

US Patent & Trademark Office

Patent Public Search | Text View

United States Patent Application Publication

20250265730

Kind Code

A1

Publication Date

August 21, 2025

Inventor(s)

Shoeb; Ali

Machine-Learned Monocular Depth Estimation and Semantic Segmentation for 6-DOF Absolute Localization of a Delivery Drone

Abstract

A method includes receiving a two-dimensional (2D) image captured by a camera on a unmanned aerial vehicle (UAV) and representative of an environment of the UAV. The method further includes applying a trained machine learning model to the 2D image to produce a semantic image of the environment and a depth image of the environment, where the semantic image comprises one or more semantic labels. The method additionally includes retrieving reference depth data representative of the environment, wherein the reference depth data includes reference semantic labels. The method also includes aligning the depth image of the environment with the reference depth data representative of the environment to determine a location of the UAV in the environment, where the aligning associates the one or more semantic labels from the semantic image with the reference semantic labels from the reference depth data.

Inventors: Shoeb; Ali (Palo Alto, CA)

Applicant: Wing Aviation LLC (Palo Alto, CA)

Family ID: 1000008578168

Appl. No.: 19/200000

Filed: May 06, 2025

Related U.S. Application Data

parent US continuation 17813188 20220718 parent-grant-document US 12307710 child US 19200000

Publication Classification

Int. Cl.: G06T7/73 (20170101); B64U50/19 (20230101); B64U101/30 (20230101); B64U101/60 (20230101); G01S19/48 (20100101); G06T7/50 (20170101)

U.S. Cl.:

CPC **G06T7/74** (20170101); **G01S19/485** (20200501); **G06T7/50** (20170101); B64U50/19 (20230101); B64U2101/30 (20230101); B64U2101/60 (20230101); G06T2207/10032 (20130101); G06T2207/20081 (20130101)

Background/Summary

CROSS-REFERENCE TO RELATED DISCLOSURE [0001] This application is a continuation of U.S. patent application Ser. No. 17/813,188, filed on Jul. 18, 2022, which is incorporated herein by reference in its entirety.

BACKGROUND

[0002] An unmanned vehicle, which may also be referred to as an autonomous vehicle, is a vehicle capable of travel without a physically-present human operator. An unmanned vehicle may operate in a remote-control mode, in an autonomous mode, or in a partially autonomous mode.

[0003] When an unmanned vehicle operates in a remote-control mode, a pilot or driver that is at a remote location can control the unmanned vehicle via commands that are sent to the unmanned vehicle via a wireless link. When the unmanned vehicle operates in autonomous mode, the unmanned vehicle typically moves based on pre-programmed navigation waypoints, dynamic automation systems, or a combination of these. Further, some unmanned vehicles can operate in both a remote-control mode and an autonomous mode, and in some instances may do so simultaneously. For instance, a remote pilot or driver may wish to leave navigation to an autonomous system while manually performing another task, such as operating a mechanical system for picking up objects, as an example.

[0004] Various types of unmanned vehicles exist for various different environments. For instance, unmanned vehicles exist for operation in the air, on the ground, underwater, and in space. Examples include quad-copters and tail-sitter UAVs, among others. Unmanned vehicles also exist for hybrid operations in which multi-environment operation is possible. Examples of hybrid unmanned vehicles include an amphibious craft that is capable of operation on land as well as on water or a floatplane that is capable of landing on water as well as on land. Other examples are also possible.

SUMMARY

[0005] The present embodiments are directed to systems and methods that may help to localize an unmanned aerial vehicle within an environment. A machine learning model may be trained to process two-dimensional (2D) images of a terrain within the environment. The model may include separate depth and semantics branches to produce both depth images and semantic images. The depth and semantic information may be used in combination to localize the vehicle in the environment relative to reference depth and semantic data.

[0006] In one aspect, a method is provided including receiving a two-dimensional (2D) image captured by a camera on a unmanned aerial vehicle (UAV) and representative of an environment of the UAV. The method further includes applying a trained machine learning model to the 2D image to produce a semantic image of the environment and a depth image of the environment, where the machine learning model has been trained with a semantics branch to produce the semantic image and a depth branch to produce the depth image, and where the semantic image comprises one or more semantic labels. The method also includes retrieving reference depth data representative of the environment, where the reference depth data includes reference semantic labels. The method additionally includes aligning the depth image of the environment with the reference depth data representative of the environment to determine a location of the UAV in the environment, where

the aligning associates the one or more semantic labels from the semantic image with the reference semantic labels from the reference depth data.

[0007] In another aspect an unmanned aerial vehicle is described, including a camera and a control system configured to receive a two-dimensional (2D) image captured by the camera and representative of an environment of the UAV; apply a trained machine learning model to the 2D image to produce a semantic image of the environment and a depth image of the environment, where the machine learning model has been trained with a semantics branch to produce the semantic image and a depth branch to produce the depth image, and where the semantic image comprises one or more semantic labels; retrieve reference depth data representative of the environment, where the reference depth data includes reference semantic labels; and align the depth image of the environment with the reference depth data representative of the environment to determine a location of the UAV in the environment, where the aligning associates the one or more semantic labels from the semantic image with the reference semantic labels from the reference depth data.

[0008] In another aspect, a non-transitory computer readable medium is described including program instructions executable by one or more processors to perform operations. The operations include receiving a two-dimensional (2D) image captured by a camera on a unmanned aerial vehicle (UAV) and representative of an environment of the UAV; applying a trained machine learning model to the 2D image to produce a semantic image of the environment and a depth image of the environment, where the machine learning model has been trained with a semantics branch to produce the semantic image and a depth branch to produce the depth image, and where the semantic image comprises one or more semantic labels; retrieving reference depth data representative of the environment, where the reference depth data includes reference semantic labels; and aligning the depth image of the environment with the reference depth data representative of the environment to determine a location of the UAV in the environment, where the aligning associates the one or more semantic labels from the semantic image with the reference semantic labels from the reference depth data.

[0009] These as well as other aspects, advantages, and alternatives will become apparent to those of ordinary skill in the art by reading the following detailed description with reference where appropriate to the accompanying drawings. Further, it should be understood that the description provided in this summary section and elsewhere in this document is intended to illustrate the claimed subject matter by way of example and not by way of limitation.

Description

BRIEF DESCRIPTION OF THE DRAWINGS

[0010] FIG. 1A is a simplified illustration of an unmanned aerial vehicle, according to an example embodiment.

[0011] FIG. 1B is a simplified illustration of an unmanned aerial vehicle, according to an example embodiment.

[0012] FIG. 1C is a simplified illustration of an unmanned aerial vehicle, according to an example embodiment.

[0013] FIG. 1D is a simplified illustration of an unmanned aerial vehicle, according to an example embodiment.

[0014] FIG. 1E is a simplified illustration of an unmanned aerial vehicle, according to an example embodiment.

[0015] FIG. 2 is a simplified block diagram illustrating components of an unmanned aerial vehicle, according to an example embodiment.

[0016] FIG. 3 is a simplified block diagram illustrating a UAV system, according to an example

embodiment.

[0017] FIG. 4 is a diagram illustrating training and inference phases of a machine learning model, in accordance with example embodiments.

[0018] FIG. 5 is a diagram illustrating application of a machine learning model, in accordance with example embodiments.

[0019] FIG. 6A illustrates an image captured by a UAV, in accordance with example embodiments.

[0020] FIG. 6B illustrates reference depth data representative of an environment of a UAV, in accordance with example embodiments.

[0021] FIG. 6C illustrates semantically labeled depth data generated by a machine learning model, in accordance with example embodiments.

[0022] FIG. 6D illustrates superimposed reference depth data and semantically labeled depth data generated by a machine learning model, in accordance with example embodiments.

[0023] FIG. 6E illustrates reference depth data aligned with semantically labeled depth data generated by a machine learning model, in accordance with example embodiments.

[0024] FIG. 7 is a block diagram of a method, in accordance with example embodiments.

DETAILED DESCRIPTION

[0025] Exemplary methods and systems are described herein. It should be understood that the word “exemplary” is used herein to mean “serving as an example, instance, or illustration.” Any implementation or feature described herein as “exemplary” or “illustrative” is not necessarily to be construed as preferred or advantageous over other implementations or features. In the figures, similar symbols typically identify similar components, unless context dictates otherwise. The example implementations described herein are not meant to be limiting. It will be readily understood that the aspects of the present disclosure, as generally described herein, and illustrated in the figures, can be arranged, substituted, combined, separated, and designed in a wide variety of different configurations, all of which are contemplated herein.

I. OVERVIEW

[0026] Determining the absolute location (longitude, latitude, and altitude) of a delivery drone or other unmanned aerial vehicle (UAV) is often critical for mission success. A common way of determining absolute location is to rely on a Global Navigation Satellite System (GNSS), such as a Global Positioning System (GPS). However, in some situations, GPS can become unavailable or the receiver can produce an erroneous location measurement. Consequently, the ability to derive an absolute location of a UAV using a different modality may serve as a complement and/or backup to GPS.

[0027] Examples described herein enable a UAV to use a machine-learned model to produce a semantically labeled three-dimensional (3D) point cloud (e.g., a 3D model of a scene). The semantically labeled 3D point can then be registered against a pre-existing point cloud. In some examples, the pre-existing point cloud may be derived from a vendor or built during a preceding visit by the UAV and/or by other UAVs. Registering the semantically labeled 3D point against the pre-existing point cloud may be performed to derive the vehicle's absolute position and orientation (e.g., to recover the vehicle's position and orientation with six degrees of freedom). In some examples, the absolute position can be used alone in the absence of GPS to complete a mission assigned to the UAV. In further examples, the absolute position can be used to cross-check GPS measurements and/or other sensor measurements. In yet further examples, the absolute position can be blended with the GPS measurement to arrive at a more certain belief of the UAV's whereabouts.

II. ILLUSTRATIVE UNMANNED VEHICLES

[0028] Herein, the terms “unmanned aerial vehicle” and “UAV” refer to any autonomous or semi-autonomous vehicle that is capable of performing some functions without a physically present human pilot.

[0029] A UAV can take various forms. For example, a UAV may take the form of a fixed-wing

aircraft, a glider aircraft, a tail-sitter aircraft, a jet aircraft, a ducted fan aircraft, a lighter-than-air dirigible such as a blimp or steerable balloon, a rotorcraft such as a helicopter or multicopter, and/or an ornithopter, among other possibilities. Further, the terms “drone,” “unmanned aerial vehicle system” (UAVS), or “unmanned aerial system” (UAS) may also be used to refer to a UAV. [0030] FIG. 1A is an isometric view of an example UAV **100**. UAV **100** includes wing **102**, booms **104**, and a fuselage **106**. Wings **102** may be stationary and may generate lift based on the wing shape and the UAV's forward airspeed. For instance, the two wings **102** may have an airfoil-shaped cross section to produce an aerodynamic force on UAV **100**. In some embodiments, wing **102** may carry horizontal propulsion units **108**, and booms **104** may carry vertical propulsion units **110**. In operation, power for the propulsion units may be provided from a battery compartment **112** of fuselage **106**. In some embodiments, fuselage **106** also includes an avionics compartment **114**, an additional battery compartment (not shown) and/or a delivery unit (not shown, e.g., a winch system) for handling the payload. In some embodiments, fuselage **106** is modular, and two or more compartments (e.g., battery compartment **112**, avionics compartment **114**, other payload and delivery compartments) are detachable from each other and securable to each other (e.g., mechanically, magnetically, or otherwise) to contiguously form at least a portion of fuselage **106**. [0031] In some embodiments, booms **104** terminate in rudders **116** for improved yaw control of UAV **100**. Further, wings **102** may terminate in wing tips **117** for improved control of lift of the UAV.

[0032] In the illustrated configuration, UAV **100** includes a structural frame. The structural frame may be referred to as a “structural H-frame” or an “H-frame” (not shown) of the UAV. The H-frame may include, within wings **102**, a wing spar (not shown) and, within booms **104**, boom carriers (not shown). In some embodiments the wing spar and the boom carriers may be made of carbon fiber, hard plastic, aluminum, light metal alloys, or other materials. The wing spar and the boom carriers may be connected with clamps. The wing spar may include pre-drilled holes for horizontal propulsion units **108**, and the boom carriers may include pre-drilled holes for vertical propulsion units **110**.

[0033] In some embodiments, fuselage **106** may be removably attached to the H-frame (e.g., attached to the wing spar by clamps, configured with grooves, protrusions or other features to mate with corresponding H-frame features, etc.). In other embodiments, fuselage **106** similarly may be removably attached to wings **102**. The removable attachment of fuselage **106** may improve quality and or modularity of UAV **100**. For example, electrical/mechanical components and/or subsystems of fuselage **106** may be tested separately from, and before being attached to, the H-frame. Similarly, printed circuit boards (PCBs) **118** may be tested separately from, and before being attached to, the boom carriers, therefore eliminating defective parts/subassemblies prior to completing the UAV. For example, components of fuselage **106** (e.g., avionics, battery unit, delivery units, an additional battery compartment, etc.) may be electrically tested before fuselage **106** is mounted to the H-frame. Furthermore, the motors and the electronics of PCBs **118** may also be electrically tested before the final assembly. Generally, the identification of the defective parts and subassemblies early in the assembly process lowers the overall cost and lead time of the UAV. Furthermore, different types/models of fuselage **106** may be attached to the H-frame, therefore improving the modularity of the design. Such modularity allows these various parts of UAV **100** to be upgraded without a substantial overhaul to the manufacturing process.

[0034] In some embodiments, a wing shell and boom shells may be attached to the H-frame by adhesive elements (e.g., adhesive tape, double-sided adhesive tape, glue, etc.). Therefore, multiple shells may be attached to the H-frame instead of having a monolithic body sprayed onto the H-frame. In some embodiments, the presence of the multiple shells reduces the stresses induced by the coefficient of thermal expansion of the structural frame of the UAV. As a result, the UAV may have better dimensional accuracy and/or improved reliability.

[0035] Moreover, in at least some embodiments, the same H-frame may be used with the wing

shell and/or boom shells having different size and/or design, therefore improving the modularity and versatility of the UAV designs. The wing shell and/or the boom shells may be made of relatively light polymers (e.g., closed cell foam) covered by the harder, but relatively thin, plastic skins.

[0036] The power and/or control signals from fuselage **106** may be routed to PCBs **118** through cables running through fuselage **106**, wings **102**, and booms **104**. In the illustrated embodiment, UAV **100** has four PCBs, but other numbers of PCBs are also possible. For example, UAV **100** may include two PCBs, one per the boom. The PCBs carry electronic components **119** including, for example, power converters, controllers, memory, passive components, etc. In operation, propulsion units **108** and **110** of UAV **100** are electrically connected to the PCBs.

[0037] Many variations on the illustrated UAV are possible. For instance, fixed-wing UAVs may include more or fewer rotor units (vertical or horizontal), and/or may utilize a ducted fan or multiple ducted fans for propulsion. Further, UAVs with more wings (e.g., an “x-wing” configuration with four wings), are also possible. Although FIG. **1** illustrates two wings **102**, two booms **104**, two horizontal propulsion units **108**, and six vertical propulsion units **110** per boom **104**, it should be appreciated that other variants of UAV **100** may be implemented with more or less of these components. For example, UAV **100** may include four wings **102**, four booms **104**, and more or less propulsion units (horizontal or vertical).

[0038] Similarly, FIG. **1B** shows another example of a fixed-wing UAV **120**. The fixed-wing UAV **120** includes a fuselage **122**, two wings **124** with an airfoil-shaped cross section to provide lift for the UAV **120**, a vertical stabilizer **126** (or fin) to stabilize the plane's yaw (turn left or right), a horizontal stabilizer **128** (also referred to as an elevator or tailplane) to stabilize pitch (tilt up or down), landing gear **130**, and a propulsion unit **132**, which can include a motor, shaft, and propeller.

[0039] FIG. **1C** shows an example of a UAV **140** with a propeller in a pusher configuration. The term “pusher” refers to the fact that a propulsion unit **142** is mounted at the back of the UAV and “pushes” the vehicle forward, in contrast to the propulsion unit being mounted at the front of the UAV. Similar to the description provided for FIGS. **1A** and **1B**, FIG. **1C** depicts common structures used in a pusher plane, including a fuselage **144**, two wings **146**, vertical stabilizers **148**, and the propulsion unit **142**, which can include a motor, shaft, and propeller.

[0040] FIG. **1D** shows an example of a tail-sitter UAV **160**. In the illustrated example, the tail-sitter UAV **160** has fixed wings **162** to provide lift and allow the UAV **160** to glide horizontally (e.g., along the x-axis, in a position that is approximately perpendicular to the position shown in FIG. **1D**). However, the fixed wings **162** also allow the tail-sitter UAV **160** to take off and land vertically on its own.

[0041] For example, at a launch site, the tail-sitter UAV **160** may be positioned vertically (as shown) with its fins **164** and/or wings **162** resting on the ground and stabilizing the UAV **160** in the vertical position. The tail-sitter UAV **160** may then take off by operating its propellers **166** to generate an upward thrust (e.g., a thrust that is generally along the y-axis). Once at a suitable altitude, the tail-sitter UAV **160** may use its flaps **168** to reorient itself in a horizontal position, such that its fuselage **170** is closer to being aligned with the x-axis than the y-axis. Positioned horizontally, the propellers **166** may provide forward thrust so that the tail-sitter UAV **160** can fly in a similar manner as a typical airplane.

[0042] Many variations on the illustrated fixed-wing UAVs are possible. For instance, fixed-wing UAVs may include more or fewer propellers, and/or may utilize a ducted fan or multiple ducted fans for propulsion. Further, UAVs with more wings (e.g., an “x-wing” configuration with four wings), with fewer wings, or even with no wings, are also possible.

[0043] As noted above, some embodiments may involve other types of UAVs, in addition to or in the alternative to fixed-wing UAVs. For instance, FIG. **1E** shows an example of a rotorcraft that is commonly referred to as a multicopter **180**. The multicopter **180** may also be referred to as a

quadcopter, as it includes four rotors **182**. It should be understood that example embodiments may involve a rotorcraft with more or fewer rotors than the multicopter **180**. For example, a helicopter typically has two rotors. Other examples with three or more rotors are possible as well. Herein, the term “multicopter” refers to any rotorcraft having more than two rotors, and the term “helicopter” refers to rotorcraft having two rotors.

[0044] Referring to the multicopter **180** in greater detail, the four rotors **182** provide propulsion and maneuverability for the multicopter **180**. More specifically, each rotor **182** includes blades that are attached to a motor **184**. Configured as such, the rotors **182** may allow the multicopter **180** to take off and land vertically, to maneuver in any direction, and/or to hover. Further, the pitch of the blades may be adjusted as a group and/or differentially, and may allow the multicopter **180** to control its pitch, roll, yaw, and/or altitude.

[0045] It should be understood that references herein to an “unmanned” aerial vehicle or UAV can apply equally to autonomous and semi-autonomous aerial vehicles. In an autonomous implementation, all functionality of the aerial vehicle is automated; e.g., pre-programmed or controlled via real-time computer functionality that responds to input from various sensors and/or pre-determined information. In a semi-autonomous implementation, some functions of an aerial vehicle may be controlled by a human operator, while other functions are carried out autonomously. Further, in some embodiments, a UAV may be configured to allow a remote operator to take over functions that can otherwise be controlled autonomously by the UAV. Yet further, a given type of function may be controlled remotely at one level of abstraction and performed autonomously at another level of abstraction. For example, a remote operator could control high level navigation decisions for a UAV, such as by specifying that the UAV should travel from one location to another (e.g., from a warehouse in a suburban area to a delivery address in a nearby city), while the UAV's navigation system autonomously controls more fine-grained navigation decisions, such as the specific route to take between the two locations, specific flight controls to achieve the route and avoid obstacles while navigating the route, and so on.

[0046] More generally, it should be understood that the example UAVs described herein are not intended to be limiting. Example embodiments may relate to, be implemented within, or take the form of any type of unmanned aerial vehicle.

III. ILLUSTRATIVE UAV COMPONENTS

[0047] FIG. 2 is a simplified block diagram illustrating components of a UAV **200**, according to an example embodiment. UAV **200** may take the form of, or be similar in form to, one of the UAVs **100**, **120**, **140**, **160**, and **180** described in reference to FIGS. 1A-1E. However, UAV **200** may also take other forms.

[0048] UAV **200** may include various types of sensors, and may include a computing system configured to provide the functionality described herein. In the illustrated embodiment, the sensors of UAV **200** include an inertial measurement unit (IMU) **202**, ultrasonic sensor(s) **204**, and a GPS **206**, among other possible sensors and sensing systems.

[0049] In the illustrated embodiment, UAV **200** also includes one or more processors **208**. A processor **208** may be a general-purpose processor or a special purpose processor (e.g., digital signal processors, application specific integrated circuits, etc.). The one or more processors **208** can be configured to execute computer-readable program instructions **212** that are stored in the data storage **210** and are executable to provide the functionality of a UAV described herein.

[0050] The data storage **210** may include or take the form of one or more computer-readable storage media that can be read or accessed by at least one processor **208**. The one or more computer-readable storage media can include volatile and/or non-volatile storage components, such as optical, magnetic, organic or other memory or disc storage, which can be integrated in whole or in part with at least one of the one or more processors **208**. In some embodiments, the data storage **210** can be implemented using a single physical device (e.g., one optical, magnetic, organic or other memory or disc storage unit), while in other embodiments, the data storage **210** can be

implemented using two or more physical devices.

[0051] As noted, the data storage **210** can include computer-readable program instructions **212** and perhaps additional data, such as diagnostic data of the UAV **200**. As such, the data storage **210** may include program instructions **212** to perform or facilitate some or all of the UAV functionality described herein. For instance, in the illustrated embodiment, program instructions **212** include a navigation module **214** and a tether control module **216**.

A. Sensors

[0052] In an illustrative embodiment, IMU **202** may include both an accelerometer and a gyroscope, which may be used together to determine an orientation of the UAV **200**. In particular, the accelerometer can measure the orientation of the vehicle with respect to earth, while the gyroscope measures the rate of rotation around an axis. IMUs are commercially available in low-cost, low-power packages. For instance, an IMU **202** may take the form of or include a miniaturized MicroElectroMechanical System (MEMS) or a NanoElectroMechanical System (NEMS). Other types of IMUs may also be utilized.

[0053] An IMU **202** may include other sensors, in addition to accelerometers and gyroscopes, which may help to better determine position and/or help to increase autonomy of the UAV **200**. Two examples of such sensors are magnetometers and pressure sensors. In some embodiments, a UAV may include a low-power, digital 3-axis magnetometer, which can be used to realize an orientation independent electronic compass for accurate heading information. However, other types of magnetometers may be utilized as well. Other examples are also possible. Further, note that a UAV could include some or all of the above-described inertia sensors as separate components from an IMU.

[0054] UAV **200** may also include a pressure sensor or barometer, which can be used to determine the altitude of the UAV **200**. Alternatively, other sensors, such as sonic altimeters or radar altimeters, can be used to provide an indication of altitude, which may help to improve the accuracy of and/or prevent drift of an IMU.

[0055] In a further aspect, UAV **200** may include one or more sensors that allow the UAV to sense objects in the environment. For instance, in the illustrated embodiment, UAV **200** includes ultrasonic sensor(s) **204**. Ultrasonic sensor(s) **204** can determine the distance to an object by generating sound waves and determining the time interval between transmission of the wave and receiving the corresponding echo off an object. A typical application of an ultrasonic sensor for unmanned vehicles or IMUs is low-level altitude control and obstacle avoidance. An ultrasonic sensor can also be used for vehicles that need to hover at a certain height or need to be capable of detecting obstacles. Other systems can be used to determine, sense the presence of, and/or determine the distance to nearby objects, such as a light detection and ranging (LIDAR) system, laser detection and ranging (LADAR) system, and/or an infrared or forward-looking infrared (FLIR) system, among other possibilities.

[0056] In some embodiments, UAV **200** may also include one or more imaging system(s). For example, one or more still and/or video cameras may be utilized by UAV **200** to capture image data from the UAV's environment. As a specific example, charge-coupled device (CCD) cameras or complementary metal-oxide-semiconductor (CMOS) cameras can be used with unmanned vehicles. Such imaging sensor(s) have numerous possible applications, such as obstacle avoidance, localization techniques, ground tracking for more accurate navigation (e.g., by applying optical flow techniques to images), video feedback, and/or image recognition and processing, among other possibilities.

[0057] UAV **200** may also include a GPS receiver **206**. The GPS receiver **206** may be configured to provide data that is typical of well-known GPS systems, such as the GPS coordinates of the UAV **200**. Such GPS data may be utilized by the UAV **200** for various functions. As such, the UAV may use its GPS receiver **206** to help navigate to the caller's location, as indicated, at least in part, by the GPS coordinates provided by their mobile device. Other examples are also possible.

B. Navigation and Location Determination

[0058] The navigation module **214** may provide functionality that allows the UAV **200** to, e.g., move about its environment and reach a desired location. To do so, the navigation module **214** may control the altitude and/or direction of flight by controlling the mechanical features of the UAV that affect flight (e.g., its rudder(s), elevator(s), aileron(s), and/or the speed of its propeller(s)).

[0059] In order to navigate the UAV **200** to a target location, the navigation module **214** may implement various navigation techniques, such as map-based navigation and localization-based navigation, for instance. With map-based navigation, the UAV **200** may be provided with a map of its environment, which may then be used to navigate to a particular location on the map. With localization-based navigation, the UAV **200** may be capable of navigating in an unknown environment using localization. Localization-based navigation may involve the UAV **200** building its own map of its environment and calculating its position within the map and/or the position of objects in the environment. For example, as a UAV **200** moves throughout its environment, the UAV **200** may continuously use localization to update its map of the environment. This continuous mapping process may be referred to as simultaneous localization and mapping (SLAM). Other navigation techniques may also be utilized.

[0060] In some embodiments, the navigation module **214** may navigate using a technique that relies on waypoints. In particular, waypoints are sets of coordinates that identify points in physical space. For instance, an air-navigation waypoint may be defined by a certain latitude, longitude, and altitude. Accordingly, navigation module **214** may cause UAV **200** to move from waypoint to waypoint, in order to ultimately travel to a final destination (e.g., a final waypoint in a sequence of waypoints).

[0061] In a further aspect, the navigation module **214** and/or other components and systems of the UAV **200** may be configured for “localization” to more precisely navigate to the scene of a target location. More specifically, it may be desirable in certain situations for a UAV to be within a threshold distance of the target location where a payload **228** is being delivered by a UAV (e.g., within a few feet of the target destination). To this end, a UAV may use a two-tiered approach in which it uses a more-general location-determination technique to navigate to a general area that is associated with the target location, and then use a more-refined location-determination technique to identify and/or navigate to the target location within the general area.

[0062] For example, the UAV **200** may navigate to the general area of a target destination where a payload **228** is being delivered using waypoints and/or map-based navigation. The UAV may then switch to a mode in which it utilizes a localization process to locate and travel to a more specific location. For instance, if the UAV **200** is to deliver a payload to a user's home, the UAV **200** may need to be substantially close to the target location in order to avoid delivery of the payload to undesired areas (e.g., onto a roof, into a pool, onto a neighbor's property, etc.). However, a GPS signal may only get the UAV **200** so far (e.g., within a block of the user's home). A more precise location-determination technique may then be used to find the specific target location.

[0063] Various types of location-determination techniques may be used to accomplish localization of the target delivery location once the UAV **200** has navigated to the general area of the target delivery location. For instance, the UAV **200** may be equipped with one or more sensory systems, such as, for example, ultrasonic sensors **204**, infrared sensors (not shown), and/or other sensors, which may provide input that the navigation module **214** utilizes to navigate autonomously or semi-autonomously to the specific target location.

[0064] As another example, once the UAV **200** reaches the general area of the target delivery location (or of a moving subject such as a person or their mobile device), the UAV **200** may switch to a “fly-by-wire” mode where it is controlled, at least in part, by a remote operator, who can navigate the UAV **200** to the specific target location. To this end, sensory data from the UAV **200** may be sent to the remote operator to assist them in navigating the UAV **200** to the specific location.

[0065] As yet another example, the UAV **200** may include a module that is able to signal to a passer-by for assistance in either reaching the specific target delivery location; for example, the UAV **200** may display a visual message requesting such assistance in a graphic display, play an audio message or tone through speakers to indicate the need for such assistance, among other possibilities. Such a visual or audio message might indicate that assistance is needed in delivering the UAV **200** to a particular person or a particular location, and might provide information to assist the passer-by in delivering the UAV **200** to the person or location (e.g., a description or picture of the person or location, and/or the person or location's name), among other possibilities. Such a feature can be useful in a scenario in which the UAV is unable to use sensory functions or another location-determination technique to reach the specific target location. However, this feature is not limited to such scenarios.

[0066] In some embodiments, once the UAV **200** arrives at the general area of a target delivery location, the UAV **200** may utilize a beacon from a user's remote device (e.g., the user's mobile phone) to locate the person. Such a beacon may take various forms. As an example, consider the scenario where a remote device, such as the mobile phone of a person who requested a UAV delivery, is able to send out directional signals (e.g., via an RF signal, a light signal and/or an audio signal). In this scenario, the UAV **200** may be configured to navigate by “sourcing” such directional signals—in other words, by determining where the signal is strongest and navigating accordingly. As another example, a mobile device can emit a frequency, either in the human range or outside the human range, and the UAV **200** can listen for that frequency and navigate accordingly. As a related example, if the UAV **200** is listening for spoken commands, then the UAV **200** could utilize spoken statements, such as “I'm over here!” to source the specific location of the person requesting delivery of a payload.

[0067] In an alternative arrangement, a navigation module may be implemented at a remote computing device, which communicates wirelessly with the UAV **200**. The remote computing device may receive data indicating the operational state of the UAV **200**, sensor data from the UAV **200** that allows it to assess the environmental conditions being experienced by the UAV **200**, and/or location information for the UAV **200**. Provided with such information, the remote computing device may determine latitudinal and/or directional adjustments that should be made by the UAV **200** and/or may determine how the UAV **200** should adjust its mechanical features (e.g., its rudder(s), elevator(s), aileron(s), and/or the speed of its propeller(s)) in order to effectuate such movements. The remote computing system may then communicate such adjustments to the UAV **200** so it can move in the determined manner.

C. Communication Systems

[0068] In a further aspect, the UAV **200** includes one or more communication systems **218**. The communications systems **218** may include one or more wireless interfaces and/or one or more wireline interfaces, which allow the UAV **200** to communicate via one or more networks. Such wireless interfaces may provide for communication under one or more wireless communication protocols, such as Bluetooth, WiFi (e.g., an IEEE 802.11 protocol), Long-Term Evolution (LTE), WiMAX (e.g., an IEEE 802.16 standard), a radio-frequency ID (RFID) protocol, near-field communication (NFC), and/or other wireless communication protocols. Such wireline interfaces may include an Ethernet interface, a Universal Serial Bus (USB) interface, or similar interface to communicate via a wire, a twisted pair of wires, a coaxial cable, an optical link, a fiber-optic link, or other physical connection to a wireline network.

[0069] In some embodiments, a UAV **200** may include communication systems **218** that allow for both short-range communication and long-range communication. For example, the UAV **200** may be configured for short-range communications using Bluetooth and for long-range communications under a CDMA protocol. In such an embodiment, the UAV **200** may be configured to function as a “hot spot;” or in other words, as a gateway or proxy between a remote support device and one or more data networks, such as a cellular network and/or the Internet. Configured as such, the UAV

200 may facilitate data communications that the remote support device would otherwise be unable to perform by itself.

[0070] For example, the UAV **200** may provide a WiFi connection to a remote device, and serve as a proxy or gateway to a cellular service provider's data network, which the UAV might connect to under an LTE or a 3G protocol, for instance. The UAV **200** could also serve as a proxy or gateway to a high-altitude balloon network, a satellite network, or a combination of these networks, among others, which a remote device might not be able to otherwise access.

D. Power Systems

[0071] In a further aspect, the UAV **200** may include power system(s) **220**. The power system **220** may include one or more batteries for providing power to the UAV **200**. In one example, the one or more batteries may be rechargeable and each battery may be recharged via a wired connection between the battery and a power supply and/or via a wireless charging system, such as an inductive charging system that applies an external time-varying magnetic field to an internal battery.

E. Payload Delivery

[0072] The UAV **200** may employ various systems and configurations in order to transport and deliver a payload **228**. In some implementations, the payload **228** of a given UAV **200** may include or take the form of a “package” designed to transport various goods to a target delivery location. For example, the UAV **200** can include a compartment, in which an item or items may be transported. Such a package may include one or more food items, purchased goods, medical items, or any other object(s) having a size and weight suitable to be transported between two locations by the UAV. In other embodiments, a payload **228** may simply be the one or more items that are being delivered (e.g., without any package housing the items).

[0073] In some embodiments, the payload **228** may be attached to the UAV and located substantially outside of the UAV during some or all of a flight by the UAV. For example, the package may be tethered or otherwise releasably attached below the UAV during flight to a target location. In some embodiments, the package may include various features that protect its contents from the environment, reduce aerodynamic drag on the system, and prevent the contents of the package from shifting during UAV flight. In other embodiments, the package may be a standard shipping package that is not specifically tailored for UAV flight.

[0074] In order to deliver the payload, the UAV may include a winch system **221** controlled by the tether control module **216** in order to lower the payload **228** to the ground while the UAV hovers above. As shown in FIG. 2, the winch system **221** may include a tether **224**, and the tether **224** may be coupled to the payload **228** by a payload retriever **226**. The tether **224** may be wound on a spool that is coupled to a motor **222** of the UAV. The motor **222** may take the form of a DC motor (e.g., a servo motor) that can be actively controlled by a speed controller. The tether control module **216** can control the speed controller to cause the motor **222** to rotate the spool, thereby unwinding or retracting the tether **224** and lowering or raising the payload retriever **226**. In practice, the speed controller may output a desired operating rate (e.g., a desired RPM) for the spool, which may correspond to the speed at which the tether **224** and payload **228** should be lowered towards the ground. The motor **222** may then rotate the spool so that it maintains the desired operating rate.

[0075] In order to control the motor **222** via the speed controller, the tether control module **216** may receive data from a speed sensor (e.g., an encoder) configured to convert a mechanical position to a representative analog or digital signal. In particular, the speed sensor may include a rotary encoder that may provide information related to rotary position (and/or rotary movement) of a shaft of the motor or the spool coupled to the motor, among other possibilities. Moreover, the speed sensor may take the form of an absolute encoder and/or an incremental encoder, among others. So in an example implementation, as the motor **222** causes rotation of the spool, a rotary encoder may be used to measure this rotation. In doing so, the rotary encoder may be used to convert a rotary position to an analog or digital electronic signal used by the tether control module **216** to determine the amount of rotation of the spool from a fixed reference angle and/or to an

analog or digital electronic signal that is representative of a new rotary position, among other options. Other examples are also possible.

[0076] Based on the data from the speed sensor, the tether control module **216** may determine a rotational speed of the motor **222** and/or the spool and responsively control the motor **222** (e.g., by increasing or decreasing an electrical current supplied to the motor **222**) to cause the rotational speed of the motor **222** to match a desired speed. When adjusting the motor current, the magnitude of the current adjustment may be based on a proportional-integral-derivative (PID) calculation using the determined and desired speeds of the motor **222**. For instance, the magnitude of the current adjustment may be based on a present difference, a past difference (based on accumulated error over time), and a future difference (based on current rates of change) between the determined and desired speeds of the spool.

[0077] In some embodiments, the tether control module **216** may vary the rate at which the tether **224** and payload **228** are lowered to the ground. For example, the speed controller may change the desired operating rate according to a variable deployment-rate profile and/or in response to other factors in order to change the rate at which the payload **228** descends toward the ground. To do so, the tether control module **216** may adjust an amount of braking or an amount of friction that is applied to the tether **224**. For example, to vary the tether deployment rate, the UAV **200** may include friction pads that can apply a variable amount of pressure to the tether **224**. As another example, the UAV **200** can include a motorized braking system that varies the rate at which the spool lets out the tether **224**. Such a braking system may take the form of an electromechanical system in which the motor **222** operates to slow the rate at which the spool lets out the tether **224**. Further, the motor **222** may vary the amount by which it adjusts the speed (e.g., the RPM) of the spool, and thus may vary the deployment rate of the tether **224**. Other examples are also possible.

[0078] In some embodiments, the tether control module **216** may be configured to limit the motor current supplied to the motor **222** to a maximum value. With such a limit placed on the motor current, there may be situations where the motor **222** cannot operate at the desired operation specified by the speed controller. For instance, as discussed in more detail below, there may be situations where the speed controller specifies a desired operating rate at which the motor **222** should retract the tether **224** toward the UAV **200**, but the motor current may be limited such that a large enough downward force on the tether **224** would counteract the retracting force of the motor **222** and cause the tether **224** to unwind instead. And as further discussed below, a limit on the motor current may be imposed and/or altered depending on an operational state of the UAV **200**.

[0079] In some embodiments, the tether control module **216** may be configured to determine a status of the tether **224** and/or the payload **228** based on the amount of current supplied to the motor **222**. For instance, if a downward force is applied to the tether **224** (e.g., if the payload **228** is attached to the tether **224** or if the tether **224** gets snagged on an object when retracting toward the UAV **200**), the tether control module **216** may need to increase the motor current in order to cause the determined rotational speed of the motor **222** and/or spool to match the desired speed.

Similarly, when the downward force is removed from the tether **224** (e.g., upon delivery of the payload **228** or removal of a tether snag), the tether control module **216** may need to decrease the motor current in order to cause the determined rotational speed of the motor **222** and/or spool to match the desired speed. As such, the tether control module **216** may be configured to monitor the current supplied to the motor **222**. For instance, the tether control module **216** could determine the motor current based on sensor data received from a current sensor of the motor or a current sensor of the power system **220**. In any case, based on the current supplied to the motor **222**, determine if the payload **228** is attached to the tether **224**, if someone or something is pulling on the tether **224**, and/or if the payload retriever **226** is pressing against the UAV **200** after retracting the tether **224**. Other examples are possible as well.

[0080] During delivery of the payload **228**, the payload retriever **226** can be configured to secure the payload **228** while being lowered from the UAV by the tether **224**, and can be further

configured to release the payload **228** upon reaching ground level. The payload retriever **226** can then be retracted to the UAV by reeling in the tether **224** using the motor **222**.

[0081] In some implementations, the payload **228** may be passively released once it is lowered to the ground. For example, a passive release mechanism may include one or more swing arms adapted to retract into and extend from a housing. An extended swing arm may form a hook on which the payload **228** may be attached. Upon lowering the release mechanism and the payload **228** to the ground via a tether, a gravitational force as well as a downward inertial force on the release mechanism may cause the payload **228** to detach from the hook allowing the release mechanism to be raised upwards toward the UAV. The release mechanism may further include a spring mechanism that biases the swing arm to retract into the housing when there are no other external forces on the swing arm. For instance, a spring may exert a force on the swing arm that pushes or pulls the swing arm toward the housing such that the swing arm retracts into the housing once the weight of the payload **228** no longer forces the swing arm to extend from the housing. Retracting the swing arm into the housing may reduce the likelihood of the release mechanism snagging the payload **228** or other nearby objects when raising the release mechanism toward the UAV upon delivery of the payload **228**.

[0082] Active payload release mechanisms are also possible. For example, sensors such as a barometric pressure based altimeter and/or accelerometers may help to detect the position of the release mechanism (and the payload) relative to the ground. Data from the sensors can be communicated back to the UAV and/or a control system over a wireless link and used to help in determining when the release mechanism has reached ground level (e.g., by detecting a measurement with the accelerometer that is characteristic of ground impact). In other examples, the UAV may determine that the payload has reached the ground based on a weight sensor detecting a threshold low downward force on the tether and/or based on a threshold low measurement of power drawn by the winch when lowering the payload.

[0083] Other systems and techniques for delivering a payload, in addition or in the alternative to a tethered delivery system are also possible. For example, a UAV **200** could include an air-bag drop system or a parachute drop system. Alternatively, a UAV **200** carrying a payload could simply land on the ground at a delivery location. Other examples are also possible.

IV. ILLUSTRATIVE UAV DEPLOYMENT SYSTEMS

[0084] UAV systems may be implemented in order to provide various UAV-related services. In particular, UAVs may be provided at a number of different launch sites that may be in communication with regional and/or central control systems. Such a distributed UAV system may allow UAVs to be quickly deployed to provide services across a large geographic area (e.g., that is much larger than the flight range of any single UAV). For example, UAVs capable of carrying payloads may be distributed at a number of launch sites across a large geographic area (possibly even throughout an entire country, or even worldwide), in order to provide on-demand transport of various items to locations throughout the geographic area. FIG. 3 is a simplified block diagram illustrating a distributed UAV system **300**, according to an example embodiment.

[0085] In the illustrative UAV system **300**, an access system **302** may allow for interaction with, control of, and/or utilization of a network of UAVs **304**. In some embodiments, an access system **302** may be a computing system that allows for human-controlled dispatch of UAVs **304**. As such, the control system may include or otherwise provide a user interface through which a user can access and/or control the UAVs **304**.

[0086] In some embodiments, dispatch of the UAVs **304** may additionally or alternatively be accomplished via one or more automated processes. For instance, the access system **302** may dispatch one of the UAVs **304** to transport a payload to a target location, and the UAV may autonomously navigate to the target location by utilizing various on-board sensors, such as a GPS receiver and/or other various navigational sensors.

[0087] Further, the access system **302** may provide for remote operation of a UAV. For instance,

the access system **302** may allow an operator to control the flight of a UAV via its user interface. As a specific example, an operator may use the access system **302** to dispatch a UAV **304** to a target location. The UAV **304** may then autonomously navigate to the general area of the target location. At this point, the operator may use the access system **302** to take control of the UAV **304** and navigate the UAV to the target location (e.g., to a particular person to whom a payload is being transported). Other examples of remote operation of a UAV are also possible.

[0088] In an illustrative embodiment, the UAVs **304** may take various forms. For example, each of the UAVs **304** may be a UAV such as those illustrated in FIGS. **1A-1E**. However, UAV system **300** may also utilize other types of UAVs without departing from the scope of the invention. In some implementations, all of the UAVs **304** may be of the same or a similar configuration. However, in other implementations, the UAVs **304** may include a number of different types of UAVs. For instance, the UAVs **304** may include a number of types of UAVs, with each type of UAV being configured for a different type or types of payload delivery capabilities.

[0089] The UAV system **300** may further include a remote device **306**, which may take various forms. Generally, the remote device **306** may be any device through which a direct or indirect request to dispatch a UAV can be made. (Note that an indirect request may involve any communication that may be responded to by dispatching a UAV, such as requesting a package delivery). In an example embodiment, the remote device **306** may be a mobile phone, tablet computer, laptop computer, personal computer, or any network-connected computing device. Further, in some instances, the remote device **306** may not be a computing device. As an example, a standard telephone, which allows for communication via plain old telephone service (POTS), may serve as the remote device **306**. Other types of remote devices are also possible.

[0090] Further, the remote device **306** may be configured to communicate with access system **302** via one or more types of communication network(s) **308**. For example, the remote device **306** may communicate with the access system **302** (or a human operator of the access system **302**) by communicating over a POTS network, a cellular network, and/or a data network such as the Internet. Other types of networks may also be utilized.

[0091] In some embodiments, the remote device **306** may be configured to allow a user to request delivery of one or more items to a desired location. For example, a user could request UAV delivery of a package to their home via their mobile phone, tablet, or laptop. As another example, a user could request dynamic delivery to wherever they are located at the time of delivery. To provide such dynamic delivery, the UAV system **300** may receive location information (e.g., GPS coordinates, etc.) from the user's mobile phone, or any other device on the user's person, such that a UAV can navigate to the user's location (as indicated by their mobile phone).

[0092] In an illustrative arrangement, the central dispatch system **310** may be a server or group of servers, which is configured to receive dispatch messages requests and/or dispatch instructions from the access system **302**. Such dispatch messages may request or instruct the central dispatch system **310** to coordinate the deployment of UAVs to various target locations. The central dispatch system **310** may be further configured to route such requests or instructions to one or more local dispatch systems **312**. To provide such functionality, the central dispatch system **310** may communicate with the access system **302** via a data network, such as the Internet or a private network that is established for communications between access systems and automated dispatch systems.

[0093] In the illustrated configuration, the central dispatch system **310** may be configured to coordinate the dispatch of UAVs **304** from a number of different local dispatch systems **312**. As such, the central dispatch system **310** may keep track of which UAVs **304** are located at which local dispatch systems **312**, which UAVs **304** are currently available for deployment, and/or which services or operations each of the UAVs **304** is configured for (in the event that a UAV fleet includes multiple types of UAVs configured for different services and/or operations). Additionally or alternatively, each local dispatch system **312** may be configured to track which of its associated

UAVs **304** are currently available for deployment and/or are currently in the midst of item transport.

[0094] In some cases, when the central dispatch system **310** receives a request for UAV-related service (e.g., transport of an item) from the access system **302**, the central dispatch system **310** may select a specific UAV **304** to dispatch. The central dispatch system **310** may accordingly instruct the local dispatch system **312** that is associated with the selected UAV to dispatch the selected UAV. The local dispatch system **312** may then operate its associated deployment system **314** to launch the selected UAV. In other cases, the central dispatch system **310** may forward a request for a UAV-related service to a local dispatch system **312** that is near the location where the support is requested and leave the selection of a particular UAV **304** to the local dispatch system **312**.

[0095] In an example configuration, the local dispatch system **312** may be implemented as a computing system at the same location as the deployment system(s) **314** that it controls. For example, the local dispatch system **312** may be implemented by a computing system installed at a building, such as a warehouse, where the deployment system(s) **314** and UAV(s) **304** that are associated with the particular local dispatch system **312** are also located. In other embodiments, the local dispatch system **312** may be implemented at a location that is remote to its associated deployment system(s) **314** and UAV(s) **304**.

[0096] Numerous variations on and alternatives to the illustrated configuration of the UAV system **300** are possible. For example, in some embodiments, a user of the remote device **306** could request delivery of a package directly from the central dispatch system **310**. To do so, an application may be implemented on the remote device **306** that allows the user to provide information regarding a requested delivery, and generate and send a data message to request that the UAV system **300** provide the delivery. In such an embodiment, the central dispatch system **310** may include automated functionality to handle requests that are generated by such an application, evaluate such requests, and, if appropriate, coordinate with an appropriate local dispatch system **312** to deploy a UAV.

[0097] Further, some or all of the functionality that is attributed herein to the central dispatch system **310**, the local dispatch system(s) **312**, the access system **302**, and/or the deployment system(s) **314** may be combined in a single system, implemented in a more complex system, and/or redistributed among the central dispatch system **310**, the local dispatch system(s) **312**, the access system **302**, and/or the deployment system(s) **314** in various ways.

[0098] Yet further, while each local dispatch system **312** is shown as having two associated deployment systems **314**, a given local dispatch system **312** may alternatively have more or fewer associated deployment systems **314**. Similarly, while the central dispatch system **310** is shown as being in communication with two local dispatch systems **312**, the central dispatch system **310** may alternatively be in communication with more or fewer local dispatch systems **312**.

[0099] In a further aspect, the deployment systems **314** may take various forms. In general, the deployment systems **314** may take the form of or include systems for physically launching one or more of the UAVs **304**. Such launch systems may include features that provide for an automated UAV launch and/or features that allow for a human-assisted UAV launch. Further, the deployment systems **314** may each be configured to launch one particular UAV **304**, or to launch multiple UAVs **304**.

[0100] The deployment systems **314** may further be configured to provide additional functions, including for example, diagnostic-related functions such as verifying system functionality of the UAV, verifying functionality of devices that are housed within a UAV (e.g., a payload delivery apparatus), and/or maintaining devices or other items that are housed in the UAV (e.g., by monitoring a status of a payload such as its temperature, weight, etc.).

[0101] In some embodiments, the deployment systems **314** and their corresponding UAVs **304** (and possibly associated local dispatch systems **312**) may be strategically distributed throughout an area such as a city. For example, the deployment systems **314** may be strategically distributed such that

each deployment system **314** is proximate to one or more payload pickup locations (e.g., near a restaurant, store, or warehouse). However, the deployment systems **314** (and possibly the local dispatch systems **312**) may be distributed in other ways, depending upon the particular implementation. As an additional example, kiosks that allow users to transport packages via UAVs may be installed in various locations. Such kiosks may include UAV launch systems, and may allow a user to provide their package for loading onto a UAV and pay for UAV shipping services, among other possibilities. Other examples are also possible.

[0102] In a further aspect, the UAV system **300** may include or have access to a user-account database **316**. The user-account database **316** may include data for a number of user accounts, and which are each associated with one or more persons. For a given user account, the user-account database **316** may include data related to or useful in providing UAV-related services. Typically, the user data associated with each user account is optionally provided by an associated user and/or is collected with the associated user's permission.

[0103] Further, in some embodiments, a person may be required to register for a user account with the UAV system **300**, if they wish to be provided with UAV-related services by the UAVs **304** from UAV system **300**. As such, the user-account database **316** may include authorization information for a given user account (e.g., a username and password), and/or other information that may be used to authorize access to a user account.

[0104] In some embodiments, a person may associate one or more of their devices with their user account, such that they can access the services of UAV system **300**. For example, when a person uses an associated mobile phone, e.g., to place a call to an operator of the access system **302** or send a message requesting a UAV-related service to a dispatch system, the phone may be identified via a unique device identification number, and the call or message may then be attributed to the associated user account. Other examples are also possible.

Machine Learning Models for Generating Inferences/Predictions

[0105] A UAV may be configured to use one or more machine learning models to facilitate perception, localization, navigation, and/or other UAV operations. FIG. **4** shows diagram **400** illustrating a training phase **402** and an inference phase **404** of trained machine learning model(s) **432**, in accordance with example embodiments. Some machine learning techniques involve training one or more machine learning algorithms on an input set of training data to recognize patterns in the training data and provide output inferences and/or predictions about (patterns in the) training data. The resulting trained machine learning algorithm can be termed as a trained machine learning model. For example, FIG. **4** shows training phase **402** where one or more machine learning algorithms **420** are being trained on training data **410** to become trained machine learning model **432**. Then, during inference phase **404**, trained machine learning model **432** can receive input data **430** and one or more inference/prediction requests **440** (perhaps as part of input data **430**) and responsively provide as an output one or more inferences and/or predictions **450**.

[0106] As such, trained machine learning model(s) **432** can include one or more models of one or more machine learning algorithms **420**. Machine learning algorithm(s) **420** may include, but are not limited to: an artificial neural network (e.g., a herein-described convolutional neural networks, a recurrent neural network, a Bayesian network, a hidden Markov model, a Markov decision process, a logistic regression function, a support vector machine, a suitable statistical machine learning algorithm, and/or a heuristic machine learning system). Machine learning algorithm(s) **420** may be supervised or unsupervised, and may implement any suitable combination of online and offline learning.

[0107] In some examples, machine learning algorithm(s) **420** and/or trained machine learning model(s) **432** can be accelerated using on-device coprocessors, such as graphic processing units (GPUs), tensor processing units (TPUs), digital signal processors (DSPs), and/or application specific integrated circuits (ASICs). Such on-device coprocessors can be used to speed up machine learning algorithm(s) **420** and/or trained machine learning model(s) **432**. In some examples, trained

machine learning model(s) **432** can be trained, reside, and execute to provide inferences on a particular UAV or other computing device, and/or otherwise can make inferences for the particular UAV or other computing device.

[0108] During training phase **402**, machine learning algorithm(s) **420** can be trained by providing at least training data **410** as training input using unsupervised, supervised, semi-supervised, and/or reinforcement learning techniques. Unsupervised learning involves providing a portion (or all) of training data **410** to machine learning algorithm(s) **420** and machine learning algorithm(s) **420** determining one or more output inferences based on the provided portion (or all) of training data **410**. Supervised learning involves providing a portion of training data **410** to machine learning algorithm(s) **420**, with machine learning algorithm(s) **420** determining one or more output inferences based on the provided portion of training data **410**, and the output inference(s) are either accepted or corrected based on correct results associated with training data **410**. In some examples, supervised learning of machine learning algorithm(s) **420** can be governed by a set of rules and/or a set of labels for the training input, and the set of rules and/or set of labels may be used to correct inferences of machine learning algorithm(s) **420**. Supervised learning of machine learning algorithm(s) **420** may include training data **410** provided by human operators and/or a trusted dataset, which may in some examples be provided by a third party.

[0109] Semi-supervised learning involves having correct results for part, but not all, of training data **410**. During semi-supervised learning, supervised learning is used for a portion of training data **410** having correct results, and unsupervised learning is used for a portion of training data **410** not having correct results. Reinforcement learning involves machine learning algorithm(s) **420** receiving a reward signal regarding a prior inference, where the reward signal can be a numerical value. During reinforcement learning, machine learning algorithm(s) **420** can output an inference and receive a reward signal in response, where machine learning algorithm(s) **420** are configured to try to maximize the numerical value of the reward signal. In some examples, reinforcement learning also utilizes a value function that provides a numerical value representing an expected total of the numerical values provided by the reward signal over time. In some examples, machine learning algorithm(s) **420** and/or trained machine learning model(s) **432** can be trained using other machine learning techniques, including but not limited to, incremental learning and curriculum learning.

[0110] In some examples, machine learning algorithm(s) **420** and/or trained machine learning model(s) **432** can use transfer learning techniques. For example, transfer learning techniques can involve trained machine learning model(s) **432** being pre-trained on one set of data and additionally trained using training data **410**. More particularly, machine learning algorithm(s) **420** can be pre-trained on data from one or more computing devices and a resulting trained machine learning model provided to a UAV, where the UAV is intended to execute the trained machine learning model during inference phase **404**. Then, during training phase **402**, the pre-trained machine learning model can be additionally trained using training data **410**, where training data **410** can be derived from the UAV. This further training of the machine learning algorithm(s) **420** and/or the pre-trained machine learning model using training data **410** of the UAV's data can be performed using either supervised or unsupervised learning. Once machine learning algorithm(s) **420** and/or the pre-trained machine learning model has been trained on at least training data **410**, training phase **402** can be completed. The trained resulting machine learning model can be utilized as at least one of trained machine learning model(s) **432**.

[0111] In particular, once training phase **402** has been completed, trained machine learning model(s) **432** can be provided to a UAV, if not already on the UAV. Inference phase **404** can begin after trained machine learning model(s) **432** are provided to the UAV.

[0112] During inference phase **404**, trained machine learning model(s) **432** can receive input data **430** and generate and output one or more corresponding inferences and/or predictions **450** about input data **430**. As such, input data **430** can be used as an input to trained machine learning

model(s) **432** for providing corresponding inference(s) and/or prediction(s) **450** to a UAV. For example, trained machine learning model(s) **432** can generate inference(s) and/or prediction(s) **450** in response to one or more inference/prediction requests **440**. Input data **430** can include data from a first UAV executing trained machine learning model(s) **432** and/or input data from one or more other UAVs.

[0113] In some examples, input data **430** can include a collection of images provided by one or more sources. The collection of images can include still images, video frames, images resident on a UAV, and/or other images. In some examples, input data **430** may include one or more two-dimensional (2D) images captured by a ground-facing camera on a UAV and representative of an environment of the UAV. The images of the environment may include, among other possibilities, a terrain beneath the UAV when the UAV is navigating from a starting location to a target location. The image of the environment may also include a delivery zone to which the UAV is tasked with delivering a payload. Other types of input data are possible as well.

[0114] Inference(s) and/or prediction(s) **450** can include output images, output intermediate images, numerical values, and/or other output data produced by trained machine learning model(s) **432** operating on input data **430** (and training data **410**). In some examples, trained machine learning model(s) **432** can use output inference(s) and/or prediction(s) **450** as input feedback **460**. Trained machine learning model(s) **432** can also rely on past inferences as inputs for generating new inferences.

[0115] In some examples, an output image included within inference(s) and/or prediction(s) **450** may be a semantic image, which includes semantic labels. The semantic labels may describe pixels or pixel areas within the image as representing different types of areas in the environment. The semantic labels may be selected from a predetermined set of labels. In some examples, the set of labels may include labels representing buildings, roads, vegetation, vehicles, driveways, lawns, and sidewalks. In further examples, other labels may be included as well or instead.

[0116] In further examples, an output image included within inference(s) and/or prediction(s) **450** may be a depth image. Each pixel of the depth image may represent depth, or distance of the pixel to an imaging device which captured the image. In examples, where the camera is a 2D camera, the depth image generated by a machine learning model may be referred to as a monocular depth image.

[0117] Output images included within inference(s) and/or prediction(s) **450** may be used by a UAV to facilitate UAV localization, navigation, payload delivery, and/or other deployment operations. In further examples, inference(s) and/or prediction(s) **450** may include other types of output data, which may be used to facilitate the same or different UAV deployment operations as well or instead.

[0118] FIG. 5 is a diagram illustrating application of a machine learning model, in accordance with example embodiments. More specifically, an example machine learning model may include encoding layers **502** and decoding layers **504** in order to operate on an input image **510**. In some examples, the input image **510** may be a single image captured by a sensor on a UAV, such as a downward-facing two-dimensional (2D) camera. The input image **510** may be representative of the terrain beneath the UAV and may be captured in order to facilitate localization of the UAV.

[0119] The encoding layers **502** and decoding layers **504** may each include one or more layers of a machine learning model as otherwise described herein, such as a neural network. In some examples, the decoding layers **504** may include a semantic branch **506** and a depth branch **508**. The semantic branch **506** may be trained to generate semantic labels of portions of the input image **510**. For example, individual pixels within the input image **510** may be labeled as representing different types of objects in the environment, such as trees, roads, or buildings. The depth branch **508** may be trained to generate a depth map representing distance to different portions of the input image **510**. In some examples, the depth branch **508** may be trained to generate a point cloud representation of the environment of the UAV.

[0120] In some examples, the machine learning model represented in FIG. 5 may be structured as a multi-task neural network to produce both semantic and depth information. One or more layers of the neural network may be shared to produce each of the two types of output information in order to leverage commonality between the task of semantic labeling and the task of generating depth information. Accordingly, the semantics branch and the depth branch of the machine learning model may operate on a commonly generated feature set.

[0121] The semantic branch 506 may be trained to generate a semantic image 512 corresponding to the input image 510. The semantic image 512 may include semantic labels for different portions of the environment represented by the input image 510. The depth branch 508 may be trained to generate a depth image 514 corresponding to the input image 510. The depth image 514 may include depth estimates for different portions of the environment represented by the input image 510. In some examples, the semantic image 512 and the depth image 514 may have the same dimensions in order to enable alignment of semantic labels and depth values. For instance, by using the same dimensions for the semantic image 512 and the depth image 514, a single semantically labeled point cloud may be generated by the machine learning model.

[0122] In order to train the machine learning model (and in particular semantic branch 506), semantic image 512 generated by the model may be compared against semantic truth 522. Semantic truth 522 may be a ground truth image with semantic labels. In some examples, the semantic labels for semantic truth 522 may be provided by a human operator (e.g., through a graphical user interface where the user identifies particular types of objects commonly represented in images of UAV environments). The ground truth semantic data may therefore be based on operator labeling of images captured by one or more UAVs. In further examples, the semantic labels for semantic truth 522 may be provided in a dataset from a vendor. In yet further examples, the semantic labels for semantic truth 522 may be determined based on past flights by the UAV and/or a different UAV using a different automated methodology for semantic labeling.

[0123] In order to further train the machine learning model (and in particular depth branch 508), depth image 514 generated by the model may be compared against depth truth 524. Depth truth 524 may be a ground truth depth image. In some examples, depth truth 524 may be a point cloud representation of a portion of an environment of a UAV. In some examples, the depth data for depth truth 524 may be provided in a dataset from a vendor. In further examples, the depth data for depth truth 524 may be determined based on past flights by the UAV and/or a different UAV. More specifically, a computer vision pipeline referred to as a structure from motion (SfM) algorithm may be used to operate on a series of camera images of a scene captured by a UAV in order to assemble a 3D representation of the scene. Notably, a machine learning model trained based on input images and corresponding SfM depth information may be trained to ultimately produce smoother and more complete depth images than the SfM depth images themselves, which may be relatively sparse in some cases. In further examples, depth truth 524 may also be based on other types of UAV depth sensors such as lidar as well or instead.

[0124] During the training phase, the depth branch of the machine learning model may be trained using different losses. In some examples, the depth branch may be trained with an absolute loss in which the network attempts to predict absolute metric depths (e.g. depth of the scene in meters or millimeters). In other examples, the depth branch of the machine learning model is trained using a scale invariant loss in which the network will produce depth maps of the scene that are correct up to a scale factor. In such examples, unscaled depth maps are the product of training a monocular depth model using a scale invariant loss.

[0125] Consequently, during the inference phase, a scale factor may be applied to such an unscaled depth map or depth image. The scale factor can be recovered in numerous ways. In some such examples, the scale factor can be a ratio of an altitude of the UAV above ground level relative to a median of the monocular depth map. Using a median of the monocular depth map may provide a convenient approximation to recover a scaled depth map when using a model that has been trained

with a scale invariant loss. In further examples, an above ground level estimate from a monocular depth map may be used instead of the median of the monocular depth map for increased precision.

V. EXAMPLE UAV LOCALIZATION SYSTEMS AND METHODS

[0126] FIG. 6A illustrates an image captured by a UAV, in accordance with example embodiments. More specifically, an image **600** may be captured by a camera on the UAV and the image **600** may represent a portion of the environment of the UAV. The image **600** may include one or more objects which may facilitate UAV localization based on depth information and/or a semantic understanding of the objects. For instance, in the illustrated example, the image **600** includes trees **602** and building **604**. In other examples, different types of objects may be captured in images instead.

[0127] In the example illustrated in FIG. 6A, the image **600** may be captured by a camera on the UAV which faces downward. The camera may therefore capture a 2D image representative of a terrain in the environment below the UAV. In general, different areas of terrain along a route of a UAV may be used to help localize the UAV during performance of a mission.

[0128] FIG. 6B illustrates reference depth data representative of an environment of a UAV, in accordance with example embodiments. More specifically, point cloud **610** may represent the same general location as illustrated by image **600**. In some examples, point cloud **610** may be provided by a vendor. In other examples, point cloud **610** may have been generated based on data from one or more UAVs during past flights of the UAVs. More specifically, point cloud **610** may have been generated using a structure from motion (SfM) algorithm applied to a series of images captured by a camera on a UAV.

[0129] Point cloud **610** may also include one or more semantically labeled areas. The semantically labeled areas may be individual points in the point cloud. For instance, each point identified as belonging to an object of interest may be semantically labeled. First area **612** may include semantically labeled points indicating that the points represent trees in the environment. Second area **614** may include semantically labeled points indicating that the points represent buildings in the environment. In other examples, larger areas in the point cloud **610** may be labeled rather than individual points. In general, not all the points in point cloud **610** may have a semantic label due to areas being unidentified and/or due to areas being identified as some type of object other than a predefined set of object types selected to assist in semantic localization of the UAV.

[0130] In some examples, the reference depth data illustrated in FIG. 6B may be retrieved by the UAV and/or provided to the UAV in advance of a flight mission to be followed by the UAV. The reference depth data may therefore represent terrain expected to be encountered by the UAV along the flight mission. Advance retrieval of relevant reference depth data may be advantageous in examples where the localization methodologies described herein are used in situations where connectivity is lost and other types of localization options such as GPS may be unavailable.

[0131] FIG. 6C illustrates semantically labeled depth data generated by a machine learning model, in accordance with example embodiments. More specifically, semantically labeled depth data **620** may be generated by applying a trained machine learning model to the image **600** illustrated in FIG. 6A. The depth data **620** may include semantic labels for one or more areas of the depth data **620**. More specifically, first area **622** may include one or more semantic labels indicating that the first area **622** represents one or more trees in the environment. Second area **624** may include one or more semantic labels indicating that the second area **624** represents one or more buildings in the environment.

[0132] In some examples, semantically labeled depth data **620** may be generated by applying a machine learning model, such as the model illustrated and described with respect to FIG. 5. In some such examples, semantically labeled depth data **620** may take the form of a separate depth image aligned with a separate semantic image, rather than a single data structure. Similar to the reference data, the machine learning model may not semantically label all of the points or areas due to areas being unidentified and/or due to areas being identified as some type of object other than a predefined set of object types selected to assist in semantic localization of the UAV. In some

examples, the same set of semantic labels may be applied both for the reference data and the data generated using a machine learning model. In other examples, there may be variations between the sets of labels (particularly where vendor data is used as reference data).

[0133] FIG. 6D illustrates superimposed reference depth data and semantically labeled depth data generated by a machine learning model, in accordance with example embodiments. More specifically, the reference data depth from FIG. 6B is illustrated superimposed with the semantically labeled depth data generated by a machine learning model from FIG. 6C. Notably, the UAV may not be entirely confident in its current position before performance of the methodologies described herein, due to inaccuracies or disruptions in other localization techniques. Accordingly, areas representing the same objects between the two sources may not be fully aligned when the UAV attempts to align the data based on other localization information. As illustrated in FIG. 6D, first area **612** from the reference depth data and first area **622** from the semantically labeled depth data generated by the machine learning model are not aligned although they both represent the same trees in the environment of the UAV. Similarly, second area **614** from the reference depth data and second area **624** from the semantically labeled depth data generated by the machine learning model are not aligned although they both represent the same building in the environment of the UAV.

[0134] FIG. 6E illustrates reference depth data aligned with semantically labeled depth data generated by a machine learning model, in accordance with example embodiments. More specifically, an alignment process may be applied to align the semantically labeled depth data generated by the machine learning model with the reference depth data. In accordance with examples described herein, semantic information about individual areas may be incorporated into the alignment process. In particular, the aligning may associate one or more semantic labels from the semantically labeled depth data generated by the machine learning model with reference semantic labels from the reference depth data. The result of the alignment process may be a 6 degree of freedom (6DOF) transformation between the two sets of data. An image captured by a UAV which is input into a pipeline as described herein may therefore yield a 6DOF position and orientation of the UAV relative to a known reference location in the environment.

[0135] In some examples, the aligning process may involve application of an iterative closest point (ICP) algorithm. The ICP algorithm may align points from the reference depth data with points from the semantically labeled depth data generated by the machine learning model such that the reference semantic labels from the reference depth data correspond to one or more semantic labels from the semantically labeled depth data. In some examples, points may only be aligned if they share the same label. In other examples, a penalty factor may be applied to points with unaligned semantic labels. In yet further examples, different semantic labels may be weighted differently in application of iterative closest point or a different aligning process. For instance, two points sharing the exact same semantic label may be given a highest weight, two points sharing similar but not the exact same semantic label (e.g., trees vs. bushes) may be given a lower weight, and two points sharing disparate semantic labels (e.g., trees vs. buildings) may be given a lowest weight. Other types of biasing may be used when factoring semantic information into determining a transformation that aligns the two sets of depth data as well or instead.

[0136] FIG. 7 is a block diagram of a method, in accordance with example embodiments. The operations illustrated in FIG. 7 may be carried out by UAV **100**, UAV **120**, UAV **140**, UAV **160**, UAV **180**, UAV **200**, or distributed UAV system **300** among other possibilities. The functions of FIG. 7 may be simplified by the removal of any one or more of the features shown therein. Further, these examples may be combined with features, aspects, and/or implementations of any of the previous figures or otherwise described herein. In some examples, the operations of FIG. 7 may each be performed by a processor executing program code stored on a non-transitory computer readable medium. In some examples, the operations of FIG. 7 may be performed fully by a UAV, fully by a computing system remote from the UAV, and/or by a combination of local and remote

computing systems.

[0137] Block **702** includes receiving a two-dimensional (2D) image captured by a camera on a unmanned aerial vehicle (UAV) and representative of an environment of the UAV.

[0138] Block **704** includes applying a trained machine learning model to the 2D image to produce a semantic image of the environment and a depth image of the environment, where the machine learning model has been trained with a semantics branch to produce the semantic image and a depth branch to produce the depth image, and where the semantic image comprises one or more semantic labels.

[0139] Block **706** includes retrieving reference depth data representative of the environment, wherein the reference depth data includes reference semantic labels.

[0140] Block **708** includes aligning the depth image of the environment with the reference depth data representative of the environment to determine a location of the UAV in the environment, wherein the aligning associates the one or more semantic labels from the semantic image with the reference semantic labels from the reference depth data.

[0141] In some examples, the operations further include controlling the UAV to navigate in the environment based on the determined location of the UAV in the environment. Such examples may involve first controlling the UAV to navigate in the environment using a Global Navigation Satellite System (GNSS) system. Subsequently, a disruption in service from the GNSS system may be detected. The location of the UAV in the environment may then be determined as described herein responsive to detecting the disruption in service from the GNSS system. Subsequent to detecting the disruption in service from the GNSS system, the operations may therefore involve controlling the UAV to navigate in the environment based on the determined location of the UAV in the environment.

[0142] Further examples may involve controlling the UAV to navigate in the environment using a GNSS system, and using the determined location of the UAV in the environment to cross-check location data from the GNSS system. Yet further examples may involve determining a GNSS location of the UAV in the environment using a GNSS system, and determining a refined location of the UAV in the environment based on the GNSS location of the UAV in the environment and the determined location of the UAV in the environment. The UAV may then be controlled to navigate in the environment based on the refined location of the UAV in the environment.

[0143] In some examples, aligning the depth image of the environment with the reference depth data representative of the environment to determine the location of the UAV in the environment may involve using an iterative closest point (ICP) algorithm. In such examples, the ICP algorithm may align points from the reference depth data with points from the depth image such that the reference semantic labels from the reference depth data correspond to the one or more semantic labels from the semantic image.

[0144] In some examples, the semantic image of the environment and the depth image of the environment produced by the machine learning model have the same dimensions. In some examples, the semantics branch and the depth branch of the machine learning model operate on a commonly generated feature set. In some examples, the machine learning model has been trained based on ground truth depth data, where the ground truth depth data is based on performance of a structure from motion (SfM) algorithm on images captured by one or more UAVs. In some examples, the machine learning model has been trained based on ground truth semantic data, where the ground truth semantic data is based on operator labeling of images captured by one or more UAVs. In some examples, the machine learning model has been trained using a scale invariant loss for training of the depth branch. Such examples may involve applying a scale factor during the inference phase. In some such examples, the scale factor may be a ratio of an altitude of the UAV above ground level relative to a median of the monocular depth map. In other examples, the scale factor may be a ratio of an altitude of the UAV above ground level relative to an above ground level estimate from an unscaled monocular depth map.

[0145] In some examples, the one or more semantic labels are selected from a predetermined set of labels. In some such examples, the predetermined set of labels includes at least the following labels: building, road, vegetation, vehicle, driveway, lawn, and sidewalk.

[0146] Some examples involve retrieving the reference depth data in advance of a flight of the UAV, where the reference depth data is selected based on a planned flight path of the UAV. Some examples involve applying a Kalman filter to the determined location of the UAV in the environment to combine with one or more other types of data in order to control navigation of the UAV in the environment.

[0147] In some examples, the camera on the UAV faces downward, and the 2D image captured by the camera is representative of the terrain in the environment below the UAV.

VI. CONCLUSION

[0148] The particular arrangements shown in the Figures should not be viewed as limiting. It should be understood that other implementations may include more or less of each element shown in a given Figure. Further, some of the illustrated elements may be combined or omitted. Yet further, an exemplary implementation may include elements that are not illustrated in the Figures.

[0149] Additionally, while various aspects and implementations have been disclosed herein, other aspects and implementations will be apparent to those skilled in the art. The various aspects and implementations disclosed herein are for purposes of illustration and are not intended to be limiting, with the true scope and spirit being indicated by the following claims. Other implementations may be utilized, and other changes may be made, without departing from the spirit or scope of the subject matter presented herein. It will be readily understood that the aspects of the present disclosure, as generally described herein, and illustrated in the figures, can be arranged, substituted, combined, separated, and designed in a wide variety of different configurations, all of which are contemplated herein.

Claims

1. A method comprising: receiving an image captured by a camera on a unmanned aerial vehicle (UAV) and representative of an environment of the UAV; determining, based on the image, a semantic image of the environment and a depth image of the environment; retrieving reference depth data representative of the environment, wherein the reference depth data includes reference semantic labels; determining a location of the UAV in the environment, wherein determining the location of the UAV in the environment comprises aligning the depth image of the environment with the reference depth data representative of the environment, wherein aligning the depth image of the environment with the reference depth data representative of the environment is based on associating one or more semantic labels from the semantic image with the reference semantic labels from the reference depth data; and controlling the UAV to navigate in the environment based on the determined location of the UAV in the environment.

2. The method of claim 1, further comprising: controlling the UAV to navigate in the environment using a Global Navigation Satellite System (GNSS) system; detecting a disruption in service from the GNSS system, wherein the location of the UAV in the environment is determined responsive to detecting the disruption in service from the GNSS system; and subsequent to detecting the disruption in service from the GNSS system, controlling the UAV to navigate in the environment based on the determined location of the UAV in the environment.

3. The method of claim 1, further comprising: controlling the UAV to navigate in the environment using a GNSS system; and using the determined location of the UAV in the environment to cross-check location data from the GNSS system.

4. The method of claim 1, further comprising: determining a GNSS location of the UAV in the environment using a GNSS system; determining a refined location of the UAV in the environment based on the GNSS location of the UAV in the environment and the determined location of the

UAV in the environment; and controlling the UAV to navigate in the environment based on the refined location of the UAV in the environment.

5. The method of claim 1, wherein the aligning the depth image of the environment with the reference depth data representative of the environment to determine the location of the UAV in the environment comprises using an iterative closest point (ICP) algorithm.
6. The method of claim 5, wherein the ICP algorithm aligns points from the reference depth data with points from the depth image such that the reference semantic labels from the reference depth data correspond to the one or more semantic labels from the semantic image.
7. The method of claim 1, wherein the semantic image of the environment and the depth image of the environment have the same dimensions.
8. The method of claim 1, wherein the camera on the UAV faces downward, and wherein the image captured by the camera is representative of a terrain in the environment below the UAV.
9. The method of claim 1, wherein the semantic image and the depth are determined using a machine learning model.
10. The method of claim 9, wherein a semantics branch and a depth branch of the machine learning model operate on a commonly generated feature set.
11. The method of claim 9, wherein the machine learning model has been trained based on ground truth depth data, wherein the ground truth depth data is based on performance of a structure from motion (SfM) algorithm on images captured by one or more UAVs.
12. The method of claim 9, wherein the machine learning model has been trained based on ground truth semantic data, wherein the ground truth semantic data is based on operator labeling of images captured by one or more UAVs.
13. The method of claim 9, wherein the machine learning model has been trained using a scale invariant loss.
14. The method of claim 13, further comprising applying a scale factor to the depth image, wherein the scale factor comprises a ratio of an altitude of the UAV above ground level relative to a median of a monocular depth map, wherein the monocular depth map is based on the reference depth data.
15. The method of claim 13, further comprising applying a scale factor to the depth image, wherein the scale factor comprises a ratio of an altitude of the UAV above ground level relative to an above ground level estimate from a monocular depth map, wherein the monocular depth map is based on the reference depth data.
16. The method of claim 1, wherein the one or more semantic labels are selected from a predetermined set of labels, wherein the predetermined set of labels comprises at least the following labels: building, road, vegetation, vehicle, driveway, lawn, and sidewalk.
17. The method of claim 1, further comprising retrieving the reference depth data in advance of a flight of the UAV, wherein the reference depth data is selected based on a planned flight path of the UAV.
18. The method of claim 1, further comprising applying a Kalman filter to the determined location of the UAV in the environment to control navigation of the UAV in the environment.
19. An unmanned aerial vehicle (UAV), comprising: a camera; and a control system configured to: receive an image captured by the camera on the UAV and representative of an environment of the UAV; determine, based on the image, a semantic image of the environment and a depth image of the environment; retrieve reference depth data representative of the environment, wherein the reference depth data includes reference semantic labels; determine a location of the UAV in the environment, wherein determining the location of the UAV in the environment comprises aligning the depth image of the environment with the reference depth data representative of the environment, wherein aligning the depth image of the environment with the reference depth data representative of the environment is based on associating one or more semantic labels from the semantic image with the reference semantic labels from the reference depth data; and control the UAV to navigate in the environment based on the determined location of the UAV in the

environment.

20. A non-transitory computer readable medium comprising program instructions executable by one or more processors to perform operations, the operations comprising: receiving an image captured by a camera on a unmanned aerial vehicle (UAV) and representative of an environment of the UAV; determining, based on the image, a semantic image of the environment and a depth image of the environment; retrieving reference depth data representative of the environment, wherein the reference depth data includes reference semantic labels; determining a location of the UAV in the environment, wherein determining the location of the UAV in the environment comprises aligning the depth image of the environment with the reference depth data representative of the environment, wherein aligning the depth image of the environment with the reference depth data representative of the environment is based on associating one or more semantic labels from the semantic image with the reference semantic labels from the reference depth data; and controlling the UAV to navigate in the environment based on the determined location of the UAV in the environment.
