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

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
09 August 2018 (09.08.2018)

WIPO | PCT

(10) International Publication Number
WO 2018/142136 A1

(51) International Patent Classification:
G01C 11/02 (2006.01) *H04N 1/00* (2006.01)
G06F 1/32 (2006.01) *F25D 23/12* (2006.01)
G06T 1/00 (2006.01)

(21) International Application Number:
PCT/GB2018/050289

(22) International Filing Date:
31 January 2018 (31.01.2018)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
1701593.4 31 January 2017 (31.01.2017) GB

(71) Applicant: **SMARTER APPLICATIONS LIMITED** [GB/GB]; 1st Floor, Asra House, 1 Long Lane, London SE1 4PG (GB).

(72) Inventors: **LANE, Christian**; Smarter Applications Limited, 1st Floor, Asra House, 1 Long Lane, London Greater London SE1 4PG (GB). **LANE, Isabella**; c/o Smarter Applications Limited, 1st Floor, Asra House, 1 Long Lane, London SE1 4PG (GB). **LANE, Andreas**; c/o Smarter Applications Limited, 1st Floor, Asra House, 1 Long Lane, London SE1 4PG (GB). **LONG, Gareth**; c/o Smarter Applications Limited, 1st Floor, Asra House, 1 Long Lane, London SE1 4PG (GB).

(74) Agent: **SLINGSBY PARTNERS LLP**; 1 Kingsway, London Greater London WC2B 6AN (GB).

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

(84) Designated States *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ,

(54) Title: IMAGING APPARATUS

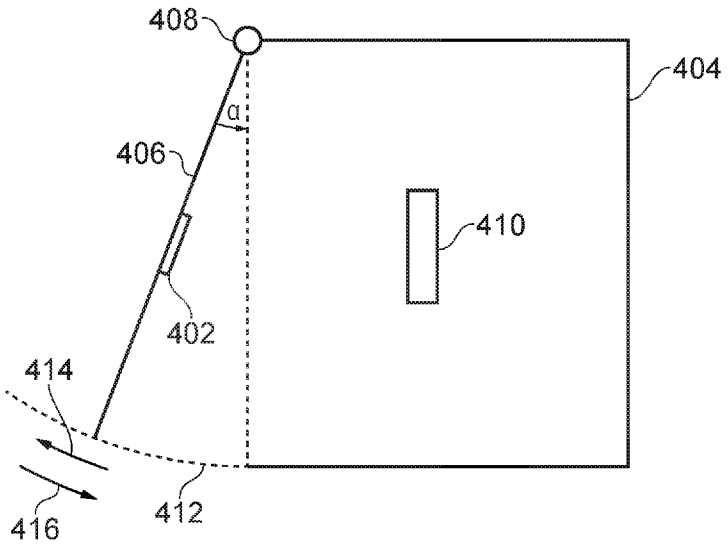

FIG. 4c

(57) Abstract: An imaging apparatus for imaging a scene, comprising an imaging device mountable to a structure which is movable relative to a scene to be imaged; a movement sensor configured to output movement data indicative of movement of the imaging device; a location sensor configured to output location data indicative of the location of the imaging device relative to the scene; and a processor configured to receive the movement data from the movement sensor and, in response, to select between a high-power mode and a low-power mode of at least one of the imaging device and the location sensor, wherein more power is consumed in the high-power mode than in the low-power mode.

WO 2018/142136 A1

[Continued on next page]

UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

Published:

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

— *before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments (Rule 48.2(h))*

IMAGING APPARATUS

This invention relates to an imaging apparatus and a method for imaging a scene. The imaging apparatus and method may enable power usage management. The imaging apparatus may provide a three-dimensional image of a scene and a method of generating a three-dimensional image of a scene.

There has been an explosion of interest in introducing wireless capability into everyday devices following the advent of the internet of things. In many instances, manufacturers have responded to this upsurge of interest by simply introducing wireless chips into domestic appliances or similar. For the user, however, this wireless capability is not always readily accessible or practically useable. Many users would like to take advantage of the new technology to make their lives easier, but they have neither the technical capability nor the money to replace their existing kitchen with a new “smart” kitchen. Therefore, there is a need for products that can introduce intelligence into people’s homes in a practical way.

According to one aspect, there is provided an imaging apparatus for providing a three-dimensional image of a scene, the imaging apparatus comprising:

an imaging device mountable to a structure which is rotatably movable relative to the scene to be imaged;

a control module for controlling the imaging device to take a plurality of images of the scene in different relative locations with respect to the scene; and

a processor configured to

receive the plurality of images from the imaging device; and

process the plurality of images to construct a three-dimensional representation of the scene.

Other aspects may include one or more of the following.

Suitably the structure is rotatable about an axis. Suitably the scene comprises an interior of an enclosure, and the axis is aligned along an edge of the enclosure. Suitably the structure comprises a door of the enclosure.

Suitably the processor is configured to process the three-dimensional representation of the scene to detect an object in the scene. Suitably the processor is configured to compare the

detected object with a list of known objects to determine a match for the detected object from the list of known objects. Suitably the list of known objects comprises at least one of a local database of objects and a remote database of objects.

Suitably the processor is configured to determine a set of difference data between the detected object and the matched object from the list of known objects, and to store the set of difference data.

Suitably the imaging device comprises a two-dimensional imaging device.

Suitably the three-dimensional representation is a depth map.

Suitably the processor is configured to process the plurality of images by determining a geometrical relationship between each of the plurality of images; aligning on or more planes of the plurality of images based on the determined geometrical relationship to form aligned images; matching one of a point and a feature in one of the aligned images with a corresponding one of a point and a feature in another of the aligned images; determining disparity information from the matched point or feature; and determining the three-dimensional representation of the scene based on the disparity information.

Suitably the processor is configured to, when aligning the one or more planes of the plurality of images, use at least one of Bouguet's algorithm and Hartley's algorithm. Suitably the processor is configured, when matching the one of the point and the feature, to use at least one of a block matching algorithm and a semi-global block matching algorithm.

Suitably the control module comprises a synchroniser configured to output a synchronisation signal, and the control module is configured to control the imaging device in dependence on the synchronisation signal, so as to synchronise the images taken by the imaging device. Suitably the synchroniser comprises at least one of a timer, an accelerometer and a position detector.

Suitably the imaging apparatus comprises at least one of an inertial measurement unit and a gyroscope configured to output data indicating location, the processor being configured to receive the data and to use the data when at least one of determining the geometrical

relationship between each of the plurality of images and aligning the one or more planes of the plurality of images.

Suitably the imaging apparatus comprises an additional imaging device arranged so that an imaging axis of the imaging device points in a different direction to an imaging axis of the additional imaging device, the additional imaging device being configured to be controlled by the control module. Suitably the additional imaging device is configured to be controlled by the control module to take an additional image in coordination with the plurality of images.

Suitably at least one of the control module and the additional imaging device is configured to output a light control signal for controlling a light. Suitably the additional imaging device comprises a light responsive to the light control signal. Suitably the imaging apparatus comprises a light sensor, and the light control signal is output in dependence on a light level sensed by the light sensor.

Suitably at least one of the control module, the imaging device and the additional imaging device comprises a wireless transceiver for communicating with one or more of a remote device and another of the control module, the imaging device and the additional imaging device. Suitably the remote device comprises the processor.

Suitably the remote device comprises memory for storing images taken by at least one of the imaging device and the additional imaging device.

Suitably the imaging apparatus comprises a further imaging device, the further imaging device being configured to image a non-visible region of the spectrum, and the further imaging device being coupled to the imaging device for imaging the scene. Suitably the further imaging device comprises an infrared sensor.

According to another aspect, there is provided a method for providing a three-dimensional image of a scene, the method comprising the steps of:

rotatably moving an imaging device relative to the scene to be imaged;

controlling the imaging device in dependence on a control module to take a plurality of images of the scene in different relative locations with respect to the scene; and

processing the plurality of images to reconstruct a three-dimensional representation of the scene.

According to another aspect, there is provided machine readable code for implementing an imaging apparatus as described herein. According to another aspect, there is provided machine readable code for implementing a method for providing a three-dimensional image as described herein.

According to another aspect, there is provided a machine-readable storage medium having encoded thereon non-transitory machine-readable code for implementing an imaging apparatus as described herein. According to another aspect, there is provided a machine-readable storage medium having encoded thereon non-transitory machine-readable code for implementing a method for providing a three-dimensional image as described herein.

According to another aspect, there is provided a method of determining objects in a scene, the method comprising the steps of:

taking a first image of the scene, the first image permitting identification of objects in the scene;

taking a second image of the scene, the second image permitting identification of objects in the scene;

determining from the first and second images which objects are present in the scene; and

storing the first and second images for later retrieval.

According to another aspect, there is provided machine readable code for implementing a method of determining locations of objects in a scene as described herein. According to another aspect, there is provided a machine-readable storage medium having encoded thereon non-transitory machine-readable code for implementing a method of determining locations of objects in a scene as described herein.

According to another aspect, there is provided an imaging apparatus for imaging a scene, comprising: an imaging device mountable to a structure which is movable relative to a scene to be imaged; a movement sensor configured to output movement data indicative of movement of the imaging device; a location sensor configured to output location data indicative of the location of the imaging device relative to the scene; and a processor configured to receive the movement data from the movement sensor and, in response, to select between a high-power

mode and a low-power mode of at least one of the imaging device and the location sensor, wherein more power is consumed in the high-power mode than in the low-power mode.

Other aspects may include one or more of the following.

The processor may be configured to select the high-power mode in response to determining that the movement data indicates movement of the imaging device. The processor may be configured to determine that the movement of the imaging device comprises vibrational movement. The processor may be configured to determine that the movement of the imaging device comprises movement relative to the scene.

The movement sensor may comprise an accelerometer. The location sensor may comprise at least one of a gyroscope and an inertial measurement unit. The location sensor may have a high-power mode and a low-power mode, and the processor may be configured to select the high-power mode of the location sensor in response to determining that the movement data indicates movement of the imaging device.

The processor may be configured to determine that the imaging device is at an imaging location with respect to the scene, and, in response, to control the imaging device to image the scene. The processor may be configured to determine that the imaging device is at the imaging location in response to receiving the location data from the location sensor. The location data may indicate that the imaging device is at the imaging location. The location data may indicate that the imaging device is moving in a particular direction past the imaging location.

The processor may be configured to control the imaging device to image the scene in response to the received movement data from the movement sensor indicating that the imaging device is moving in a particular direction.

Imaging the scene may comprise capturing a still image using the imaging device. Imaging the scene may comprise capturing a still image from a stream of images using the imaging device.

The processor may be configured to access a memory at which data indicating the imaging location is stored.

The processor may be configured to zero the location sensor when selecting between the high-power mode and the low-power mode.

The processor may be configured, in response to determining that the imaging device is at one end of a range of movement relative to the scene to be imaged, to select the low-power mode.

The imaging apparatus may comprise an additional imaging device arranged so that an imaging axis of the imaging device points in a different direction to an imaging axis of the additional imaging device, and the processor may be configured to control the additional imaging device. The additional imaging device may be configured to be controlled to capture an additional image in coordination with the image captured by the imaging device. The processor may be configured to track a plurality of imaging locations for a plurality of imaging devices, with each imaging location being for a respective imaging device.

The processor may be configured to output a light control signal for controlling a light. At least one of the imaging device and the additional imaging device may comprise a light responsive to the light control signal. The imaging apparatus may comprise a light sensor, and the processor may be configured to output the light control signal in dependence on a light level sensed by the light sensor.

The imaging apparatus may be configured to capture an image using the imaging device when the imaging device is at a first location and to capture an image using the additional imaging device when the imaging device is at a second, different, location.

The additional imaging device may be remote from the imaging device.

The structure may be rotatable about an axis. The scene may comprise an interior of an enclosure, and the axis may be aligned along an edge of the enclosure. The structure may comprise a door of the enclosure.

According to another aspect, there is provided a method for managing power usage in an imaging apparatus, the method comprising: receiving movement data indicative of movement of an imaging device for imaging a scene; and selecting between a high-power mode and a

low-power mode of at least one of the imaging device and a location sensor configured to output location data indicative of the location of the imaging device relative to the scene; wherein more power is consumed in the high-power mode than in the low-power mode.

According to another aspect, there is provided machine readable code for implementing a method for managing power usage in an imaging apparatus as described herein. According to another aspect, there is provided a machine-readable storage medium having encoded thereon non-transitory machine-readable code for implementing a method for managing power usage in an imaging apparatus as described herein.

The present invention will now be described by way of example with reference to the accompanying drawings. In the drawings:

Figure 1a shows a schematic example of an imaging apparatus;
Figure 1b shows another schematic example of an imaging apparatus;
Figure 2 shows an example of a method for providing a three-dimensional image;
Figures 3a to 3c show views of a camera;
Figures 4a and 4b show examples of a camera mounted relative to an enclosure in different positions;
Figure 4c shows another example of a camera mounted relative to an enclosure;
Figure 5a shows an example of components within a camera;
Figure 5b shows another example of components within a camera;
Figure 6 shows an example of a control module;
Figure 7 shows an example of a system incorporating the control module;
Figure 8 shows an example of a method for setting up the camera;
Figure 9a shows an example of a typical method for using the camera;
Figure 9b shows another example of a method for using the camera;
Figure 10 shows an example of a system incorporating the camera and an additional camera;
Figure 11 shows an example method of determining locations of objects in a scene; and
Figure 12 shows an example system for determining locations of objects in a scene.

An example of one tool for introducing intelligence into an existing kitchen is shown in figure 1a. The tool comprises an imaging apparatus, shown generally at 100. The imaging apparatus 100 comprises an imaging device 102. Suitably, the imaging device is a camera. Suitably the

camera is configured to image a scene in the visible portion of the electromagnetic (EM) spectrum. Other imaging devices could be used as well as, or instead of, a camera. For example, one or more of the other imaging devices can be used to image different parts of the EM spectrum from that part imaged by the camera. The camera 102 comprises a lens 104. The lens 104 may be a wide-angle lens or a fish eye lens.

The imaging apparatus further comprises a control module 106. The control module 106 is for controlling the camera 102 to take a plurality of images of a scene. The imaging apparatus further comprises a processor 108. The processor is configured to receive the plurality of images from the camera and to process the plurality of images to reconstruct a three-dimensional representation of the scene.

The camera 102 is arranged to be rotatably mountable with respect to the scene to be imaged. Suitably the camera 102 is arranged to be mountable to a structure which is rotatably movable relative to the scene.

This permits the camera to be moved in a rotatable manner, such as being moved along an arc, relative to the scene to be imaged. Moving the camera 102 along an arc permits images to be taken of the scene from differing perspectives. Taking multiple images from different perspectives allows three-dimensional information to be extracted from the images. This information extraction is suitably performed using stereoscopic image processing techniques. Suitably, two images are taken by the camera and used to extract three-dimensional information. In other examples, more than two images may be used. The use of more than two images can provide redundancy checks to be made in the extracted information, and/or it can permit an enhancement of the accuracy of the extracted information. This can improve the resulting quality of the three-dimensional (3D) information, and therefore also any calculations based on the 3D information.

Further, knowledge of the rotation of the camera relative to the scene allows image quality enhancement. Where the arc along which the camera moves is known, geometrical differences in the images taken by the camera can be calculated. This can allow for an improvement in the subsequent processing of the images, for example, by permitting more accurate edge detection. It can be difficult to accurately detect edges in images of crowded scenes. However, by processing more than one image, where the images are taken at positions spaced apart from one another along an arc of movement of the camera, the edges

can be more readily identified. This is possible as a result of the change of orientation of the camera as it moves along the arc relative to the scene. Such a movement of the camera provides images from which enhanced information can be determined, when compared to images taken, for example, from a single perspective, or from a single camera orientation (albeit at a different position).

A method for providing a three-dimensional image of a scene will now be discussed with reference to figure 2. The imaging device, or camera, is rotatably moved relative to a scene to be imaged 202. The control module is used to control the camera 204. In response to the control of the camera by the control module, the camera takes a plurality of images of the scene 206. The plurality of images are then processed 208, and a 3D representation of the scene is reconstructed 210.

In some examples, the control module can be used to control the camera by causing the output of a camera control signal. The camera is suitably responsive to the camera control signal to take an image. The camera control signal can cause the camera to take two or more images spaced apart by a particular time period (where the particular time period can be predetermined or adjustable, for example by the control module and/or by a user). The control module may be configured to output more than one camera control signal, where each camera control signal is arranged to cause the camera to take an image.

Suitably, the camera takes two images of the scene. In other examples, more than two images of the scene can be taken and processed.

Another example of a tool for introducing intelligence into an existing kitchen is shown in figure 1b. The tool comprises an imaging apparatus, shown generally at 150. The imaging apparatus 150 comprises an imaging device 152. Suitably the imaging device is a camera, such as a camera that is configured to image a scene in the visible portion of the EM spectrum. Other imaging devices could be used as well as, or instead of, a camera. For example, one or more of the other imaging device can be used to image different parts of the EM spectrum from that part imaged by the camera. The camera 152 comprises a lens 154. The lens may be a wide-angle lens or a fish eye lens.

The imaging apparatus further comprises a processor 156. The processor couples to the imaging device 152, and to a movement sensor 158 and a location sensor 160. The movement

sensor is configured to output movement data which indicates movement of the imaging device. The movement sensor may comprise an accelerometer or a vibration sensor. The accelerometer may be a one-axis accelerometer, a two-axis accelerometer or a three-axis accelerometer. As discussed herein, in at least some examples the imaging device is arranged for movement about an arc. Use of a one-axis accelerometer may therefore be sufficient to be able to determine the movement of the imaging device. Thus the movement of the imaging device can be determined at low cost. The location sensor may comprise at least one of a gyroscope and an inertial measurement unit. At least one of the movement sensor 158 and the location sensor 160 may be fast with the imaging device 152. In this way the movement sensor and/or the location sensor will move with the imaging device.

The imaging apparatus may also comprise a memory 162 for storing computer instructions for execution by the processor 156, for storing data such as the movement data and/or the location data, and/or for storing images captured by the imaging device 152. The imaging apparatus 150 may comprise a transmitter or a transceiver 164 for coupling the imaging apparatus to another device. The connection between the imaging apparatus 150 and the other device may be wired or wireless. Suitably, the transceiver 164 is configured to communicate using a wireless protocol, such as a Bluetooth protocol, an IEEE 802.11 (WiFi) protocol (for example 802.11a, 802.11b/g/n or 802.11ac), GSM, LTE and so on. The transceiver may be provided as part of the imaging device.

The imaging device 152 is arranged to image a scene. For example, the imaging device can be mounted to a structure which is movable relative to the scene to be imaged. The structure may be rotatably movable relative to the scene. Mounting the imaging device in this way permits it to be moved in a repeatable manner, such as being moved along an arc, relative to the scene. Repeatable movement of the imaging device enables it to be used to capture images of the scene in a reproducible manner. For example, a plurality of images can be taken of the scene from the same perspective, and at the same distance from the imaged scene. This reproducibility of the captured images enables comparisons to be carried out more easily between the captured images.

The imaging device or camera 102, 152 can be of any suitable form. An example is shown in figure 3. The illustrated example comprises a camera 302 which is circular in cross-section (figure 3a) and comprises a central lens 304. The camera comprises a button 306. The button can be used to switch the camera on and off. The camera further comprises a back-plate 308.

The back-plate is suitably mountable to a structure so as to mount the camera to the structure. The pivoting of the back-plate relative to a remainder of the camera 310 permits easy orientation of the camera such that the lens 304 is directed towards the scene to be imaged. This provides a tolerance in the mounting of the back-plate to the structure, permitting the camera to be usefully mountable in a wider variety of locations than might otherwise be the case.

Further examples of the camera will now be discussed with reference to figures 4a and 4b. Figures 4a and 4b illustrate an enclosure or container 404. The enclosure 404 is provided with a door 406. The door 406 is rotatably mounted to the container 404 about an axis (schematically illustrated at 408). The enclosure is shown in plan view, with the axis 408 being aligned along an edge of the enclosure adjacent an opening in the enclosure. The door is rotatable about the axis 408 to pivot relative to the enclosure and to move from one position in which the door is open, and another position in which the door is closed, closing the opening of the enclosure.

The camera 402 is, in the illustrated example, mounted to an interior face of the door 406, i.e. a face of the door that is interior to the enclosure when the door is closed. This face of the door faces the interior of the enclosure. Thus, mounting the camera 402 in this way permits the camera to face into the interior of the enclosure. Thus the camera 402 is enabled to take images or pictures of the interior of the enclosure, and in particular of an object 410 inside the enclosure. Mounting the camera 402 to the door permits the camera to rotate about the axis 408.

Figure 4a shows the door 406 in a position in which it is opened more than the position shown in figure 4b. Referring to figures 4a and 4b, it can be seen that the orientation of the camera 402 will change as the door is opened and closed. The camera is mounted to the door such that it moves with the door as the door is opened and closed. Thus the camera 402 moves about an arc relative to the enclosure 404.

The door 406 is repeatably movable relative to the enclosure 404. The door moves along the same path as it is opened and closed. This means that the camera 402 will also move along a repeatable path as the door is opened and closed. This permits the control module to control the camera 402 to take images at repeatable positions relative to the scene, i.e. relative to the interior of the enclosure.

Taking the images in this repeatable way permits an easier comparison of the images. For example, the images can be processed more efficiently, without requiring additional image processing to orient the images. This can help in reducing the processing power needed, which in turn can save cost of processing and/or power (which will mean that a battery can last longer before needing to be replaced or recharged).

Dashed lines in figures 4a and 4b illustrate the differing views the camera 402 will have of the object 410 in the two illustrated positions of the door 406. This illustrates that the change in orientation of the camera 402 as it moves rotatably relative to the enclosure interior permits a series of images to be taken from which 3D information can be calculated.

Reference is also made to figure 4c. Like reference numbers in figures 4a, 4b and 4c illustrate like elements. A description of these elements is not repeated; reference is made to the description of figures 4a and 4b above. The door 406 of the enclosure 404 can open and close about the axis 408. The edge of the door distal from the axis traces out an arc 412. The edge of the door is movable in an opening direction 414 and a closing direction 416. As the door opens and closes, the angle that the door makes to the closed position of the door (i.e. the front of the enclosure as illustrated) will change. At the position of the door illustrated in figure 4c, the door makes an angle of α degrees to the enclosure.

The camera may be provided in a camera module 500, illustrated in figure 5a. The camera module 500 may comprise at least a portion of the control module 106. The camera module 500 may comprise at least a portion of the processor 108. An example of components of the camera module will now be discussed. The camera module 500 comprises a lens 504. The lens can be a wide-angle or fish-eye lens. The lens is coupled to an image sensor 506. Suitably the image sensor is a charge coupled device (CCD). The image sensor is, in this example, an OmniVision OV5642 5MP sensor. Suitably the camera comprises the lens and the image sensor.

The image sensor 506 is coupled to an image processing module 508. In the illustrated example, the image processing module comprises an image compression integrated circuit (IC). As illustrated in figure 5a, the image compression IC can comprise a JPEG compression IC, such as a low-power JPEG compression IC. This enables the images sensed by the image sensor to be processed for storage, transmittal, and/or further processing. Compressing the

images can permit them to be subsequently processed at lower power and/or processing cost. For example, a compressed image will be smaller and so will take up less storage space. Similarly, a compressed image will be transferable more quickly due to its reduced size.

The image processing module is coupled to a transceiver 509. In some examples a transmitter or a transmitter and receiver may be provided in place of the transceiver. The transceiver 509 is configured to transmit the images received at the image sensor 506 and processed by the image processing module 508. Suitably the transceiver 509 is a wireless transceiver. The transceiver suitably comprises an IC. The wireless transceiver is suitably configured to transmit over a wireless protocol. The wireless protocol suitably comprises at least one of Bluetooth, Wi-Fi (for example 802.11a, 802.11b/g/n or 802.11ac), GSM, LTE etc.

The camera module 500 comprises a power management system comprising a battery 510, a power management module 512 and a power connector 514. In other examples, only one of the battery 510 and the power connector 514 need to be provided. In the illustrated example, the battery 510 comprises a Li-ion 4.2V 2000 mAh battery. Other battery types could be provided instead of or as well as this battery type. The illustrated power connector 514 is a USB-C 5V connector. Other types of connector may additionally or alternatively be provided.

The power management module 512 controls power in the camera module 500. The power management module 512 is suitably configured to control power in the camera module 500 in dependence on at least one of the power level in the battery 510 and the power connection status of the power connector 514. The power management module controls power to the image sensor 506, the image processing module 508 and the transceiver 509.

The camera module 500 illustrated in figure 5a further comprises one or both of a light and a vibration sensor 516. Outputs from one or both of these sensors can be used to determine when the camera is to take an image. The illustrated camera module 500 also comprises a temperature sensor 518 for sensing the temperature at the camera module, for example the temperature within the enclosure when the door is closed. The illustrated camera module 500 also comprises an accelerometer 520, such as a one-axis accelerometer. The accelerometer is configured to sense the acceleration of the camera module 500. Information about the acceleration of the camera module 500 can be used to determine the location of the camera module 500 with respect to the enclosure as it is moved about an arc relative to the enclosure, as discussed above.

Another example of a camera module is illustrated at 550 in figure 5b. The camera module 550 may comprise at least a portion of the processor 156. As illustrated, the camera module 550 comprises a camera module processor, at 557. Components in this example of the camera module 550 will now be discussed. The camera module 550 comprises a lens 554. The lens can be a wide-angle or fish-eye lens. The lens is coupled to an image sensor 556. Suitably the image sensor is a charge coupled device (CCD). The image sensor may comprise an OmniVision OV5642 5MP sensor. Suitably the imaging device, or camera, comprises the lens and the image sensor.

The image sensor 556 is coupled to the camera module processor 557. The camera module processor 557 comprises an image processor or image processing module 558. The image processing module may comprise an image compression integrated circuit (IC). The image compression IC can comprise a JPEG compression IC, such as a low-power JPEG compression IC.

The camera module processor 557 is coupled to a transceiver 559. In some examples a transmitter or a transmitter and receiver may be provided as well as or in place of the transceiver. The transceiver 559 is configured to transmit the images received at the image sensor 556, and/or processed versions of those images as processed by the image processing module 558. Suitably the transceiver 559 is a wireless transceiver. The transceiver suitably comprises an IC. The wireless transceiver is suitably configured to transmit over a wireless protocol. The wireless protocol suitably comprises at least one of Bluetooth, Wi-Fi (for example 802.11a, 802.11b/g/n or 802.11ac), GSM, LTE and so on.

The camera module 550 comprises a power management system comprising a battery 560, a power management module 562 and a power connector 564. In other examples, only one of the battery 560 and the power connector 564 need be provided. The battery 560 may comprise a Li-ion 4.2V 2000 mAh battery. Other battery types could be provided instead of or as well as this battery type. The power connector 564 may comprise a USB-C 5V connector. Other types of connector may additionally or alternatively be provided. Whilst the power management module 562 is shown in figure 5b as a separate element, the functionality of this module may be carried out by the camera module processor 557. The power management module 562 need not be provided as a separate element. This is illustrated in figure 5b by the battery 560 and the power connector 564 being shown coupled to the camera module

processor 557 and to the power management module 562. In some examples, the battery and/or the power connector are coupled to one of the camera module processor 557 and, where present, the power management module 562.

The power management module 562, and/or the camera module processor 557, controls power in the camera module 550. The power management module 562, and/or the camera module processor 557, is suitably configured to control power in the camera module 550 in dependence on at least one of the power level in the battery 560 and the power connection status of the power connector 564. The power management module 562, and/or the camera module processor 557, controls power to the image sensor 556, the image processing module 558 and the transceiver 559.

The camera module 550 illustrated in figure 5b further comprises a sensor bank 570. The sensor bank comprises a movement sensor 572 and a location sensor 574. The movement sensor 572 may comprise an accelerometer. The movement sensor 572 may comprise a vibration sensor. The location sensor 574 may comprise a gyroscope. The location sensor 574 may comprise an inertial measurement unit. The sensor bank comprises a temperature sensor 576 and a light sensor 578. The temperature sensor 576 is for sensing the temperature at the camera module, for example the temperature within the enclosure when the door is closed. The light sensor 578 is for sensing light levels at the camera module 550, for example the luminosity of light in the environment of the camera module. Not all of these sensors need be provided in all examples. In some examples, additional and/or alternative sensors may be provided.

A signal received at the transceiver 559 may be used to determine when to capture an image using the imaging device. A signal from the sensor bank 570 may be used to determine when to capture an image using the imaging device. The signal from the sensor bank 570 may comprise one or more outputs from the movement sensor, the location sensor, the temperature sensor, the light sensor and/or any other sensor present. The camera module processor 557 may be configured to determine when to take an image using the imaging device, in dependence on at least one of the signal received at the transceiver 559 and the signal from the sensor bank 570.

The movement sensor 572 is configured to output movement data. The location sensor 574 is configured to output location data. Movement of the imaging device can be determined in

dependence on at least one of the movement data and the location data (for example, a change in the location data with time can indicate movement). Suitably the movement data is used to determine movement of the imaging device. The location of the imaging device can be determined in dependence on at least one of the movement data and the location data (for example, knowledge of how the movement data changes with time can be used to determine a location, as discussed briefly elsewhere herein). Suitably the location data is used to determine the location of the imaging device.

In preferred examples, at least one of the imaging device, the image processor, the transceiver and one or more sensor of the sensor bank has a high-power mode and a low-power mode, where a greater amount of power is typically consumed in the high-power mode than in the low-power mode. For example, the high-power mode can comprise a powered-up or active mode, in which the respective element is fully operational. The low-power mode can comprise a powered-down or inactive mode, in which the respective element is not fully operational. The low-power mode may comprise a standby mode and/or a sleep mode, in which the respective element is not fully operational. The respective element may retain some operational capabilities in the low-power mode. For example, in a sleep or standby mode, the respective element may not be able to carry out its primary function (for example, the imaging device may not be able to capture an image, or the location sensor may not be able to output location data) but may still be in a partially-powered state so as to be able to wake from that state into the high-power state quickly. In this way, the respective elements can save energy by being in the low-power state when they are not needed to carry out their primary function, but can still be switched to the high-power state quickly when needed. Suitably, the respective elements are configured to change from the low-power mode to the high-power mode on an activation signal. The activation signal may be received by the transceiver 559. The activation signal may be output from a sensor, such as the movement sensor 572. The camera module processor 557 may output the activation signal, for example in dependence on one or more of a signal received by the transceiver and a sensor output. The respective element in the low-power mode can be woken from that low-power mode into the high-power mode, for example by the camera module processor, using the activation signal.

In some examples, at least one of the location sensor, the temperature sensor and the light sensor can have high- and low-power modes. Suitably at least the location sensor has high- and low-power modes.

In one configuration, the camera module processor 557 can receive movement data from the movement sensor 572. In dependence on that movement data, the camera module processor 557 can select between the high-power mode and the low-power mode of the imaging device, such as the camera, and/or the location sensor, such as the gyroscope.

Not all of the illustrated components need be provided in every example.

In a preferred example, the enclosure is a fridge, and the camera is mountable to the door of the fridge. The camera is then able to take pictures or images of the contents of the fridge.

The following discussion will be made in the context of the example of the imaging apparatus shown in figure 1a.

An example of a control module is illustrated in figure 6. The control module, shown generally at 600, comprises a synchroniser 602 and a module for outputting a control signal 604. Suitably the control module is configured to output the control signal to the camera.

The synchroniser comprises, in one example, at least one of a timer, the accelerometer and a position detector. The synchroniser is suitably configured to output a synchronisation signal to permit synchronisation of the images taken by the camera. For example, the image sensor may be responsive to receiving the synchronisation signal to capture, store and/or transmit an image. For example, receiving the synchronisation signal can cause the image sensor to take an image.

Images can be considered to be synchronised where they are taken with the camera in particular locations. For example, it might be desired to take one image when the door is being closed and is at a particular angle (say, 10 degrees) relative to the closed position of the door. It might then be desirable to take another image when the door is at another, different angle (say, 5 degrees) relative to the closed position of the door. The particular angles can be selected in dependence on the image quality at those angles. The separation of the angles can be selected in dependence on a quality of a 3D representation obtained at a particular angular separation of positions for the images. For example, where it has been determined that a separation of 5 degrees in the position of the door is optimal to permit generation of a 3D representation of the scene, this angular separation can be chosen.

Suitably the control module causes the images to be repeatably synchronised. This enables subsequent sets of images to be compared against one another, and a picture built up of the contents of the enclosure (for example the fridge) over time.

Some or all of the control module might also be implemented by a distributed processing system, e.g. on a smartphone, in the cloud or elsewhere. The control module, or some aspects of it, could be implemented in hardware but it is most likely to be implemented by a processor acting under software control. Suitably the control module comprises a wireless module 606, for communicating with the camera, the processor, another device and/or the cloud. The wireless module 606 is preferably capable of transmitting and/or receiving data via a suitable wireless communication protocol such as Bluetooth, Wi-Fi, GSM, LTE etc. The wireless module is thus able to wirelessly upload data to an external controller (whether in a smartphone, the cloud or elsewhere). The other device may be a smartphone, tablet, laptop or desktop computer or the like. In many implementations, at least part of the control module is likely to be a smartphone running an app.

For example, as the imaging device captures an image (or more than one image), the captured image(s) can be sent wirelessly to a remote device, such as a device in the cloud. If the wireless communication path between the imaging apparatus and the cloud is disrupted, for example by losing WiFi signal at the imaging device within the fridge, data such as the captured image(s) can be stored locally to the imaging device, and transmitted once the wireless communication path is re-established. For instance, in some situations, a camera will capture an image as the fridge door is nearly closed. As the door closes, the WiFi connection from the camera may be lost. The image can be stored locally to the camera, and transmitted when the fridge door is next opened. Note that, suitably, this does not require much if any additional memory local to the camera. Since the image can be sent as the fridge door is opening, it will be sent before the next image is captured (on closing of the door). Hence the memory space used by the stored image can be freed up before the next image is taken. The local memory suitably comprises flash memory.

The processor may comprise a remote processer in addition to the local processor. The processing of the images can be performed at the local processor, at the remote processor, or at a combination of the local and remote processor, i.e. the images can be partially processed locally and partially processed remotely.

Similarly, in the context of the example of the imaging apparatus shown in figure 1b, some or all of the processor might be implemented by a distributed processing system, e.g. on a smartphone, in the cloud or elsewhere. Some aspects of the functionality of the processor could be implemented in hardware but it is most likely to be implemented in software.

The transceiver may be configured for communicating with the camera, the processor, another device and/or the cloud. The transceiver is preferably capable of transmitting and/or receiving data via a suitable wireless communication protocol such as Bluetooth, Wi-Fi, GSM, LTE and so on. The transceiver is thus able to wirelessly upload data to an external controller (whether in a smartphone, the cloud or elsewhere). The other device may be a smartphone, tablet, laptop or desktop computer or the like. In many implementations, at least part of the processor is likely to be at a smartphone running an app.

A schematic representation of such a system is illustrated in figure 7. The control module 702 is able to communicate bi-directionally with another device 704 and the cloud 708. Similarly, the other device 704 and the cloud 708 are able to communicate bi-directionally with each other. The other device is typically a user device such as a smartphone, tablet computer, PC, laptop etc. It comprises a wireless module 705 that is capable of transmitting and/or receiving data via a suitable wireless communication protocol such as Bluetooth, Wi-Fi, GSM, LTE etc. It also comprises a user interface 706, which may include one or more of a display, keyboard, touchscreen etc.

In this example the other device 704 also comprises a database 707. The database may include details relating to one or more items for identification. The database may store details of commonly used items. For example, the imaging apparatus can be used to obtain images of a fridge interior. In this case, the items likely to be images are those items typically found in fridges, such as milk, yoghurt, vegetables and so on. The database may therefore suitably comprise details of these types of items. For example, the database might comprise details regarding the size and/or shape of a milk container, and the colour of the container (which might indicate whether it contains full fat milk or semi-skimmed milk, for example).

The database 707 is shown as being part of the other device 704 in figure 7 but it may be stored externally of the device and accessed when needed. For example, it might be stored in the cloud, illustrated at 709. If the database is stored externally, the other device 704 may include a cache to store frequently accessed information locally. Similarly, the control module

may additionally or alternatively comprise a local cache to store frequently accessed information local to the control module.

In the context of the example of the imaging device shown in figure 1b, the camera module processor may be provided as well as or in place of the control module 702 of figure 7. In a similar manner to the discussion above, the camera module processor 557 may be configured to communicate bi-directionally with the other device 704 and the cloud 708. The remainder of the description above in respect of figure 7 is applicable to this example too. It is not being repeated here for brevity.

Suitably the imaging device comprises a two-dimensional imaging device. For example, the camera may be a 2D camera. The use of a 2D camera to obtain 3D vision can be done without increasing hardware cost.

Aspects of the operation of the camera will now be discussed with reference to figures 8 and 9. Figure 8 illustrates an example of a method for setting up the camera. In this example, the camera is a wireless camera that takes a picture of the contents in a user's fridge upon the fridge door being opened or closed. Suitably the camera includes a light sensor and/or motion sensor to detect opening or closing of the fridge door. The device suitably operates in sleep mode until it senses light and/or motion. This automatically triggers the camera to take a picture upon a forwards *y*-axis motion.

The fridge camera may use high-speed sampling of a one-axis accelerometer to calculate the relative location of the camera, based upon the understanding that: (i) distance equals speed multiplied by time; and (ii) that the camera is movable about a fixed hinge. From this information it will take a picture of the fridge only when the camera is moving in the correct direction and as it passes the optimal selected location. When the fridge door is closed, accelerometer positioning is reset, avoiding any cumulative drift in readings.

The fridge camera may use a location sensor such as a gyroscope to calculate the location of the camera relative to the fridge body.

In setting up the camera, the camera is first located in a desired location in the fridge 802. This can be as directed or as recommended in instructions. The camera (here termed a FridgeCam) is then paired with a remote device 804. Suitably the pairing occurs over a

wireless protocol such as Bluetooth, Wi-Fi, GSM and/or LTE. The remote device is suitably a smartphone running an app. Once paired, a setup sequence is initiated.

As part of the setup process initiated at 804, the camera will be configured to send frequent still shots or images to the remote device 806, for example the app. The app will direct a user to adjust the angle of the fridge door until a suitable image is obtained. The user will then confirm that they are happy with the selected angle 808. This position of the camera, i.e. the position of the door to which the camera is mounted, is saved internally, and will be the reference position that is used in normal operation of the camera 810. I.e. during a normal opening or closing of the fridge door, the camera will take a picture when it is at the position selected by the user in the setup process. This process can be followed for each position at which an image should be taken. For example, where two images are to be taken, two positions can be selected. This process helps ensure that the view or views desired by the user is/are used for each image or series of images captured by the camera.

In some examples, the camera is configured to send a live feed of the scene as imaged by the camera. The live feed may comprise an image feed which is imaged at or up to about 30 frames per second. The camera is mountable centrally between the door hinge and the distal edge of the door. Mounting the camera in this way permits a single calibration procedure to be carried out irrespective of whether the door opens to the left, to the right, or to both sides.

Figure 9a shows an example of a typical method for using a camera such as a fridge camera. At 902, motion of the camera or light is detected. This causes at least one of a wireless module CPU and an image compression IC, such as a JPEG image compression IC, to wake from a sleep or low-power state 904. The at least one of the wireless module CPU and the image compression IC are typically put into the sleep or low-power state to conserve energy. This can help the lifetime of a battery powering the system to be extended.

The system then monitors, for example by using an accelerometer, for a change in acceleration over time, which can be used to identify the door position 906. When the door is closing and is in the desired position, for example the position selected during the setup process, a still image is captured by the camera 908. The image is suitably a 5MP JPEG image, typically 512 KB in size.

In this example, the system then connects wirelessly to the cloud to upload the image to a storage region in the cloud 910, 912. Suitably the wireless connection is one of Bluetooth, Wi-Fi, GSM and LTE. Once the image has been uploaded, the system returns to the low-power or sleep state to conserve energy 910.

A notification of the uploaded image can be sent to a designated device, such as a user's mobile telephone, by using a push notification service 914. This ensures that the user is kept aware of the operation of the fridge camera. This step is suitably controllable by a user. For example, the user may adjust the number and/or frequency with which notifications are sent to their device. The user may switch off the push notification service in favour of user polling. Alternatively, the push notification service may be used in combination with user polling.

Figure 9b shows another example method for using a camera such as a fridge camera. At 950, motion of the camera, or light, is detected. Motion of the camera can be detected by the movement sensor. Light can be detected by the light sensor. Suitably, the method comprises detecting movement of the camera in dependence on the movement data from the movement sensor. In response to detecting movement of the camera, the location sensor is caused to wake from a low-power state or mode into a high-power state or mode 952. The location sensor can be kept in the low-power mode when the door is closed to save energy usage of the imaging apparatus. The location sensor can be woken into the high-power mode when it is needed. The location of the camera, and its direction of movement can be monitored 954, for example by analysing the movement data and/or the location data. When it is determined that the camera is at a location at which it is desired to capture an image, and the camera is moving in the correct direction, for example in a closing direction 416, an image can be captured 956.

In one example, the high-power mode of the location sensor is selected in dependence on the movement data. The selection of the high-power mode can be made before the door opens. For example, the movement sensor can detect vibrational movement as well as movement of the camera relative to the scene (i.e. the fridge body). Vibrational movement, or vibrational motion, can be characterised by constrained motion about the same position. Vibrational movement can be differentiated from movement of the imaging device relative to the scene. When experiencing vibrational movement, the camera may move about a position which does not vary with respect to the scene to be imaged. One way of imagining this is to consider the camera wobbling. In this case, the camera will undergo vibrational motion, but

its position or location will not change, or will not substantially change, with respect to the scene.

The occurrence of vibrational movement of the camera can indicate that movement of the camera relative to the scene is about to occur. For instance, where the camera is attached to the door of a fridge, as a user grasps the door, the camera is likely to undergo vibrational movement. After grasping the door, the user can open the door. The camera will then undergo movement relative to the scene (for example a fridge interior) as the door opens.

Selecting the high-power mode for the location sensor based on detecting vibrational motion of the camera enables the powering-up of the location sensor before the door opens. This can mean that the location sensor is fully operational as the door starts to move, which can increase the accuracy of the location data.

Powering-up the location sensor before the camera moves relative to the scene has a further benefit. The location sensor reading, or the location data, may have a tendency to drift over time. Readings, or location data, output from the location sensor can become inaccurate as time progresses. It is therefore desirable to zero the readings of the location sensor, for example periodically. One way in which this can be done is to zero the location sensor when it is detected that the door has been closed. Such a detection can be made in dependence on the movement data and/or the location data. Suitably the detection is made in dependence on the movement data, to avoid inaccuracies in the location data leading to inappropriately zeroing the location sensor. Zeroing the location sensor after the door has been closed can be performed before selecting, for example by the processor, the low-power mode for the location sensor. Thus the location sensor can be zeroed, or reset, before being put into a low-power mode, such as a standby or sleep mode.

Another, preferred, approach, is to zero the location sensor before the door opens. This is possible in the present system due to the movement data indicating that the camera is about to move, for example by detecting vibrational movement before detecting movement relative to the fridge body. The time window, which is typically up to a second or more, between the user grasping the fridge door handle, and pulling on the handle to open the door, is sufficient to power up the location sensor, and to reset or zero the location sensor. This approach enables the location data subsequently output by the location sensor to be highly accurate. The location sensor may be zeroed after a threshold time has elapsed since the last time it

was zeroed (for example when, or just before, the door next opens after the threshold time has elapsed), and/or after a threshold number of door openings and closings since the last time it was zeroed.

Thus, in an imaging apparatus where the movement sensor comprises an accelerometer and the location sensor comprises a gyroscope, movement data from the accelerometer can be continually or periodically sampled, for example up to every 500 msec, up to every 200 msec or up to every 100 msec. When the movement data indicates that the camera is moving, for example vibrating, the gyroscope and/or the camera can be powered up. Optionally the gyroscope can be zeroed at that point. Location data of the gyroscope can then be sampled to determine the location of the camera relative to the scene to be imaged. Where data is sampled periodically, the sampling rates may differ when at least one component of the system is in a low-power mode (such as when the door is shut) compared to when the components are in high-power modes. For example, the movement sensor may be sampled periodically when the location sensor and/or the imaging device is in a low-power mode, and the movement sensor may be sampled at a higher rate, or continuously, when the location sensor and/or the imaging device is in a high-power mode. This approach can offer energy/processing savings during potentially long periods of time when, in the example herein, the fridge door is not open. Such savings can be achieved whilst maintaining a high accuracy of readings when the fridge door is open, or is likely to be opened.

As discussed above, during a set-up phase, a determination can be made of an imaging location (of the camera relative to the scene) at which it is desired to take an image of the scene. For example, where the camera is mounted to the door of a fridge, the user can choose to take an image when the fridge door is at an angle, α, which permits an optimal view of the interior of the fridge. Preferably the imaging apparatus is configured to capture the image as the door is closed past this imaging location. This helps to ensure that the image captured will be up-to-date. Since a user often puts things in a fridge or takes things out of a fridge after opening the door, imaging the fridge as the door closes is more likely to provide up-to-date images. Thus the imaging apparatus is suitably configured to determine both the direction of movement of the camera, for example whether the door is opening 414 or closing 416, and the location of the camera relative to the scene (i.e. the fridge interior). The direction of movement of the camera can be determined in dependence on the location data (for example a time-variance of the location data) and/or the movement data.

The camera can be configured to capture a still image of the scene. For example, the camera can take an image as it passes the imaging location in a closing direction. The camera can take a single image as it passes the imaging location in a closing direction. The camera can be configured to take a series of images, for example an image stream such as a video stream. The series of images may be taken at or up to about 30 frames per second. The still image can be taken from the series of images. The processor can control the camera to image the scene by capturing a still image. The processor can control the camera to image the scene by capturing a still image from a stream or series of images.

The series of images will comprise images taken at locations about the imaging location, i.e. before the imaging location and after the imaging location. The separation in location at which each image in the series is taken will depend on the speed of movement of the camera. This can be determined, for example by the processor, in dependence on at least one of the movement data and the location data. One of the images in the series may be taken at the imaging location. This image may be captured as the still image. In some examples, all images in the series, or stream, of images which are imaged at positions within a threshold distance of the imaging location can be considered candidate images. The threshold distance may be less than about 10mm, or less than about 5mm, or less than about 2mm. The threshold distance may depend on the speed at which the camera is moving. The threshold distance may decrease as the speed at which the camera moves decreases. A still image can be selected from the candidate images. This selection can be performed by the processor. This selection can be performed by the image processor. This selection can be performed in dependence on an image quality metric. The image quality metric can comprise at least one of image sharpness, illumination level, distance from the imaging location at which the image was taken, and so on. In some examples, the selection can be performed in dependence on minimising or maximising a value associated with the image quality metric. For example, the sharpest image may be selected. In another example, the image taken closest to the imaging location can be selected. Combinations of these, and optionally other factors, can be taken into account. For example, an image of the series of images can be selected from a group of candidate images having at least a threshold sharpness and at least a minimum illumination level, the image of the group having been taken at a position closer to the imaging location than any other image of the group. Thus if an image taken at the imaging location is poor quality, it can be discarded, and another

image selected in its place. This enables a useful image to be selected even where one or more images of the series of images are of poor quality.

The captured image, and/or the series of images, can be stored at a memory. The memory may be local to the camera. The memory may be remote from the camera. In some examples a memory can be provided local to the camera and another memory provided remote from the camera. The local memory and/or the remote memory is suitably accessible to the processor. Data indicating the imaging location may be stored at the local and/or remote memory. In some examples, the processor is configured, for example during the set-up phase, to store the data indicating the imaging location at the local and/or remote memory. Suitably the data indicating the imaging location is stored in the local memory, for speed of access. Suitably the captured image, and/or the series of images, is stored at the remote memory, to save memory space local to the camera.

The processor may be configured to select the low-power mode of the camera and/or a sensor such as the location sensor. The processor may be configured to select the low-power mode in dependence on determining that the camera is at one end of a range of movement relative to the scene to be imaged. The selection of the low-power mode may be made in dependence on at least one of the movement data and the location data indicating that the camera is not moving. For example, the camera and/or at least one of the sensors can be put into a low-power mode when the movement data indicates a lack of movement of the camera, and/or the location data indicates that the door is closed. As mentioned above, the processor may be configured to zero the location sensor before selecting the low-power mode of the location sensor.

The imaging apparatus may comprise an additional imaging device arranged so that an imaging axis of the imaging device points in a different direction to an imaging axis of the additional imaging device. The processor may be configured to control the additional imaging device. The area to be imaged by the additional imaging device is at least partially non-overlapping with the area to be imaged by the imaging device. For example, the imaging device and the additional imaging device can be arranged to image different areas. For example, the imaging device can be arranged to image the interior of a fridge (i.e. looking into the fridge) and the additional imaging device can be arranged to image the door of the fridge (i.e. looking out of the fridge). In another example, the additional imaging device can be arranged to image an inside of a tray in the fridge. The view of the tray interior may be at

least partially obscured in the view of the imaging device, so the imaging of the tray interior by the additional imaging device enables additional aspects of the scene to be imaged, compared to using only the imaging device. Two or more additional imaging devices may be provided. Thus additional areas, for example of the fridge, can be imaged.

The additional imaging device can be controlled to capture an additional image in coordination with the image captured by the imaging device. The images captured by the imaging device and the additional imaging device can be coordinated under the control of the processor. The imaging device and the additional imaging device can have a master/slave relationship. The imaging device can be considered to be a master device and the additional imaging device can be considered to be a slave device. The (slave) additional imaging device can be controlled to capture an image in dependence on the capturing of an image by the (master) imaging device. Coordinating the capturing of images by the imaging device and the additional imaging device can comprise taking the images at the same time, or in succession. For example, the additional imaging device can be controlled to capture its image immediately after the imaging device captures its image. The additional imaging device can be controlled to capture its image a predefined time after the imaging device captures its image. For example, the processor can control the imaging device to take an image. A predefined time later, the processor can control the additional imaging device to take an image.

The processor may be configured to track a plurality of imaging locations for a plurality of imaging devices, with each imaging location being for a respective imaging device. The processor may be configured to track more than one imaging location so that each camera can image at a camera-specific or optimal location, angle and/or timing.

The imaging apparatus can comprise a light for illuminating the scene to be imaged. The imaging device can comprise the light. Where present, the additional imaging device (or one of a plurality of additional imaging devices) can comprise the light. Two or more of the imaging device and the one or more additional imaging device can comprise a light. The processor may be configured to output a light control signal for controlling the light or a plurality of lights. The light or lights can, for example take the form of a flash for illuminating the fridge interior, the fridge door and/or a fridge tray. The light or lights are responsive to the light control signal.

As mentioned above, the imaging apparatus comprises a light sensor. The processor may be configured to output the light control signal in dependence on a light level sensed by the light sensor. Thus, the light emitted by the light or lights can be controlled in dependence on the ambient light levels present, as detected by the light sensor. This enables the power or intensity of the at least one light to be controlled so that an image can be captured at a desired light level. The light can therefore be controlled so that the image is captured at a desired luminosity level. In some examples, the processor is configured to control the at least one light so that each image is taken at the same luminosity level. This can assist in enhancing the consistency between the captured images. In some examples, the light sensor only need be powered-up once the fridge door has opened by at least a threshold amount. This is because fridges typically have an interior light, which illuminates once the door has opened part-way. The processor may be configured to determine when the door has opened by at least a threshold amount (distance and/or angle), and to select the high-power mode of the light sensor in dependence on that determination. The processor may be configured to make this determination in dependence on at least one of the movement data and the location data. Selecting a high-power mode for the light sensor in dependence on the movement data and/or the location data can therefore mean that the light sensor can remain in the low-power mode when it is not needed, helping to conserve energy. In this manner battery life may be prolonged.

The imaging apparatus may be configured to capture an image using the imaging device when the imaging device is at a first location and to capture an image using the additional imaging device when the imaging device is at a second, different, location. This enables the position of the camera to determine when each of two (or more) cameras are controlled to take an image.

Additionally or alternatively, the imaging apparatus may be configured to capture an image using the imaging device at a first time, and to capture an image using the additional imaging device at a different, second time. The second time may be later than the first time. The second time may be a predefined time later than the first time. The imaging apparatus may be configured to capture an image using the imaging device when the imaging device is at the imaging location, and to capture an image using the additional imaging device when the imaging device has moved a predetermined distance from the imaging location. The difference between the first time and the second time may be determined in dependence on the speed of movement of the imaging device. The difference between the first time and the

second time may be determined in dependence on at least one of the movement data and the location data. For example, the movement data and/or the location data can be used, for example by the processor, to determine the speed at which the imaging device is moving with respect to the scene.

Arranging the imaging apparatus to capture two (or more) images at different times (using two, or more, cameras), or to ensure that two images are not captured at the same time, can reduce the interference between two imaging devices. For example, where one of the imaging devices comprises a light which illuminates when that imaging device captures its image, that light may shine towards the other imaging device. Capturing an image using that other imaging device when the light is shining on it may cause glare or other undesired optical effects or artefacts to occur in the captured image. Thus it may be desirable to avoid taking images using different imaging devices at the same time.

One way of achieving this is to assign a different imaging location to each imaging device. Not all of the imaging devices need move. In some examples, where at least one imaging device moves, locations along the path of movement of that imaging device can be assigned as locations which trigger the capture of images by different imaging devices. For example, where the imaging apparatus comprises an imaging device mountable to the door of a fridge and an additional imaging device mountable on a shelf in the body of the fridge, the imaging device will move as the door opens and closes but the additional imaging device will not. As the imaging device (on the door) moves past a first imaging location in a closing direction, an image (for example of the fridge interior) can be taken by the imaging device. As the imaging device moves past a second imaging location in a closing location, an image (for example of the fridge door or of a tray in the fridge) can be taken by the additional imaging device. Where the additional imaging device does not move, the timing of the image capture by that additional imaging device is less critical. In this example, there may therefore be a larger tolerance around determining when to capture the image using the additional imaging device than there is around determining when to capture the image using the imaging device.

The additional imaging device is preferably remote from the imaging device. At least one of the imaging device and the additional imaging device may comprise a wireless transceiver for communicating with one or more of a remote device and another of the imaging device and the additional imaging device. The imaging apparatus can comprise the remote device.

The remote device can comprise at least a portion of the processor. The remote device may comprise the remote memory.

A discussion will now be provided of the processing of the images taken by the camera so as to reconstruct a 3D representation of the imaged scene.

As mentioned above, the camera is configured to take a series of images (which suitably comprises at least a pair of synchronised images). These images are passed to the processor for processing.

The series of images can comprise a video sequence of images. The video sequence may be a video sequence of up to 10 seconds. Thus two or more still images, or a video sequence, can be passed to the processor for processing.

The processor is configured to process the three-dimensional representation of the scene to detect an object in the scene. The processor analyses the images, and can make use of image detection algorithms, such as edge detection algorithms, to detect one or more objects in the image (i.e. in the imaged scene).

The processor is suitably configured to compare the detected object with a list of known objects to determine a match for the detected object from the list of known objects. Image recognition algorithms, such as 3D image recognition algorithms can be used to enable to processor to detect and/or identify the object. This allows the imaging apparatus to be able to identify what objects are in the imaged scene, or in one example, what items are in the fridge.

The items can be identified by comparing the processed image against a set of known markers (for example, identifying at least one of the size, shape, colour, logo, barcode and so on associated with an item). The closest match from the list of known objects to these known markers can be selected as the identified item. For example, where a milk container is being imaged, information can be obtained about the size and shape of the container, and the colour of the container or of a part of the container. The object in the list of objects that is the closest match (for example, a 2 pint container of semi-skimmed milk) is selected, and the imaged item identified as that object (in this example, the imaged item is identified as a 2 pint container of semi-skimmed milk).

Identifying the item in the fridge enables the imaging apparatus to determine additional information associated with that item. This associated information can be obtained from an information store. The information store can be provided locally or remotely from the imaging apparatus.

The list of known objects suitably comprises at least one of a local database of objects and a remote database of objects. Storage may be an issue when considering a local database. This can be due to size of the storage device and/or power consumed by the storage device. However, a local database will be much faster to access, meaning that item identification can occur more quickly. A remote database can be provided so that local storage requirements for the database are not an issue. However, it may take longer to access the remote database, which can increase the time taken to identify the item in the fridge. Such delays may be perceptible to a user, and so it is desirable to avoid, or reduce, these delays.

However, in some examples, at least some of the processing of the images is done in the cloud, or at a remote device. In these examples, it is convenient to carry out at least some of the processing where the database is held. For example, where the database is held on a remote device, the processing can be done at that remote device. Where the database is held in the cloud, the processing can be done in the cloud. Thus access times are less of an issue, and efficient processing can be achieved.

In one example, a relatively smaller local database is provided, and a relatively larger remote database is also provided. The remote database can complement the local database. For example, the local database can store information associated with the most frequently used items, enabling a faster identification for items that are typically imaged in the fridge. The provision of the remote database ensures that other items, such as unusual items, can still be identified. This arrangement can provide a balance between speed of identification and having a complete inventory of items that might be imaged.

Suitably, the local database is updatable to reflect the items actually imaged in the fridge. For example, the local database can ‘learn’ which are the most frequently identified items for a particular fridge, and can ensure that associated information is stored locally for those items.

In one example, the processor is configured to determine a set of difference data between the detected object and the matched object from the list of known objects, and to store the set of difference data. The object, or item, may be identified based on a closest match between the data associated with the object determined from the images, and the data associated with an item from the database. In some cases, a match can be made with a high level of confidence. In other cases, a lower level of confidence may be accepted to be able to identify the item. The differences between the data in the database and the data determined from the images may be due to one or more factors which may include that the images were taken under different lighting conditions, at different angles, the objects are at different depths, the objects may be partially obscured, and so on. Where it is accepted that the identification of the item is accurate, for example by a user confirmation of the identification and/or a later validation of the identification from a subsequent set of images, the processor can update a local store so that a future identification of that item can be made with a higher confidence. In some examples, the processor is configured to update the database to include the additional data to enable a future match to be made.

Thus the processor may be configured to improve the 3D models against which items are matched by a learning process.

Suitably the three-dimensional representation is a depth map. This permits the extraction of depth information from the images. This can be useful in reconstructing the scene, and in identifying items.

Suitably the processor is configured to process the plurality of images by determining a geometrical relationship between each of the plurality of images. The processor is suitably configured to process the plurality of images by aligning one or more planes of the plurality of images based on the determined geometrical relationship to form aligned images. The processor may be configured to process the plurality of images by matching one of a point and a feature in one of the aligned images with a corresponding one of a point and a feature in another of the aligned images. The processor may also be configured to process the plurality of images by determining disparity information from the matched point or feature; and determining the three-dimensional representation of the scene based on the disparity information.

This group of functions can together achieve 'structure from motion'. The group of functions are executed on the image sequence (i.e. the two or more images taken by the camera) to reconstruct the scene and extract a 3D depth map of the scene. The flow above illustrates a typical flow of function operation.

Computing the geometrical relationship between each of the plurality of images is a way of performing stereo calibration on the images. Computing the geometrical relationship between each picture in space allows images to be mapped onto one another to permit comparison of the images.

Stereo calibration can be performed to calibrate the camera and to get the required intrinsic and extrinsic parameters to be used in subsequent processing. Intrinsic parameters include focal length, image format and the principal point. Extrinsic parameters can be used to represent a coordinate transformation from the scene to the projection of the scene imaged by the camera.

Aligning one or more planes of the plurality of images based on the determined geometrical relationship to form aligned images is a way of performing stereo rectification. It involves aligning one or more planes of multiple pictures to permit comparison between the aligned images. In the rectification process, a transformation of each image plane is typically derived so that after the transformation corresponding lines in the planes of each image are collinear, and are usually also parallel to an axis of the image (for example the *x*-axis). Corresponding points in the images will therefore be found along lines in each of the images which will have the same vertical coordinates in the images. Rectification simplifies the determination of matching points in the images, by reducing it to a 1D search problem (for example along a particular line in an image). This can lead to an increase in the speed of the matching operation.

Suitably, the processor is configured to, when aligning one or more planes of the plurality of images, use at least one of Bouguet's algorithm and Hartley's algorithm. Bouguet's algorithm is provided as part of a MATLAB toolbox for camera calibration. Bouguet's algorithm is suitably used if calibrated.

Hartley's algorithm is a description of a normalised eight-point algorithm, which can be used to estimate the fundamental matrix or essential matrix **E** (which relates corresponding points

in stereo images). The algorithm typically uses eight or more corresponding image points (i.e. eight pairs of points, one of each pair being in a respective one of a pair of images) Fewer than eight point pairs can be sufficient when using variations of this algorithm. Since **E** has five degrees of freedom, it can be sufficient to use 5 pairs of points. Hartley's algorithm is suitably used if uncalibrated.

Matching one of a point and a feature in one of the aligned images with a corresponding one of a point and a feature in another of the aligned images is a way of performing stereo correspondence. It permits the matching of 3D points in multiple pictures or images. The point can correspond to a single pixel, or to a group of pixels. The feature can correspond to a feature in the image, such as a detected edge and so on. Correspondence between two images can be found in one of several ways. In a correlation-based way, checks can be made to see if a location in one image appears to match a location in another image. In a feature-based way, one or more feature can be found in an image, and the arrangement of aspects of that feature can be compared with a feature in another image.

Suitably, the processor is configured, for example when matching the one of the point and the feature, to use at least one of a block matching algorithm and a semi-global block matching algorithm. Images typically comprise a number of blocks, such as macroblocks, and subsequent images in the sequence will usually comprise the same or similar blocks, offset from the position of the block in the preceding image by a displacement vector. A block matching algorithm enables matching blocks to be located in the sequence of images. Analysing blocks rather than individual pixels can reduce the processing load required in this step.

Thus, the above group of functions permits reconstruction and extraction of a 3D depth map for a scene. The 3D information can then be used to enable additional techniques such as 3D object matching and 3D model learning to enhance performance.

Benefits of the approach described above include:

- the exploitation of the natural, guided motion of a door mounted to a container enables reliable and repeatable 3D reconstruction without direct user intervention,
- the performance of object matching and classification can be enhanced, and
- the depth information can be displayed to the end-user allowing them access to this additional information.

The following discussion is made in the context of the example imaging apparatus of figure 1a.

Additional enhancements can be achieved by providing one or more additional units as part of or for coupling to the imaging apparatus. These will be discussed below.

In one example, the imaging apparatus further comprises at least one of an inertial measurement unit (IMU) and a gyroscope configured to output data indicating location. The processor is configured to receive the data and to use the data when at least one of determining the geometrical relationship between each of the plurality of images and aligning one or more planes of the plurality of images. Using the location data from an IMU or gyroscope can enhance the performance of the stereo calibration and/or rectification.

The data indicating location might comprise location data and/or data from which the location can be calculated. For example, a position detector can be configured to output position data (for example, after a suitable calibration process). In another example, an accelerometer can be configured to output data regarding the acceleration of the accelerometer with time. From this data, and a knowledge of the extent of movement of the accelerometer (for example, the extent of movement of the fridge door to which the accelerometer is mounted), the location of the accelerometer, and hence of the camera, can be determined.

In one example, data can be extracted from an IMU physically mounted to the camera, where the IMU has been zeroed (or otherwise calibrated) at the door close position, to retrieve the relative position of the camera to the enclosure or container. This information is then used to guide the function utilised to perform stereo calibration and/or stereo rectification. This can help to reduce the number of pixels that must be processed to complete those functions. This results in more efficient processing. This can reduce the time required to extract 3D information from a sequence of images and/or reduce the processing power and thus cost required to extract 3D information from a sequence of images. It can also help save battery power in computing the functions by reducing the load on the processor.

Additional enhancements can be achieved by providing for additional views to be provided by the imaging apparatus. The camera is arranged to image a particular scene, for example

the inside of a fridge. To do this, the camera is pointed towards the inside of the fridge. The camera suitably has an imaging axis along which it points. Thus, in this example, the imaging axis of the camera will point towards the inside of the fridge.

Suitably, the imaging apparatus further comprises an additional imaging device 1002 arranged so that an imaging axis of the imaging device 1004 points in a different direction to an imaging axis of the additional imaging device 1002, the additional imaging device being configured to be controlled by the control module 1006. The additional imaging device (camera) 1002 can, in one example, be mountable within the enclosure. For example, the additional camera can be mounted within the fridge, such as by being mounted to the front of a shelf of the fridge. The camera is suitably directed towards the inside of the fridge to be able to take images of the contents of the fridge. The additional camera is suitably directed towards the fridge door to be able to take images of items stored in the fridge door. Suitably, the camera and the additional camera take images of different scenes. This means that the resulting images cover a larger area/volume to be imaged, which helps provide more information from the camera and the additional camera.

The additional camera is also controlled by the control module. The additional camera can comprise, or can be coupled to, a receiver such as a wireless receiver. Suitably the additional camera comprises or is coupled to a wireless transceiver 1008. The receiver or transceiver permits the additional camera to communicate with the control module 1006. The additional camera and the control module can suitably communicate over a communication network comprising one or more of a Bluetooth, Wi-Fi, GSM, LTE etc. communication network. The control module is suitably configured to transmit a control signal to the additional camera over this communication network.

Enabling control of the additional camera by the control module, as well as enabling control of the camera by the control module, permits coordination of the additional camera with the camera. The additional imaging device is suitably configured to be controlled by the control module to take an additional image in coordination with the plurality of images. Thus, the additional camera can be controlled so as to take an image (for example of the fridge door) at the same time that the camera takes an image (for example one of the plurality of images), or when the camera (and hence the door) is in a suitable position such that the image taken by the additional camera can reveal more information about what items are in the door. Since, in the example of taking images of a fridge interior, many frequently used

items are stored in the fridge door, it is important to be able to image this region too, so as to obtain a full picture of the content of the fridge. The coupling together of the additional camera with the camera by using a wireless protocol enables a completely wireless solution to be achieved.

In one example, the additional camera is suitably controlled to take an image of the door when the door is nearly closed. This will mean that the image of the door will be as flat as possible, i.e. that the image of the door is as planar as possible with respect to the plane of the door itself, which can reduce or minimise distortion when processing this image. Taking an image in this way means that processing of the image to be able to identify items in the image can be reduced. In one example, the additional camera is controlled to take an image of the door when the door is at a particular angle from its closed position. The particular angle is suitably about 5 degrees. The particular angle may be 5 degrees. The angle can be automatically determined and/or user-selectable.

The interior of the fridge, or more generally, the enclosure, may be dark when the additional camera is controlled to take the image. Suitably at least one of the control module and the additional imaging device is configured to output a light control signal for controlling a light 1010. The light may be a separate light within the enclosure, such as a fridge light, or the light may be provided as an additional light coupled to at least one of the control module and the additional camera so as be able to receive the light signal and to turn on in response. The additional light may be a flash, such as a camera flash. In one example, the additional imaging device comprises the light responsive to the light control signal.

Suitably the imaging apparatus comprises a light sensor 1012, and the light control signal is output in dependence on a light level sensed by the light sensor. The light sensor 1012 may be provided coupled to the additional camera 1002. This means that the light is only turned on when necessary, i.e. when the additional camera determines via the light sensor, that the light level is insufficient to take an image. The light sensor may be coupled to the control module 1006. More than one light sensor may be provided.

The provision of the light sensor and the light control signal means that the light level when the additional camera takes an image will always be a certain minimum light level. If the light level is lower than this minimum light level, then the light control signal will ensure that a light

is turned on to achieve at least the minimum light level. This means that the images taken by the additional camera will provide a good image of the items in the door.

In the above discussion, a camera is configured to image a scene such as the interior of a fridge. The image provided information about the contents of the fridge. It is also possible to obtain further information about the contents of the fridge by providing a further imaging device which can image a different part of the spectrum from the camera. Thus, where the camera is configured to image the visible part of the spectrum, a further imaging device can be provided which is configured to image a non-visible part of the spectrum. This will be discussed below.

In one example, the imaging apparatus comprises a further imaging device 1014, the further imaging device being configured to image a non-visible region of the spectrum, and the further imaging device 1014 being coupled to the imaging device 1004 for imaging the scene. The further imaging device 1014 may, in some examples, be coupled to the imaging device 1004 via the control module 1006. The provision of the further imaging device 1014 allows additional information to be obtained, for example information that cannot be obtained by the camera 1004 or the additional camera 1002. This can enhance the analysis of the data obtained. Suitably the field of view of the further imaging device 1014 is substantially overlapping with the field of view of the camera 1004. This permits the further imaging device to provide information about the items that are visible in the images obtained by the camera 1004.

The further imaging device 1014 suitably comprises an infrared sensor 1016. The infrared sensor permits thermal information relating to the enclosure interior to be obtained. This can particularly useful in the context of a fridge, since temperature fluctuations in the fridge itself, or in particular items in the fridge, can reveal information about whether the items are being stored under the correct conditions (i.e. cool enough to prevent food items spoiling too quickly).

For example, as food goes off, bacteria can tend to multiply even within sealed packets. This bacterial action is often associated with a change in the thermal properties of the item. For example, bacterial action can increase the temperature of an item. An infrared sensor will permit such thermal fluctuations, which might otherwise be imperceptible (for example, to a human observer), to be determined.

Often, a person will assess whether food in a fridge is still safe to consume by using their senses: taste, smell, sight and touch. This is not always a reliable method. Out of caution, people often throw food away if it passes the displayed expiry date of the packaging. Sometimes this food would still be edible, but there is a risk that this is not the case. With the relative inaccuracy of using human senses to check whether food is still safe, more food is typically thrown away than necessary. This is wasteful.

The provision of an infrared sensor enables a more accurate determination of the freshness of food, and whether food is safe to eat. This is achievable whether or not the food item is in a sealed packet or not. Thermal fluctuations will still be detectable through packaging. Thus using the infrared sensor permits a useful determination of whether the food is still suitable for consumption.

The thermal reading from an identified food item can be compared to a known thermal profile for that item, calibrated as appropriate for the temperature of the fridge. The known thermal profile for an item can be included in the database of information associated with that item. This comparison can indicate whether the food item is safe to eat or not.

Suitably the infrared sensor also permits a determination of the calorie contents of a food item/its calorie volume.

As discussed above, the use of the camera, and optionally also the additional camera and/or the further imaging device, permits a user to remotely monitor the content of an enclosure such as a fridge. A beneficial feature of the methods and apparatus described above is that they enable the user to be provided with information about products related to identified items in the fridge. For example, the imaging apparatus may link an item that the user has in their fridge with other foodstuffs that could be used to make a meal. Those related items might be items that the user has in stock or not. Typically the linked items would form ingredients for a recipe. The ingredient list could then be provided to the user, with an indication of which items the user already has and which might have to be purchased. The imaging apparatus could also provide the user with the recipe.

The provision of such recipes can be updated, for example if it is determined that an item in the fridge has recently gone off. The user could be prompted to use an alternative item in the fridge instead, or they could be prompted to purchase a replacement of the expired item.

Suitably, the information extracted from the image provided by the further imaging device, such as the infrared sensor, will permit a determination that the food item will shortly expire. In response to making this determination, the user can be prompted with a recipe that makes use of that item before it goes off. Doing this can help reduce wastage by prompting the user to use up food before it goes off.

A determination that a food item is about to go off could also be made based on a length of time that that item has been in the fridge. If an item of a certain type typically has a lifetime in the fridge of 7 days, then if that item is identified as being newly added to the fridge, then if 6 days, say, have elapsed and that item is still in the fridge, the user can be prompted to use the item.

Another advantageous feature of the methods and apparatus described above is that it allows additional functionality of a user device, such as a smartphone, and/or the functionality of other applications, such as If This Then That, to be combined with the functionality of the camera. As an example, the imaging apparatus, for example the control module, may provide the user with alerts about items that are running low when the user's phone indicates that the user is outside a supermarket. The imaging apparatus may thus provide the user with information not only when the remaining amount of an item is below a certain threshold, such as a pre-determined or user-selectable threshold, but also when some completely independent factor indicates that the user might be interested in receiving that information at that moment.

The methods and apparatus described above have been primarily described with reference to a domestic setting in which they are used to help a user keep track of foodstuffs in a fridge. This is just one example of a suitable application and it should be understood that the camera may be used to keep track of any substance or object.

Determining locations of objects in a scene

When imaging a scene, and more particularly objects in the scene, over time, the content of the scene is likely to change. For example, where the scene is the interior of an enclosure

such as a fridge, changes in the objects imaged will occur when items are put into the fridge and/or taken out of the fridge. Further, putting items into the fridge may obscure other items already in the fridge. It may not necessarily be clear whether an item has been taken out of the fridge, or is simply being obscured by another item subsequently added to the fridge. It would be desirable to have a way of being able to distinguish between these situations. This can be done by determining objects in the fridge (i.e. in the scene to be imaged).

One example of a method for determining objects in a scene comprises the steps of: taking a first image of the scene, the first image permitting identification of objects in the scene; taking a second image of the scene, the second image permitting identification of objects in the scene; determining from the first and second images which objects are present in the scene; and storing the first and second images for later retrieval.

The images may be taken at different times. In this case, the images can be stored to build up a picture of the contents of the scene (which might be the inside of a fridge) over time. A user may access the store to view stored images, and can therefore 'look' into the scene at different points in time.

The images may be taken by different imaging devices, such as different cameras. The different imaging devices suitably image different parts of the EM spectrum and/or image along different imaging axes. This enables alternative views of the scene to be compared.

The above techniques can allow the user (or an automated system, such as one based on image recognition techniques, for example one or more of the techniques discussed above) to determine which objects are present, including objects that might be hidden behind, or obscured by, other objects in at least one image.

The images may be processed to extract more information, including depth information. This can allow the determination of the locations of objects in a scene, such as a fridge interior.

One example of a method for determining locations of objects in a scene comprises the steps of: taking a first image of the scene 1102, identifying in the first image a first object 1104, and determining from the first image a depth of the first object in the scene 1106. The method comprises taking a second image of the scene 1108, identifying in the second image a second object 1110, and determining from the second image a depth of the second object

in the scene 1112. The method further comprises determining that the first object is not identifiable in the second image 1114. In dependence on making such a determination, the method comprises determining in dependence on the depth of the first object and the depth of the second object whether the second object is located in front of the first object or whether the second object replaces the first object 1116.

The first image may be taken at a first time. The second image may be taken at a second time later than the first time.

Typically, a single image of a fridge interior may only allow in the order of 40-50% of the content to be directly imaged, due to some items obscuring other items. Thus, by building up a sequence of images, such as a 'layered' view of the content of the scene, it is possible to more accurately determine what items are still in the scene, even if those items are not all visible in a particular image.

For example, a user may put a small item, such as a tub of butter, in the fridge, and then put a large item, such as a jug of milk, in front of the butter. Simply looking at a face-on image of the fridge will therefore only show the milk, and not the butter. The user can determine that the butter is there using one or other of the techniques above. In one technique, images are taken at different points in time. Thus, where one image is taken after putting the butter in the fridge, but before putting in the milk, the butter will be visible in that image. A second image can be taken after putting the milk in the fridge too. The butter will not be visible in this second image. A comparison of the two images can then be used to find out that there is some butter behind the milk.

Alternatively, where images are taken from different orientations, such as along different imaging axes, one image might be a face-on image. In this image, the milk will be visible but the butter will not. In another image, taken at a different angle and/or from a different location, the butter may be visible, by 'looking behind' the milk. Thus taking multiple images at different angles and/or locations can help reveal the contents of the scene (for example the fridge contents).

Suitably these techniques can use images taken to provide a three-dimensional image of a scene, as discussed above. Preferably, the images that are taken are stored for later retrieval (either before or after image processing) and are not discarded.

The first image and/or the second image may be a composite image or a series of images. This can allow 3D depth information to be extracted from the first image and/or the second image. The 3D depth information is suitably extracted from the image as described above.

The images are suitably taken by an imaging device 1202 such as a camera. The camera is suitably coupled to a processor 1204 for processing the images. The processor is suitably coupled to a transceiver 1206 such as a wireless transceiver, for example one that is configured to communication using at least one of Bluetooth, Wi-Fi, GSM and LTE. The transceiver permits the processor to communicate, for example bi-directionally, with a remote device 1208 and/or the cloud 1210. The processor may also be coupled to a local storage device 1212.

The images, for example the layered images, can be stored in the local storage device 1212 and/or remotely. For example, the images can be stored at a storage region 1214 in the remote device 1208 and/or at a storage region 1216 in the cloud 1210.

The processing of the images to identify objects in the images and determine the depth of the objects in the scene may be performed at the processor 1204, at the remote device 1208, and/or in the cloud 1210.

The method suitably comprises determining that the first depth is greater than the second depth. This information can be useful in determining that the second object is obscuring the first object.

Object recognition, for example 3D object recognition, can be used to determine the depth of the first object and/or the second object in the image. Suitably depth in the image can be determined by considering differences in two or more images of the same scene from a different perspective. For example, an algorithm that considers 5 points in each image is suitably used to determine depth information. In one example, Hartley's algorithm (introduced briefly above) can be used.

At least some of the structures shown in the figures (any block apparatus diagrams included herein) are intended to correspond to a number of functional blocks in an apparatus. This is for illustrative purposes only. The figures are not intended to define a strict division between

different parts of hardware on a chip or between different programs, procedures or functions in software. In some embodiments, some or all of the algorithms described herein may be performed wholly or partly in hardware. In many implementations, at least the controller, zone identifier and product weight manager will be implemented by a processor acting under software control. Any such software is preferably stored on a non-transient computer readable medium, such as a memory (RAM, cache, hard disk etc) or other storage means (USB stick, CD, FLASH, ROM, disk etc).

The applicant hereby discloses in isolation each individual feature described herein and any combination of two or more such features, to the extent that such features or combinations are capable of being carried out based on the present specification as a whole in the light of the common general knowledge of a person skilled in the art, irrespective of whether such features or combinations of features solve any problems disclosed herein, and without limitation to the scope of the claims. The applicant indicates that aspects of the present invention may consist of any such individual feature or combination of features. In view of the foregoing description it will be evident to a person skilled in the art that various modifications may be made within the scope of the invention.

Claims

1. An imaging apparatus for imaging a scene, comprising:

an imaging device mountable to a structure which is movable relative to a scene to be imaged;

a movement sensor configured to output movement data indicative of movement of the imaging device;

a location sensor configured to output location data indicative of the location of the imaging device relative to the scene; and

a processor configured to receive the movement data from the movement sensor and, in response, to select between a high-power mode and a low-power mode of at least one of the imaging device and the location sensor, wherein more power is consumed in the high-power mode than in the low-power mode.

2. An imaging apparatus according to claim 1, in which the processor is configured to select the high-power mode in response to determining that the movement data indicates movement of the imaging device.

3. An imaging apparatus according to claim 2, in which the processor is configured to determine that the movement of the imaging device comprises vibrational movement.

4. An imaging apparatus according to claim 2 or claim 3, in which the processor is configured to determine that the movement of the imaging device comprises movement relative to the scene.

5. An imaging apparatus according to any preceding claim, in which the movement sensor comprises an accelerometer.

6. An imaging apparatus according to any preceding claim, in which the location sensor comprises at least one of a gyroscope and an inertial measurement unit.

7. An imaging apparatus according to any preceding claim, in which the location sensor has a high-power mode and a low-power mode, and the processor is configured to select the high-power mode of the location sensor in response to determining that the movement data indicates movement of the imaging device.

8. An imaging apparatus according to any preceding claim, in which the processor is configured to determine that the imaging device is at an imaging location with respect to the scene, and, in response, to control the imaging device to image the scene.

9. An imaging apparatus according to claim 8, in which the processor is configured to determine that the imaging device is at the imaging location in response to receiving the location data from the location sensor.

10. An imaging apparatus according to claim 9, in which the location data indicates that the imaging device is at the imaging location.

11. An imaging apparatus according to claim 9 or claim 10, in which the location data indicates that the imaging device is moving in a particular direction past the imaging location.

12. An imaging apparatus according to any of claims 8 to 11, in which the processor is configured to control the imaging device to image the scene in response to the received movement data from the movement sensor indicating that the imaging device is moving in a particular direction.

13. An imaging apparatus according to any of claims 8 to 12, in which imaging the scene comprises capturing a still image using the imaging device.

14. An imaging apparatus according to any of claims 8 to 13, in which imaging the scene comprises capturing a still image from a stream of images using the imaging device.

15. An imaging apparatus according to any of claims 8 to 14, in which the processor is configured to access a memory at which data indicating the imaging location is stored.

16. An imaging apparatus according to any preceding claim, in which the processor is configured to zero the location sensor when selecting between the high-power mode and the low-power mode.

17. An imaging apparatus according to any preceding claim, in which the processor is configured, in response to determining that the imaging device is at one end of a range of movement relative to the scene to be imaged, to select the low-power mode.

18. An imaging apparatus according to any preceding claim, comprising an additional imaging device arranged so that an imaging axis of the imaging device points in a different direction to an imaging axis of the additional imaging device, and the processor is configured to control the additional imaging device.

19. An imaging apparatus according to claim 18, in which the additional imaging device is configured to be controlled to capture an additional image in coordination with the image captured by the imaging device.

20. An imaging apparatus according to claim 18 or claim 19, in which the processor is configured to track a plurality of imaging locations for a plurality of imaging devices, with each imaging location being for a respective imaging device.

21. An imaging apparatus according to any preceding claim, in which the processor is configured to output a light control signal for controlling a light.

22. An imaging apparatus according to claim 21, in which at least one of the imaging device and the additional imaging device comprises a light responsive to the light control signal.

23. An imaging apparatus according to claim 21 or claim 22, comprising a light sensor, the processor being configured to output the light control signal in dependence on a light level sensed by the light sensor.

24. An imaging apparatus according to any of claims 18 to 23, configured to capture an image using the imaging device when the imaging device is at a first location and to capture an image using the additional imaging device when the imaging device is at a second, different, location.

25. An imaging apparatus according to any of claims 18 to 24, in which the additional imaging device is remote from the imaging device.

26. An imaging apparatus according to any preceding claim, in which the structure is rotatable about an axis.

27. An imaging apparatus according to claim 26, in which the scene comprises an interior of an enclosure, and the axis is aligned along an edge of the enclosure.

28. An imaging apparatus according to claim 27, in which the structure comprises a door of the enclosure.

29. A method for managing power usage in an imaging apparatus, the method comprising:

receiving movement data indicative of movement of an imaging device for imaging a scene; and

selecting between a high-power mode and a low-power mode of at least one of the imaging device and a location sensor configured to output location data indicative of the location of the imaging device relative to the scene;

wherein more power is consumed in the high-power mode than in the low-power mode.

30. Machine readable code for implementing a method according to claim 29.

31. An imaging apparatus for providing a three-dimensional image of a scene, the imaging apparatus comprising:

an imaging device mountable to a structure which is rotatably movable relative to the scene to be imaged;

a control module for controlling the imaging device to take a plurality of images of the scene in different relative locations with respect to the scene; and

a processor configured to

receive the plurality of images from the imaging device; and

process the plurality of images to construct a three-dimensional representation of the scene.

32. An imaging apparatus according to claim 31, in which the structure is rotatable about an axis.

33. An imaging apparatus according to claim 32, in which the scene comprises an interior of an enclosure, and the axis is aligned along an edge of the enclosure.

34. An imaging apparatus according to claim 33, in which the structure comprises a door of the enclosure.

35. An imaging apparatus according to any of claims 31 to 34, in which the processor is configured to:

process the three-dimensional representation of the scene to detect an object in the scene.

36. An imaging apparatus according to claim 35, in which the processor is configured to

compare the detected object with a list of known objects to determine a match for the detected object from the list of known objects.

37. An imaging apparatus according to claim 36, in which the list of known objects comprises at least one of a local database of objects and a remote database of objects.

38. An imaging apparatus according to claim 36 or claim 37, in which the processor is configured to determine a set of difference data between the detected object and the matched object from the list of known objects, and to store the set of difference data.

39. An imaging apparatus according to any of claims 31 to 38, in which the imaging device comprises a two-dimensional imaging device.

40. An imaging apparatus according to any of claims 31 to 39, in which the three-dimensional representation is a depth map.

41. An imaging apparatus according to any of claims 31 to 40, in which the processor is configured to process the plurality of images by:

determining a geometrical relationship between each of the plurality of images;

aligning on or more planes of the plurality of images based on the determined geometrical relationship to form aligned images;

matching one of a point and a feature in one of the aligned images with a corresponding one of a point and a feature in another of the aligned images;

determining disparity information from the matched point or feature; and

determining the three-dimensional representation of the scene based on the disparity information.

42. An imaging apparatus according to claim 41, in which the processor is configured to, when aligning the one or more planes of the plurality of images, use at least one of Bouguet's algorithm and Hartley's algorithm.

43. An imaging apparatus according to claim 41 or claim 42, in which the processor is configured, when matching the one of the point and the feature, to use at least one of a block matching algorithm and a semi-global block matching algorithm.

44. An imaging apparatus according to any of claims 31 to 43, in which the control module comprises a synchroniser configured to output a synchronisation signal, and the control module is configured to control the imaging device in dependence on the synchronisation signal, so as to synchronise the images taken by the imaging device.

45. An imaging apparatus according to claim 44, in which the synchroniser comprises at least one of a timer, an accelerometer and a position detector.

46. An imaging apparatus according to any of claims 41 to 45, comprising at least one of an inertial measurement unit and a gyroscope configured to output data indicating location, the processor being configured to receive the data and to use the data when at least one of determining the geometrical relationship between each of the plurality of images and aligning the one or more planes of the plurality of images.

47. An imaging apparatus according to any of claim 31 to 46, comprising an additional imaging device arranged so that an imaging axis of the imaging device points in a different direction to an imaging axis of the additional imaging device, the additional imaging device being configured to be controlled by the control module.

48. An imaging apparatus according to claim 47, in which the additional imaging device is configured to be controlled by the control module to take an additional image in coordination with the plurality of images.

49. An imaging apparatus according to claim 47 or claim 48, in which at least one of the control module and the additional imaging device is configured to output a light control signal for controlling a light.

50. An imaging apparatus according to any of claims 47 to 49, in which the additional imaging device comprises a light responsive to the light control signal.

51. An imaging apparatus according to any of claims 47 to 50, in which the imaging apparatus comprises a light sensor, and the light control signal is output in dependence on a light level sensed by the light sensor.

52. An imaging apparatus according to any of claims 31 to 51 in which at least one of the control module, the imaging device and the additional imaging device comprises a wireless transceiver for communicating with one or more of a remote device and another of the control module, the imaging device and the additional imaging device.

53. An imaging apparatus according to claim 52, in which the remote device comprises the processor.

54. An imaging apparatus according to claim 52 or claim 53, in which the remote device comprises memory for storing images taken by at least one of the imaging device and the additional imaging device.

55. An imaging apparatus according to any of claims 31 to 54, comprising a further imaging device, the further imaging device being configured to image a non-visible region of the spectrum, and the further imaging device being coupled to the imaging device for imaging the scene.

56. An imaging apparatus according to claim 55, in which the further imaging device comprises an infrared sensor.

57. A method for providing a three-dimensional image of a scene, the method comprising the steps of:

rotatably moving an imaging device relative to the scene to be imaged;

controlling the imaging device in dependence on a control module to take a plurality of images of the scene in different relative locations with respect to the scene; and

processing the plurality of images to reconstruct a three-dimensional representation of the scene.

58. Machine readable code for implementing a method according to claim 57.

59. A method of determining objects in a scene, the method comprising the steps of:

taking a first image of the scene, the first image permitting identification of objects in the scene;

taking a second image of the scene, the second image permitting identification of objects in the scene;

determining from the first and second images which objects are present in the scene; and

storing the first and second images for later retrieval.

60. Machine readable code for implementing a method according to claim 59.

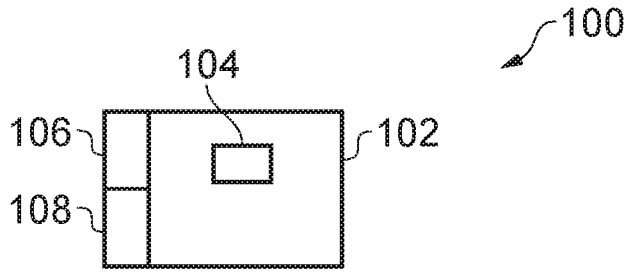

FIG. 1a

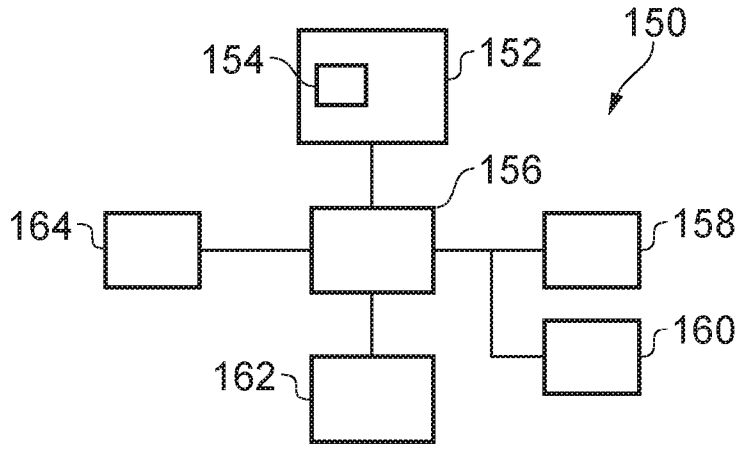

FIG. 1b

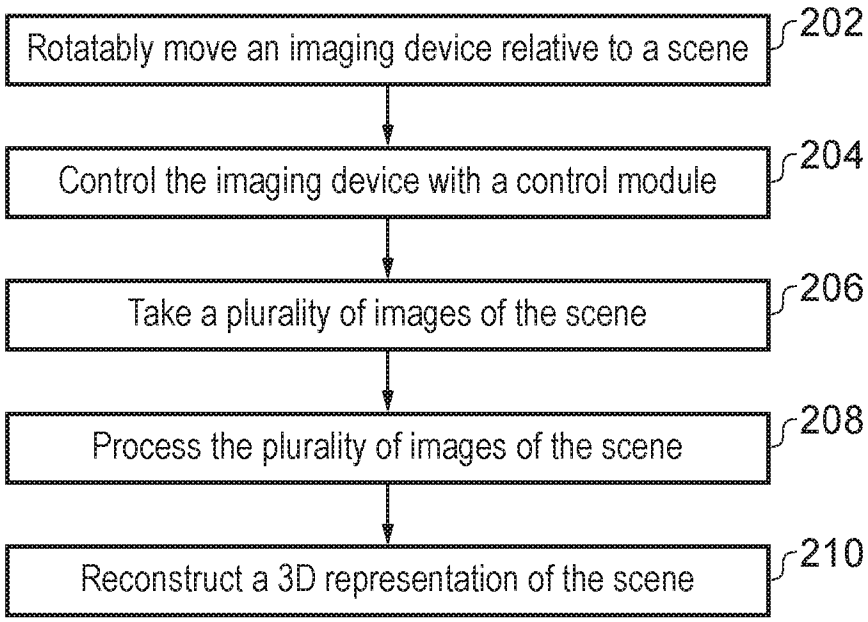

FIG. 2

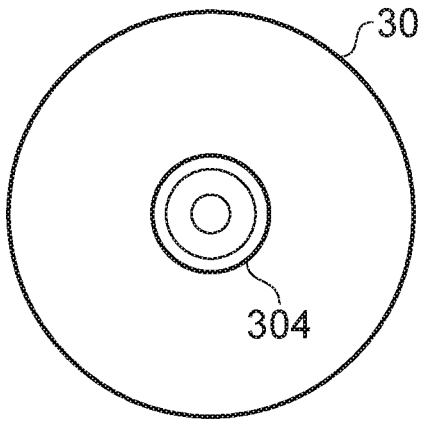

FIG. 3a

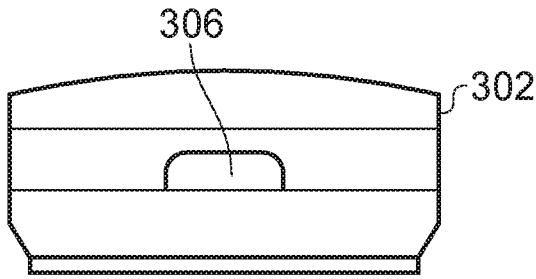

FIG. 3b

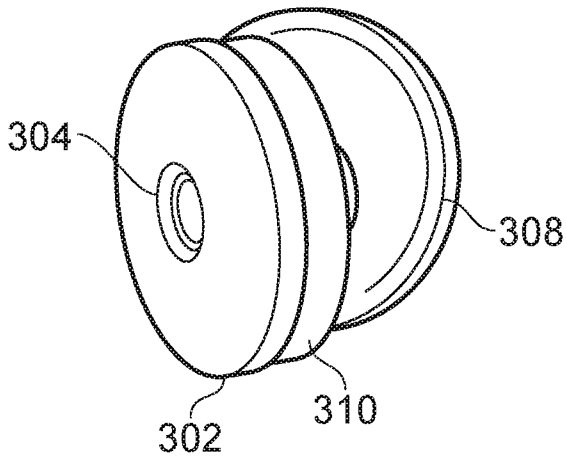

FIG. 3c

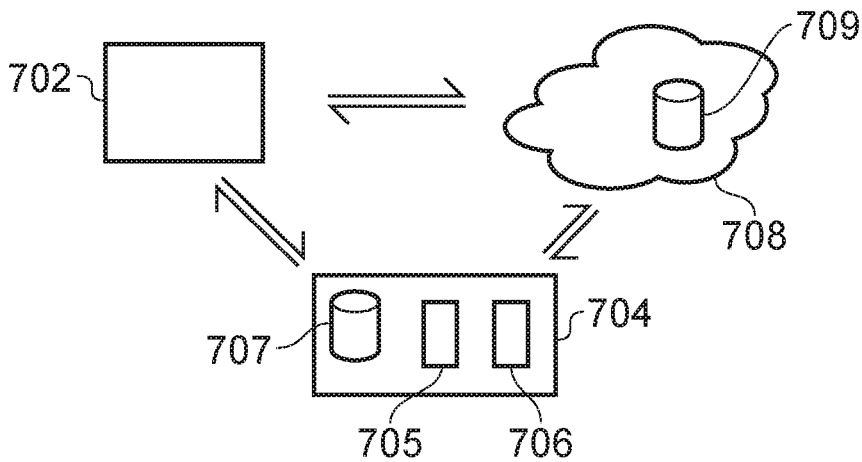

FIG. 7

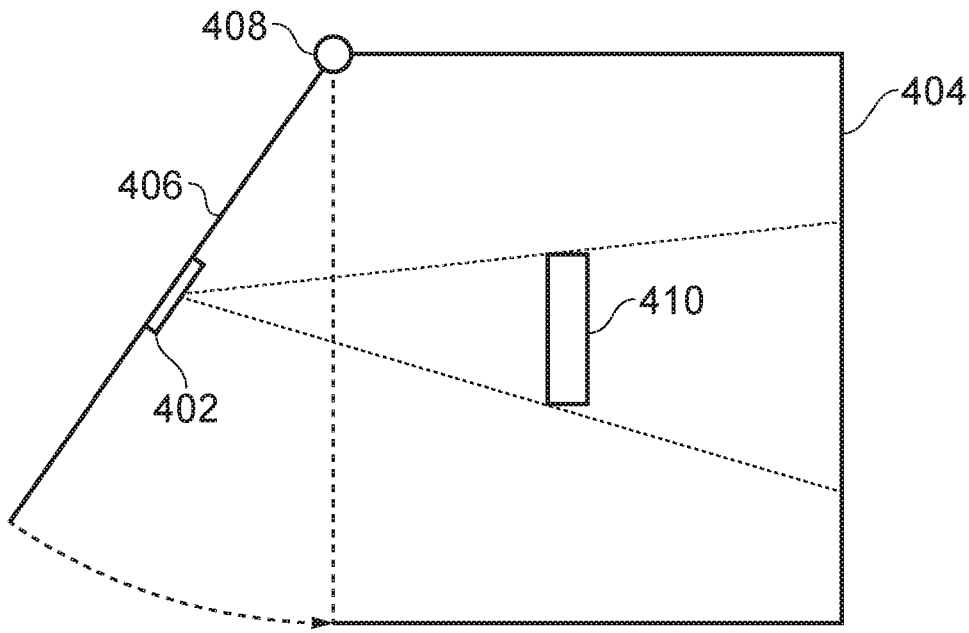

FIG. 4a

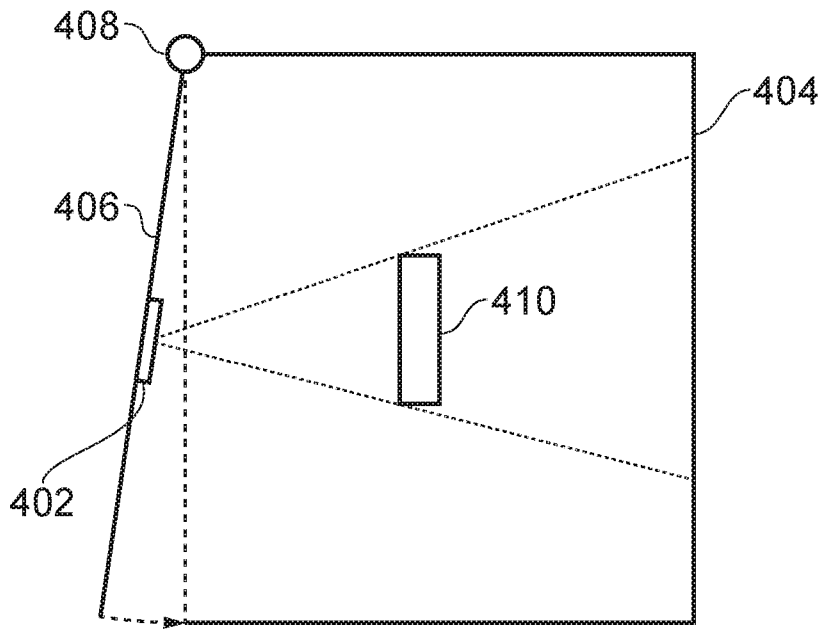

FIG. 4b

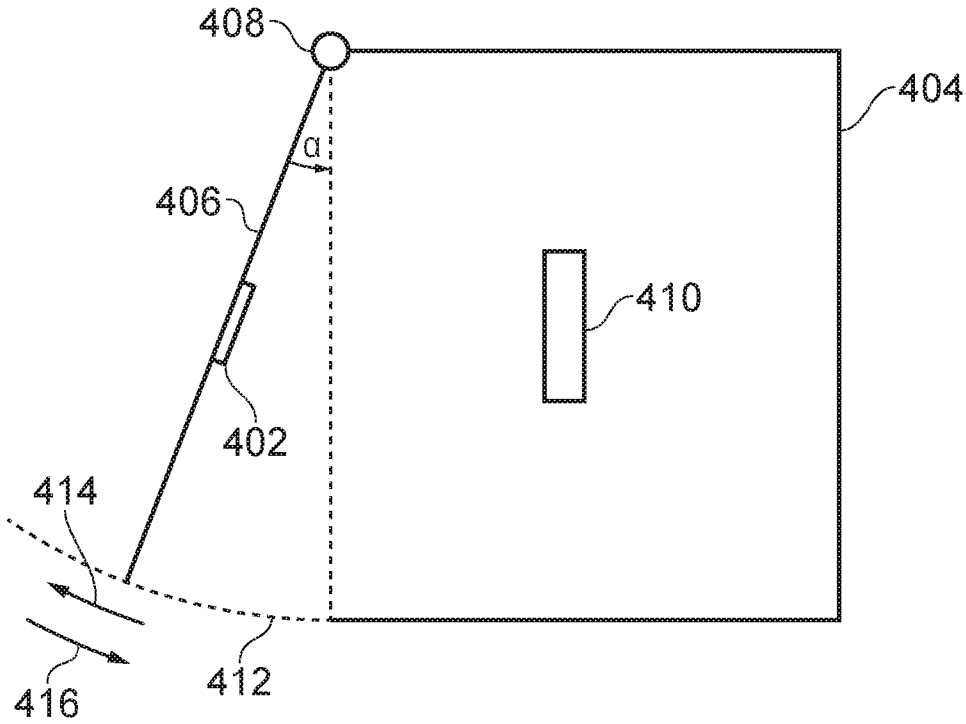

FIG. 4c

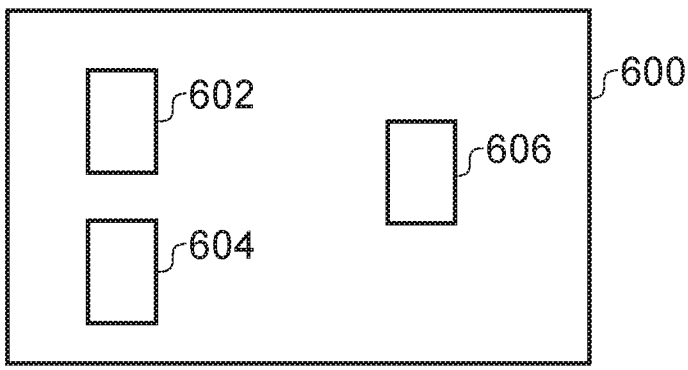

FIG. 6

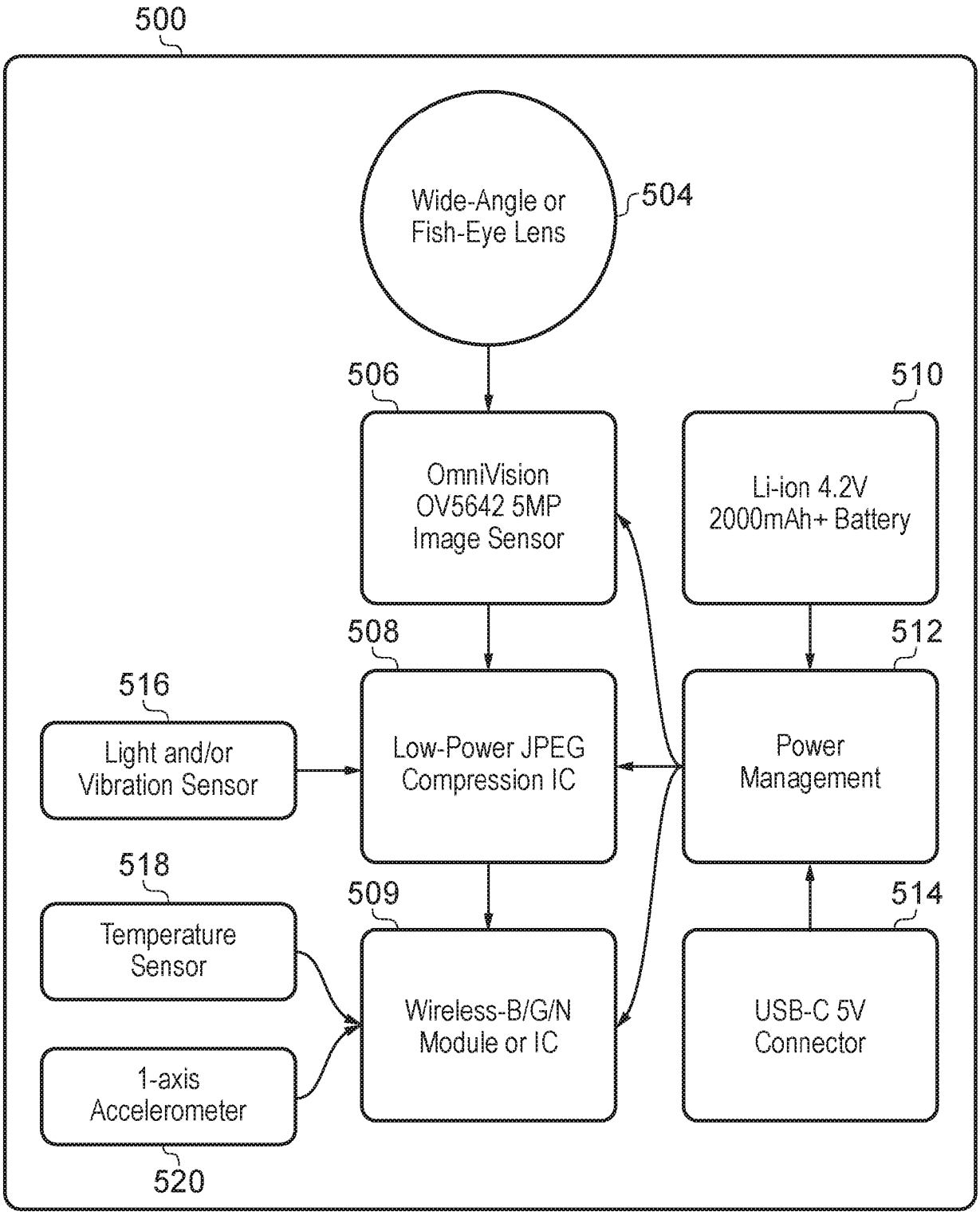
500
Wide-Angle or Fish-Eye Lens
504
506
OmniVision OV5642 5MP Image Sensor
510
Li-ion 4.2V 2000mAh+ Battery
508
512
516
Light and/or Vibration Sensor
Low-Power JPEG Compression IC
Power Management
518
509
514
Temperature Sensor
Wireless-B/G/N Module or IC
USB-C 5V Connector
1-axis Accelerometer
520

FIG. 5a

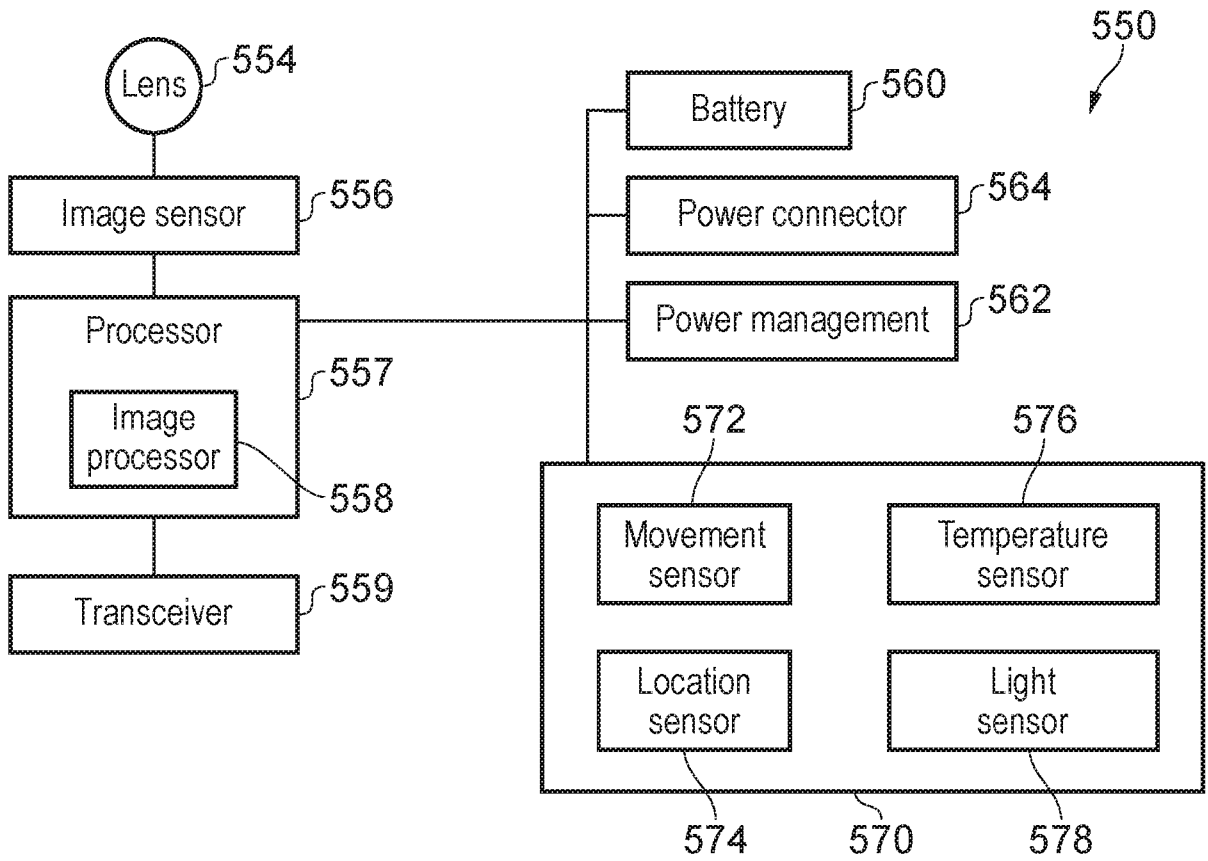
550
Lens
554
Image sensor
556
Processor
557
Image processor
558
Transceiver
559
Battery
560
Power connector
564
Power management
562
572
576
Movement sensor
Temperature sensor
Location sensor
Light sensor
574
570
578

FIG. 5b

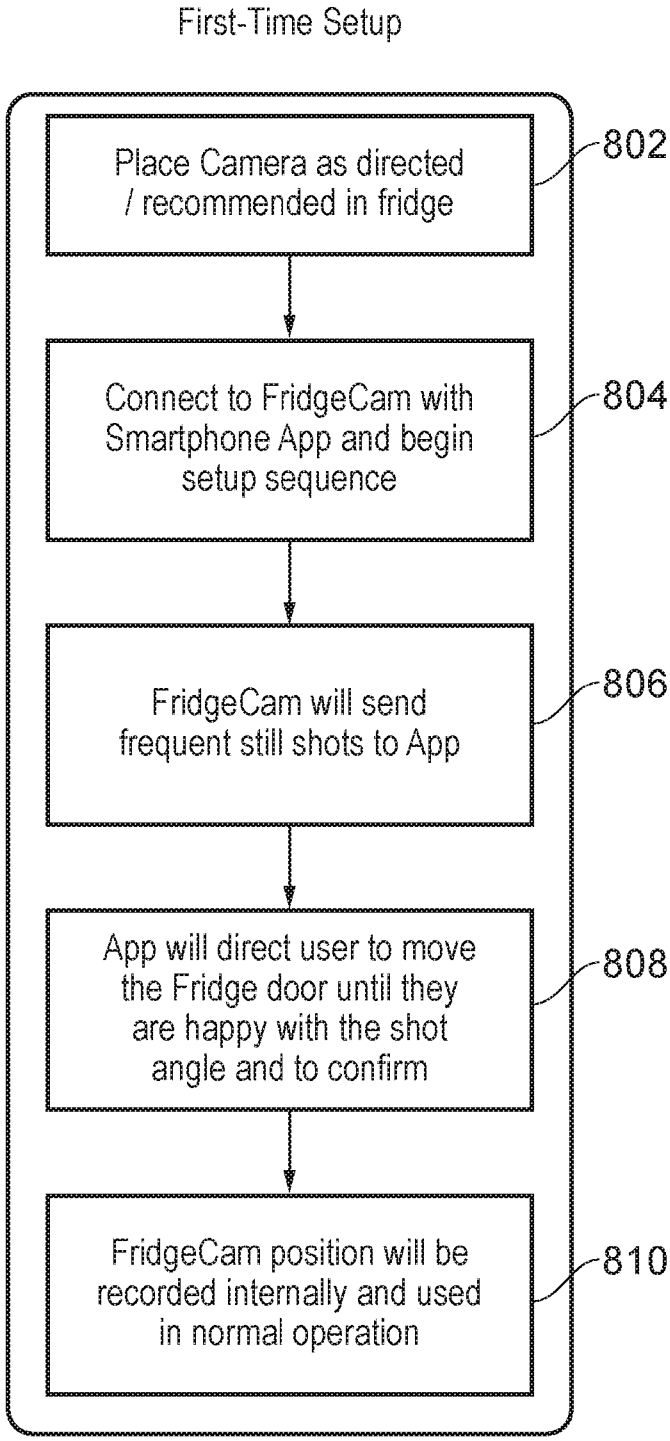
First-Time Setup
Place Camera as directed / recommended in fridge
802
Connect to FridgeCam with Smartphone App and begin setup sequence
804
FridgeCam will send frequent still shots to App
806
App will direct user to move the Fridge door until they are happy with the shot angle and to confirm
808
FridgeCam position will be recorded internally and used in normal operation
810

FIG. 8

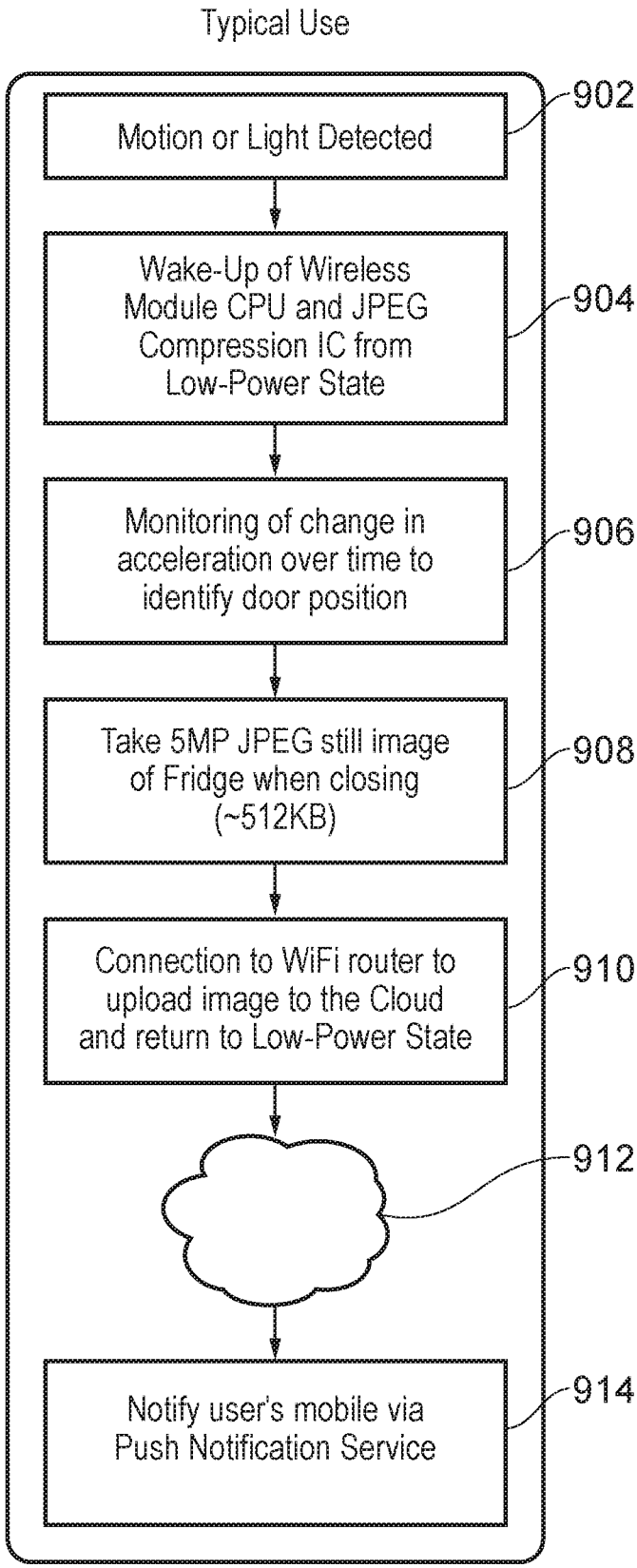
Typical Use
Motion or Light Detected
902
Wake-Up of Wireless Module CPU and JPEG Compression IC from Low-Power State
904
Monitoring of change in acceleration over time to identify door position
906
Take 5MP JPEG still image of Fridge when closing (~512KB)
908
Connection to WiFi router to upload image to the Cloud and return to Low-Power State
910
912
Notify user's mobile via Push Notification Service
914

FIG. 9a

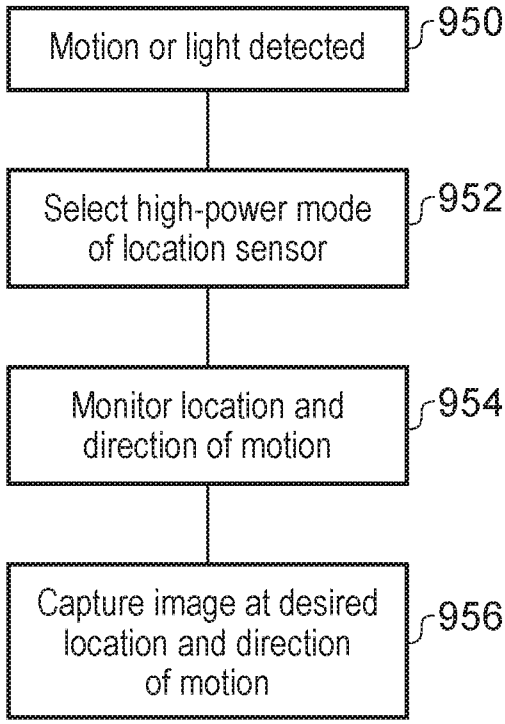
Motion or light detected
950
Select high-power mode of location sensor
952
Monitor location and direction of motion
954
Capture image at desired location and direction of motion
956

FIG. 9b

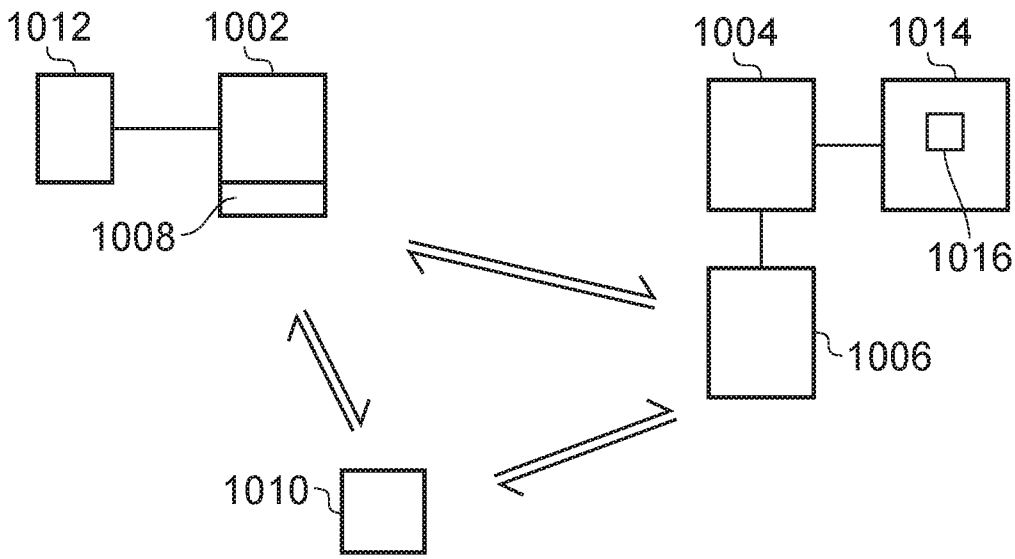
1012
1002
1008
1004
1014
1016
1006
1010

FIG. 10

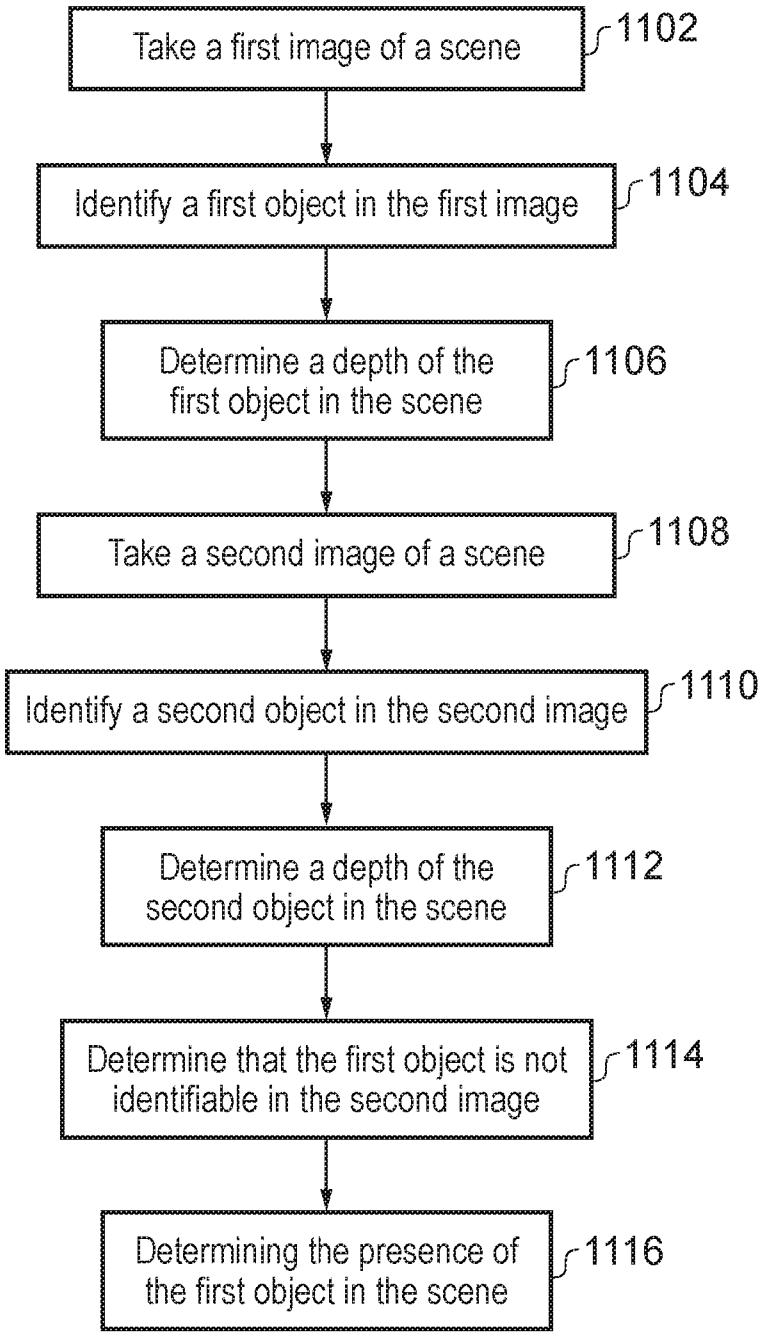
Take a first image of a scene
1102
Identify a first object in the first image
1104
Determine a depth of the first object in the scene
1106
Take a second image of a scene
1108
Identify a second object in the second image
1110
Determine a depth of the second object in the scene
1112
Determine that the first object is not identifiable in the second image
1114
Determining the presence of the first object in the scene
1116

FIG. 11

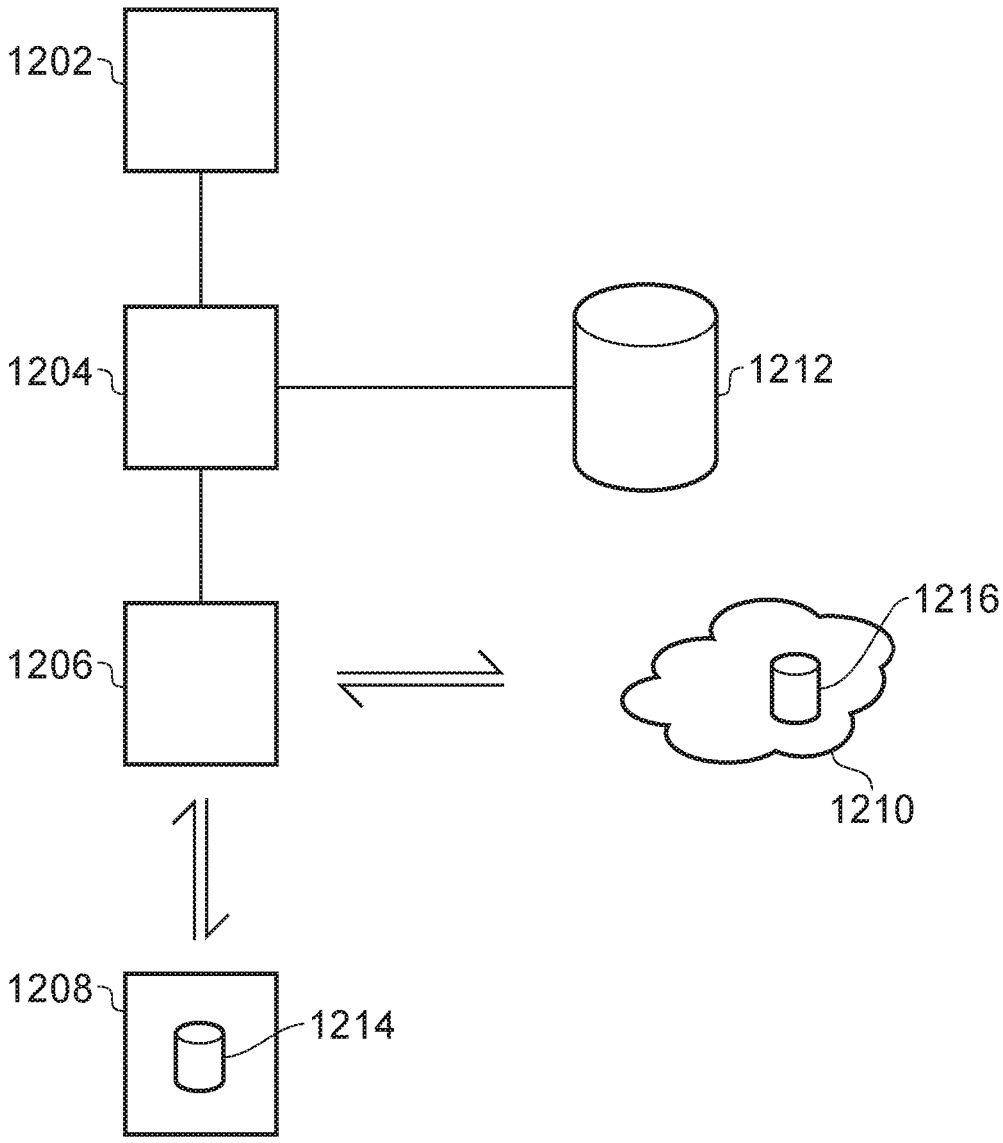

FIG. 12

INTERNATIONAL SEARCH REPORT

International application No
PCT/GB2018/050289

A. CLASSIFICATION OF SUBJECT MATTER
INV. G01C11/02 G06F1/32 G06T1/00 H04N1/00 F25D23/12
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)
G01C G06F G06T H04N F25D

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, WPI Data

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	US 2016/057394 A1 (MARUTANI YUUKI [JP] ET AL) 25 February 2016 (2016-02-25) figures 2,13A-13B,14 paragraphs [0085], [0091], [0094] - [0095], [0119], [0129] - [0131], [0150], [0179], [0219], [0376] -----	1-7,29, 30

☐ Further documents are listed in the continuation of Box C.

☒ See patent family annex.

* Special categories of cited documents :

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

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

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

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

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

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

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

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

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
25 April 2018	13/07/2018

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	Faivre, Olivier

INTERNATIONAL SEARCH REPORT

International application No.
PCT/GB2018/050289

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.:

1-7, 29, 30

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: 1-7, 29, 30

   saving power of an imaging device

   ---

2. claims: 8-17

   controlling the position where an image is taken

   ---

3. claims: 18-20, 24, 25

   imaging different areas simulataneously

   ---

4. claims: 21-23

   controlling the illumination of an image

   ---

5. claims: 26-28

   moving an imaging device in a reproducible manner

   ---

6. claims: 31-58

   acquiring a three-dimensional representation of a scene

   ---

7. claims: 59, 60

   identifying objects in a scene

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No
PCT/GB2018/050289

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
US 2016057394 A1	25-02-2016	CN 105074365 A	18-11-2015
		EP 2975342 A1	20-01-2016
		JP 2015111026 A	18-06-2015
		KR 20150128901 A	18-11-2015
		KR 20170052705 A	12-05-2017
		TW 201447204 A	16-12-2014
		TW 201700932 A	01-01-2017
		US 2016057394 A1	25-02-2016
		WO 2014142120 A1	18-09-2014