

(43) International Publication Date
15 September 2016 (15.09.2016)(51) International Patent Classification:
H04N 5/232 (2006.01)(21) International Application Number:
PCT/US2016/019142(22) International Filing Date:
23 February 2016 (23.02.2016)

(25) Filing Language: English

(26) Publication Language: English

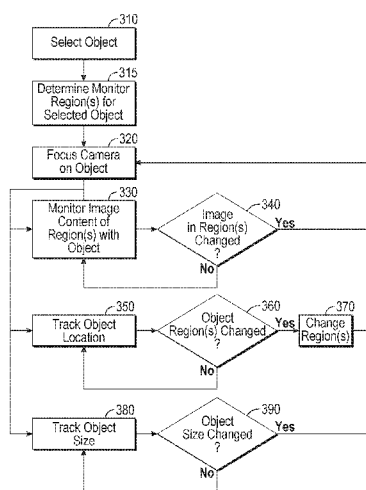
(30) Priority Data:
14/643,876 10 March 2015 (10.03.2015) US(71) Applicant: **QUALCOMM INCORPORATED** [US/US];
ATTN: International IP Administration, 5775 Morehouse
Drive, San Diego, California 92121-1714 (US).(72) Inventors: **DENG, Fan**; 5775 Morehouse Drive, San
Diego, California 92121-1714 (US). **MA, Lei**; 5775 More-
house Drive, San Diego, California 92121-1714 (US).
CHAN, Leung Chun; 5775 Morehouse Drive, San Diego,
California 92121-1714 (US).(74) Agent: **HERMANSON, Gregory A.**; KNOBBE
MARTENS OLSON & BEAR LLP, 2040 Main Street,
Fourteenth Floor, Irvine, California 92614 (US).(81) Designated States (unless otherwise indicated, for every
kind of national protection available): AE, AG, AL, AM,
AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY,BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM,
DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT,
HN, HR, HU, ID, IL, IN, IR, IS, JP, KE, KG, KN, KP, KR,
KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG,
MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM,
PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC,
SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN,
TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.(84) Designated States (unless otherwise indicated, for every
kind of regional protection available): ARIPO (BW, GH,
GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ,
TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU,
TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE,
DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU,
LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK,
SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ,
GW, KM, ML, MR, NE, SN, TD, TG).**Declarations under Rule 4.17:**— as to applicant's entitlement to apply for and be granted a
patent (Rule 4.17(ii))— as to the applicant's entitlement to claim the priority of the
earlier application (Rule 4.17(iii))**Published:**

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

(54) Title: SYSTEMS AND METHODS FOR CONTINUOUS AUTO FOCUS (CAF)

FIG. 3

300



(57) **Abstract:** Systems and methods for continuous automatic focus are disclosed. In some aspects, an object to focus on is selected from a field of view of a camera. At least one region of the field of view that includes the object is monitored for change. Regions of the field of view that do not include the object may not be monitored, thus improving efficiency and speed. If a change in the image content of the monitored region(s) occurs, the method may trigger a focus event. If the object moves to different region(s) of the field of view, the method detects the movement and adjusts the regions of the image being monitored to maintain focus on the object. In some aspects, the size of the object is also monitored. If the size of the object changes, a focus event may be triggered, depending on whether the amount of change exceeds predetermined parameters.



WO 2016/144532 A1

SYSTEMS AND METHODS FOR CONTINUOUS AUTO FOCUS (CAF)

FIELD

[0001] This application relates generally to imaging systems and more specifically to tracking objects for continuous automatic focus.

BACKGROUND

[0002] Autofocus systems automatically adjust the distance between a lens element and an image sensor to keep objects in focus. Many of these systems determine the best focus based on high frequency content within the image being focused. For example, a lens position that maximizes the high frequency content may represent an image with the best contrast ratio, and thus the best focus. However, existing autofocus systems may lose focus when an object moves relative to the camera. Therefore, there is a need for systems and methods that continuously and automatically maintain focus on moving objects.

SUMMARY

[0003] A summary of sample aspects of the disclosure follows. For convenience, one or more aspects of the disclosure may be referred to herein simply as “some aspects.”

[0004] Methods and apparatuses or devices being disclosed herein each have several aspects, no single one of which is solely responsible for its desirable attributes. Without limiting the scope of this disclosure, for example, as expressed by the claims which follow, its more prominent features will now be discussed briefly. After considering this discussion, and particularly after reading the section entitled “Detailed Description” one will understand how the features being described provide advantages that include allowing tracking objects for continuous autofocus.

[0005] One aspect is a method of focusing a camera. In various embodiments the method may include sequentially capturing a first and second image of a scene. The method may further include determining a first size of a first object in the first image. The method may further include determining a second size of the first object in the second image. The method may further include focusing the camera based on a difference between the first size and the second size.

[0006] For some implementations, the method may further include determining a first bounding region includes the first object in the first image. For some implementations, the method may further include determining a second bounding region includes the first object in the second image. For some implementations, the method may further include focusing the camera based on a difference between the first bounding region and the second bounding region.

[0007] For some implementations, the method may further include focusing the camera based on a difference between an area of the first bounding region and an area of the second bounding region.

[0008] For some implementations, the method may further include determining a ratio between an area of the second bounding region and an area of the first bounding region. For some implementations, the method may further include focusing the camera based on the ratio being less than a size decrease threshold or exceeding a size increase threshold.

[0009] For some implementations, the method may further include focusing the camera if the difference exceeds a region difference threshold. For some implementations, the method may further include maintaining the focus of the camera if the difference does not exceed the region difference threshold.

[0010] For some implementations, the method may further include segmenting the first and second image into a plurality of corresponding regions. For some implementations, the method may further include determining a first region in the first image includes the first object, and a corresponding second region in the second image. For some implementations, the method may further include determining a first difference between image content in the first region and the second region. For some implementations, the method may further include focusing the camera based on the first difference in image content

[0011] For some implementations, the method may further include excluding from consideration when focusing the camera, regions of the plurality of regions that are determined to not include the first object.

[0012] For some implementations, the method may further include determining a third region in the second image includes the first object. For some implementations, the method may further include capturing a third image of the object at a third time, the third time after the second time. For some implementations, the method

may further include determining a fourth region in the third image corresponding to the third region. For some implementations, the method may further include determining a second difference between image content in the third region and the fourth region. For some implementations, the method may further include focusing the camera in response to the second difference in image content.

[0013] Another aspect is an imaging apparatus. The imaging apparatus may include a camera configured to sequentially capture a first and second image of a scene. The imaging apparatus may include a processor in communication with the camera. The processor may be configured to determine a first size of a first object in the first image. The processor may be further configured to determine a second size of the first object in the second image. The imaging apparatus may include a camera controller configured to focus the camera based on a difference between the first size and the second size.

[0014] For some implementations, the processor is further configured to determine a first bounding region includes the first object in the first image. For some implementations, the processor is further configured to determine a second bounding region includes the first object in the second image. For some implementations, the processor is further configured to focus the camera based on a difference between the first bounding region and the second bounding region.

[0015] For some implementations, the camera controller is further configured to focus the camera based on a difference between an area of the first bounding region and an area of the second bounding region.

[0016] For some implementations, the processor is further configured to determine a ratio between an area of the second bounding region and an area of the first bounding region. For some implementations, the camera controller is further configured to focus the camera based on the ratio being less than a size decrease threshold or exceeding a size increase threshold.

[0017] For some implementations, the camera controller is further configured to focus the camera if the difference exceeds a region difference threshold. For some implementations, the camera controller is configured to maintain the focus of the camera if the difference does not exceed the region difference threshold.

[0018] For some implementations, the processor is further configured to segment the first and second image into a plurality of corresponding regions. For some implementations, the processor is further configured to determine a first region in the first

image includes the first object, and a corresponding second region in the second image. For some implementations, the processor is further configured to determine a first difference between image content in the first region and the second region. For some implementations, the camera controller is further configured to focus the camera based on the first difference in image content.

[0019] For some implementations, the processor is further configured to exclude from consideration, when focusing the camera, regions of the plurality of regions that are determined to not include the first object.

[0020] For some implementations, the processor is further configured to determine a third region in the second image includes the first object. For some implementations, the processor is further configured to capture a third image of the object at a third time, the third time after the second time. For some implementations, the processor is further configured to determine a fourth region in the third image corresponding to the third region, wherein the processor is further configured to determine a second difference between image content in the third region and the fourth region, and wherein the camera controller is further configured to focus the camera in response to the second difference in image content.

[0021] Another aspect is an imaging apparatus. The apparatus may include means for sequentially capturing a first and second image of a scene. The apparatus may include means for determining a first size of a first object in the first image. The apparatus may include means for determining a second size of the first object in the second image. The apparatus may include means for focusing the capturing means based on a difference between the first size and the second size.

[0022] For some implementations, the capturing means may include a camera. For some implementations, the first size determining means may include a processor. For some implementations, the second size determining means may include a processor. For some implementations, the focusing means may include a camera controller.

[0023] Another aspect is a non-transitory computer-readable medium storing instructions for focusing a camera, the instructions when executed that, when executed, perform a method. The method may include sequentially capturing a first and second image of a scene. The method may include determining a first size of a first object in the first image. The method may include determining a second size of the first object in the

second image. The method may include focusing the camera based on a difference between the first size and the second size.

BRIEF DESCRIPTION OF THE DRAWINGS

[0024] FIG. 1 illustrates an example of an apparatus (for example, a mobile communication device) that includes an imaging system that can record images of a scene.

[0025] FIG. 2 is a block diagram representation of an example of an embodiment of an imaging system that tracks an object for continuous autofocus.

[0026] FIG. 3 is a flowchart that illustrates an example of a method for continuous autofocus.

[0027] FIG. 4 is a representation of a scene including an object within a monitor region at three different times: after focusing on the object, before the object moves, and after the object moved to a new position.

[0028] FIG. 5 is a representation of an object moving from a first position in a first monitor region in a first image, to a second position in a second monitor region in a second image, for which the change in monitor region triggers a switch in monitor regions as well as a focus event.

[0029] FIG. 6 shows a representation of an object within a bounding box at different size ratios as compared to a first size of the object, for which a focus event is triggered when the change in size ratio is less than a bounding box ratio decrease threshold, or when the change in size ratio is greater than a bounding box ratio increase threshold.

[0030] FIG. 7 is a flowchart that illustrates a method of a process for continuous autofocus according to the embodiments described herein.

DETAILED DESCRIPTION OF CERTAIN EMBODIMENTS

[0031] The following detailed description is directed to certain specific embodiments of the invention. However, the invention can be embodied in a multitude of different ways. It should be apparent that the aspects herein may be embodied in a wide variety of forms and that any specific structure, function, or both being disclosed herein is merely representative. Based on the teachings herein one skilled in the art should appreciate that an aspect disclosed herein may be implemented independently of any

other aspects and that two or more of these aspects may be combined in various ways. For example, an apparatus may be implemented or a method may be practiced using any number of the aspects set forth herein. In addition, such an apparatus may be implemented or such a method may be practiced using other structure, functionality, or structure and functionality in addition to or other than one or more of the aspects set forth herein.

[0032] Further, the systems and methods described herein may be implemented on a variety of different computing devices that hosts a camera. These include mobile phones, tablets, dedicated cameras, wearable computers, personal computers, photo booths or kiosks, personal digital assistants, ultra-mobile personal computers, and mobile internet devices. They may use general purpose or special purpose computing system environments or configurations. Examples of computing systems, environments, and/or configurations that may be suitable for use with the invention include, but are not limited to, personal computers, server computers, hand-held or laptop devices, multiprocessor systems, microprocessor-based systems, programmable consumer electronics, network PCs, minicomputers, mainframe computers, distributed computing environments that include any of the above systems or devices, and the like.

[0033] Autofocus systems measure a focus parameter at different distances between a lens element and an image sensor, and select the distance that optimizes the focus parameter. For example, the focus parameter may be a measure of high frequency content in a portion of a field of view, since in-focus images typically have more high frequency content than out-of-focus images. Existing systems reduce the computational load of measuring the focus parameter of the entire field of view by determining the focus parameter based on a portion of the field of view. The portion may be a center region of the field of view, a region selected by a user, or a region including an object. The object may be detected automatically (for example, using automatic face recognition) or selected by a user. Whether the center field of view, a region selected by a user, a region including an object detected automatically, or a region including an object selected by a user, focus measures for this portion of the field of view are computed at different distances between a lens element and an image sensor, and the distance that optimizes the focus parameter is selected. A controller moves the lens to the distance with the optimum focus parameter to focus the image. However, existing autofocus systems may lose focus when objects move.

[0034] An object may move within a portion of the field of view monitored by the autofocus system, or to a portion of the field of view not monitored by the autofocus system. In addition to lateral movement, an object may appear larger as it moves closer to the camera, or appear smaller as it moves farther from the camera. Each of these events may cause an object to lose focus. Therefore, a continuous autofocus system disclosed herein may trigger a focus event when the object moves.

[0035] The object to be focused on may be selected using a touchscreen, or determined automatically using image processing and pattern recognition techniques. The object may be enclosed in an enclosure, such as a bounding box or a convex hull. A bounding box, typically a small, rectangular shape that includes the object, may be used to measure focus parameters. The bounding box may be a minimum bounding box. In some aspects, the sides of the bounding box may be aligned with rows and columns of the image.

[0036] The continuous autofocus system may divide its field of view into non-overlapping monitor regions. For example, some implementations may have monitor regions consisting of four rows and four columns (4x4). In some implementations, the number of rows and columns may be equal, such as 2x2, 3x3, 4x4, 8x8, 16x16, or in some other implementations, the monitor regions may be asymmetrical, and be arranged, for example, as 2x4 or 4x2. The continuous autofocus system may monitor a subset of these regions to determine focus parameters instead of monitoring the entire field of view. Some implementations may measure one, two, three, or four regions for example. In implementations with 12 regions, for example, if only one region is monitored, just one-twelfth of the field of view is monitored.

[0037] The disclosed continuous autofocus system may first identify a subset of all the regions that may include the object. For some implementations, “including an object” may be defined to allow for an object to be included in just one region. These implementations may, for example, define including an object as including one of a centroid of the object, a center of the bounding region, a larger portion of the area of the object than any other region, or a larger portion of the area of the bounding region than any other region. For some implementations, “including an object” may be defined to allow for an object to be included in more than one region. These implementations may, for example, define including an object as including one of a portion of the object or a portion of the bounding region.

[0038] After identifying a subset of all the regions that may include the object, the disclosed system focuses the object using focus parameters for the subset. Other regions not in the subset may be excluded from consideration when focusing the camera. Once focused, the continuous autofocus system compares regions in the subset for the focused frame to corresponding region(s) in subsequent images. When there is a difference, for example, when the difference exceeds a threshold, the continuous autofocus system may trigger a focus event. If the object moves to a different region that is not in the current subset of monitored regions, then the continuous autofocus system may include the new region in the subset of regions. Additionally, other regions may be removed from the subset of regions being monitored. A focus event may also be triggered when the object moves to a different region.

[0039] In some aspects, a single region may be monitored. In these aspects, if the object spans more than one region, the monitored region for an object may be assigned based on criteria such as the region with the object centroid, a region with the center of a bounding box, or a region having the largest percentage of the object, or the region with the largest percentage of the bounding box/enclosure of the object.

[0040] The continuous autofocus system may track the size of an object since a previous focus event. The size may change, for example, if the object moves closer to or farther from the camera. The size of the object may be measured within each image directly, or in other words, based on a boundary outline of the object itself. Alternatively, a size of a bounding box/enclosure for the object may be used to estimate the relative size of the object. The change in an area of a bounding box may be used as an estimate of the change in object size since a previous focus event. For some implementations, the continuous autofocus system uses a ratio or a percent change in size since the last focus event.

[0041] Thus, some implementations trigger focus events to improve focus on moving objects when there is a change in the size of an image of an object, there is a change in which region(s) is/are being monitored, i.e. in which region(s) the object is positioned, or there is a change in the image content of the monitored region(s). The focus may also be accomplished more effectively than existing systems, and with greater computational efficiency.

[0042] FIG. 1 illustrates an example of an apparatus 100 (for example, a mobile communication device) that includes an imaging system that can record images of

a scene. The apparatus 100 includes a display 120. The apparatus 100 may also include a camera on the reverse side of the apparatus, which is not shown. The display 120 may display images captured within the field of view 130 of the camera. Processing logic within the apparatus 100 may divide the field of view 130 into a plurality of regions with region boundaries 140. FIG. 1 shows object 150, a person, within the field of view 130. The object 150 is also within the region 145. Processing logic within the apparatus 100 may detect the object 150 and determine a bounding region that encloses the object 150. In some aspects, the determined region may be a polygon or a rectangle. As shown, processing logic within the apparatus 100 determined a bounding region (bounding box) 160 that encloses the object 150. In this example, bounding box 160 is aligned with the rows and columns of imaging elements that make up the field of view 130. In some aspects, the bounding region may be only large enough to enclose the object 150, and thus be a minimum bounding box.

[0043] FIG. 2 is an internal block diagram an embodiment of the apparatus 100 shown in FIG 1. The apparatus 100 includes a camera 110, a display 120, user input device 240, an electronic processor 250, and an electronic memory 260. The processor 250 is operably connected to the display 120, camera 110, user input device 240, and memory 260. The camera 110 includes a lens 210, a camera controller 220, and an imaging sensor 230. The camera controller 220 and imaging sensor are in communication with a processor 250.

[0044] The imaging sensor (image sensor, sensor) 230 captures images at a spatial resolution limited by the resolution of the sensor 230. The sensor 116 may comprise rows and columns of picture elements (pixels) that may use semiconductor technology, such as charged couple device (CCD) or complementary metal oxide semiconductors (CMOS) technology, that determine an intensity of incident light at each pixel during an exposure period for each image frame. In some embodiments, incident light may be filtered to one or more spectral ranges to take color images. For example, a Bayer filter mosaic on the sensor 116 may filter light using red, green and blue filters capture full color, three band images.

[0045] The camera controller 230 may operably control movement of the lens 210 (or at least one lens element) for focusing, control the size of the aperture and/or how long the aperture is open to control exposure (and/or the exposure period), and/or control sensor 230 properties (for example, gain).

[0046] The memory 260 may store processor instructions that configure the processor 250 to perform functions. For example, the instructions may configure the processor to capture one or more images using the camera 110, and to display the captured images on the display 120. The instructions may also configure the processor 250 to perform a variety of image processing tasks, including the imaging processing methods discussed herein.

[0047] For some embodiments, the processor 160 may perform (or process) functionality associated with blocks in the flowchart depicted in FIG. 3 below. For example, the processor 250 may be configured to perform (or process) select object 310 functionality, focus object 320 functionality, monitor region with object 330 functionality, determine if image in region(s) changed functionality 340, track object location 350 functionality, determine if object region(s) changed 360 functionality, change region(s) functionality, track object size 380 functionality, and/or determine if object size changed 390 functionality. In some aspects, the processor may be configured by instructions stored in the memory 260 to perform one or more functions discussed below with respect to Figure. 7.

[0048] The display 120 is configured to display images that are communicated to the display 120 from the processor 160. For some implementations, the display 120 displays user feedback, for example, annotations for touch-to-focus indicating selected frame segments. For some implementations, the display 120 displays menus prompting user input.

[0049] The user input device 240 is configured to accept user input. In some embodiments, the user input device 240 is a touch screen included in display 120. In some implementations, the touch screen 240 accepts multi-touch input from a user. In some implementations, instructions in the memory 260 may configure the processor to accept input from the user input device 240 that selects a region within an image. The selected region may contain an object the user wishes the camera to focus on. In some implementations, a plurality of regions may be superimposed onto an image displayed on the display 120. The user input device 240 may receive input selecting one of the superimposed regions. The selected region may then be used by the processor 250 to determine a focus of the camera 110.

[0050] FIG. 3 is a flowchart of an example method 300 for continuous autofocus. In block 310, method 300 selects an object. The method 300 may select the

object either automatically or manually. In some aspects, processor 250 may select the object automatically using an object detection or object recognition system based on a plurality of intensity values in the image. For some implementations, the processor 250 may use face recognition techniques to identify the object. The approach taken to automatically identify an object may depend on the application or context, which may be user selectable through user input device 240.

[0051] In some aspects, the selection of the object may be based on input received from the user input device 240. For example, as discussed above, a user may select an object or a region of an image using the user input device 240. In some aspects, the input may specify the corners of a bounding box substantially enclosing the object. In implementations that allow a user to select a region of an image, the processor 250 may then perform object detection within the selected region to select the object in block 310.

[0052] In block 315, method 300 determines a monitor region for the object. In some aspects, a field of view of an imaging sensor may be divided into a plurality of regions. One or more of those regions may be selected as a monitor region(s) for the selected object. As discussed above, the monitored region(s) may include all or a substantial portion of the selected object. In some aspects, the method 300 may enclose the object; for example, by generating a bounding region around the selected object within the one or more selected regions. In some aspects, the bounding region may be a minimum bounding box, and/or be shaped like a polygon or a rectangle. In some aspects, the bounding region may be aligned with the rows and columns of the image.

[0053] In block 320, method 300 focuses the camera on the object. In some aspects, the camera is focused on the object based on the monitor region(s) determined in block 315. For example, image content of the monitor regions may be evaluated to focus the camera. In an embodiment, in block 320, method 300 may step through different focal plane distances, measuring a focus parameter of the monitored regions at each of the different focus plane distances, and selecting the focus plane distance that optimizes the focus parameter to be the focus plane that focuses the camera. For example, the focus parameter may measure high frequency content at a particular focal distance. In some aspects, the camera may be focused in block 320 without consideration of one or more regions that are not part of the currently monitored regions determined in block 315 and/or 370 (discussed below). Some regions of an image may be excluded from consideration to reduce computational overhead associated with determining the best

focus position for the camera. By reducing the amount of image data that needs to be processed, autofocus may be performed more quickly and efficiently.

[0054] In block 330, method 300 monitors image content of the region(s) including the object. As discussed above, method 300 may divide the captured images into non-overlapping regions. One or more of those regions may be selected for monitoring based on the location of the object selected in block 310. Other regions may be excluded from monitoring.

[0055] In decision block 340, if the image in the region has changed, the camera is refocused on the object in block 320. In some aspects, whether the region has changed in block 340 may be based on a comparison of an amount of changed image content to a threshold. For example, if the number of changed pixels exceeds the threshold, the image content of the region is considered to be changed in some implementations of decision block 340. Some aspects may determine a difference image between a current image and an image to which a previous focus was applied. Absolute values of the difference image may be summed and the sum compared to a region difference threshold to determine whether an image has sufficiently changed to warrant a refocus by block 320. In some aspects, the absolute value function may be replaced by other functions, such as squaring the difference values. In some embodiments, method 300 determines a difference in one or more sensor 230 information statistics, such as exposure or focus value. In some embodiments, it is not necessary to register images, as global motion (i.e. from jitter) would trigger refocus. If decision block 340 determines that the image in the monitored region(s) has not changed substantially (or more than a threshold), then method 300 returns to block 330, which continues to monitor the region(s) via additional sequentially captured images.

[0056] Block 350 tracks the selected object's location within the plurality of regions discussed above. Tracking the selected object's location may include generating a series of outputs that indicate a current location of the object in a corresponding series of image frames. This series of outputs may then be evaluated by decision block 360, discussed in detail below.

[0057] Block 350 may track the object's location within the plurality of regions using a variety of methods, depending on the implementation. For example, in some implementations, the object's location may be based on a computed centroid of the

object. Whichever region holds the centroid of the object is the region that is monitored in these implementations. All other regions may not be monitored in this implementation.

[0058] Other aspects may track multiple regions. For example, in some aspects, all regions including at least a portion of the object or a region bounding the object may be monitored. Regions that contain no portions of the object (or region bounding the object) may not be monitored in these implementations.

[0059] In other implementations, a series of bounding regions for the object may be determined for a corresponding series of image frames. Center points of the bounding regions may be used to determine in which region(s) the object is positioned. For example, if the center enters a new region, the region(s) containing the object may be adjusted to include the region in which the center is located. Other regions may be removed from a set of region(s) including the object, for example, if those regions no longer include any portion of the selected object.

[0060] The size of the selected object may also affect the regions determined to include the object. For example, if the selected object moves toward the camera 110, it may grow larger within the field of view of the camera 110 and encompass more region(s) of the field of view 130. In this case, the region(s) containing the object may be adjusted to include new regions. Alternatively, if the selected object moves further from the camera 110, it may be included in fewer regions than previous image frames. Thus, in some implementations the region(s) determined to include the object may be adjusted to remove regions that no longer include any portion of the selected object. In other implementations, as the object moves further away, its centroid and/or center of a region bounding the object may change regions.

[0061] Decision block 360 determines whether the location of the object has changed to a different region, based on the output generated by monitoring block 330. If the current object regions generated by block 330 represent a change, the regions being monitored may be updated in block 370. In some aspects, block 370 updates a set of monitored regions to be consistent with the regions block 350 determined to include the object. Thus, in block 370, one or more other regions may be removed from the set of monitored region(s) in block 370 if, for example, they no longer include any portion of the selected object, the centroid of the object has moved to a new region, or the center of a bounding region has moved to a new region. As discussed above, regions may be added to a set of monitored regions in block 370 based on a location of a centroid of the

object, a center of a bounding region substantially enclosing the object, or based on regions that include at least a portion of the object, based on implementation.

[0062] If decision block 360 determines that the object continues to exist within the currently monitored regions, then method 300 returns to block 350, which continues to track the location of the object via additional sequentially captured images.

[0063] In block 380, method 300 tracks the object size. The size of the object may be determined using a variety of methods depending on implementations. In some embodiments, the method 300 tracks the area of a bounding region enclosing the object, such as a minimum bounding box that encloses the object. The size is then based on the area of the bounding region. In some aspects, the size is determined based on an actual size of the object itself, for example, based on a number of pixels or image elements that are included in the object.

[0064] In decision block 390, method 300 determines if the object size changed between a previous focus event and the current image frame. For some embodiments, if the size of the object increased relative to a previous size measurement in a previously focused image frame by more than a size increase threshold, method 300 triggers a focus event by moving to block 320. The size increase threshold may be 1.01, 1.05, 1.1, 1.15, 1.2, 1.25, 1.3, 1.35, 1.4 or any threshold in between. For example, a threshold of 1.2 may correspond to a 20% increase in size since a previous focus event. For some embodiments, if the size of the object decreased by more than a size decrease threshold, method 300 triggers a focus event by moving to block 320. For example, in some implementations, the size decrease threshold may be 0.85, corresponding to a 15% decrease in area of the size since the last focus event. Other implementations may utilize a size decrease threshold of .99, .95, .9, .8, .75, .7 or any size decrease threshold in between these values. If decision block 390 determines that the size of the object has not changed substantially (or more than a size threshold as discussed above), then method 300 returns to block 380, which continues to track the object's size by comparing the object's size in additional sequentially captured images to the object's size at a previous focus event.

[0065] FIG. 3 illustrates parallel operation of blocks 330, 350, and 380. In other implementations, these operations may be done in sequence, or in a different order. As one focus trigger is sufficient to initiate the focus operation, there may be no need to

monitor the region(s), track object location, and track object size during a focus operation.

[0066] In some implementations, if method 300 can no longer identify the selected object (for example, the object moves out of view, or the object gets too small to be tracked), process 300 may return to block 310, where a new object may be selected.

[0067] FIG. 4 is a representation of a scene including an object 150/450a/450b within a monitor region 145 at three different times: after focusing on the object 150, before the object 450a moves, and after the object 450b moves. The object 150 starts at position 151 after focusing the object. The same object, represented as object 450a, remains in position 151 before moving. The same object, represented as object 450b, moves to position 152. As the image in region 145 did not change while object 450a stayed in position 151, the object stayed in focus and the processor 250 did not trigger a focus event. However, when the object moved from position 151 to position 152 the image in region 145 did change. This change triggered a focus event because the difference in image content exceeds an image difference threshold.

[0068] FIG. 5 is a representation of an object moving from a first position in a first monitor region in a first image, to a second position in a second monitor region in a second image, for which the change in monitor region triggers a switch in monitor regions as well as a focus event. FIG. 5 shows two images of an object 150 in a scene. After focus, object 150 in bounding box 160 with center point 170 is in monitor region 145. In a subsequent image, object 150 moves along arrow 180 to monitor region 545. This second image is object 154, which corresponds to the same person in the scene. This second image shows object 154 in bounding box 164 with center point 174. Because the center point of bounding box 174 is in region 545, the continuous autofocus system may switch monitor regions from region 145 to region 545 and trigger a focus event.

[0069] FIG. 6 illustrates how a change in a size of a selected object 150 can trigger a focus event. In FIG. 6, objects 150a-g are shown as the same object depicted in a variety of sizes. Each of the objects 150a-g is enclosed by a corresponding bounding box 660a-g. The object 150d represents the starting size (or position). In this example, a camera may be initially focused on object 150d. The initial focus on object 150d is represented by hash mark 620. Objects to the right of object 150d, i.e., 150e-g, are relatively larger than object 150d, while objects to the left of object 150d, i.e., 150a-c, are relatively smaller than object 150d.

[0070] In some implementations, as the size of object 150d changes in size (for example represented by positions to the right, i.e. objects 150e-g) some implementations may determine a size ratio between the object 150d and the size of, for example, object 150e. In some aspects, the size ratio is based on an area of the subjects 150d and 150e themselves. In other aspects, the size ratio is based on a ratio of an area of bounding boxes 660d and 660e. In some aspects, when the change in size or ratio is greater than a threshold, a focus event is triggered, as discussed above with respect to FIG. 4 and blocks 380 and 390.

[0071] In FIG. 6, while object 150e is larger than object 150d, the size is not substantial enough to exceed the size threshold in the illustrated implementation. Thus, the change in size between objects 150d and 150e does not trigger a focus event. However, when the object increases further in size as represented by object 150f, the difference in size between object 150d and 150f exceeds the size increase threshold, and a focus event is triggered. This is represented by hash mark 630. When the object increases further in size from 150f to 150g, no focus event is triggered. This may be due to a size ratio between the size of object 150f and 150g being less than the size ratio used in the illustrated implementation. Note that the size ratio is based on the size of the object when the last focus event was performed. Thus, the size of object 150e and the size of object 150f are both compared to the size of object 150d, since object 150d was the last object to which a focus event was applied, as represented by hash mark 620. In contrast, the size of object 150g is compared to the size of object 150f, since that was the last object to which a focus was applied, as represented by hash mark 640.

[0072] A similar process is used when an object decreases in size. For example, if object 150d changes in size to a size represented by object 150c, a size ratio may be determined between the sizes of object 150d and object 150c. This difference may be compared to a size decrease threshold, and if the difference exceeds the threshold, a focus event may be triggered. In this example, no event is triggered based on the change from 150d to 150c, based on the lack of a hash mark between objects 150d and 150c. Similarly, a further reduction in size from 150c to 150b still does not result in a focus event, as the difference in size between object 150d and 150b still does not exceed the example implementation's size threshold. However, when the object is further reduced in size as represented by object 150a, the difference between object 150a and 150d exceeds the size threshold, and a focus event is triggered, as represented by hash

mark 610. Note that while both an increase in size and a decrease in size may result in a comparison between a size difference and a threshold, different thresholds may be used for each case. For example, some implementations may use a 10% increase in size, or a ratio of 1.1, to trigger a focus event and an 8% decrease in size, or a ratio of 0.92, to trigger a focus event. Other implementations may use a similar threshold for both increases and decreases in size. For example, some implementations may use a 9% increase or decrease in size to trigger a focus event.

[0073] FIG. 7 is a flowchart that illustrates a method 700 of continuous autofocus according to the embodiments described herein. In some aspects, instructions stored in the memory 260 of FIG. 2 configure the processor 250 to perform the functions of method 700. In some aspects, method 700 and method 300 of FIG. 3 are integrated or part of the same method. For example, the “track object size” blocks 380 and 390 of method 300 may correspond to at least blocks 720-740 of method 700, discussed below.

[0074] At block 710, method 700 sequentially captures a first and second image of a scene. In some implementations, at least some of the functionality of block 710 may be performed by camera 110 illustrated in FIG. 2. For example, in some aspects, processor instructions stored in the memory 260 may configure the processor 250 to control the camera 110 to capture the first and second images.

[0075] At block 720, method 700 determines a first size of a first object in the first image. In some implementations, the functionality of block 720 may be performed by the processor 250 illustrated in FIG. 2. As discussed above, in some aspects, the first size is based on an area of a region that bounds the first object within the first image, or based on an area of the first object itself in the first image. For example, the area may be based on a number of pixels or image elements that represent the first object within the first image.

[0076] At block 730, method 700 determines a second size of the first object in the second image. In some implementations, the functionality of block 730 may be performed by processor 250 illustrated in FIG. 2. As discussed above, in some aspects, the second size is based on an area of a region that bounds the first object within the second image, or based on an area of the first object itself in the second image.

[0077] At block 740, method 700 focuses the camera based on a difference between the first size and the second size. For example, block 740 may focus the camera as described above with respect to blocks 380 and 390 of FIG. 3 and/or FIG. 6.

[0078] In some aspects, a ratio between the size of the first image and the second image is compared to one or more thresholds. For example, in some implementations, a first threshold may be less than one, and control whether an object has been reduced in size in the second image as compared to the first image to a degree that the camera should be refocused. In this implementation the ratio may represent the size of the first object in the second image as compared to the size of the object in the first image. If this ratio is smaller than the first threshold, a focus event may be triggered. Similarly, this ratio may be compared to a second threshold larger than one. If the ratio exceeds the second threshold, some implementations may determine that the size of the object in the second image is larger than the object in the first image to such a degree that the camera should be refocused. In some implementations, one threshold could be compared to two computed ratios. For example, a first ratio could represent the first size / second size, while a second ratio represents the second size / first size.

[0079] One of skill in the art would understand that other implementations may reverse the relationships between the first and second sizes that is represented by a single computed ratio, i.e. a reverse ratio in other implementations may represent the size of the first object in the first image as compared to the size of the object in the second image. In these implementations, one of ordinary skill would understand that two thresholds could be identified to accomplish focus events as described above.

[0080] In some implementations, the functionality of block 720 may be performed by camera controller 220 illustrated in FIG. 2. In some implementations, block 720 may be performed by the processor 250 in combination with the camera controller 220. For example, the processor 250 may determine the differences between the first size and second size, and depending on how the difference compares to one or more thresholds, may command the camera controller 110 to focus the camera.

[0081] Some aspects of method 700 may include one or more of the functions discussed above with respect to blocks 330/340, and/or 350/360/370.

[0082] It should be understood that any reference to an element herein using a designation such as “first,” “second,” and so forth does not generally limit the quantity or order of those elements. Rather, these designations may be used herein as a convenient method of distinguishing between two or more elements or instances of an element. Thus, a reference to first and second elements does not mean that only two elements may be employed there or that the first element must precede the second element in some

manner. Also, unless stated otherwise a set of elements may comprise one or more elements. In addition, terminology of the form “at least one of: A, B, or C” used in the description or the claims means “A or B or C or any combination of these elements.”

[0083] As used herein, the term “determining” encompasses a wide variety of actions. For example, “determining” may include calculating, computing, processing, deriving, investigating, looking up (e.g., looking up in a table, a database or another data structure), ascertaining and the like. Also, “determining” may include receiving (e.g., receiving information), accessing (e.g., accessing data in a memory) and the like. Also, “determining” may include resolving, selecting, choosing, establishing and the like.

[0084] As used herein, a phrase referring to “at least one of” a list of items refers to any combination of those items, including single members. As an example, “at least one of: a, b, or c” is intended to cover: a, b, c, a-b, a-c, b-c, and a-b-c.

[0085] The various operations of methods described above may be performed by any suitable means capable of performing the operations, such as various hardware and/or software component(s), circuits, and/or module(s). Generally, any operations illustrated in the Figures may be performed by corresponding functional means capable of performing the operations.

[0086] The various illustrative logical blocks, modules and circuits described in connection with the present disclosure may be implemented or performed with a general purpose processor, a digital signal processor (DSP), an application specific integrated circuit (ASIC), a field programmable gate array signal (FPGA) or other programmable logic device (PLD), discrete gate or transistor logic, discrete hardware components or any combination thereof designed to perform the functions described herein. A general purpose processor may be a microprocessor, but in the alternative, the processor may be any commercially available processor, controller, microcontroller or state machine. A processor may also be implemented as a combination of computing devices, e.g., a combination of a DSP and a microprocessor, a plurality of microprocessors, one or more microprocessors in conjunction with a DSP core, or any other such configuration.

[0087] In one or more aspects, the functions described may be implemented in hardware, software, firmware, or any combination thereof. If implemented in software, the functions may be stored on or transmitted over as one or more instructions or code on a computer-readable medium. Computer-readable media includes both computer storage

media and communication media including any medium that facilitates transfer of a computer program from one place to another. A storage media may be any available media that can be accessed by a computer. By way of example, and not limitation, such computer-readable media can comprise RAM, ROM, EEPROM, CD-ROM or other optical disk storage, magnetic disk storage or other magnetic storage devices, or any other medium that can be used to carry or store desired program code in the form of instructions or data structures and that can be accessed by a computer. Also, any connection is properly termed a computer-readable medium. For example, if the software is transmitted from a website, server, or other remote source using a coaxial cable, fiber optic cable, twisted pair, digital subscriber line (DSL), or wireless technologies such as infrared, radio, and microwave, then the coaxial cable, fiber optic cable, twisted pair, DSL, or wireless technologies such as infrared, radio, and microwave are included in the definition of medium. Disk and disc, as used herein, includes compact disc (CD), laser disc, optical disc, digital versatile disc (DVD), floppy disk and Blu-ray disc where disks usually reproduce data magnetically, while discs reproduce data optically with lasers. Thus, in some aspects computer readable medium may comprise non-transitory computer readable medium (e.g., tangible media). In addition, in some aspects computer readable medium may comprise transitory computer readable medium (e.g., a signal). Combinations of the above should also be included within the scope of computer-readable media.

[0088] The methods disclosed herein comprise one or more steps or actions for achieving the described method. The method steps and/or actions may be interchanged with one another without departing from the scope of the claims. In other words, unless a specific order of steps or actions is specified, the order and/or use of specific steps and/or actions may be modified without departing from the scope of the claims.

[0089] Further, it should be appreciated that modules and/or other appropriate means for performing the methods and techniques described herein can be downloaded and/or otherwise obtained by a user terminal and/or base station as applicable. For example, such a device can be coupled to a server to facilitate the transfer of means for performing the methods described herein. Alternatively, various methods described herein can be provided via storage means (e.g., RAM, ROM, a physical storage medium such as a compact disc (CD) or floppy disk, etc.), such that a user terminal and/or base

station can obtain the various methods upon coupling or providing the storage means to the device. Moreover, any other suitable technique for providing the methods and techniques described herein to a device can be utilized.

[0090] It is to be understood that the claims are not limited to the precise configuration and components illustrated above. Various modifications, changes and variations may be made in the arrangement, operation and details of the methods and apparatus described above without departing from the scope of the claims.

WHAT IS CLAIMED IS:

1. A method of focusing a camera, comprising:
 - sequentially capturing a first and second image of a scene;
 - determining a first size of a first object in the first image;
 - determining a second size of the first object in the second image;
 - focusing the camera based on a difference between the first size and the second size.
2. The method of claim 1, further comprising:
 - determining a first bounding region includes the first object in the first image;
 - determining a second bounding region includes the first object in the second image;
 - focusing the camera based on a difference between the first bounding region and the second bounding region.
3. The method of claim 2, further comprising focusing the camera based on a difference between an area of the first bounding region and an area of the second bounding region.
4. The method of claim 2, further comprising:
 - determining a ratio between an area of the second bounding region and an area of the first bounding region;
 - focusing the camera in response to the ratio being less than a size decrease threshold or exceeding a size increase threshold.
5. The method of claim 2, further comprising:
 - focusing the camera if the difference exceeds a region difference threshold; and
 - maintaining the focus of the camera if the difference does not exceed the region difference threshold.
6. The method of claim 1, further comprising:

segmenting the first and second image into a plurality of corresponding regions;

determining a first region in the first image includes the first object, and a corresponding second region in the second image;

determining a first difference between image content in the first region and the second region; and

focusing the camera based on the first difference in image content.

7. The method of claim 6, further comprising excluding from consideration when focusing the camera, regions of the plurality of regions that are determined to not include the first object.

8. The method of claim 6, further comprising:

determining a third region in the second image includes the first object;

capturing a third image of the object at a third time, the third time after the second time;

determining a fourth region in the third image corresponding to the third region;

determining a second difference between image content in the third region and the fourth region; and

focusing the camera in response to the second difference in image content.

9. An imaging apparatus comprising:

a camera configured to sequentially capture a first and second image of a scene;

a processor in communication with the camera configured to determine a first size of a first object in the first image, the processor further configured to determine a second size of the first object in the second image; and

a camera controller configured to focus the camera based on a difference between the first size and the second size.

10. The imaging apparatus of claim 9, wherein the processor is further configured to determine a first bounding region containing the first object in the first image, and

determine a second bounding region containing the first object in the second image, and wherein the camera controller is further configured to focus the camera based on a difference between the first bounding region and the second bounding region.

11. The imaging apparatus of claim 10, wherein the camera controller is further configured to focus the camera based on a difference between an area of the first bounding region and an area of the second bounding region.

12. The imaging apparatus of claim 11, wherein the processor is further configured to determine a ratio between an area of the second bounding region and an area of the first bounding region, and wherein the camera controller is further configured to focus the camera based on the ratio being less than a size decrease threshold or exceeding a size increase threshold.

13. The imaging apparatus of claim 12, wherein the camera controller is further configured to focus the camera if the difference exceeds a region difference threshold, and to maintain the focus of the camera if the difference does not exceed the region difference threshold.

14. The imaging apparatus of claim 9, wherein the processor is further configured to:

segment the first and second image into a plurality of corresponding regions,

determine a first region in the first image includes the first object, and a corresponding second region in the second image,

determine a first difference between image content in the first region and the second region, and

wherein the camera controller is further configured to focus the camera based on the first difference in image content.

15. The apparatus of claim 14, wherein the processor is further configured to exclude from consideration, when focusing the camera, regions of the plurality of regions that are determined to not include the first object.

16. The apparatus of claim 14, wherein the processor is further configured to:

- determine a third region in the second image includes the first object,
- capture a third image of the object at a third time, the third time after the second time,
- determine a fourth region in the third image corresponding to the third region, and
- determine a second difference between image content in the third region and the fourth region,

and wherein the camera controller is further configured to focus the camera in response to the second difference in image content.

17. An imaging apparatus comprising:

- means for sequentially capturing a first and second image of a scene;
- means for determining a first size of a first object in the first image;
- means for determining a second size of the first object in the second image; and
- means for focusing the capturing means based on a difference between the first size and the second size.

18. The apparatus of claim 17, wherein the capturing means comprises a camera, wherein the first size determining means comprises a processor, wherein the second size determining means comprises the processor, and wherein the focusing means comprises a camera controller.

19. A non-transitory computer-readable medium storing instructions that when executed cause a processor to perform a method of focusing a camera, the method comprising:

- sequentially capturing a first and second image of a scene;
- determining a first size of a first object in the first image;
- determining a second size of the first object in the second image; and
- focusing the camera based on a difference between the first size and the second size.

1/7

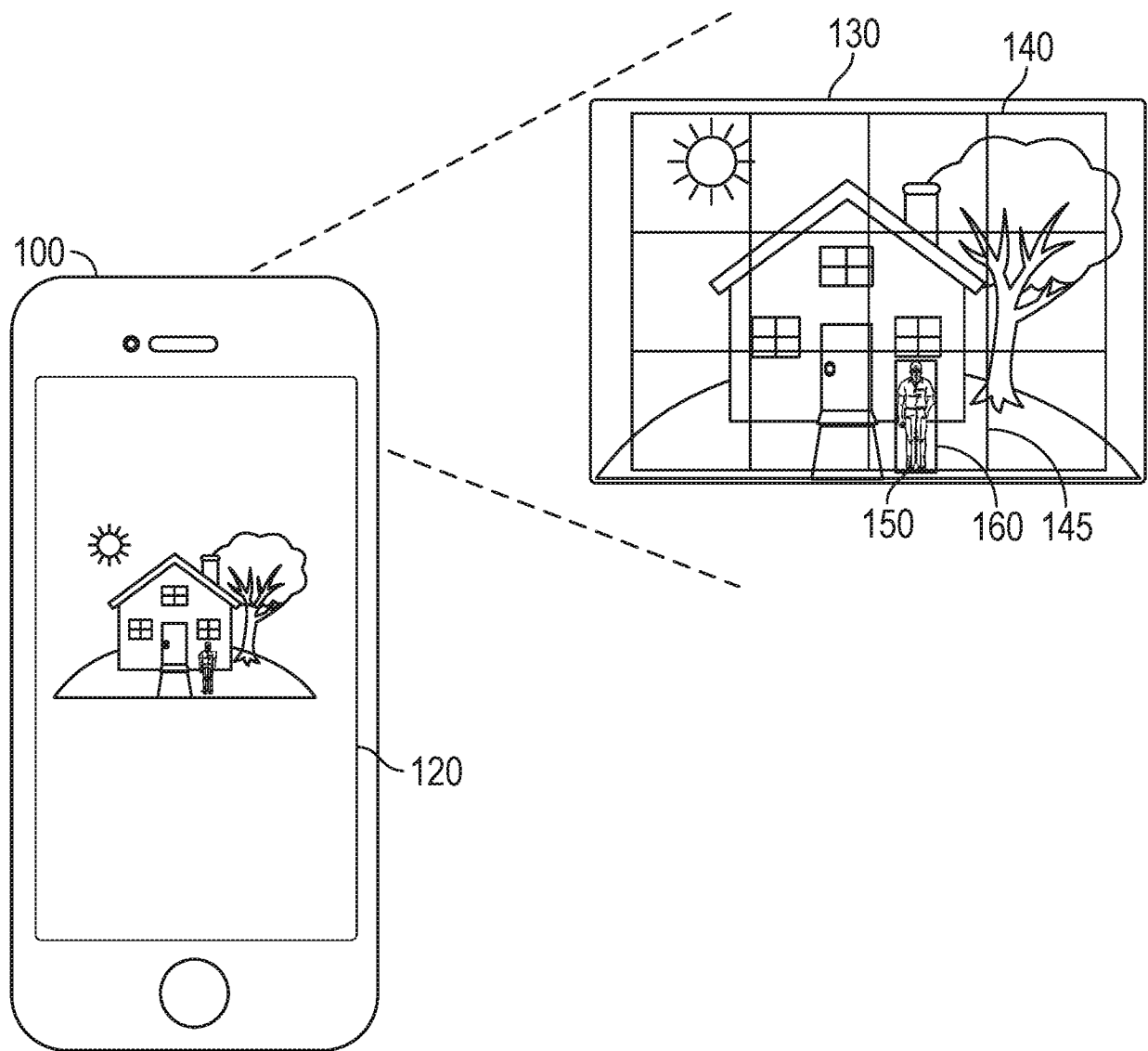


FIG. 1

2/7

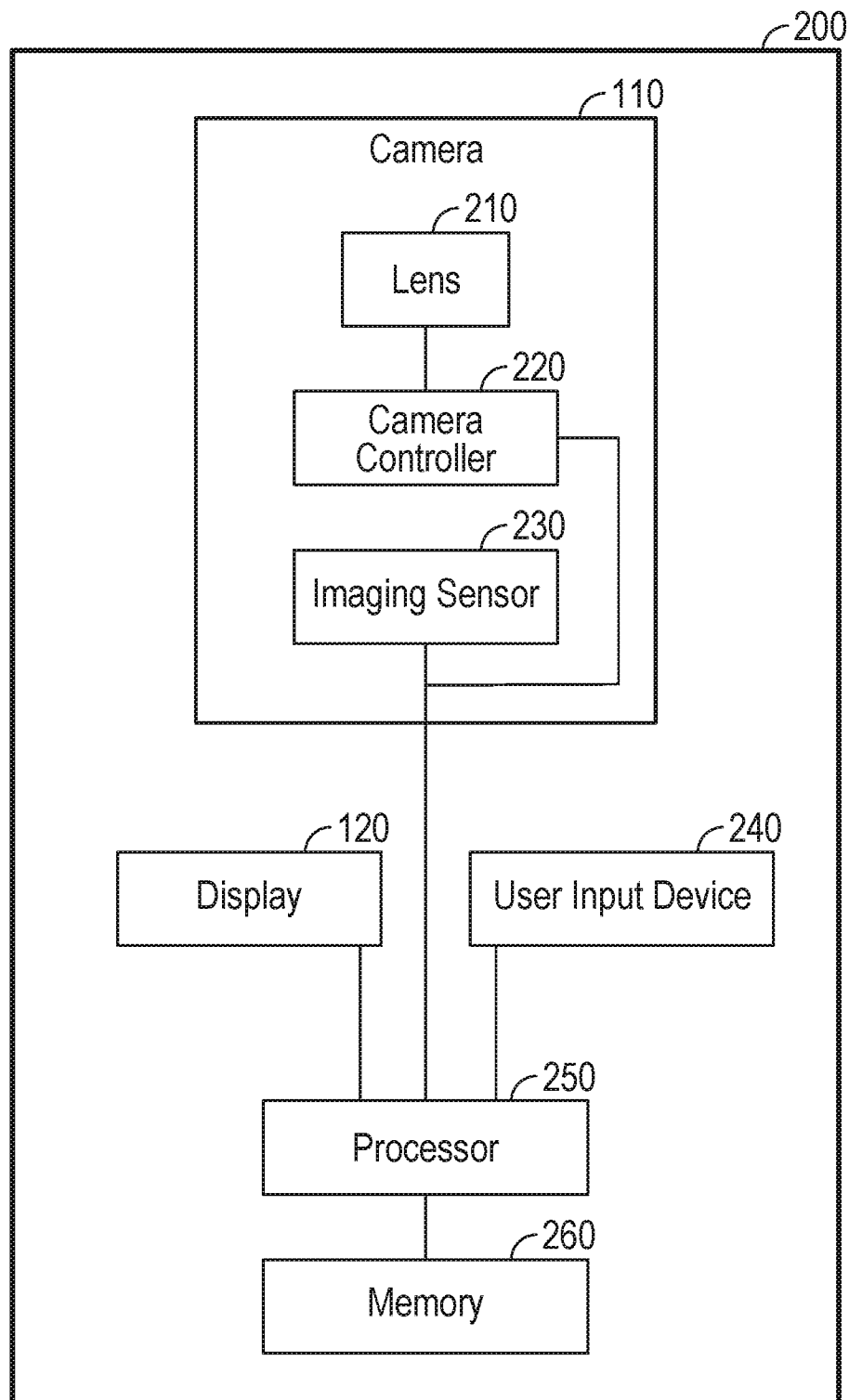


FIG. 2

3/7

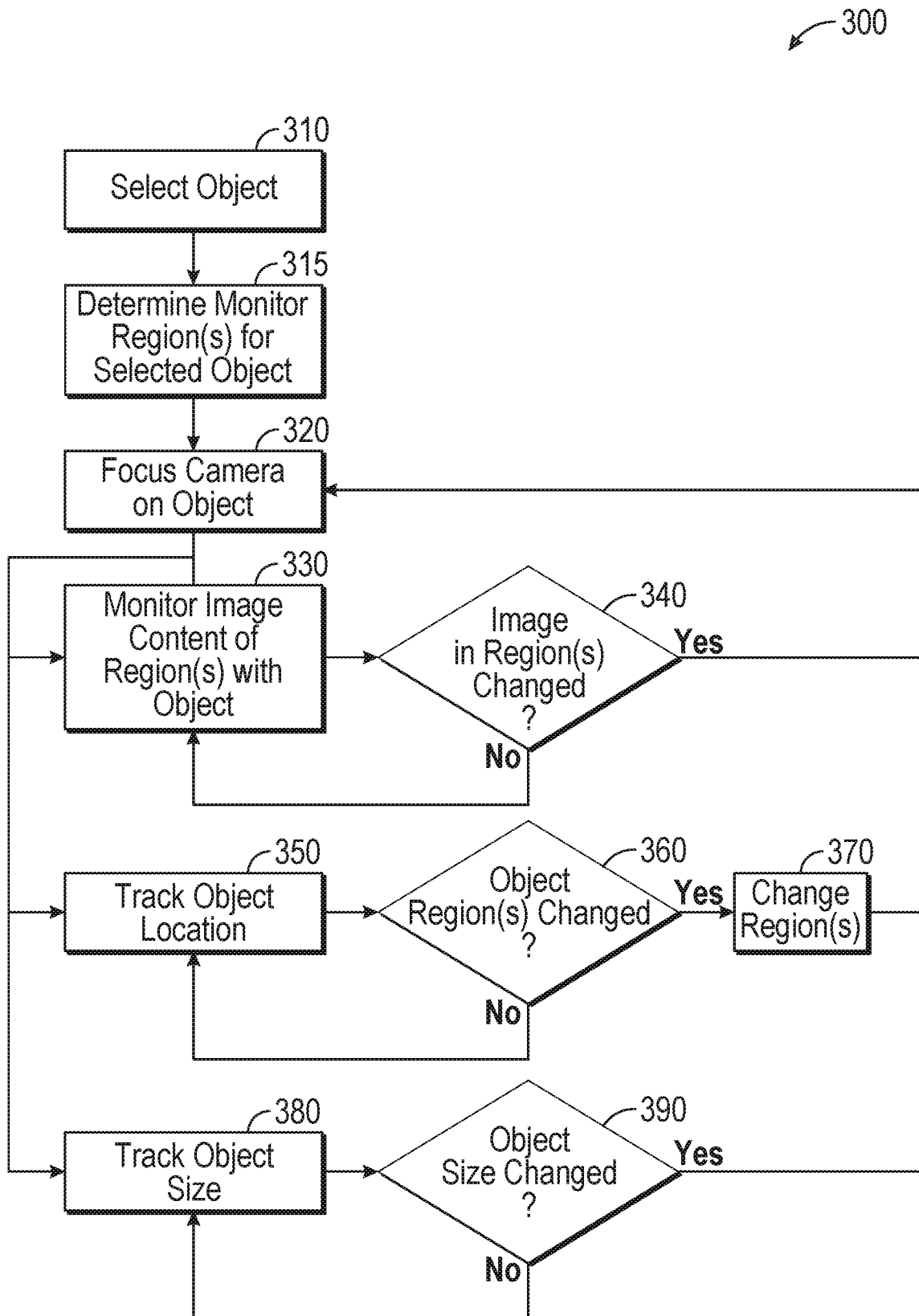


FIG. 3

4/7

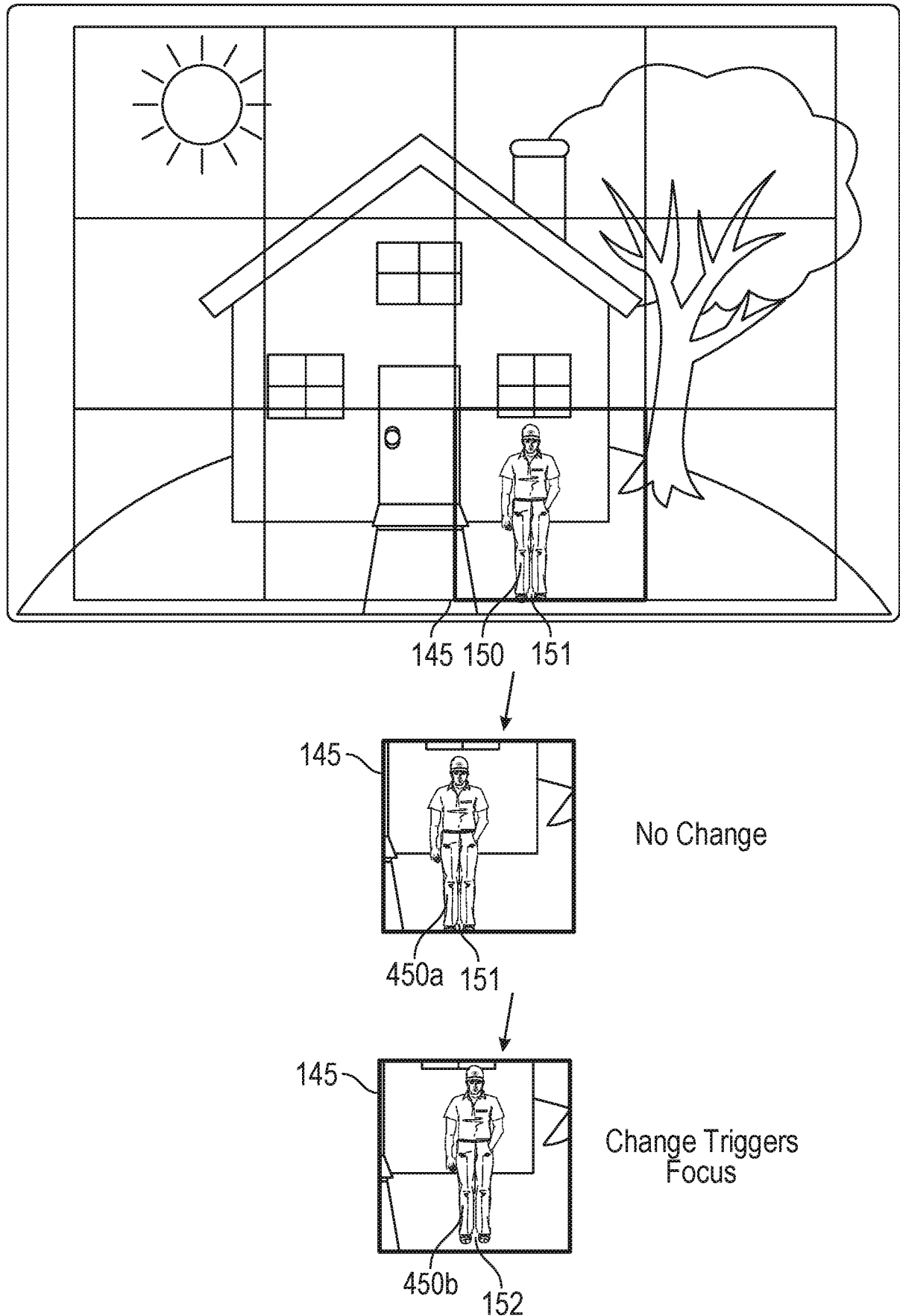


FIG. 4

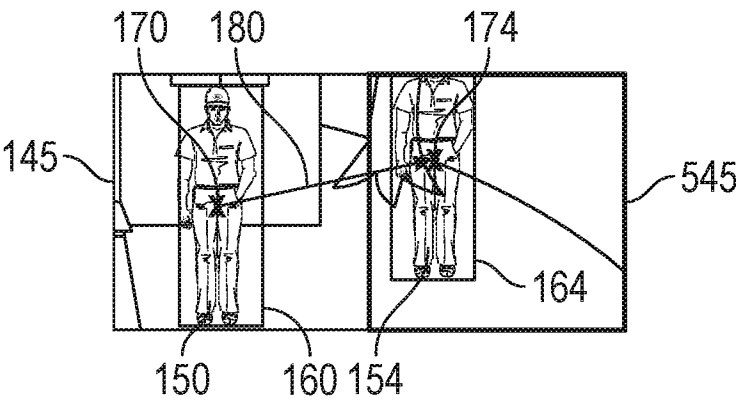


FIG. 5

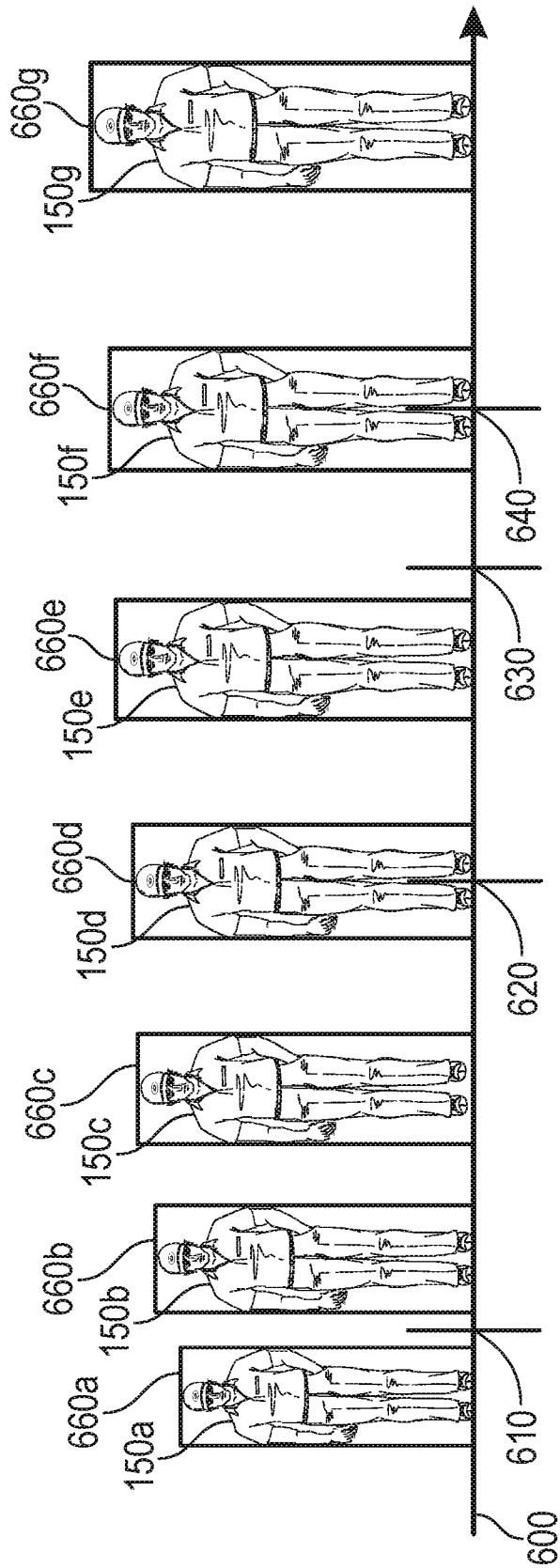


FIG. 6

7/7

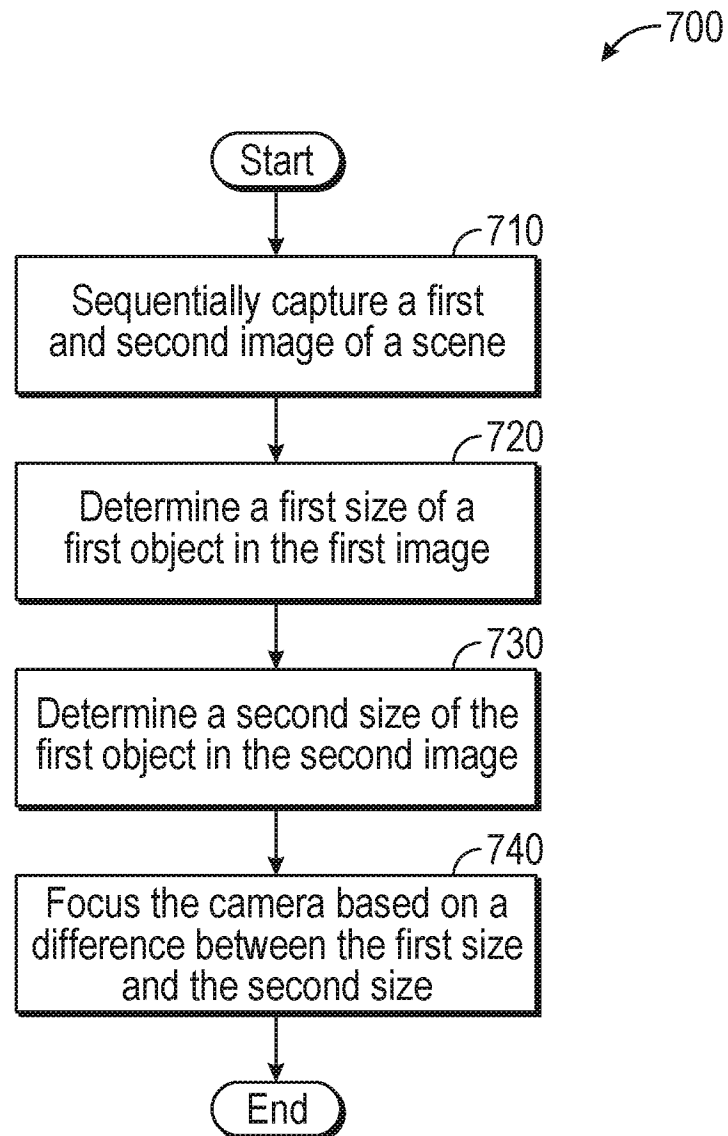


FIG. 7

INTERNATIONAL SEARCH REPORT

International application No
PCT/US2016/019142

A. CLASSIFICATION OF SUBJECT MATTER

INV. H04N5/232
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)

H04N

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)

EP0-Internal

C. DOCUMENTS CONSIDERED TO BE RELEVANT

| Category* | Citation of document, with indication, where appropriate, of the relevant passages | Relevant to claim No. |
|-----------|--|------------------------|
| X | US 2009/009651 A1 (TAKAYANAGI WATARU [JP]) 8 January 2009 (2009-01-08) | 1-5, 9-13,17, 18 |
| Y | paragraph [0060] - paragraph [0089] paragraph [0109] - paragraph [0129] paragraph [0145] - paragraph [0168] figures 1-18 | 6-8, 14-16 |
| Y | ----- JP 2000 188713 A (RICOH KK) 4 July 2000 (2000-07-04) paragraph [0009] - paragraph [0065] figures 1-10 | 6-8, 14-16 |
| X | ----- US 2011/043681 A1 (SHIBAGAMI GENJIRO [JP]) 24 February 2011 (2011-02-24) paragraph [0023] - paragraph [0052] figures 1-4 ----- -/- | 1-5, 9-13, 17-19 |

☒ 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

26 May 2016

Date of mailing of the international search report

01/06/2016

Name and mailing address of the ISA/

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

Authorized officer

Rocca, Delphine

INTERNATIONAL SEARCH REPORT

International application No

PCT/US2016/019142

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

| Category* | Citation of document, with indication, where appropriate, of the relevant passages | Relevant to claim No. |
|-----------|--|------------------------|
| X | US 2014/253785 A1 (CHAN WEI-KAI [TW] ET AL) 11 September 2014 (2014-09-11) paragraph [0024] - paragraph [0032] figures 2-12 ----- | 1-3, 9-11, 17-19 |
| A | US 2013/293741 A1 (NISHI TAKESHI [JP]) 7 November 2013 (2013-11-07) paragraph [0025] - paragraph [0061] figures 1-4 ----- | 1,6-9, 14-17,19 |

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No

PCT/US2016/019142

| Patent document cited in search report | Publication date | Patent family member(s) | Publication date |
|---|---------------------|--|--|
| US 2009009651 A1 | 08-01-2009 | NONE | |
| JP 2000188713 A | 04-07-2000 | JP 3732665 B2 JP 2000188713 A | 05-01-2006 04-07-2000 |
| US 2011043681 A1 | 24-02-2011 | CN 101995733 A JP 5213813 B2 JP 2011039433 A US 2011043681 A1 | 30-03-2011 19-06-2013 24-02-2011 24-02-2011 |
| US 2014253785 A1 | 11-09-2014 | CN 104038691 A US 2014253785 A1 | 10-09-2014 11-09-2014 |
| US 2013293741 A1 | 07-11-2013 | US 2013293741 A1 WO 2012093663 A1 | 07-11-2013 12-07-2012 |