over the new 2D point registrations that are near the loop closure pose corrections for all of the updated registrations of the robot's sensors from the multi-sensor data registration to provide reiterated loop closure pose corrections;

update map pose data of each of the robot's sensors using the reiterated loop closure pose corrections;

re-update the updated registrations of each of the robot's sensors with the updated map pose data, thereby resulting in enhanced registrations; and

re-update, probabilistically with the enhanced registrations, respective cells in the enhanced common 2D occupancy grid, thereby resulting in the navigation global map.