

(51) International Patent Classification:  
*G05D 1/02* (2006.01)

(21) International Application Number:

PCT/US2016/059171

(22) International Filing Date:

27 October 2016 (27.10.2016)

(25) Filing Language:

English

(26) Publication Language:

English

(30) Priority Data:

62/267,886 15 December 2015 (15.12.2015) US  
15/249,250 26 August 2016 (26.08.2016) US(71) Applicant: **QUALCOMM INCORPORATED** [US/US];  
Attn: International IP Administration, 5775 Morehouse  
Drive, San Diego, California 92121-1714 (US).(72) Inventors: **WIERZYNSKI, Casimir Matthew**; 5775  
Morehouse Drive, San Diego, California 92121-1714 (US).  
**BEHABADI, Bardia Fallah**; 5775 Morehouse Drive, San  
Diego, California 92121-1714 (US). **GIBSON, Sarah  
Paige**; 5775 Morehouse Drive, San Diego, California  
92121-1714 (US). **AGHAMOHAMMADI, Aliakbar**;  
5775 Morehouse Drive, San Diego, California 92121-1714  
(US). **AGARWAL, Saurav**; 1907 Dartmouth Street,  
Apartment 1404, San Diego, Texas 77840 (US).(74) Agents: **LENKIN, Alan M.** et al.; Seyfarth Shaw LLP,  
Suite 3500, 2029 Century Park East, Los Angeles, Califor-  
nia 90067-3021 (US).(81) Designated States (unless otherwise indicated, for every  
kind of national protection available): AE, AG, AL, AM,  
AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY,  
BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DJ, DK, DM,  
DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT,  
HN, HR, HU, ID, IL, IN, IR, IS, JP, KE, KG, KN, KP, KR,  
KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME,  
MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ,  
OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA,  
SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM,  
TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM,  
ZW.(84) Designated States (unless otherwise indicated, for every  
kind of regional protection available): ARIPO (BW, GH,  
GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ,  
TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU,  
TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE,  
DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU,  
LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK,  
SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ,  
GW, KM, ML, MR, NE, SN, TD, TG).

Published:

— with international search report (Art. 21(3))

(54) Title: AUTONOMOUS VISUAL NAVIGATION

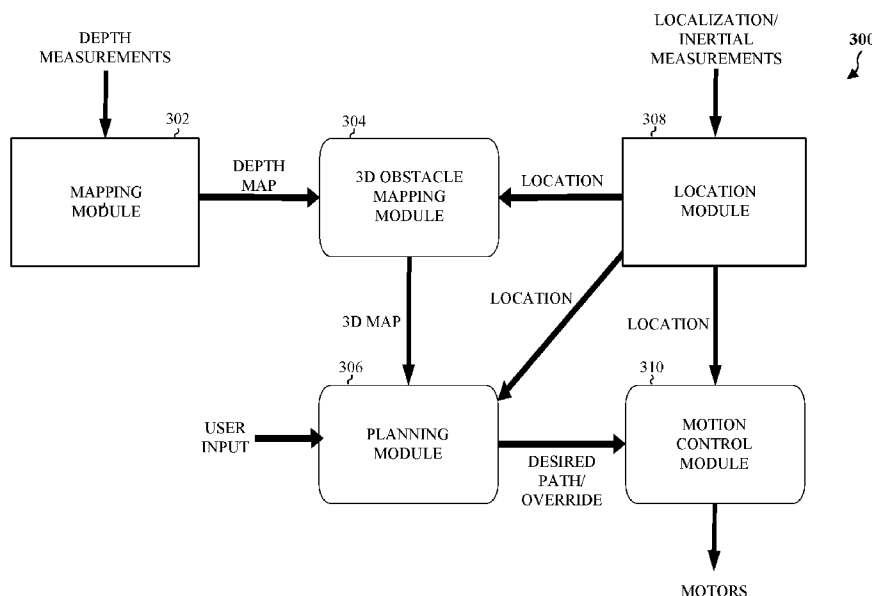


FIG. 3

(57) Abstract: A method of visual navigation for a robot includes integrating a depth map with localization information to generate a three-dimensional (3D) map. The method also includes motion planning based on the 3D map, the localization information, and/or a user input. The motion planning overrides the user input when a trajectory and/or a velocity, received via the user input, is predicted to cause a collision.

## **AUTONOMOUS VISUAL NAVIGATION**

### **CROSS REFERENCE TO RELATED APPLICATION**

**[0001]** This application claims the benefit under 35 U.S.C. § 119(e) to United States Provisional Patent Application No. 62/267,886, entitled “AUTONOMOUS VISUAL NAVIGATION,” filed on December 15, 2015, the disclosure of which is expressly incorporated herein by reference in its entirety.

### **BACKGROUND**

#### **Field**

**[0002]** Certain aspects of the present disclosure generally relate to robotics and, more particularly, to onboard vision-based localization, mapping, and planning.

#### **Background**

**[0003]** Robots may be designed to perform behaviors or tasks with a high degree of autonomy. A robot may use different modules and components for performing various tasks. For example, the robot may have different components for localization, mapping and planning. Localization is directed to solving the problem of determining where the robot is located. The robot receives input from its sensors to understand where the robot is located within its environment.

**[0004]** Mapping is directed to building a representation of the environment. For example, mapping is used to determine which portion of the environment is occupied and which parts are free space. Furthermore, mapping may prevent the robot from colliding with obstacles.

**[0005]** Planning is directed to determining how to perform a task after the robot knows the layout of the environment and how it will travel from point A to B. That is, in some cases, prior to moving from a current position to a target, it is desirable to determine the trajectory (e.g., path) to the target with the lowest cost from multiple candidate trajectories evaluated during a planning phase.

## SUMMARY

**[0006]** In one aspect of the present disclosure, a method of visual navigation for a robot is disclosed. The method includes integrating a depth map with localization information to generate a three-dimensional (3D) map. The method also includes motion planning based on the 3D map, the localization information, or a user input. The motion planning overrides the user input when the user input is predicted to cause a collision.

**[0007]** Another aspect of the present disclosure is directed to an apparatus including means for integrating a depth map with localization information to generate a three-dimensional map. The apparatus also includes means for motion planning based on the 3D map, the localization information, or a user input. The motion planning overrides the user input when the user input is predicted to cause a collision.

**[0008]** In another aspect of the present disclosure, a non-transitory computer-readable medium with non-transitory program code recorded thereon is disclosed. The program code for visual navigation for a robot is executed by a processor and includes program code to integrate a depth map with localization information to generate a three-dimensional map. The program code also includes program code to motion plan based on the 3D map, the localization information, or a user input. The motion planning overrides the user input when the user input is predicted to cause a collision.

**[0009]** Another aspect of the present disclosure is directed to an apparatus having a memory unit and one or more processors coupled to the memory unit. The processor(s) is configured to integrate a depth map with localization information to generate a three-dimensional map. The processor(s) is also configured to motion plan based on the 3D map, the localization information, or a user input. The motion planning overrides the user input when the user input is predicted to cause a collision.

**[0010]** Additional features and advantages of the disclosure will be described below. It should be appreciated by those skilled in the art that this disclosure may be readily utilized as a basis for modifying or designing other structures for carrying out the same purposes of the present disclosure. It should also be realized by those skilled in the art that such equivalent constructions do not depart from the teachings of the disclosure as set forth in the appended claims. The novel features, which are believed to be

characteristic of the disclosure, both as to its organization and method of operation, together with further objects and advantages, will be better understood from the following description when considered in connection with the accompanying figures. It is to be expressly understood, however, that each of the figures is provided for the purpose of illustration and description only and is not intended as a definition of the limits of the present disclosure.

### **BRIEF DESCRIPTION OF THE DRAWINGS**

**[0011]** The features, nature, and advantages of the present disclosure will become more apparent from the detailed description set forth below when taken in conjunction with the drawings in which like reference characters identify correspondingly throughout.

**[0012]** FIGURE 1 illustrates an example implementation of autonomous visual navigation using a system-on-a-chip (SOC), including a general-purpose processor in accordance with certain aspects of the present disclosure.

**[0013]** FIGURE 2 illustrates an example implementation of a system in accordance with aspects of the present disclosure.

**[0014]** FIGURE 3 is a block diagram illustrating different modules/means/components in an exemplary apparatus.

**[0015]** FIGURE 4 is a block diagram illustrating a location module according to aspects of the present disclosure.

**[0016]** FIGURE 5 illustrates an example of stereo-vision cameras according to aspects of the present disclosure.

**[0017]** FIGURES 6A, 6B, 6C, 6D, 6E, and 7 illustrate examples of a robot in an environment according to aspects of the present disclosure.

**[0018]** FIGURE 8 is a block diagram illustrating different modules/means/components in an exemplary apparatus.

**[0019]** FIGURE 9 illustrates a flow diagram for a method for autonomous visual navigation by a robot according to aspects of the present disclosure.

**[0020]** FIGURE 10 illustrates a flow diagram for a method for motion planning according to aspects of the present disclosure.

### **DETAILED DESCRIPTION**

**[0021]** The detailed description set forth below, in connection with the appended drawings, is intended as a description of various configurations and is not intended to represent the only configurations in which the concepts described herein may be practiced. The detailed description includes specific details for the purpose of providing a thorough understanding of the various concepts. However, it will be apparent to those skilled in the art that these concepts may be practiced without these specific details. In some instances, well-known structures and components are shown in block diagram form in order to avoid obscuring such concepts.

**[0022]** Based on the teachings, one skilled in the art should appreciate that the scope of the disclosure is intended to cover any aspect of the disclosure, whether implemented independently of or combined with any other aspect of the disclosure. For example, an apparatus may be implemented or a method may be practiced using any number of the aspects set forth. In addition, the scope of the disclosure is intended to cover such an apparatus or method practiced using other structure, functionality, or structure and functionality in addition to or other than the various aspects of the disclosure set forth. It should be understood that any aspect of the disclosure disclosed may be embodied by one or more elements of a claim.

**[0023]** The word “exemplary” is used herein to mean “serving as an example, instance, or illustration.” Any aspect described herein as “exemplary” is not necessarily to be construed as preferred or advantageous over other aspects.

**[0024]** Although particular aspects are described herein, many variations and permutations of these aspects fall within the scope of the disclosure. Although some benefits and advantages of the preferred aspects are mentioned, the scope of the disclosure is not intended to be limited to particular benefits, uses or objectives. Rather, aspects of the disclosure are intended to be broadly applicable to different technologies, system configurations, networks and protocols, some of which are illustrated by way of example in the figures and in the following description of the preferred aspects. The detailed description and drawings are merely illustrative of the disclosure rather than

limiting, the scope of the disclosure being defined by the appended claims and equivalents thereof.

**[0025]** Aspects of the present disclosure are directed to an integrated solution that combines localization mapping and motion planning. The integrated solution achieves motion planning based on a depth map and localization information. A localization module, such as a visual-inertial odometer, may be used for localization. Obstacle mapping may be achieved by volumetric integration of successive, multi-view depth measurements. Planning may be achieved by sampling-based geometric planning combined with fast re-planning.

**[0026]** For systems, such as robots, that may be either manually or autonomously controlled, it is desirable to plan a motion of the robot based on an integrated depth map. Furthermore, the motion planning may include overriding user input.

**[0027]** Additionally, it is desirable to safely pilot a robot both manually (e.g., stick control) and autonomously (e.g., waypoint control). Operators further desire to apply information gathered during manual operation to autonomous operation, and vice versa. Furthermore, planning may be specified to prevent a robot from colliding into obstacles, both in manual and autonomous modes. Planning may also be specified to plan the lowest cost route when the robot is operating in either a manual or an autonomous mode.

**[0028]** Furthermore, a robot may localize itself, such that the robot does not rely on external systems or signals (e.g., global positioning systems and beacons) for localization. In one configuration, the robot includes sensors that improve light and energy use.

**[0029]** Aspects of the present disclosure are directed to a robot with onboard vision-based localization, mapping, and planning such that the robot can motion plan to target locations. In one configuration, the motion planning overrides a user input.

**[0030]** FIGURE 1 illustrates an example implementation of the aforementioned method of motion planning by a robot using a system-on-a-chip (SOC) 100, which may include a general-purpose processor (CPU) or multi-core general-purpose processors (CPUs) 102 in accordance with certain aspects of the present disclosure. Variables

(e.g., neural signals and synaptic weights), system parameters associated with a computational device (e.g., neural network with weights), delays, frequency bin information, and task information may be stored in a memory block associated with a neural processing unit (NPU) 108, in a memory block associated with a CPU 102, in a memory block associated with a graphics processing unit (GPU) 104, in a memory block associated with a digital signal processor (DSP) 106, in a dedicated memory block 118, or may be distributed across multiple blocks. Instructions executed at the general-purpose processor 102 may be loaded from a program memory associated with the CPU 102 or may be loaded from a dedicated memory block 118.

**[0031]** The SOC 100 may also include additional processing blocks tailored to specific functions, such as a GPU 104, a DSP 106, a connectivity block 110, which may include fourth generation long term evolution (4G LTE) connectivity, unlicensed Wi-Fi connectivity, USB connectivity, Bluetooth connectivity, and the like, and a multimedia processor 112 that may, for example, detect and recognize gestures. In one implementation, the NPU is implemented in the CPU, DSP, and/or GPU. The SOC 100 may also include a sensor processor 114, image signal processors (ISPs), and/or navigation 120, which may include a global positioning system.

**[0032]** The SOC 100 may be based on an ARM instruction set. In an aspect of the present disclosure, the instructions loaded into the general-purpose processor 102 may comprise code for determining a probability distribution function (PDF) of an occupancy level for a location. The instructions loaded into the general-purpose processor 102 may also comprise code for integrating a depth map with localization information to generate a three-dimensional (3D) map. The instructions loaded into the general-purpose processor 102 may also comprise code for planning motion based on the 3D map, the localization information, and/or a user input. In one configuration, the motion planning overrides the user input when the user input is predicted to cause a collision.

**[0033]** FIGURE 2 illustrates an example implementation of a system 200 in accordance with certain aspects of the present disclosure. As illustrated in FIGURE 2, the system 200 may have multiple local processing units 202 that may perform various operations of methods described herein. Each local processing unit 202 may comprise a local state memory 204 and a local parameter memory 206 that may store parameters of

a neural network. In addition, the local processing unit 202 may have a local (neuron) model program (LMP) memory 208 for storing a local model program, a local learning program (LLP) memory 210 for storing a local learning program, and a local connection memory 212. Furthermore, as illustrated in FIGURE 2, each local processing unit 202 may interface with a configuration processor unit 214 for providing configurations for local memories of the local processing unit, and with a routing connection processing unit 216 that provides routing between the local processing units 202.

**[0034]** In one configuration, a navigation model is configured for integrating a depth map with localization information to generate a three-dimensional (3D) map and motion planning based on the 3D map, the localization information, and/or a user input. In one aspect, the aforementioned means may be the general-purpose processor 102, program memory associated with the general-purpose processor 102, memory block 118, local processing units 202, and or the routing connection processing units 216 configured to perform the functions recited. In another configuration, the aforementioned means may be any module or any apparatus configured to perform the functions recited by the aforementioned means.

**[0035]** According to certain aspects of the present disclosure, each local processing unit 202 may be configured to determine parameters of the model based upon desired one or more functional features of the model, and develop the one or more functional features towards the desired functional features as the determined parameters are further adapted, tuned and updated.

## AUTONOMOUS VISUAL NAVIGATION

**[0036]** In conventional systems, during manual operation, robots may use a virtual bumper system based on a reactive (e.g., memoryless) system, such as sonar or ultrasound, to halt the robot if the range between the robot and an obstacle is less than a threshold. The virtual bumper may not be useful during autonomous operation because the virtual bumper system is memoryless. Therefore, an autonomous robot cannot plan based on previous measurements. Furthermore, the virtual bumper system does not integrate sensor readings across time and pose. Thus, the system has increased noise. Furthermore, because the virtual bumper system is reactive, the virtual bumper system cannot plan a new trajectory and/or velocity to avoid a collision. Rather, the robot halts operation to avoid a collision and is not redirected around obstacles.



**[0037]** Additionally, in conventional systems for autonomous operation, a waypoint-based scheme is specified to allow the user to specify location-based waypoints, such as GPS waypoints, for navigation. In this system, a robot is specified to move from one waypoint to another waypoint. Still, the waypoint-based scheme uses a location system, such as GPS, which may be unreliable in some locations, such as indoor locations or urban locations. Furthermore, the waypoint-based scheme does not redirect the robot around obstacles.

**[0038]** Accordingly, it is desirable to combine different components into one robot to improve navigation and an understanding of the robot's environment. The different components may include a localization module, such as a visual-inertial odometer, for determining location of the robot in an environment; a mapping module for obtaining multi-view depth measurements; and a planning module for sampling-based geometric planning combined with re-planning. That is, as previously discussed, aspects of the present disclosure are directed to a robot that combines localization mapping and motion planning. More specifically, motion planning may use the depth map and the localization information.

**[0039]** According to aspects of the present disclosure, the combined components may be used for motion planning based on an obstacle map generated from a mapping module, localization information obtained from the localization module, and planning based on a user input and/or the planning module. In one configuration, the robot receives user input for a trajectory and/or velocity. In this configuration, the robot may override the user input to avoid collisions based on obstacles determined from the mapping module. The robot may also override the user input to improve the route, such as selecting a route that uses less fuel and/or selecting a route that is short in distance and/or time.

**[0040]** The robot may include multiple sensor modalities, such as a vision sensor and a localization sensor. The mapping module may use a stereo-vision sensor, such as multiple forward facing cameras to obtain a depth map. The depth measurements may be performed at each time step to determine a distance of multiple points from the robot. Furthermore, the mapping module may be specified for a volumetric integration of successive depth measurements (e.g., multi-view depth measurements). The robot may obtain localization information from the inertial measurement unit, such as

accelerometers and/or gyroscopes. The localization information may also be obtained from sensors, such as downward facing sensors. Localization refers to the location of the robot within an environment. In one configuration, the downward facing sensor is a sonar sensor and/or a camera.

**[0041]** In one aspect of the disclosure, the robot integrates each received depth measurement with the measurements of a localization module to place the depth measurements into an appropriate position in a map, such as a three-dimensional (3D) map. For example, the robot may obtain depth measurements at a first location and the robot associates the measurements with the first location of the robot. Additionally, the robot may move from the first location to a second location and obtain new depth measurements. Furthermore, the robot may associate the new depth measurements with the second location. Accordingly, the map, such as a three-dimensional map, is generated based on the depth measurements and localization measurements as the robot moves through an environment.

**[0042]** During the planning stage, the robot may provide motion control based on a mode of operation, such as a manual mode or an autonomous mode. For example, in the manual mode of operation, the autonomous system extrapolates a pilot's command (e.g., user input) for velocity and/or trajectory into the three-dimensional map and checks for collisions. Furthermore, the robot may override the user input to avoid a collision or to plan an improved route (e.g., faster route). In the autonomous mode of operation, the robot uses an accumulated map to plan a collision-free path to a saved goal. Furthermore, the autonomous system may re-plan along the route if map updates render a plan unsafe.

**[0043]** FIGURE 3 is a block diagram of a robot 300 according to an aspect of the present disclosure. As shown in FIGURE 3, the robot 300 includes a mapping module 302, which receives depth measurements from a depth sensor. The depth measurements measure the robot's distance from obstacles. In one configuration, the depth sensor includes multiple forward facing vision sensors, such as cameras. The mapping module 302 may generate a depth map based on the depth measurements. Furthermore, the depth map may be input to a 3D obstacle mapping module 304. The input of the depth map may be substantially instantaneous.

**[0044]** As shown in FIGURE 3, a location module 308 receives localization measurements and/or inertial measurements. The localization measurements may be obtained from a downward facing sensor, such as a downward facing camera. Furthermore, the inertial measurements may be obtained from the inertial measurement unit, such as accelerometers and/or gyroscopes. Based on the received localization measurements and/or inertial measurements, the location module 308 determines the robot's location and outputs the location to the 3D obstacle mapping module 304. The input of the location information may be substantially instantaneous.

**[0045]** The 3D obstacle mapping module 304 associates the received location information with the received depth map to generate a 3D map of the robot's current location. The 3D map may be referred to as an integrated depth map. Over time, as the robot 300 moves to different locations and obtains depth measurements from the different locations, a 3D map of the visited locations is generated. The 3D map is output to a planning module 306. In one configuration, the robot 300 receives a user input to move the robot 300 at a specific velocity and/or trajectory. Additionally, the user input may set a goal for the robot 300 to reach. The user input may be received from a remote device, such as a computer or remote controller.

**[0046]** In one configuration, when the user input sets a goal for the robot 300, the planning module 306 plans a route from the robot's current position to the goal. The desired path is output to the motion control module 310, which controls the robot's motors to move along the planned route at the desired trajectory and/or velocity. The motors may control a locomotion component, such as engines, rotary blades, and/or wheels. Furthermore, as shown in FIGURE 3, the location module 308 transmits the obtained location to both the planning module 306 and the motion control module 310 so that both modules are aware of the robot's current location. Additionally, as the location of the robot 300 changes, the 3D map generated from the 3D obstacle mapping module 304 may be updated. Thus, the planning module may update the desired path based on updates to the 3D map.

**[0047]** In another configuration, as discussed above, the robot 300 receives a user input to move the robot 300 at a specific velocity and/or trajectory. In this configuration, the planning module 306 overrides the user input to avoid a collision or to improve the specified velocity and/or trajectory based on the current location in the

3D map. The override information may be output to the motion control module 310 to control the robot's motors based on the override information. Of course, in some cases, the planning module may not override the user input. Thus, the user input is output from the planning module 306 to the motion control module 310 as the desired path.

**[0048]** FIGURE 4 illustrates an example of a location module 400 according to an aspect of the present discourse. The location module 400 may be a sensor fusion module that combines inertial measurement unit (IMU) data and sensor data to compute a robot's pose in a global/world coordinate frame. In one configuration, the sensor for the location module is a downward facing sensor. Still, aspects of the present disclosure are not limited to a downward facing sensor. As shown in FIGURE 4, the location module 400 may include a feature detector 402, a tracker 406, and a filter 404. In one configuration, the filter 404 is an extended Kalman filter. According to an aspect of the present disclosure, the feature detector receives frames from a downward facing sensor, such as a camera. In one configuration, the camera is a monocular camera.

**[0049]** The feature detector 402 may be configured to detect features of a received frame. Furthermore, the tracker 406 may track the detected changes in the position of the features across received frames to estimate the robot's movement. For example, the feature detector 402 may detect ten features in a first frame. Additionally, in this example, the tracker 406 may track the detected ten features in subsequent frames. Thus, when a second frame is received at the feature detector 402, the feature detector 402 attempts to detect the ten features of the first frame from the second frame and the tracker 406 determines the change in position of the ten features of the first frame from the ten features of the second frame. The change in position provides an estimate of the robot's movement (e.g., how far the robot has moved). Of course, ten features is an example and aspects of the present disclosure are not limited to detecting a specific number of features.

**[0050]** Additionally, as shown in FIGURE 4, a filter 404 may receive input from a first sensor and a second sensor. The first sensor may be an accelerometer to sense linear acceleration due to self-motion and the force of gravity. The second sensor may be a gyroscope to obtain an orientation (e.g., angles) of a three-dimensional global location. For visual-inertial odometry, the world frame may be centered on a location of a first captured image. Thus, relative to a robot's starting position, visual-inertial

odometry provides an estimate of a three-dimensional global location and corresponding orientation. For example, the three-dimensional global location and corresponding orientation information may be input to a trajectory controller for waypoint navigation or velocity control. The three-dimensional global location and corresponding orientation information may also be input to a mapping of a subsystem for voxel map integration.

**[0051]** In one configuration, the filter combines the input of the first sensor, the second sensor, and the feature information from the feature detector 402 and tracker 406 to output a location of the robot. The robot's location may be output as a six-degree of freedom (6 DOF) pose. The six-degree of freedom pose may refer to forward/back, up/down, left/right, pitch, yaw, and roll. In one configuration, the robot is configured to fly. Of course, aspects of the present disclosure are not limited to a flying robot and are contemplated for a robot that may be land and/or sea based.

**[0052]** FIGURE 5 illustrates an example of a depth sensor 500 according to an aspect of the present disclosure. As shown in FIGURE 5, the depth sensor 500 includes a first camera 502 and a second camera 504 that are spaced apart a known distance 506. As shown in FIGURE 5, both cameras 502, 504 are focused on a specific point 512 of an object 514. Furthermore, the location of a first pixel 508 corresponding to the specific point 512 in a first image 516 is determined for the first camera 502 and the location of a second pixel 510 corresponding to the specific point 512 in a second image 518 is determined for the second camera 504. Based the location of the first pixel 508 and second pixel 510, using triangulation, the robot may determine the depth of the specific point 512 of the object 514 from the cameras 502, 504. The process of determining the depth may be repeated for multiple pixels (e.g., each pixel) of the images obtained by the camera to create a depth map of the robot's current location.

**[0053]** Although not shown in FIGURE 5, the cameras 502, 504 may be mounted on the robot's nose (e.g., front). A three-dimensional reconstruction may be developed based on stereo depth perception. In one configuration, the process discretizes the world into a three-dimensional grid. Each location in the grid may be referred to as a voxel. A depth image from the stereo camera and a six-degree of freedom pose from the localization process may be combined into a three-dimensional voxel map as the robot navigates around an environment and observes obstacles. In one configuration,

the voxel map is generated based on the observation. Additionally, the mapping process may run at the stereo camera frame rate, such that the map is updated as new information is observed by the cameras. The update may be substantially instantaneous. Furthermore, the voxel map may be used by a motion planner to plan safe and reliable paths in real-time.

**[0054]** Visual-inertial odometry may provide localization information to the planner, which may be updated in real-time. In one configuration, when a target is specified for the robot, the planner uses a latest state estimate as a starting state and the target as a goal to generate a plan for reaching the goal. The robot may execute the plan by following a computed trajectory in a closed-loop manner. As the plan is executed, new information is observed that may invalidate the plan, in which case the planner adopts a new plan.

**[0055]** FIGURE 6A illustrates an example of an environment 600 that includes a mapped area 602 and an un-mapped area 604. The mapped area 602 may be a 3D mapped area based on depth measurements and localization. That is, as previously discussed, the 3D map may be generated by integrating the depth map with the localization information. The depth map and localization information may be obtained while the robot is operated in a manual mode (e.g., via user input) or an autonomous mode. As shown in FIGURE 6A, based on the depth measurements, the robot may determine that certain voxels are occupied by an object 608. Thus, the planning module is specified to avoid the object 608 when the robot is navigated in a manual mode or an autonomous mode. The un-mapped area 604 may be an area that has yet to be explored by the robot.

**[0056]** In one configuration, the robot may be at a first location 606 and may receive user input to autonomously navigate to a target location 610. As shown in FIGURE 6A, the target location 610 may be in an un-mapped area 604. In response to the user input, the robot selects candidate points 612 in unoccupied areas of the mapped area 602. The candidate points 612 may be randomly selected from the unoccupied areas of the mapped area 602. The candidate points 612 may be points on the mapped area 602 that are deemed safe for travel. Thus, the candidate points 612 may be used by the planning module to select a collision-free trajectory from the first location 606 to the target location 610. That is, the robot plans trajectories 614 (e.g., edges) between each of the

candidate points 612. As shown in FIGURE 6A, the robot may select candidate points 612 and trajectories 614 in the un-mapped area 604. Still, because the un-mapped area 604 has yet to be explored, the candidate points 612 and trajectories 614 are predicted candidate points 612 and predicted trajectories 614.

**[0057]** FIGURE 6B illustrates an example of a selected path 620 between the first location 606 and the target location 610 according to an aspect of the present disclosure. As shown in FIGURE 6B, after determining the candidate points 612 and trajectories 614, the planning module may search the various paths available between the first location 606 and the target location 610 based on the determined candidate points 612 and trajectories 614. The search may be a Dykstra search. Based on a search of the different paths, the planning module selects a path 620 for the robot to travel. The selected path is a collision free path. Furthermore, the selected path 620 may use fewer resources, such as gas, and/or time in comparison to other paths. The selected path 620 may be referred to as the minimum cost path.

**[0058]** FIGURE 6C illustrates an example of the robot moving along the selected path 620 according to an aspect of the present disclosure. As shown in FIGURE 6C, based on the selected path 620, the robot moves from a first location 606 to a second location 622 along the path 620. The second location 622 is one of the candidate points 612 selected during the planning phase. Additionally, as shown in FIGURE 6C, as the robot moves along the path 620, a portion of the un-mapped area 604 becomes mapped as the area comes into view of the robot's sensors. That is, as an example, as shown in FIGURE 6C, an area 624, which was previously included in the un-mapped area 604, becomes included in the mapped area 602 based on the robot's movement. Thus, as the robot moves along the path 620, the robot may alter the selected path based on updated depth maps and location information.

**[0059]** FIGURE 6D illustrates an example of the robot moving along the selected path 620 according to an aspect of the present disclosure. As shown in FIGURE 6D, based on the selected path 620, the robot moves from the second location 622 to a third location 626 along the path 620. The third location 626 is one of the candidate points 612 selected during the planning phase. Additionally, as shown in FIGURE 6D, as the robot moves along the path 620, a portion of the un-mapped area 604 becomes a mapped area as that area comes into view of the robot's sensors. In this example, based

on the depth measurements, the robot determines that an object 628 occupies the voxels along the planned path 620. Thus, if the robot were to continue along the planned path 620, a collision would occur. Accordingly, given that an object 628 overlaps the planned path 620, the planning module of the robot selects a new path based on the updated depth measurements and location information.

**[0060]** FIGURE 6E illustrates an example of a robot selecting a new path 630 based on an updated 3D map. As previously discussed, the planning module of the robot selects a new path based on the updated depth measurements and location information (e.g., the updated 3D map). As shown in FIGURE 6E, when the robot is at the third location 626, given the knowledge of the object 628, the planning module searches for new paths between the third location 626 and the target location 610. That is, the robot selects a new path when an obstacle, such as the object 628, comes into view. Based on a search of the different paths, the planning module selects a new path 630 for the robot to travel. As previously discussed, the selected path 630 may be referred to as the minimum cost path. For brevity, FIGURE 6E does not illustrate an un-mapped area 604. Still, in this example, when the robot is in the third location 626, a portion of the environment 600 may still be un-mapped. Furthermore, the new path 630 may be updated again as the robot moves receives updated depth maps and location information as the robot moves along different points of the path 630.

**[0061]** FIGURE 7 illustrates an example of a robot 702 in a mapped environment 700 according to an aspect of the present disclosure. The mapped environment 700 may be a 3D mapped area based on depth measurements and localization. That is, as previously discussed, the 3D map may be generated by integrating the depth map with the localization information. The depth map and localization information may be obtained while the robot is operated in a manual mode (e.g., via user input) or an autonomous mode. As shown in FIGURE 7, based on the depth measurements, the robot 702 may determine that certain voxels are occupied by a first object 706 and a second object 704. Thus, the planning module is specified to avoid the first object 706 and the second object 704 when the robot is navigated in a manual mode or an autonomous mode.

**[0062]** In one configuration, the robot 702 receives user input to move along a first trajectory 708. The first trajectory 708 may be specified to reach a target 720.



Alternatively, the first trajectory 708 may not be target oriented. Additionally, or alternatively, the user input may specify a velocity. Based on the generated 3D map, the robot 702 has knowledge that the first trajectory 708 received via user input will cause a collision with the second object 704.

**[0063]** In one configuration, the planning module overrides the user input and plans a second trajectory 722 to avoid the second object 704. The second trajectory 722 may be randomly selected to avoid the obstacle. Furthermore, the second trajectory 722 may not be target oriented. Furthermore, the second trajectory 722 may cause the motors of the robot 702 to navigate on the second trajectory 722 instead of the first trajectory 708. Additionally, or alternatively, the override may adjust the velocity of the robot 702. If the environment is a 3D environment, the planning module may still maintain the first trajectory 708. However, the planning module may override the planned altitude so that the robot 702 travels over or under the first object, if possible.

**[0064]** In another configuration, the robot performs motion planning to avoid the collision. The motion planning may include randomly selecting safe points in the mapped area 700 (e.g., the 3D map). For example, the robot may randomly select candidate points 710 as alternate navigation points. The randomly selected candidate points 710 may be points on the map that are deemed safe for travel (e.g., collision-free areas of the map). The randomly selected candidate points 710 may be candidate points within a pre-determined proximity to the user designated path (e.g., the first trajectory 708). Of course, aspects of the present disclosure are not limited to selecting candidate points within a pre-determined proximity to the user designated path, as points outside the pre-determined proximity may also be selected.

**[0065]** Additionally, as shown in FIGURE 7, to override the user-input to avoid the collision, the robot may determine one or more collision-free paths 712 between the randomly selected candidate points 710. To avoid the collision, the robot may travel along one of the collision-free paths from an initial position of the robot 702 to one of the randomly selected candidate points 710. Additionally, after reaching one of the randomly selected candidate points 710, the robot may select another path to other randomly selected candidate points 710, until the robot 702 reaches the target 720 or until the robot 702 receives a new user input.

**[0066]** Moreover, in one configuration, after determining one or more collision-free paths 712, the robot may determine a minimum cost path between the initial position and the target 720. The minimum cost path may use fewer resources, such as gas, and/or time, in comparison to other paths. In the example of FIGURE 7, the robot selects a minimum cost path 714 to avoid the collision predicted from the first trajectory 708. Finally, although not shown in FIGURE 7, the robot may change the selected path if a new obstacle were to materialize when the robot is traveling along one of the collision-free paths 712, such as the minimum cost path 714.

**[0067]** FIGURE 8 is a diagram illustrating an example of a hardware implementation for an apparatus 800, such as a robot, employing a processing system 820. The processing system 820 may be implemented with a bus architecture, represented generally by the bus 824. The bus 824 may include any number of interconnecting buses and bridges depending on the specific application of the processing system 820 and the overall design constraints. The bus 824 links together various circuits including one or more processors and/or hardware modules, represented by the processor 804 the communication module 808, location module 806, sensor module 802, locomotion module 810, and the computer-readable medium 814. The bus 824 may also link various other circuits such as timing sources, peripherals, voltage regulators, and power management circuits, which are well known in the art, and therefore, will not be described any further.

**[0068]** The apparatus 800 includes a processing system 820 coupled to a transceiver 816. The transceiver 816 is coupled to one or more antennas 818. The transceiver 816 enables communicating with various other apparatus over a transmission medium. For example, the transceiver 816 may receive user input via transmissions from the user. The processing system 820 includes a processor 804 coupled to a computer-readable medium 814. The processor 804 is responsible for general processing, including the execution of software stored on the computer-readable medium 814. The software, when executed by the processor 804, causes the processing system 820 to perform the various functions described for any particular apparatus. The computer-readable medium 814 may also be used for storing data that is manipulated by the processor 804 when executing software.

**[0069]** The sensor module 802 may be used to obtain measurements via a first sensor 828 and a second sensor 826. The first sensor 828 may be a stereo-vision sensor, for performing measurements, such as a stereoscopic camera. The second sensor 826 may be a camera and/or an inertial measurement unit. Of course, aspects of the present disclosure are not limited to a stereo-vision sensor as other types of sensors, such as, for example, radar, thermal, sonar, and/or lasers are also contemplated for performing measurements. The measurements of the first sensor 828 and the second sensor 826 may be processed by one or more of the processor 804 the communication module 808, location module 806, locomotion module 810, the computer-readable medium 814, and other modules 832 834. As previously discussed, the measurements from the first sensor 828 may be used to obtain depth measurements. Furthermore, the measurements from the second sensor 826 may be used for localization. For example, the measurements from the second sensor 826 may be used by to location module 806 to determine a location of the apparatus 800. Furthermore, the measurements of the first sensor 828 and the second sensor 826 may be transmitted to an external device by the transceiver 816. The first sensor 828 and the second sensor 826 are not limited to being defined external to the apparatus 800, as shown in FIGURE 8, the first sensor 828 and the second sensor 826 may also be defined within the apparatus 800.

**[0070]** The location module 806 may be used to determine a location of the apparatus 800. The communication module 808 may use the transceiver 816 to send and receive information, such as the location of the apparatus 800, to an external device. The locomotion module 810 may be used to provide locomotion to the apparatus 800. As an example, locomotion may be provided via rotary blades 812. Of course, aspects of the present disclosure are not limited to providing locomotion via rotary blades 812 and are contemplated for any other type of component for providing locomotion, such as propellers, wheels, treads, fins, and/or jet engines.

**[0071]** The processing system 820 includes an integrating module 832 for integrating a depth map with localization information to generate a three-dimensional (3D) map. The processing system 820 also includes a planning module 834 for planning a motion based at least in part on the 3D map, the localization information, and/or a user input. In one configuration, the motion planning overrides the user input when the user input is expected (e.g., predicted) to cause a collision. The modules may

be software modules running in the processor 804, resident/stored in the computer-readable medium 814, one or more hardware modules coupled to the processor 804, or some combination thereof.

**[0072]** FIGURE 9 illustrates a method 900 for visual navigation for a robot. In block 902, in an optional configuration, the robot determines a depth map from multiple cameras and obtains localization information from multiple sensors. That is, the robot may include multiple sensor modalities. The depth measurements may be performed at each time step to determine a distance of multiple points from the robot. A mapping module may use a sensor, such as a stereo-vision sensor, to obtain a depth map. Thus, in one optional configuration, at block 904, the robot generates the depth map based on measurements obtained from a stereo camera.

**[0073]** Furthermore, the robot may obtain localization information from the inertial sensor information. Thus, in one optional configuration, at block 906, the robot obtains inertial sensor information from a gyroscope and/or an accelerometer unit. The localization information may also be obtained from sensors, such as downward facing sensors. Localization refers to the location of the robot within an environment. In one configuration, the downward facing sensor is a sonar sensor and/or a camera.

**[0074]** Additionally, in block 908, a robot integrates a depth map with localization information to generate a 3D map. For example, the robot may obtain depth measurements at a first location and the robot associates the measurements with the first location of the robot. Additionally, the robot may move from the first location to a second location and obtain new depth measurements. Furthermore, the robot may associate the new depth measurements with the second location. Accordingly, the map, such as a 3D map, is generated based on the depth measurements and localization measurements as the robot moves through an environment.

**[0075]** Furthermore, in block 910, the robot plans motion based on the 3D map, the localization information, and/or user input. In one configuration, when a target is specified for the robot, a motion planner of the robot uses a latest state estimate as a starting state and the target as a goal to generate a plan for reaching the goal. The robot may execute the plan by following a computed trajectory in a closed-loop manner. As the plan is executed, new information is observed that may invalidate the plan, in which

case the planner adopts a new plan. Additionally, or alternatively, the motion planning may override the user input when the user input is predicted to cause a collision. Moreover, the motion planning may override the user input when a trajectory and/or a velocity, received via the user input is predicted to cause a collision

**[0076]** In an optional configuration, in block 912, the robot selects a new trajectory and/or a new velocity when overriding the user input. For example, the user input may specify a velocity and/or a trajectory. Based on the generated 3D map, the robot has knowledge that the user selected velocity and/or a trajectory will cause a collision with an object. Thus, in this configuration, the planning module overrides the user input and plans a new trajectory and/or velocity to avoid the second object.

**[0077]** FIGURE 10 illustrates a more detailed method 1000 for motion planning. In block 1002, which is similar to block 910 of FIGURE 9, the robot plans motion based on the 3D map, the localization information, and/or user input. In block 1004, the motion planning overrides the user input when the user input is predicted to cause a collision. Additionally, in an optional configuration, in block 1006, the motion planning overrides the user input when a trajectory and/or a velocity, received via the user input, is predicted to cause the collision.

**[0078]** In another optional configuration, in block 1008, the motion planning randomly selects multiple collision-free points in the 3D map and determines one or more collision-free paths between two or more of the multiple collision-free points. Additionally, in the optional configuration, in block 1010, the motion planning further determines a minimum-cost path to a target based on the one or more collision-free path. Moreover, after determining the minimum-cost path, in block 1012, the motion planning may further alter the minimum-cost path when an obstacle is viewed along the minimum-cost path.

**[0079]** In yet another optional configuration, in block 1014, the motion planning randomly selects multiple predicted collision-free points in an unmapped area and determines one or more predicted collision-free paths between two or more of the multiple predicted collision-free points.

**[0080]** In some aspects, the methods 900 and 1000 may be performed by the SOC 100 (FIGURE 1) the system 200 (FIGURE 2), or the apparatus 800 (FIGURE 8). That

is, each of the elements of methods 900 and 1000 may, for example, but without limitation, be performed by the SOC 100, the system 200, the apparatus 800, or one or more processors (e.g., CPU 102, local processing unit 202, processor 804) and/or other components included therein.

**[0081]** The various operations of methods described above may be performed by any suitable means capable of performing the corresponding functions. The means may include various hardware and/or software component(s) and/or module(s), including, but not limited to, a circuit, an application specific integrated circuit (ASIC), or processor. Generally, where there are operations illustrated in the figures, those operations may have corresponding counterpart means-plus-function components with similar numbering.

**[0082]** As used herein, the term “determining” encompasses a wide variety of actions. For example, “determining” may include calculating, computing, processing, deriving, investigating, looking up (e.g., looking up in a table, a database or another data structure), ascertaining and the like. Additionally, “determining” may include receiving (e.g., receiving information), accessing (e.g., accessing data in a memory) and the like. Furthermore, “determining” may include resolving, selecting, choosing, establishing and the like.

**[0083]** As used herein, a phrase referring to “at least one of” a list of items refers to any combination of those items, including single members. As an example, “at least one of: a, b, or c” is intended to cover: a, b, c, a-b, a-c, b-c, and a-b-c.

**[0084]** The various illustrative logical blocks, modules and circuits described in connection with the present disclosure may be implemented or performed with a general-purpose processor, a digital signal processor (DSP), an application specific integrated circuit (ASIC), a field programmable gate array signal (FPGA) or other programmable logic device (PLD), discrete gate or transistor logic, discrete hardware components or any combination thereof designed to perform the functions described herein. A general-purpose processor may be a microprocessor, but in the alternative, the processor may be any commercially available processor, controller, microcontroller or state machine. A processor may also be implemented as a combination of computing devices, e.g., a combination of a DSP and a microprocessor, a plurality of

microprocessors, one or more microprocessors in conjunction with a DSP core, or any other such configuration.

**[0085]** The steps of a method or algorithm described in connection with the present disclosure may be embodied directly in hardware, in a software module executed by a processor, or in a combination of the two. A software module may reside in any form of storage medium that is known in the art. Some examples of storage media that may be used include random access memory (RAM), read only memory (ROM), flash memory, erasable programmable read-only memory (EPROM), electrically erasable programmable read-only memory (EEPROM), registers, a hard disk, a removable disk, a CD-ROM and so forth. A software module may comprise a single instruction, or many instructions, and may be distributed over several different code segments, among different programs, and across multiple storage media. A storage medium may be coupled to a processor such that the processor can read information from, and write information to, the storage medium. In the alternative, the storage medium may be integral to the processor.

**[0086]** The methods disclosed herein comprise one or more steps or actions for achieving the described method. The method steps and/or actions may be interchanged with one another without departing from the scope of the claims. In other words, unless a specific order of steps or actions is specified, the order and/or use of specific steps and/or actions may be modified without departing from the scope of the claims.

**[0087]** The functions described may be implemented in hardware, software, firmware, or any combination thereof. If implemented in hardware, an example hardware configuration may comprise a processing system in a device. The processing system may be implemented with a bus architecture. The bus may include any number of interconnecting buses and bridges depending on the specific application of the processing system and the overall design constraints. The bus may link together various circuits including a processor, machine-readable media, and a bus interface. The bus interface may be used to connect a network adapter, among other things, to the processing system via the bus. The network adapter may be used to implement signal processing functions. For certain aspects, a user interface (e.g., keypad, display, mouse, joystick, etc.) may also be connected to the bus. The bus may also link various other circuits such as timing sources, peripherals, voltage regulators, power management

circuits, and the like, which are well known in the art, and therefore, will not be described any further.

**[0088]** The processor may be responsible for managing the bus and general processing, including the execution of software stored on the machine-readable media. The processor may be implemented with one or more general-purpose and/or special-purpose processors. Examples include microprocessors, microcontrollers, DSP processors, and other circuitry that can execute software. Software shall be construed broadly to mean instructions, data, or any combination thereof, whether referred to as software, firmware, middleware, microcode, hardware description language, or otherwise. Machine-readable media may include, by way of example, random access memory (RAM), flash memory, read only memory (ROM), programmable read-only memory (PROM), erasable programmable read-only memory (EPROM), electrically erasable programmable Read-only memory (EEPROM), registers, magnetic disks, optical disks, hard drives, or any other suitable storage medium, or any combination thereof. The machine-readable media may be embodied in a computer-program product. The computer-program product may comprise packaging materials.

**[0089]** In a hardware implementation, the machine-readable media may be part of the processing system separate from the processor. However, as those skilled in the art will readily appreciate, the machine-readable media, or any portion thereof, may be external to the processing system. By way of example, the machine-readable media may include a transmission line, a carrier wave modulated by data, and/or a computer product separate from the device, all which may be accessed by the processor through the bus interface. Alternatively, or in addition, the machine-readable media, or any portion thereof, may be integrated into the processor, such as the case may be with cache and/or general register files. Although the various components discussed may be described as having a specific location, such as a local component, they may also be configured in various ways, such as certain components being configured as part of a distributed computing system.

**[0090]** The processing system may be configured as a general-purpose processing system with one or more microprocessors providing the processor functionality and external memory providing at least a portion of the machine-readable media, all linked together with other supporting circuitry through an external bus architecture.



Alternatively, the processing system may comprise one or more neuromorphic processors for implementing the neuron models and models of neural systems described herein. As another alternative, the processing system may be implemented with an application specific integrated circuit (ASIC) with the processor, the bus interface, the user interface, supporting circuitry, and at least a portion of the machine-readable media integrated into a single chip, or with one or more field programmable gate arrays (FPGAs), programmable logic devices (PLDs), controllers, state machines, gated logic, discrete hardware components, or any other suitable circuitry, or any combination of circuits that can perform the various functionality described throughout this disclosure. Those skilled in the art will recognize how best to implement the described functionality for the processing system depending on the particular application and the overall design constraints imposed on the overall system.

**[0091]** The machine-readable media may comprise a number of software modules. The software modules include instructions that, when executed by the processor, cause the processing system to perform various functions. The software modules may include a transmission module and a receiving module. Each software module may reside in a single storage device or be distributed across multiple storage devices. By way of example, a software module may be loaded into RAM from a hard drive when a triggering event occurs. During execution of the software module, the processor may load some of the instructions into cache to increase access speed. One or more cache lines may then be loaded into a general register file for execution by the processor. When referring to the functionality of a software module below, it will be understood that such functionality is implemented by the processor when executing instructions from that software module. Furthermore, it should be appreciated that aspects of the present disclosure result in improvements to the functioning of the processor, computer, machine, or other system implementing such aspects.

**[0092]** If implemented in software, the functions may be stored or transmitted over as one or more instructions or code on a computer-readable medium. Computer-readable media include both computer storage media and communication media including any medium that facilitates transfer of a computer program from one place to another. A storage medium may be any available medium that can be accessed by a computer. By way of example, and not limitation, such computer-readable media can

comprise RAM, ROM, EEPROM, CD-ROM or other optical disk storage, magnetic disk storage or other magnetic storage devices, or any other medium that can be used to carry or store desired program code in the form of instructions or data structures and that can be accessed by a computer. Additionally, any connection is properly termed a computer-readable medium. For example, if the software is transmitted from a website, server, or other remote source using a coaxial cable, fiber optic cable, twisted pair, digital subscriber line (DSL), or wireless technologies such as infrared (IR), radio, and microwave, then the coaxial cable, fiber optic cable, twisted pair, DSL, or wireless technologies such as infrared, radio, and microwave are included in the definition of medium. Disk and disc, as used herein, include compact disc (CD), laser disc, optical disc, digital versatile disc (DVD), floppy disk, and Blu-ray® disc where disks usually reproduce data magnetically, while discs reproduce data optically with lasers. Thus, in some aspects computer-readable media may comprise non-transitory computer-readable media (e.g., tangible media). In addition, for other aspects computer-readable media may comprise transitory computer-readable media (e.g., a signal). Combinations of the above should also be included within the scope of computer-readable media.

**[0093]** Thus, certain aspects may comprise a computer program product for performing the operations presented herein. For example, such a computer program product may comprise a computer-readable medium having instructions stored (and/or encoded) thereon, the instructions being executable by one or more processors to perform the operations described herein. For certain aspects, the computer program product may include packaging material.

**[0094]** Further, it should be appreciated that modules and/or other appropriate means for performing the methods and techniques described herein can be downloaded and/or otherwise obtained by a user terminal and/or base station as applicable. For example, such a device can be coupled to a server to facilitate the transfer of means for performing the methods described herein. Alternatively, various methods described herein can be provided via storage means (e.g., RAM, ROM, a physical storage medium such as a compact disc (CD) or floppy disk, etc.), such that a user terminal and/or base station can obtain the various methods upon coupling or providing the storage means to the device. Moreover, any other suitable technique for providing the methods and techniques described herein to a device can be utilized.

**[0095]** It is to be understood that the claims are not limited to the precise configuration and components illustrated above. Various modifications, changes and variations may be made in the arrangement, operation and details of the methods and apparatus described above without departing from the scope of the claims.

## CLAIMS

### WHAT IS CLAIMED IS:

1. A method of visual navigation for a robot, comprising:  
integrating a depth map with a localization information to generate a three-dimensional (3D) map; and  
motion planning based at least in part on the 3D map, the localization information, or a user input, the motion planning overriding the user input when the user input is predicted to cause a collision.
2. The method of claim 1, further comprising:  
determining the depth map from a plurality of cameras; and  
obtaining the localization information from a plurality of sensors.
3. The method of claim 2, in which the localization information comprises at least one of an image information, an inertial sensor information, or a combination thereof.
4. The method of claim 3, further comprising obtaining the inertial sensor information from at least one of a gyroscope, an accelerometer, or a combination thereof.
5. The method of claim 1, further comprising generating the depth map based on measurements obtained from a stereo camera.
6. The method of claim 1, further comprising selecting at least one of a new trajectory, a new velocity, or a combination thereof, when overriding the user input.
7. The method of claim 1, further comprising overriding the user input when at least one of a trajectory, a velocity, or a combination thereof, received via the user input, is predicted to cause the collision.

8. The method of claim 1, in which the motion planning comprises:
  - randomly selecting a plurality of collision-free points in the 3D map;
  - determining at least one collision-free path between at least two of the plurality of collision-free points; and
  - determining a minimum-cost path to a target based at least in part on the at least one collision-free path.
9. The method of claim 8, in which the motion planning further comprises altering the minimum-cost path when an obstacle is viewed along the minimum-cost path.
10. The method of claim 8, in which the motion planning further comprises:
  - randomly selecting a plurality of predicted collision-free points in an unmapped area; and
  - determining at least one predicted collision-free path between at least two of the plurality of predicted collision-free points.
11. An apparatus, comprising:
  - a memory; and
  - at least one processor coupled to the memory, the at least one processor configured:
    - to integrate a depth map with a localization information to generate a three-dimensional (3D) map; and
    - to motion plan based at least in part on the 3D map, the localization information, or a user input, the motion planning overriding the user input when the user input is predicted to cause a collision.
12. The apparatus of claim 11, in which the at least one processor is further configured:
  - to determine the depth map from a plurality of cameras; and
  - to obtain the localization information from a plurality of sensors.
13. The apparatus of claim 12, in which the localization information comprises at least one of an image information, an inertial sensor information, or a combination thereof.

14. The apparatus of claim 13, in which the at least one processor is further configured to obtain the inertial sensor information from at least one of a gyroscope, an accelerometer, or a combination thereof.

15. The apparatus of claim 11, in which the at least one processor is further configured to generate the depth map based on measurements obtained from a stereo camera.

16. The apparatus of claim 11, in which the at least one processor is further configured to select at least one of a new trajectory, a new velocity, or a combination thereof, when overriding the user input.

17. The apparatus of claim 11, in which the at least one processor is further configured to override the user input when at least one of a trajectory, a velocity, or a combination thereof, received via the user input, is predicted to cause the collision.

18. The apparatus of claim 11, in which the at least one processor is further configured:

- to randomly select a plurality of collision-free points in the 3D map;
- to determine at least one collision-free path between at least two of the plurality of collision-free points; and
- to determine a minimum-cost path to a target based at least in part on the at least one collision-free path.

19. The apparatus of claim 18, in which the at least one processor is further configured to alter the minimum-cost path when an obstacle is viewed along the minimum-cost path.

20. The apparatus of claim 18, in which the at least one processor is further configured:

- to randomly select a plurality of predicted collision-free points in an unmapped area; and
- to determine at least one predicted collision-free path between at least two of the plurality of predicted collision-free points.

21. An apparatus, comprising:  
means for integrating a depth map with a localization information to generate a three-dimensional (3D) map; and  
means for motion planning based at least in part on the 3D map, the localization information, or a user input, the motion planning overriding the user input when the user input is predicted to cause a collision.
22. The apparatus of claim 21, further comprising:  
means for determining the depth map from a plurality of cameras; and  
means for obtaining the localization information from a plurality of sensors.
23. The apparatus of claim 22, in which the localization information comprises at least one of an image information, an inertial sensor information, or a combination thereof.
24. The apparatus of claim 23, further comprising means for obtaining the inertial sensor information from at least one of a gyroscope, an accelerometer, or a combination thereof.
25. The apparatus of claim 21, further comprising means for generating the depth map based on measurements obtained from a stereo camera.
26. The apparatus of claim 21, further comprising means for selecting at least one of a new trajectory, a new velocity, or a combination thereof, when overriding the user input.
27. The apparatus of claim 21, further comprising means for overriding the user input when at least one of a trajectory, a velocity, or a combination thereof, received via the user input, is predicted to cause the collision.
28. The apparatus of claim 21, in which the means for motion planning further comprises:  
means for randomly selecting a plurality of collision-free points in the 3D map;  
means for determining at least one collision-free path between at least two of the plurality of collision-free points; and

means for determining a minimum-cost path to a target based at least in part on the at least one collision-free path.

29. The apparatus of claim 28, in which the means for motion planning further comprises means for altering the minimum-cost path when an obstacle is viewed along the minimum-cost path.

30. The apparatus of claim 28, in which the means for motion planning further comprises:

means for randomly selecting a plurality of predicted collision-free points in an unmapped area; and

means for determining at least one predicted collision-free path between at least two of the plurality of predicted collision-free points.

31. A non-transitory computer-readable medium having program code recorded thereon for providing visual navigation to a robot, the program code executed by a processor and comprising:

program code to integrate a depth map with a localization information to generate a three-dimensional (3D) map; and

program code to motion plan based at least in part on the 3D map, the localization information, or a user input, the motion planning overriding the user input when the user input is predicted to cause a collision.

32. The non-transitory computer-readable medium of claim 31, in which the program code further comprises:

program code to determine the depth map from a plurality of cameras; and

program code to obtain the localization information from a plurality of sensors.

33. The non-transitory computer-readable medium of claim 32, in which the localization information comprises at least one of image information, an inertial sensor information, or a combination thereof.

34. The non-transitory computer-readable medium of claim 33, in which the program code further comprises program code to obtain the inertial sensor information from at least one of a gyroscope, an accelerometer, or a combination thereof.



35. The non-transitory computer-readable medium of claim 31, in which the program code further comprises program code to generate the depth map based on measurements obtained from a stereo camera.

36. The non-transitory computer-readable medium of claim 31, in which the program code further comprises program code to select at least one of a new trajectory, a new velocity, or a combination thereof, when overriding the user input.

37. The non-transitory computer-readable medium of claim 31, in which the program code to motion plan further comprises program code to override the user input when at least one of a trajectory, a velocity, or a combination thereof, received via the user input, is predicted to cause the collision.

38. The non-transitory computer-readable medium of claim 31, in which the program code to motion plan further comprises:

program code to randomly select a plurality of collision-free points in the 3D map;

program code to determine at least one collision-free path between at least two of the plurality of collision-free points; and

program code to determine a minimum-cost path to a target based at least in part on the at least one collision-free path.

39. The non-transitory computer-readable medium of claim 38, in which the program code to motion plan further comprises program code to alter the minimum-cost path when an obstacle is viewed along the minimum-cost path.

40. The non-transitory computer-readable medium of claim 38, in which the program code to motion plan further comprises:

program code to randomly select a plurality of predicted collision-free points in an unmapped area; and

program code to determine at least one predicted collision-free path between at least two of the plurality of predicted collision-free points.

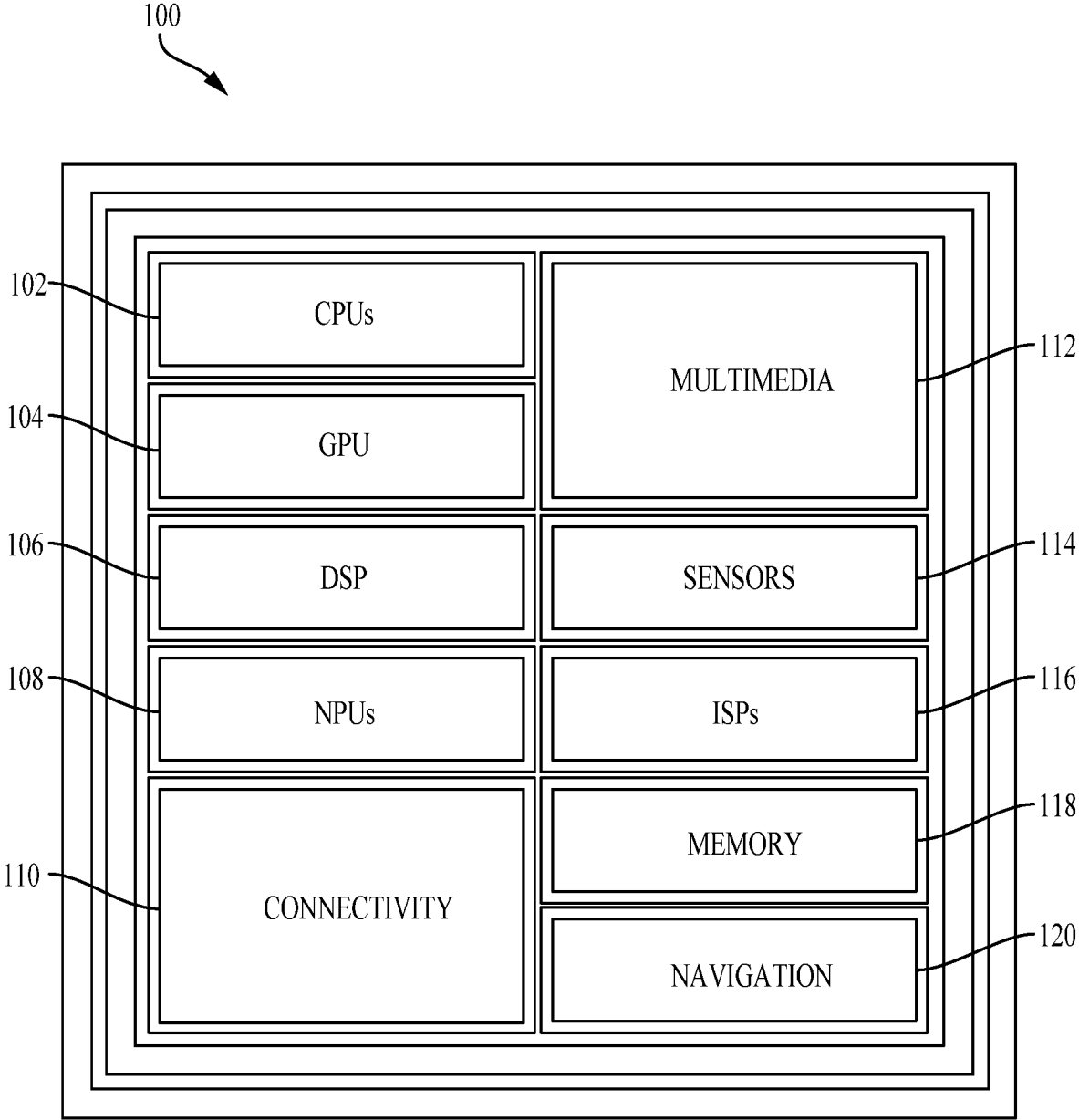


FIG. 1

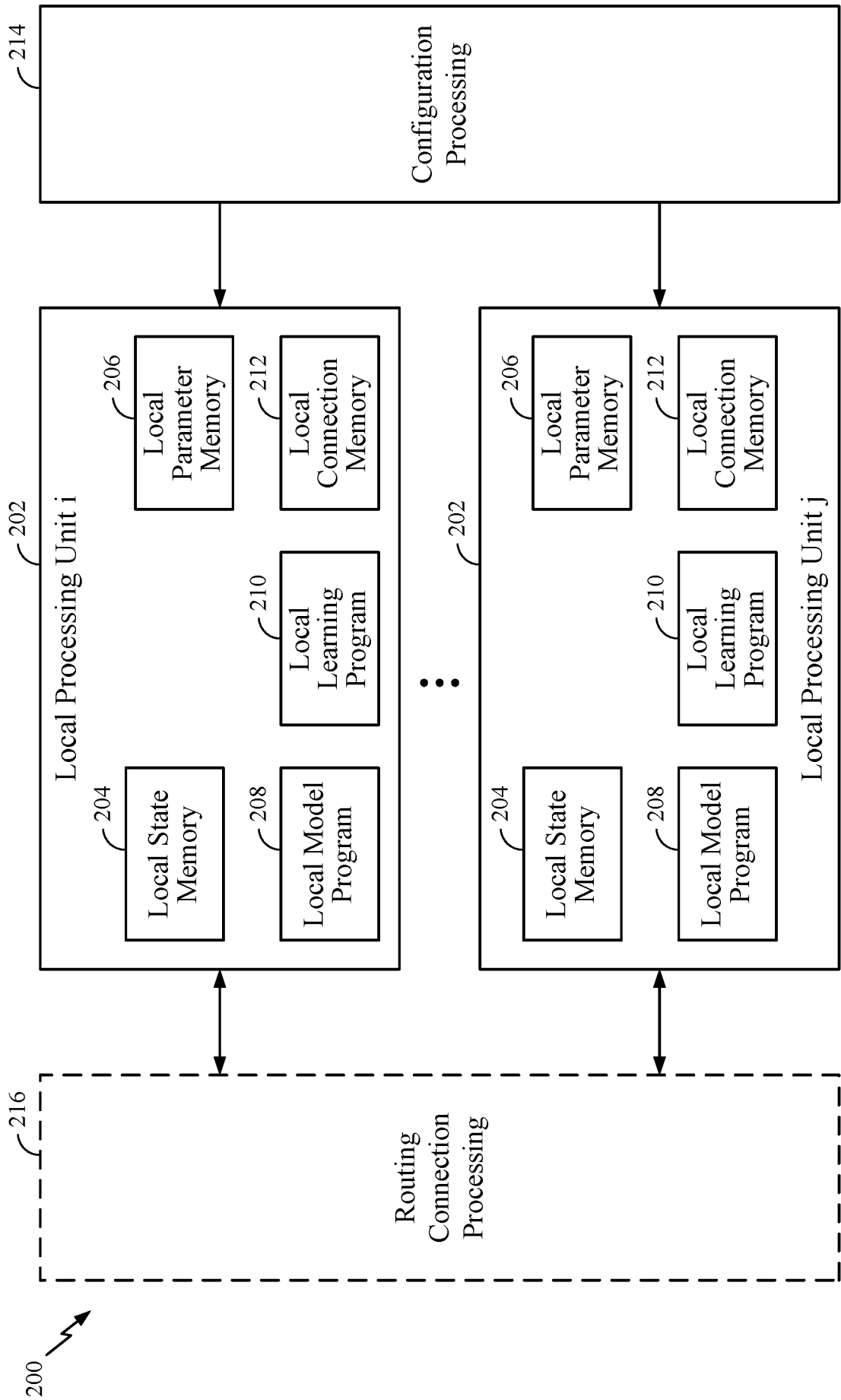


FIG. 2

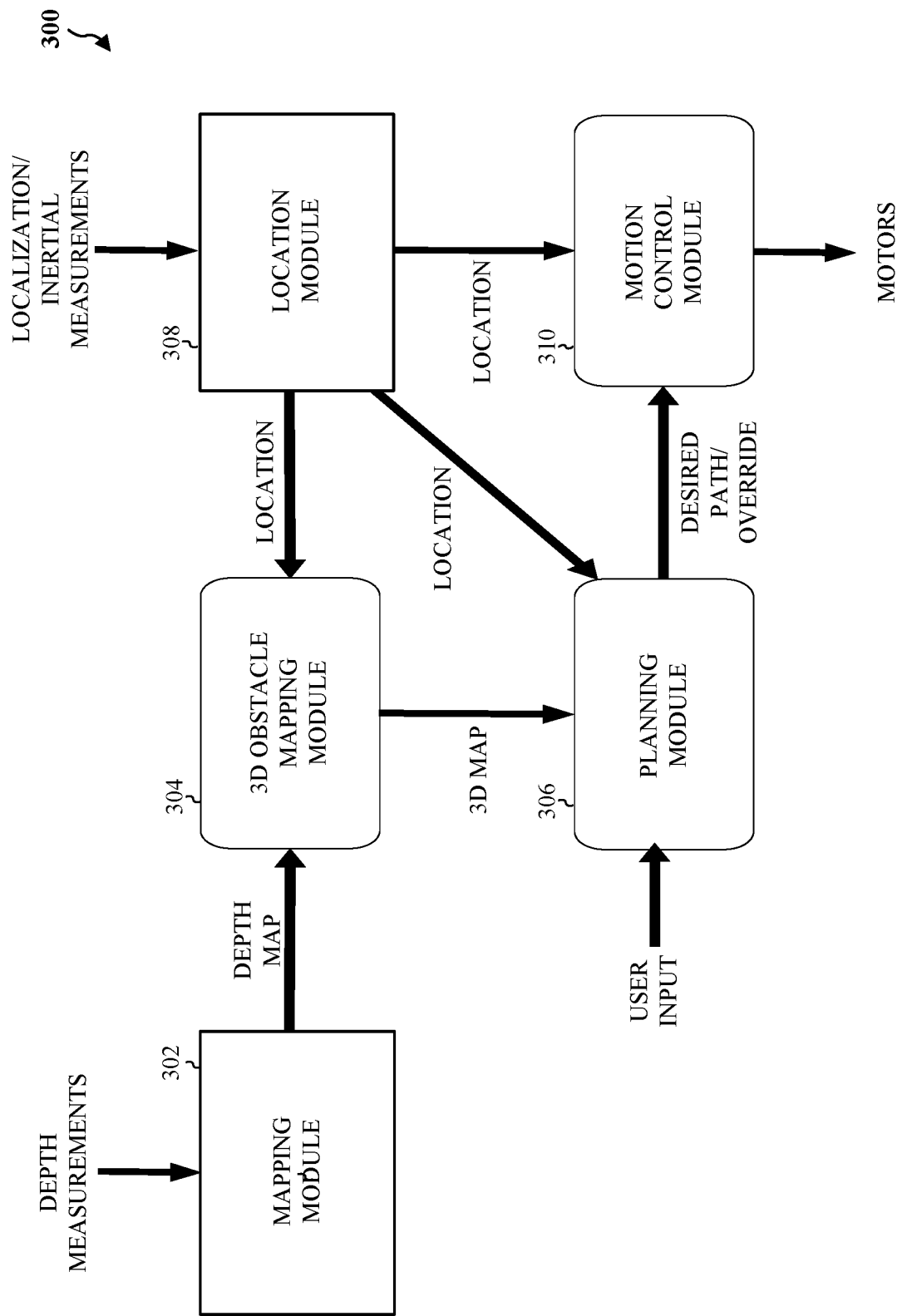
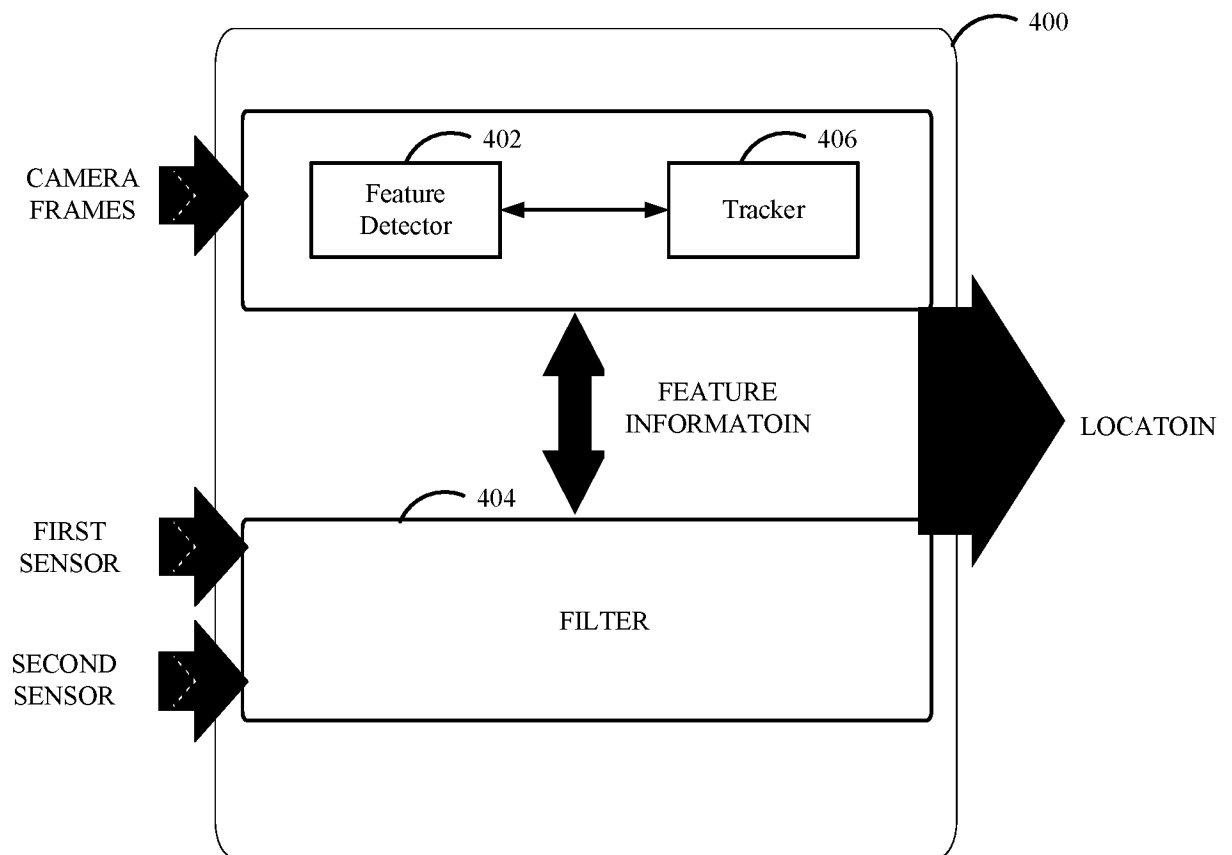
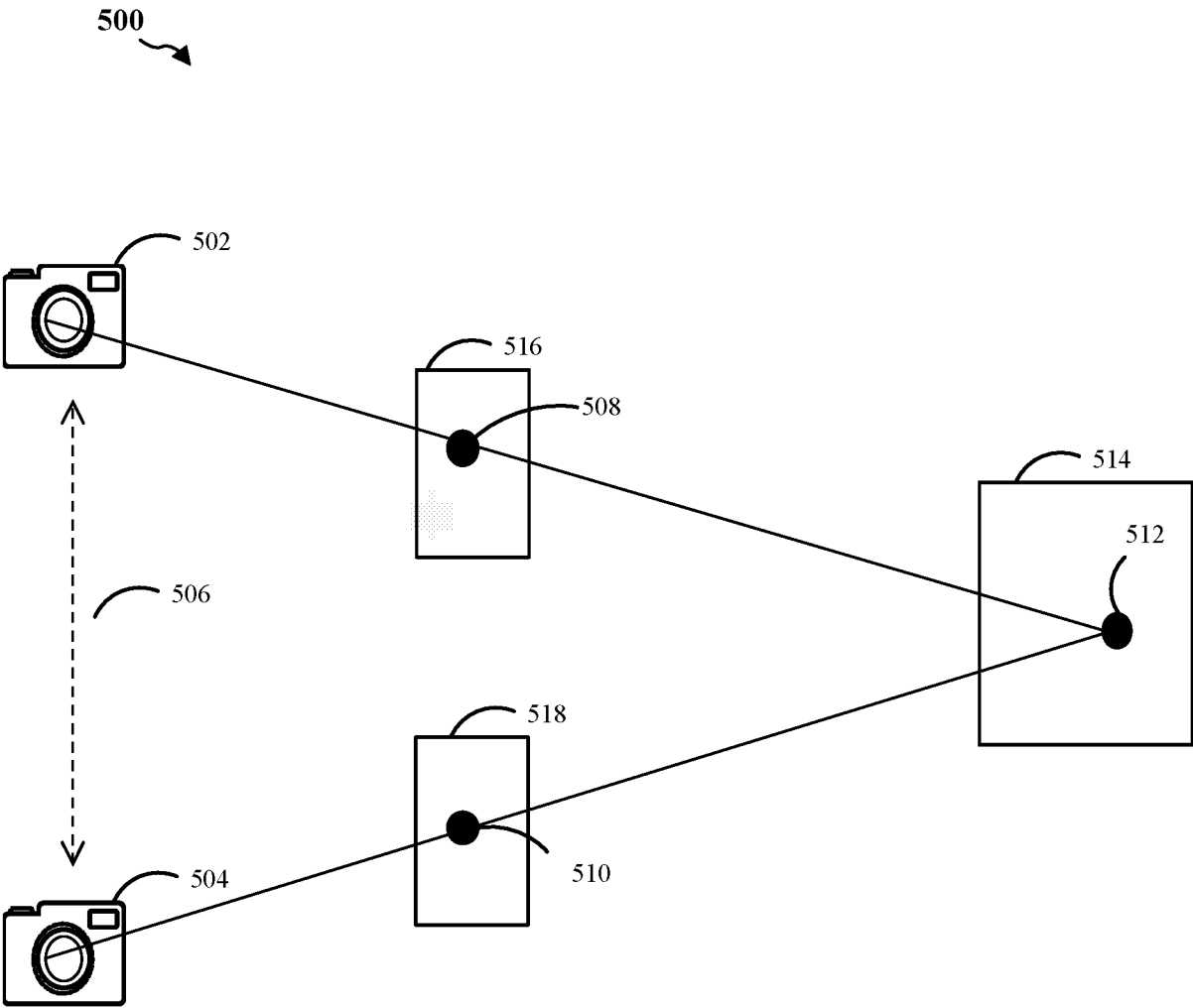


FIG. 3

**FIG. 4**



**FIG. 5**

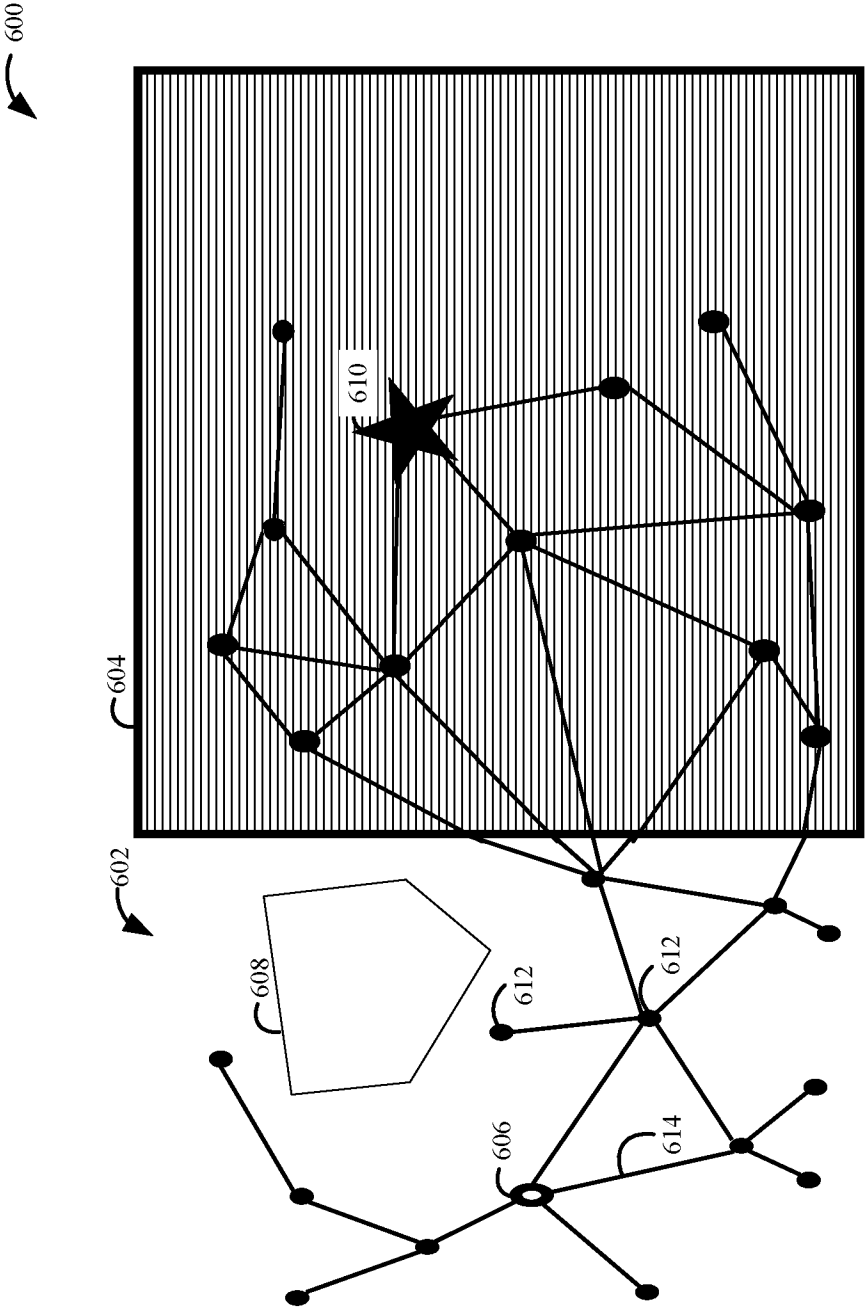


FIG. 6A

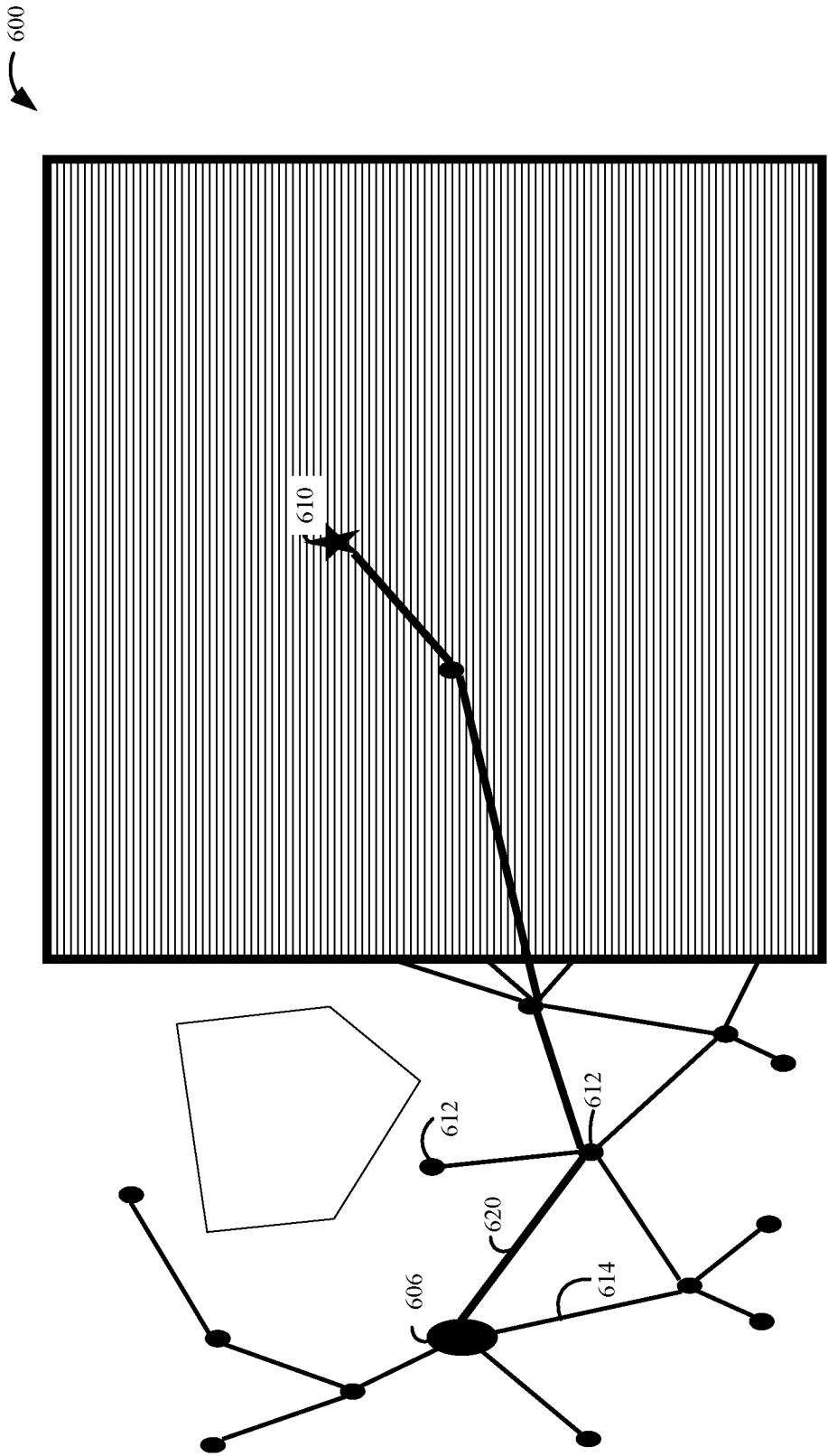


FIG. 6B



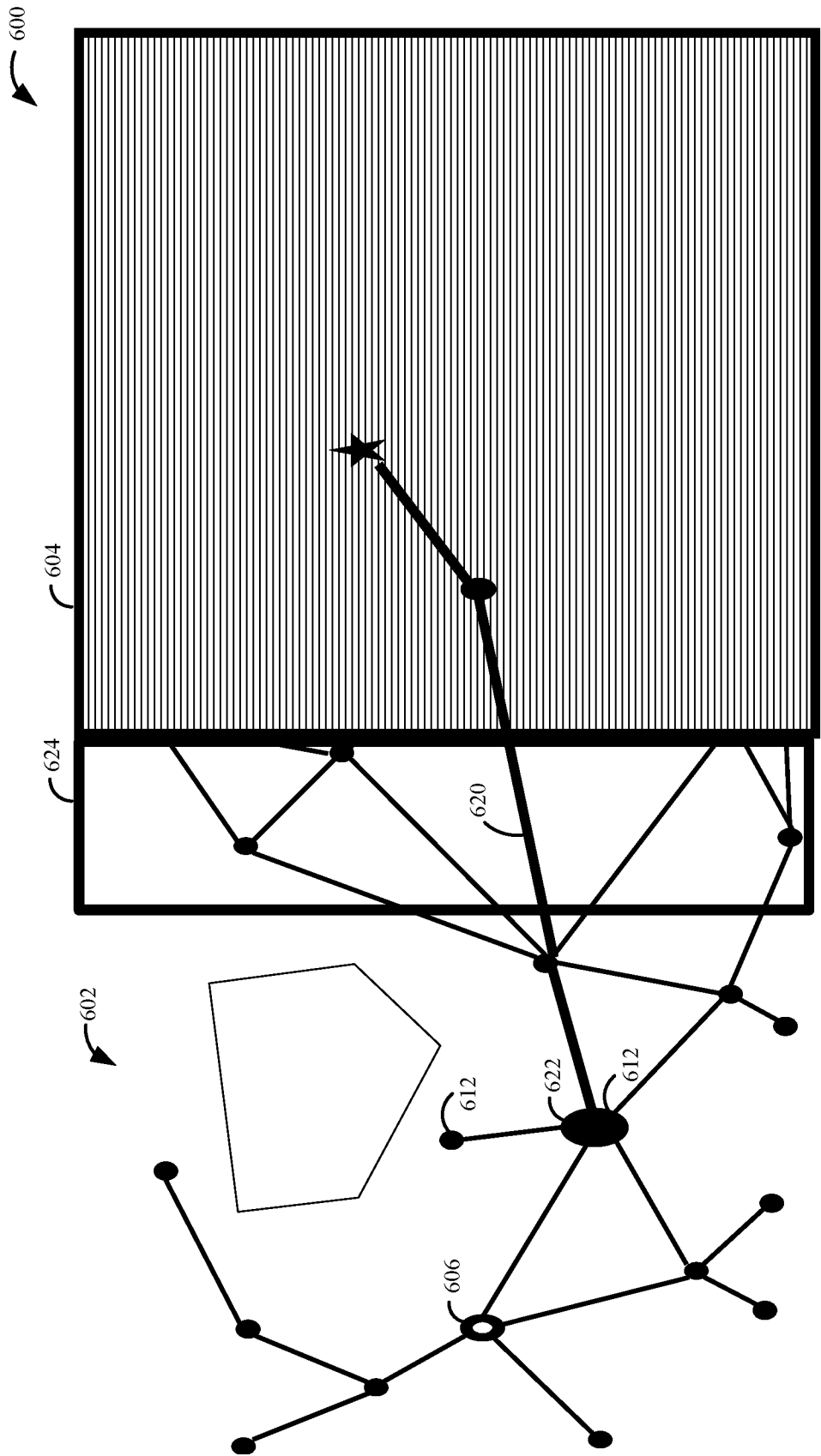


FIG. 6C

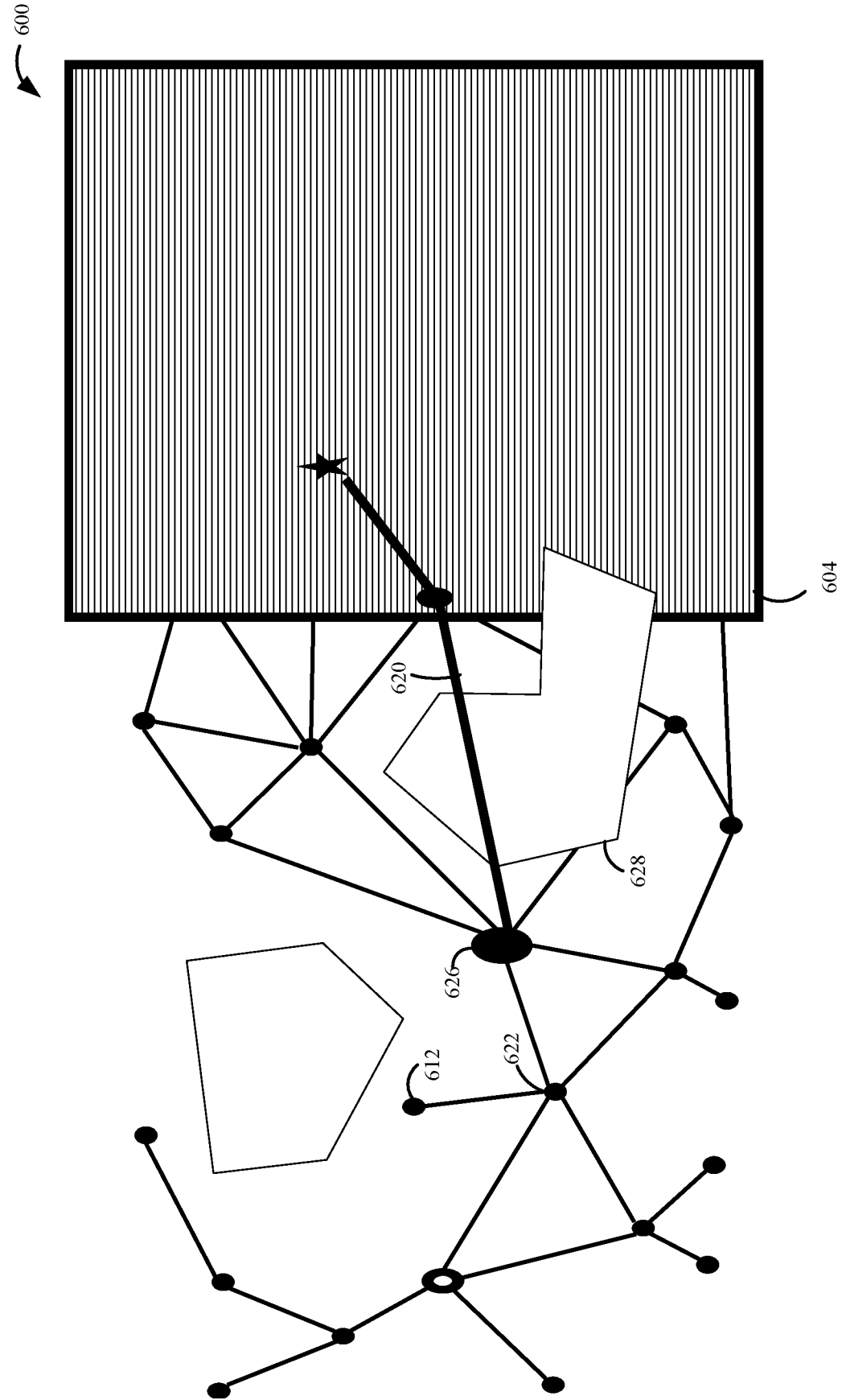


FIG. 6D

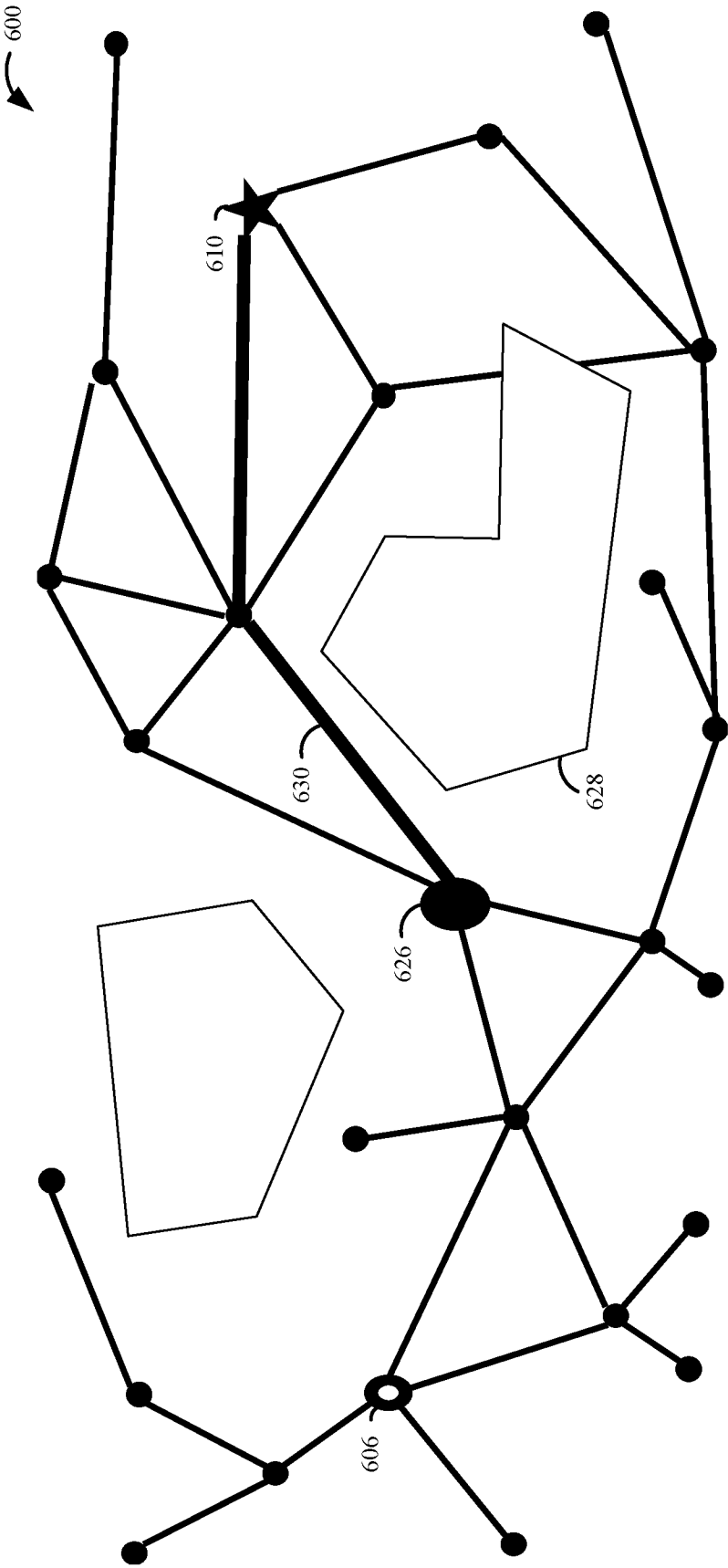


FIG. 6E

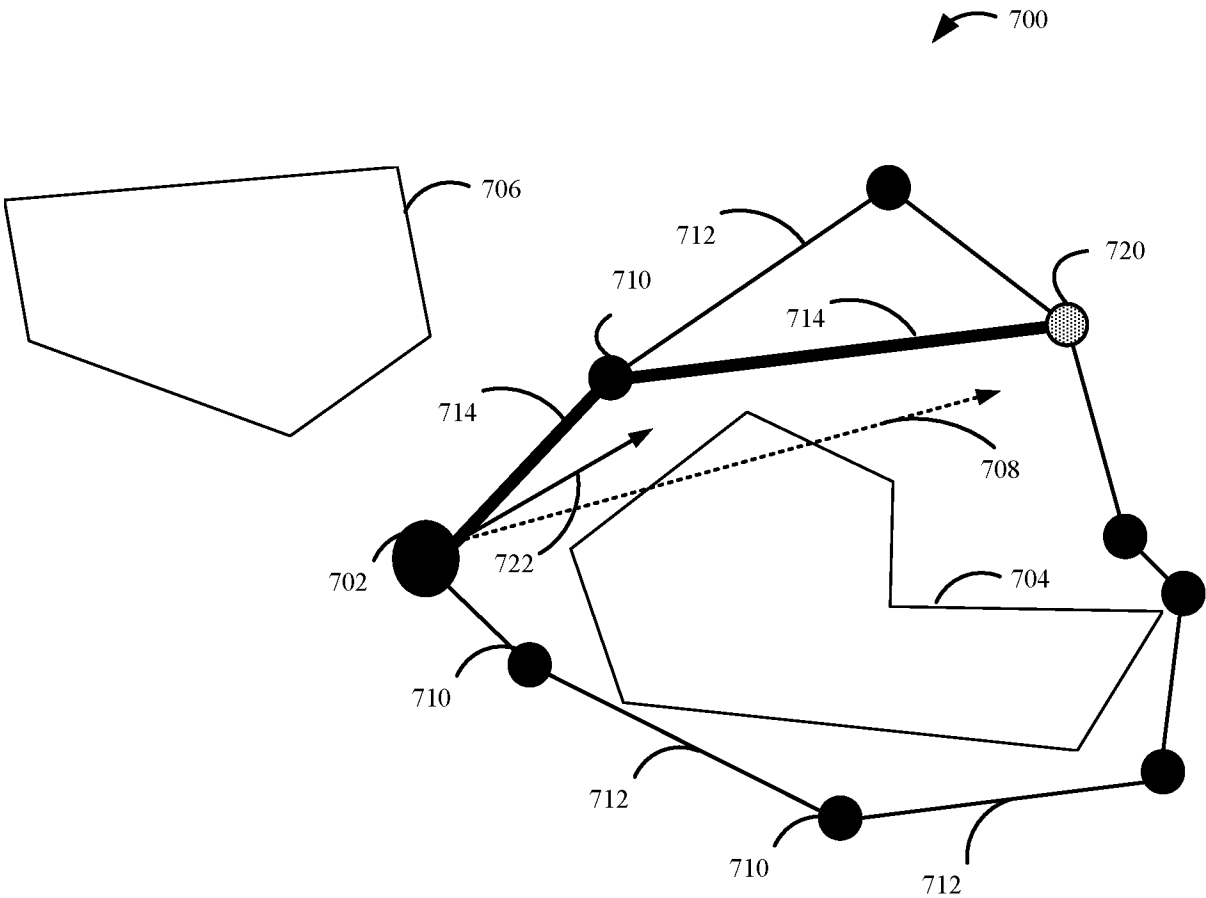
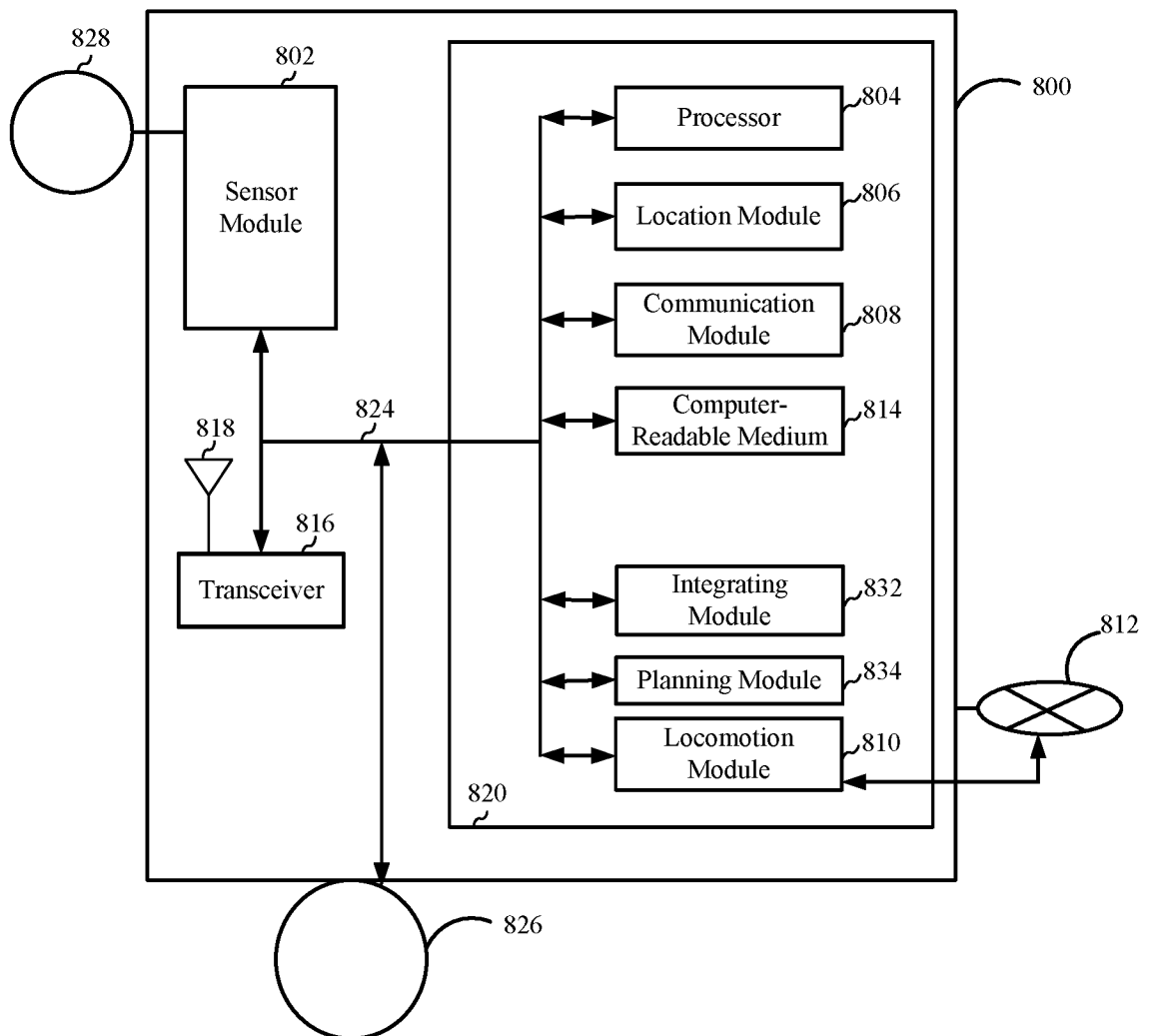


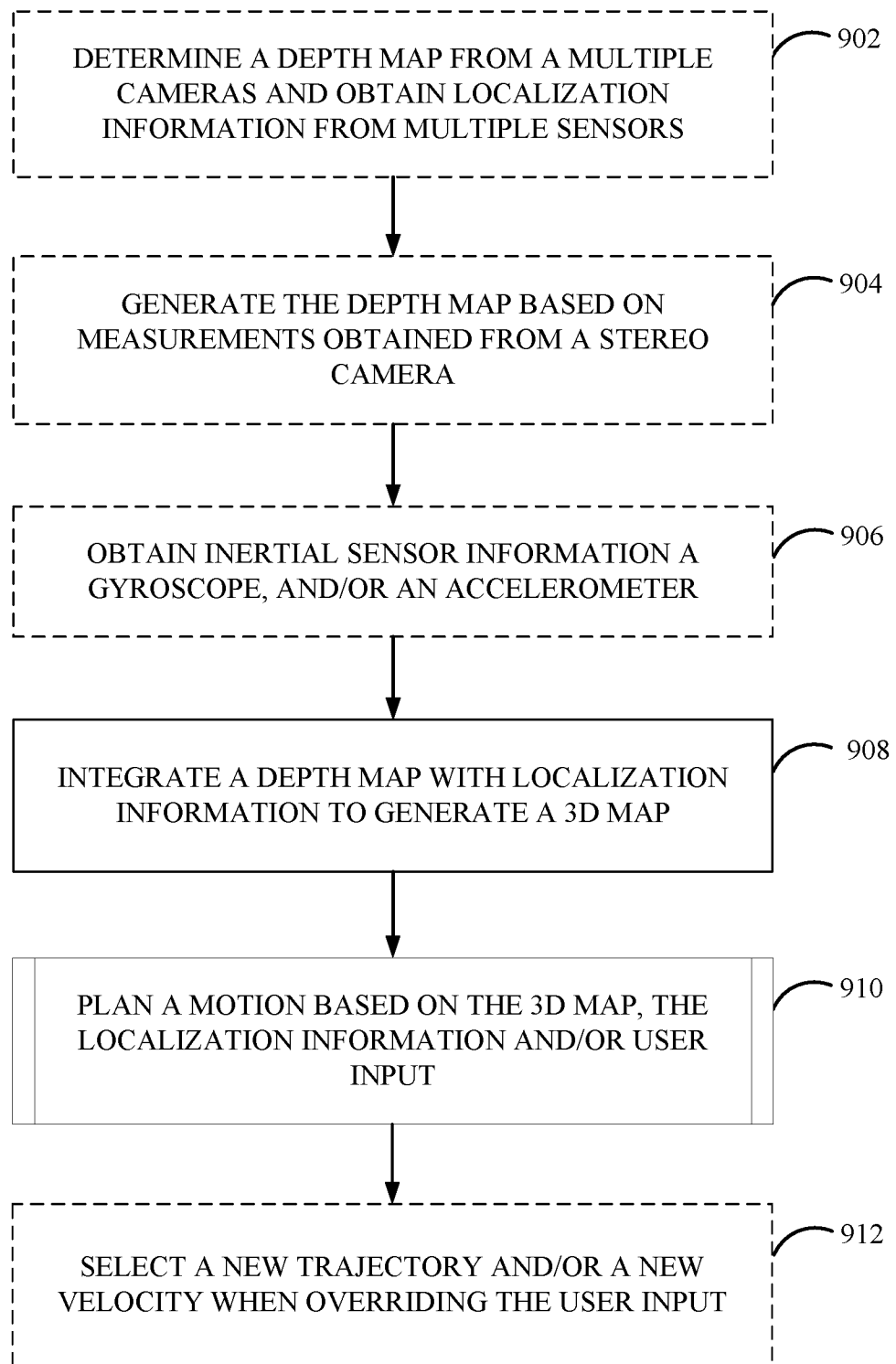
FIG. 7

12/14

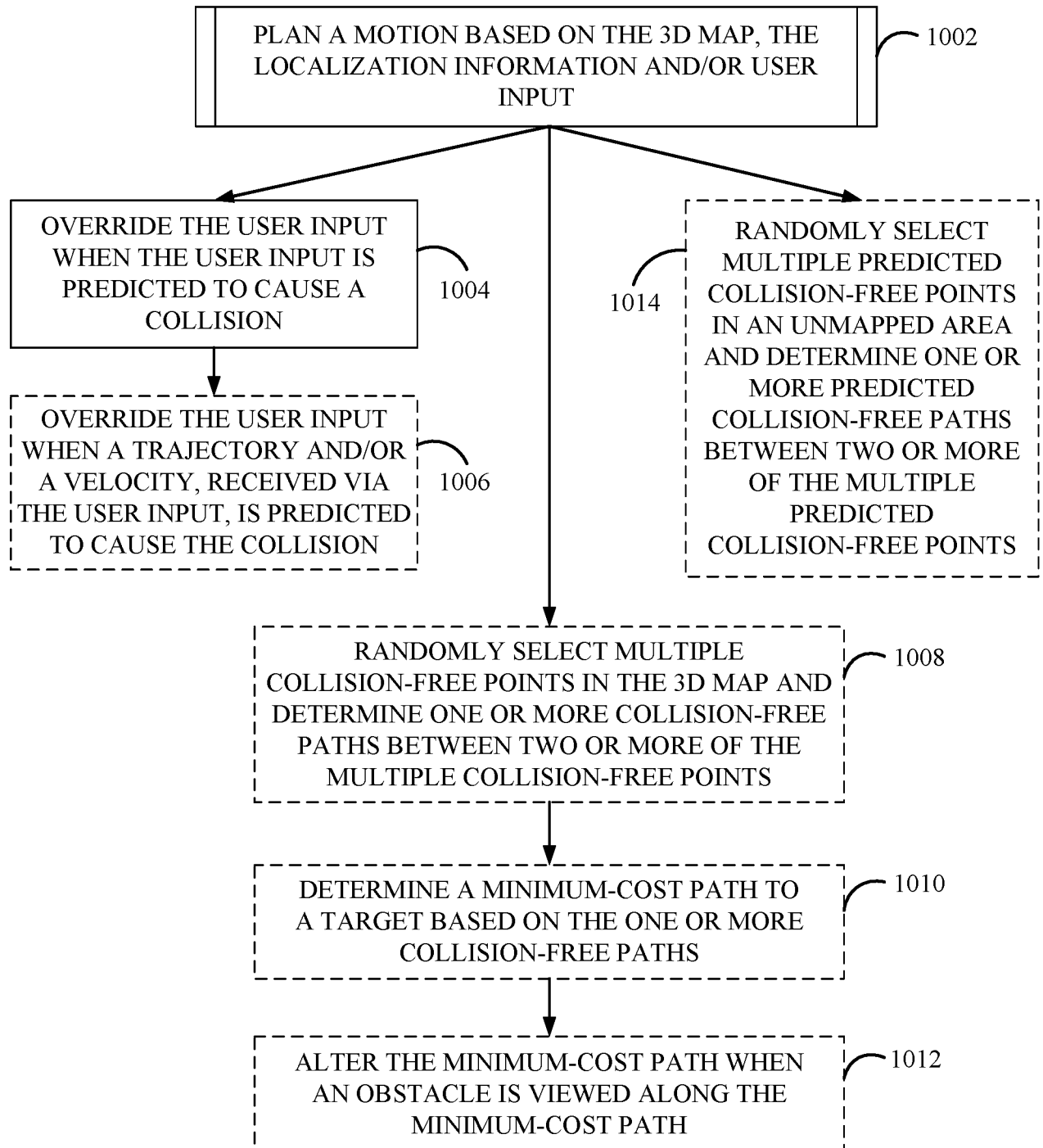
**FIG. 8**

13/14

900

**FIG. 9**

14/14

**FIG. 10**

# INTERNATIONAL SEARCH REPORT

International application No  
PCT/US2016/059171

A. CLASSIFICATION OF SUBJECT MATTER  
INV. G05D1/02  
ADD.

According to International Patent Classification (IPC) or to both national classification and IPC

## B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)  
G05D

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

EPO-Internal

## C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	US 2012/253582 A1 (CHRYSTHAKOPOULOS GEORGIOS [US] ET AL) 4 October 2012 (2012-10-04) paragraphs [0006], [0009], [0040] -----	1-40
A	US 2014/320593 A1 (PIRCHHEIM CHRISTIAN [AT] ET AL) 30 October 2014 (2014-10-30) paragraph [0005] -----	1-40
A	US 2010/063663 A1 (TOLSTEDT JONATHAN LOUIS [US] ET AL) 11 March 2010 (2010-03-11) paragraphs [0029], [0037] -----	1-40



Further documents are listed in the continuation of Box C.



See patent family annex.

\* Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search

17 January 2017

Date of mailing of the international search report

24/01/2017

Name and mailing address of the ISA/

European Patent Office, P.B. 5818 Patentlaan 2  
NL - 2280 HV Rijswijk  
Tel. (+31-70) 340-2040,  
Fax: (+31-70) 340-3016

Authorized officer

Polednicek, Milos



# INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No

PCT/US2016/059171

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
US 2012253582	A1	04-10-2012	NONE
-----			
US 2014320593	A1	30-10-2014	CN 105210113 A 30-12-2015
			EP 2992505 A1 09-03-2016
			JP 2016526313 A 01-09-2016
			KR 20160003066 A 08-01-2016
			US 2014320593 A1 30-10-2014
			WO 2014179349 A1 06-11-2014
-----			
US 2010063663	A1	11-03-2010	EP 2169506 A2 31-03-2010
			US 2010063663 A1 11-03-2010
-----			