(19) Europäisches Patentamt
European Patent Office
Office européen des brevets

(11) **EP 2 987 323 B1**

(12) **EUROPEAN PATENT SPECIFICATION**

(45) Date of publication and mention of the grant of the patent:
21.10.2020 Bulletin 2020/43

(21) Application number: **14724934.6**

(22) Date of filing: **14.04.2014**

(51) Int Cl.:
G01B 11/25 *(2006.01)* ***G02B 27/42*** *(2006.01)*
G06T 1/60 *(2006.01)* ***G06T 7/586*** *(2017.01)*
H04N 5/225 *(2006.01)* ***H04N 5/33*** *(2006.01)*
H04N 9/04 *(2006.01)* ***H04N 13/239*** *(2018.01)*
H04N 13/25 *(2018.01)* ***H04N 13/254*** *(2018.01)*
H04N 13/271 *(2018.01)* ***H04N 17/00*** *(2006.01)*
H04N 13/00 *(2018.01)*

(86) International application number:
PCT/US2014/033919

(87) International publication number:
WO 2014/172231 (23.10.2014 Gazette 2014/43)

(54) **ACTIVE STEREO WITH SATELLITE DEVICE OR DEVICES**

AKTIVES STEREO MIT EINER ODER MEHREREN SATELLITENVORRICHTUNGEN

STÉRÉO ACTIVE AVEC UN OU PLUSIEURS DISPOSITIFS SATELLITES

(84) Designated Contracting States:
AL AT BE BG CH CY CZ DE DK EE ES FI FR GB GR HR HU IE IS IT LI LT LU LV MC MK MT NL NO PL PT RO RS SE SI SK SM TR

(30) Priority: **15.04.2013 US 201361812232 P**
21.06.2013 US 201313924475

(43) Date of publication of application:
24.02.2016 Bulletin 2016/08

(73) Proprietor: **Microsoft Technology Licensing, LLC**
Redmond, WA 98052 (US)

(72) Inventors:
- **KIRK, Adam G.**
  Redmond, Washington 98052-6399 (US)
- **WHYTE, Oliver A.**
  Redmond, Washington 98052-6399 (US)
- **RHEMANN, Christoph**
  Redmond, Washington 98052-6399 (US)
- **IZADI, Shahram**
  Redmond, Washington 98052-6399 (US)

(74) Representative: **CMS Cameron McKenna Nabarro Olswang LLP**
Cannon Place
78 Cannon Street
London EC4N 6AF (GB)

(56) References cited:
KR-A- 20110 132 260
US-A1- 2002 136 444
US-A1- 2012 056 982

Note: Within nine months of the publication of the mention of the grant of the European patent in the European Patent Bulletin, any person may give notice to the European Patent Office of opposition to that patent, in accordance with the Implementing Regulations. Notice of opposition shall not be deemed to have been filed until the opposition fee has been paid. (Art. 99(1) European Patent Convention).

EP 2 987 323 B1

Description

BACKGROUND

[0001] In active depth sensing, such as used by active stereo systems, a projector projects patterns of light such as infrared (IR) dots or lines to illuminate a scene being sensed. Note that the projector may use a laser that consumes on the order of 1W of power, which means that the projector consumes too much power to be practically built into in a small satellite device such as a smartphone or tablet, and instead is typically part of a device plugged into a wall socket.

[0002] The projected patterns are then captured by a camera / sensor (two or more in stereo systems), with the image (or images) processed to compute a depth map or the like. For example, in stereo systems, stereo cameras capture two images from different viewpoints. Then, for example, one way to perform depth estimation with a stereo pair of images is to find correspondences between the images, e.g., to correlate projected and sensed dots in the left image with counterpart dots in the right image. Once matched, the projected patterns within the images may be correlated with one another, and triangulation, possibly along with disparities between one or more features of the correlated dots (e.g., including their intensities) may be used to estimate a depth to an object struck by that particular projected dot.

[0003] In most scenarios, the cameras that capture the stereo images are arranged to sense a relatively large space, such as a room, which can be relatively large. As a result, an object in the distance, such as a face, appears in only a relatively small number of camera pixels. Sufficient detail and/or accurate depth estimation is thus not available as needed by many applications, such as face recognition and so forth.

US 2012/0056982 A1 (Katz et al.) discloses a depth camera system that uses a structured light illuminator and multiple sensors such as infrared light detectors. One sensor can be optimized for shorter range detection while another sensor is optimized for longer range detection. The sensors can have a different baseline distance from the illuminator, as well as a different spatial resolution, exposure time and sensitivity. In one approach, depth values are obtained from each sensor by matching to the structured light pattern, and the depth values are merged to obtain a final depth map which is provided as an input to an application. The sensors are optimized to address variables such as lighting conditions, surface textures and colors, and the potential for occlusions. The optimization includes optimizing placement of the sensors relative to one another and relative to an illuminator. US 2002/0136444 A1 (Brown et al.) discloses a system for characterizing at least a portion of a surface of an object and generally complex surfaces such as automobile parts, airplane fuselages and wings, and antenna dishes. It comprises the steps of projecting from a projector having no predetermined pose a plurality of images onto a surface desired to be characterized. The images comprise a first set of images and a second set of images distinguishable from the first set. All the first and the second set of images have known two-dimensional coordinates relative to the projector. The method further comprises determining a pose of a first and a second camera positioned in spaced relation from each other by measurement of known reference points placed on the object to characterize. Using these known reference points, a first pose for a first camera and a second pose for a second camera are computed. The pose of projector is then deducted based on the calculated 3D coordinates of the elements of the first set of images.

SUMMARY

[0004] The scope of invention is defined by the appended claims. Further embodiments are defined in the dependent claims. The embodiments which do not fall within the scope of the appended set of claims are to be interpreted as example embodiments or background information, useful only for understanding the invention. According to aspects of the present invention there is provided a method, media and system as defined in the accompanying claims. The following Summary is provided to introduce a selection of representative concepts in a simplified form that are further described below in the Detailed Description. This Summary is not intended to identify key features or essential features of the claimed subject matter, nor is it intended to be used in any way that would limit the scope of the claimed subject matter.

[0005] Briefly, one or more of various aspects of the subject matter described herein are directed towards using information from a satellite (e.g., mobile) device to enhance depth data computed by another device, such as a base station. One or more aspects are directed towards receiving image-related data from a satellite device at another device, and enhancing a first set of depth data based at least in part upon the image-related data and pose information of the satellite device.

[0006] In one or more aspects, a base station is configured to determine pose information of a satellite device and capture one or more images. The base station receives image-related data from the satellite device, and compute a depth map based at least in part upon the pose information and the image-related data.

[0007] One or more aspects are directed towards receiving image-related data from a satellite device at a base station, and capturing actively illuminated stereo images at the base station. Pose information corresponding to the satellite device is determined. The actively illuminated stereo images are processed to determine a depth map from the base station's perspective that is enhanced by depth information corresponding to the image-related data.

[0008] Other advantages may become apparent from the following detailed description when taken in conjunction with the drawings.

BRIEF DESCRIPTION OF THE DRAWINGS

[0009] The present invention is illustrated by way of example and not limited in the accompanying figures in which like reference numerals indicate similar elements and in which:

FIGURE 1 is a representation of a base station that projects a light pattern into a scene, in which a satellite device captures at least part of the scene, according to one or more example implementations.

FIG. 2 is a block diagram representing example components that may be used to project and capture images for processing into depth data, according to one or more example implementations.

FIG. 3 is a block diagram representing example processing of base station data and satellite data into depth data, according to one or more example implementations.

FIG. 4 is a block diagram representing example processing of one satellite device's data and another satellite device's data into depth data, according to one or more example implementations.

FIG. 5 is a representation of satellite device configured to signal a base station, according to one or more example implementations.

FIGS. 6A and 6B are representations of ways in which depth data may be determined from a single satellite device camera, according to one or more example implementations.

FIG. 7 is a flow diagram representing example steps in obtaining image-related data from a satellite device for use in determining a depth map, according to one or more example implementations.

FIG. 8 is a block diagram representing an exemplary non-limiting computing system or operating environment, in the form of a mobile device, into which one or more aspects of various embodiments described herein can be implemented.

DETAILED DESCRIPTION

[0010] Various aspects of the technology described herein are generally directed towards using satellite (e.g., mobile) devices such as smartphones and tablet computers to communicate information to and from an active illumination base station, such as to improve depth sensing, provide more image detail and/or estimate pose. At the same time, the satellite devices leverage the base station's active illumination for their own active sensing, whereby the satellite devices do not need to consume battery power to project the light pattern.

[0011] In one or more alternatives, multiple satellite devices may use projected light from an external source to communicate image / depth / position data computed therefrom among one another. Depth and other useful data with respect to a scene thus may be available to the devices without the need for a fixed base station.

[0012] It should be understood that any of the examples herein are non-limiting. For instance, while satellite devices are generally exemplified herein as being mobile, an auxiliary device (or more than one) may be positioned as a "fixed satellite" which may be calibrated with the base station if desired. This may be used to provide a camera that is closer to some significant part of a scene, for example. Thus, as used herein, the term "satellite" need not be mobile and may include any auxiliary device. As another example, time-of-flight cameras may be used to determine at least some data. As such, the present invention is not limited to any particular embodiments, aspects, concepts, structures, functionalities or examples described herein. Rather, any of the embodiments, aspects, concepts, structures, functionalities or examples described herein are non-limiting, and the present invention may be used various ways that provide benefits and advantages in active depth sensing and image capture and processing in general.

[0013] FIG. 1 exemplifies one implementation, in which a base station 100 includes stereo (e.g., infrared) cameras 101 and 102 that capture images of a scene 104 from different perspectives. A projector 106 illuminates the scene 104 with a light pattern, (e.g., on the order of 100,000 points). The light pattern may be pseudorandom, and may be generated by emitting laser light through a diffractive optical element. For purposes of brevity, the points / components of the emitted pattern are generally referred to as dots herein, although line patterns, spots of other shapes and so forth instead may be generated and sensed.

[0014] The cameras 102 and 103 capture the dots as they reflect off of object surfaces in the scene 222 and (possibly) the background. In general, the dots sensed by one camera (e.g., 102) may be correlated with the dots sensed by the other camera (e.g., 103) via an image processing 108 (e.g., a system or subsystem), thereby providing (e.g., via triangulation) the distance to the reflective surface upon which each dot was incident. Note that FIG. 1 is not intended to be to scale, nor convey any sizes, distance, dot distribution pattern, dot density and so on.

[0015] The base station 100 also may include a clean IR camera, e.g., one with a filter that blocks the active illumination pattern and/or an RGB camera that captures visible light; if both are present, they may be combined into a single camera. Camera 112 in FIG. 1 represents any of these alternatives.

[0016] As shown in FIG. 1, one of the objects sensed in the scene 104 may be a satellite device 110, such as a mobile smartphone or tablet computing device. At the same time, the satellite device 110 may be sensing other objects in the scene, e.g., including the depth thereto, such as via stereo IR cameras that sense the illumination pattern, as well as possibly clean IR and/or RGB images thereof.

[0017] FIG. 2 shows an example base station 100 in which the stereo cameras 102 and 103 of an image cap-

turing system or subsystem 224 capture images synchronized in time (e.g., the cameras are "genlocked"). In one implementation the cameras capture infrared (IR) images, as IR does not affect the visible appearance of the scene (which is highly advantageous, such as in video conferencing and object modeling applications). As can be readily appreciated, in some scenarios such as studio environments, more than two IR depth-sensing cameras may be present. Further, one or more other cameras may be present in a given system, such as RGB cameras, and such other cameras may be used to help align images, correlate dot pairs in different stereo images, and so forth, for example.

[0018] In FIG. 2, the projector 106 projects an IR pattern onto a scene. By illuminating the scene with a relatively large number of distributed infrared dots, the cameras 102 and 103 capture more texture data as part of the infrared image data. Note that the placement of the projector 106 may be outside the cameras (e.g., FIG. 1), or in between the cameras (FIG. 2) or at another location, such as above or below one or both of the cameras. The examples herein are in no way limiting of where the cameras 102, 103 and/or projector 106 are located relative to one another, and similarly, the cameras may be positioned at different positions relative to each other.

[0019] In one implementation the example image capturing system or subsystem 224 includes a controller 228 that via a camera interface 230 controls the operation of the cameras 102 and 103, and if present, the camera 112. The exemplified controller via a projector interface 232 also controls the operation of the projector 106. For example, the cameras 102 and 103 are synchronized (genlocked) to capture stereo images at the same time, such as by a controller signal (or different signals for each camera). The projector 106 may be turned on or off, pulsed, and otherwise have one or more parameters controllably varied, for example.

[0020] The images captured by the cameras 102 and 103 (and camera 112, if present) are provided to the image processing 108 system or subsystem, comprising logic implemented in hardware and/or software (e.g., as computer-readable instructions). In some implementations, the image processing system 108 and image capturing system or subsystem 224, or parts thereof, may be combined into a single device. For example a home entertainment device may include all of the components shown in FIG. 1 (as well as others not shown). In other implementations, parts (or all) of the image capturing system or subsystem 224, such as the cameras and projector, may be a separate device that couples to a gaming console, personal computer, satellite device, dedicated processing device and/or the like.

[0021] In one or more implementations, the image processing system or subsystem 108 includes a processor 240 and a memory 242 containing one or more image processing algorithms 244. One or more depth maps may 246 be obtained via the algorithms 124 such as by extracting matching features (such as dots and/or lines). For example, as is known, such as described in U.S. published patent application no. 20130100256, different dots or other projected elements have different features when captured, including intensity (brightness), depending on the distance from the projector to the reflective surfaces and/or the distance from the camera to the reflective surfaces. As is also known, the dots in different images taken at the same time (e.g., with genlocked stereo cameras) may be correlated with one another, such as by matching small (e.g., RGB) patches between RGB images of the same scene captured at the same instant. Thus, with captured images, known algorithms can determine individual depth-related features (depth maps) by matching projected light components (e.g., dots) in each image, using triangulation / disparities of certain features between matched dots to determine depths. This is one way in which a depth map may be obtained via stereo image processing.

[0022] Also shown in FIG. 1 are one or more interfaces 132 to the base station, such as for connecting a keyboard, game controller, display, pointing device microphone for speech commands and/or the like as appropriate for a user to interact with an application or the like that uses the depth map. At least one interface 132 allows the satellite device 110 to communicate (e.g., wirelessly) with the based station as described herein.

[0023] One or more implementations comprise a base station 100 in a room emitting the laser pattern to calculate depth. When the user uses a satellite device 110, such as a smartphone or tablet, the camera or stereo cameras on the satellite device 110 can observe the base unit's emitted laser pattern. The satellite device 110 thus can compute stereo based on that externally-generated laser pattern, as well as communicate information to the base station 100.

[0024] As described herein, the satellite device's data may be used to compute a depth map at higher resolution than the base station can compute; (note that "higher resolution" as used herein is not referring to any of the cameras' resolutions, but rather to the ability to capture an object with more pixels than a more distant camera at the base station because the satellite device is closer to the object, e.g., the user's face). Moreover, not only may the spatial resolution be increased, but also the depth resolution / precision. The image information or data corresponding to the information (e.g., a depth map computed at the satellite device) may be transmitted to the base unit. As used herein, "image-related data" refers to actual image data (of one or more IR and/or RGB actively illuminated images and/or clean IR and/or RGB images, in whole or in part,) any associated metadata, and/or any corresponding information processed from that image data, e.g., a depth map. Thus, image-related data may be communicated to and from a base station, and to and from satellite devices.

[0025] Additionally, the images and/or depth maps taken at the satellite device and at the base unit can be correlated. This provides for six degrees of freedom

(6DoF) pose estimation of the mobile unit's location relative to the base unit.

[0026] Pose estimation may be determined by image alignment and/or by pattern matching / dot correlation. By way of example, if the satellite device captures an image that the base station can also sense (e.g., user or device is not blocking the base station from "seeing" the same area), then the pixels (or corresponding surrounding patches of each pixel) can be matched in known ways. Alignment also may be established by aligning the depth data computed from the base station with the depth data computed from the satellite device. The pose of the satellite device thus may be estimated.

[0027] This is generally represented in FIG. 3, where a base station 300 projects a light pattern from a projector 306 that is sensed by base station cameras 301 and 302 as well as by one or more cameras on a satellite device 330 (more than one satellite device may be present). The cameras 301 and 302 provide base station images 332 that are fed into image processing along with satellite device data 334, which may be images or information such as depth maps locally processed from the images on each satellite device.

[0028] Note that as indicated via the dashed lines from the device 330, an optional (e.g., low power) projector may be included in any or all of the one or more satellite devices. The data projected by the satellite device augments the data projected from the base station. The projector in the satellite device 330 may be lower-power because it is limited in distance, limited spatially (e.g. a very sparse pattern or only focused at a small image region, and/or limited temporally (e.g., emitted only every few frames).

[0029] The base station images 332 may be processed to determine the pose / 3D location of the satellite device 330. With this information, image processing 308 outputs depth data 336, which may be conventional depth data from the base station's own cameras enhanced by the satellite device data 334. In general, the projected dots captured by the base station cameras may correlated with dots captured by the satellite device, after adjusting for size differences and the like due to the resolution differences. The pose data 338 also may be output, such as to other satellite devices.

[0030] Note that instead of the satellite device providing the base station with data for enhancing its image data, the base station may send an approximate depth map (along with the device's pose) to the satellite device. This may be used to increase the accuracy, and/or decrease the required computation for the satellite device's own depth estimation.

[0031] The base station's depth map (as well as RGB data) may be enhanced by downsampling the higher resolution data into a partial depth map that is combined with the initial base station depth map at the appropriate coordinates, after adjusting the depths for the pose, to improve accuracy at those coordinates, for example. Another way that the base station's depth map may be enhanced is to maintain the original depth map in association with the higher resolution depth map (e.g., after adjusting depths based upon pose), for example. In this way, an application using the depth map may "zoom-in" to the more fine-grained depth data when desired, e.g., the application need not be limited to the original depth map's pixel resolution.

[0032] Note that some mobile devices may have front and rear facing cameras. If so, one scenario is to use the cameras for different purposes. For example, the rear facing camera may be used to capture images for pose computations, while the front facing camera may capture images of the user's face, for example, which may be used for enhanced depth computations.

[0033] FIG. 4 shows another alternative, in which a plurality of satellite devices 440(1) - 440 (n) act as peers and thus can compute depth information for a scene without a base station. However, because satellite devices are battery powered, a light pattern from an external projector 406 (or more than one) may be leveraged. Note that a base station is basically just another device, except generally not mobile (a base station device may be relatively compact and portable but uses a large or fixed power source, and thus is not typically carried around, unlike a tablet device or smartphone).

[0034] As shown in FIG. 4, each of the satellite devices 440(1) - 440 (n) exchanges its data with one another. One of the devices 440(2) is shown as having image processing 408 that processes its own image data 442 and other satellite device data 444 to obtain depth data 446 enhanced by one or more other satellite devices' data; it is understood that any satellite device may have similar image processing capabilities, and/or may receive depth data from another satellite device.

[0035] In another aspect, FIG. 5 shows satellite devices 550 and 552 configured to provide a signal such as to flash an identifier to a base station 500 via infrared light. The identifier may include or be used to look up any type of information, including device type, capabilities and so forth. The signal also may help in more efficiently determining satellite device pose.

[0036] Note that instead of the base station identifying a satellite device, and/or estimating its pose, satellite devices may identify the base station, and estimate its own pose based on that information. This allows a satellite device to determine its pose without necessarily communicating with the base station, which may be useful in some situations (e.g. if there is a relatively large number of satellite devices).

[0037] In another aspect, FIG. 6A shows how a satellite device moving over time can compute depth data with a single camera. Because the dots are stationary, any movement of the device changes the dots' positions from previous to the new perspective of the camera.

[0038] FIG. 6B shows two projectors 606 and 607 sensed by one camera on a satellite mobile device. The projected patterns may be different, whereby the satellite device may distinguish one source from another, or the

projectors can alternate which one is projecting at a time to provide two different perspectives of the same object / scene being sensed.

[0039] Indeed, a projector may have knowledge of the light pattern, and if so is basically a camera in reverse. Thus a single projector and single satellite device camera may be used to compute stereo depth data.

[0040] FIG. 7 is a simplified flow diagram representing an example set of steps that may be taken to obtain enhanced depth information (and possibly other information such as RGB images) via a satellite device. For example, consider that an application program running on or coupled to a base station wants to produce a detailed close-up representation of an object that is relatively far from the base station and not particularly easy or desirable to move closer. Although a base station is described in the example steps below, it is understood that another mobile device or set of devices may be used instead of a base station.

[0041] At step 702, a base station program (or a program coupled thereto) communicates a message to a user instructing the user to aim a satellite device at the object to obtain some close-up video frames. At step 704, the user does so, whereby the satellite device starts transmitting data (e.g., a stream of images or depth data processed therefrom) to the base station. At each frame, at step 706 the base station processes its own captured images to determine the satellite camera pose, and compute an initial depth map.

[0042] Step 708 represents receiving data from the satellite device; note that at least some of these steps may be occurring in parallel and/or in different orders, e.g., some or all of the data may be received from the satellite device before, during or after base station processing. Further note that instead of per frame, some slower sampling rate may be used depending on the application.

[0043] Step 710 represents using the satellite data to enhance the depth map with more accurate depth data, for example, and/or to maintain a separate depth map for the captured object. Step 712 repeats the process until done, at which time the user is notified.

[0044] The user may be asked to capture video of the object from different directions. If projected light is being sensed by the satellite device, the user may be instructed to turn the object so that different parts of the object face the projector at different times (note that the user cannot see the infrared light pattern, and thus may not know it is being projected, whereby the instruction may be to rotate the object to face the base station device). Alternatively multiple projectors from different directions may be present. In this way, a full three-dimensional mesh or the like of the object may be generated (with which RGB image data may be combined), at a significantly greater level of detail than the base station is able to capture.

[0045] The satellite device and base station (or other mobile device) may operate together without human intervention. For example, a mobile device application may detect something interesting via the device camera and communicate image data to the base station. Periodically or on some other schedule the base station may request one or more images from a satellite device. The base station (e.g., as instructed by an application program) may want to have an improved image of something that (as known via pose data) that the satellite device's camera is aiming at. For example, this may be to obtain a higher resolution image whenever desired.

[0046] As another example, data from the satellite device may be used to replace data in the image map captured by a satellite device or devices. As a particular example, consider that something is blocking a desired part of a scene from the base station's perspective, e.g., a person has walked in front of an interesting part of a scene that the satellite device seeks to capture. Via a satellite device or set of satellite devices, the scene may be captured (from a different angle and possibly different resolution), whereby the scene depths (and RGB or clean IR data) can be recomputed as if the person was not blocking that part of the scene.

EXAMPLE OPERATING ENVIRONMENT

[0047] FIG. 8 illustrates an example of a suitable mobile device 800 on which aspects of the subject matter described herein may be implemented. The mobile device 800 is only one example of a device and is not intended to suggest any limitation as to the scope of use or functionality of aspects of the subject matter described herein. Neither should the mobile device 800 be interpreted as having any dependency or requirement relating to any one or combination of components illustrated in the example mobile device 800.

[0048] With reference to FIG. 8, an example device for implementing aspects of the subject matter described herein includes a mobile device 800. In some embodiments, the mobile device 800 comprises a cell phone, a handheld device that allows voice communications with others, some other voice communications device, or the like. In these embodiments, the mobile device 800 may be equipped with a camera for taking pictures, although this may not be required in other embodiments. In other embodiments, the mobile device 800 may comprise a personal digital assistant (PDA), hand-held gaming device, notebook computer, printer, appliance including a set-top, media center, or other appliance, other mobile devices, or the like. In yet other embodiments, the mobile device 800 may comprise devices that are generally considered non-mobile such as personal computers, servers, or the like.

[0049] The mobile device may comprise a hand-held remote control of an appliance or toy, with additional circuitry to provide the control logic along with a way to input data to the remote control. For example, an input jack or other data receiving sensor may allow the device to be repurposed for non-control code data transmission. This may be accomplished without needing to store much of the data to transmit, e.g., the device may act as a data

relay for another device (possibly with some buffering), such as a smartphone.

[0050] Components of the mobile device 800 may include, but are not limited to, a processing unit 805, system memory 810, and a bus 815 that couples various system components including the system memory 810 to the processing unit 805. The bus 815 may include any of several types of bus structures including a memory bus, memory controller, a peripheral bus, and a local bus using any of a variety of bus architectures, and the like. The bus 815 allows data to be transmitted between various components of the mobile device 800.

[0051] The mobile device 800 may include a variety of computer-readable media. Computer-readable media can be any available media that can be accessed by the mobile device 800 and includes both volatile and nonvolatile media, and removable and non-removable media. By way of example, and not limitation, computer-readable media may comprise computer storage media and communication media. Computer storage media includes volatile and nonvolatile, removable and non-removable media implemented in any method or technology for storage of information such as computer-readable instructions, data structures, program modules, or other data. Computer storage media includes, but is not limited to, RAM, ROM, EEPROM, flash memory or other memory technology, CD-ROM, digital versatile disks (DVD) or other optical disk storage, magnetic cassettes, magnetic tape, magnetic disk storage or other magnetic storage devices, or any other medium which can be used to store the desired information and which can be accessed by the mobile device 800.

[0052] Communication media typically embodies computer-readable instructions, data structures, program modules, or other data in a modulated data signal such as a carrier wave or other transport mechanism and includes any information delivery media. The term "modulated data signal" means a signal that has one or more of its characteristics set or changed in such a manner as to encode information in the signal. By way of example, and not limitation, communication media includes wired media such as a wired network or direct-wired connection, and wireless media such as acoustic, RF, Bluetooth®, Wireless USB, infrared, Wi-Fi, WiMAX, and other wireless media. Combinations of any of the above should also be included within the scope of computer-readable media.

[0053] The system memory 810 includes computer storage media in the form of volatile and/or nonvolatile memory and may include read only memory (ROM) and random access memory (RAM). On a mobile device such as a cell phone, operating system code 820 is sometimes included in ROM although, in other embodiments, this is not required. Similarly, application programs 825 are often placed in RAM although again, in other embodiments, application programs may be placed in ROM or in other computer-readable memory. The heap 830 provides memory for state associated with the operating system 820 and the application programs 825. For example, the operating system 820 and application programs 825 may store variables and data structures in the heap 830 during their operations.

[0054] The mobile device 800 may also include other removable/non-removable, volatile/nonvolatile memory. By way of example, FIG. 8 illustrates a flash card 835, a hard disk drive 836, and a memory stick 837. The hard disk drive 836 may be miniaturized to fit in a memory slot, for example. The mobile device 800 may interface with these types of non-volatile removable memory via a removable memory interface 831, or may be connected via a universal serial bus (USB), IEEE 8394, one or more of the wired port(s) 840, or antenna(s) 865. In these embodiments, the removable memory devices 835-437 may interface with the mobile device via the communications module(s) 832. In some embodiments, not all of these types of memory may be included on a single mobile device. In other embodiments, one or more of these and other types of removable memory may be included on a single mobile device.

[0055] In some embodiments, the hard disk drive 836 may be connected in such a way as to be more permanently attached to the mobile device 800. For example, the hard disk drive 836 may be connected to an interface such as parallel advanced technology attachment (PATA), serial advanced technology attachment (SATA) or otherwise, which may be connected to the bus 815. In such embodiments, removing the hard drive may involve removing a cover of the mobile device 800 and removing screws or other fasteners that connect the hard drive 836 to support structures within the mobile device 800.

[0056] The removable memory devices 835-437 and their associated computer storage media, discussed above and illustrated in FIG. 8, provide storage of computer-readable instructions, program modules, data structures, and other data for the mobile device 800. For example, the removable memory device or devices 835-437 may store images taken by the mobile device 800, voice recordings, contact information, programs, data for the programs and so forth.

[0057] A user may enter commands and information into the mobile device 800 through input devices such as a key pad 841 and the microphone 842. In some embodiments, the display 843 may be touch-sensitive screen and may allow a user to enter commands and information thereon. The key pad 841 and display 843 may be connected to the processing unit 805 through a user input interface 850 that is coupled to the bus 815, but may also be connected by other interface and bus structures, such as the communications module(s) 832 and wired port(s) 840. Motion detection 852 can be used to determine gestures made with the device 800.

[0058] A user may communicate with other users via speaking into the microphone 842 and via text messages that are entered on the key pad 841 or a touch sensitive display 843, for example. The audio unit 855 may provide electrical signals to drive the speaker 844 as well as re-

ceive and digitize audio signals received from the microphone 842.

[0059] The mobile device 800 may include a video unit 860 that provides signals to drive a camera 861. The video unit 860 may also receive images obtained by the camera 861 and provide these images to the processing unit 805 and/or memory included on the mobile device 800. The images obtained by the camera 861 may comprise video, one or more images that do not form a video, or some combination thereof.

[0060] The communication module(s) 832 may provide signals to and receive signals from one or more antenna(s) 865. One of the antenna(s) 865 may transmit and receive messages for a cell phone network. Another antenna may transmit and receive Bluetooth® messages. Yet another antenna (or a shared antenna) may transmit and receive network messages via a wireless Ethernet network standard.

[0061] Still further, an antenna provides location-based information, e.g., GPS signals to a GPS interface and mechanism 872. In turn, the GPS mechanism 872 makes available the corresponding GPS data (e.g., time and coordinates) for processing.

[0062] In some embodiments, a single antenna may be used to transmit and/or receive messages for more than one type of network. For example, a single antenna may transmit and receive voice and packet messages.

[0063] When operated in a networked environment, the mobile device 800 may connect to one or more remote devices. The remote devices may include a personal computer, a server, a router, a network PC, a cell phone, a media playback device, a peer device or other common network node, and typically includes many or all of the elements described above relative to the mobile device 800.

[0064] Aspects of the subject matter described herein are operational with numerous other general purpose or special purpose computing system environments or configurations. Examples of well known computing systems, environments, and/or configurations that may be suitable for use with aspects of the subject matter described herein include, but are not limited to, personal computers, server computers, hand-held or laptop devices, multiprocessor systems, microcontroller-based systems, set top boxes, programmable consumer electronics, network PCs, minicomputers, mainframe computers, distributed computing environments that include any of the above systems or devices, and the like.

[0065] Aspects of the subject matter described herein may be described in the general context of computer-executable instructions, such as program modules, being executed by a mobile device. Generally, program modules include routines, programs, objects, components, data structures, and so forth, which perform particular tasks or implement particular abstract data types. Aspects of the subject matter described herein may also be practiced in distributed computing environments where tasks are performed by remote processing devices that are linked through a communications network. In a distributed computing environment, program modules may be located in both local and remote computer storage media including memory storage devices.

[0066] Furthermore, although the term server may be used herein, it will be recognized that this term may also encompass a client, a set of one or more processes distributed on one or more computers, one or more stand-alone storage devices, a set of one or more other devices, a combination of one or more of the above, and the like.

CONCLUSION

[0067] While the invention is susceptible to various modifications and alternative constructions, certain illustrated embodiments thereof are shown in the drawings and have been described above in detail. It should be understood, however, that there is no intention to limit the invention to the specific forms disclosed, but on the contrary, the intention is to cover all modifications and alternative constructions falling within the scope of disclosure as defined by the appended claims.

Claims

1. A method comprising:

   projecting a light pattern towards a scene (104);
   capturing, by a base station, at least one image (332; 442) of the scene (104) illuminated by the light pattern;
   receiving (708), at the base station, image-related data from a satellite device (110; 330; 440), the image-related data comprising data of the scene illuminated by the light pattern, the satellite device being in the illuminated scene;
   correlating an image of the at least one image (332; 442) with an image of the image-related data, or a depth map computed from the at least one image (332; 442) with a depth map computed from the image-related data, to determine pose information of the satellite device (110; 330; 440);
   computing a first set of depth data associated with the at least one image; and
   enhancing (710) the first set of depth data based at least in part upon the image-related data and pose information of the satellite device (110; 330; 440).

2. The method of claim 1, wherein the base station comprises a projector (106; 306) for the projecting.

3. The method of claim 1 wherein enhancing (710) the first set of depth data comprises replacing at least some of the data in the first set with other depth data corresponding to at least part of the image-related

data.

4. The method of claim 1 wherein enhancing (710) the first set of depth data comprises maintaining a first depth map corresponding to the first set of depth data in association with a second depth map corresponding to the image-related data.

5. The method of claim 1 further comprising, computing at least some depth data, including processing one or more actively illuminated images captured by the satellite device (110; 330; 440), or processing one or more actively illuminated stereo images captured by the satellite device, including based upon at least some active illumination emitted from the satellite device, or both processing one or more actively illuminated images captured by the satellite device, and processing one or more actively illuminated stereo images captured by the satellite device, including based upon at least some active illumination emitted from the satellite device.

6. The method of claim 1, wherein correlating an image of the at least one image (332; 442) with an image of the image-related data, or a depth map computed from the at least one image (332; 442) with a depth map computed from the image-related data, comprises performing an image alignment and/or a pattern matching to determine the pose information of the satellite device.

7. One or more computer-readable storage media or logic having computer-executable instructions, which when executed perform a method according to any one of claims 1 to 6.

8. A system comprising:

a projector (406; 606; 607) for projecting a light pattern towards a scene (104); and
a base station (100; 300; 500), the base station configured to:

capture at least one image (332; 442) of the scene illuminated by the light pattern;
receive (708) image-related data from a satellite device (110; 330; 440), the image-related data comprising data of the scene illuminated by the light pattern, the satellite device being in the illuminated scene;
correlate an image of the at least one image (332; 442) with an image of the image-related data, or a depth map computed from the at least one image (332; 442) with a depth map computed from the image-related data, to determine (706) pose information of the satellite device;
compute a first set of depth data associated with the at least one image, and
enhance (710) the first set of depth data based at least in part upon the image-related data and pose information of the satellite device.

9. The system of claim 8 wherein the satellite device (110; 330; 440) is configured to compute (706) depth data and communicate the depth data to the base station as at least part of the image-related data.

Patentansprüche

1. Verfahren, umfassend:

Projizieren eines Lichtmusters in Richtung einer Szene (104);
Aufnehmen, durch eine Basisstation, mindestens eines Bildes (332; 442) der Szene (104), die durch das Lichtmuster beleuchtet wird;
Empfangen (708), an der Basisstation, von bildbezogenen Daten von einer Satellitenvorrichtung (110; 330; 440), wobei die bildbezogenen Daten Daten der Szene umfassen, die von dem Lichtmuster beleuchtet wird, wobei die Satellitenvorrichtung in der beleuchteten Szene ist;
Korrelieren eines Bildes des mindestens einen Bildes (332; 442) mit einem Bild der bildbezogenen Daten, oder einer Tiefenkarte, die aus dem mindestens einen Bild (332; 442) berechnet wurde, mit einer Tiefenkarte, die aus den bildbezogenen Daten berechnet wurde, um Poseninformationen der Satellitenvorrichtung (110; 330; 440) zu bestimmen;
Berechnen eines ersten Satzes von Tiefendaten, die dem mindestens einen Bild zugeordnet sind; und
Verbessern (710) des ersten Satzes von Tiefendaten zumindest teilweise basierend auf den bildbezogenen Daten und Poseninformationen der Satellitenvorrichtung (110; 330; 440).

2. Verfahren nach Anspruch 1, wobei die Basisstation einen Projektor (106; 306) für das Projizieren umfasst.

3. Verfahren nach Anspruch 1, wobei Verbessern (710) des ersten Satzes von Tiefendaten Ersetzen mindestens mancher der Daten in dem ersten Satz mit anderen Tiefendaten, die mindestens einem Teil der bildbezogenen Daten entsprechen, umfasst.

4. Verfahren nach Anspruch 1, wobei Verbessern (710) des ersten Satzes von Tiefendaten Aufrechterhalten einer ersten Tiefenkarte, die dem ersten Satz von Tiefendaten in Zuordnung zu einer zweiten Tiefenkarte, die den bildbezogenen Daten entspricht, um-

fasst.

5. Verfahren nach Anspruch 1, weiter umfassend: Berechnen mindestens mancher Tiefendaten, einschließlich Verarbeiten von einem oder mehreren aktiv beleuchteten Bildern, die von der Satellitenvorrichtung (110; 330; 440) aufgenommen wurden, oder Verarbeiten von einem oder mehreren aktiv beleuchteten Stereobildern, die von der Satellitenvorrichtung aufgenommen wurden, einschließlich, basierend auf mindestens einiger aktiver Beleuchtung, die von der Satellitenvorrichtung emittiert wird, oder beides, Verarbeiten von einem oder mehreren aktiv beleuchteten Bildern, die von der Satellitenvorrichtung aufgenommen wurden, und Verarbeiten von einem oder mehreren aktiv beleuchteten Stereobildern, die von der Satellitenvorrichtung aufgenommen wurden, basierend auf mindestens einiger aktiver Beleuchtung beinhaltend, die von der Satellitenvorrichtung emittiert wird.

6. Verfahren nach Anspruch 1, wobei Korrelieren eines Bildes des mindestens einen Bildes (332; 442) mit einem Bild der bildbezogenen Daten, oder einer Tiefenkarte, die aus dem mindestens einen Bild (332; 442) berechnet wurde, mit einer Tiefenkarte, die aus den bildbezogenen Daten berechnet wurde, Durchführen einer Bildausrichtung und/oder eines Musterabgleichs umfasst, um die Poseninformationen der Satellitenvorrichtung zu bestimmen.

7. Ein oder mehrere computerlesbare Speichermedien oder Logik mit computerausführbaren Anweisungen, die, wenn sie ausgeführt werden, ein Verfahren nach einem der Ansprüche 1 bis 6 durchführen.

8. System, umfassend:

einen Projektor (406; 606; 607) zum Projizieren eines Lichtmusters in Richtung einer Szene (104); und
eine Basisstation (100; 300; 500), wobei die Basisstation für das Folgende konfiguriert ist:

Aufnehmen mindestens eines Bildes (332; 442) der Szene, die von dem Lichtmuster beleuchtet wird;
Empfangen (708) von bildbezogenen Daten von einer Satellitenvorrichtung (110; 330; 440), wobei die bildbezogenen Daten Daten der Szene umfassen, die von dem Lichtmuster beleuchtet wird, wobei die Satellitenvorrichtung in der beleuchteten Szene ist;
Korrelieren eines Bildes des mindestens einen Bildes (332; 442) mit einem Bild der bildbezogenen Daten, oder einer Tiefenkarte, die aus dem mindestens einen Bild (332; 442) berechnet wurde, mit einer Tiefenkarte, die aus den bildbezogenen Daten berechnet wurde, um Poseninformationen der Satellitenvorrichtung zu bestimmen (706);
Berechnen eines ersten Satzes von Tiefendaten, die dem mindestens einen Bild zugeordnet sind, und
Verbessern (710) des ersten Satzes von Tiefendaten zumindest teilweise basierend auf den bildbezogenen Daten und Poseninformationen der Satellitenvorrichtung.

9. System nach Anspruch 8, wobei die Satellitenvorrichtung (110; 330; 440) dafür konfiguriert ist, Tiefendaten zu berechnen (706) und die Tiefendaten an die Basisstation als mindestens Teil der bildbezogenen Daten zu kommunizieren.

Revendications

1. Procédé comprenant :

la projection d'un motif de lumière vers une scène (104) ;
la capture, par une station de base, d'au moins une image (332 ; 442) de la scène (104) illuminée par le motif de lumière ;
la réception (708), au niveau de la station de base, de données relatives à l'image à partir d'un dispositif satellite (110 ; 330 ; 440), les données relatives à l'image comprenant des données de la scène illuminée par le motif de lumière, le dispositif satellite étant dans la scène illuminée ;
la corrélation d'une image de l'au moins une image (332; 442) avec une image des données relatives à l'image, ou d'une carte de profondeur calculée à partir de l'au moins une image (332 ; 442) avec une carte de profondeur calculée à partir des données relatives à l'image, pour déterminer des informations de pose du dispositif satellite (110; 330; 440) ;
le calcul d'un premier ensemble de données de profondeur associées à l'au moins une image ; et
l'amélioration (710) du premier ensemble de données de profondeur sur la base au moins en partie des données relatives à l'image et aux informations de pose du dispositif satellite (110 ; 330 ; 440).

2. Procédé selon la revendication 1, dans lequel la station de base comprend un projecteur (106 ; 306) pour la projection.

3. Procédé selon la revendication 1, dans lequel l'amélioration (710) du premier ensemble de données de

profondeur comprend le remplacement d'au moins certaines des données dans le premier ensemble par d'autres données de profondeur correspondant à au moins une partie des données relatives à l'image.

4. Procédé selon la revendication 1, dans lequel l'amélioration (710) du premier ensemble de données de profondeur comprend le maintien d'une première carte de profondeur correspondant au premier ensemble de données de profondeur en association avec une seconde carte de profondeur correspondant aux données relatives à l'image.

5. Procédé selon la revendication 1, comprenant en outre le calcul d'au moins certaines données de profondeur, incluant le traitement d'une ou plusieurs images illuminées activement capturées par le dispositif satellite (110 ; 330; 440), ou le traitement d'une ou plusieurs images stéréo illuminées activement par le dispositif satellite, y compris sur la base d'au moins une certaine illumination active émise à partir du dispositif satellite, ou à la fois le traitement d'une ou plusieurs images illuminées activement capturées par le dispositif satellite et le traitement d'une ou plusieurs images stéréo illuminées activement par le dispositif satellite, y compris sur la base d'au moins une certaine illumination active émise à partir du dispositif satellite.

6. Procédé selon la revendication 1, dans lequel la corrélation d'une image de l'au moins une image (332 ; 442) avec une image des données relatives à l'image, ou d'une carte de profondeur calculée à partir de l'au moins une image (332 ; 442) avec une carte de profondeur calculée à partir des données relatives à l'image, comprend l'exécution d'un alignement d'image et/ou d'une mise en correspondance de motif pour déterminer les informations de pose du dispositif satellite.

7. Un ou plusieurs supports de stockage lisibles par ordinateur ou logique ayant des instructions exécutables par ordinateur qui, quand elles sont exécutées, effectuent un procédé selon l'une quelconque des revendications 1 à 6.

8. Système comprenant :

un projecteur (406 ; 606 ; 607) pour projeter un motif de lumière vers une scène (104) ; et
une station de base (100 ; 300 ; 500), la station de base étant configurée pour :

capturer au moins une image (332; 442) de la scène illuminée par le motif de lumière ;
recevoir (708) des données relatives à l'image à partir d'un dispositif satellite (110 ; 330 ; 440), les données relatives à l'image comprenant des données de la scène illuminée par le motif de lumière, le dispositif satellite étant dans la scène illuminée ;
mettre en corrélation une image de l'au moins une image (332 ; 442) avec une image des données relatives à l'image, ou une carte de profondeur calculée à partir de l'au moins une image (332 ; 442) avec une carte de profondeur calculée à partir des données relatives à l'image, pour déterminer (706) des informations de pose du dispositif satellite ;
calculer un premier ensemble de données de profondeur associées à l'au moins une image, et
améliorer (710) le premier ensemble de données de profondeur sur la base au moins en partie des données relatives à l'image et aux informations de pose du dispositif satellite.

9. Système selon la revendication 8, dans lequel le dispositif satellite (110 ; 330 ; 440) est configuré pour calculer (706) des données de profondeur et communiquer les données de profondeur à la station de base comme au moins une partie des données relatives à l'image.

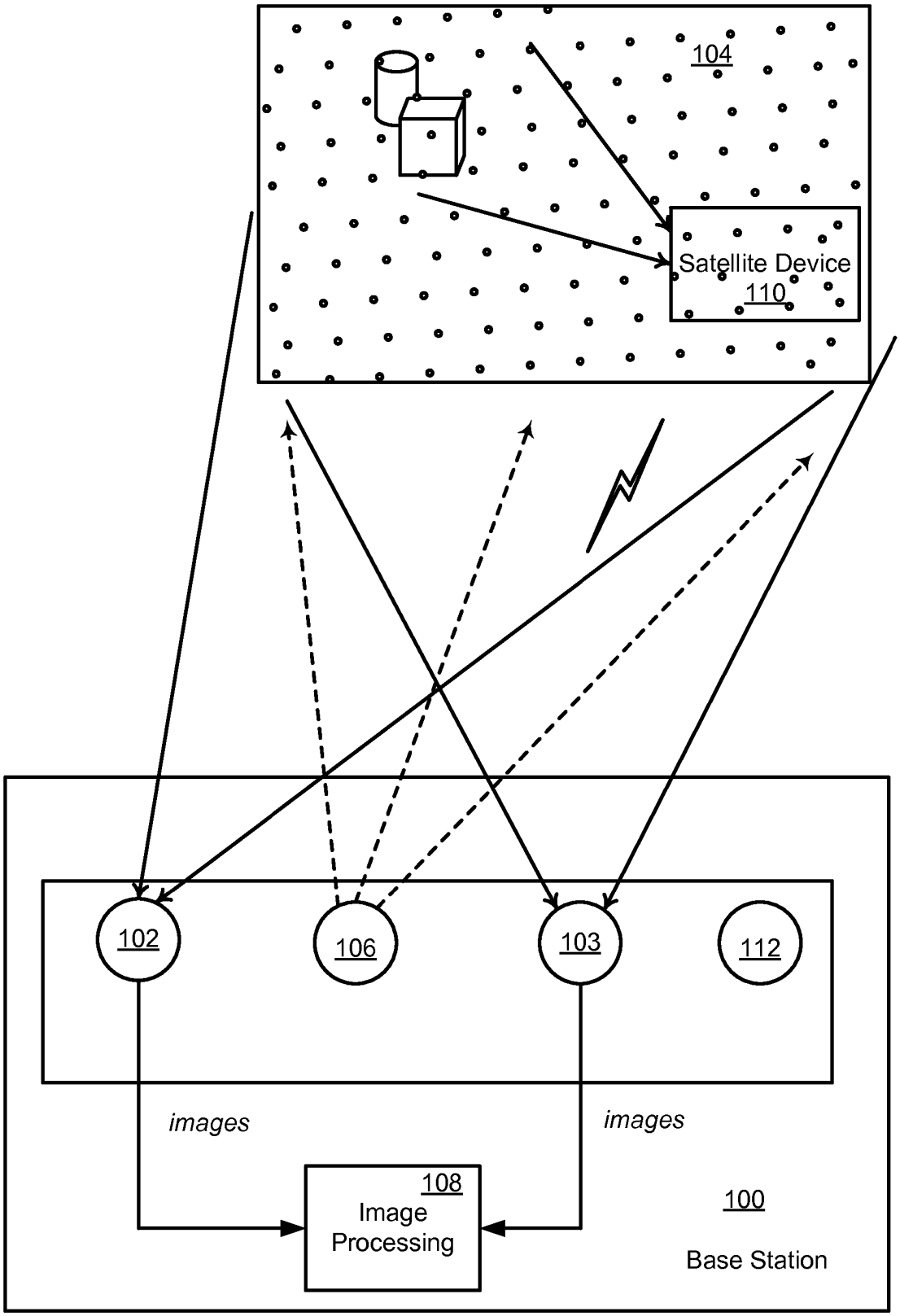
104
Satellite Device
110
102
106
103
112
images
images
108
Image
Processing
100
Base Station

FIG. 1

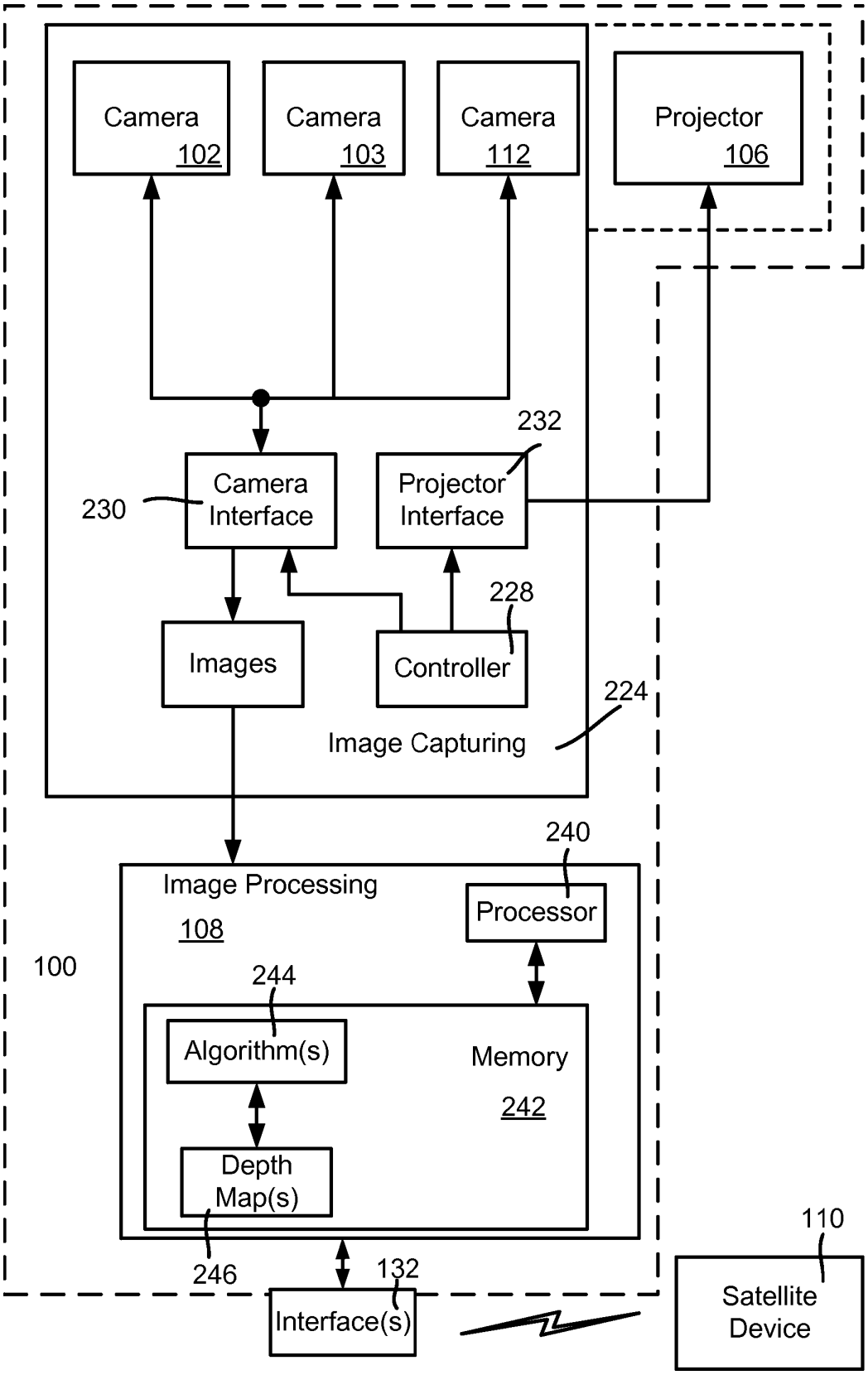
Camera
102
Camera
103
Camera
112
Projector
106
232
Camera
Interface
230
Projector
Interface
228
Images
Controller
224
Image Capturing
240
Image Processing
108
Processor
100
244
Algorithm(s)
Memory
242
Depth
Map(s)
246
132
110
Interface(s)
Satellite
Device

FIG. 2

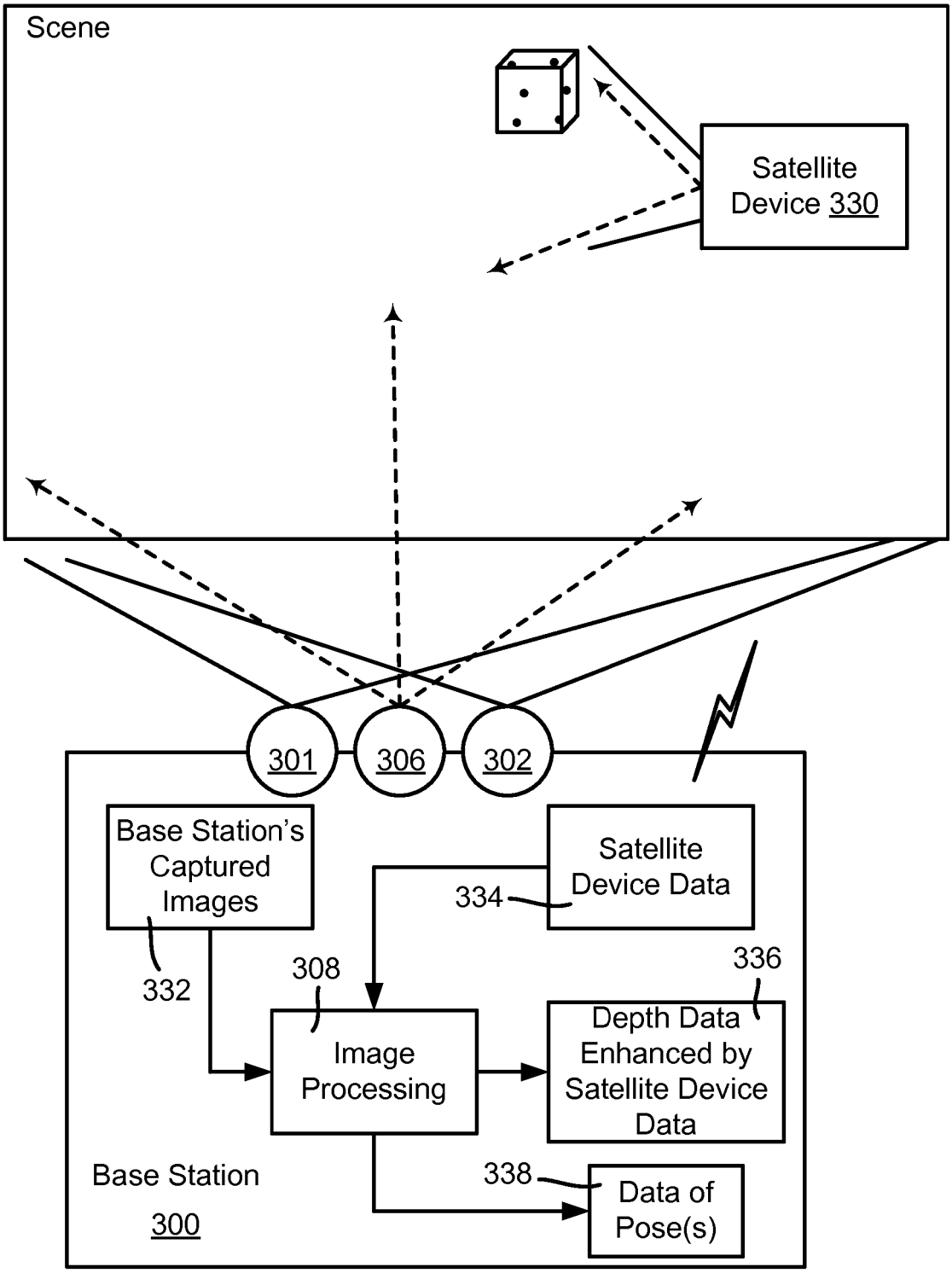
Scene
Satellite Device 330
301
306
302
Base Station's Captured Images
332
Satellite Device Data
334
308
336
Image Processing
Depth Data Enhanced by Satellite Device Data
Base Station
300
338
Data of Pose(s)

FIG. 3

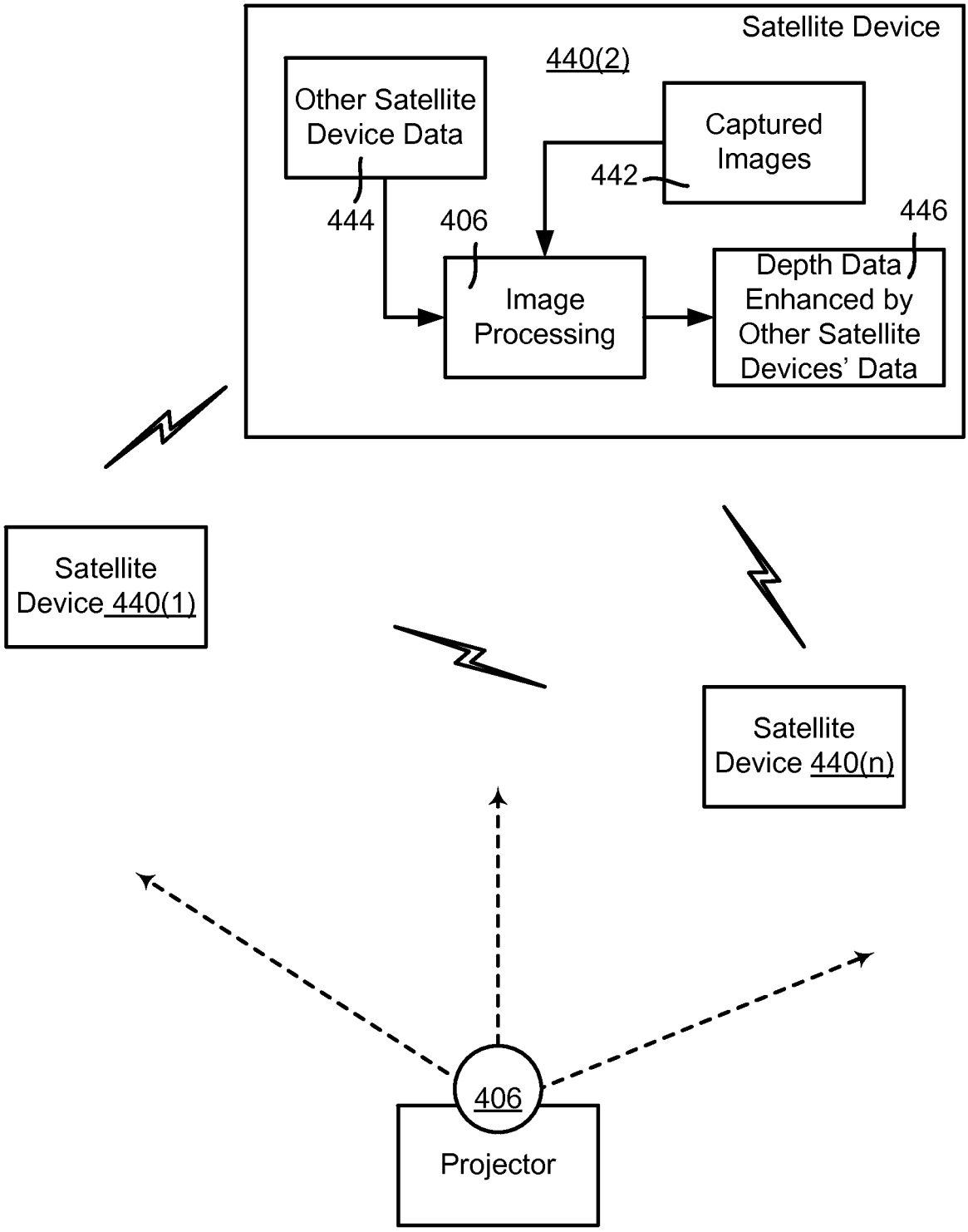
Satellite Device
440(2)
Other Satellite Device Data
Captured Images
442
444
406
446
Image Processing
Depth Data Enhanced by Other Satellite Devices' Data
Satellite Device 440(1)
Satellite Device 440(n)
406
Projector

FIG. 4

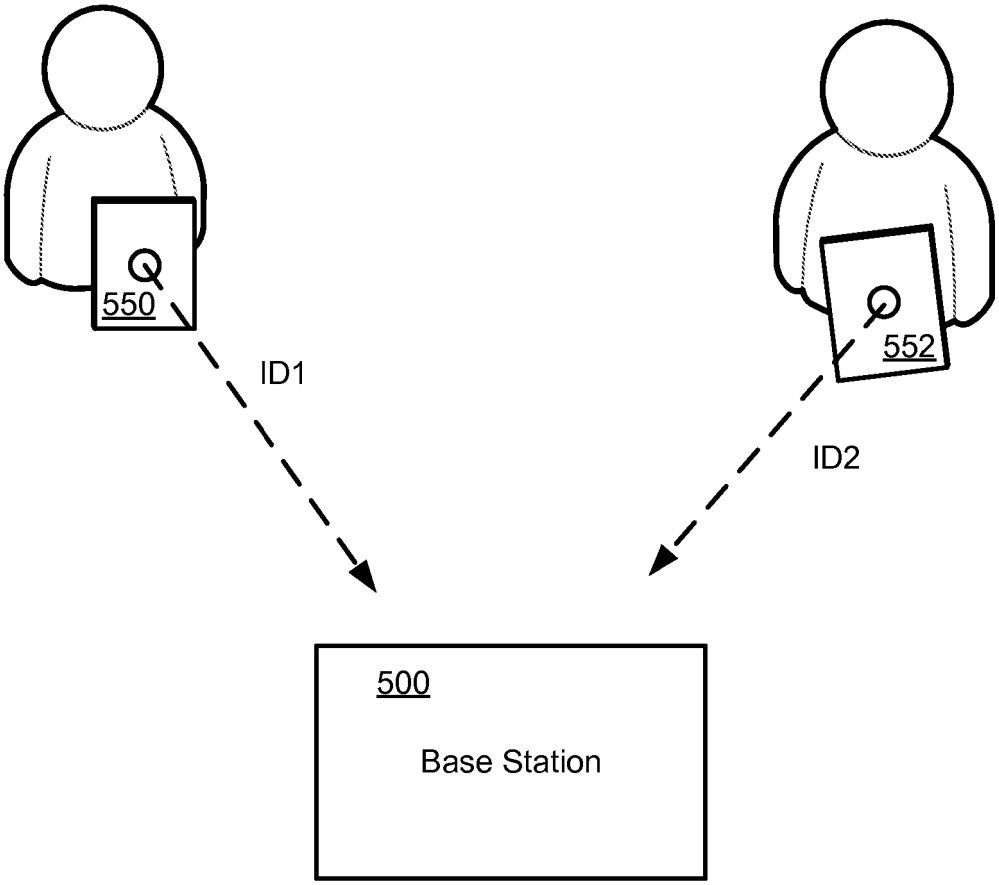

FIG. 5

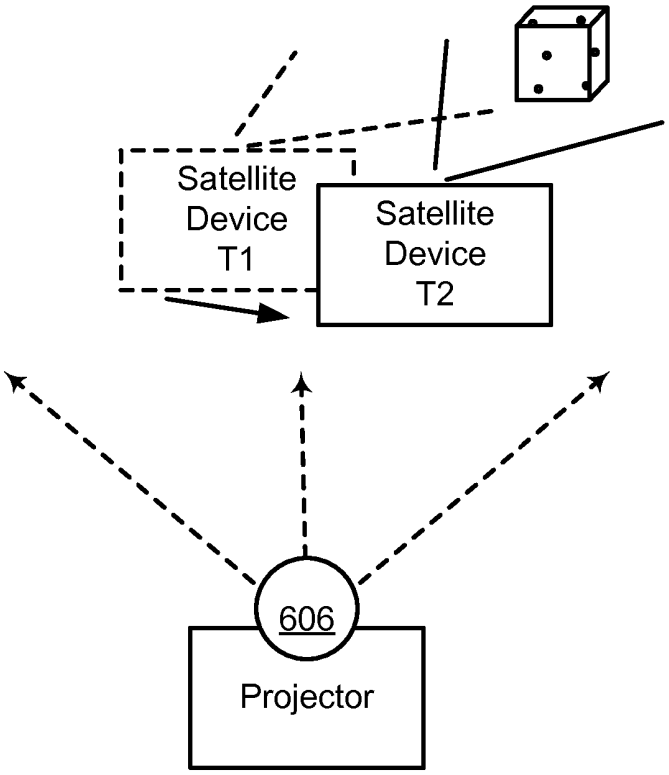
Satellite
Device
T1
Satellite
Device
T2
606
Projector

FIG. 6A

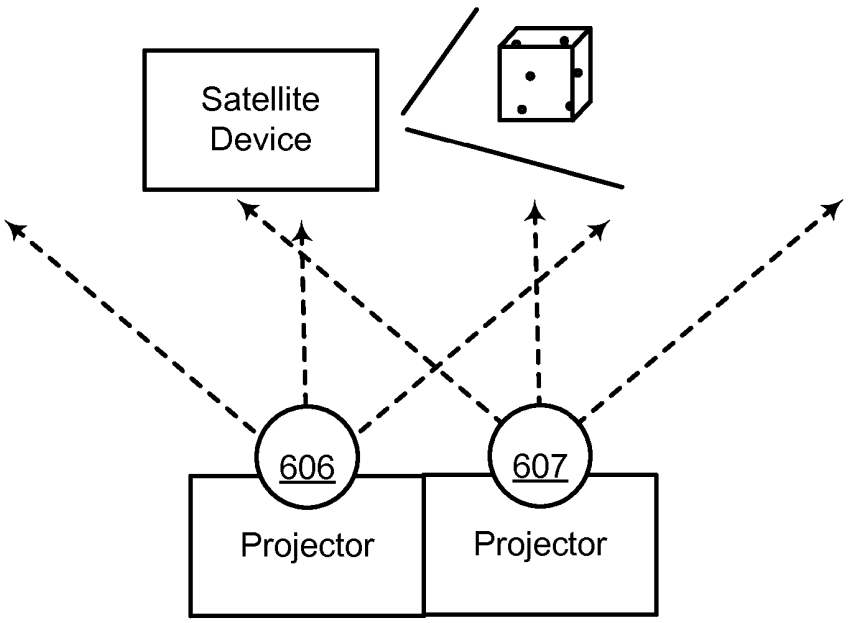
Satellite
Device
606
607
Projector
Projector

FIG. 6B

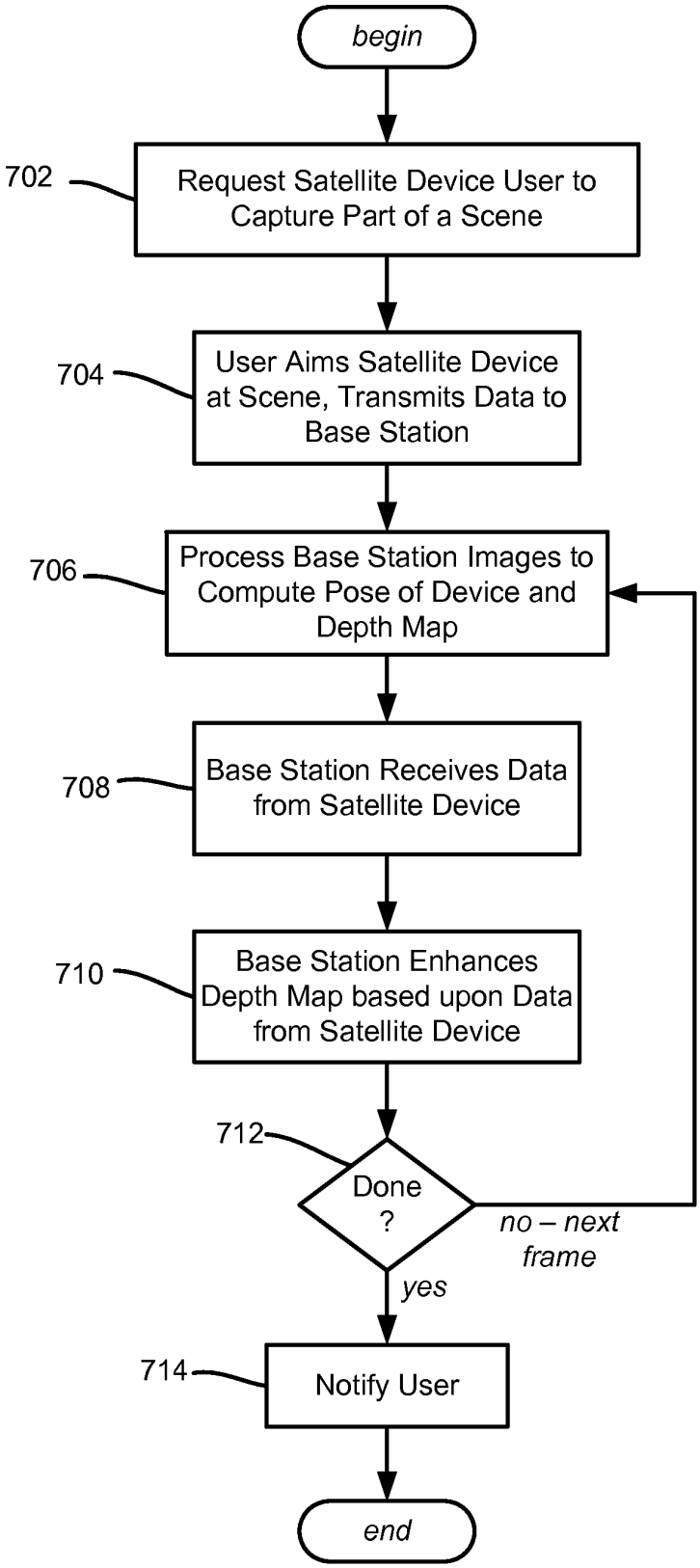
begin
702
Request Satellite Device User to Capture Part of a Scene
704
User Aims Satellite Device at Scene, Transmits Data to Base Station
706
Process Base Station Images to Compute Pose of Device and Depth Map
708
Base Station Receives Data from Satellite Device
710
Base Station Enhances Depth Map based upon Data from Satellite Device
712
Done ?
no – next frame
yes
714
Notify User
end

FIG. 7

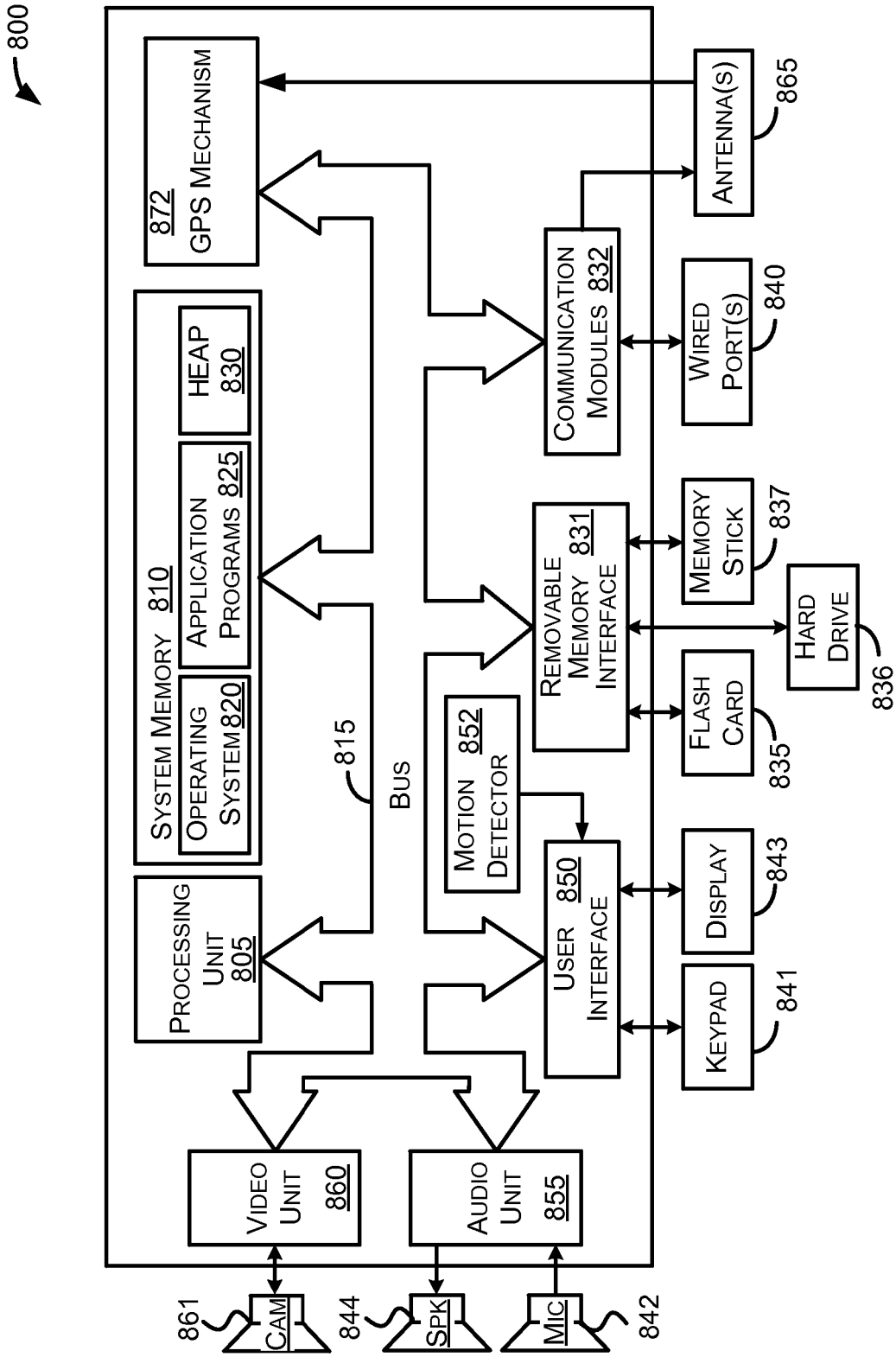
800
872 GPS MECHANISM
ANTENNA(S) 865
SYSTEM MEMORY 810
OPERATING SYSTEM 820
APPLICATION PROGRAMS 825
HEAP 830
COMMUNICATION MODULES 832
WIRED PORT(S) 840
815
BUS
REMOVABLE MEMORY INTERFACE 831
MEMORY STICK 837
HARD DRIVE 836
MOTION DETECTOR 852
FLASH CARD 835
PROCESSING UNIT 805
USER INTERFACE 850
DISPLAY 843
KEYPAD 841
VIDEO UNIT 860
AUDIO UNIT 855
861
CAM
844
SPK
MIC
842

FIG. 8

REFERENCES CITED IN THE DESCRIPTION

This list of references cited by the applicant is for the reader's convenience only. It does not form part of the European patent document. Even though great care has been taken in compiling the references, errors or omissions cannot be excluded and the EPO disclaims all liability in this regard.

Patent documents cited in the description

- US 20120056982 A1, Katz **[0003]**
- US 20020136444 A1, Brown **[0003]**
- US 20130100256 A **[0021]**