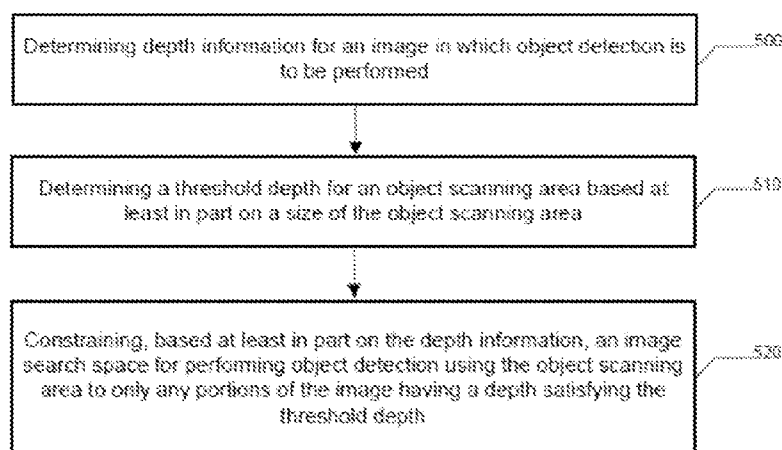




- (51) International Patent Classification:
G06K 9/00 (2006.01) *H04N 5/232* (2006.01)
G06T 7/00 (2006.01)
- (21) International Application Number: PCT/FI2011/050976
- (22) International Filing Date:
4 November 2011 (04.11.2011)
- (25) Filing Language: English
- (26) Publication Language: English
- (30) Priority Data:
4025/CHE/2010 30 December 2010 (30.12.2010) IN
- (71) Applicant (for all designated States except US): **NOKIA CORPORATION** [FI/FI]; Keilalahdentie 4, FI-02150 Espoo (FI).
- (72) Inventors; and
- (75) Inventors/Applicants (for US only): **MISHRA, Pranav** [IN/IN]; RA 402, Purva Riviera, Whitefield Road, Bangalore 560037 (IN). **GOVINDARAO, Krishna** [IN/IN]; 663, 5th Cross, 2nd Block, Banashankari 1st Stage, Bangalore 560050 (IN). **PUTRAYA, Gururaj** [IN/IN]; Flat # 115, Oceanus Ebony, Bellandur, Bangalore 560103 (IN).
- (74) Agents: **NOKIA CORPORATION** et al.; IPR Department, Ari Aarnio, Keilalahdentie 4, FI-02150 Espoo (FI).
- (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, 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, IS, JP, KE, KG, KM, 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, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, 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, MD, 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, ML, MR, NE, SN, TD, TG).

[Continued on next page]

(54) Title: METHODS AND APPARATUSES FOR PERFORMING OBJECT DETECTION



(57) Abstract: Methods and apparatuses are provided for performing object detection. A method may include determining depth information for an image in which object detection is to be performed. The method may further include constraining, based at least in part on the depth information and a defined size of an object scanning area, an image search space for performing object detection in the image using the object scanning area. Corresponding apparatuses are also provided.

FIG. 5



Published:

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

METHODS AND APPARATUSES FOR PERFORMING OBJECT DETECTION

TECHNOLOGICAL FIELD

[0001] Example embodiments of the present invention relate generally to object
5 detection technology and and, more particularly, relate to methods and apparatuses for
performing object detection.

BACKGROUND

[0002] Object detection, including, for example, face detection, is finding an increasing
10 number of uses and applications. The increase in potential applications for facial analyses
has partly occurred as a result of the continuously increasing speed and capabilities of
modern microprocessors. As a result, face detection can be used in a number of settings
for various applications including biometrics, user interface applications, gaming
application, social networking and other interpersonal communications applications. The
15 advancement in computing power of microprocessors has also made object detection
functionality available on mobile devices, such as cell phones and other smart devices.

[0003] Although object detection techniques continue to improve, many current
methods require either a high computation capability or suffer from limited object
detection performance. As such, performing object detection in an image using current
20 methods may impose a significant burden on computing resources and may be relatively
time consuming due to the computational complexity of many object detection techniques.

BRIEF SUMMARY

[0004] Methods, apparatuses, and computer program products are herein provided for
25 performing object detection. Methods, apparatuses, and computer program products in
accordance with various embodiments may provide several advantages to computing
devices and computing device users. Some example embodiments provide for constraining
an image search space for performing object detection in an image. More particularly,
some example embodiments use a relationship between object size and depth to constrain
30 the image search space. In this regard, given an object size, some example embodiments
may determine a threshold depth for the object size and may constrain the image search
space for the object size to only portions of an image having a depth that satisfies the
threshold depth. Accordingly, some example embodiments may utilize depth information

for an image to reduce the time required for performing object detection in an image by limiting the areas of the image searched for a given object size based at least in part on a relationship between object size and depth.

[0005] In a first example embodiment, a method is provided, which comprises

5 determining depth information for an image in which object detection is to be performed. The method of this example embodiment further comprises constraining, based at least in part on the depth information and a defined size of an object scanning area, an image search space for performing object detection in the image using the object scanning area.

[0006] In another example embodiment, an apparatus comprising at least one processor

10 and at least one memory storing computer program code is provided. The at least one memory and stored computer program code are configured, with the at least one processor, to cause the apparatus of this example embodiment to at least determine depth information for an image in which object detection is to be performed. The at least one memory and stored computer program code are configured, with the at least one processor, to further
15 cause the apparatus of this example embodiment to constrain, based at least in part on the depth information and a defined size of an object scanning area, an image search space for performing object detection in the image using the object scanning area.

[0007] In another example embodiment, a computer program product is provided. The

computer program product of this example embodiment includes at least one computer-
20 readable storage medium having computer-readable program instructions stored therein.

The program instructions of this example embodiment comprise program instructions configured to determine depth information for an image in which object detection is to be performed. The program instructions of this example embodiment further comprise
program instructions configured to constrain, based at least in part on the depth

25 information and a defined size of an object scanning area, an image search space for performing object detection in the image using the object scanning area.

[0008] In another example embodiment, an apparatus is provided that comprises means for determining depth information for an image in which object detection is to be

performed. The apparatus of this example embodiment further comprises means for

30 constraining, based at least in part on the depth information and a defined size of an object scanning area, an image search space for performing object detection in the image using the object scanning area.

[0009] The above summary is provided merely for purposes of summarizing some example embodiments of the invention so as to provide a basic understanding of some aspects of the invention. Accordingly, it will be appreciated that the above described example embodiments are merely examples and should not be construed to narrow the scope or spirit of the invention in any way. It will be appreciated that the scope of the invention encompasses many potential embodiments, some of which will be further described below, in addition to those here summarized.

BRIEF DESCRIPTION OF THE DRAWINGS

[0010] Having thus described embodiments of the invention in general terms, reference will now be made to the accompanying drawings, which are not necessarily drawn to scale, and wherein:

[0011] FIG. 1 illustrates a block diagram of an object detection apparatus according to an example embodiment;

[0012] FIG. 2 is a schematic block diagram of a mobile terminal according to an example embodiment;

[0013] FIG. 3 illustrates segmentation of an image according to an example embodiment;

[0014] FIG. 4 illustrates a flowchart according to an example method for performing object detection according to an example embodiment;

[0015] FIG. 5 illustrates a flowchart according to an example method for performing object detection according to an example embodiment; and

[0016] FIG. 6 illustrates a flowchart according to an example method for performing object detection according to an example embodiment.

DETAILED DESCRIPTION

[0017] Some embodiments of the present invention will now be described more fully hereinafter with reference to the accompanying drawings, in which some, but not all embodiments of the invention are shown. Indeed, the invention may be embodied in many different forms and should not be construed as limited to the embodiments set forth herein; rather, these embodiments are provided so that this disclosure will satisfy applicable legal requirements. Like reference numerals refer to like elements throughout.

[0018] As used herein, the terms “data,” “content,” “information” and similar terms may be used interchangeably to refer to data capable of being transmitted, received, displayed and/or stored in accordance with various example embodiments. Thus, use of any such terms should not be taken to limit the spirit and scope of the disclosure. Further, where a computing device is described herein to receive data from another computing device, it will be appreciated that the data may be received directly from the another computing device or may be received indirectly via one or more intermediary computing devices, such as, for example, one or more servers, relays, routers, network access points, base stations, and/or the like.

[0019] The term “computer-readable medium” as used herein refers to any medium configured to participate in providing information to a processor, including instructions for execution. Such a medium may take many forms, including, but not limited to a non-transitory computer-readable storage medium (e.g., non-volatile media, volatile media), and transmission media. Transmission media include, for example, coaxial cables, copper wire, fiber optic cables, and carrier waves that travel through space without wires or cables, such as acoustic waves and electromagnetic waves, including radio, optical and infrared waves. Signals include man-made transient variations in amplitude, frequency, phase, polarization or other physical properties transmitted through the transmission media. Examples of computer-readable media include a floppy disk, a flexible disk, hard disk, magnetic tape, any other magnetic medium, a compact disc read only memory (CD-ROM), compact disc compact disc-rewritable (CD-RW), digital versatile disc (DVD), Blu-Ray, any other optical medium, punch cards, paper tape, optical mark sheets, any other physical medium with patterns of holes or other optically recognizable indicia, a random access memory (RAM), a programmable read only memory (PROM), an erasable programmable read only memory (EPROM), a FLASH-EPROM, any other memory chip or cartridge, a carrier wave, or any other medium from which a computer can read. The term computer-readable storage medium is used herein to refer to any computer-readable medium except transmission media. However, it will be appreciated that where embodiments are described to use a computer-readable storage medium, other types of computer-readable mediums may be substituted for or used in addition to the computer-readable storage medium in alternative embodiments.

[0020] Additionally, as used herein, the term ‘circuitry’ refers to (a) hardware-only circuit implementations (e.g., implementations in analog circuitry and/or digital circuitry);

(b) combinations of circuits and computer program product(s) comprising software and/or firmware instructions stored on one or more computer readable memories that work together to cause an apparatus to perform one or more functions described herein; and (c) circuits, such as, for example, a microprocessor(s) or a portion of a microprocessor(s), that require software or firmware for operation even if the software or firmware is not physically present. This definition of 'circuitry' applies to all uses of this term herein, including in any claims. As a further example, as used herein, the term 'circuitry' also includes an implementation comprising one or more processors and/or portion(s) thereof and accompanying software and/or firmware. As another example, the term 'circuitry' as used herein also includes, for example, a baseband integrated circuit or applications processor integrated circuit for a mobile phone or a similar integrated circuit in a server, a cellular network device, other network device, and/or other computing device.

[0021] Object detection may generally be performed by scanning a rectangular unit area through an image in search for an object. The rectangular unit area may, for example, be represented by a template or a model. For example, in face detection a 20x20 face box may be scanned through an entire image at every pixel in the image. Then a 25x25 face box may be scanned through the entire image at every pixel. This process may be continued up to an MxM face box size. Accordingly, face detection may be performed based on a series of templates or models of varying sizes. Assuming P such face box sizes are evaluated, the total number of points computed may be expressed as $O(N^2)*P$, where N is the total number of pixels in an image at which the face box is placed to detect whether or not a face is present. As such, conventional object detection may be quite time consuming and may be computationally intensive.

[0022] Accordingly, some example embodiments provided herein may enhance object detection by constraining an image search space for performing object detection in an image using a given object scanning area (e.g., a template or model). In this regard, rather than scanning the entire image using each object scanning area used for performing object detection, some example embodiments may leverage depth information for an image to constrain the search space for a given object scanning area size (e.g., the search space for an object of a given size). As such, some example embodiments may improve computational efficiency and reduce the time required for performing object detection in images carrying depth information, such as three-dimensional (3-D) images.

[0023] Example embodiments, which constrain the image search space may provide several advantages. For example, if an image is of a landscape scenery where most objects are at a large distance (e.g., a large depth) from the camera capturing, example embodiments may skip performing face detection due to the depth of the image. In this regard, it may be considered that given the depth of the image, any faces that are present will be insignificant and small. Further, as some example embodiments may reduce the search space for a given object size and/or a number of sizes of objects for which detection is performed, the number of false positives that might otherwise result may be reduced. Additionally, in some example embodiments wherein object detection is implemented on a mobile device, battery life may be increased.

[0024] FIG. 1 illustrates a block diagram of an object detection apparatus 102 for performing object detection according to an example embodiment. It will be appreciated that the object detection apparatus 102 is provided as an example of some embodiments and should not be construed to narrow the scope or spirit of the invention in any way. In this regard, the scope of the disclosure encompasses many potential embodiments in addition to those illustrated and described herein. As such, while FIG. 1 illustrates one example of a configuration of an apparatus for performing object detection, other configurations may also be used to implement embodiments of the present invention.

[0025] The object detection apparatus 102 may be embodied as a desktop computer, laptop computer, mobile terminal, mobile computer, mobile phone, mobile communication device, one or more servers, one or more network nodes, game device, digital camera/camcorder, audio/video player, television device, radio receiver, digital video recorder, positioning device, chipset, a computing device comprising a chipset, any combination thereof, and/or the like. In this regard, the object detection apparatus 102 may comprise any computing device or other apparatus that is configured to perform object detection in accordance with one or more example embodiments disclosed herein. In an example embodiment, the object detection apparatus 102 is embodied as a mobile computing device, such as a mobile terminal, such as that illustrated in FIG. 2.

[0026] In this regard, FIG. 2 illustrates a block diagram of a mobile terminal 10 representative of one embodiment of an object detection apparatus 102. It should be understood, however, that the mobile terminal 10 illustrated and hereinafter described is merely illustrative of one type of object detection apparatus 102 that may implement and/or benefit from various embodiments of the invention and, therefore, should not be

taken to limit the scope of the disclosure. While several embodiments of the electronic device are illustrated and will be hereinafter described for purposes of example, other types of electronic devices, such as mobile telephones, mobile computers, portable digital assistants (PDAs), pagers, laptop computers, desktop computers, gaming devices, 5 televisions, and other types of electronic systems, may employ various embodiments of the invention.

[0027] As shown, the mobile terminal 10 may include an antenna 12 (or multiple antennas 12) in communication with a transmitter 14 and a receiver 16. The mobile terminal 10 may also include a processor 20 configured to provide signals to and receive 10 signals from the transmitter and receiver, respectively. The processor 20 may, for example, be embodied as various means including circuitry, one or more microprocessors with accompanying digital signal processor(s), one or more processor(s) without an accompanying digital signal processor, one or more coprocessors, one or more multi-core processors, one or more controllers, processing circuitry, one or more computers, various 15 other processing elements including integrated circuits such as, for example, an ASIC (application specific integrated circuit) or FPGA (field programmable gate array), or some combination thereof. Accordingly, although illustrated in FIG. 2 as a single processor, in some embodiments the processor 20 comprises a plurality of processors. These signals sent and received by the processor 20 may include signaling information in accordance 20 with an air interface standard of an applicable cellular system, and/or any number of different wireline or wireless networking techniques, comprising but not limited to Wi-Fi, wireless local access network (WLAN) techniques such as Institute of Electrical and Electronics Engineers (IEEE) 802.11, 802.16, and/or the like. In addition, these signals may include speech data, user generated data, user requested data, and/or the like. In this 25 regard, the mobile terminal may be capable of operating with one or more air interface standards, communication protocols, modulation types, access types, and/or the like. More particularly, the mobile terminal may be capable of operating in accordance with various first generation (1G), second generation (2G), 2.5G, third-generation (3G) communication protocols, fourth-generation (4G) communication protocols, Internet Protocol Multimedia 30 Subsystem (IMS) communication protocols (e.g., session initiation protocol (SIP)), and/or the like. For example, the mobile terminal may be capable of operating in accordance with 2G wireless communication protocols IS-136 (Time Division Multiple Access (TDMA)), Global System for Mobile communications (GSM), IS-95 (Code Division Multiple Access

(CDMA)), and/or the like. Also, for example, the mobile terminal may be capable of operating in accordance with 2.5G wireless communication protocols General Packet Radio Service (GPRS), Enhanced Data GSM Environment (EDGE), and/or the like.

Further, for example, the mobile terminal may be capable of operating in accordance with

5 3G wireless communication protocols such as Universal Mobile Telecommunications System (UMTS), Code Division Multiple Access 2000 (CDMA2000), Wideband Code Division Multiple Access (WCDMA), Time Division-Synchronous Code Division Multiple Access (TD-SCDMA), and/or the like. The mobile terminal may be additionally capable of operating in accordance with 3.9G wireless communication protocols such as
10 Long Term Evolution (LTE) or Evolved Universal Terrestrial Radio Access Network (E-UTRAN) and/or the like. Additionally, for example, the mobile terminal may be capable of operating in accordance with fourth-generation (4G) wireless communication protocols and/or the like as well as similar wireless communication protocols that may be developed in the future.

15 **[0028]** Some Narrow-band Advanced Mobile Phone System (NAMPS), as well as Total Access Communication System (TACS), mobile terminals may also benefit from embodiments of this invention, as should dual or higher mode phones (e.g., digital/analog or TDMA/CDMA/analog phones). Additionally, the mobile terminal 10 may be capable of operating according to Wi-Fi or Worldwide Interoperability for Microwave Access
20 (WiMAX) protocols.

[0029] It is understood that the processor 20 may comprise circuitry for implementing audio/video and logic functions of the mobile terminal 10. For example, the processor 20 may comprise a digital signal processor device, a microprocessor device, an analog-to-digital converter, a digital-to-analog converter, and/or the like. Control and signal
25 processing functions of the mobile terminal may be allocated between these devices according to their respective capabilities. The processor may additionally comprise an internal voice coder (VC) 20a, an internal data modem (DM) 20b, and/or the like. Further, the processor may comprise functionality to operate one or more software programs, which may be stored in memory. For example, the processor 20 may be capable of operating a
30 connectivity program, such as a web browser. The connectivity program may allow the mobile terminal 10 to transmit and receive web content, such as location-based content, according to a protocol, such as Wireless Application Protocol (WAP), hypertext transfer protocol (HTTP), and/or the like. The mobile terminal 10 may be capable of using a

Transmission Control Protocol/Internet Protocol (TCP/IP) to transmit and receive web content across the internet or other networks.

[0030] The mobile terminal 10 may also comprise a user interface including, for example, an earphone or speaker 24, a ringer 22, a microphone 26, a display 28, a user input interface, and/or the like, which may be operationally coupled to the processor 20. In this regard, the processor 20 may comprise user interface circuitry configured to control at least some functions of one or more elements of the user interface, such as, for example, the speaker 24, the ringer 22, the microphone 26, the display 28, and/or the like. The processor 20 and/or user interface circuitry comprising the processor 20 may be configured to control one or more functions of one or more elements of the user interface through computer program instructions (e.g., software and/or firmware) stored on a memory accessible to the processor 20 (e.g., volatile memory 40, non-volatile memory 42, and/or the like). Although not shown, the mobile terminal may comprise a battery for powering various circuits related to the mobile terminal, for example, a circuit to provide mechanical vibration as a detectable output. The display 28 of the mobile terminal may be of any type appropriate for the electronic device in question with some examples including a plasma display panel (PDP), a liquid crystal display (LCD), a light-emitting diode (LED), an organic light-emitting diode display (OLED), a projector, a holographic display or the like. The user input interface may comprise devices allowing the mobile terminal to receive data, such as a keypad 30, a touch display (not shown), a joystick (not shown), and/or other input device. In embodiments including a keypad, the keypad may comprise numeric (0-9) and related keys (#, *), and/or other keys for operating the mobile terminal.

[0031] In some example embodiments, the mobile terminal 10 may include a media capturing element, such as a camera, video and/or audio module, in communication with the processor 20. The media capturing element may be any means for capturing an image, video and/or audio for storage, display or transmission. For example, in an example embodiment in which the media capturing element is a camera module 36, the camera module 36 may include a digital camera capable of forming a digital image file from a captured image. In addition, the digital camera of the camera module 36 may be capable of capturing a video clip. As such, the camera module 36 may include all hardware, such as a lens or other optical component(s), and software necessary for creating a digital image file from a captured image as well as a digital video file from a captured video clip. Alternatively, the camera module 36 may include only the hardware needed to view an

image, while a memory device of the mobile terminal 10 stores instructions for execution by the processor 20 in the form of software necessary to create a digital image file from a captured image. As yet another alternative, an object or objects within a field of view of the camera module 36 may be displayed on the display 28 of the mobile terminal 10 to illustrate a view of an image currently displayed which may be captured if desired by the user. As such, as referred to hereinafter, an image may be either a captured image or an image comprising the object or objects currently displayed by the mobile terminal 10, but not necessarily captured in an image file. In some example embodiments, the camera module 36 may comprise a 3-D camera configured to capture 3-D images. Additionally or alternatively, the camera module 36 may be configured to capture depth information for images (e.g., two-dimensional images having depth information about a distance to objects captured in the images). In some example embodiments, the camera module 36 may be configured to compute depth based on two or more images captured in quick succession (e.g., two images captured with a little shift between the two images). In an example embodiment, the camera module 36 may further include a processing element such as a co-processor which assists the processor 20 in processing image data and an encoder and/or decoder for compressing and/or decompressing image data. The encoder and/or decoder may encode and/or decode according to, for example, a joint photographic experts group (JPEG) standard, a moving picture experts group (MPEG) standard, or other format.

[0032] As shown in FIG. 2, the mobile terminal 10 may also include one or more means for sharing and/or obtaining data. For example, the mobile terminal may comprise a short-range radio frequency (RF) transceiver and/or interrogator 64 so data may be shared with and/or obtained from electronic devices in accordance with RF techniques. The mobile terminal may comprise other short-range transceivers, such as, for example, an infrared (IR) transceiver 66, a BluetoothTM (BT) transceiver 68 operating using BluetoothTM brand wireless technology developed by the BluetoothTM Special Interest Group, a wireless universal serial bus (USB) transceiver 70 and/or the like. The BluetoothTM transceiver 68 may be capable of operating according to ultra-low power BluetoothTM technology (e.g., WibreeTM) radio standards. In this regard, the mobile terminal 10 and, in particular, the short-range transceiver may be capable of transmitting data to and/or receiving data from electronic devices within a proximity of the mobile terminal, such as within 10 meters, for example. Although not shown, the mobile terminal may be capable of transmitting and/or receiving data from electronic devices according to

various wireless networking techniques, including Wi-Fi, WLAN techniques such as IEEE 802.11 techniques, IEEE 802.15 techniques, IEEE 802.16 techniques, and/or the like.

[0033] The mobile terminal 10 may comprise memory, such as a subscriber identity module (SIM) 38, a removable user identity module (R-UIM), and/or the like, which may store information elements related to a mobile subscriber. In addition to the SIM, the mobile terminal may comprise other removable and/or fixed memory. The mobile terminal 10 may include volatile memory 40 and/or non-volatile memory 42. For example, volatile memory 40 may include Random Access Memory (RAM) including dynamic and/or static RAM, on-chip or off-chip cache memory, and/or the like. Non-volatile memory 42, which may be embedded and/or removable, may include, for example, read-only memory, flash memory, magnetic storage devices (e.g., hard disks, floppy disk drives, magnetic tape, etc.), optical disc drives and/or media, non-volatile random access memory (NVRAM), and/or the like. Like volatile memory 40 non-volatile memory 42 may include a cache area for temporary storage of data. One or more of the volatile memory 40 or non-volatile memory 42 may be embodied as a tangible, non-transitory memory. The memories may store one or more software programs, instructions, pieces of information, data, and/or the like which may be used by the mobile terminal for performing functions of the mobile terminal. For example, the memories may comprise an identifier, such as an international mobile equipment identification (IMEI) code, capable of uniquely identifying the mobile terminal 10.

[0034] Returning to FIG. 1, in an example embodiment, the object detection apparatus 102 includes various means for performing the various functions herein described. These means may comprise one or more of a processor 110, memory 112, communication interface 114, user interface 116, or object detection circuitry 118. The means of the object detection apparatus 102 as described herein may be embodied as, for example, circuitry, hardware elements (e.g., a suitably programmed processor, combinational logic circuit, and/or the like), a computer program product comprising computer-readable program instructions (e.g., software or firmware) stored on a computer-readable medium (e.g. memory 112) that is executable by a suitably configured processing device (e.g., the processor 110), or some combination thereof.

[0035] In some example embodiments, one or more of the means illustrated in FIG. 1 may be embodied as a chip or chip set. In other words, the object detection apparatus 102 may comprise one or more physical packages (e.g., chips) including materials, components

and/or wires on a structural assembly (e.g., a baseboard). The structural assembly may provide physical strength, conservation of size, and/or limitation of electrical interaction for component circuitry included thereon. In this regard, the processor 110, memory 112, communication interface 114, user interface 116, and/or object detection circuitry 118 may be embodied as a chip or chip set. The object detection apparatus 102 may therefore, in some cases, be configured to or may comprise component(s) configured to implement embodiments of the present invention on a single chip or as a single “system on a chip.” As such, in some cases, a chip or chipset may constitute means for performing one or more operations for providing the functionalities described herein and/or for enabling user interface navigation with respect to the functionalities and/or services described herein.

[0036] The processor 110 may, for example, be embodied as various means including one or more microprocessors with accompanying digital signal processor(s), one or more processor(s) without an accompanying digital signal processor, one or more coprocessors, one or more multi-core processors, one or more controllers, processing circuitry, one or more computers, various other processing elements including integrated circuits such as, for example, an ASIC (application specific integrated circuit) or FPGA (field programmable gate array), one or more other types of hardware processors, or some combination thereof. Accordingly, although illustrated in FIG. 1 as a single processor, in some embodiments the processor 110 comprises a plurality of processors. The plurality of processors may be in operative communication with each other and may be collectively configured to perform one or more functionalities of the object detection apparatus 102 as described herein. The plurality of processors may be embodied on a single computing device or distributed across a plurality of computing devices collectively configured to function as the object detection apparatus 102. In embodiments wherein the object detection apparatus 102 is embodied as a mobile terminal 10, the processor 110 may be embodied as or comprise the processor 20. In some example embodiments, the processor 110 is configured to execute instructions stored in the memory 112 or otherwise accessible to the processor 110. These instructions, when executed by the processor 110, may cause the object detection apparatus 102 to perform one or more of the functionalities of the object detection apparatus 102 as described herein. As such, whether configured by hardware or software methods, or by a combination thereof, the processor 110 may comprise an entity capable of performing operations according to one or more example embodiments while configured accordingly. Thus, for example, when the processor 110 is

embodied as an ASIC, FPGA or the like, the processor 110 may comprise specifically configured hardware for conducting one or more operations described herein.

Alternatively, as another example, when the processor 110 is embodied as an executor of instructions, such as may be stored in the memory 112, the instructions may specifically
5 configure the processor 110 to perform one or more algorithms and operations described herein.

[0037] The memory 112 may comprise, for example, volatile memory, non-volatile memory, or some combination thereof. In this regard, the memory 112 may comprise a non-transitory computer-readable storage medium. Although illustrated in FIG. 1 as a
10 single memory, the memory 112 may comprise a plurality of memories. The plurality of memories may be embodied on a single computing device or may be distributed across a plurality of computing devices collectively configured to function as the object detection apparatus 102. In various example embodiments, the memory 112 may comprise a hard
15 disk, random access memory, cache memory, flash memory, a compact disc read only memory (CD-ROM), digital versatile disc read only memory (DVD-ROM), an optical disc, circuitry configured to store information, or some combination thereof. In embodiments wherein the object detection apparatus 102 is embodied as a mobile terminal 10, the
memory 112 may comprise the volatile memory 40 and/or the non-volatile memory 42. The memory 112 may be configured to store information, data, applications, instructions,
20 or the like for enabling the object detection apparatus 102 to carry out various functions in accordance with various example embodiments. For example, in some example embodiments, the memory 112 is configured to buffer input data for processing by the processor 110. Additionally or alternatively, the memory 112 may be configured to store
program instructions for execution by the processor 110. The memory 112 may store
25 information in the form of static and/or dynamic information. The stored information may include, for example, images, information correlating object size to depth, and/or the like. This stored information may be stored and/or used by the object detection circuitry 118 during the course of performing its functionalities.

[0038] The communication interface 114 may be embodied as any device or means
30 embodied in circuitry, hardware, a computer program product comprising computer readable program instructions stored on a computer readable medium (e.g., the memory 112) and executed by a processing device (e.g., the processor 110), or a combination thereof that is configured to receive and/or transmit data from/to another computing

device. In an example embodiment, the communication interface 114 is at least partially embodied as or otherwise controlled by the processor 110. In this regard, the communication interface 114 may be in communication with the processor 110, such as via a bus. The communication interface 114 may include, for example, an antenna, a transmitter, a receiver, a transceiver and/or supporting hardware or software for enabling communications with one or more remote computing devices. The communication interface 114 may be configured to receive and/or transmit data using any protocol that may be used for communications between computing devices. In this regard, the communication interface 114 may be configured to receive and/or transmit data using any protocol that may be used for transmission of data over a wireless network, wireline network, some combination thereof, or the like by which the object detection apparatus 102 and one or more computing devices may be in communication. The communication interface 114 may additionally be in communication with the memory 112, user interface 116, and/or object detection circuitry 118, such as via a bus.

[0039] The user interface 116 may be in communication with the processor 110 to receive an indication of a user input and/or to provide an audible, visual, mechanical, or other output to a user. As such, the user interface 116 may include, for example, a keyboard, a mouse, a joystick, a display, a touch screen display, a microphone, a speaker, and/or other input/output mechanisms. In embodiments wherein the user interface 116 comprises or is in communication with a display, the display may comprise, for example, a cathode ray tube (CRT) display, a plasma display panel (PDP), a liquid crystal display (LCD), a light-emitting diode (LED), an organic light-emitting diode display (OLED), a projector (e.g., a projector configured to project a display on a projection screen, wall, and/or other object), a holographic display, or the like. In embodiments wherein the user interface 116 comprises a touch screen display, the user interface 116 may additionally be configured to detect and/or receive an indication of a touch gesture or other input to the touch screen display. The user interface 116 may be in communication with the memory 112, communication interface 114, and/or object detection circuitry 118, such as via a bus.

[0040] The object detection circuitry 118 may be embodied as various means, such as circuitry, hardware, a computer program product comprising computer readable program instructions stored on a computer readable medium (e.g., the memory 112) and executed by a processing device (e.g., the processor 110), or some combination thereof and, in some embodiments, is embodied as or otherwise controlled by the processor 110. In

embodiments wherein the object detection circuitry 118 is embodied separately from the processor 110, the object detection circuitry 118 may be in communication with the processor 110. The object detection circuitry 118 may further be in communication with one or more of the memory 112, communication interface 114, or user interface 116, such as via a bus.

[0041] In some example embodiments, the object detection circuitry 118 may be configured to compute and/or access pre-computed information defining a relation between object size and depth. Object size may comprise a size of an object in an image (e.g., at a given depth). In some example embodiments, the object may comprise a face. In this regard, some example embodiments may be applied to face detection. However, example embodiments are not limited to face detection and may be applied to detection of any object. As such, where examples are described herein with respect to face detection, faces should be regarded as an example of one type of object that may be detected using various example embodiments disclosed herein. Depth may, for example, define a distance of an object captured in an image from a camera used to capture the image. In this regard, depth may be defined in terms of actual distance, relative distance, some combination thereof, or the like.

[0042] The information defining a relation between object size and depth may, for example, comprise a table correlating object size to depth. In this regard, the relationship information may define a threshold object size for a given depth. The threshold object size may be defined by a given size, T . Alternatively, the threshold object size may be defined by a minimum size, $T1$, and a maximum size, $T2$. The relationship information may additionally or alternatively define a threshold depth for a given object size. The threshold depth may be defined by a given depth value, D . Alternatively, the threshold depth may be defined by a minimum depth, $D1$, and a maximum depth, $D2$.

[0043] The thresholds for a given object size and/or for a given depth may, for example, be determined using heuristic algorithms. Additionally or alternatively, the thresholds may be determined based at least in part on statistical analysis of a large number of objects captured at various distances from a camera, where object size and depth are known. Accordingly, object size may be determined as a function of depth and depth may be determined as a function of object size. As such, the statistical analysis may be used to determine thresholds correlating object size and depth.

[0044] In some example embodiments, information defining a relation between object size and depth may be determined based at least in part on a type of device used to capture a given image. In this regard, a relation between object size and depth may vary given the type of camera, camera optics, image sensor, and/or the like used to capture the image.

5 Accordingly, in some example embodiments wherein the object detection apparatus 102 comprises a mobile terminal 10, the relationship between depth and object size may be determined based at least in part on a type of the camera module 36.

[0045] Information defining a relationship between object size and depth may, for example, be computed by the object detection circuitry 118 and stored in the memory 112
10 for use when performing object detection. As another example, the information may be obtained or otherwise accessed from a third party source, such as server accessible by the communication interface 114 via a network. In this regard, a camera manufacturer or other third party may provide pre-computed information relating object size and depth for a particular image set, camera device, or the like.

15 [0046] The object detection circuitry 118 may be configured to determine depth information for an image in which object detection is to be performed. In this regard, the object detection circuitry 118 may be configured to access depth information associated with the image. This depth information may be pre-computed or may be computed and accessed by the object detection circuitry 118. The object detection circuitry 118 may be
20 configured to determine depth information based on a pixel level depth map for the image. The pixel level depth map may be generated by a camera which captured the image or may be computed by the object detection circuitry 118 based on available depth information associated with the image. The pixel level depth map may express a depth (e.g., a distance from the camera that captured the image) of each pixel in the image. The depth map may,
25 for example, be structured as a two dimensional array having a size corresponding to the width and height of the image, in pixels. Accordingly, the depth of a pixel ($\text{Depth}(x,y)$) may be determined from the array.

[0047] In some example embodiments, the object detection circuitry 118 may use pixel level depth information to segment an image into a plurality of contiguous regions, with
30 each region having a respective depth. In this regard, some depth maps may not be very accurate at a pixel level resolution (e.g. a depth map of same dimensions as the image). However, pixels that belong to a region may be expected to have the same depth. Therefore, given a depth map, all pixels belonging to a neighborhood which have the same

depth may be grouped as one region or segment. For example, in the case of a portrait photo, the face of the person may be in the foreground while the rest of the image is comprised of the background. The segmentation based on depth may provide two segments or regions, the foreground and background.

5 **[0048]** The object detection circuitry 118 may be configured to implement any appropriate segmentation method for performing image segmentation. By way of non-limiting example, the object detection circuitry 118 may be configured to perform image segmentation using a quantization method. For example, assume the depth values in a depth map have a dynamic range of 8 bits (e.g., [0,255]). It may be desirable to divide the
10 image into relatively few regions. As such, 256 depth levels may not be needed.

Accordingly, the pixel level depth values may be divided by some number to yield a desired number of possible depth levels. For example, assuming a dynamic range of 8 depth levels is desired, the pixel level depth values may be divided by 32. Accordingly, all pixels which have depths in a similar range will be assigned the same value. For
15 example, a pixel having a depth value of 67 and a pixel having a depth value of 75 may be assigned the same value (2), since $\text{floor}(67/32) = 2$ and $\text{floor}(75/32) = 2$.

[0049] Given the dynamic range of 8 provided in this example, it may be expected to have objects at up to 8 possible distances from the camera capturing the image. While neighboring pixels belonging to a region may be expected to have the same depth value,
20 this may not be the case in reality due to noise, incorrect estimation of depth at a few locations, and/or the like. Accordingly, for example, a situation may result wherein a central pixel in a group of pixels has a depth of 7 while its neighbors have a depth of 4. As such, it may be accepted that the pixel having the depth of 7 is in error. Therefore, it may be desirable to “smooth” out occurrences of such errors. The object detection circuitry
25 118 may accordingly be configured to apply morphological filtering and/or other technique(s) to smooth such errors.

[0050] FIG. 3 illustrates an example segmentation of an image that may result from performance of the segmentation techniques described above. In this regard, FIG. 3 illustrates five regions resulting from segmentation of the image. The region 302 has a
30 depth value of 7. The region 304 has a depth value of 4. The region 306 has a depth value of 2. The region 308 has a depth value of 6. The region 310 has a depth value of 3. Accordingly, it may be seen that through segmentation, a relatively low number of regions of pixels having a common depth may be determined. It will be appreciated, however, that

FIG. 3 is provided as an abstract example. In practice, a region may not have a rectangular or other uniform shape, as illustrated in the example of Fig. 3.

[0051] When performing object detection, the object detection circuitry 118 may be configured to iteratively scan the image using a series of object scanning areas (e.g., models or templates) having various sizes. In this regard, each respective size of object scanning area may be configured for detecting a different object size. For an object scanning area having a given size, the object detection circuitry 118 may be configured to constrain the image search space for performing object detection in the image using the object scanning area based at least in part on depth information for the image and the size of the object scanning area. In this regard, for a given pixel location or region of an image, the object detection circuitry 118 may perform object detection using the object scanning area only if the depth of the pixel or region satisfies a relationship criterion with respect to the size of the object scanning area. This relationship criterion may be defined by the pre-computed information defining a relationship between object size and depth.

[0052] As an example, the object detection circuitry 118 may scan an image with an object scanning area having a size of $M*N$. The object detection circuitry 118 may determine based on the pre-computed relationship information a threshold depth for an object scanning area of size $M*N$. The object detection circuitry 118 may accordingly constrain the image search space for performing object detection using the object scanning area of size $M*N$ to only any portions of the image having a depth satisfying the threshold depth corresponding to an object scanning area of size $M*N$.

[0053] In example embodiments wherein the threshold depth is a single depth value, the object detection circuitry 118 may, for example, perform object detection at a pixel location or in a region using the object scanning area of size $M*N$ only if the depth of the pixel location or region equals the threshold depth. Alternatively, as another example, the object detection circuitry 118 may perform object detection at the pixel location or region if the depth of the pixel location or region is within a predefined tolerance range of the threshold depth. Thus, for example, if the threshold depth is 4 and the tolerance range is ± 1 , object detection may be performed if the depth at the location is 3, 4, or 5. In example embodiments wherein the threshold depth includes a minimum threshold depth and a maximum threshold depth, the object detection circuitry 118 may perform object detection at a pixel location or in a region using the object scanning area of size $M*N$ if the

depth of the pixel location or region is not less than the minimum depth threshold and is not greater than the maximum depth threshold.

[0054] The object detection circuitry 118 may, for example, iteratively determine for each pixel of the image the number of pixels lying within a region of size $M*N$ that satisfy the threshold depth for the object scanning area of size $M*N$. This number of pixels may be determined by the object detection circuitry 118 through brute force counting, an integral image technique, and/or the like. If the number or percentage of pixels satisfying the threshold depth is greater than a defined threshold (for example, 50% or more of the pixels within the $M*N$ region), then the object detection circuitry 118 may perform object detection at the pixel using the object scanning area of size $M*N$. Otherwise, object detection at the pixel may be skipped. Accordingly, the number of pixel locations at which object detection is performed may be constrained for each object scanning area size, which may result in a significant increase in detection speed and may reduce computational burden.

[0055] In example embodiments wherein an image is segmented into regions, the object detection circuitry 118 may determine on a region-by-region basis whether to perform object detection in a given region using an object scanning area of size $M*N$. In this regard, object detection may be performed in a region using the object scanning area of size $M*N$ if the depth of the region satisfies the threshold depth for the object scanning area of size $M*N$. Otherwise, object detection in the region using the object scanning area of size $M*N$ may be skipped. Accordingly, the image search space for a given object scanning area size may be significantly reduced if one or more regions have a depth that does not satisfy a threshold depth for the object scanning area size.

[0056] In some example embodiments, the object detection circuitry 118 may be configured to constrain the image search space on the basis of a threshold object size in addition to or in lieu of a depth threshold. In this regard, an image may be segmented into M regions, which may have different respective depths. The object detection circuitry 118 may accordingly determine based on pre-computed relation information a threshold object size corresponding to a depth of a respective region. Object detection in the region may accordingly be constrained to only those object scanning area sizes satisfying the threshold object size. In example embodiments wherein the threshold object size is a single size value, the object detection circuitry 118 may, for example, perform object detection in a given region only with the object scanning area having a size $M*N$ that satisfies the

threshold object size. Alternatively, there may be a tolerance range for sizes greater than and less than the threshold object size. Accordingly, the object detection circuitry 118 may perform object detection in a given region with those object scanning area sizes that fall within the tolerance range. In example embodiments wherein the threshold object size includes a minimum threshold object size, $T1$, and a maximum threshold object size, $T2$, the object detection circuitry 118 may perform object detection in a region using those object scanning area sizes that are at least $T1$ and not greater than $T2$. As such, given a region having depth, d , only those object sizes that are permissible at that depth, as determined based on pre-computed relation information defining a relationship between depth and object size, may be evaluated. Thus, with reference to FIG. 3, only those object sizes permissible at a depth of 3 may be evaluated in the region 310.

[0057] Some example embodiments may be combined with additional techniques for performing object detection and/or for constraining an image search space. For example, skin filtering may be applied in some example embodiments to make face detection computationally more efficient. It will be appreciated that other techniques may be used in addition to or in lieu of skin filtering in other example embodiments.

[0058] FIG. 4 illustrates a flowchart according to an example method for performing object detection according to an example embodiment. The operations illustrated in and described with respect to FIG. 4 may, for example, be performed by, with the assistance of, and/or under the control of one or more of the processor 110, memory 112, communication interface 114, user interface 116, or object detection circuitry 118. Operation 400 may comprise determining depth information for an image in which object detection is to be performed. In this regard, operation 400 may comprise accessing pre-computed depth information and/or computing depth information, such as based at least in part on depth information associated with the image. The processor 110, memory 112, and/or object detection circuitry 118 may, for example, provide means for performing operation 400. Operation 410 may comprise constraining, based at least in part on the depth information and a size of an object scanning area, an image search space for performing object detection in the image using the object scanning area. The processor 110, memory 112, and/or object detection circuitry 118 may, for example, provide means for performing operation 410.

[0059] FIG. 5 illustrates a flowchart according to another example method for performing object detection according to an example embodiment. The operations

illustrated in and described with respect to FIG. 5 may, for example, be performed by, with the assistance of, and/or under the control of one or more of the processor 110, memory 112, communication interface 114, user interface 116, or object detection circuitry 118. Operation 500 may comprise determining depth information for an image in which object detection is to be performed. In this regard, operation 500 may comprise accessing pre-computed depth information and/or computing depth information, such as based at least in part on depth information associated with the image. The processor 110, memory 112, and/or object detection circuitry 118 may, for example, provide means for performing operation 500. Operation 510 may comprise determining a threshold depth for an object scanning area based at least in part on a size of the object scanning area. The processor 110, memory 112, and/or object detection circuitry 118 may, for example, provide means for performing operation 510. Operation 520 may comprise constraining, based at least in part on the depth information, an image search space for performing object detection using the object scanning area to only any portions of the image having a depth satisfying the threshold depth. The processor 110, memory 112, and/or object detection circuitry 118 may, for example, provide means for performing operation 520.

[0060] FIG. 6 illustrates a flowchart according to another example method for performing object detection according to an example embodiment. The operations illustrated in and described with respect to FIG. 6 may, for example, be performed by, with the assistance of, and/or under the control of one or more of the processor 110, memory 112, communication interface 114, user interface 116, or object detection circuitry 118. Operation 600 may comprise determining depth information for an image in which object detection is to be performed. In this regard, operation 600 may comprise accessing pre-computed depth information and/or computing depth information, such as based at least in part on depth information associated with the image. The processor 110, memory 112, and/or object detection circuitry 118 may, for example, provide means for performing operation 600. Operation 610 may comprise determining a threshold object size corresponding to a depth of a region of the image. The processor 110, memory 112, and/or object detection circuitry 118 may, for example, provide means for performing operation 610. Operation 620 may comprise performing object detection in the region using an object scanning area having a defined size only in an instance in which the size of the object scanning area satisfies the threshold object size. The processor 110, memory 112, and/or

object detection circuitry 118 may, for example, provide means for performing operation 620.

[0061] FIGs. 4-6 each illustrate a flowchart of a system, method, and computer program product according to an example embodiment. It will be understood that each

5 block of the flowcharts, and combinations of blocks in the flowcharts, may be implemented by various means, such as hardware and/or a computer program product comprising one or more computer-readable mediums having computer readable program instructions stored thereon. For example, one or more of the procedures described herein may be embodied by computer program instructions of a computer program product. In
10 this regard, the computer program product(s) which embody the procedures described herein may be stored by one or more memory devices of a mobile terminal, server, or other computing device (for example, in the memory 112) and executed by a processor in the computing device (for example, by the processor 110). In some embodiments, the computer program instructions comprising the computer program product(s) which
15 embody the procedures described above may be stored by memory devices of a plurality of computing devices. As will be appreciated, any such computer program product may be loaded onto a computer or other programmable apparatus (for example, an object detection apparatus 102) to produce a machine, such that the computer program product including the instructions which execute on the computer or other programmable apparatus creates
20 means for implementing the functions specified in the flowchart block(s). Further, the computer program product may comprise one or more computer-readable memories on which the computer program instructions may be stored such that the one or more computer-readable memories can direct a computer or other programmable apparatus to function in a particular manner, such that the computer program product comprises an
25 article of manufacture which implements the function specified in the flowchart block(s). The computer program instructions of one or more computer program products may also be loaded onto a computer or other programmable apparatus (for example, an object detection apparatus 102) to cause a series of operations to be performed on the computer or other programmable apparatus to produce a computer-implemented process such that the
30 instructions which execute on the computer or other programmable apparatus implement the functions specified in the flowchart block(s).

[0062] Accordingly, blocks of the flowcharts support combinations of means for performing the specified functions. It will also be understood that one or more blocks of

the flowcharts, and combinations of blocks in the flowcharts, may be implemented by special purpose hardware-based computer systems which perform the specified functions, or combinations of special purpose hardware and computer program product(s).

[0063] The above described functions may be carried out in many ways. For example, any suitable means for carrying out each of the functions described above may be employed to carry out embodiments of the invention. In one embodiment, a suitably configured processor (for example, the processor 110) may provide all or a portion of the elements. In another embodiment, all or a portion of the elements may be configured by and operate under control of a computer program product. The computer program product for performing the methods of an example embodiment of the invention includes a computer-readable storage medium (for example, the memory 112), such as the non-volatile storage medium, and computer-readable program code portions, such as a series of computer instructions, embodied in the computer-readable storage medium.

[0064] Many modifications and other embodiments of the inventions set forth herein will come to mind to one skilled in the art to which these inventions pertain having the benefit of the teachings presented in the foregoing descriptions and the associated drawings. Therefore, it is to be understood that the embodiments of the invention are not to be limited to the specific embodiments disclosed and that modifications and other embodiments are intended to be included within the scope of the invention. Moreover, although the foregoing descriptions and the associated drawings describe example embodiments in the context of certain example combinations of elements and/or functions, it should be appreciated that different combinations of elements and/or functions may be provided by alternative embodiments without departing from the scope of the invention. In this regard, for example, different combinations of elements and/or functions than those explicitly described above are also contemplated within the scope of the invention. Although specific terms are employed herein, they are used in a generic and descriptive sense only and not for purposes of limitation.

We Claim:

1. A method comprising:
determining depth information for an image in which object detection is to be
5 performed; and
constraining, based at least in part on the depth information and a defined size of an
object scanning area, an image search space for performing object detection in the image
using the object scanning area.
- 10 2. The method of Claim 1, wherein determining depth information for the
image comprises determining depth information based at least in part on a pixel depth map
for the image.
- 15 3. The method of any of Claims 1-2, wherein:
determining depth information for the image comprises segmenting, based at least
in part on a depth map, the image into a plurality of contiguous regions, each region having
a respective depth; and
constraining the image search space comprises performing object detection in a
region using the object scanning area only in an instance in which the depth of the region
20 satisfies a relationship criterion with respect to the size of the object scanning area.
- 25 4. The method of any of Claims 1-3, wherein constraining the image search
space comprises constraining the image search space based at least in part on a pre-
computed relationship between object size and depth.
5. The method of any of Claims 1-4, wherein constraining the image search
space comprises constraining the image search space based at least in part on a type of
device used to capture the image.
- 30 6. The method of any of Claims 1-5, further comprising:
determining a threshold depth for the object scanning area based at least in part on
the defined size of the object scanning area; and

wherein constraining the image search space comprises constraining the image search space for performing object detection using the object scanning area to only any portions of the image having a depth satisfying the threshold depth.

- 5 7. The method of Claim 6, wherein:
 determining the threshold depth comprises determining a minimum depth threshold
 and a maximum depth threshold; and
 constraining the image search space comprises constraining the image search space
 to only any portions of the image having a depth that is not less than the minimum depth
10 threshold and that is not greater than the maximum depth threshold.

8. The method of any of Claims 1-5, wherein constraining the image search
 space comprises:
 determining a threshold object size corresponding to a depth of a region of the
15 image; and
 performing object detection in the region using the object scanning area only in an
 instance in which the defined size of the object scanning area satisfies the threshold object
 size.

- 20 9. The method of any of Claims 1-8, wherein performing object detection
 comprises performing face detection.

10. A computer program product comprising at least one computer-readable
 storage medium having computer-readable program instructions stored therein, the
25 computer-readable program instructions comprising program instructions configured to
 cause an apparatus to perform a method according to any of Claims 1-9.

11. An apparatus comprising at least one processor and at least one memory
 storing computer program code, wherein the at least one memory and stored computer
30 program code are configured, with the at least one processor, to cause the apparatus to at
 least:
 determine depth information for an image in which object detection is to be
 performed; and

constrain, based at least in part on the depth information and a defined size of an object scanning area, an image search space for performing object detection in the image using the object scanning area.

5 12. The apparatus of Claim 11, wherein the at least one memory and stored computer program code are configured, with the at least one processor, to cause the apparatus to determine depth information for the image by determining depth information based at least in part on a pixel depth map for the image.

10 13. The apparatus of any of Claims 11-12, wherein the at least one memory and stored computer program code are configured, with the at least one processor, to cause the apparatus to:

 determine depth information for the image at least in part by segmenting, based at least in part on a depth map, the image into a plurality of contiguous regions, each region
15 having a respective depth; and

 constrain the image search space by performing object detection in a region using the object scanning area only in an instance in which the depth of the region satisfies a relationship criterion with respect to the size of the object scanning area.

20 14. The apparatus of any of Claims 11-13, wherein the at least one memory and stored computer program code are configured, with the at least one processor, to cause the apparatus to constrain the image search space based at least in part on a pre-computed relationship between object size and depth.

25 15. The apparatus of any of Claims 11-14, wherein the at least one memory and stored computer program code are configured, with the at least one processor, to cause the apparatus to constrain the image search space based at least in part on a type of device used to capture the image.

30 16. The apparatus of any of Claims 11-15, wherein the at least one memory and stored computer program code are configured, with the at least one processor, to further cause the apparatus to:

determine a threshold depth for the object scanning area based at least in part on the defined size of the object scanning area; and

constrain the image search space at least in part by constraining the image search space for performing object detection using the object scanning area to only any portions
5 of the image having a depth satisfying the threshold depth.

17. The apparatus of Claim 16, wherein the at least one memory and stored computer program code are configured, with the at least one processor, to cause the apparatus to:

10 determine the threshold depth by determining a minimum depth threshold and a maximum depth threshold; and

constrain the image search space at least in part by constraining the image search space to only any portions of the image having a depth that is not less than the minimum depth threshold and that is not greater than the maximum depth threshold.

15

18. The apparatus of any of Claims 11-15, wherein the at least one memory and stored computer program code are configured, with the at least one processor, to cause the apparatus to constrain the image search space at least in part by:

determining a threshold object size corresponding to a depth of a region of the
20 image; and

performing object detection in the region using the object scanning area only in an instance in which the defined size of the object scanning area satisfies the threshold object size.

25 19. The apparatus of any of Claims 11-18, wherein performing object detection comprises performing face detection.

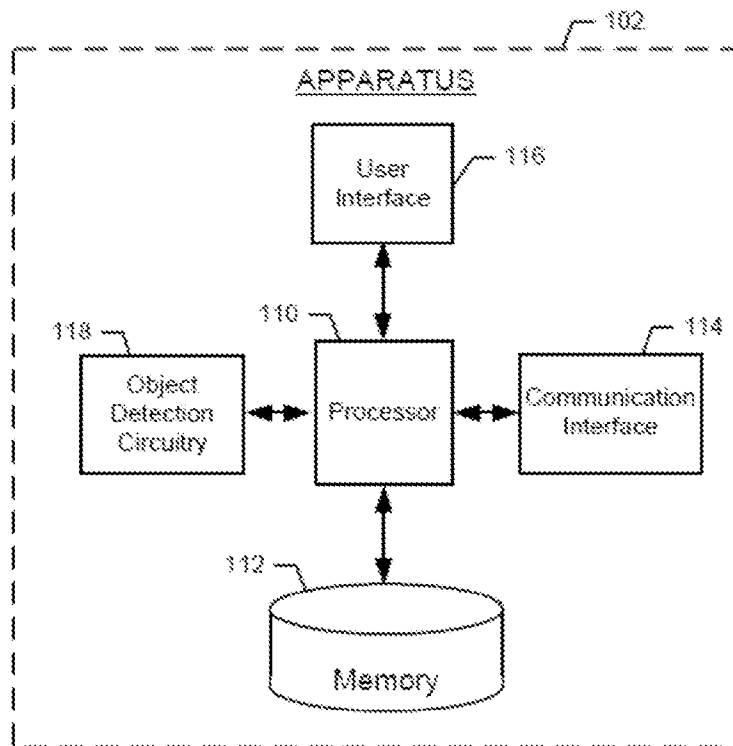
20. The apparatus of any of Claims 11-19, wherein the apparatus comprises or is embodied on a mobile computing device, the mobile computing device comprising user
30 interface circuitry and user interface software stored on one or more of the at least one memory, wherein the user interface circuitry and user interface software are configured to:

facilitate user control of at least some functions of the mobile computing device through use of a display; and

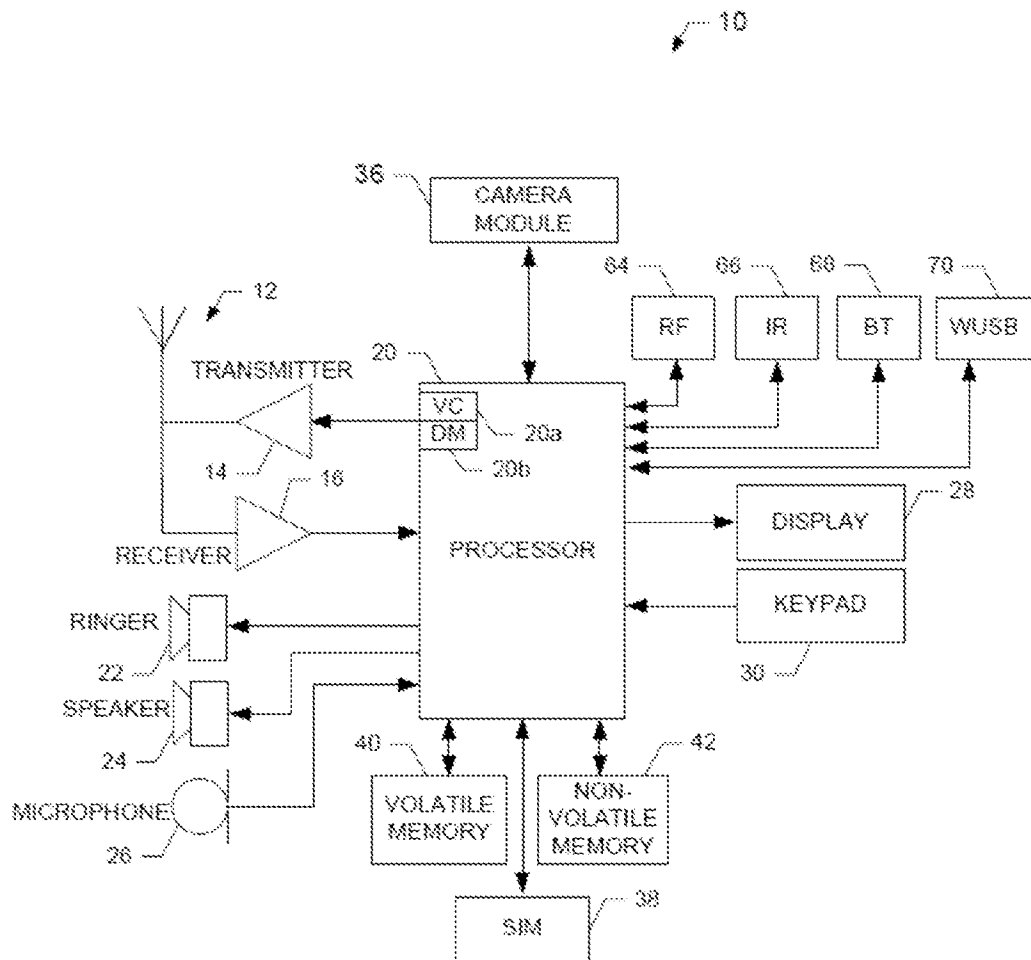
cause at least a portion of a user interface of the mobile computing device to be displayed on the display to facilitate user control of at least some functions of the mobile computing device.

- 5 21. An apparatus comprising:
- means for determining depth information for an image in which object detection is
- to be performed; and
- means for constraining, based at least in part on the depth information and a
- 10 defined size of an object scanning area, an image search space for performing object
- detection in the image using the object scanning area.

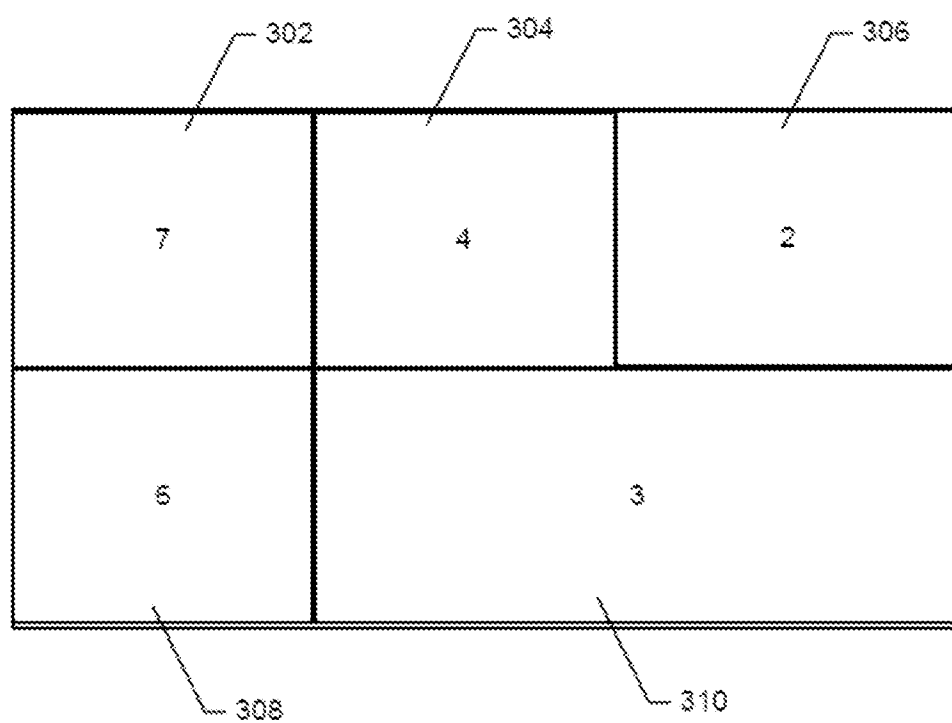
1/6

FIG. 1

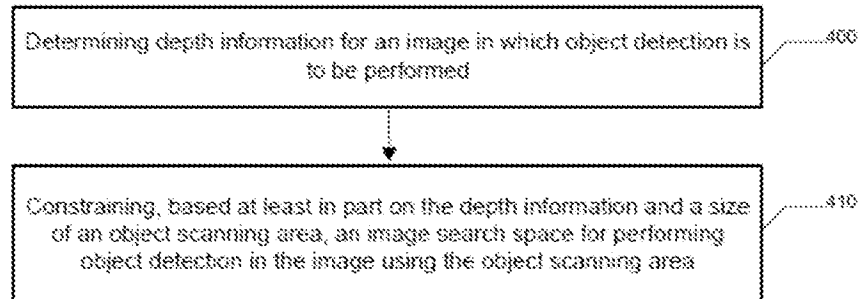
2/6

**FIG. 2**

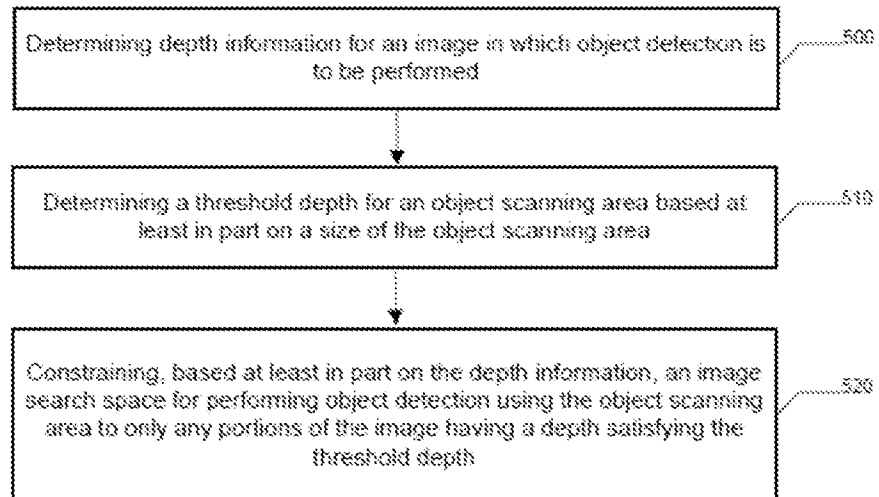
3/6

FIG. 3

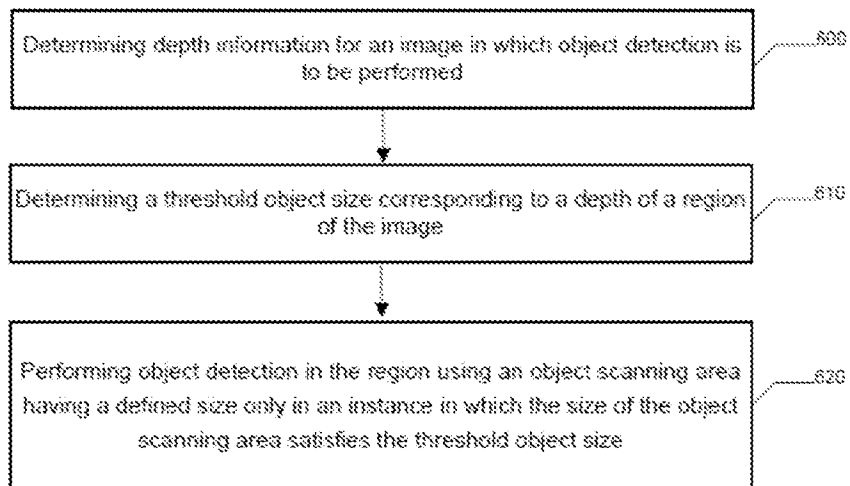
4/5

FIG. 4

5/6

FIG. 5

6/6

FIG. 6

INTERNATIONAL SEARCH REPORT

International application No.

PCT/FI2011/050976

A. CLASSIFICATION OF SUBJECT MATTER

See extra sheet

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)

IPC: G06K, G06T, H04N

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched
FI, SE, NO, DK

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

EPO-Internal, WPI, IEEE Xplore, Google Scholar

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	US 2007121094 A1 (GALLAGHER ANDREW et al.) 31 May 2007 (31.05.2007) abstract; paragraphs [0003]-[0004], [0039]-[0041], [0082]-[0088]; claim 1; Figs. 2c, 3, 4	1-21
X	US 2004252862 A1 (CAMUS THEODORE ARMAND et al.) 16 December 2004 (16.12.2004) abstract; par. [0004]-[0005], [0007], [0014]-[0024]; claim 1; Figs. 1-4	1-4, 10-14, 21
X	US 2003235341 A1 (GOKTURK SALIH BURAK et al.) 25 December 2003 (25.12.2003) abstract; par. [0018]-[0019], [0033], [0035]-[0036], [0039]-[0042], [0051], [0055]; claim 1; Fig. 2	1-4, 10-14, 21
X	US 2008219505 A1 (MORIMITSU NOBORU) 11 September 2008 (11.09.2008) abstract; par. [0012]-[0013], [0061]-[0075]; claim 1; Figs. 1, 4	1-4, 10-14, 21

☒ 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
06 March 2012 (06.03.2012)

Date of mailing of the international search report
08 March 2012 (08.03.2012)

Name and mailing address of the ISA/FI
National Board of Patents and Registration of Finland
P.O. Box 1160, FI-00101 HELSINKI, Finland
Facsimile No. +358 9 6939 5328

Authorized officer
Timo Laakso
Telephone No. +358 9 6939 500

INTERNATIONAL SEARCH REPORT

International application No.

PCT/FI2011/050976

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	EP 2219133 A1 (AUTOLIV DEV) 18 August 2010 (18.08.2010) the whole document	1-21
A	EP 1892149 A1 (HARMAN BECKER AUTOMOTIVE SYS) 27 February 2008 (27.02.2008) the whole document	1-21

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.
PCT/FI2011/050976

Patent document cited in search report	Publication date	Patent family members(s)	Publication date
US 2007121094 A1	31/05/2007	WO 2007064465 A1	07/06/2007
.....			
US 2004252862 A1	16/12/2004	US 2008159620 A1	03/07/2008
		US 2006210117 A1	21/09/2006
		WO 2004114202 A1	29/12/2004
		EP 1639519 A1	29/03/2006
		US 2004258279 A1	23/12/2004
		US 2004252864 A1	16/12/2004
		US 2004252863 A1	16/12/2004
.....			
US 2003235341 A1	25/12/2003	None	
.....			
US 2008219505 A1	11/09/2008	DE 102008012840 A1	06/11/2008
		JP 2008216177 A	18/09/2008
.....			
EP 2219133 A1	18/08/2010	US 2011311103 A1	22/12/2011
		AT 527620T T	15/10/2011
		WO 2010094401 A1	26/08/2010
.....			
EP 1892149 A1	27/02/2008	US 2008049975 A1	28/02/2008
		JP 2008052741 A	06/03/2008
		AT 422185T T	15/02/2009
.....			

INTERNATIONAL SEARCH REPORT

International application No.

PCT/FI2011/050976

CLASSIFICATION OF SUBJECT MATTER

Int.Cl.

G06K 9/00 (2006.01)

G06T 7/00 (2006.01)

H04N 5/232 (2006.01)