(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) **World Intellectual Property Organization**
International Bureau

(43) **International Publication Date**
20 June 2024 (20.06.2024)

WIPO | PCT

(10) **International Publication Number**
WO 2024/129121 A1

(51) **International Patent Classification:**
G06F 3/01 (2006.01)

(21) **International Application Number:**
PCT/US2022/081530

(22) **International Filing Date:**
14 December 2022 (14.12.2022)

(25) **Filing Language:** English

(26) **Publication Language:** English

(71) **Applicant: GOOGLE LLC** [US/US]; 1600 Amphitheatre Parkway, Mountain View, California 94043 (US).

(72) **Inventors: PUROHIT, Aveek**; GOOGLE LLC, 1600 Amphitheatre Parkway, Mountain View, California 94043 (US). **SHIN, Dongeek**; GOOGLE LLC, 1600 Amphitheatre Parkway, Mountain View, California 94043 (US).

(74) **Agent: GUENTHER, Brett** et al.; BRAKE HUGHES BELLERMANN LLP, P.O. Box 1077, Middletown, Maryland 21769 (US).

(81) **Designated States** *(unless otherwise indicated, for every kind of national protection available)*: AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CV, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IQ, IR, IS, IT, JM, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, WS, ZA, ZM, ZW.

(84) **Designated States** *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, CV, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SC, SD, SL, ST, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, ME, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

Declarations under Rule 4.17:
— *of inventorship (Rule 4.17(iv))*

Published:
— *with international search report (Art. 21(3))*

(54) **Title:** GESTURE INPUT FOR HEAD-MOUNTED DEVICE

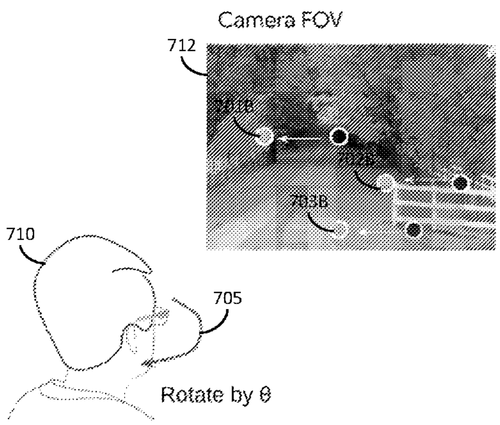

FIG. 7B

(57) **Abstract:** A head-mounted device worn by a user may be configured to recognize a selection of an item in a user interface based on a movement of the head of the user. The recognition may include capturing a plurality of images using a world-facing camera of the head-mounted device. The plurality of images may be analyzed to recognize an object that should be stationary but that appears to be moving in the plurality of images as a result of a head-turn. The movement of the stationary object in the plurality of images may be characterized to detect a head-turn gesture, which can be used to select an item on the user interface.

WO 2024/129121 A1

GESTURE INPUT FOR HEAD-MOUNTED DEVICE

FIELD OF THE DISCLOSURE

[0001] The present disclosure relates to a gesture input for a head-mounted device and more specifically to inverse head-movement gesture based on images captured by the head-mounted device.

BACKGROUND

[0002] A head-mounted device may be configured to provide an augmented-reality (AR) environment. The user may interact with the AR environment using a user interface (UI) presented in a field of view (FOV) of the user. The interaction may require a gesture performed by the user to move a cursor in the UI.

SUMMARY

[0003] In at least one aspect, the present disclosure generally describes a mode of gesture detection for a head-mounted device that does not rely on tactile sensing, eye-tracking, or inertial measurements. Instead, images from a world-facing camera of the head worn device may detect a motion of the head-mounted device by the apparent motion of otherwise stationary landmarks defined in the images. The apparent motion of the stationary landmarks may be analyzed over time (i.e., over images) to detect a head-turn gesture, which may be used alone, or in combination with other gestures/inputs, to interact with the head-mounted device.

[0004] In some aspects, the techniques described herein relate to a method for detecting a gesture including: capturing a first image using a world-facing camera of a head-mounted device; recognizing a stationary object in the first image; identifying a plurality of features on the stationary object in the first image; capturing a second image using the world-facing camera of the head-mounted device; locating the plurality of features in the second image; processing the first image and the second image to determine a plurality of movements of the plurality of features between the first image and the second image; detecting the gesture based on the plurality of movements; and controlling a user-interface based on the gesture.

[0005] In some aspects, the techniques described herein relate to a head-mounted device including: a world-facing camera configured to capture images of an environment from a point-of-view (POV) of a user wearing the head-mounted device; a heads-up display (HUD) configured to present a user-interface to the user wearing the head-mounted device; and a processor configured by software instructions to: receive a plurality of images from the world-facing camera; recognize a stationary object in the plurality of images; determine a movement of the stationary object in the plurality of images; detect a head-turn gesture based on the movement of the stationary object in the plurality of images; and highlighting an item in the user-interface presented on the HUD to indicate a selection of the item based on the head-turn gesture.

[0006] The foregoing illustrative summary, as well as other exemplary objectives and/or advantages of the disclosure, and the manner in which the same are accomplished, are further explained within the following detailed description and its accompanying drawings.

BRIEF DESCRIPTION OF THE DRAWINGS

[0007] FIG. 1A is a user-interface of a head-mounted device having a first item selected based on a first head turn of a user according to a possible implementation of the present disclosure.

[0008] FIG. 1B is the user-interface of a heads-mounted device, as in FIG. 1A, having a second item selected based on a second head turn of a user according to a possible implementation of the present disclosure.

[0009] FIG. 2 illustrates possible gestures for a user wearing a head-mounted device according to possible implementations of the present disclosure.

[0010] FIG. 3 is a perspective view of a head-mounted device according to a possible implementation of the present disclosure.

[0011] FIG. 4 is a system block diagram of a head-mounted device according to a possible implementation of the present disclosure.

[0012] FIG. 5 is a state diagram illustrating possible gesture-detection modes of a head-mounted device according to a possible implementation of the present disclosure.

[0013] FIG. 6 is a flowchart of a method for detecting a gesture for interaction with a user-interface of head-mounted device according to a possible implementation of the present disclosure.

[0014] FIG. 7A illustrates a plurality of features of stationary objects in a first image captured by a head-mounted device according to a possible implementation of the present disclosure.

[0015] FIG. 7B illustrates the apparent movement of the plurality of features of stationary objects in a second captured by a head-mounted device according to a possible implementation of the present disclosure.

[0016] FIG. 8A illustrates a first apparent movement of a stationary object at a first range according to a possible implementation of the present disclosure.

[0017] FIG. 8B illustrates a second apparent movement of a stationary object at a second range according to a possible implementation of the present disclosure.

[0018] The components in the drawings are not necessarily to scale relative to each other. Like reference numerals designate corresponding parts throughout the several views.

DETAILED DESCRIPTION

[0019] A head-mounted device (HMD), such as AR glasses, may be configured with a variety of sensors to monitor a user and an environment of the user to provide an AR environment. The AR environment may include a user interface and interaction with the user interface may require detecting a gesture corresponding to a focus of the user. In some implementations, it may be desirable for the accuracy of gesture detection to be approximately 100%, but variations in the sensor, the environment, and/or the gesture itself may make accurately detecting the gesture difficult.

[0020] An eye-tracking camera of the HMD may capture images of an eye (or eyes) of a user to detect a gesture. For example, a focus (i.e., selection) of the user may be determined by tracking the eye (or eyes) of the user in the captured images. Eye tracking may not be accurate, or may not be possible, for all users, however, for a few reasons. First, operation/appearance of the eye (or eyes) of the user may not accurately track a focus of the user due to a condition of the user (e.g., amblyopia). Second, the eye-tracking camera may not accurately capture images of the eye (or eyes) of the user due to a condition of the image capture (e.g., obscuration). Third, an algorithm may not accurately recognize and convert features (e.g., landmarks) of the eye into a focus position due to a condition of the HMD (e.g., misalignment, bad calibration, etc.). Accordingly, eye-tracking using camera images (i.e., vision) alone may not be useful in all gesturing scenarios.

[0021] An inertial measurement unit (IMU) may capture movement of the head of the user wearing the HMD to detect a gesture. For example, the focus (i.e., selection) of the user may be determined from IMU sensor data if the head of the user can be tracked. Head tracking may not be accurate for all users, however, for a few reasons. For example, when a user is in a moving vehicle, the IMU may measure the movement of both the user and the vehicle. As a result, movements corresponding to gestures may be obscured by other movements. What is more, the sensitivity of the IMU sensors may require a movement of the user performing the gesture to be uncomfortably large for the user. Accordingly, head-tracking using the IMU alone may not be useful in all gesturing scenarios.

[0022] While a combination of eye tracking and head tracking can be used to improve gesture detection, even this combination may not be useful in all gesturing scenarios. For example, IMU/vision gesture sensing may require calibration, which can be affected by changes in the HMD over time and/or its environment. The present disclosure describes a gesture sensing approach that does not require an IMU or an eye tracking camera and can therefor replace or supplement the gesture detection approaches (i.e., gesture-detection modes) described above. The disclosed approach may have the technical advantage of improving an overall accuracy of gesture detection, which can improve an AR experience for a user.

[0023] As part of the AR experience, a UI may be presented in a FOV of a user on a heads-up display (i.e., HUD). The UI may include controls (e.g., buttons) that may be interacted with to configure a function or response in the AR experience. Interacting with a control may include selecting the control (e.g., highlighting the control) and then activating the control (e.g., pressing the control). The present disclosure describes techniques for interacting with the control that includes a gesture made by the user turning his/her head (i.e., head-turn gesture).

[0024] FIG. 1A is a user-interface of a head-mounted device having a first item selected based on a first head turn of a user according to a possible implementation of the present disclosure. The user interface 100 can be displayed in a heads-up display shown to a user wearing the head-mounted device and can include messages for conveying information from the HMD to the user and controls for conveying information from the user the HMD. As shown, the messages can include a clock 110A indicating the time and a prompt 110B requesting input from a user. As shown, the controls can include a microphone control 111, which when activated, can be configured the HMD to collect sounds from a microphone of the HMD. The control can further include a camera control, which when activated can configure the HMD to collect images from a camera of the HMD. Either control can be selected (or no control can be selected) by a gesture. Because only two items may be selected, the UI of FIG. 1A may be said to have a discrete cursor that can move between one of two positions.

[0025] FIG. 1A includes a highlighted camera control 114 (i.e., selected camera control) indicating that the camera control is selected. While various highlighting techniques may be used, the camera control shown in FIG. 1A is surrounded by a circular graphic to indicate that it is selected. The selection of the camera control is prompted by a head-turn of a user 101 wearing the HMD in a first direction 102 towards the camera control in the UI.

[0026] FIG. 1B includes a highlighted microphone control 112 (i.e., selected microphone control) and a camera control 113 that is not highlighted indicating that the microphone control is selected, and the camera control is not selected. The selection of the microphone control is prompted by a head-turn of a user 101 wearing the HMD in a second direction 103 towards the microphone control in the UI.

[0027] FIG. 2 illustrates possible gestures for a user wearing a head-mounted device according to possible implementations of the present disclosure. FIG. 2 illustrates a user wearing a head-mounted device 210 on their head 200. The head-mounted device may be configured to detect a head-turn gesture corresponding to movements of the head 200 of the user.

[0028] The head-turn gesture may include a direction. For example, a head-turn gesture may be rotation around a first axis 201, which can be sensed as a UP-to-DOWN head-turn gesture or as a DOWN-to-UP head-turn gesture. In another example, a head-turn gesture may be rotation around a second axis 202, which can be sensed (i.e., detected) as a RIGHT-to-LEFT head-turn gesture or as a LEFT-to-RIGHT head-turn gesture. Other directions exists (e.g., UPPER-RIGHT-to-LOWER-LEFT, UPPER-LEFT-to-LOWER-RIGHT, LOWER-RIGHT-to-UPPER-LEFT, LOWER-LEFT-to-UPPER-RIGHT, etc.) and a head-turn gesture may be defined for any of these directions. Further, a head-turn gesture may be a combination of these directions. The direction of a head-turn gesture may be aligned with a direction on the user-interface to facilitate the interaction illustrated by FIGS. 1A-1B.

[0029] The head-turn gesture may include a magnitude. For example, a first angle of rotation around the first axis 201 may correspond to a magnitude of an UP/DOWN head-turn gesture, and a second angle of rotation around the second axis 202 may correspond to a magnitude of a RIGHT/LEFT head-turn gesture, where larger angles correspond to larger head-turn gestures. Detection of a head-turn gesture may include comparing a magnitude of the head-turn gesture to a threshold. For example, when the magnitude of the head-turn gesture is above a threshold then a head-turn gesture may be detected, otherwise the head-turn gesture may not be detected. The threshold may be settable and/or adjustable. For example, a user may initially train a threshold as part of a gesture-detection setup process in which the user is prompted to perform head-turn gestures to set the threshold level (i.e., sensitivity) of the head-turn gesture detection. Additionally, or alternatively the user may be provided with a control to adjust the threshold level (i.e., sensitivity) of the head-turn gesture detection.

[0030] A detected head-turn gesture may be combined with another gesture to activate a function of the HMD. For example, while the head-turn gesture may be sufficient to select a control on a user-interface another gesture may be necessary to activate the selected control. In a possible implementation, a tap gesture may be performed while an item (i.e., control) is highlighted (i.e., selected) in the UI and then activating the highlighted item based on the tap gesture. As shown in FIG. 2, a tap gesture may include a user momentarily pressing a finger 220 to a side 221 of the head-mounted device 210.

[0031] FIG. 3 is a perspective view of a head-mounted device according to a possible implementation of the present disclosure. As shown, the head-mounted device may be implemented as smart glasses configured for augmented reality (i.e., augmented reality (AR) glasses). The AR glasses 300 can be configured to be worn on the head and face of a user. The AR glasses 300 include a right earpiece 301 and a left earpiece 302 that are supported by the ears of a user. The AR glasses 300 further include a bridge portion 303 that is supported by the nose of the user so that a left lens 304 and a right lens 305 can be positioned in front a left eye of the user and a right eye of the user respectively. The portions of the AR glasses 300 can be collectively referred to as the frame of the AR glasses. The frame of the AR glasses can contain electronics to enable function. For example, the frame may include a battery, a processor, a memory (e.g., non-transitory computer readable medium), electronics to support sensors (e.g., cameras, depth sensors, etc.), at least one position sensor (e.g., an inertial measurement unit) and interface devices (e.g., speakers, display, network adapter, etc.). The AR glasses may display and sense an environment relative to a coordinate system 330. The coordinate system 330 can be aligned with the head of a user wearing the AR glasses. For example, the eyes of the user may be along a line in a horizontal (e.g., LEFT/RIGHT, X-axis) direction of the coordinate system 330.

[0032] A user wearing the AR glasses 300 can experience information displayed in an area corresponding to the lens (or lenses) so that the user can view virtual elements within their natural field of view. Accordingly, the AR glasses 300 can further include a heads-up display (i.e., HUD) configured to display visual information at a lens (or lenses) of the AR glasses. As shown, the heads-up display may present AR data (e.g., images, graphics, text, icons, etc.) on a portion 315 of a lens (or lenses) of the AR glasses so that a user may view the AR data as the user looks through a lens of the AR glasses. In this way, the AR data can overlap with the user's view of the environment. In a possible implementation, the portion 315 can correspond to (i.e., substantially match) area(s) of the right lens 305 and/or left lens 304.

[0033] The AR glasses 300 can include an IMU that is configured to track motion of the head of a user wearing the AR glasses. The IMU may be disposed within the frame of the AR glasses and aligned with the coordinate system 330 of the AR glasses 300.

[0034] The AR glasses 300 can include a first camera 310 that is directed to a first camera field-of-view that overlaps with the natural field-of-view of the eyes of the user when the glasses are worn. In other words, the first camera 310 (i.e., world-facing camera) can capture images of a view aligned with a point-of-view (POV) of a user (i.e., an egocentric view of the user).

[0035] In a possible implementation, the AR glasses 300 can further include a depth sensor 311. The depth sensor 311 may be implemented as a second camera that is directed to a second field-of-view that overlaps with the natural field-of-view of the eyes of a user when the glasses are worn. The second camera and the first camera 310 may be configured to capture stereoscopic images of the field of view of the user that include depth information about objects in the field of view of the user. The depth information may be generated using visual odometry and used as part of the camera measurement corresponding to the motion of the head-mounted device. In other implementations the depth sensor 311 can be implemented as another type of depth (i.e., range) sensing device, including (but not limited to) a structured light depth sensor or a lidar depth sensor. The depth sensor 311 can be configured to capture a depth image corresponding to the field-of-view of the user. The depth image includes pixels having pixel values that correspond to depths (i.e., ranges) to objects measured at positions corresponding to the pixel positions in the depth image.

[0036] In a possible implementation, the AR glasses 300 can further include an illuminator 312 to help the imaging and/or depth sensing. For example, the illuminator 312 can be implemented as an infra-red (IR) projector configured to transmit IR light (e.g., near-infra-red light) into the environment of the user to help the first camera 310 capture images and/or the depth sensor 311 to determine a range of an object.

[0037] The AR glasses 300 can further include an eye-tracking sensor. The eye tracking sensor can include a right-eye camera and/or a left-eye camera 321. As shown, a left-eye camera 321 can be located in a portion of the frame so that a left FOV 323 of the left-eye camera 321 includes the left eye of the user when the AR glasses are worn.

[0038] The AR glasses 300 can further include one or more microphones. The one or more microphones can be spaced apart on the frames of the AR glasses. As shown in FIG. 3, the AR glasses can include a first microphone 331 and a second microphone 332. The microphones may be configured to operate together as a microphone array. The microphone array can be configured to apply sound localization to determine directions of the sounds relative to the AR glasses.

[0039] The AR glasses may further include a left speaker 341 and a right speaker 342 configured to transmit audio to the user. Additionally, or alternatively, transmitting audio to a user may include transmitting the audio over a wireless communication link 345 to a listening device (e.g., hearing aid, earbud, etc.). For example, the AR glasses may transmit audio to a left wireless earbud 346 and to a right earbud 347.

[0040] FIG. 4 is a system block diagram of a head-mounted device according to a possible implementation of the present disclosure. As mentioned, the head-mounted device 400 (i.e., head-worn device) may be implemented as smart glasses worn by a user. In a possible implementation, the smart glasses may provide a user with (and enable a user to interact with) an augmented reality (AR) environment. In these implementations, the smart glasses may be referred to as AR glasses. While AR glasses are not the only possible head-mounted device that can be implemented using the disclosed systems and methods (e.g., virtual reality headset), the disclosure may refer to the AR glasses implementation as the head-mounted device throughout the disclosure.

[0041] The head-mounted device 400 may be worn on the head of a user (i.e., wearer) and can be configured to monitor a position and orientation (i.e., pose) of the head of the user. Additionally, the head-mounted device 400 may be configured to monitor the environment of the user. The head-mounted device may be further configured to determine a frame coordinate system based on the pose of the user and a world coordinate system based on the environment of the user. The relationship between the frame coordinate system and the world coordinate system may be used to visually anchor digital objects to real objects in the environment. The digital objects may be displayed to a user in a heads-up display. For these functions the head-mounted device 400 may include a variety of sensors and subsystems.

[0042] The head-mounted device 400 may include a world-facing camera 410. The world-facing camera (i.e., front-facing camera) may be configured to capture images of a front field-of-view 415. The front field-of-view 415 may be aligned with a user's field of view so that the front-facing camera captures images from a point-of-view of the user. The camera may be a color camera based on a charge coupled device (CCD) or complementary metal oxide semiconductor (CMOS) sensor.

[0043] The head-mounted device 400 may further include an eye-tracking camera 420 (or cameras). The eye tracking camera may be configured to capture images of an eye field-of-view 425. The eye field of view 425 may be aligned with an eye of the user so that the eye-tracking camera 420 captures images of the user's eye. The eye images may be analyzed to determine a gaze direction of a user, which may be included in analysis to refine the frame coordinate system to better align with a direction in which the user is looking. When the head-mounted device 400 is implemented as smart glasses, the eye-tracking camera 420 may be integrated with a portion of a frame of the glasses surrounding a lens to directly image the eye. In a possible implementation the head-mounted device 400 includes an eye-tracking camera 420 for each eye and a gaze direction may be based on the images of both eyes.

[0044] The head-mounted device 400 may further include a depth sensor 416 (i.e., range detector) configured to measure a range of objects in (at least) the front-field-of-view 415 and an illuminator configured to transmit light (e.g., near infra-red light) into (at least) the front field-of-view 415 to aid function of the world-facing camera 410 and/or the depth sensor 416.

[0045] The head-mounted device 400 may further include location sensors 430. The location sensors may be configured to determine a position of the head-mounted device (i.e., the user) on the planet, in a building (or other designated area), or relative to another device. For this, the head-mounted device 400 may communicate with other devices over a wireless communication link 435. For example, a user's position may be determined within a building based on communication between the head-mounted device 400 and a wireless router 431 or indoor positioning unit. In another example, a user's position may be determined on the planet based on a global positioning system (GPS) link between the head-mounted device 400 and a GPS satellite 432 (e.g., a plurality of GPS satellites). In another example, a user's position may be determined relative to a device (e.g., mobile phone 433) based on ultra-wide band (UWB) communication or Bluetooth communication between the mobile phone 433 and the location sensors 430.

[0046] The head-mounted device 400 may further include a display 440. For example, the display 440 may be a heads-up display (i.e., HUD) displayed on a portion (e.g., the entire portion) of a lens of AR glasses. In a possible implementation, a projector positioned in a frame arm of the glasses may project light to a surface of a lens, where it is reflected to an eye of the user. In another possible implementation, the head-mounted device 400 may include a display for each eye.

[0047] The head-mounted device 400 may further include a battery 450. The battery may be configured to provide energy to the subsystems, modules, and devices of the head-mounted device 400 to enable their operation. The battery may be rechargeable and have an operating life (e.g., lifetime) between charges. The head-mounted device 400 may include circuitry or software to monitor a battery level of the battery 450.

[0048] The head-mounted device 400 may further include a communication interface 460. The communication interface may be configured to communicate information digitally over a wireless communication link (e.g., WiFi, Bluetooth, etc.). For example, the head-mounted device 400 may be communicatively coupled to a network 461 (i.e., the cloud) or a device (e.g., the mobile phone 433) over a wireless communication link 435. The wireless communication link may allow operations of a computer-implemented method to be divided between devices (i.e., split-processing). Additionally, the communication link may allow a device to communicate a condition, such as a battery level or a power mode (e.g., low-power mode). In this regard, devices communicatively coupled to the head-mounted device 400 may be considered as accessories to the head-mounted device 400 and therefore each may be referred to as an accessory device. In a possible implementation, an accessory device (e.g., mobile phone 433, tablet) may be configured to detect a gesture and then communicate this detection to the head-mounted device 400 to trigger a response from the head-mounted device 400.

[0049] The head-mounted device 400 may further include a memory 470. The memory may be a non-transitory computer readable medium (i.e., CRM). The memory may be configured to store a computer program product. The computer program can instruct a processor to perform computer implemented methods (i.e., computer programs). These computer programs (also known as modules, programs, software, software applications or code) include machine instructions for a programmable processor and can be implemented in a high-level procedural and/or object-oriented programming language, and/or in assembly/machine language. As used herein, the terms “machine-readable medium” or “computer-readable medium” refers to any computer program product, apparatus and/or device (e.g., magnetic discs, optical disks, memory, Programmable Logic Devices (PLDs)) used to provide machine instructions and/or data to a programmable processor, including a machine-readable medium that receives machine instructions as a machine-readable signal. The term “machine-readable signal” refers to any signal used to provide machine instructions and/or data to a programmable processor.

[0050] The head-mounted device 400 may further include a processor 480. The processor may be configured to carry out instructions (e.g., software, applications, etc.) to configure the functions of the head-mounted device 400. In a possible implementation the processor may include multiple processor cores. In a possible implementation, the head-mounted computing device may include multiple processors. In a possible implementation, processing for the head-mounted computing device may be carried out over a network 461.

[0051] The head-mounted device 400 may further include an inertial measurement unit (IMU 490). The IMU may include a plurality of sensor modules to determine its position, orientation, and/or movement. The IMU may have a frame coordinate system (X,Y,Z) and each sensor module may output values relative to each direction of the frame coordinate system.

[0052] The head-mounted device 400 may further include a light sensor 499 configured to measure an ambient light level. In a possible implementation the light sensor 499 output may trigger a low-light condition when the measured ambient light is at or below a low-light threshold. In a possible implementation, the low-light condition can trigger the illuminator 417.

[0053] The head-mounted device may be configured to detect gestures based on images of the environment from a point-of-view (POV) of the user captured by the world-facing camera. This mode of gesture detection may be referred to as a world-image gesture-detection mode. This mode of gesture detection may not require eye-tracking or inertial measurements, which may be advantageous in certain circumstances. This mode of gesture detection, however, may not be optimal for all operating conditions. Accordingly, in a possible implementation, the head-mounted device may be configured to detect gestures in different modes based on conditions.

[0054] FIG. 5 is a state diagram illustrating possible gesture-detection modes of a head-mounted device according to a possible implementation of the present disclosure. The head-mounted device may be configured to change between three gesture-detection modes. The head-mounted device may be configured in an eye-tracking gesture-detection mode 520. In this mode, the positions of the eye (or eyes) may be sensed and measured to perform a selection of an item in a user interface (e.g., see FIGS. 1A-1B).

[0055] In the eye-tracking gesture-detection mode 520, a tracking condition 515 (i.e., quality) of the eye-tracking may be monitored. If the tracking condition indicates that eye tracking is inaccurate/impossible then a gesture-detection of the head-mounted device may change from the eye-tracking gesture-detection mode 520 to the world-image gesture-detection mode 510. For example, this change may occur when the eye-tracking camera is blocked/damaged, when a user's eye is patched, or the like. When the tracking condition 515 is restored (e.g., eye-tracking has an accuracy above a threshold), then the head-mounted device may change its gesture-detection from the world-image gesture-detection mode 510 to the eye-tracking gesture-detection mode 520.

[0056] In the world-image gesture-detection mode 510, a low-light condition 525 may be monitored. For example, an ambient light sensor may be used to detect when the ambient light is below a threshold. Alternatively, images captured by the world facing camera may be analyzed to detect that low-light condition 525 exists. When the low-light condition 525 is detected, then the head-mounted device may change the gesture-detection mode from the world-image gesture-detection mode 510 to an IMU gesture-detection mode 530. In this mode, head-turn gestures may be based on a plurality of movements measured by the IMU. The lighting is restored above the low-light condition (i.e., made brighter) then the head-mounted device may change its gesture-detection from the IMU gesture-detection mode 530 to the world-image gesture-detection mode.

[0057] FIG. 6 is a flowchart of a method for detecting a gesture for interaction with a user-interface of head-mounted device according to a possible implementation of the present disclosure. For example, the method shown in FIG 6 may be implemented by the head-mounted device while in the world-image gesture-detection mode 510.

[0058] The method 600 includes capturing 610 a plurality of images using a world-facing camera of the head-mounted device. The plurality of images may be images captured in a sequence of a video stream. For example, a first image may be captured at a first time and a second image may be captured at a second time that is after the first time. The time between the images may be selected based on an expected rate of movement corresponding to a head-turn gesture so that the images capture the motion.

[0059] The method 600 further includes recognizing 620 a stationary object in the plurality of images. Recognizing the stationary object may include processing the images to recognize 621 (i.e., classify) objects. For example, an image of an outdoor environment may be analyzed to recognize objects as buildings, street signs, cars, people, etc., while an image of an indoor environment may be analyzed to recognize objects as chairs, sofas, people, etc.. The recognition of an object may help classify the object as likely moving or likely not moving. For example, a person recognized in an image may be determines as likely moving based on the person's stance (e.g., walking stance) in the image, while a street signal may be determined as likely not moving (i.e., stationary) inherently. In a possible implementation, the images may be segmented 622 to remove portions that include objects likely moving. For example, a road recognized in the images may be removed because all objects in the road are likely moving. This may make recognizing a stationary object (or objects) in the plurality of images easier.

[0060] In a possible implementation, the recognition uses a neural network trained to recognize objects in images. What is more, it is also possible to turn the head movement estimation problem into a supervised learning problem. In a possible implementation, a first image taken at a first time (i.e., image_(t)) and a second image captured at a second time (i.e., image_(t+N)) are inputs to the neural network, with the output being a direct estimation of the rotational angle of a head-turn gesture. While possible, this approach can require a large amount of training data as the differential image properties will depend completely on the scene itself. Accordingly, it may be advantageous in some implementations to simplify the movement determination by identifying features (i.e., landmarks) in the image and tracking their relative movement (in pixels) over time (i.e., between images).

[0061] The method 600 further includes determining 630 (apparent) movement of the recognized stationary object in the plurality of images. This may include analyzing 631 the images to locate pixel positions of plurality of features (e.g., edges, points, corners, shapes, patterns, words etc.) of the stationary object(s) in a first image at a first time and locating pixel positions of the plurality of features of the stationary objects(s) in a second image at a second time and determining a head movement based on the relative change in the locations (i.e., pixel positions between the images).

[0062] FIG. 7A illustrates a plurality of features of stationary objects in a first image captured by a head-mounted device according to a possible implementation of the present disclosure. A first image 711 captured by a head-mounted device while the head 710 of a user (i.e., wearer) is in a first position. The first image 711 is taken at a first time (t1). A plurality of stationary landmarks are located in the first image 711. For example, a tree is recognized in the first image 711, and a first feature (e.g., bark pattern) on the tree is located at a first pixel location 701A in the first image 711. A fence is recognized in the first image 711, and a second feature (e.g., top of post) on the fence is located at a second pixel location 702A in the first image 711. A road is recognized in the first image 711, and a third feature (e.g., edge) on the road is located at a third pixel location 703A in the first image 711.

[0063] FIG. 7B illustrates the apparent movement of the plurality of features of stationary objects in a second captured by a head-mounted device according to a possible implementation of the present disclosure. A second image 712 captured by a head-mounted device while the head 710 of a user (i.e., wearer) is in a second position. The second position is rotated by an angle (θ) relative to the first position. The second image 712 is taken at a second time (t2), which is a period after the first time (t1). The plurality of stationary landmarks are located in the second image 712 and compared to their corresponding locations in the first image 711. .For example, the first feature (e.g., bark pattern) on the tree is identified and determined to be at a first pixel location 701B in the second image 712. The second feature (e.g., top of post) on the fence is identified and determined to be at a second pixel location 702B in the second image 712. The third feature (e.g., edge) on the road is identified and determined to be at a third pixel location 703B in the second image 712.

[0064] A first optical-flow vector may be computed between the first pixel location 701A in the first image and the first pixel location 701B in the second image. A second optical-flow vector may be computed between the second pixel location 702A in the first image and the second pixel location 702B in the second image. A third optical-flow vector may be computed between the third pixel location 703A in the first image and the third pixel location 703B in the second image.

[0065] The optical-flow vector can collectively define a movement. Each optical-flow vector may have a magnitude measured as a pixel length and a direction measured as angle relative to an axis, or edge (e.g., horizontal, vertical), of the images. A head-turn gesture may be detected based on the magnitude (e.g., average magnitude) of the optical-flow vectors. A head-turn gesture may also be detected based on a direction (e.g., average direction) of the optical-flow vectors. The direction of the optical-flow vectors is opposite a movement direction 705 of the head 710 of the user. Detection of the head-turn gesture may also be aided by additional analysis of the optical-flow vectors. For example, variation between the optical-flow vectors may help to determine a head-turn gesture from other movement.

[0066] Returning to FIG. 6, the method 600 further includes detecting 640 a head-turn gesture based on the (apparent) movement of the stationary object in the plurality of images. This may include determining the magnitude and direction of the movement, as described above. The gesture may be determined by comparing 642 the magnitude to a threshold (e.g., 10 pixels) and detecting the gesture based on the comparison. As mentioned, the threshold may be adjusted. Accordingly, in a possible implementation the method 600 further includes adjusting the sensitivity 643 of the detection (i.e., adjusting the level of the threshold).

[0067] The magnitude of the apparent movement of a stationary object measured in the images may be related to the range between the head-mounted device and the (stationary) object in the image. Accordingly, detecting the head-turn gesture may include estimating 641 a range to the stationary object and then adjusting the threshold for detection and/or the magnitude of the movement based on the range.

[0068] FIG. 8A illustrates a first apparent movement of a stationary object at a first range according to a possible implementation of the present disclosure. A field of view in a first position 801A captures an image of a first object 821 at a first range 811. When the field of view is rotated to a second position 801B the first object 821 appears to move by a first distance 831.

[0069] FIG. 8B illustrates a second apparent movement of a stationary object at a second range according to a possible implementation of the present disclosure. A field of view in a first position 802A captures an image of a second object 822 at a first range 812. When the field of view is rotated to a second position 802B the second object 822 appears to move by a second distance 832.

[0070] The second distance 832 is larger than the first distance 831 even though the rotations of the fields-of-view are the same. Accordingly, precisely determining a head-turn magnitude may require knowledge of the range of the stationary object used to measure the head-turn. It should be noted that if the first distance 831 and the second distance 832 are both greater than a threshold then a head-turn gesture could be detected without knowing first distance 831 or second distance 832 exactly.

[0071] Returning to FIG. 6, the method includes highlighting 650 (i.e., selecting) an item (e.g., control) in the user interface to indicate a selection of the item based on the head-turn gesture. In a possible implementation, the method 600 may further include receiving 660 an additional gesture (e.g., a tap gesture) while the item is highlighted in the user-interface to activate the item (e.g., press a button).

[0072] In the following, some examples of the disclosure are described.

[0073] Example 1. A method for detecting a gesture comprising: capturing a first image using a world-facing camera of a head-mounted device; recognizing a stationary object in the first image; identifying a plurality of features on the stationary object in the first image; capturing a second image using the world-facing camera of the head-mounted device; locating the plurality of features in the second image; processing the first image and the second image to determine a plurality of movements of the plurality of features between the first image and the second image; detecting the gesture based on the plurality of movements; and controlling a user-interface based on the gesture.

[0074] Example 2. The method as in example 1, where detecting the gesture based on the plurality of movements of the plurality of features between the first image and the second image includes: determining a magnitude of a head-turn gesture based on the plurality of movements.

[0075] Example 3. The method as in example 2, where determining the magnitude of the head-turn gesture based on the plurality of movements includes: estimating a range to the stationary object.

[0076] Example 4. The method as in example 2, where detecting the gesture based on the plurality of movements of the plurality of features between the first image and the second image includes: comparing the magnitude of the head-turn gesture to a threshold to obtain a comparison; and detecting the head-turn gesture based on the comparison.

[0077] Example 5. The method as in example 2, where detecting the gesture based on the plurality of movements of the plurality of features between the first image and the second image includes: receiving a sensitivity adjustment from a user; adjusting a threshold based on the sensitivity adjustment to obtain an adjusted threshold; comparing the magnitude of the head-turn gesture to the adjusted threshold to obtain a comparison; and detecting the head-turn gesture based on the comparison.

[0078] Example 6. The method as in example 1, where processing the first image and the second image to determine the plurality of movements of the plurality of features includes: generating optical-flow vectors based on the plurality of features in the first image and the plurality of features in the second image; and determining a head-turn gesture based on the optical-flow vectors, the head-turn gesture including a direction of the head-turn gesture and a magnitude of the head-turn gesture.

[0079] Example 7. The method as in example 6, where controlling the user-interface based on the gesture includes: relating the head-turn gesture to a plurality of controls on the user-interface; and selecting a control of the plurality of controls based on the head-turn gesture.

[0080] Example 8. The method as in example 7, where selecting the control of the plurality of controls based on the gesture includes: generating a graphic to generate a highlighted control in the user-interface.

[0081] Example 9. The method as in example 8, further including: activating the highlighted control by tapping the head-mounted device.

[0082] Example 10. The method as in example 1, where processing the first image and the second image to determine a plurality of movement of the plurality of features includes: inputting the first image and the second image to a neural network configured to output an estimate of a head-turn gesture including a direction and a magnitude.

[0083] Example 11. The method as in example 1, where processing the first image and the second image to determine the plurality of movements of the plurality of features between the first image and the second image includes: transmitting the plurality of features of the first image and the second image from the head-mounted device to a computing device to determine the plurality of movements at the computing device.

[0084] Example 12. The method as in example 1, where recognizing the stationary object in the first image includes: identifying a plurality of objects in the first image; classifying the plurality of objects to recognize objects that are likely moving; segmenting the first image to create a segmented image in which portions of the first image including the objects recognized as likely moving are removed; and recognizing the stationary object in the segmented image.

[0085] Example 13. A head-mounted device comprising: a world-facing camera configured to capture images of an environment from a point-of-view (POV) of a user wearing the head-mounted device; a heads-up display (HUD) configured to present a user-interface to the user wearing the head-mounted device; and a processor configured by software instructions to: receive a plurality of images from the world-facing camera; recognize a stationary object in the plurality of images; determine a movement of the stationary object in the plurality of images; detect a head-turn gesture based on the movement of the stationary object in the plurality of images; and highlighting an item in the user-interface presented on the HUD to indicate a selection of the item based on the head-turn gesture.

[0086] Example 14. The head-mounted device as in example 13, further including an inertial measurement unit (IMU) and an ambient-light detector, wherein the processor is configured to: detecting a low-light condition based on an output of the ambient-light detector; change a gesture-detection mode of the head-mounted device from a world-image gesture-detection mode to an IMU gesture-detection mode based on the low-light condition; and detect the head-turn gesture in the IMU gesture-detection mode based on a plurality of movements measured by the IMU.

[0087] Example 15. The head-mounted device as in example 13, further including an eye-tracking camera, wherein the processor is configured to: detecting a tracking condition based on an output of the eye-tracking camera; change a gesture-detection mode of the head-mounted device from a world-image gesture-detection mode to an eye-tracking gesture-detection mode based on the tracking condition; and detect the head-turn gesture in the eye-tracking gesture-detection mode based on a plurality of eye movements measured by the eye-tracking camera.

[0088] Example 16. The head-mounted device as in example 13, where the head-mounted device includes a depth sensor configured to detect a range of the stationary object, wherein to detect the head-turn gesture based on the movement, the processor is configured to: determine a magnitude of the head-turn gesture based on the movement of the stationary object in the plurality of images and the range of the stationary object.

[0089] Example 17. The head-mounted device as in example 13, where the processor is further configured to: receive a tap gesture while the item is highlighted in the user-interface; and activate the item highlighted in the user-interface based on the tap gesture.

[0090] Example 18. The head-mounted device as in example 13, where to detect the head-turn gesture based on the movement of the stationary object in the plurality of images, the processor is further configured to: determine an amplitude of the head-turn gesture based on the movement of the stationary object; compare the amplitude of the head-turn gesture to a threshold to obtain a comparison; and detect the head-turn gesture based on the comparison.

[0091] Example 19. The head-mounted device as in example 18, where the threshold is adjustable by the user of the head-mounted device.

[0092] Example 20. The head-mounted device as in example 13, where the head-mounted device is augmented-reality glasses.

[0093] In the specification and/or figures, typical embodiments have been disclosed. The present disclosure is not limited to such exemplary embodiments. The use of the term "and/or" includes any and all combinations of one or more of the associated listed items. The figures are schematic representations and so are not necessarily drawn to scale. Unless otherwise noted, specific terms have been used in a generic and descriptive sense and not for purposes of limitation.

[0094] Some implementations may be implemented using various semiconductor processing and/or packaging techniques. Some implementations may be implemented using various types of semiconductor processing techniques associated with semiconductor substrates including, but not limited to, for example, Silicon (Si), Gallium Arsenide (GaAs), Gallium Nitride (GaN), Silicon Carbide (SiC) and/or so forth.

[0095] While certain features of the described implementations have been illustrated as described herein, many modifications, substitutions, changes and equivalents will now occur to those skilled in the art. It is, therefore, to be understood that the appended claims are intended to cover all such modifications and changes as fall within the scope of the implementations. It should be understood that they have been presented by way of example only, not limitation, and various changes in form and details may be made. Any portion of the apparatus and/or methods described herein may be combined in any combination, except mutually exclusive combinations. The implementations described herein can include various combinations and/or sub-combinations of the functions, components and/or features of the different implementations described.

[0096] It will be understood that, in the foregoing description, when an element is referred to as being on, connected to, electrically connected to, coupled to, or electrically coupled to another element, it may be directly on, connected or coupled to the other element, or one or more intervening elements may be present. In contrast, when an element is referred to as being directly on, directly connected to or directly coupled to another element, there are no intervening elements present. Although the terms directly on, directly connected to, or directly coupled to may not be used throughout the detailed description, elements that are shown as being directly on, directly connected or directly coupled can be referred to as such. The claims of the application, if any, may be amended to recite exemplary relationships described in the specification or shown in the figures.

[0097] As used in this specification, a singular form may, unless definitely indicating a particular case in terms of the context, include a plural form. Spatially relative terms (e.g., over, above, upper, under, beneath, below, lower, and so forth) are intended to encompass different orientations of the device in use or operation in addition to the orientation depicted in the figures. In some implementations, the relative terms above and below can, respectively, include vertically above and vertically below. In some implementations, the term adjacent can include laterally adjacent to or horizontally adjacent to.

CLAIMS

1. A method for detecting a gesture comprising:
capturing a first image using a world-facing camera of a head-mounted device;
recognizing a stationary object in the first image;
identifying a plurality of features on the stationary object in the first image;
capturing a second image using the world-facing camera of the head-mounted device;
locating the plurality of features in the second image;
processing the first image and the second image to determine a plurality of movements of the plurality of features between the first image and the second image;
detecting the gesture based on the plurality of movements; and
controlling a user-interface based on the gesture.

2. The method according to claim 1, wherein detecting the gesture based on the plurality of movements of the plurality of features between the first image and the second image includes:
determining a magnitude of a head-turn gesture based on the plurality of movements.

3. The method according to claim 2, wherein determining the magnitude of the head-turn gesture based on the plurality of movements includes:
estimating a range to the stationary object.

4. The method according to claim 2 or 3, wherein detecting the gesture based on the plurality of movements of the plurality of features between the first image and the second image includes:
comparing the magnitude of the head-turn gesture to a threshold to obtain a comparison; and
detecting the head-turn gesture based on the comparison.

5. The method according to claim 2 or 3, wherein detecting the gesture based on the plurality of movements of the plurality of features between the first image and the second image includes:
receiving a sensitivity adjustment from a user;
adjusting a threshold based on the sensitivity adjustment to obtain an adjusted threshold;
comparing the magnitude of the head-turn gesture to the adjusted threshold to obtain a comparison; and
detecting the head-turn gesture based on the comparison.

6. The method according to any of claims 1 to 5, wherein processing the first image and the second image to determine the plurality of movements of the plurality of features includes:
generating optical-flow vectors based on the plurality of features in the first image and the plurality of features in the second image; and
determining a head-turn gesture based on the optical-flow vectors, the head-turn gesture including a direction of the head-turn gesture and a magnitude of the head-turn gesture.

7. The method according to claim 6, wherein controlling the user-interface based on the gesture includes:
relating the head-turn gesture to a plurality of controls on the user-interface; and
selecting a control of the plurality of controls based on the head-turn gesture.

8. The method according to claim 7, wherein selecting the control of the plurality of controls based on the gesture includes:
generating a graphic to generate a highlighted control in the user-interface.

9. The method according to claim 8, further including:
activating the highlighted control by tapping the head-mounted device.

10. The method according to any of claims 1 to 9, wherein processing the first image and the second image to determine a plurality of movement of the plurality of features includes:
inputting the first image and the second image to a neural network configured to output an estimate of a head-turn gesture including a direction and a magnitude.

11. The method according to any of claims 1 to 10, wherein processing the first image and the second image to determine the plurality of movements of the plurality of features between the first image and the second image includes:
transmitting the plurality of features of the first image and the second image from the head-mounted device to a computing device to determine the plurality of movements at the computing device.

12. The method according to any of claims 1 to 11, wherein recognizing the stationary object in the first image includes:
identifying a plurality of objects in the first image;
classifying the plurality of objects to recognize objects that are likely moving;
segmenting the first image to create a segmented image in which portions of the first image including the objects recognized as likely moving are removed; and
recognizing the stationary object in the segmented image.

13. A head-mounted device comprising:

a world-facing camera configured to capture images of an environment from a point-of-view (POV) of a user wearing the head-mounted device;

a heads-up display (HUD) configured to present a user-interface to the user wearing the head-mounted device; and

a processor configured by software instructions to:

receive a plurality of images from the world-facing camera;

recognize a stationary object in the plurality of images;

determine a movement of the stationary object in the plurality of images;

detect a head-turn gesture based on the movement of the stationary object in the plurality of images; and

highlighting an item in the user-interface presented on the HUD to indicate a selection of the item based on the head-turn gesture.

14. The head-mounted device according to claim 13, further including an inertial measurement unit (IMU) and an ambient-light detector, wherein the processor is configured to:

detecting a low-light condition based on an output of the ambient-light detector;

change a gesture-detection mode of the head-mounted device from a world-image gesture-detection mode to an IMU gesture-detection mode based on the low-light condition; and

detect the head-turn gesture in the IMU gesture-detection mode based on a plurality of movements measured by the IMU.

15. The head-mounted device according to claim 13 or 14, further including an eye-tracking camera, wherein the processor is configured to:

detecting a tracking condition based on an output of the eye-tracking camera;

change a gesture-detection mode of the head-mounted device from a world-image gesture-detection mode to an eye-tracking gesture-detection mode based on the tracking condition; and

detect the head-turn gesture in the eye-tracking gesture-detection mode based on a plurality of eye movements measured by the eye-tracking camera.

16. The head-mounted device according to any of claims 13 to 15, wherein the head-mounted device includes a depth sensor configured to detect a range of the stationary object, wherein to detect the head-turn gesture based on the movement, the processor is configured to:
determine a magnitude of the head-turn gesture based on the movement of the stationary object in the plurality of images and the range of the stationary object.

17. The head-mounted device according to any of claims 13 to 16, wherein the processor is further configured to:
receive a tap gesture while the item is highlighted in the user-interface; and
activate the item highlighted in the user-interface based on the tap gesture.

18. The head-mounted device according to any of claims 13 to 17, wherein to detect the head-turn gesture based on the movement of the stationary object in the plurality of images, the processor is further configured to:
determine an amplitude of the head-turn gesture based on the movement of the stationary object;
compare the amplitude of the head-turn gesture to a threshold to obtain a comparison; and
detect the head-turn gesture based on the comparison.

19. The head-mounted device according to claim 18, wherein the threshold is adjustable by the user of the head-mounted device.

20. The head-mounted device according to any of claims 13 to 19, wherein the head-mounted device is augmented-reality glasses.

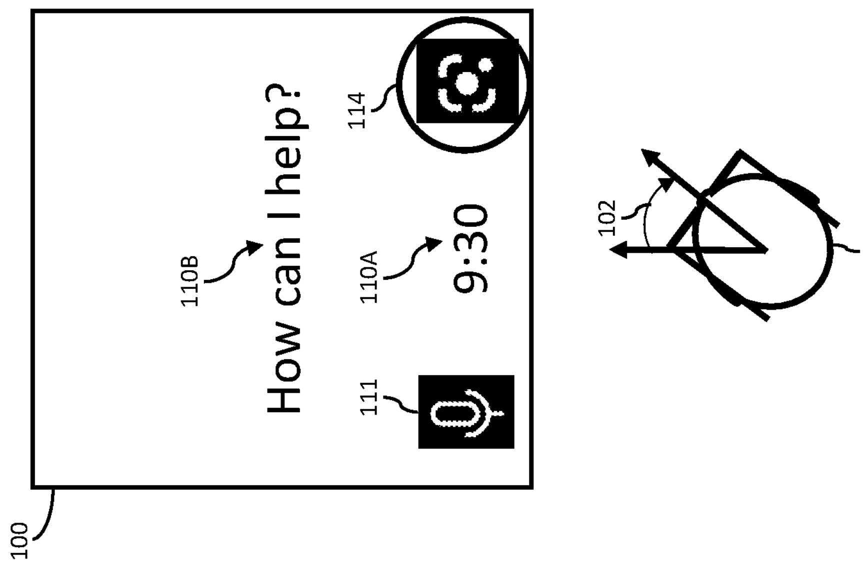
100
How can I help?
110B
9:30
110A
111
114
102
101

FIG. 1A

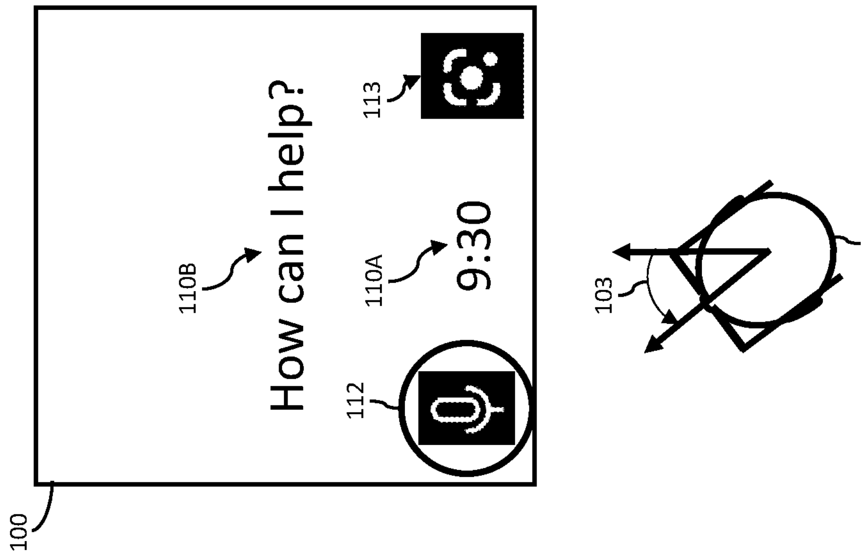
100
How can I help?
110B
9:30
110A
112
113
103
101

FIG. 1B

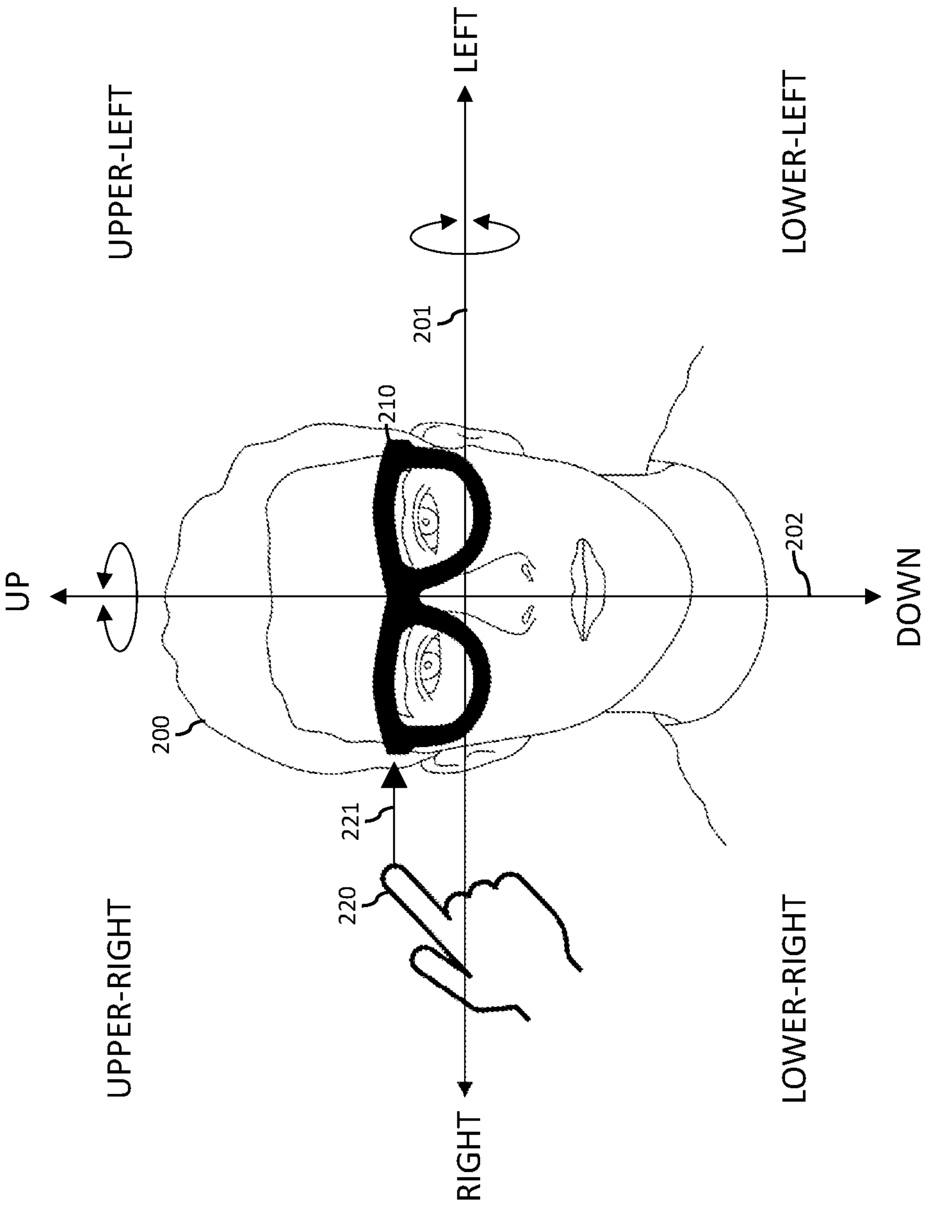
UP
DOWN
LEFT
RIGHT
UPPER-LEFT
LOWER-LEFT
UPPER-RIGHT
LOWER-RIGHT
200
201
202
210
220
221

FIG. 2

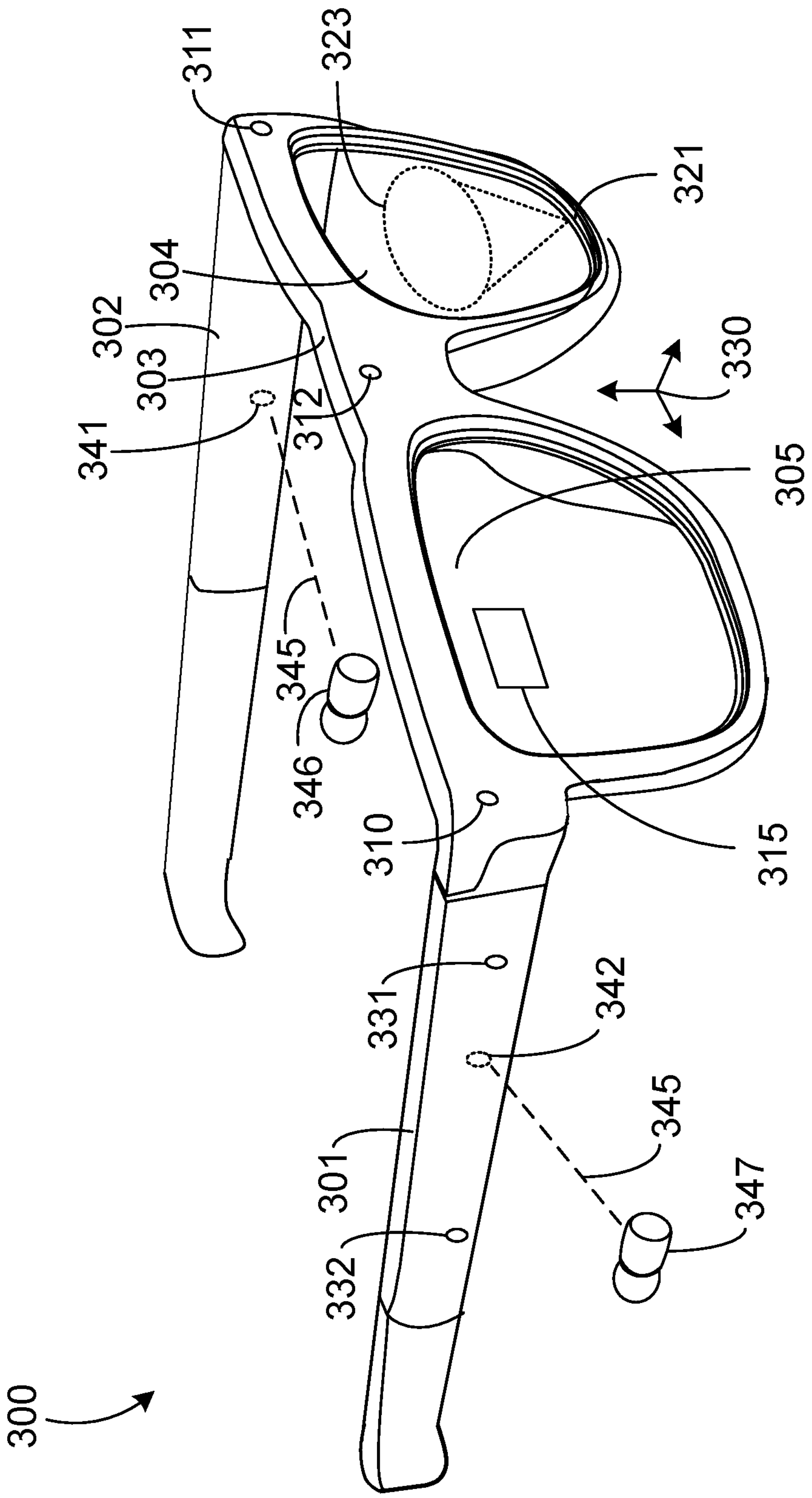
311
323
321
304
302
303
330
341
312
305
345
346
310
315
331
342
345
301
347
332
300

FIG. 3

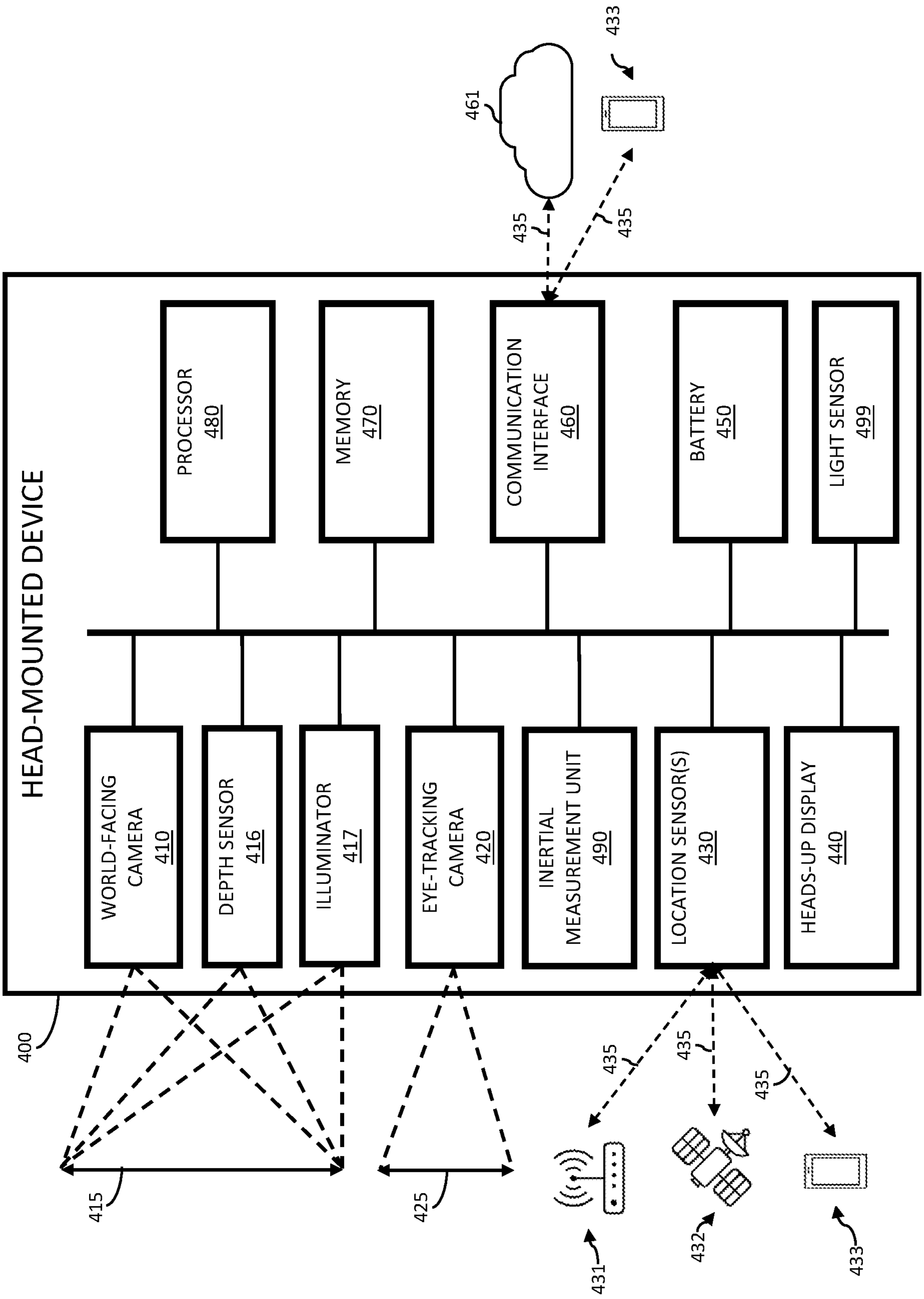
HEAD-MOUNTED DEVICE
PROCESSOR 480
MEMORY 470
COMMUNICATION INTERFACE 460
BATTERY 450
LIGHT SENSOR 499
WORLD-FACING CAMERA 410
DEPTH SENSOR 416
ILLUMINATOR 417
EYE-TRACKING CAMERA 420
INERTIAL MEASUREMENT UNIT 490
LOCATION SENSOR(S) 430
HEADS-UP DISPLAY 440
400
415
425
431
432
433
435
461

FIG. 4

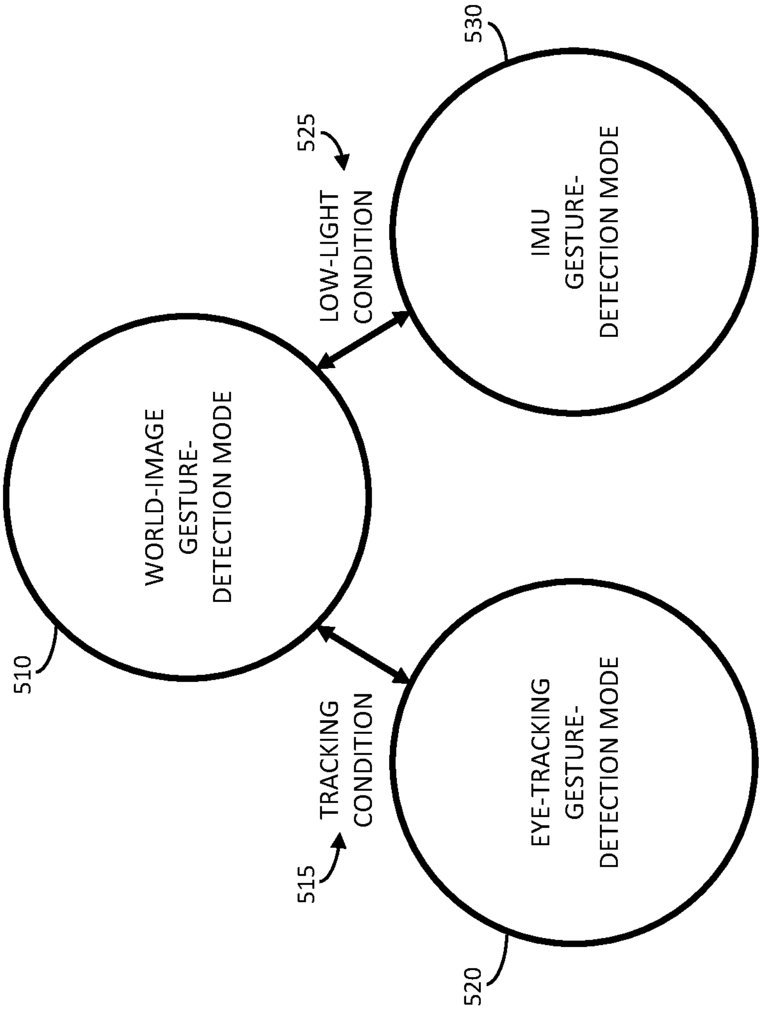
530
525
LOW-LIGHT CONDITION
IMU GESTURE-DETECTION MODE
WORLD-IMAGE GESTURE-DETECTION MODE
510
TRACKING CONDITION
515
EYE-TRACKING GESTURE-DETECTION MODE
520

FIG. 5

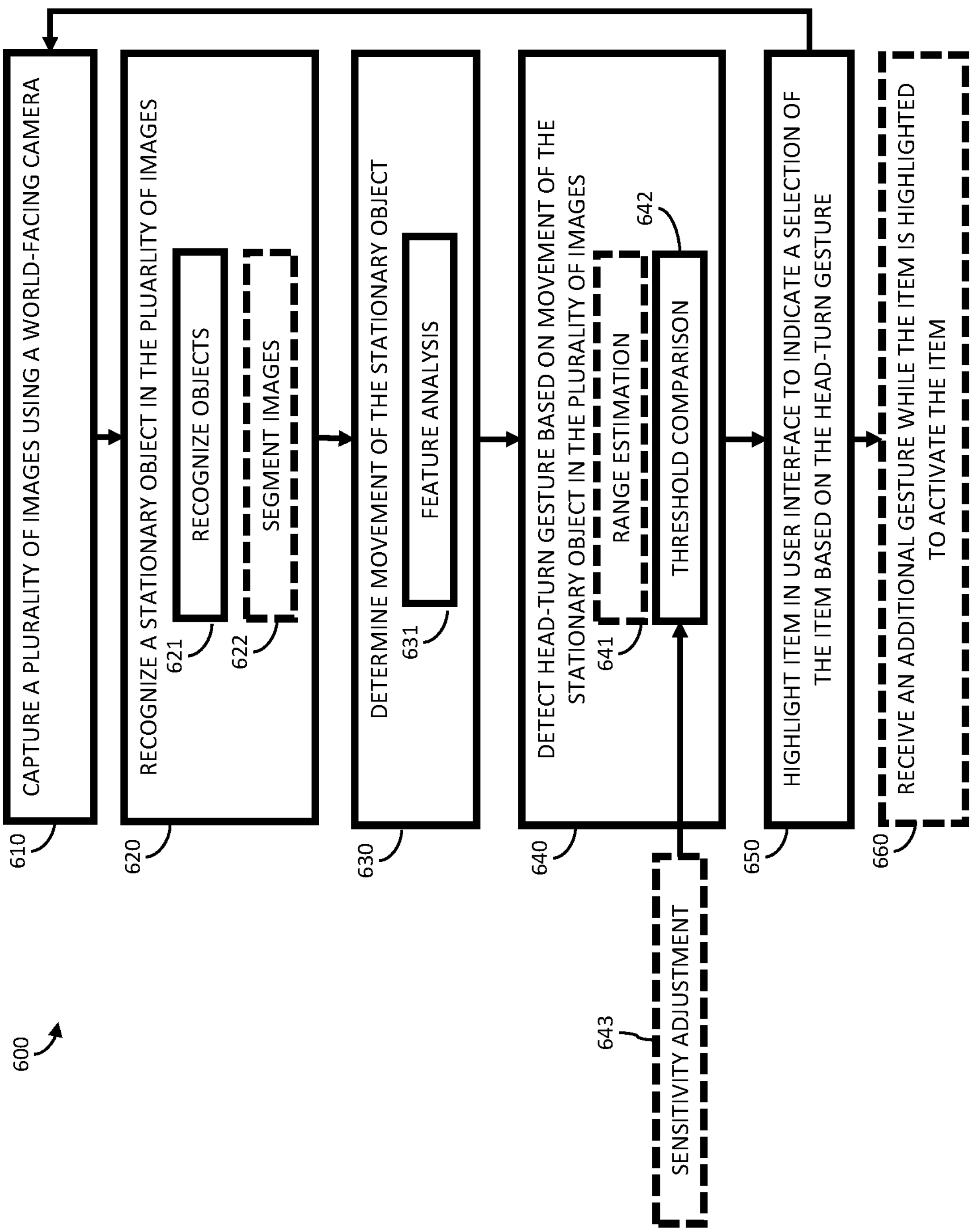
600
610
CAPTURE A PLURALITY OF IMAGES USING A WORLD-FACING CAMERA
620
RECOGNIZE A STATIONARY OBJECT IN THE PLUARLITY OF IMAGES
621
RECOGNIZE OBJECTS
622
SEGMENT IMAGES
630
DETERMINE MOVEMENT OF THE STATIONARY OBJECT
631
FEATURE ANALYSIS
640
DETECT HEAD-TURN GESTURE BASED ON MOVEMENT OF THE STATIONARY OBJECT IN THE PLURALITY OF IMAGES
641
RANGE ESTIMATION
642
THRESHOLD COMPARISON
643
SENSITIVITY ADJUSTMENT
650
HIGHLIGHT ITEM IN USER INTERFACE TO INDICATE A SELECTION OF THE ITEM BASED ON THE HEAD-TURN GESTURE
660
RECEIVE AN ADDITIONAL GESTURE WHILE THE ITEM IS HIGHTLIGHTED TO ACTIVATE THE ITEM

FIG. 6

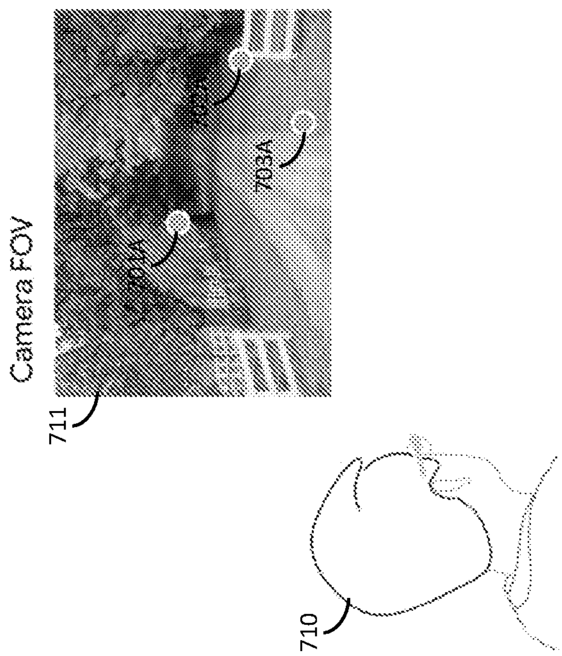
Camera FOV
711
701A
702A
703A
710

FIG. 7A

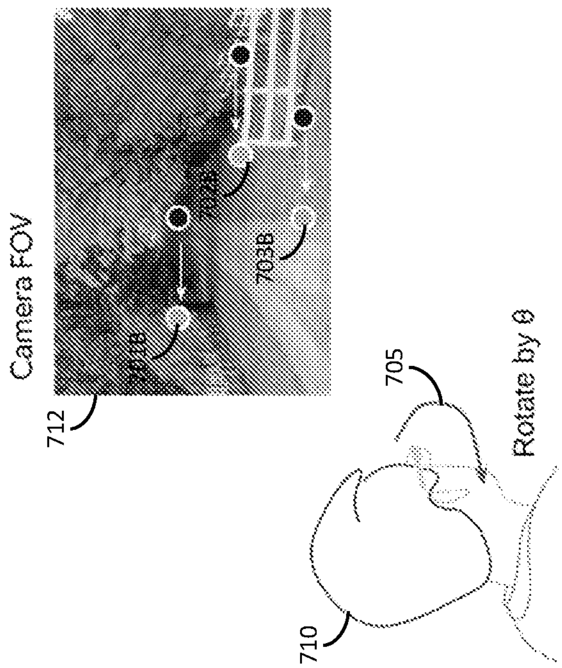
Camera FOV
712
701B
702B
703B
705
Rotate by θ
710

FIG. 7B

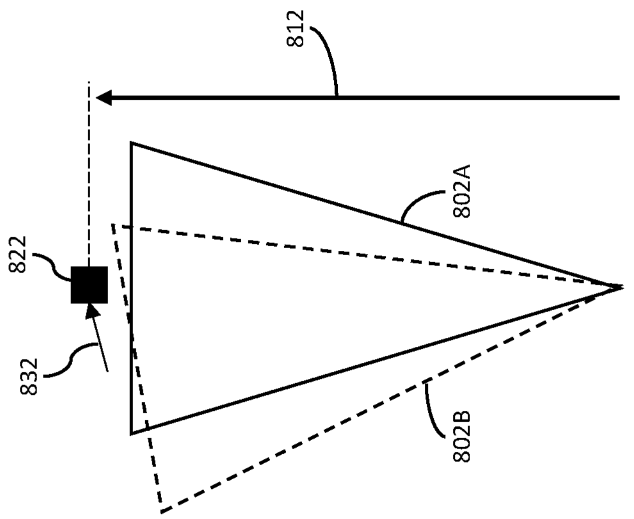

FIG. 8B

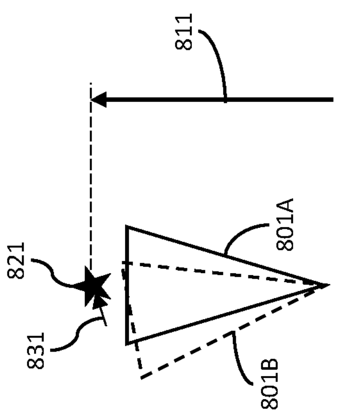

FIG. 8A

INTERNATIONAL SEARCH REPORT

International application No
PCT/US2022/081530

A. CLASSIFICATION OF SUBJECT MATTER
INV. G06F3/01
ADD.

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

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
G06F

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

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

EPO-Internal

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	US 2019/220090 A1 (HUDMAN JOSHUA MARK [US]) 18 July 2019 (2019-07-18)	1,2,4-6, 10-12, 18-20
Y	abstract paragraph [0005]	7-9, 13-15,17
A	paragraph [0025] - paragraph [0026] paragraph [0029] - paragraph [0032] paragraph [0034] - paragraph [0035] paragraph [0040] paragraph [0046] - paragraph [0051] -----	3,16
X	US 2021/183343 A1 (BEITH SCOTT [US] ET AL) 17 June 2021 (2021-06-17)	1-6, 10-12, 18-20
Y	paragraph [0041] paragraph [0043] paragraph [0048] - paragraph [0050] -----	13-16
	-/--	

[x] Further documents are listed in the continuation of Box C.

[x] See patent family annex.

* Special categories of cited documents :

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

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

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

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

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

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

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

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

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
17 July 2023	25/07/2023

Name and mailing address of the ISA/	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Fax: (+31-70) 340-3016	Bedarida, Alessandro

2

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	US 9 459 692 B1 (LI ADAM [US]) 4 October 2016 (2016-10-04)	1-6, 10-12, 18-20
Y	abstract	13,14,16
A	column 4, line 48 - line 63 column 5, line 1 - line 2 column 5, line 7 - line 8 column 5, line 11 - line 14 column 5, line 18 - line 21 column 5, line 29 - line 33 column 5, line 52 - line 63 column 6, line 4 - line 23 column 6, line 29 - line 39 claims 1, 3, 4 -----	15
X	US 2016/342206 A1 (SHAZLY TARKE A [US] ET AL) 24 November 2016 (2016-11-24)	1,2,4-6, 10-12, 15,18-20
Y	abstract	13,14
A	paragraph [0011] - paragraph [0013] paragraph [0031] paragraph [0060] claim 1 -----	3,16
X	US 2020/089333 A1 (BENSON SIMON MARK [GB] ET AL) 19 March 2020 (2020-03-19)	1,2,4-6, 10-12, 15,18-20
Y	abstract	13,14
A	paragraph [0033] paragraph [0037] paragraph [0079] - paragraph [0080] paragraph [0102] -----	3,16
Y	US 2021/312160 A1 (SHIN EUI HYUN [KR] ET AL) 7 October 2021 (2021-10-07) paragraph [0028] - paragraph [0030] paragraph [0036] -----	7-9,13, 17

2

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US2022/081530

Box No. II Observations where certain claims were found unsearchable (Continuation of item 2 of first sheet)

This international search report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

2. ☐ Claims Nos.:
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:

3. ☐ Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box No. III Observations where unity of invention is lacking (Continuation of item 3 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

see additional sheet

1. ☒ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.

2. ☐ As all searchable claims could be searched without effort justifying an additional fees, this Authority did not invite payment of additional fees.

3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:

4. ☐ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims;; it is covered by claims Nos.:

Remark on Protest

☐ The additional search fees were accompanied by the applicant's protest and, where applicable, the payment of a protest fee.

☐ The additional search fees were accompanied by the applicant's protest but the applicable protest fee was not paid within the time limit specified in the invitation.

☒ No protest accompanied the payment of additional search fees.

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

This International Searching Authority found multiple (groups of) inventions in this international application, as follows:

1. claims: 1-6, 10-12, 14-16, 18-20(completely); 13(partially)

   Detection of HMD pose.

   ---

2. claims: 7-9, 17(completely); 13(partially)

   User interface.

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No
PCT/US2022/081530

Patent document cited in search report		Publication date	Patent family member(s)			Publication date
US 2019220090	A1	18-07-2019	CN	111602082	A	28-08-2020
			CN	115857173	A	28-03-2023
			EP	3714318	A1	30-09-2020
			JP	7207809	B2	18-01-2023
			JP	2021511699	A	06-05-2021
			KR	20200106547	A	14-09-2020
			US	2019220090	A1	18-07-2019
			US	2021149481	A1	20-05-2021
			WO	2019143793	A1	25-07-2019
US 2021183343	A1	17-06-2021	CN	114761909	A	15-07-2022
			EP	4073618	A1	19-10-2022
			KR	20220110205	A	05-08-2022
			TW	202127105	A	16-07-2021
			US	2021183343	A1	17-06-2021
			WO	2021118745	A1	17-06-2021
US 9459692	B1	04-10-2016	US	9459692	B1	04-10-2016
			WO	2017172984	A2	05-10-2017
US 2016342206	A1	24-11-2016	US	2016342206	A1	24-11-2016
			WO	2015116640	A1	06-08-2015
US 2020089333	A1	19-03-2020	EP	3552052	A1	16-10-2019
			GB	2557593	A	27-06-2018
			JP	7177054	B2	22-11-2022
			JP	2020501263	A	16-01-2020
			US	2020089333	A1	19-03-2020
			WO	2018104732	A1	14-06-2018
US 2021312160	A1	07-10-2021	CN	112912824	A	04-06-2021
			KR	20200046301	A	07-05-2020
			US	2021312160	A1	07-10-2021
			WO	2020085821	A1	30-04-2020