



- (51) International Patent Classification:
G06T 7/20 (2006.01) *H05B 37/02* (2006.01)
- (21) International Application Number:
PCT/US2013/069231
- (22) International Filing Date:
8 November 2013 (08.11.2013)
- (25) Filing Language: English
- (26) Publication Language: English
- (30) Priority Data:
61/724,073 8 November 2012 (08.11.2012) US
13/744,810 18 January 2013 (18.01.2013) US
- (71) Applicant: LEAP MOTION, INC. [US/US]; 333 Bryant Street, LL150, San Francisco, CA 94107 (US).
- (72) Inventor; and
- (71) Applicant : HOLZ, David [US/US]; 219 Brannan Street, Apt. 12J, San Francisco, CA 94107 (US).
- (74) Agent: BEFFEL, Ernest, J., Jr.; Haynes Beffel & Wolfeld LLP, P.O. Box 366, Half Moon Bay, CA 94019 (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, LT, 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, 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).

[Continued on next page]

(54) Title: OBJECT DETECTION AND TRACKING WITH VARIABLE-FIELD ILLUMINATION DEVICES

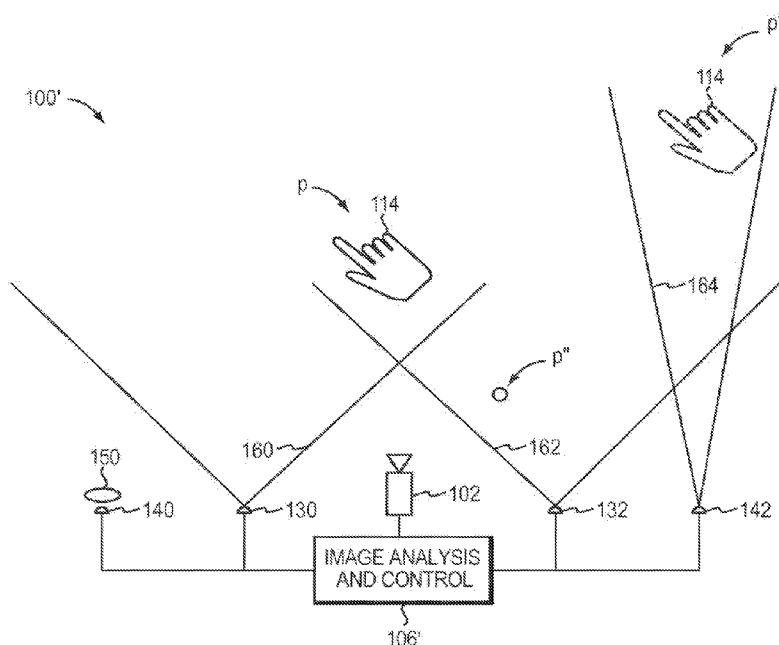


FIG. 1B

(57) Abstract: Imaging systems and methods optimize illumination of objects for purposes of detection, recognition and/or tracking by tailoring the illumination to the position of the object within the detection space. For example, feedback from a tracking system may be used to control and aim the lighting elements so that the illumination can be reduced or increased depending on the need.



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))

OBJECT DETECTION AND TRACKING WITH VARIABLE-FIELD ILLUMINATION DEVICES

CROSS-REFERENCE TO RELATED APPLICATIONS

[0001] This application claims priority to U.S. Application Serial No. 13/744,810 filed on January 18, 2013, which claims the benefits of U.S. Provisional Application Serial No. 61/724,073, filed on November 8, 2012, the entire disclosures of which are hereby incorporated by reference.

BACKGROUND

[0002] The present disclosure relates generally to imaging systems and in particular to three-dimensional (3D) object detection, tracking and characterization using optical imaging.

[0003] Motion-capture systems are used in a variety of contexts to obtain information about the conformation and motion of various objects, including objects with articulating members, such as human hands or human bodies. Such systems generally include cameras to capture sequential images of an object in motion and computers to analyze the images to create a reconstruction of an object's volume, position and motion. For 3D motion capture, at least two cameras are typically used.

[0004] Image-based motion-capture systems rely on the ability to distinguish an object of interest from other objects or background. This is often achieved using image-analysis algorithms that detect edges, typically by comparing pixels to detect abrupt changes in color and/or brightness. Adequate illumination of the object of interest is obviously key to successful detection. For systems operative over a broad detection space, high illumination intensities are typically necessary; in effect, the entire monitored region is bathed in light (which may be visible or outside the visible spectrum, e.g., infrared radiation), so that "worst case" object positions receive sufficient illumination. This means that the unitary light intensity is almost always greater than necessary, imposing excessive requirements in terms of hardware design, the number and/or luminance of lighting elements, and power consumption.

SUMMARY

[0005] The present invention relates to imaging systems and methods that optimize illumination of objects for purposes of detection, recognition and/or tracking by tailoring the illumination to the position of the object within the detection space. For example, feedback from a tracking system may be used to control and aim the lighting elements so that the illumination can be reduced or increased depending on the need.

[0006] In general, systems in accordance with the present invention utilize an array of lighting elements (e.g., light-emitting diodes (LEDs), incandescent bulbs, lasers, etc.) having different or variable illumination angles (also known as beam angles, i.e., the angle between the two directions opposed to each other over the beam axis for which the luminous intensity is half the maximum luminous intensity). The greater the angle of illumination, the faster will be the intensity falloff with distance. Accordingly, as the object is tracked, its range and position determine which illumination devices are turned on, or their illumination angles.

[0007] For example, the detection space may initially be lit with one or more wide-beam lighting elements with a collective field of view similar to that of the tracking device, e.g., a camera. Once the object's position is obtained, the wide-beam lighting element(s) may be turned off and one or more narrow-beam lighting elements, pointing in the direction of the object, activated. As the object moves, different ones of the narrow-beam lighting elements are activated. In many implementations, these directional lighting elements only need to be located in the center of the field of view of the camera; for example, in the case of hand tracking, people will not often try to interact with the camera from a wide angle and a large distance simultaneously.

[0008] If the tracked object is at a large angle to the camera (i.e., far to the side of the motion-tracking device), it is likely relatively close to the device. Accordingly, a low-power, wide-beam lighting element may be suitable. As a result, the lighting array may include only one or a small number of wide-beam lighting elements close to the camera along with an equal or larger number of narrow-beam devices (e.g., collectively covering the center-field region of space in front of the camera — for example, within a 30° or 45° cone around the normal to the camera). Thus, it is possible to decrease or minimize the number of lighting elements required to illuminate a space in which motion is detected by using a small number

of wide-beam elements and a larger (or equal) number of narrow-beam elements directed toward the center field.

[0009] It is also possible to cover a wide field of view with many narrow-beam LEDs pointing in different directions. These may be operated so as to scan the monitored space in order to identify the elements actually spotlighting the object; only these are kept on and the others turned off. In some embodiments, the motion system computes a predicted trajectory of the tracked object, and this trajectory is used to anticipate which illumination elements should be activated as the object moves. The trajectory is revised, along with the illumination pattern, as new tracking information is obtained.

[00010] Accordingly, in a first aspect, the invention pertains to an image capture and analysis system comprising a camera oriented toward a field of view; a plurality of individually controllable lighting elements directed toward the field of view, the lighting elements including at least one wide-beam lighting element and a plurality of narrow-beam lighting elements; and a controller coupled to the camera. In various embodiments, the controller and the light sources are configured to operate the camera to capture a sequence of images including an object moving in the field of view; and operate the lighting elements to illuminate the moving object with sufficient intensity to facilitate image capture by the camera but with minimum supplied power.

[00011] In various embodiments, the lighting elements include at least one wide-beam lighting element and a plurality of narrow-beam lighting elements. The narrow-beam lighting elements may have a beam angle of approximately 60° and the wide-beam lighting element(s) may have a beam angle of approximately 120°. The controller may be configured to track the object and activate different ones of the narrow-beam lighting elements to maintain the object in illumination as it moves. In addition, the controller may be configured to deactivate lighting elements not necessary to illuminate the object. In some embodiments, the controller steers the narrow-beam lighting elements to maintain the object in illumination by at least one such element as the object moves. For example, the controller may be configured to analyze images from the camera to predict a trajectory of the object and to activate or steer lighting elements in accordance with the predicted trajectory.

[00012] In a second aspect, the invention pertains to an image capture and analysis system. In various embodiments, the system comprises a camera oriented toward a field of view; a

plurality of individually controllable lighting elements and controllable optics associated with at least some of the lighting elements, the optics controllably altering a beam angle of the associated lighting elements; and a controller coupled to the camera and the light sources. The controller is configured to operate the camera to capture a sequence of images including an object moving in the field of view; and operate the lighting elements and the optics to illuminate the moving object with sufficient intensity to facilitate image capture by the camera but with minimum supplied power. In some embodiments, the optics are tunable lenses. The controller may control the optics to vary a beam angle of a lighting element and a drive current to vary an output luminance of the lighting element to minimize the drive current.

[00013] In a third aspect, the invention pertains to a method of image capture and analysis for use with a camera and a plurality of lighting elements oriented toward a field of view; the lighting elements include at least one wide-beam lighting element and a plurality of narrow-beam lighting elements. In particular, the lighting elements may include at least one wide-beam lighting element and a plurality of narrow-beam lighting elements. In various embodiments, the method comprises operating the camera to capture a sequence of images including an object moving in the field of view; and operating the lighting elements to illuminate the moving object with sufficient intensity to facilitate image capture by the camera but with minimum supplied power. In some implementations the narrow-beam lighting elements have a beam angle of approximately 60° and the wide-beam lighting element(s) have a beam angle of approximately 120°.

[00014] In various embodiments, the object is tracked and different ones of the narrow-beam lighting elements are activated to maintain the object in illumination as it moves; lighting elements not necessary to illuminate the object may be deactivated. Moreover, the narrow-beam lighting elements may be steered to maintain the object in illumination by at least one such element as the object moves. Images from the camera may be analyzed to predict a trajectory of the object, and lighting elements may be activated and/or steered in accordance with the predicted trajectory.

[00015] In a fourth aspect, the invention relates to a method of image capture and analysis for use with a camera and a plurality of individually controllable lighting elements oriented toward a field of view and controllable optics associated with at least some of the lighting

elements; the optics controllably alter a beam angle of the associated lighting elements. In various embodiments, the method comprises operating the camera to capture a sequence of images including an object moving in the field of view; and operating the lighting elements and the optics to illuminate the moving object with sufficient intensity to facilitate image capture by the camera but with minimum supplied power.

[00016] As used herein, the term “substantially” or “approximately” means $\pm 10\%$ (e.g., by weight or by volume), and in some embodiments, $\pm 5\%$. The term “consists essentially of” means excluding other materials that contribute to function, unless otherwise defined herein. The term “light” refers to any form of electromagnetic radiation and not merely, for example, to visible light. Reference throughout this specification to “one example,” “an example,” “one embodiment,” or “an embodiment” means that a particular feature, structure, or characteristic described in connection with the example is included in at least one example of the present technology. Thus, the occurrences of the phrases “in one example,” “in an example,” “one embodiment,” or “an embodiment” in various places throughout this specification are not necessarily all referring to the same example. Furthermore, the particular features, structures, routines, steps, or characteristics may be combined in any suitable manner in one or more examples of the technology. The headings provided herein are for convenience only and are not intended to limit or interpret the scope or meaning of the claimed technology.

[00017] The following detailed description together with the accompanying drawings will provide a better understanding of the nature and advantages of the present invention.

BRIEF DESCRIPTION OF THE DRAWINGS

[00018] FIG. 1A illustrates a representative system for capturing image data.

[00019] FIG. 1B illustrates a system for capturing image data according to an embodiment of the present invention.

[00020] FIG. 2 is a simplified block diagram of a computer system implementing an image analysis apparatus according to an embodiment of the present invention.

[00021] FIGS. 3A-3C are graphs of brightness data for rows of pixels that may be obtained according to an embodiment of the present invention.

[00022] FIG. 4 is a flow diagram of a process for identifying the location of an object in an image according to an embodiment of the present invention.

[00023] FIG. 5 illustrates a timeline in which light sources pulsed on at regular intervals according to an embodiment of the present invention.

[00024] FIG. 6 illustrates a timeline for pulsing light sources and capturing images according to an embodiment of the present invention.

[00025] FIG. 7 is a flow diagram of a process for identifying object edges using successive images according to an embodiment of the present invention.

[00026] FIG. 8 is a top view of a computer system incorporating a motion detector as a user input device according to an embodiment of the present invention.

[00027] FIG. 9 is a front view of a tablet computer illustrating another example of a computer system incorporating a motion detector according to an embodiment of the present invention.

[00028] FIG. 10 illustrates a goggle system incorporating a motion detector according to an embodiment of the present invention.

[00029] FIG. 11 is a flow diagram of a process for using motion information as user input to control a computer system or other system according to an embodiment of the present invention.

[00030] FIG. 12 illustrates a system for capturing image data according to another embodiment of the present invention.

[00031] FIG. 13 illustrates a system for capturing image data according to still another embodiment of the present invention.

DETAILED DESCRIPTION

[00032] Refer first to FIG. 1A, which illustrates a system 100 for capturing image data that does not implement the present invention. System 100 includes a pair of cameras 102, 104 coupled to an image-analysis system 106. Cameras 102, 104 can be any type of camera, including cameras sensitive across the visible spectrum or, more typically, with enhanced sensitivity to a confined wavelength band (e.g., the infrared (IR) or ultraviolet bands); more

generally, the term “camera” herein refers to any device (or combination of devices) capable of capturing an image of an object and representing that image in the form of digital data. For example, line sensors or line cameras rather than conventional devices that capture a two-dimensional (2D) image can be employed. The term “light” is used generally to connote any electromagnetic radiation, which may or may not be within the visible spectrum, and may be broadband (e.g., white light) or narrowband (e.g., a single wavelength or narrow band of wavelengths).

[00033] Cameras 102, 104 are preferably capable of capturing video images (i.e., successive image frames at a constant rate of at least 15 frames per second), although no particular frame rate is required. The capabilities of cameras 102, 104 are not critical to the invention, and the cameras can vary as to frame rate, image resolution (e.g., pixels per image), color or intensity resolution (e.g., number of bits of intensity data per pixel), focal length of lenses, depth of field, etc. In general, for a particular application, any cameras capable of focusing on objects within a spatial volume of interest can be used. For instance, to capture motion of the hand of an otherwise stationary person, the volume of interest might be defined as a cube approximately one meter on a side.

[00034] The illustrated system 100 includes a pair of light sources 108, 110, which can be disposed to either side of cameras 102, 104, and controlled by image-analysis system 106. Light sources 108, 110 can be infrared light sources of generally conventional design, e.g., infrared light-emitting diodes (LEDs), and cameras 102, 104 can be sensitive to infrared light. Filters 120, 122 can be placed in front of cameras 102, 104 to filter out visible light so that only infrared light is registered in the images captured by cameras 102, 104. In some embodiments where the object of interest is a person’s hand or body, use of infrared light can allow the motion-capture system to operate under a broad range of lighting conditions and can avoid various inconveniences or distractions that may be associated with directing visible light into the region where the person is moving. However, a particular wavelength or region of the electromagnetic spectrum is required.

[00035] It should be stressed that the arrangement shown in FIG. 1A is representative and not limiting. For example, lasers or other light sources can be used instead of LEDs. For laser setups, additional optics (e.g., a lens or diffuser) may be employed to widen the laser beam (and make its field of view similar to that of the cameras). Useful arrangements can

also include short- and wide-angle illuminators for different ranges. Light sources are typically diffuse rather than specular point sources; for example, packaged LEDs with light-spreading encapsulation are suitable.

[00036] In operation, cameras 102, 104 are oriented toward a region of interest 112 in which an object of interest 114 (in this example, a hand) and one or more background objects 116 can be present. Light sources 108, 110 are arranged to illuminate region 112. In some embodiments, one or more of the light sources 108, 110 and one or more of the cameras 102, 104 are disposed below the motion to be detected, e.g., where hand motion is to be detected, beneath the spatial region where that motion takes place. This is an optimal location because the amount of information recorded about the hand is proportional to the number of pixels it occupies in the camera images, the hand will occupy more pixels when the camera's angle with respect to the hand's "pointing direction" is as close to perpendicular as possible. Because it is uncomfortable for a user to orient his palm toward a screen, the optimal positions are either from the bottom looking up, from the top looking down (which requires a bridge) or from the screen bezel looking diagonally up or diagonally down. In scenarios looking up there is less likelihood of confusion with background objects (clutter on the user's desk, for example) and if it is directly looking up then there is little likelihood of confusion with other people out of the field of view (and also privacy is enhanced by not imaging faces). Image-analysis system 106, which can be, e.g., a computer system, can control the operation of light sources 108, 110 and cameras 102, 104 to capture images of region 112. Based on the captured images, image-analysis system 106 determines the position and/or motion of object 114.

[00037] For example, as a step in determining the position of object 114, image-analysis system 106 can determine which pixels of various images captured by cameras 102, 104 contain portions of object 114. In some embodiments, any pixel in an image can be classified as an "object" pixel or a "background" pixel depending on whether that pixel contains a portion of object 114 or not. With the use of light sources 108, 110, classification of pixels as object or background pixels can be based on the brightness of the pixel. For example, the distance (r_o) between an object of interest 114 and cameras 102, 104 is expected to be smaller than the distance (r_b) between background object(s) 116 and cameras 102, 104. Because the intensity of light from sources 108, 110 decreases as $1/r^2$, object 114 will be more brightly lit than background 116, and pixels containing portions of object 114

(i.e., object pixels) will be correspondingly brighter than pixels containing portions of background 116 (i.e., background pixels). For example, if $r_B/r_O = 2$, then object pixels will be approximately four times brighter than background pixels, assuming object 114 and background 116 are similarly reflective of the light from sources 108, 110, and further assuming that the overall illumination of region 112 (at least within the frequency band captured by cameras 102, 104) is dominated by light sources 108, 110. These assumptions generally hold for suitable choices of cameras 102, 104, light sources 108, 110, filters 120, 122, and objects commonly encountered. For example, light sources 108, 110 can be infrared LEDs capable of strongly emitting radiation in a narrow frequency band, and filters 120, 122 can be matched to the frequency band of light sources 108, 110. Thus, although a human hand or body, or a heat source or other object in the background, may emit some infrared radiation, the response of cameras 102, 104 can still be dominated by light originating from sources 108, 110 and reflected by object 114 and/or background 116.

[00038] In this arrangement, image-analysis system 106 can quickly and accurately distinguish object pixels from background pixels by applying a brightness threshold to each pixel. For example, pixel brightness in a CMOS sensor or similar device can be measured on a scale from 0.0 (dark) to 1.0 (fully saturated), with some number of gradations in between depending on the sensor design. The brightness encoded by the camera pixels scales standardly (linearly) with the luminance of the object, typically due to the deposited charge or diode voltages. In some embodiments, light sources 108, 110 are bright enough that reflected light from an object at distance r_O produces a brightness level of 1.0 while an object at distance $r_B = 2r_O$ produces a brightness level of 0.25. Object pixels can thus be readily distinguished from background pixels based on brightness. Further, edges of the object can also be readily detected based on differences in brightness between adjacent pixels, allowing the position of the object within each image to be determined. Correlating object positions between images from cameras 102, 104 allows image-analysis system 106 to determine the location in 3D space of object 114, and analyzing sequences of images allows image-analysis system 106 to reconstruct 3D motion of object 114 using conventional motion algorithms.

[00039] FIG. 1B illustrates a system 100' implementing the invention, with a single camera 102 illustrated for ease of presentation. The illustrated system 100' includes a pair of wide-beam LEDs 130, 132 located proximate to camera 102, and a pair of narrow-beam LEDs 140, 142 located further away from camera 102. Typically, a "wide beam" is about

120° wide and a narrow beam is approximately 60° wide, although these are representative figures only and may vary with the application; more generally, a wide beam can have a beam angle anywhere from >90° to 180°, and a narrow beam can have a beam angle anywhere from >0° to 90°. Image-analysis system 106 also controls the operation of LEDs 130, 132, 140, 142. It should be understood that four LEDs are shown for exemplary purposes only; the system may contain more or fewer lighting elements as described below and dependent upon the application, and furthermore, the lighting elements need not be LEDs; once again that term is used generically for convenience.

[00040] As illustrated, the wide beams 160, 162 of LEDs 130, 132 intercept the object 114 in its initial position P, which is central and relatively close to the camera 102. The LEDs 130, 132 may be low-power LEDs used to illuminate only objects close to camera 102, or may be of sufficient power to illuminate the entire detection space in order to facilitate initial object detection, in which case it is desirable to keep LEDs 130, 132 active only for so long as is necessary. As the object 114 moves away to a second position P', it comes within the illumination field of narrow-beam LED 142. Image-analysis system 106' performs the functions described above but also includes functionality to activate and deactivate the LEDs based on the images recorded by camera 102. Here, system 106' registers this transition and both activates LED 102 and deactivates wide-beam LEDs 130, 132. Accordingly, as the object 114 is tracked, its range determines which LEDs are turned on.

[00041] In some embodiments, the LEDs 130, 132, 140, 142 are identical devices but have an associated electrically or electromechanically controllable optic 150 (shown, for convenience, only over LED 140) that is controlled by system 106', which also controls the power (e.g., drive current) supplied to the LEDs using conventional drive and control circuitry. For example, the optic 150 may be a tunable lens. In this case, system 106' determines which of the LEDs can illuminate object 114 with the narrowest beam angle so as to draw the least power, and both drives the LED at the determined power level and adjusts the associated optic 150 to produce the necessary beam angle. As the object 114 is tracked, system 106' may cause the beam angle of the selected LED to widen and increase its drive current until the object 114 enters the field of an adjacent LED that can illuminate it with less power consumption, at which point system 106' activates the new LED and turns off the previously selected LED.

[00042] For example, the detection space may initially be lit with one or more wide-beam lighting elements with a collective field of view similar to that of camera 102. Once the position of object 114 is acquired by system 106', it may turn off the wide-beam LED(s) and activate one or more narrow-beam LEDs pointing in the direction of the object 114. As the object 114 moves, different ones of the narrow-beam LEDs are activated. In many implementations, these directional LEDs need only be located in the center of the field of view of camera 102; for example, in the case of hand tracking, people will not often try to interact with the camera from a wide angle and a large distance simultaneously.

[00043] If the tracked object is at a large angle to the camera (i.e., far to the side of the motion-tracking device at position P'), it is likely relatively close to the camera 102. Accordingly, a low-power, wide-beam LED may be suitable. As a result, the lighting array may include only one or a small number of wide-beam LEDs close to camera 102 along with an equal or larger number of narrow-beam LEDs (e.g., collectively covering the center-field region of space in front of camera 102 — for example, within a 30° or 45° cone around the normal to the camera). Thus, it is possible to decrease or minimize the number of LEDs required to illuminate a space in which motion is detected by using a small number of wide-beam LEDs and a larger (or equal) number of narrow-beam LEDs directed toward the center field.

[00044] It is also possible to cover a wide field of view with many narrow-beam LEDs pointing in different directions. These may be operated by system 106' so as to scan the monitored space in order to identify the LEDs actually illuminating object 114; system 106' keeps power supplied to these LEDs and deactivates the others. In some embodiments, system 106' computes a predicted trajectory of the tracked object 114, using conventional predictive tracking techniques, and this trajectory is used to anticipate which LEDs should be activated as the object 114 moves. As object 114 moves through space as recorded by successive image frames of camera 102, system 106' plots the actual object trajectory and compares it, in real time, with the predicted trajectory, revising the latter — along with the illumination pattern — as new images are acquired and analyzed and tracking information derived therefrom. In other implementations, the LEDs are themselves physically or optically steerable. In such implementations, system 106' uses the predicted trajectory to steer beams to keep the object illuminated, activating new LEDs only when the object passes out of the steerable range of the LED currently tracking it. Any steering arrangement

amenable to computer control may be utilized for this purpose. For example, the LED itself may be mounted to a flexible substrate that is pivotably suspended over an angled fulcrum; a piezoelectric or spring-loaded electromagnetic actuator can be used to vary the angle of the LED across an axis. Alternatively, the LED output may pass through a steerable lens or other optical element that varies the orientation of the beam. Suitable arrangements for physically or optically steering LED output are well known in the scanning and display arts.

[00045] Image-analysis system 106, 106' (also referred to as an image analyzer) can include or consist of any device or device component that is capable of capturing and processing image data, e.g., using techniques described herein. FIG. 2 is a simplified block diagram of a computer system 200, implementing image-analysis system 106, 106' according to an embodiment of the present invention. Computer system 200 includes a processor 202, a memory 204, a camera interface 206, a display 208, speakers 209, a keyboard 210, and a mouse 211.

[00046] Memory 204 can be used to store instructions to be executed by processor 202 as well as input and/or output data associated with execution of the instructions. In particular, memory 204 contains instructions, conceptually illustrated as a group of modules described in greater detail below, that control the operation of processor 202 and its interaction with the other hardware components. An operating system directs the execution of low-level, basic system functions such as memory allocation, file management and operation of mass storage devices. The operating system may be or include a variety of operating systems such as Microsoft WINDOWS operating system, the Unix operating system, the Linux operating system, the Xenix operating system, the IBM AIX operating system, the Hewlett Packard UX operating system, the Novell NETWARE operating system, the Sun Microsystems SOLARIS operating system, the OS/2 operating system, the BeOS operating system, the MACINTOSH operating system, the APACHE operating system, an OPENSTEP operating system or another operating system of platform.

[00047] The computing environment may also include other removable/nonremovable, volatile/nonvolatile computer storage media. For example, a hard disk drive may read or write to nonremovable, nonvolatile magnetic media. A magnetic disk drive may read from or writes to a removable, nonvolatile magnetic disk, and an optical disk drive may read from or write to a removable, nonvolatile optical disk such as a CD-ROM or other optical media.

Other removable/nonremovable, volatile/nonvolatile computer storage media that can be used in the exemplary operating environment include, but are not limited to, magnetic tape cassettes, flash memory cards, digital versatile disks, digital video tape, solid state RAM, solid state ROM, and the like. The storage media are typically connected to the system bus through a removable or non-removable memory interface.

[00048] Processor 202 may be a general-purpose microprocessor, but depending on implementation can alternatively be a microcontroller, peripheral integrated circuit element, a CSIC (customer-specific integrated circuit), an ASIC (application-specific integrated circuit), a logic circuit, a digital signal processor, a programmable logic device such as an FPGA (field-programmable gate array), a PLD (programmable logic device), a PLA (programmable logic array), an RFID processor, smart chip, or any other device or arrangement of devices that is capable of implementing the processes of the invention.

[00049] Camera interface 206 can include hardware and/or software that enables communication between computer system 200 and cameras such as cameras 102, 104 shown in FIG. 1A, as well as associated light sources such as light sources 108, 110 of FIG. 1A. Thus, for example, camera interface 206 can include one or more data ports 216, 218 to which cameras can be connected, as well as hardware and/or software signal processors to modify data signals received from the cameras (e.g., to reduce noise or reformat data) prior to providing the signals as inputs to a conventional motion-capture (“mocap”) program 214 executing on processor 202. In some embodiments, camera interface 206 can also transmit signals to the cameras, e.g., to activate or deactivate the cameras, to control camera settings (frame rate, image quality, sensitivity, etc.), or the like. Such signals can be transmitted, e.g., in response to control signals from processor 202, which may in turn be generated in response to user input or other detected events.

[00050] Camera interface 206 can also include controllers 217, 219, to which light sources (e.g., light sources 108, 110) can be connected. In some embodiments, controllers 217, 219 supply operating current to the light sources, e.g., in response to instructions from processor 202 executing mocap program 214. In other embodiments, the light sources can draw operating current from an external power supply (not shown), and controllers 217, 219 can generate control signals for the light sources, e.g., instructing the light sources to be turned

on or off or changing the brightness. In some embodiments, a single controller can be used to control multiple light sources.

[00051] Instructions defining mocap program 214 are stored in memory 204, and these instructions, when executed, perform motion-capture analysis on images supplied from cameras connected to camera interface 206. In one embodiment, mocap program 214 includes various modules, such as an object detection module 222 and an object analysis module 224; again, both of these modules are conventional and well-characterized in the art. Object detection module 222 can analyze images (e.g., images captured via camera interface 206) to detect edges of an object therein and/or other information about the object's location. Object analysis module 224 can analyze the object information provided by object detection module 222 to determine the 3D position and/or motion of the object. Examples of operations that can be implemented in code modules of mocap program 214 are described below. Memory 204 can also include other information and/or code modules used by mocap program 214.

[00052] Display 208, speakers 209, keyboard 210, and mouse 211 can be used to facilitate user interaction with computer system 200. These components can be of generally conventional design or modified as desired to provide any type of user interaction. In some embodiments, results of motion capture using camera interface 206 and mocap program 214 can be interpreted as user input. For example, a user can perform hand gestures that are analyzed using mocap program 214, and the results of this analysis can be interpreted as an instruction to some other program executing on processor 200 (e.g., a web browser, word processor, or other application). Thus, by way of illustration, a user might use upward or downward swiping gestures to "scroll" a webpage currently displayed on display 208, to use rotating gestures to increase or decrease the volume of audio output from speakers 209, and so on.

[00053] It will be appreciated that computer system 200 is illustrative and that variations and modifications are possible. Computer systems can be implemented in a variety of form factors, including server systems, desktop systems, laptop systems, tablets, smart phones or personal digital assistants, and so on. A particular implementation may include other functionality not described herein, e.g., wired and/or wireless network interfaces, media playing and/or recording capability, etc. In some embodiments, one or more cameras may be

built into the computer rather than being supplied as separate components. Further, an image analyzer can be implemented using only a subset of computer system components (e.g., as a processor executing program code, an ASIC, or a fixed-function digital signal processor, with suitable I/O interfaces to receive image data and output analysis results).

[00054] While computer system 200 is described herein with reference to particular blocks, it is to be understood that the blocks are defined for convenience of description and are not intended to imply a particular physical arrangement of component parts. Further, the blocks need not correspond to physically distinct components. To the extent that physically distinct components are used, connections between components (e.g., for data communication) can be wired and/or wireless as desired.

[00055] Execution of object detection module 222 by processor 202 can cause processor 202 to operate camera interface 206 to capture images of an object and to distinguish object pixels from background pixels by analyzing the image data. FIGs. 3A-3C are three different graphs of brightness data for rows of pixels that may be obtained according to various embodiments of the present invention. While each graph illustrates one pixel row, it is to be understood that an image typically contains many rows of pixels, and a row can contain any number of pixels; for instance, an HD video image can include 1080 rows having 1920 pixels each.

[00056] FIG. 3A illustrates brightness data 300 for a row of pixels in which the object has a single cross-section, such as a cross-section through a palm of a hand. Pixels in region 302, corresponding to the object, have high brightness while pixels in regions 304 and 306, corresponding to background, have considerably lower brightness. As can be seen, the object's location is readily apparent, and the locations of the edges of the object (at 308 at 310) are easily identified. For example, any pixel with brightness above 0.5 can be assumed to be an object pixel, while any pixel with brightness below 0.5 can be assumed to be a background pixel.

[00057] **[0038]** FIG. 3B illustrates brightness data 320 for a row of pixels in which the object has multiple distinct cross-sections, such as a cross-section through fingers of an open hand. Regions 322, 323, and 324, corresponding to the object, have high brightness while pixels in regions 326-329, corresponding to background, have low brightness. Again, a

simple threshold cutoff on brightness (e.g., at 0.5) suffices to distinguish object pixels from background pixels, and the edges of the object can be readily ascertained.

[00058] FIG. 3C illustrates brightness data 340 for a row of pixels in which the distance to the object varies across the row, such as a cross-section of a hand with two fingers extending toward the camera. Regions 342 and 343 correspond to the extended fingers and have highest brightness; regions 344 and 345 correspond to other portions of the hand and are slightly less bright; this can be due in part to being farther away in part to shadows cast by the extended fingers. Regions 348 and 349 are background regions and are considerably darker than hand-containing regions 342-345. A threshold cutoff on brightness (e.g., at 0.5) again suffices to distinguish object pixels from background pixels. Further analysis of the object pixels can also be performed to detect the edges of regions 342 and 343, providing additional information about the object's shape.

[00059] It will be appreciated that the data shown in FIGs. 3A-3C is illustrative. In some embodiments, it may be desirable to adjust the intensity of light sources 108, 110 such that an object at an expected distance (e.g., r_0 in FIG. 1A) will be overexposed – that is, many if not all of the object pixels will be fully saturated to a brightness level of 1.0. (The actual brightness of the object may in fact be higher.) While this may also make the background pixels somewhat brighter, the $1/r^2$ falloff of light intensity with distance still leads to a ready distinction between object and background pixels as long as the intensity is not set so high that background pixels also approach the saturation level. As FIGs. 3A-3C illustrate, use of lighting directed at the object to create strong contrast between object and background allows the use of simple and fast algorithms to distinguish between background pixels and object pixels, which can be particularly useful in real-time motion-capture systems. Simplifying the task of distinguishing background and object pixels can also free up computing resources for other motion-capture tasks (e.g., reconstructing the object's position, shape, and/or motion).

[00060] Refer now to FIG. 4, which illustrates a process 400 for identifying the location of an object in an image according to an embodiment of the present invention. Process 400 can be implemented, e.g., in system 100 of FIG. 1A. At block 402, light sources 108, 110 are turned on. At block 404, one or more images are captured using cameras 102, 104. In some embodiments, one image from each camera is captured. In other embodiments, a sequence of images is captured from each camera. The images from the two cameras can be closely

correlated in time (e.g., simultaneous to within a few milliseconds) so that correlated images from the two cameras can be used to determine the 3D location of the object.

[00061] At block 406, a threshold pixel brightness is applied to distinguish object pixels from background pixels. Block 406 can also include identifying locations of edges of the object based on transition points between background and object pixels. In some embodiments, each pixel is first classified as either object or background based on whether it exceeds the threshold brightness cutoff. For example, as shown in FIGs. 3A-3C, a cutoff at a saturation level of 0.5 can be used. Once the pixels are classified, edges can be detected by finding locations where background pixels are adjacent to object pixels. In some embodiments, to avoid noise artifacts, the regions of background and object pixels on either side of the edge may be required to have a certain minimum size (e.g., 2, 4 or 8 pixels).

[00062] In other embodiments, edges can be detected without first classifying pixels as object or background. For example, $\Delta\beta$ can be defined as the difference in brightness between adjacent pixels, and $|\Delta\beta|$ above a threshold (e.g., 0.3 or 0.5 in terms of the saturation scale) can indicate a transition from background to object or from object to background between adjacent pixels. (The sign of $\Delta\beta$ can indicate the direction of the transition.) In some instances where the object's edge is actually in the middle of a pixel, there may be a pixel with an intermediate value at the boundary. This can be detected, e.g., by computing two brightness values for a pixel i : $\beta_L = (\beta_i + \beta_{i-1})/2$ and $\beta_R = (\beta_i + \beta_{i+1})/2$, where pixel $(i-1)$ is to the left of pixel i and pixel $(i+1)$ is to the right of pixel i . If pixel i is not near an edge, $|\beta_L - \beta_R|$ will generally be close to zero; if pixel i is near an edge, then $|\beta_L - \beta_R|$ will be closer to 1, and a threshold on $|\beta_L - \beta_R|$ can be used to detect edges.

[00063] In some instances, one part of an object may partially occlude another in an image; for example, in the case of a hand, a finger may partly occlude the palm or another finger. Occlusion edges that occur where one part of the object partially occludes another can also be detected based on smaller but distinct changes in brightness once background pixels have been eliminated. FIG. 3C illustrates an example of such partial occlusion, and the locations of occlusion edges are apparent.

[00064] Detected edges can be used for numerous purposes. For example, as previously noted, the edges of the object as viewed by the two cameras can be used to determine an approximate location of the object in 3D space. The position of the object in a 2D plane

transverse to the optical axis of the camera can be determined from a single image, and the offset (parallax) between the position of the object in time-correlated images from two different cameras can be used to determine the distance to the object if the spacing between the cameras is known.

[00065] Further, the position and shape of the object can be determined based on the locations of its edges in time-correlated images from two different cameras, and motion (including articulation) of the object can be determined from analysis of successive pairs of images. Examples of techniques that can be used to determine an object's position, shape and motion based on locations of edges of the object are described in co-pending U.S. Serial No. 13/414,485, filed March 7, 2012, the entire disclosure of which is incorporated herein by reference. Those skilled in the art with access to the present disclosure will recognize that other techniques for determining position, shape and motion of an object based on information about the location of edges of the object can also be used.

[00066] In some embodiments, light sources 108, 110 can be operated in a pulsed mode rather than being continually on. This can be useful, e.g., if light sources 108, 110 have the ability to produce brighter light in a pulse than in a steady-state operation. FIG. 5 illustrates a timeline in which light sources 108, 110 are pulsed on at regular intervals as shown at 502. The shutters of cameras 102, 104 can be opened to capture images at times coincident with the light pulses as shown at 504. Thus, an object of interest can be brightly illuminated during the times when images are being captured.

[00067] In some embodiments, the pulsing of light sources 108, 110 can be used to further enhance contrast between an object of interest and background by comparing images taken with lights 108, 110 on and images taken with lights 108, 110 off. FIG. 6 illustrates a timeline in which light sources 108, 110 are pulsed on at regular intervals as shown at 602, while shutters of cameras 102, 104 are opened to capture images at times shown at 604. In this case, light sources 108, 110 are "on" for every other image. If the object of interest is significantly closer than background regions to light sources 108, 110, the difference in light intensity will be stronger for object pixels than for background pixels. Accordingly, comparing pixels in successive images can help distinguish object and background pixels.

[00068] FIG. 7 is a flow diagram of a process 700 for identifying object edges using successive images according to an embodiment of the present invention. At block 702, the

light sources are turned off, and at block 704 a first image (A) is captured. Then, at block 706, the light sources are turned on, and at block 708 a second image (B) is captured. At block 710, a “difference” image $B - A$ is calculated, e.g., by subtracting the brightness value of each pixel in image A from the brightness value of the corresponding pixel in image B. Since image B was captured with lights on, it is expected that $B - A$ will be positive for most pixels.

[00069] At block 712, a threshold can be applied to the difference image ($B - A$) to identify object pixels, with ($B - A$) above a threshold being associated with object pixels and ($B - A$) below the threshold being associated with background pixels. Object edges can then be defined by identifying where object pixels are adjacent to background pixels, as described above. Object edges can be used for purposes such as position and/or motion detection, as described above.

[00070] Contrast-based object detection as described herein can be applied in any situation where objects of interest are expected to be significantly closer (e.g., half the distance) to the light source(s) than background objects. One such application relates to the use of motion-detection as user input to interact with a computer system. For example, the user may point to the screen or make other hand gestures, which can be interpreted by the computer system as input.

[00071] A computer system 800 incorporating a motion detector as a user input device according to an embodiment of the present invention is illustrated in FIG. 8. Computer system 800 includes a desktop box 802 that can house various components of a computer system such as processors, memory, fixed or removable disk drives, video drivers, audio drivers, network interface components, and so on. A display 804 is connected to desktop box 802 and positioned to be viewable by a user. A keyboard 806 is positioned within easy reach of the user's hands. A motion-detector unit 808 is placed near keyboard 806 (e.g., behind, as shown or to one side), oriented toward a region in which it would be natural for the user to make gestures directed at display 804 (e.g., a region in the air above the keyboard and in front of the monitor). Cameras 810, 812 (which can be similar or identical to cameras 102, 104 described above) are arranged to point generally upward, and light sources 814, 816 (which can be similar or identical to light sources 108, 110 described above) are arranged to either side of cameras 810, 812 to illuminate an area above motion-detector unit 808. In

typical implementations, the cameras 810, 812 and the light sources 814, 816 are substantially coplanar. This configuration prevents the appearance of shadows that can, for example, interfere with edge detection (as can be the case were the light sources located between, rather than flanking, the cameras). A filter, not shown, can be placed over the top of motion-detector unit 808 (or just over the apertures of cameras 810, 812) to filter out all light outside a band around the peak frequencies of light sources 814, 816.

[00072] In the illustrated configuration, when the user moves a hand or other object (e.g., a pencil) in the field of view of cameras 810, 812, the background will likely consist of a ceiling and/or various ceiling-mounted fixtures. The user's hand can be 10–20 cm above motion detector 808, while the ceiling may be five to ten times that distance (or more). Illumination from light sources 814, 816 will therefore be much more intense on the user's hand than on the ceiling, and the techniques described herein can be used to reliably distinguish object pixels from background pixels in images captured by cameras 810, 812. If infrared light is used, the user will not be distracted or disturbed by the light.

[00073] Computer system 800 can utilize the architecture shown in FIG. 1A. For example, cameras 810, 812 of motion-detector unit 808 can provide image data to desktop box 802, and image analysis and subsequent interpretation can be performed using the processors and other components housed within desktop box 802. Alternatively, motion-detector unit 808 can incorporate processors or other components to perform some or all stages of image analysis and interpretation. For example, motion-detector unit 808 can include a processor (programmable or fixed-function) that implements one or more of the processes described above to distinguish between object pixels and background pixels. In this case, motion-detector unit 808 can send a reduced representation of the captured images (e.g., a representation with all background pixels zeroed out) to desktop box 802 for further analysis and interpretation. A particular division of computational tasks between a processor inside motion-detector unit 808 and a processor inside desktop box 802 is not required.

[00074] It is not always necessary to discriminate between object pixels and background pixels by absolute brightness levels; for example, where knowledge of object shape exists, the pattern of brightness falloff can be utilized to detect the object in an image even without explicit detection of object edges. On rounded objects (such as hands and fingers), for example, the $1/r^2$ relationship produces Gaussian or near-Gaussian brightness distributions

near the centers of the objects; imaging a cylinder illuminated by an LED and disposed perpendicularly with respect to a camera results in an image having a bright center line corresponding to the cylinder axis, with brightness falling off to each side (around the cylinder circumference). Fingers are approximately cylindrical, and by identifying these Gaussian peaks, it is possible to locate fingers even in situations where the background is close and the edges are not visible due to the relative brightness of the background (either due to proximity or the fact that it may be actively emitting infrared light). The term “Gaussian” is used broadly herein to connote a bell-shaped curve that is typically symmetric, and is not limited to curves explicitly conforming to a Gaussian function.

[00075] FIG. 9 illustrates a tablet computer 900 incorporating a motion detector according to an embodiment of the present invention. Tablet computer 900 has a housing, the front surface of which incorporates a display screen 902 surrounded by a bezel 904. One or more control buttons 906 can be incorporated into bezel 904. Within the housing, e.g., behind display screen 902, tablet computer 900 can have various conventional computer components (processors, memory, network interfaces, etc.). A motion detector 910 can be implemented using cameras 912, 914 (e.g., similar or identical to cameras 102, 104 of FIG. 1A) and light sources 916, 918 (e.g., similar or identical to light sources 108, 110 of FIG. 1A) mounted into bezel 904 and oriented toward the front surface so as to capture motion of a user positioned in front of tablet computer 900.

[00076] When the user moves a hand or other object in the field of view of cameras 912, 914, the motion is detected as described above. In this case, the background is likely to be the user’s own body, at a distance of roughly 25–30 cm from tablet computer 900. The user may hold a hand or other object at a short distance from display 902, e.g., 5–10 cm. As long as the user’s hand is significantly closer than the user’s body (e.g., half the distance) to light sources 916, 918, the illumination-based contrast enhancement techniques described herein can be used to distinguish object pixels from background pixels. The image analysis and subsequent interpretation as input gestures can be done within tablet computer 900 (e.g., leveraging the main processor to execute operating-system or other software to analyze data obtained from cameras 912, 914). The user can thus interact with tablet 900 using gestures in 3D space.

[00077] A goggle system 1000, as shown in FIG. 10, may also incorporate a motion detector according to an embodiment of the present invention. Goggle system 1000 can be used, e.g., in connection with virtual-reality and/or augmented-reality environments. Goggle system 1000 includes goggles 1002 that are wearable by a user, similar to conventional eyeglasses. Goggles 1002 include eyepieces 1004, 1006 that can incorporate small display screens to provide images to the user's left and right eyes, e.g., images of a virtual reality environment. These images can be provided by a base unit 1008 (e.g., a computer system) that is in communication with goggles 1002, either via a wired or wireless channel. Cameras 1010, 1012 (e.g., similar or identical to cameras 102, 104 of FIG. 1A) can be mounted in a frame section of goggles 1002 such that they do not obscure the user's vision. Light sources 1014, 1016 can be mounted in the frame section of goggles 1002 to either side of cameras 1010, 1012. Images collected by cameras 1010, 1012 can be transmitted to base unit 1008 for analysis and interpretation as gestures indicating user interaction with the virtual or augmented environment. (In some embodiments, the virtual or augmented environment presented through eyepieces 1004, 1006 can include a representation of the user's hand, and that representation can be based on the images collected by cameras 1010, 1012.)

[00078] When the user gestures using a hand or other object in the field of view of cameras 1008, 1010, the motion is detected as described above. In this case, the background is likely to be a wall of a room the user is in, and the user will most likely be sitting or standing at some distance from the wall. As long as the user's hand is significantly closer than the user's body (e.g., half the distance) to light sources 1012, 1014, the illumination-based contrast enhancement techniques described herein facilitate distinguishing object pixels from background pixels. The image analysis and subsequent interpretation as input gestures can be done within base unit 1008.

[00079] It will be appreciated that the motion-detector implementations shown in FIGs. 8–10 are illustrative and that variations and modifications are possible. For example, a motion detector or components thereof can be combined in a single housing with other user input devices, such as a keyboard or trackpad. As another example, a motion detector can be incorporated into a laptop computer, e.g., with upward-oriented cameras and light sources built into the same surface as the laptop keyboard (e.g., to one side of the keyboard or in front of or behind it) or with front-oriented cameras and light sources built into a bezel surrounding the laptop's display screen. As still another example, a wearable motion

detector can be implemented, e.g., as a headband or headset that does not include active displays or optical components.

[00080] As illustrated in FIG. 11, motion information can be used as user input to control a computer system or other system according to an embodiment of the present invention. Process 1100 can be implemented, e.g., in computer systems such as those shown in FIGs. 8–10. At block 1102, images are captured using the light sources and cameras of the motion detector. As described above, capturing the images can include using the light sources to illuminate the field of view of the cameras such that objects closer to the light sources (and the cameras) are more brightly illuminated than objects farther away.

[00081] At block 1104, the captured images are analyzed to detect edges of the object based on changes in brightness. For example, as described above, this analysis can include comparing the brightness of each pixel to a threshold, detecting transitions in brightness from a low level to a high level across adjacent pixels, and/or comparing successive images captured with and without illumination by the light sources. At block 1106, an edge-based algorithm is used to determine the object's position and/or motion. This algorithm can be, for example, any of the tangent-based algorithms described in the above-referenced '485 application; other algorithms can also be used.

[00082] At block 1108, a gesture is identified based on the object's position and/or motion. For example, a library of gestures can be defined based on the position and/or motion of a user's fingers. A "tap" can be defined based on a fast motion of an extended finger toward a display screen. A "trace" can be defined as motion of an extended finger in a plane roughly parallel to the display screen. An inward pinch can be defined as two extended fingers moving closer together and an outward pinch can be defined as two extended fingers moving farther apart. Swipe gestures can be defined based on movement of the entire hand in a particular direction (e.g., up, down, left, right) and different swipe gestures can be further defined based on the number of extended fingers (e.g., one, two, all). Other gestures can also be defined. By comparing a detected motion to the library, a particular gesture associated with detected position and/or motion can be determined.

[00083] At block 1110, the gesture is interpreted as user input, which the computer system can process. The particular processing generally depends on application programs currently executing on the computer system and how those programs are configured to respond to

particular inputs. For example, a tap in a browser program can be interpreted as selecting a link toward which the finger is pointing. A tap in a word-processing program can be interpreted as placing the cursor at a position where the finger is pointing or as selecting a menu item or other graphical control element that may be visible on the screen. The particular gestures and interpretations can be determined at the level of operating systems and/or applications as desired, and no particular interpretation of any gesture is required.

[00084] Full-body motion can be captured and used for similar purposes. In such embodiments, the analysis and reconstruction advantageously occurs in approximately real-time (e.g., times comparable to human reaction times), so that the user experiences a natural interaction with the equipment. In other applications, motion capture can be used for digital rendering that is not done in real time, e.g., for computer-animated movies or the like; in such cases, the analysis can take as long as desired.

[00085] Embodiments described herein provide efficient discrimination between object and background in captured images by exploiting the decrease of light intensity with distance. By brightly illuminating the object using one or more light sources that are significantly closer to the object than to the background (e.g., by a factor of two or more), the contrast between object and background can be increased. In some instances, filters can be used to remove light from sources other than the intended sources. Using infrared light can reduce unwanted “noise” or bright spots from visible light sources likely to be present in the environment where images are being captured and can also reduce distraction to users (who presumably cannot see infrared).

[00086] The embodiments described above provide two light sources, one disposed to either side of the cameras used to capture images of the object of interest. This arrangement can be particularly useful where the position and motion analysis relies on knowledge of the object’s edges as seen from each camera, as the light sources will illuminate those edges. However, other arrangements can also be used. For example, FIG. 12 illustrates a system 1200 with a single camera 1202 and two light sources 1204, 1206 disposed to either side of camera 1202. This arrangement can be used to capture images of object 1208 and shadows cast by object 1208 against a flat background region 1210. In this embodiment, object pixels and background pixels can be readily distinguished. In addition, provided that background 1210 is not too far from object 1208, there will still be enough contrast between pixels in the

shadowed background region and pixels in the unshadowed background region to allow discrimination between the two. Position and motion detection algorithms using images of an object and its shadows are described in the above-referenced '485 application and system 1200 can provide input information to such algorithms, including the location of edges of the object and its shadows.

[00087] FIG. 13 illustrates another system 1300 with two cameras 1302, 1304 and one light source 1306 disposed between the cameras. System 1300 can capture images of an object 1308 against a background 1310. System 1300 is generally less reliable for edge illumination than system 100 of FIG. 1A; however, not all algorithms for determining position and motion rely on precise knowledge of the edges of an object. Accordingly, system 1300 can be used, e.g., with edge-based algorithms in situations where less accuracy is required. System 1300 can also be used with non-edge-based algorithms.

[00088] While the invention has been described with respect to specific embodiments, one skilled in the art will recognize that numerous modifications are possible. The number and arrangement of cameras and light sources can be varied. The cameras' capabilities, including frame rate, spatial resolution, and intensity resolution, can also be varied as desired. The light sources can be operated in continuous or pulsed mode. The systems described herein provide images with enhanced contrast between object and background to facilitate distinguishing between the two, and this information can be used for numerous purposes, of which position and/or motion detection is just one among many possibilities.

[00089] Threshold cutoffs and other specific criteria for distinguishing object from background can be adapted for particular cameras and particular environments. As noted above, contrast is expected to increase as the ratio r_B/r_O increases. In some embodiments, the system can be calibrated in a particular environment, e.g., by adjusting light-source brightness, threshold criteria, and so on. The use of simple criteria that can be implemented in fast algorithms can free up processing power in a given system for other uses.

[00090] Any type of object can be the subject of motion capture using these techniques, and various aspects of the implementation can be optimized for a particular object. For example, the type and positions of cameras and/or light sources can be optimized based on the size of the object whose motion is to be captured and/or the space in which motion is to be captured. Analysis techniques in accordance with embodiments of the present invention

can be implemented as algorithms in any suitable computer language and executed on programmable processors. Alternatively, some or all of the algorithms can be implemented in fixed-function logic circuits, and such circuits can be designed and fabricated using conventional or other tools.

[00091] Computer programs incorporating various features of the present invention may be encoded on various computer readable storage media; suitable media include magnetic disk or tape, optical storage media such as compact disk (CD) or DVD (digital versatile disk), flash memory, and any other non-transitory medium capable of holding data in a computer-readable form. Computer-readable storage media encoded with the program code may be packaged with a compatible device or provided separately from other devices. In addition program code may be encoded and transmitted via wired optical, and/or wireless networks conforming to a variety of protocols, including the Internet, thereby allowing distribution, e.g., via Internet download.

[00092] Thus, although the invention has been described with respect to specific embodiments, it will be appreciated that the invention is intended to cover all modifications and equivalents within the scope of the following claims.

What is claimed is:

CLAIMS

1. An object illumination and image capture part of a person-machine interface comprising:
 - a camera oriented toward a field of view;
 - a plurality of individually controllable lighting elements directed toward the field of view, at least some of the lighting elements having different beam angles, including at least one wide-beam lighting element and one or more narrow-beam lighting elements; and
 - a controller coupled to the camera and light sources and configured to:
 - operate the camera to capture a sequence of images including a control object arriving at and moving in the field of view;
 - operate the lighting elements to illuminate the field of view with the wide-beam lighting element while detecting the arrival of the control object and to activate the narrow-beam lighting elements and follow the control object with the narrow-beam lighting elements as it moves within the field of view; and
 - illuminate the moving control object using the narrow-beam lighting elements with sufficient intensity to facilitate detection of control gestures by the control object within the field of view.
2. The person-machine interface of claim 1, wherein a plurality of lighting elements are narrow-beam lighting elements.
3. The person-machine interface of claim 2, wherein the narrow-beam lighting elements have a beam angle of approximately 60°.
4. The person-machine interface of claim 2, wherein the at least one wide-beam lighting element has a beam angle of approximately 120°.
5. The person-machine interface of claim 2, wherein the controller is configured to track the object and activate different ones of the narrow-beam lighting elements to maintain the object in illumination as it moves.
6. The person-machine interface of claim 5, wherein the controller is configured to deactivate lighting elements not necessary to illuminate the object.
7. The person-machine interface of claim 2, wherein the controller is configured to steer the narrow-beam lighting elements to maintain the object in illumination by at least one such element as the object moves.

1 8. The person-machine interface of claim 1, wherein the controller is configured to analyze
2 images from the camera to predict a trajectory of the object and to activate or steer lighting
3 elements in accordance with predicted trajectory.

1 9. An object illumination and image capture system part of a person-machine interface
2 comprising:

3 a camera oriented toward a field of view;

4 a plurality of individually controllable lighting elements and controllable optics
5 associated with at least some of the lighting elements, the optics controllably altering a beam
6 angle of associated lighting elements; and

7 a controller coupled to the camera and light sources and configured to:

8 operate the camera to capture a sequence of images including a control object
9 moving in the field of view;

10 operate the lighting elements and the optics to illuminate the field of view with at
11 least one broad beam while detecting arrival of the control object and with narrower
12 beams while following the control object as it moves within the field of view; and

13 illuminate the moving control object with sufficient intensity to facilitate
14 application of object axis detection to the sequence of images of the control object.

1 10. The system of claim 9, wherein the optics are tunable lenses.

1 11. The system of claim 10, wherein the controller controls the optics to vary a beam angle of
2 a lighting element and a drive current to vary an output luminance of the lighting element .

1 12. A method of object illumination and image capture as a part of a person-machine
2 interface used with a camera and a plurality of lighting elements oriented toward a field of view,
3 at least some of the lighting elements having different beam angles, the method comprising
4 actions of:

5 operating the camera to capture a sequence of images including a control object moving
6 in the field of view; and

7 operating the lighting elements to illuminate the field of view with wide-beam lighting
8 element while detecting arrival of the control object and to activate narrow-beam lighting
9 elements and follow the control object with the narrow-beam lighting elements as it moves
10 within the field of view; and

11 illuminating the moving object with sufficient intensity to facilitate application of object
12 axis detection to the sequence of images of the control object.

1 13. The method of claim 12, wherein the lighting elements include a plurality of narrow-
2 beam lighting elements.

1 14. The method of claim 13, wherein the narrow-beam lighting elements have a beam angle
2 of approximately 60°.

1 15. The method of claim 13, wherein the at least one wide-beam lighting element has a beam
2 angle of approximately 120°.

1 16. The method of claim 13, wherein the object is tracked and different ones of the narrow-
2 beam lighting elements are activated to maintain the object in illumination as it moves.

1 17. The method of claim 16, further comprising the step of deactivating lighting elements not
2 necessary to illuminate the object.

1 18. The method of claim 13, further comprising the step of steering the narrow-beam lighting
2 elements to maintain the object in illumination by at least one such element as the object moves.

1 19. The method of claim 12, further comprising the step of analyzing images from the
2 camera to predict a trajectory of the object and activating or steering lighting elements in
3 accordance with predicted trajectory.

1 20. A method of object illumination and image capture as a part of a person-machine
2 interface used with a camera and a plurality of individually controllable lighting elements
3 oriented toward a field of view and controllable optics associated with at least some of the
4 lighting elements, the optics controllably altering a beam angle of associated lighting elements,
5 the method comprising actions of:

6 operating the camera to capture a sequence of images including a control object moving
7 in the field of view;

8 operating the lighting elements and the optics to illuminate the field of view with wide-
9 beam lighting element while detecting arrival of the control object and to activate narrow-
10 beam lighting elements and follow the control object with the narrow-beam lighting elements
11 as it moves within the field of view; and

12 illuminating the moving object with sufficient intensity to facilitate detection of control
13 gestures by the control object within the field of view.

1 21. An image capture and analysis system comprising:

2 a camera oriented toward a field of view;

3 a plurality of individually controllable lighting elements directed toward the field of
4 view, at least some of the lighting elements having different beam angles; and

5 a controller coupled to the camera and light sources and configured to:

6 operate the camera to capture a sequence of images including an object moving in the
7 field of view; and

8 operate the lighting elements to illuminate the moving object with sufficient intensity to facilitate
9 image capture by the camera but with minimum supplied power.

1 22. The system of claim 21, wherein at least one of the lighting elements is a wide-beam
2 lighting element and a plurality of lighting elements are narrow-beam lighting elements.

1 23. The system of claim 22, wherein the narrow-beam lighting elements have a beam angle
2 of approximately 60°.

1 24. The system of claim 22, wherein the at least one wide-beam lighting element has a beam
2 angle of approximately 120°.

1 25. The system of claim 22, wherein the controller is configured to track the object and
2 activate different ones of the narrow-beam lighting elements to maintain the object in
3 illumination as it moves.

1 26. The system of claim 25, wherein the controller is configured to deactivate lighting
2 elements not necessary to illuminate the object.

1 27. The system of claim 22, wherein the controller is configured to steer the narrow-beam
2 lighting elements to maintain the object in illumination by at least one such element as the object
3 moves.

1 28. The system of claim 21, wherein the controller is configured to analyze images from the
2 camera to predict a trajectory of the object and to activate or steer lighting elements in
3 accordance with predicted trajectory.

1 29. An image capture and analysis system comprising:

2 a camera oriented toward a field of view;

3 a plurality of individually controllable lighting elements and controllable optics
4 associated with at least some of the lighting elements, the optics controllably altering a beam
5 angle of associated lighting elements; and

6 a controller coupled to the camera and light sources and configured to:

7 operate the camera to capture a sequence of images including an object moving in the
8 field of view; and

9 operate the lighting elements and the optics to illuminate the moving object with sufficient
10 intensity to facilitate image capture by the camera but with minimum supplied power.

1 30. The system of claim 29, wherein the optics are tunable lenses.

1 31. The system of claim 30, wherein the controller controls the optics to vary a beam angle of
2 a lighting element and a drive current to vary an output luminance of the lighting element to
3 minimize the drive current.

1 32. A method of image capture and analysis for use with a camera and a plurality of lighting
2 elements oriented toward a field of view, at least some of the lighting elements having different
3 beam angles, the method comprising:

4 operating the camera to capture a sequence of images including an object moving in the
5 field of view; and

6 operating the lighting elements to illuminate the moving object with sufficient intensity to
7 facilitate image capture by the camera but with minimum supplied power.

1 33. The method of claim 32, wherein the lighting elements include at least one wide-beam
2 lighting element and a plurality of narrow-beam lighting elements.

1 34. The method of claim 33, wherein the narrow-beam lighting elements have a beam angle
2 of approximately 60°.

1 35. The method of claim 33, wherein the at least one wide-beam lighting element has a beam
2 angle of approximately 120°.

1 36. The method of claim 33, wherein the object is tracked and different ones of the narrow-
2 beam lighting elements are activated to maintain the object in illumination as it moves.

1 37. The method of claim 36, further comprising the step of deactivating lighting elements not
2 necessary to illuminate the object.

1 38. The method of claim 33, further comprising the step of steering the narrow-beam lighting
2 elements to maintain the object in illumination by at least one such element as the object moves.

1 39. The method of claim 32, further comprising the step of analyzing images from the
2 camera to predict a trajectory of the object and activating or steering lighting elements in
3 accordance with predicted trajectory.

1 40. A method of image capture and analysis for use with a camera and a plurality of
2 individually controllable lighting elements oriented toward a field of view and controllable
3 optics associated with at least some of the lighting elements, the optics controllably altering a
4 beam angle of associated lighting elements, the method comprising:

5 operating the camera to capture a sequence of images including an object moving in the
6 field of view;

7 operating the lighting elements and the optics to illuminate the moving object with
8 sufficient intensity to facilitate image capture by the camera but with minimum supplied
9 power.

1/9

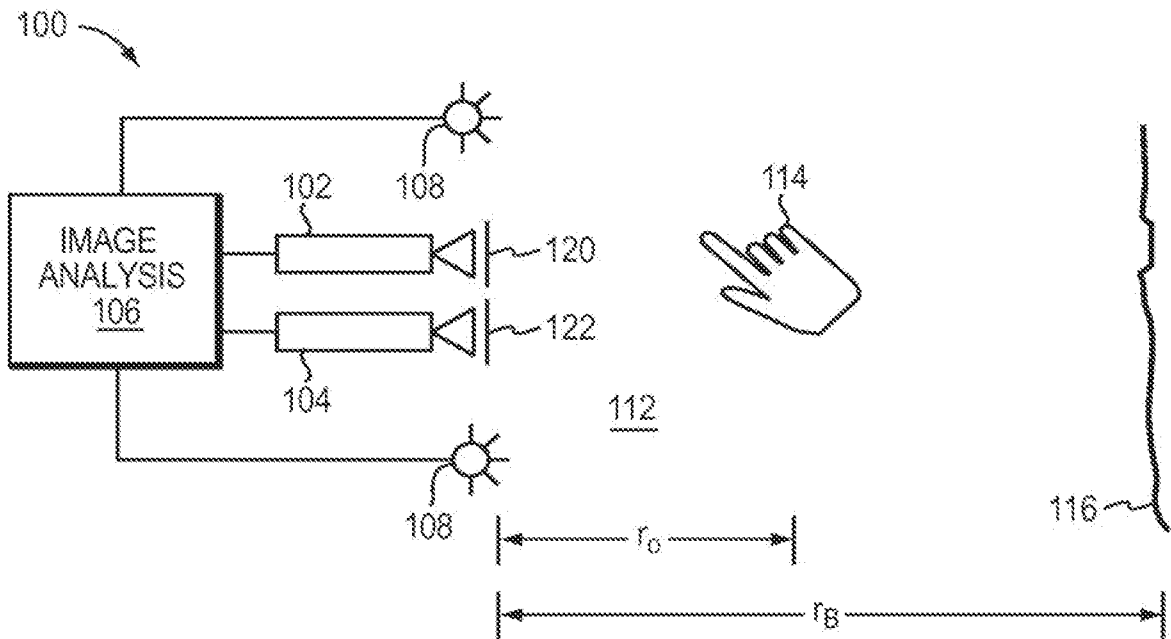


FIG. 1A

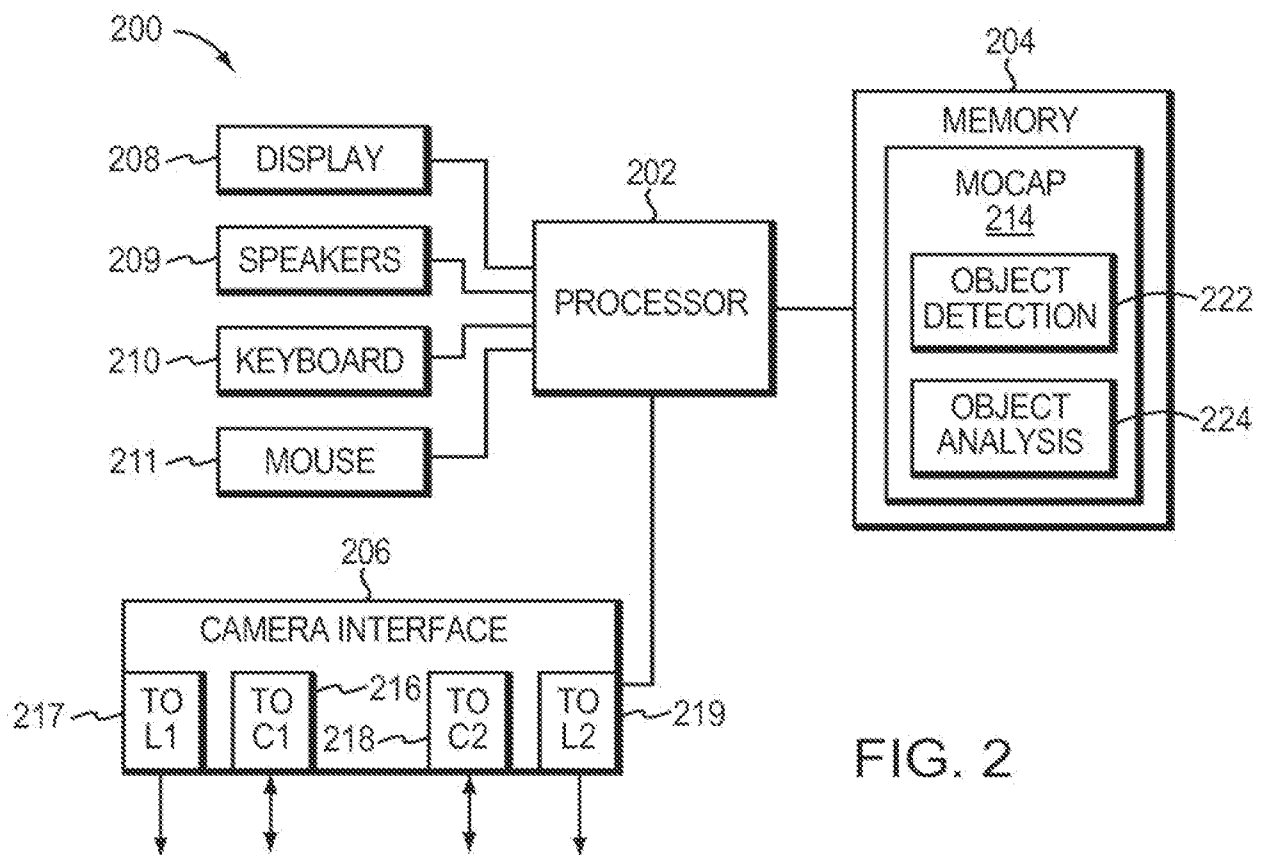


FIG. 2

2/9

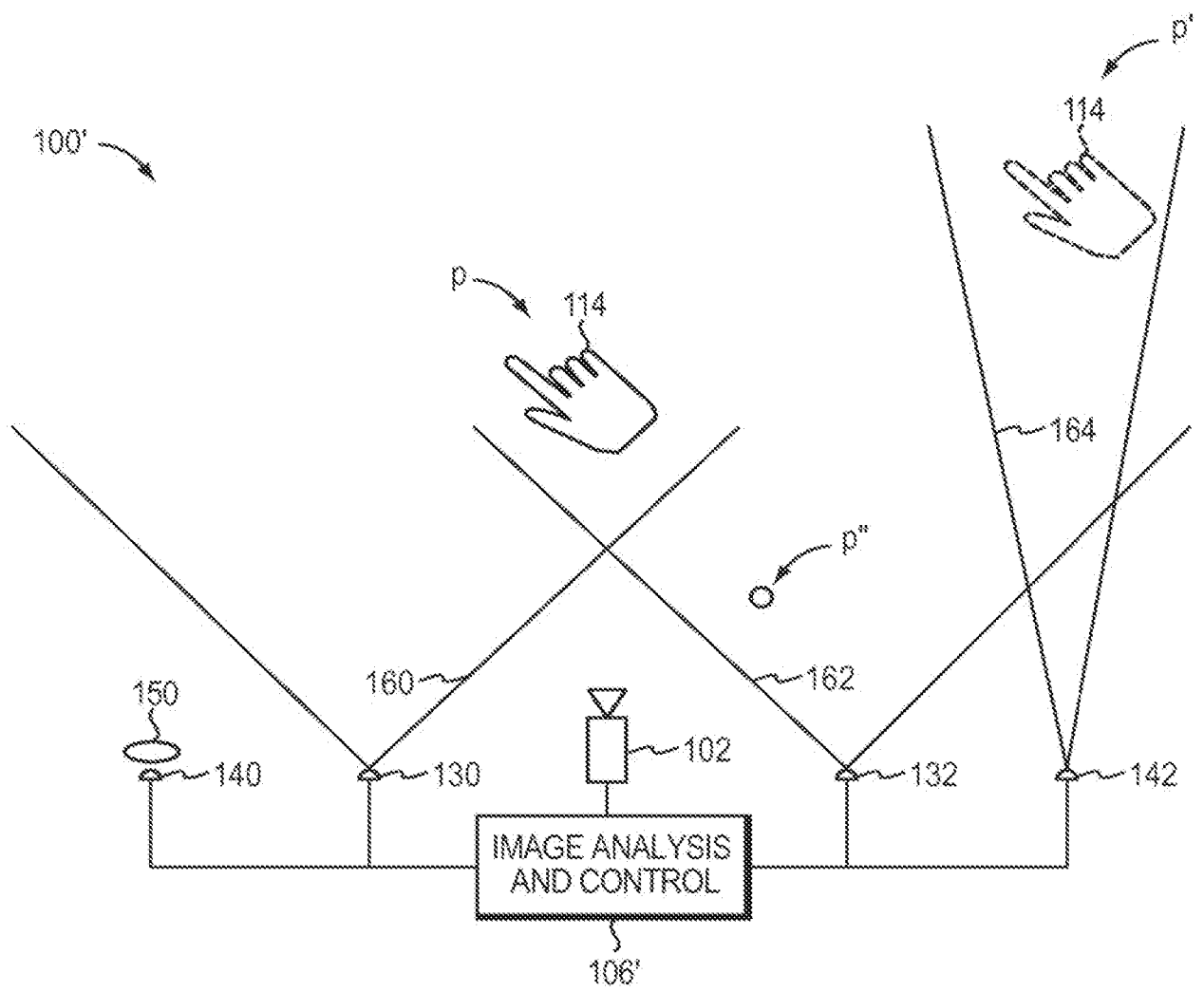


FIG. 1B

3/9

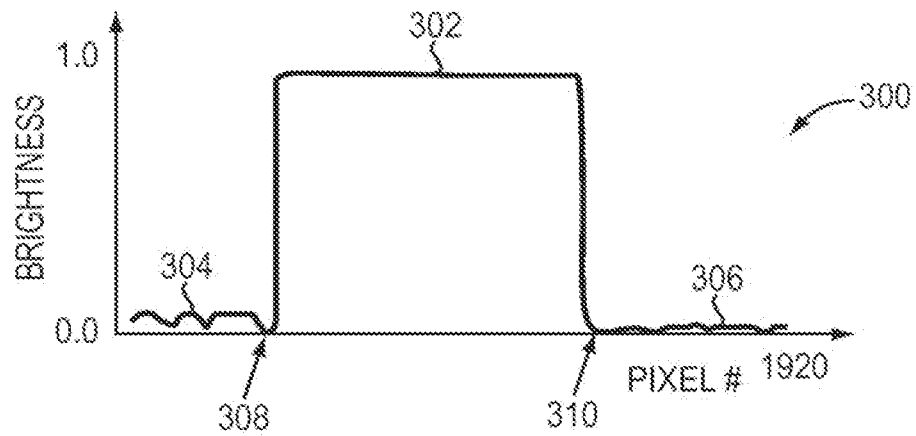


FIG. 3A

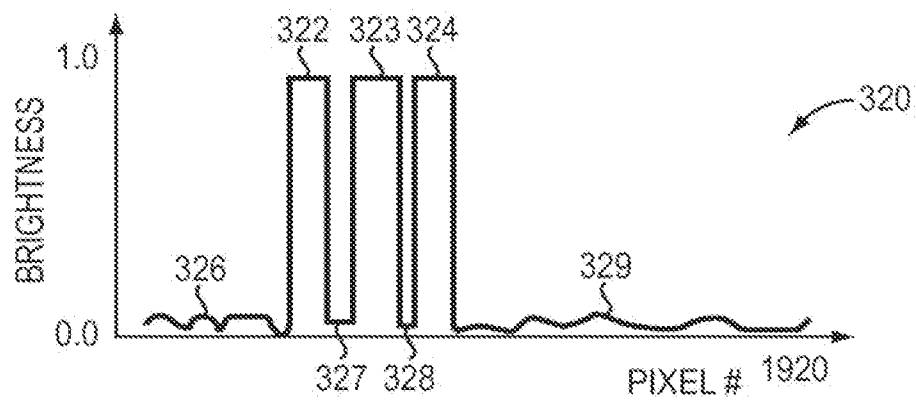


FIG. 3B

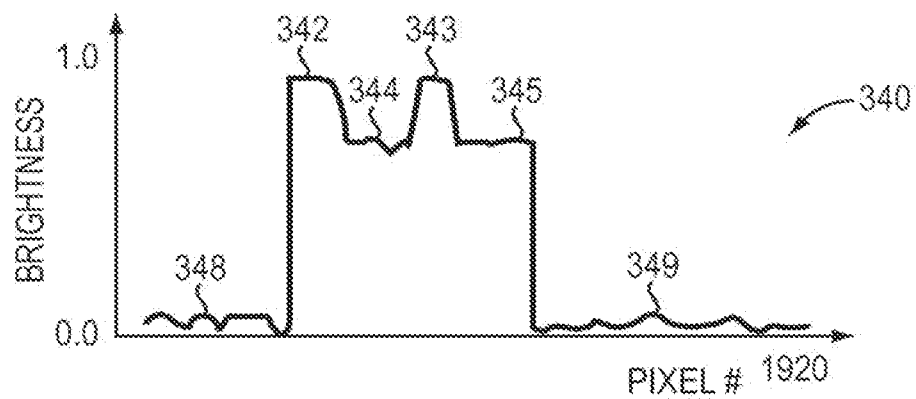


FIG. 3C

4/9

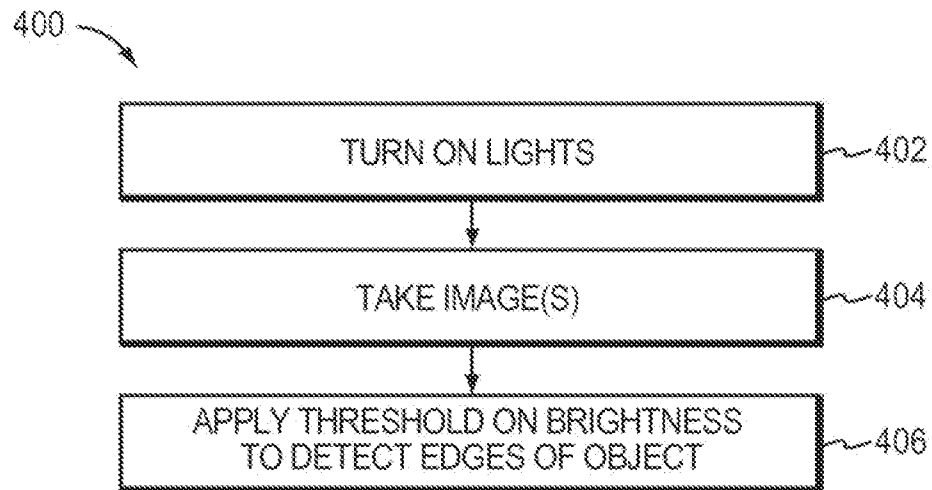


FIG. 4

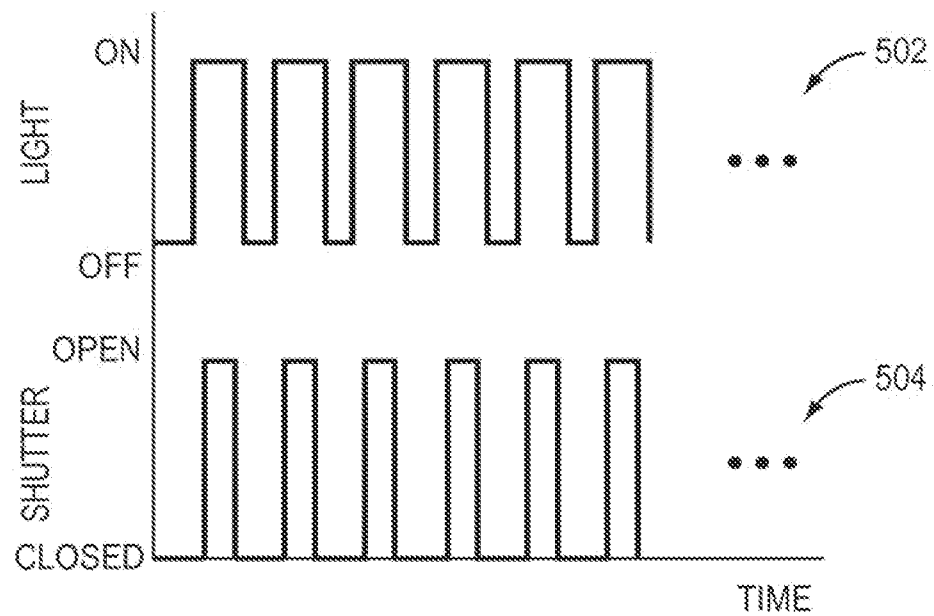


FIG. 5

5/9

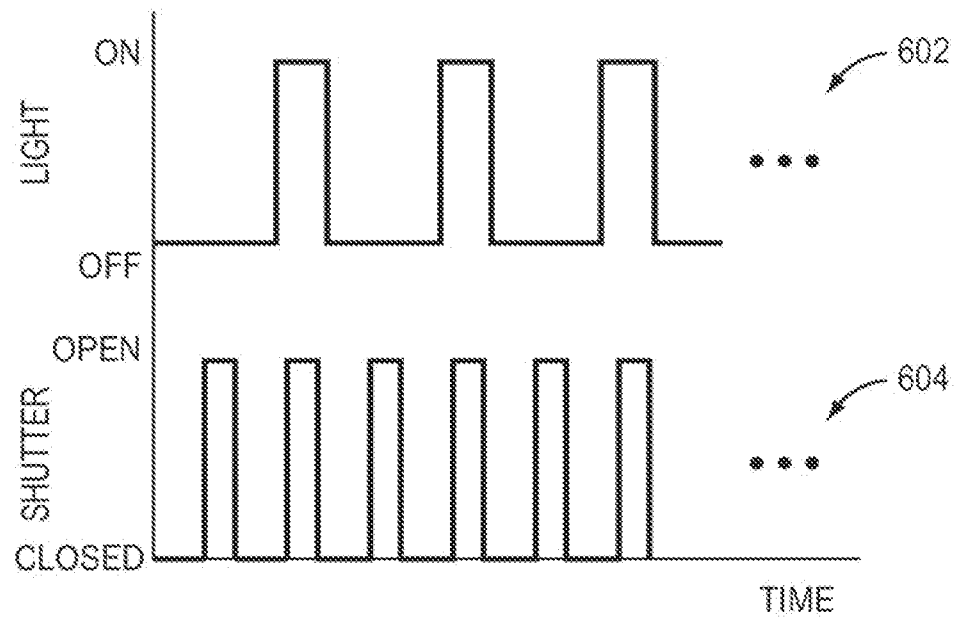


FIG. 6

6/9

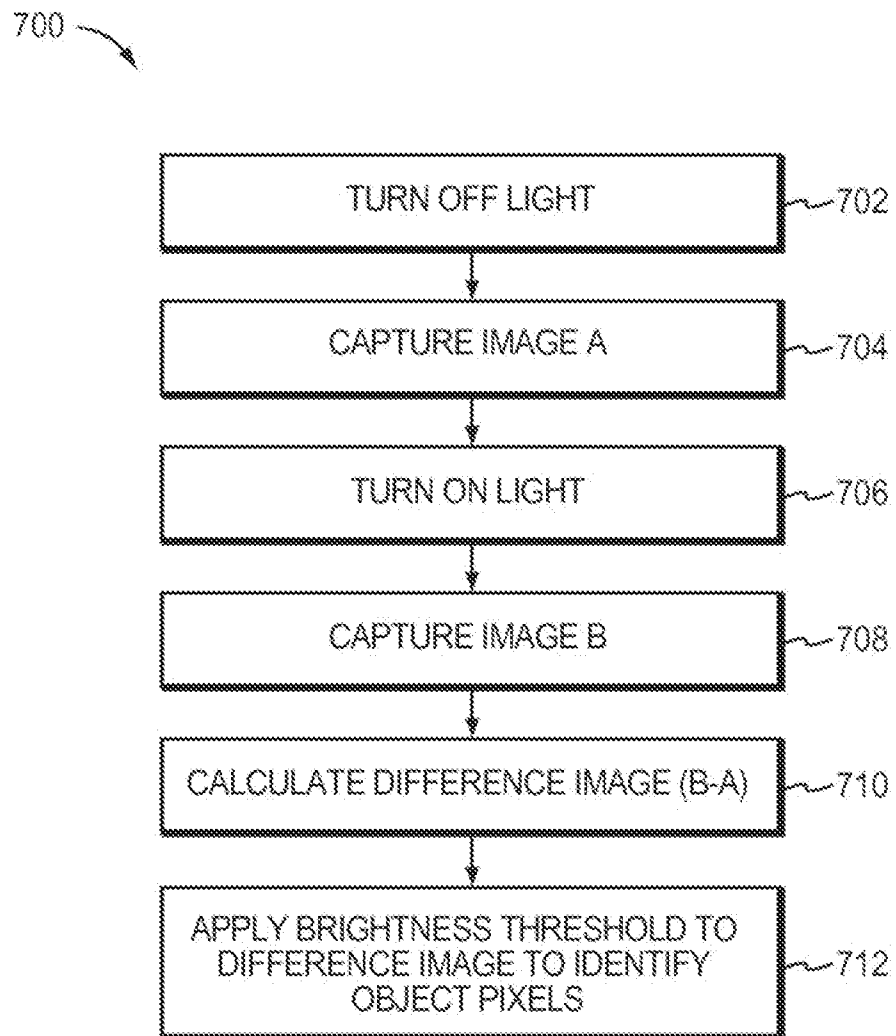


FIG. 7

7/9

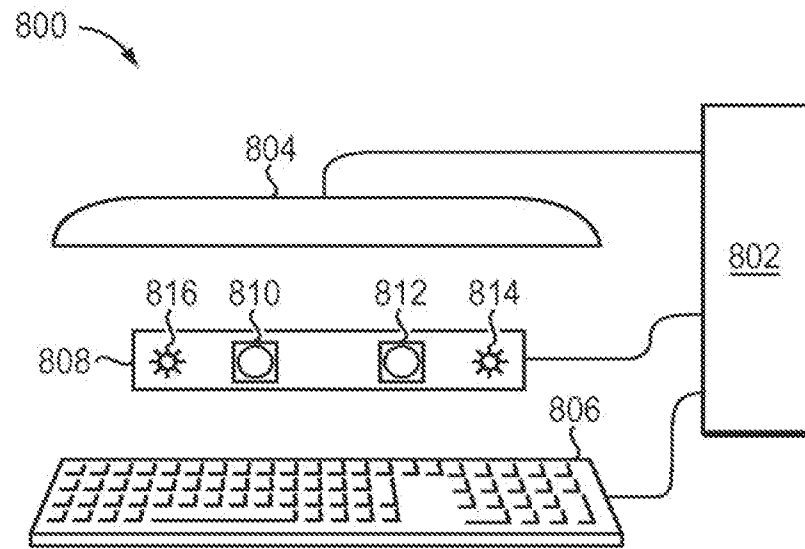


FIG. 8

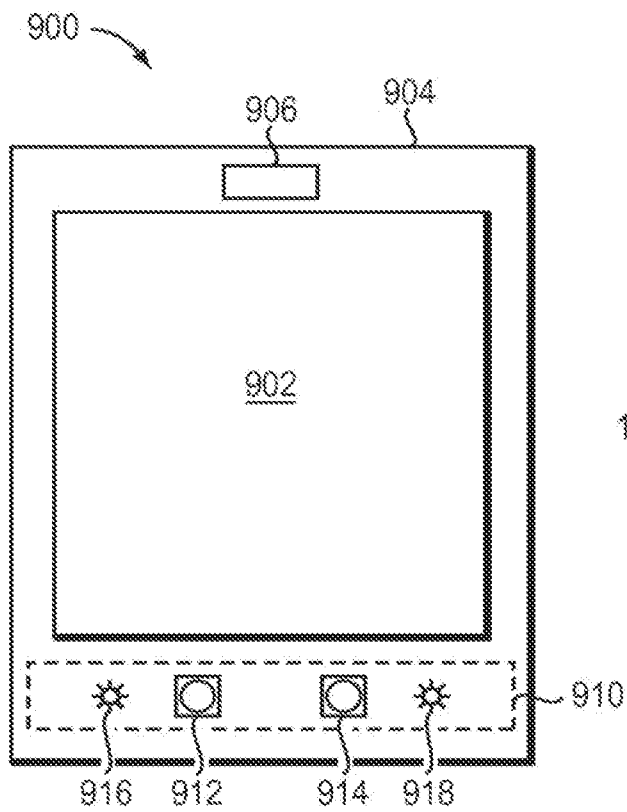


FIG. 9

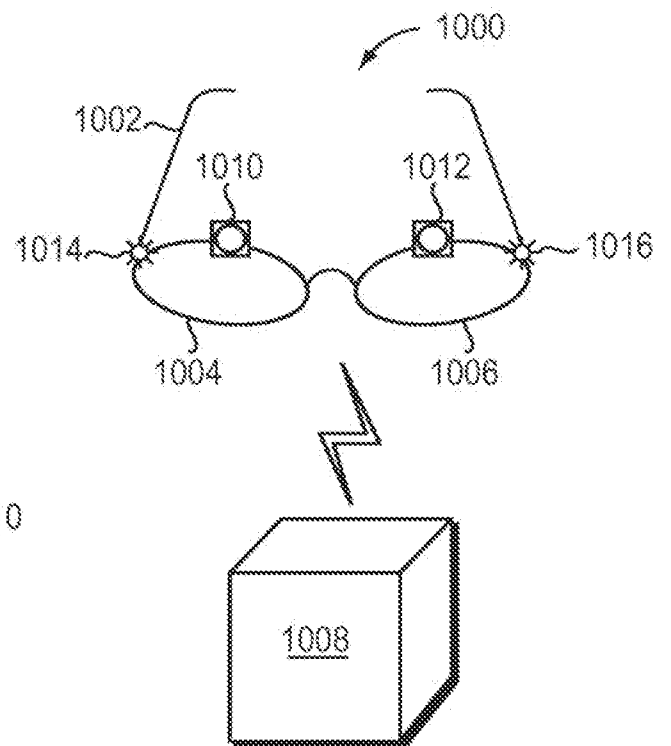


FIG. 10

8/9

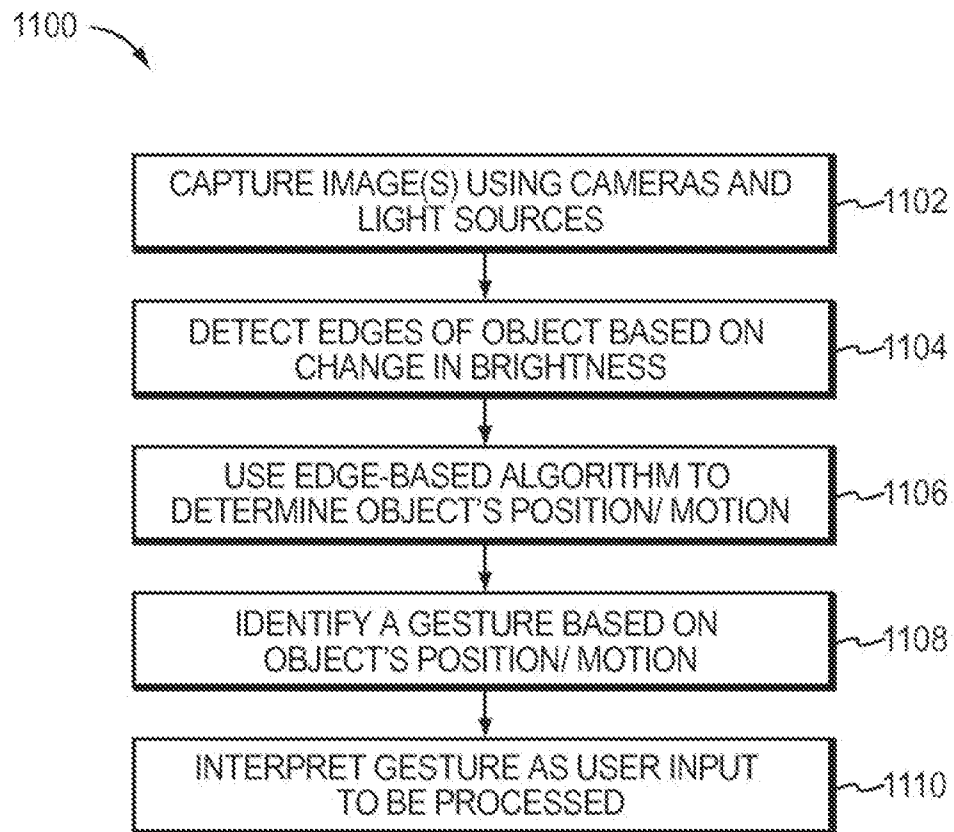


FIG. 11

9/9

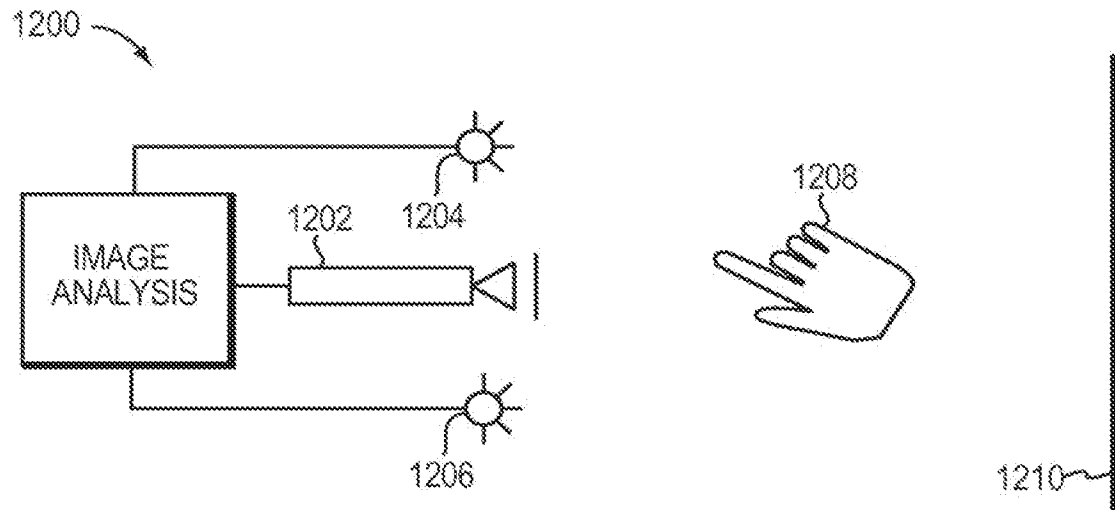


FIG. 12

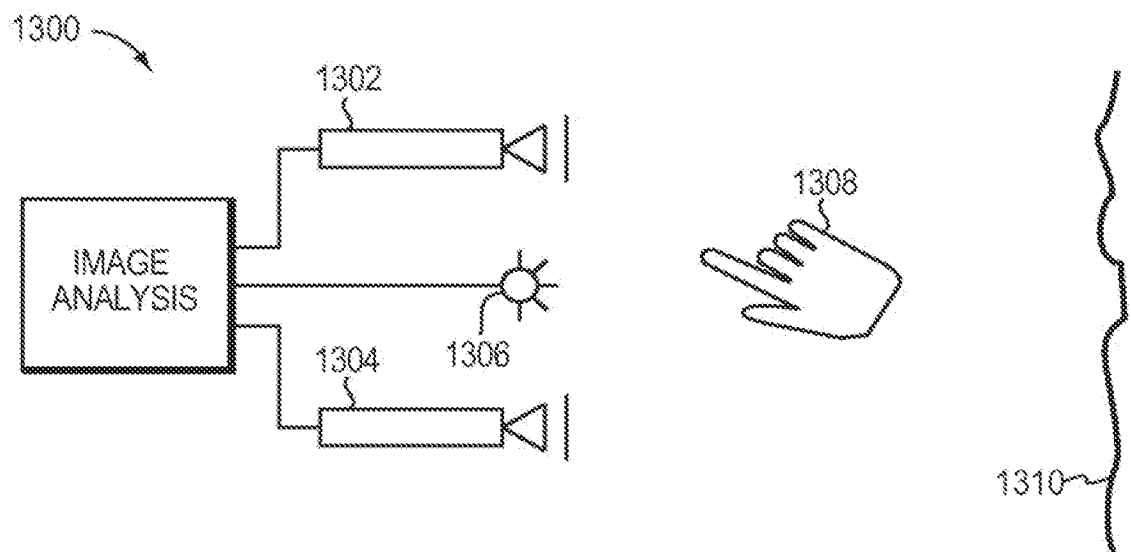


FIG. 13

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US2013/069231**A. CLASSIFICATION OF SUBJECT MATTER****G06T 7/20(2006.01)i, H05B 37/02(2006.01)i**

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)

G06T 7/20; G06F 3/048; G09G 5/08; H04N 5/225; H04N 15/00; G06K 9/34; G09G 5/00; H05B 37/02

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

Korean utility models and applications for utility models

Japanese utility models and applications for utility models

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

eKOMPASS(KIPO internal) & keywords: motion, light, image, camera, bright

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	US 2008-0244468 A1 (H. KEITH NISHIHARA et al.) 02 October 2008 See abstract; paragraphs [0017]–[0033]; claim 1; and figures 1-3.	21,32
A		1-20, 22-31, 33-40
Y	US 2011-0310220 A1 (SCOTT MCELLOWNEY) 22 December 2011 See abstract; paragraphs [0023]–[0063]; and figures 1-6.	21,32
A	US 2009-0116742 A1 (H. KEITH NISHIHARA) 07 May 2009 See paragraphs [0017]–[0029]; claim 1; and figures 1-3.	1-40
A	US 2011-0080490 A1 (IAN CLARKSON et al.) 07 April 2011 See abstract; paragraphs [0027]–[0046]; claim 1; and figures 1-2.	1-40
A	US 2005-0156888 A1 (TONG XIE et al.) 21 July 2005 See paragraphs [0013]–[0024]; claim 1; and figures 1-3.	1-40

☐ 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

12 March 2014 (12.03.2014)

Date of mailing of the international search report

13 March 2014 (13.03.2014)

Name and mailing address of the ISA/KR

International Application Division
Korean Intellectual Property Office
189 Cheongsa-ro, Seo-gu, Daejeon Metropolitan City, 302-701,
Republic of Korea

Facsimile No. +82-42-472-7140

Authorized officer

HWANG, Yun Koo

Telephone No. +82-42-481-5715



INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No.

PCT/US2013/069231

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
US 2008-0244468 A1	02/10/2008	DE 102009023875 A1 EP 1879130 A2 EP 1879130 A3 EP 2645303 A2 EP 2645303 A3 GB 0909597 D0 GB 2460542 A GB 2460542 B US 2008-0013826 A1 US 8180114 B2 US 8589824 B2	04/02/2010 16/01/2008 15/12/2010 02/10/2013 04/12/2013 15/07/2009 09/12/2009 01/12/2010 17/01/2008 15/05/2012 19/11/2013
US 2011-0310220 A1	22/12/2011	CN 102231037 A	02/11/2011
US 2009-0116742 A1	07/05/2009	EP 2068230 A2 EP 2068230 A3 EP 2068230 B1 US 8139110 B2	10/06/2009 22/02/2012 12/12/2012 20/03/2012
US 2011-0080490 A1	07/04/2011	US 8515128 B1 US 8547327 B2	20/08/2013 01/10/2013
US 2005-0156888 A1	21/07/2005	CN 1910647 A CN 1910647 B CN 1910647 C0 EP 1706862 A2 EP 1706862 A4 JP 2007-524158 A KR 10-1099820 B1 KR 102007-0007269 A US 7184022 B2 WO 2005-072500 A2 WO 2005-072500 A3	07/02/2007 31/08/2011 07/02/2007 04/10/2006 09/04/2008 23/08/2007 27/12/2011 15/01/2007 27/02/2007 11/08/2005 06/07/2006