



- (51) **International Patent Classification:**
G01B 11/02 (2006.01) *H04N 13/02* (2006.01)
- (21) **International Application Number:**
PCT/US2016/068203
- (22) **International Filing Date:**
22 December 2016 (22.12.2016)
- (25) **Filing Language:** English
- (26) **Publication Language:** English
- (30) **Priority Data:**
15/069,491 14 March 2016 (14.03.2016) US
- (71) **Applicant:** SYMBOL TECHNOLOGIES, LLC
[US/US]; 1 Zebra Plaza, Holtsville, NY 11742 (US).
- (72) **Inventor:** ZOLOTOV, Serguei; 62 Mattawa Crescent, Ot-
tawa, Ontario K2M 2E8 (CA).
- (74) **Agents:** KAPMAR, Dmitry et al.; Zebra Technologies,
Corp., 3 Overlook Point, Lincolnshire, IL 60069 (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, DJ, 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, KH, KN,
KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA,
MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG,
NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS,
RU, RW, SA, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY,
TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN,
ZA, ZM, ZW.

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

Declarations under Rule 4.17:

- *as to applicant's entitlement to apply for and be granted a
patent (Rule 4.17(ii))*

Published:

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

(54) **Title:** DEVICE AND METHOD OF DIMENSIONING USING DIGITAL IMAGES AND DEPTH DATA

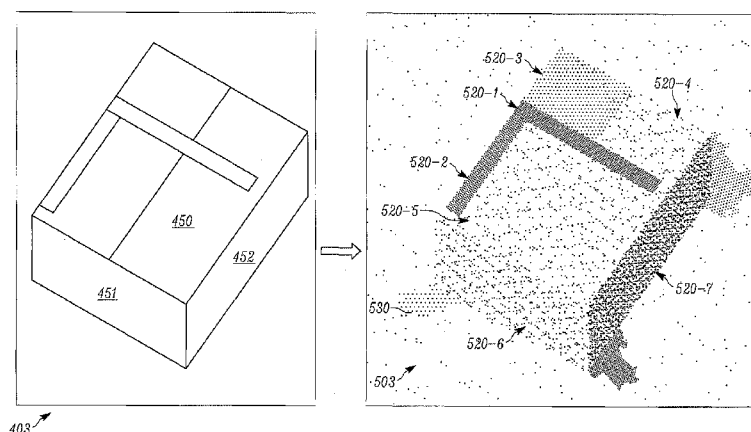


FIG. 5

(57) **Abstract:** A device and method of dimensioning using digital images and depth data is provided. The device includes a camera and a depth sensing device whose fields of view generally overlap. Segments of shapes belonging to an object identified in a digital image from the camera are identified. Based on respective depth data, from the depth sensing device, associated with each of the segments of the shapes belonging to the object, it is determined whether each of the segments is associated with a same shape belonging to the object. Once all the segments are processed to determine their respective associations with the shapes of the object in the digital image, dimensions of the object are computed based on the respective depth data and the respective associations of the shapes.

DEVICE AND METHOD OF DIMENSIONING USING DIGITAL IMAGES AND DEPTH DATA

BACKGROUND

[0001] Digital cameras are commonly used on mobile devices. Depth sensing cameras are generally not used on mobile devices; however depth sensing cameras (such as time-of-flight cameras, structured light cameras, stereo cameras) can be used to measure object dimensions. Measuring object dimensions with depth sensing cameras, however, can require high quality and/or expensive components to deliver high precision data to compete with common ruler measurements. Furthermore, many of such depth sensing cameras have inherent problems. For example, depth sensing cameras that use structured light are subject to interference with ambient light; in addition, their precision declines with the distance to a physical object being dimensioned; and an ASIC (application-specific integrated circuit) is required to calculate depth data from the many structured light images that are acquired. Furthermore, stereo cameras can require a significant baseline for accurate measurements, that are then subject to low light degradation; and image correspondence problems can necessitate hardware acceleration. Time-of-flight cameras can be subject to noise, reflections and light interference; furthermore, accurate dimensioning measurements can require precise corners of the physical object being measured, and the time-of-flight can further provide poor results at edges of the physical object.

BRIEF DESCRIPTION OF THE SEVERAL VIEWS OF THE DRAWINGS

[0002] The accompanying figures, where like reference numerals refer to identical or functionally similar elements throughout the separate views, together with the detailed description below, are incorporated in and form part of the specification, and serve to further illustrate implementations of concepts described herein, and explain various principles and advantages of those implementations.

[0003] FIG. 1 depicts front and rear views of a hybrid device for dimensioning using digital images and depth data, in accordance with some implementations.

[0004] FIG. 2 depicts a schematic block diagram of the device of FIG. 1, in accordance with some implementations.

[0005] FIG. 3 is a flowchart of a method of dimensioning using digital images and depth data, in accordance with some implementations.

[0006] FIG. 4 depicts the device of FIG. 2 acquiring a digital image and depth data of a physical object, in accordance with some implementations.

[0007] FIG. 5 depicts an example of processing of the data acquired in FIG. 4, in accordance with some implementations.

[0008] FIG. 6 depicts an example of further processing of the data acquired in FIG. 4, in accordance with some implementations.

[0009] FIG. 7 depicts an example of further processing of the data acquired in FIG. 4, in accordance with some implementations.

[0010] FIG. 8 depicts an example of further processing of the data acquired in FIG. 4, in accordance with some implementations.

[0011] FIG. 9 depicts the device of FIG. 1 rendering a digital image of a physical object as well as dimensions thereof, in accordance with some implementations.

[0012] FIG. 10 is a flowchart of a method of dimensioning using digital images and depth data, in accordance with some implementations.

[0013] Skilled artisans will appreciate that elements in the figures are illustrated for simplicity and clarity and have not necessarily been drawn to scale. For example, the dimensions of some of the elements in the figures may be exaggerated relative to other elements to help to improve understanding of implementations of the present specification.

[0014] The apparatus and method components have been represented where appropriate by conventional symbols in the drawings, showing only those specific details that are pertinent to understanding the implementations of the present specification so as not to obscure the disclosure with details that will be readily apparent to those of ordinary skill in the art having the benefit of the description herein.

DETAILED DESCRIPTION

[0015] An aspect of the specification provides a device comprising: a camera device configured to acquire a digital image in a field of view; a depth sensing device configured to acquire depth data across at least a portion of the field of view of the camera device; and, a dimensioning processor configured to: determine segments of shapes belonging to an object identified in the digital image; based on respective depth data, from the depth sensing device, associated with each of the segments of the shapes belonging to the object, determine whether each of the segments is associated with a same shape belonging to the object; and, once all the segments are processed to determine their respective associations with the shapes of the object in the digital image, compute dimensions of the object based on the respective depth data and the respective associations of the shapes.

[0016] The dimensioning processor can be further configured to determine whether each of the segments is associated with the same shape belonging to the object by: identifying adjacent segments of the shapes belonging to the object identified in the digital image; and, comparing at least a portion of the respective depth data of the adjacent segments.

[0017] The dimensioning processor can be further configured to determine whether each of the segments is associated with the same shape belonging to the object by: iterating through all adjacent segments of the segments to determine whether the adjacent segments can be joined into a larger super-segment of the same shape.

[0018] The dimensioning processor can be further configured to determine whether each of the segments is associated with the same shape belonging to the object by: determining, based on the respective depth data, whether adjacent segments are co-located with the same shape.

[0019] The dimensioning processor can be further configured to filter out one or more of identified shapes and identified segments that are not associated with the object.

[0020] The device can further comprise a display device, and the dimensioning processor can be further configured to control the display device to one or more of: render the digital image; identify the object in the digital image; and render the dimensions of the object.

[0021] The device can further comprise a memory storing data defining the object, and the dimensioning processor can be further configured to: identify the object in the digital image by comparing the digital image with the data defining the object. The data defining the object can be stored in a container in the memory, the container being one of a plurality of containers stored in the memory, each of the plurality of containers defining respective objects. The data defining the object can comprise stereometric properties of the object, the stereometric properties comprising a set of one or more of quantitative descriptions and qualitative descriptions of the object that uniquely characterize the object in three-dimensions.

[0022] The dimensioning processor can be further configured to determine geometry of the object based on relative location of the segments of the object in the digital image and in the depth data.

[0023] Each of the shapes can comprises one or more of: a primitive shape, a plane, a triangle, a square, a rectangle, a circle, a sphere, and a polygon.

[0024] The dimensioning processor can be further configured to preprocess the digital image to one or more of: filter image noise; detect edges of the object; segment color; determine morphological features in the digital image; and determine orientation of lines in the digital image.

[0025] The dimensioning processor can be further configured to preprocess the depth data to one or more of: correct for distortions; correct for interference; filter noise; and convert the depth data to point cloud data format.

[0026] The camera device can comprise an RGB (red-green-blue) digital camera.

[0027] The depth sensing device can comprise one or more of a structured light camera, a time-of-flight camera, stereo vision camera, an active stereo vision camera, a passive stereo vision camera, a stereoscopic camera, and a LIDAR (light detection and ranging) device.

[0028] Another aspect of the specification provides a method comprising: at a device comprising: a camera device configured to acquire a digital image in a field of view; a depth sensing device configured to acquire depth data across at least a portion of the field of view of the camera device; and, a dimensioning processor, determining, at the dimensioning processor, segments of shapes belonging to an object identified in the

digital image; based on respective depth data, from the depth sensing device, associated with each of the segments of the shapes belonging to the object, determining, at the dimensioning processor, whether each of the segments is associated with a same shape belonging to the object; and, once all the segments are processed to determine their respective associations with the shapes of the object in the digital image, computing, at the dimensioning processor, dimensions of the object based on the respective depth data and the respective associations of the shapes.

[0029] The determining whether each of the segments is associated with the same shape belonging to the object can occur by: identifying adjacent segments of the shapes belonging to the object identified in the digital image; and, comparing at least a portion of the respective depth data of the adjacent segments.

[0030] The determining whether each of the segments is associated with the same shape belonging to the object can occur by: iterating through all adjacent segments of the segments to determine whether the adjacent segments can be joined into a larger super-segment of the same shape.

[0031] The determining whether each of the segments is associated with the same shape belonging to the object can occur by: determining, based on the respective depth data, whether adjacent segments are co-located with the same shape.

[0032] The method can further comprise filtering out one or more of identified shapes and identified segments that are not associated with the object.

[0033] The device can further comprise a display device, and the method can further comprise controlling the display device to one or more of: render the digital image; identify the object in the digital image; and render the dimensions of the object.

[0034] The device can further comprise a memory storing data defining the object, and the method can further comprise: identifying the object in the digital image by comparing the digital image with the data defining the object. The data defining the object can be stored in a container in the memory, the container being one of a plurality of containers stored in the memory, each of the plurality of containers defining respective objects. The data defining the object can comprise stereometric properties of the object, the stereometric properties comprising a set of one or more of

quantitative descriptions and qualitative descriptions of the object that uniquely characterize the object in three-dimensions.

[0035] The method can further comprise determining geometry of the object based on relative location of the segments of the object in the digital image and in the depth data.

[0036] Each of the shapes can comprises one or more of: a primitive shape, a plane, a triangle, a square, a rectangle, a circle, a sphere, and a polygon.

[0037] The method can further comprise preprocessing the digital image to one or more of: filter image noise; detect edges of the object; segment color; determine morphological features in the digital image; and determine orientation of lines in the digital image.

[0038] The method can further comprise preprocessing the depth data to one or more of: correct for distortions; correct for interference; filter noise; and convert the depth data to point cloud data format.

[0039] The camera device can comprise an RGB (red-green-blue) digital camera.

[0040] The depth sensing device can comprise one or more of a structured light camera, a time-of-flight camera, stereo vision camera, an active stereo vision camera, a passive stereo vision camera, a stereoscopic camera, and a LIDAR (light detection and ranging) device.

[0041] Yet a further aspect of the speciation provides a non-transitory computer-readable medium storing a computer program, wherein execution of the computer program is for: at a device comprising: a camera device configured to acquire a digital image in a field of view; a depth sensing device configured to acquire depth data across at least a portion of the field of view of the camera device; and, a dimensioning processor, determining, at the dimensioning processor, segments of shapes belonging to an object identified in the digital image; based on respective depth data, from the depth sensing device, associated with each of the segments of the shapes belonging to the object, determining, at the dimensioning processor, whether each of the segments is associated with a same shape belonging to the object; and, once all the segments are processed to determine their respective associations with the shapes of the object in

the digital image, computing, at the dimensioning processor, dimensions of the object based on the respective depth data and the respective associations of the shapes.

[0042] FIG. 1 depicts a rear perspective view 1-I and a front perspective view 1-II of a dimensioning device 101; FIG. 2 depicts a schematic block diagram of device 101. With reference to both FIG. 1 and FIG. 2, the dimensioning device 101 comprises: a camera device 103 configured to acquire a digital image in a field of view; a depth sensing device 105 configured to acquire depth data across at least a portion of the field of view of the camera device 103; and, a dimensioning processor 120 configured to: determine segments of shapes belonging to an object identified in the digital image; based on respective depth data, from depth sensing device 105, associated with each of the segments of the shapes belonging to the object, determine whether each of the segments is associated with a same shape belonging to the object; and, once all the segments are processed to determine their respective associations with the shapes of the object in the digital image, compute dimensions of the object based on the respective depth data and the respective associations of the shapes. Camera device 103 will be interchangeably referred to hereafter as camera 103; and dimensioning processor 120 will be interchangeably referred to hereafter as dimensioning processor 120.

[0043] As depicted, device 101 further comprises a memory 122, a communication interface 124 (interchangeably referred to as interface 124), a display device 126, at least one input device 128, a speaker 132 and a microphone 134.

[0044] Device 101, and its components, will now be described in further detail.

[0045] Device 101 can comprise a computing device specially configured for object dimensioning, including but not limited to a graphics processing unit (GPU), a graphics processing device, a graphics processing engine, a video processing device, and the like. Indeed, device 101 can be any type of electronic device that can be used in a self-contained manner to acquire images and depth data using, respectively, camera 103 and depth sensing device 105, to dimension objects in a field of view of camera 103 and depth sensing device 105.

[0046] Device 101 can include, but is not limited to, any suitable combination of electronic devices, communications devices, computing devices, personal computers, laptop computers, portable electronic devices, mobile computing devices, portable computing devices, tablet computing devices, laptop computing devices, desktop phones, telephones, PDAs (personal digital assistants), cellphones, smartphones, e-readers, internet-enabled appliances, mobile camera devices and the like that are specially configured for object dimensioning. Other suitable devices are within the scope of present implementations. For example, device 101 need not comprise a mobile communication device, but rather can comprise a device with specialized functions, for example a device having warehouse inventory tracking and/or other data acquisition functionality, such as a mobile scanner having one or more of radio frequency identification (RFID) reader, Near Field Communication (NFC) reader, imager, and/or laser-based scanner data acquisition components.

[0047] While a specific physical configuration of device 101 is depicted in FIG. 1, other physical configurations of device 101 are within the scope of present implementations. For example, device 101 can further include one or more handles, such as a handle below the display 126, as well as a trigger for triggering the data acquisition components and the like.

[0048] Camera 103 can comprise a digital camera, an RGB (red-green-blue) digital camera, and the like, configured to acquire digital images, including, but not limited to, images in a video stream. While details of camera 103 are not depicted, it is assumed that camera 103 comprises components for acquiring digital images including, but not limited to, respective charge coupled devices (CCD) and the like, as well as respective lenses, respective focusing devices (including, but not limited to voice coils and the like), etc. Hence, data from camera 103 generally comprises two-dimensional data, and specifically a two-dimensional array of values which can comprise an array of two-dimensional color coordinates and/or brightness coordinates.

[0049] Depth sensing device 105 can comprise comprises one or more of a structured light camera, a time-of-flight camera, stereo vision camera, an active stereo vision camera (which can project its own light, including, but not limited to, infrared light), a passive stereo vision camera (which relies on ambient light), a stereoscopic camera,

and a LIDAR (light detection and ranging) device. However, other depth sensing devices are within the scope of present implementations. In some implementations, depth sensing device 105 can comprise a pair of cameras, which can include camera 103, that form a stereo vision camera and/or a stereoscopic camera. When depth sensing device 105 comprises a structured light camera, depth sensing device 105 can include a device configured to project structured light and a camera, which can include camera 103 configured to capture images of physical objects illuminated by the structured light. When depth sensing device 105 comprises a time-of-flight (ToF) camera or a LIDAR device, depth sensing device 105 comprises components for implanting such functionality. Furthermore, depth sensing device 105 can include components for acquiring depth data across a field of view including, but not limited to, respective charge coupled devices (CCD) and the like, as well as respective lenses, and respective focusing devices (including, but not limited to voice coils and the like).

[0050] Depth sensing device 105 is configured to sense depth across a field of view, such that a distance from depth sensing device 105 to surfaces of a physical object can be determined. In particular, depth data from depth sensing device 105 could be processed to produce a three-dimensional map of the field of view of depth sensing device 105, however such processing to produce a three-dimensional map of the field of view of depth sensing device 105 can be obviated by present implementations. Hence, data from depth sensing device 105 can be referred to as three-dimensional data and can comprise a two-dimensional array of depth coordinates (i.e. each element in the two-dimensional array is associated with a corresponding depth).

[0051] Hence, depth data from depth sensing device 105 generally represents a distance from the depth sensing device 105 to portions of objects in the field of view of depth sensing device 105. In some implementations the depth data can comprise, and/or be converted to, “real world coordinates” which can comprise three-dimensional coordinates in a selected reference frame and/or a given reference frame; in some implementation such a reference frame can be relative to depth sensing device 105 (e.g. depth sensing device 105 can comprise an origin of the reference frame), while in other implementations, such a reference frame can be relative to a fixed coordinate, for example a point in warehouse and/or a geographical coordinate

determined using a Global Positioning System (GPS) device. In the latter implementations, device 101 can comprise a GPS device, and coordinates of depth data from depth sensing device 105 can be determined relative to an origin defined with respect to GPS data.

[0052] As depicted in FIG. 1, external components of each of camera 103 and depth sensing device 105 can be located on a rear side of device 101, and display device 126 can be looked on a front side of device 101, such that digital images and depth data can be captured at a rear side of device 101. Put another way, fields of view of each of camera 103 and depth sensing device 105 overlap as both devices 103, 105 are disposed at a rear side of device 101. However, both devices 103, 105 need not be disposed on a same side of device 101 as long as fields of view of each of devices 103, 105 overlap such that an object being dimensioned by dimensioning device 101 is in a field of view of each of devices 103, 105.

[0053] Furthermore, as also depicted in FIG. 1, image capturing components and/or data capturing components of each of camera 103 and depth sensing device 105, such as lenses and the like, can be separated by a given distance such that images from camera 103 and depth data from depth sensing device 105 generally image and/or sense depth across a similar field of view. In other words, respective fields of view of camera 103 and depth sensing device 105 generally overlap, and indeed, in some implementations, lenses, and the like, of each of camera 103 and depth sensing device 105 are as close together as possible both to reduce parallax therebetween and to ensure that depth data can be captured for an object imaged by camera 103.

[0054] It is appreciated that the locations of each of camera 103 and depth sensing device 105 in each of FIG. 1 and FIG. 2 are merely schematic and do not necessarily represent actual relative positions of each of camera 103 and depth sensing device 105; in other words, in FIG.1, while each of camera 103 and depth sensing device 105 are depicted as side by side across a transverse longitudinal of device 101, each of camera 103 and depth sensing device 105 can be located anywhere on device 101, presuming that their fields of view at least partially overlap such that physical objects in their respective fields of view can be both imaged by camera 103 and sensed by depth sensing device 105.

[0055] In an ideal device, camera 103 and depth sensing device 105 would occupy the same point in space such that their respective fields of view would be identical; however, as that is not practical, one implementation includes an RGB camera (e.g. camera 103) and a ToF camera . Furthermore, such a configuration leads to preprocessing of respective digital images and depth data to align corresponding areas and/or pixels, to attempt to minimize “pixel shadowing” and “dark regions”. For example, due to parallax in view points between camera 103 and depth sensing device 105, in some cases of arrangements of object in their field of view, regions that are visible to camera 103 are not visible to depth sensing device 105, or vice versa. This can cause: an absence of image data in regions for which there is depth data; and/or an absence of depth data in regions for which there is image data. Such regions can be referred to as “dark” as it is generally difficult to associate them with any segment. Larger “dark regions” can hence distort data being processed using techniques described herein, so closest possible co-location of camera 103 and depth sensing device 105 is desired in such an implementation.

[0056] Dimensioning processor 120 can comprise a processor and/or a plurality of processors, including but not limited to one or more central processors (CPUs) and/or one or more processing units and/or one or more graphic processing units (GPUs); either way, dimensioning processor 120 comprises a hardware element and/or a hardware processor. Indeed, in some implementations, dimensioning processor 120 can comprise an ASIC (application-specific integrated circuit) and/or an FPGA (field-programmable gate array) specifically configured to implement the functionality of device 101. Hence, device 101 is preferably not a generic computing device, but a device specifically configured to implement specific functionality including dimensioning using digital images and depth data as described in further detail below. For example, device 101 and/or dimensioning processor 120 can specifically comprise a computer executable engine configured to dimension objects in a field of view of camera 103 and depth sensing device 105 using digital images and depth data.

[0057] Memory 122 can comprise a non-volatile storage unit (e.g. Erasable Electronic Programmable Read Only Memory (“EEPROM”), Flash Memory) and a volatile storage unit (e.g. random access memory (“RAM”)). Programming instructions that

implement the functional teachings of device 101 as described herein are typically maintained, persistently, in memory 122 and used by dimensioning processor 120 which makes appropriate utilization of volatile storage during the execution of such programming instructions. Those skilled in the art recognize that memory 122 is an example of computer readable media that can store programming instructions executable on dimensioning processor 120. Furthermore, memory 122 is also an example of a memory unit and/or memory module and/or a non-volatile memory.

[0058] In particular, memory 122 can store an application 136 that, when executed by dimensioning processor 120, enables dimensioning processor 120 to: determine segments of shapes belonging to an object identified in the digital image; based on respective depth data, from depth sensing device 105, associated with each of the segments of the shapes belonging to the object, determine whether each of the segments is associated with a same shape belonging to the object; and, once all the segments are processed to determine their respective associations with the shapes of the object in the digital image, compute dimensions of the object based on the respective depth data and the respective associations of the shapes.

[0059] As depicted, memory 122 further stores data 140 defining an object, for example an object in a digital image acquired by camera 103; dimensioning processor 120 can hence be further configured to: identify the object in the digital image by comparing the digital image with data 140 defining the object. For example, data 140 defining the object can be stored in a container 141 in memory 122; in some of these implementations, as depicted, container 141 can be one of a plurality of containers 141-n stored in memory 122, each of the plurality of containers 141-n defining respective objects, for example in respective data 140-n stored therein.

[0060] In some implementations, data 140, 140-n defining the object can comprise stereometric properties of the object, the stereometric properties comprising a set of one or more of quantitative descriptions and qualitative descriptions of the object that uniquely characterize the object in three-dimensions (i.e. relating the object represented in the digital image and the depth data to a physical object being imaged).

[0061] Qualitative descriptions can include one or more sets of rules that can be used in an object (and/or target) description algorithm. For example, when a target object

has straight lines (as in the case of a box), a segment extraction algorithm can include, but is not limited to, using a probabilistic Hough transform to analyze lines and to outline object segments.

[0062] Quantitative descriptions can include one or more sets numerically expressed parameters that can describe an object. For example, when a target object has straight angles (as in the case of a box), a segment extraction algorithm can be used which indicates 90 degrees as a primary basket for classifying line relations in an image.

[0063] For example, planes can be characterized by one or more of: a three-dimensional normal vector having a length that is the distance from the center of the reference frame; and/or as an unbounded continuous three-dimensional object with any four points belonging to a quadrilateral. Indeed, any description of an object can be defined in data 140, 140-n presuming that such a description distinctively defines the object and is configured to be used with methods and processes described herein.

[0064] While "n" containers of data are depicted in FIG. 1, it is appreciated that any number of containers of data can be stored at memory 122 (and/or it is appreciated that data can also be stored without containers. In particular, memory 122 can be preconfigured with data 140 storing definitions of one or more objects for which dimensions are to be determined; for example, in warehousing applications, it can be assumed that boxes are to be dimensioned, hence data 140 can define a box (e.g. an object having six sides that are joined by about 90° angles). However, other types of objects that are to be dimensioned can be defined by data 140-n, including objects that are of more complex shapes than boxes. In any event, it will be assumed hereafter that boxes are to be dimensioned and data 140 defines a box.

[0065] Dimensioning processor 120 can be further configured to communicate with display device 126, which comprises any suitable one of, or combination of, flat panel displays (e.g. LCD (liquid crystal display), plasma displays, OLED (organic light emitting diode) displays, capacitive or resistive touchscreens, CRTs (cathode ray tubes) and the like. In particular, dimensioning processor 120 can be configured to control display device 126 to one or more of: render a digital image from camera 103; identify an object in the digital image; and render the dimensions of the object,

presuming such dimensions have been determined, as described in further detail below.

[0066] Device 101 generally comprises at least one input device 128 configured to receive input data, and can comprise any suitable combination of input devices, including but not limited to a keyboard, a keypad, a pointing device, a mouse, a track wheel, a trackball, a touchpad, a touch screen (e.g. integrated with display device 126), and the like. Other suitable input devices are within the scope of present implementations. In some implementations, one or more of input device 128 and display device 126 can be external to device 101, with dimensioning processor 120 in communication with any external components via a suitable connection and/or link.

[0067] As depicted, device 101 further comprises an optional speaker 132 and an optional microphone 134 (either of which can alternatively be external to device 101). Speaker 132 comprises any suitable speaker for converting audio data to sound to provide one or more of audible alerts, audible communications from remote communication devices, and the like. Microphone 134 comprises any suitable microphone for receiving sound and converting to audio data. Speaker 132 and microphone 134 can be used in combination to implement telephone and/or communication functions at device 101.

[0068] As depicted, dimensioning processor 120 also connects to optional interface 124, which can be implemented as one or more radios and/or connectors and/or network adaptors, configured to wirelessly communicate with one or more communication networks (not depicted). It will be appreciated that interface 124 is configured to correspond with network architecture that is used to implement one or more communication links to the one or more communication networks, including but not limited to any suitable combination of USB (universal serial bus) cables, serial cables, wireless links, cell-phone links, cellular network links (including but not limited to 2G, 2.5G, 3G, 4G+ such as UMTS (Universal Mobile Telecommunications System), GSM (Global System for Mobile Communications), CDMA (Code division multiple access), FDD (frequency division duplexing), LTE (Long Term Evolution), TDD (time division duplexing), TDD-LTE (TDD-Long Term Evolution), TD-SCDMA (Time Division Synchronous Code Division Multiple Access) and the like,

wireless data, Bluetooth™ links, NFC (near field communication) links, WLAN (wireless local area network) links, WiFi links, WiMax links, packet based links, the Internet, analog networks, the PSTN (public switched telephone network), access points, and the like, and/or a combination.

[0069] While not depicted, device 101 further comprises a power supply, including, but not limited to, a battery, a power pack and the like, and/or a connection to a mains power supply and/or a power adaptor (e.g. and AC-to-DC (alternating current to direct current) adaptor). In general, the power supply powers components of device 101.

[0070] Hence, it should be understood that in general a wide variety of configurations for device 101 are contemplated.

[0071] Attention is now directed to FIG. 3 which depicts a block diagram of a flowchart of a method 300 for dimensioning using digital images and depth data, according to non-limiting implementations. In order to assist in the explanation of method 300, it will be assumed that method 300 is performed using device 101, and specifically by dimensioning processor 120 and when dimensioning processor 120 processes instructions stored at memory 122, for example application 136. Indeed, method 300 is one way in which device 101 can be configured. Furthermore, the following discussion of method 300 will lead to a further understanding of device 101, and its various components. However, it is to be understood that device 101 and/or method 300 can be varied, and need not work exactly as discussed herein in conjunction with each other, and that such variations are within the scope of present implementations.

[0072] Regardless, it is to be emphasized, that method 300 need not be performed in the exact sequence as shown, unless otherwise indicated; and likewise various blocks may be performed in parallel rather than in sequence; hence the elements of method 300 are referred to herein as “blocks” rather than “steps”. It is also to be understood, however, that method 300 can be implemented on variations of device 101 as well.

[0073] At blocks 301, 303, which can be performed sequentially and/or in parallel, dimensioning processor 120 receives a digital image from camera device 103 and corresponding depth data from depth sensing device 105.

[0074] At block 305, dimensioning processor 120 determines segments of shapes belonging to an object identified in the digital image. For example, a segment can comprise an object identified in the digital image during an intermediate stage of segmentation. In some instances, a segment can include, but is not limited to, blob of indeterminate shape on a side and/or a surface of an object in the digital image that is identified in an initial processing of the digital image, which is identified for future analysis. As will be described hereafter, several segments can be combined into a given shape belonging to the object in the digital image, and the given shape can define a side of the object in the digital image; indeed, it can sometimes be assumed that one shape is associated with one side of the object, unless, for example, the object is curved (e.g. a sphere) with no distinguishable sides, where segments can nonetheless still be combined into shapes.

[0075] At block 307, based on respective depth data, from the depth sensing device, associated with each of the segments of the shapes associated with the object, dimensioning processor 120 determines whether each of the segments is associated with a same shape associated with the object. Put another way, at block 307, based on respective depth data corresponding to each of the segments of the shapes that define a given object, dimensioning processor 120 determines whether each of the segments is associated with a same shape that forms at least a part of the object.

[0076] At block 309, dimensioning processor 120 determines whether all the segments are processed to determine their respective associations with the shapes of the object in the digital image. When all the segments have not been processed (e.g. a "NO" decision at block 309), block 307 continues.

[0077] However, once all the segments are processed to determine their respective associations with the shapes of the object in the digital image (e.g. a "YES" decision at block 309), at block 311, dimensioning processor 120 computes dimensions of the object based on the respective depth data and the respective associations of the shapes.

[0078] At an optional block 313, dimensioning processor 120 can control display device 126 to one or more of: render the digital image; identify the object in the digital image; and render the dimensions of the object. Alternatively, and/or in addition to block 313, dimensioning processor 120 can store, at memory 122, one or

more of the digital image and the dimensions of the object. Alternatively, and/or in addition to block 313, dimensioning processor 120 can transmit, using interface 124, one or more of the digital image and the dimensions of the object, for example to an external device and/or a remote device; such transmission of one or more of the digital image and the dimensions of the object can occur in conjunction with cloud-based warehousing applications and/or cloud-based warehousing functionality.

[0079] Method 300 will now be discussed with reference to FIGS. 4 to 10. Attention is hence next directed to FIG. 4, which is substantially similar to FIG. 2, with like elements having like numbers. In particular, FIG. 4 depicts a non-limiting implementation of blocks 301, 303, in which dimensioning processor 120 receives a digital image 403 and depth data 405, respectively, from camera 103 and depth sensing device 105, for example in response to processing application 136. FIG. 4 also depicts a box 410 (i.e. a physical object) in a field of view of both camera 103 and depth sensing device 105. Box 410 comprises six sides (three of them visible to both camera 103 and depth sensing device 105: specifically a top 450 and sides 451, 452 extending downward from top 450). Furthermore, as depicted, a top 450 of box 410 visible to both camera 103 and depth sensing device 105 includes labels 411, 412, or the like, that are of a different color than box 410; as well, the top 450 of box 410 includes flaps such that a line visually bisects the top 450 of box 410.

[0080] Attention is next directed to FIG. 5 which depicts a non-limiting visual rendering of digital image 403 which results when camera 103 senses box 410 (and/or a physical object); for example the visual rendering of digital image 403 could be rendered at display device 126 (e.g. see FIG. 9). At block 305, dimensioning processor 120 processes digital image 403 to determine segments of shapes belonging to an object identified in the digital image; the term "object" in this instance is used to denote a group of areas and/or a group of pixels in digital image 403 which appear to represent a physical object, for example box 410.

[0081] In particular, dimensioning processor 120 processes digital image 403 to search for segments of shapes that can include, but are not limited to, a primitive shape, a plane, a triangle, a square, a rectangle, a circle, a sphere, a polygon, and the like; for example, shapes that can be used to define an object as stored in data 140. In

digital image 403 in FIG. 5, each of the top 450 and two visible sides 451, 452 can correspond to polygons and/or rectangles, and at block 305 dimensioning processor 120 determines segments which can make up the polygons or rectangles.

[0082] Hence, FIG. 5 also depicts a rendering 503 of digital image 403 after block 305 has been implemented, where segments of shapes identified in block 305 are indicated. It is appreciated that rendering 503 merely represents an intermediate step in the determining of dimensions of the physical object represented by the object in digital image 403, and rendering 503 is not necessarily displayed at display device 126.

[0083] In particular, in rendering 503, dimensioning processor 120 has identified a plurality of segments 520-1, 520-2, 520-3, 520-4, 520-5, 520-6, 520-7 in digital image 403, which will be interchangeably referred to hereafter, collectively, as segments 520, and generically as a segment 520. As seen in FIG. 5, edges of segments 520 are not necessarily well defined as image processing algorithms can cause edges and/or corners of box 410 to be unclear, depending on the contrast between sides that form the edges and/or corners, and/or the contrast between box 410 and items in the background of digital image 403.

[0084] In any event, dimensioning processor 120 processes digital image 403 to identify segments 520. In some implementations, dimensioning processor 120 can be further configured to preprocess digital image 403 to one or more of: filter image noise; detect edges of the object; segment color; determine morphological features in digital image 403; and determine orientation of lines in digital image 403. Such preprocessing can be performed in order to better identify edges between segments 520 and the like.

[0085] In yet further implementations, dimensioning processor 120 can be further configured to preprocess depth data 405 to one or more of: correct for distortions; correct for interference; filter noise; and convert depth data 405 to a point cloud data format. In these latter implementations, further processing of depth data 405 can comprise processing depth data 405 in the point cloud data format.

[0086] In some implementations, dimensioning processor 120 can be further configured to filter out one or more of identified shapes and identified segments that are not associated with the object. For example, in rendering 503, various shapes and/or segments are shown, such as segment 530, that are not associated with box 410, which can occur due to lighting of box 410 and/or items adjacent box 410, noise and the like. Such shapes and/or segments can be filtered out.

[0087] Indeed, in some implementations, segments 520, 530, and the like, can be compared to data 140-n, including data 140 defining various objects and their corresponding properties, to determine which segments 520, 530 to filter out and which segments 520, 530 can be used in method 300. In yet further implementations, data from digital image 403 and depth data 405 are merged to identify segments 520, 530. For example, one or more of data 140, 140-n, and application 136, can store properties and/or rules which define objects, and such rules can be used to identify segments 520, 530. For example, such properties can include, but are not limited to, position (e.g. a top plane, a floor plane, side planes, and the like), geometry (a cuboid frame, corner locations, and the like), groupings of similar textures/colors, relations of particular features between corners, between edges/lines, between sides, between texture areas, and the like. Hence, identifying segments can include identifying portions of digital image 403 that delineate different types of areas (e.g. based on color, texture, position), which can be augmented by comparing corresponding depth data 405 to determine whether the segments are part of a same shape belonging to the object or not, as described in further detail hereafter.

[0088] In any event, segments 520-1, 520-2 respectively correspond to labels 411, 412, segment 520-3 corresponds to an area defined by label 411 and a flap of the top 450 of box 410, segments 520-4, 520-5 correspond to remaining areas of the top 450 of box 410, and segments 520-6, 520-7 correspond to two sides 451, 452 of box 410 visible to camera 103 and depth sensing device 105. However, from digital image 403 alone, dimensioning processor 120 is generally unable to determine and/or confirm which segments 520 are associated with which shapes of the object; hence dimensioning processor 120 is generally unable to determine and/or confirm which segments 520 are associated with sides 451, 452 and/or a top 450 of box 410; for

example while each of segments 520-1, 520-2, 520-3, 520-4, 520-5, together, correspond to a same shape and/or a top 450 of box 410, it is not easily possible to make such a conclusion on the basis of digital image 403 alone.

[0089] Hence, at block 307, dimensioning processor 120 processes respective depth data 405 associated with each of segments 520 of the shapes belonging to the object to determine whether each of segments 520 are associated with a same shape. For example, attention is next directed to FIG. 6 which depicts a non-limiting example of an implementation of block 305.

[0090] In particular, dimensioning processor 120 can identify adjacent segments 520 of the shapes belonging to the object identified in digital image 403, and compare at least a portion of the respective depth data of the adjacent segments 520. Depth data 405 can be parsed to associate portions of depth data 405 with segments identified in digital image 403, without performing a full three-dimensional mapping using depth data 405, which can be time consuming and/or use excessive processing resources. For example, segments 520-1, 520-2 can be determined to be adjacent shapes as there is no other segment 520 there between. Furthermore, respective depth data associated with each segment 520-1, 520-2 can be compared to determine a relative distance from depth sensing device 105 and/or a relative position in a reference frame.

[0091] For example, as depicted, depth data 405 associated with an area 601 of segment 520-1 is compared to depth data 405 associated with an area 602 of adjacent segment 520-2 to determine whether segments 520-1, 520-2 can be combined. Hence, areas 601, 602 of digital image 403 are mapped to depth data 405. While each of areas 601, 602 is depicted as an ellipse, areas 601, 602 can be any shape within a respective segment 520-1, 520-2, and be any size from a pixel to the entirety of a respective segment 520-1, 520-2. Indeed, any portion of respective depth data 405 associated with segments 520-1, 520-2 can be compared to determine whether segments 520-1, 520-2 are associated with a same shape.

[0092] For example, portions of respective depth data 405 associated with segments 520-1, 520-2 can be compared to determine whether they form a subset of a same plane (e.g. a top 450 of box 410); as such, respective depth data 405 associated with segments 520-1, 520-2 will generally be about continuous, and/or lie in a same plane

and/or are co-planar. In other words, depth data 405 associated with segments 520-1, 520-2 is processed to determine whether it meets criteria for joining them together into a larger super-segment of the same shape belonging to the object. For example, in some implementations, dimensioning processor 120 can be further configured to determine whether each of segments 520-1, 520-2 are associated with a same shape belonging to the object, by: determining, based on respective depth data 405, whether adjacent segments 520-1, 520-2 are co-located with the same shape. Hence, again, respective depth data 405 of adjacent segments 520-1, 520-2 is processed to determine whether the respective depth data 405 is co-planar.

[0093] Such a super-segment 720 is depicted in FIG. 7, which is substantially similar to FIG. 6, with like elements having like numbers; however, FIG. 7 depicts a non-limiting visual rendering 703 of digital image 403 which results when, segments 520-1, 520-2 have been joined into a larger super-segment 720; specifically, rendering 703 is derived from rendering 503, but in rendering 703 segments 520-1, 520-2 have been joined into a larger super-segment 720. It is appreciated that rendering 703 merely represents an intermediate step in the determining of dimensions of the physical object represented by the object in digital image 403, and rendering 703 is not necessarily displayed at display device 126.

[0094] FIG. 7 also depicts areas 605, 606, similar to areas 601, 602, but associated respectively with segments 520-5, 520-6. Again, areas 605, 606 can be larger or smaller than depicted. Furthermore, portions of respective depth data 405 associated with segments 520-5, 520-6 can be compared to determine whether they form a subset of a same plane (e.g. a top 450 of box 410). However, as respective depth data 405 associated with segments 520-5, 520-6 indicate, for example, that segments 520-5, 520-6 are not co-planar, it is determined that segments 520-5, 520-6 are part of different shapes, and not part of the same shape.

[0095] In any event, dimensioning processor 120 can be configured to iterate through all adjacent segments of segments 520 to determine whether the adjacent segments can be joined into a larger super-segment of the same shape. As such, dimensioning processor 120 can determine that all of segments 520-1, 520-2, 520-3, 520-4, 520-5 can be joined together into a larger super-segment. For example, attention is next

directed to FIG. 8 which depicts a non-limiting visual rendering 803 of digital image 403 which results once all segments 520 are processed to determine their respective associations with the shapes of the object in digital image 403. It is appreciated that rendering 803 merely represents an intermediate step in the determining of dimensions of the physical object represented by the object in digital image 403, and rendering 803 is not necessarily displayed at display device 126.

[0096] Specifically, rendering 803 is derived from renderings 503, 703 but in rendering 803, segments 520-1, 520-2, 520-3, 520-4, 520-5 have been joined into a larger super-segment 820.

[0097] For example, in FIG. 8, dimensioning processor 120 has iterated through all of segments 520-1, 520-2, 520-3, 520-4, 520-5 (as well as segments 520-6, 520-7) and determined that segments 520-1, 520-2, 520-3, 520-4, 520-5 are all associated with a same shape, which can be associated with a top 450 of box 410. Alternatively, after super-segment 720 is formed, dimensioning processor 120 iterates through the remaining segments and determines that segments 520-3, 520-4, 520-5 are to be joined with super-segment 720 to form super-segment 820.

[0098] During the iteration, dimensioning processor 120 has further determined that segments 520-6, 520-7 are not associated with a same shape as segments 520-1, 520-2, 520-3, 520-4, 520-5 (and/or super-segment 720). Hence, in FIG. 8, dimensioning processor 120 has further determined that segments 520-6, 520-7 each, independently form a respective shape, and have hence not been joined with other segments 520.

[0099] Once all segments 520 are processed to determine their respective associations with the shapes of the object in digital image 403 (e.g. as determined at block 309), at block 311 dimensioning processor 120 computes dimensions of the object based on the respective depth data and the respective associations of the shapes. Dimensioning processor 120 can compare data represented by rendering 803 with data 140 to determine that the physical object represented by the object in digital image 403 represents a box (e.g. box 410), and determine dimensions of the box using any dimensioning algorithm.

[00100] For example, again with attention to FIG. 8, sides of segments 520-6, 520-7, 820 of rendering 803 corresponding to sides of the box can be identified and dimensions thereof determined from depth data 405. In FIG. 8 edges of segments 520-6, 520-7, 820 corresponding to height (H), length (L) and width (W) of the box are determined and indicated (for example, based on identifying intersections of adjacent non-coplanar shapes of the object from the depth data), and dimensions of edges H, L, W can be determined from depth data 405, by the correlation between depth data 405 and digital image 403 and/or a correlation between depth data 405 and edges H, L, W identified in rendering 803. In other words, dimensioning processor 120 can be further configured to determine geometry of the object based on relative location of segments 520 of the object in digital image 403 and in depth data 405. From the geometry, the dimensions can be determined.

[00101] When the dimensions have been determined, dimensioning processor 120 can render the dimensions at display device 126. For example, attention is next directed to FIG. 9 which depicts a perspective front view of device 101 in use; specifically, box 410 is in a field of view of camera 103 and depth sensing device 105 (on a rear side of device 101, and hence not visible in FIG. 9), and method 300 has been used to determine dimensions of box 410. Dimensioning processor 120 hence controls display device 126 to render digital image 403 showing box 410, as well as determined dimensions of box 410. As depicted, lines corresponding to a height, a length and width of box 410 are also rendered at display device 126 with respective dimensions thereof also shown. As those of skill in the art will realize, object volume may also be computed and displayed based on the determined dimensions as well as data 140 defining the object.

[00102] Attention is next directed to FIG. 10, which depicts an alternative of method 300; specifically FIG. 10 depicts a block diagram of a flowchart of a method 1000 for dimensioning using digital images and depth data, according to non-limiting implementations, according to non-limiting implementations. In order to assist in the explanation of method 1000, it will be assumed that method 1000 is performed using device 101, and specifically by dimensioning processor 120 and when dimensioning processor 120 processes instructions stored at memory 122, for example application

136. Indeed, method 1000 is one way in which device 101 can be configured. Furthermore, the following discussion of method 1000 will lead to a further understanding of device 101, and its various components. However, it is to be understood that device 101 and/or method 1000 can be varied, and need not work exactly as discussed herein in conjunction with each other, and that such variations are within the scope of present implementations.

[00103] Regardless, it is to be emphasized, that method 1000 need not be performed in the exact sequence as shown, unless otherwise indicated; and likewise various blocks may be performed in parallel rather than in sequence; hence the elements of method 1000 are referred to herein as “blocks” rather than “steps”. It is also to be understood, however, that method 1000 can be implemented on variations of device 101 as well.

[00104] At blocks 1001, 1003, dimensioning processor 120 receives, respectively from camera 103 and depth sensing device 105, image data and depth data. Blocks 1001, 1003 respectively correspond to blocks 301, 303 of method 300. Furthermore, in method 1000, image data can be interchangeably referred to as two-dimensional (2D) data, and depth data can be interchangeably referred to as three-dimensional (3D) data. Indeed, in general, data from camera 103 is two-dimensional and can comprise an array of two-dimensional color coordinates and/or brightness coordinates; similarly, in general, data from depth sensing device 105 is three-dimensional and can comprise a two-dimensional array of depth coordinates (i.e. each element in the two-dimensional array is associated with a corresponding depth).

[00105] At blocks 1004-1, 1004-2, dimensioning processor 120 can respectively preprocess the 2D data and the 3D data. For example, preprocessing the 2D digital image at block 1004-1 can include, but is not limited to, image noise filtering, edges detection, color segmentation, morphological feature segmentation (e.g. line orientation, edge orientation, etc.), and the like. Preprocessing the 3D depth data at block 1004-2 can include, but is not limited to, distortion corrections, interference corrections, noise filtering, phase (depth) map to point cloud data (PCD) conversion, and the like

[00106] At block 1005, dimensioning processor 120 locates and/or extracts an object of interest from the 2D data and the 3D data by merging the 2D data and the 3D data and locating the object based on properties stored, for example, in data 140 and/or based on properties configured in application 136. Such properties can include, but are not limited to, position (e.g. a top plane, a floor plane, side planes, and the like), geometry (a cuboid frame, corner locations, and the like), groupings of similar textures/colors, relations of particular features between corners, between edges/lines, between sides, between texture areas, and the like. Indeed, locating objects and/or segments associated with an object in a digital image can hence include processing both the digital image and the depth data to find such segments and/or objects. Hence block 1005 corresponds to block 305 of method 300.

[00107] At block 1007, dimensioning processor 120 can map an object area, for example in the digital image, the 3D depth data (e.g. referred to in block 1007 as the 3D data set). Such mapping can include, but is not limited to, extracting spatial data from the 3D data set, which can include extracting at least two sides, extracting one side and one main diagonal, and the like; in other words, sufficient spatial data is extracted to characterize the object without performing a full three-dimensional mapping. The degree of data that is sufficient to characterize a given type of object can be defined in data 140-n; in other words, for given type of object, data 140-n includes rules for determining a degree of mapping that is to occur to map an object area. Further at block 1007, dimensioning processor 120 iterates over one or more best fits of segments, including but not limited to best rectangular fits, based on edge locations in the 2D digital image and the 3D depth data. In general, block 1007 corresponds to block 307 of method 300.

[00108] While not depicted, it is assumed in method 1000 that the iteration of block 1007 occurs until all identified segments have been processed as described above; hence, block 1007 inherently also corresponds to block 309 of method 300.

[00109] At block 1011, dimensioning processor 120 can then calculate and/or determine dimensions of the physical object represented by the object in the 2D digital image and the 3D depth data. Hence, block 1011 corresponds to block 311 of method 300.

[00110] Thereafter, dimensioning processor 120 can control display device 126 to one or more of: render the digital image; identify the object in the digital image; and render the dimensions of the object.

[00111] Hence, described herein is a device which uses a combination of a digital image and corresponding depth data, respectively acquired using a camera device and a depth sensing device, to determine dimensions of a physical object in a field of view of the camera device and the depth sensing device. By correlating features in the digital image and the depth data, the problems with dimensioning using only a 2D digital image or only 3D depth data can be obviated.

[00112] In the foregoing specification, specific implementations have been described. However, one of ordinary skill in the art appreciates that various modifications and changes can be made without departing from the scope of the specification as set forth in the claims below. Accordingly, the specification and figures are to be regarded in an illustrative rather than a restrictive sense, and all such modifications are intended to be included within the scope of present teachings.

[00113] The benefits, advantages, solutions to problems, and any element(s) that may cause any benefit, advantage, or solution to occur or become more pronounced are not to be construed as a critical, required, or essential features or elements of any or all the claims. The invention is defined solely by the appended claims including any amendments made during the pendency of this application and all equivalents of those claims as issued.

[00114] Moreover in this document, relational terms such as first and second, top and bottom, and the like may be used solely to distinguish one entity or action from another entity or action without necessarily requiring or implying any actual such relationship or order between such entities or actions. The terms "comprises," "comprising," "has", "having," "includes", "including," "contains", "containing" or any other variation thereof, are intended to cover a non-exclusive inclusion, such that a process, method, article, or apparatus that comprises, has, includes, contains a list of elements does not include only those elements but may include other elements not expressly listed or inherent to such process, method, article, or apparatus. An element preceded by "comprises ...a", "has ...a", "includes ...a", "contains ...a" does not,

without more constraints, preclude the existence of additional identical elements in the process, method, article, or apparatus that comprises, has, includes, contains the element. The terms “a” and “an” are defined as one or more unless explicitly stated otherwise herein. The terms “substantially”, “essentially”, “approximately”, “about” or any other version thereof, are defined as being close to as understood by one of ordinary skill in the art, and in one non-limiting implementation the term is defined to be within 10%, in another implementation within 5%, in another implementation within 1% and in another implementation within 0.5%. The term “coupled” as used herein is defined as connected, although not necessarily directly and not necessarily mechanically. A device or structure that is “configured” in a certain way is configured in at least that way, but may also be configured in ways that are not listed.

[00115] It will be appreciated that some implementations may be comprised of one or more generic or specialized processors (or “processing devices”) such as microprocessors, digital signal processors, customized processors and field programmable gate arrays (FPGAs) and unique stored program instructions (including both software and firmware) that control the one or more processors to implement, in conjunction with certain non-processor circuits, some, most, or all of the functions of the method and/or apparatus described herein. Alternatively, some or all functions could be implemented by a state machine that has no stored program instructions, or in one or more application specific integrated circuits (ASICs), in which each function or some combinations of certain of the functions are implemented as custom logic. Of course, a combination of the two approaches could be used.

[00116] Moreover, an implementation can be implemented as a computer-readable storage medium having computer readable code stored thereon for programming a computer (e.g., comprising a processor) to perform a method as described and claimed herein. Examples of such computer-readable storage mediums include, but are not limited to, a hard disk, a CD-ROM, an optical storage device, a magnetic storage device, a ROM (Read Only Memory), a PROM (Programmable Read Only Memory), an EPROM (Erasable Programmable Read Only Memory), an EEPROM (Electrically Erasable Programmable Read Only Memory) and a Flash memory. Further, it is expected that one of ordinary skill, notwithstanding possibly

significant effort and many design choices motivated by, for example, available time, current technology, and economic considerations, when guided by the concepts and principles disclosed herein will be readily capable of generating such software instructions and programs and ICs with minimal experimentation.

[00117] The Abstract of the Disclosure is provided to allow the reader to quickly ascertain the nature of the technical disclosure. It is submitted with the understanding that it will not be used to interpret or limit the scope or meaning of the claims. In addition, in the foregoing Detailed Description, it can be seen that various features are grouped together in various implementations for the purpose of streamlining the disclosure. This method of disclosure is not to be interpreted as reflecting an intention that the claimed implementations require more features than are expressly recited in each claim. Rather, as the following claims reflect, inventive subject matter lies in less than all features of a single disclosed implementation. Thus the following claims are hereby incorporated into the Detailed Description, with each claim standing on its own as a separately claimed subject matter.

Claims

We claim:

1. A device comprising:

a camera device configured to acquire a digital image in a field of view;

5 a depth sensing device configured to acquire depth data across at least a portion of the field of view of the camera device; and,

a dimensioning processor configured to:

determine segments of shapes belonging to an object identified in the digital image;

10 based on respective depth data, from the depth sensing device, associated with each of the segments of the shapes belonging to the object, determine whether each of the segments is associated with a same shape belonging to the object; and

15 once all the segments are processed to determine their respective associations with the shapes of the object in the digital image, compute dimensions of the object based on the respective depth data and the respective associations of the shapes.

2. The device of claim 1, wherein the dimensioning processor is further configured to determine whether each of the segments is associated with the same shape belonging to the object by:

20 identifying adjacent segments of the shapes belonging to the object identified in the digital image; and,
comparing at least a portion of the respective depth data of the adjacent segments.

3. The device of claim 1, wherein the dimensioning processor is further
25 configured to determine whether each of the segments is associated with the same shape belonging to the object by: iterating through all adjacent segments of the

segments to determine whether the adjacent segments can be joined into a larger super-segment of the same shape.

4. The device of claim 1, wherein the dimensioning processor is further configured to determine whether each of the segments is associated with the same shape belonging to the object by: determining, based on the respective depth data, whether adjacent segments are co-located with the same shape.

5. The device of claim 1, wherein the dimensioning processor is further configured to filter out one or more of identified shapes and identified segments that are not associated with the object.

6. The device of claim 1, further comprising a display device, the dimensioning processor further configured to control the display device to one or more of: render the digital image; identify the object in the digital image; and render the dimensions of the object.

7. The device of claim 1, further comprising a memory storing data defining the object, the dimensioning processor is further configured to: identify the object in the digital image by comparing the digital image with the data defining the object.

8. The device of claim 7, wherein the data defining the object is stored in a container in the memory, the container being one of a plurality of containers stored in the memory, each of the plurality of containers defining respective objects.

9. The device of claim 7, wherein the data defining the object comprises stereometric properties of the object, the stereometric properties comprising a set of one or more of quantitative descriptions and qualitative descriptions of the object that uniquely characterize the object in three-dimensions.

10. The device of claim 1, wherein the dimensioning processor is further configured to determine geometry of the object based on relative location of the segments of the object in the digital image and in the depth data.

11. The device of claim 1, wherein each of the shapes comprises one or more of: a primitive shape, a plane, a triangle, a square, a rectangle, a circle, a sphere, and a polygon.

12. The device of claim 1, wherein the dimensioning processor is further configured to preprocess the digital image to one or more of: filter image noise; detect

edges of the object; segment color; determine morphological features in the digital image; and determine orientation of lines in the digital image.

13. The device of claim 1, wherein the dimensioning processor is further configured to preprocess the depth data to one or more of: correct for distortions;
5 correct for interference; filter noise; and convert the depth data to point cloud data format.

14. The device of claim 1, wherein the camera device comprises an RGB (red-green-blue) digital camera.

15. The device of claim 1, wherein the depth sensing device comprises one or
10 more of a structured light camera, a time-of-flight camera, stereo vision camera, an active stereo vision camera, a passive stereo vision camera, a stereoscopic camera, and a LIDAR (light detection and ranging) device.

16. A method comprising:

15 at a device comprising: a camera device configured to acquire a digital image in a field of view; a depth sensing device configured to acquire depth data across at least a portion of the field of view of the camera device; and, a dimensioning processor, determining, at the dimensioning processor, segments of shapes belonging to an object identified in the digital image;

20 based on respective depth data, from the depth sensing device, associated with each of the segments of the shapes belonging to the object, determining, at the dimensioning processor, whether each of the segments is associated with a same shape belonging to the object; and,

once all the segments are processed to determine their respective associations with the shapes of the object in the digital image,

25 computing, at the dimensioning processor, dimensions of the object based on the respective depth data and the respective associations of the shapes.

17. The method of claim 16, wherein the determining whether each of the segments is associated with the same shape belonging to the object occurs by:

identifying adjacent segments of the shapes belonging to the object identified in the digital image; and,
comparing at least a portion of the respective depth data of the adjacent segments.

5 18. The method of claim 16, wherein the determining whether each of the segments is associated with the same shape belonging to the object occurs by: iterating through all adjacent segments of the segments to determine whether the adjacent segments can be joined into a larger super-segment of the same shape.

10 19. The method of claim 16, wherein the determining whether each of the segments is associated with the same shape belonging to the object occurs by: determining, based on the respective depth data, whether adjacent segments are co-located with the same shape.

20. A non-transitory computer-readable medium storing a computer program, wherein execution of the computer program is for:

15 at a device comprising: a camera device configured to acquire a digital image in a field of view; a depth sensing device configured to acquire depth data across at least a portion of the field of view of the camera device; and, a dimensioning processor, determining, at the dimensioning processor, segments of shapes belonging to an object identified in the digital image;

20 based on respective depth data, from the depth sensing device, associated with each of the segments of the shapes belonging to the object, determining, at the dimensioning processor, whether each of the segments is associated with a same shape belonging to the object; and, once all the segments are processed to determine their respective associations with the shapes of the object in the digital image,

25 computing, at the dimensioning processor, dimensions of the object based on the respective depth data and the respective associations of the shapes.

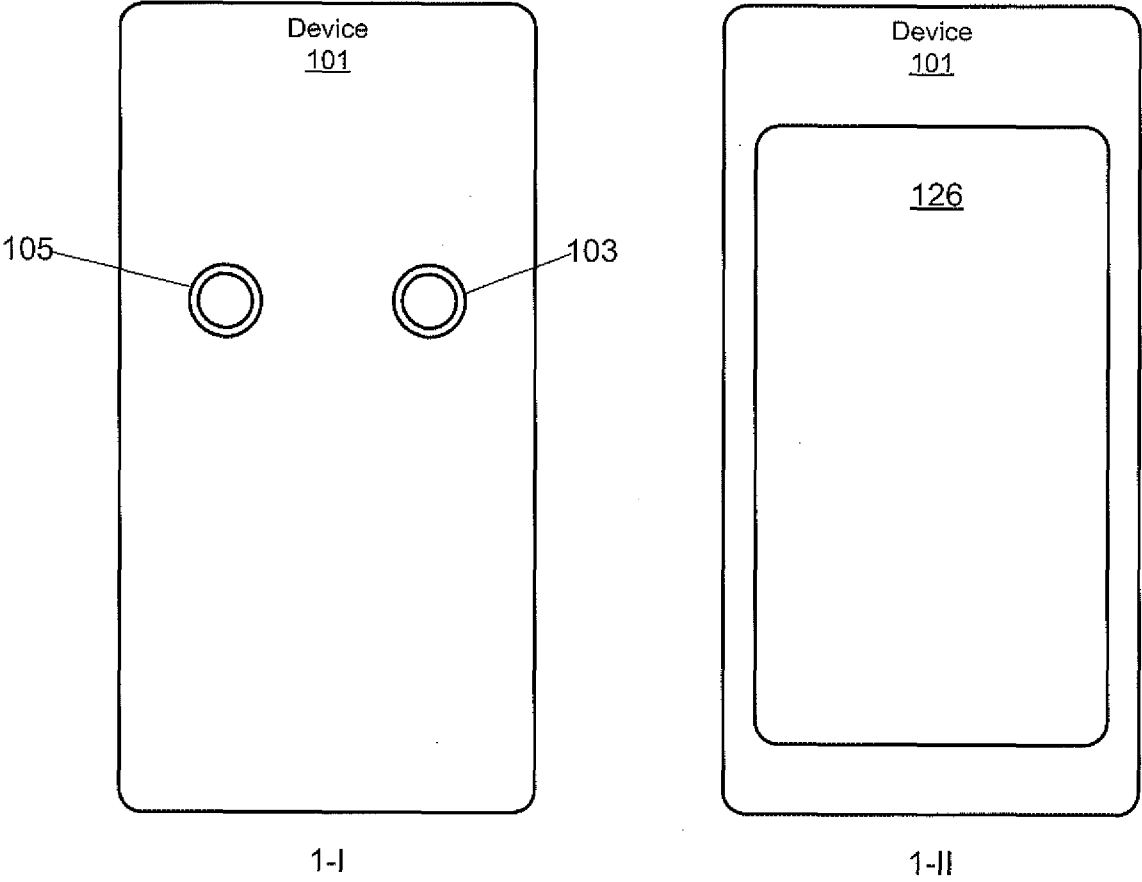


Fig. 1

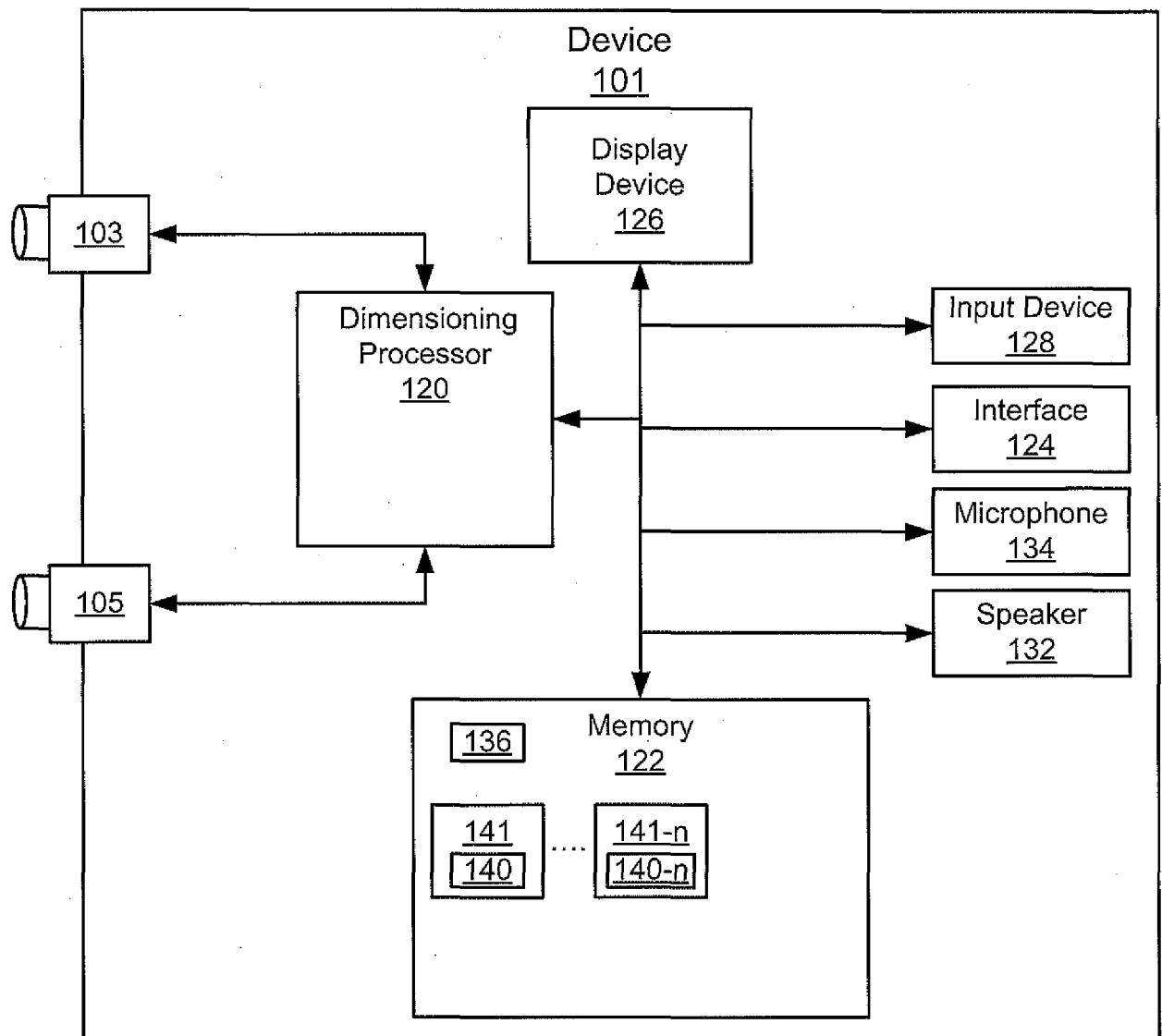


Fig. 2

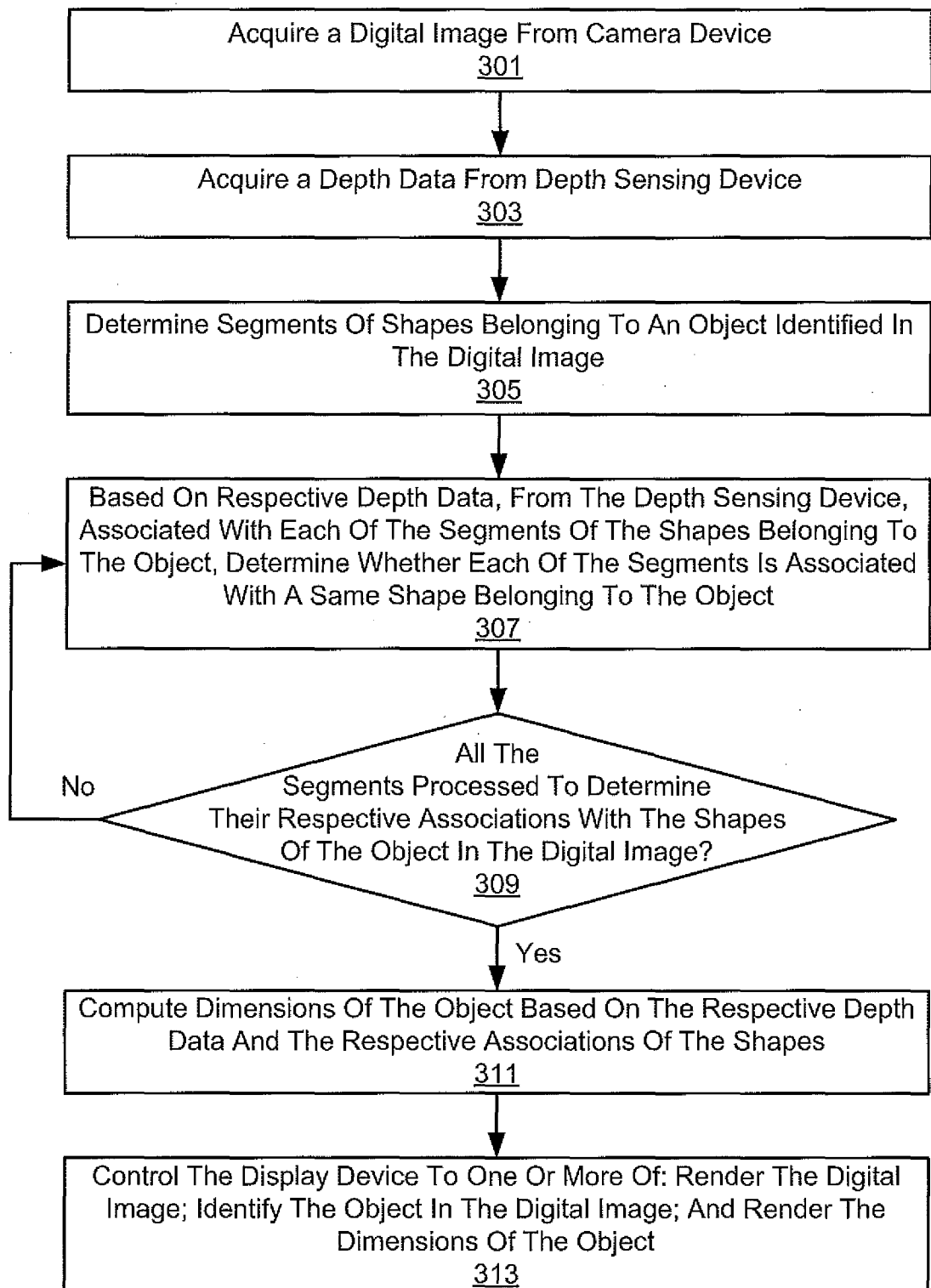


Fig. 3

300

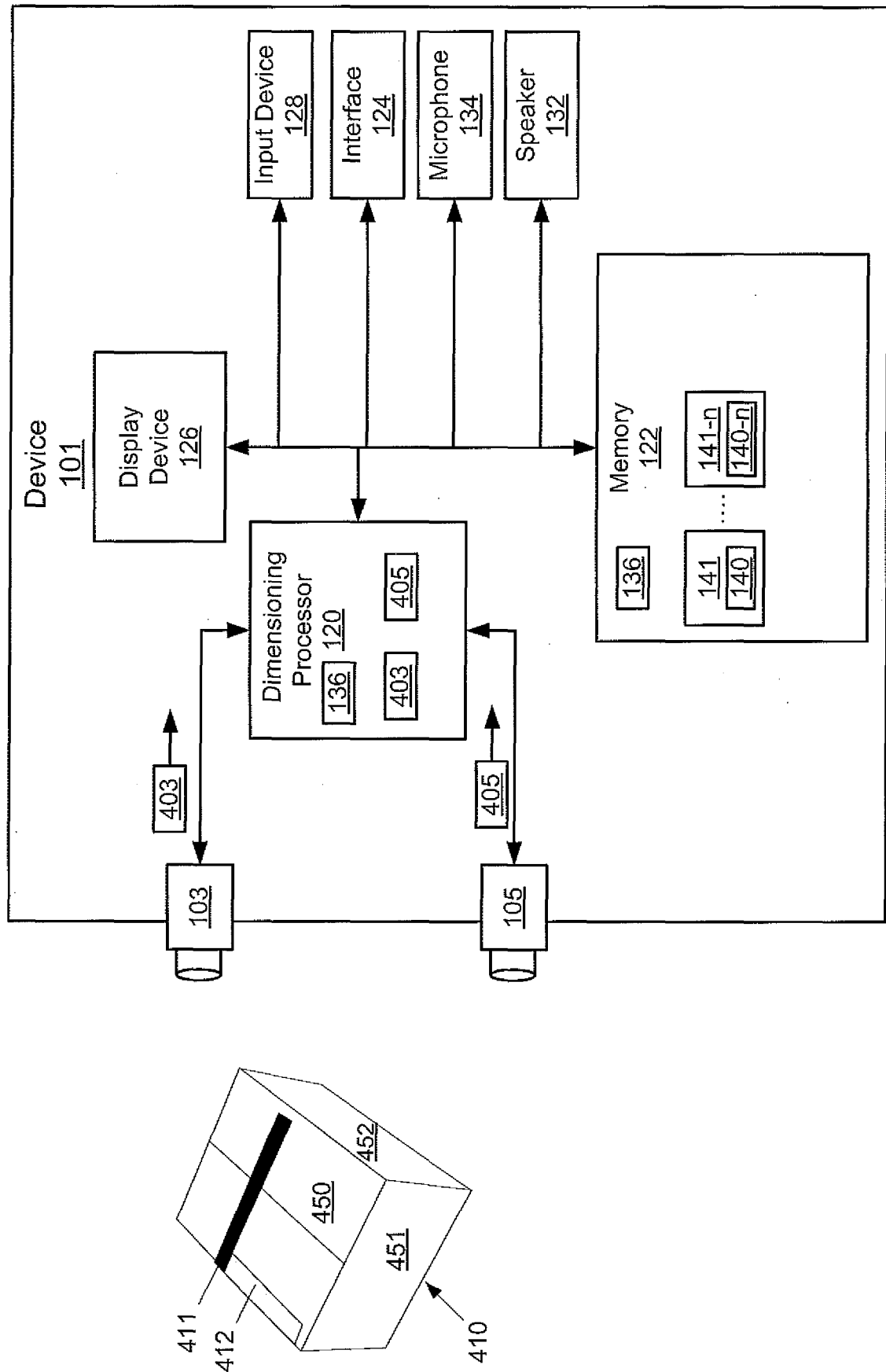


Fig. 4

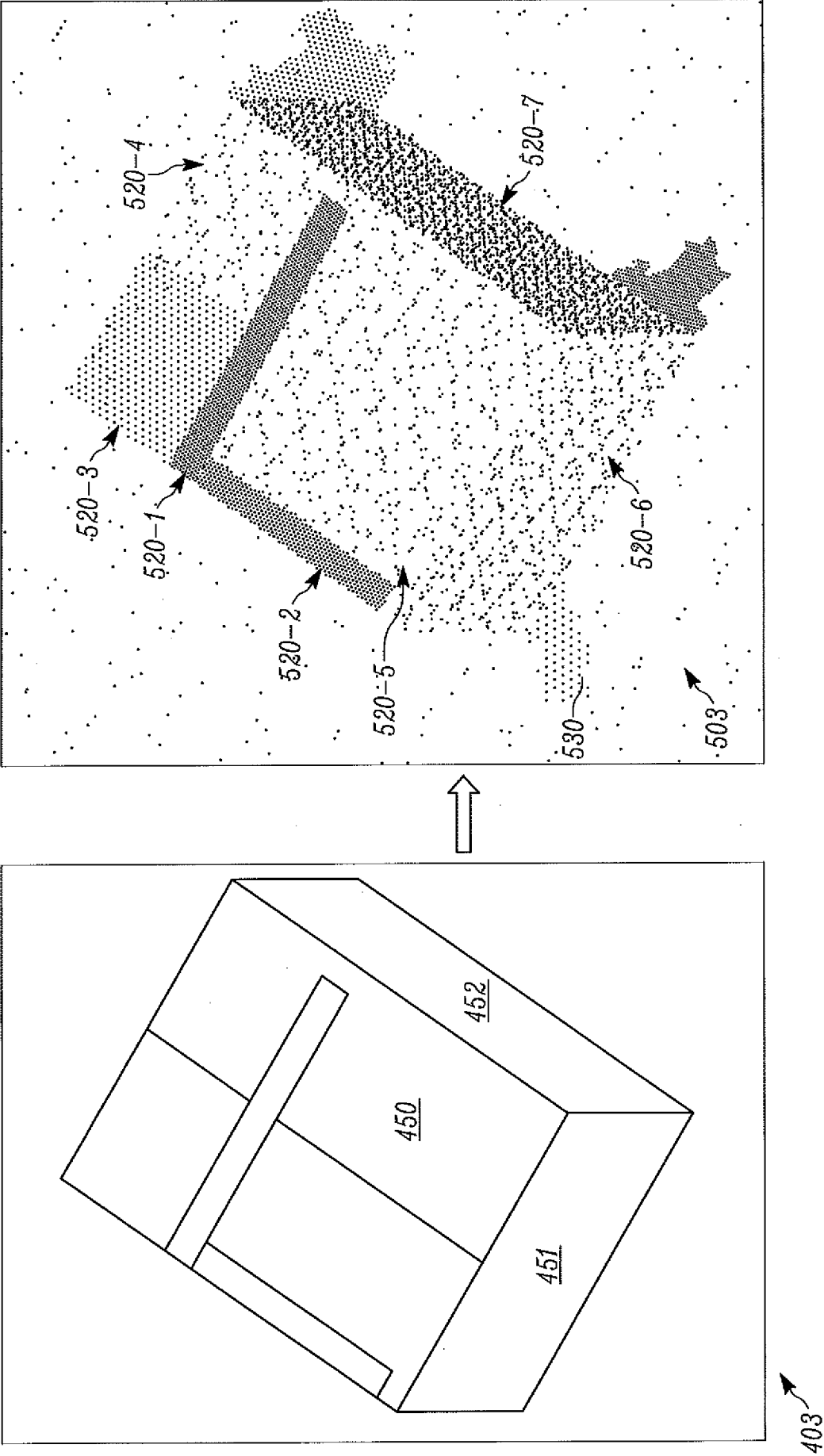


FIG. 5

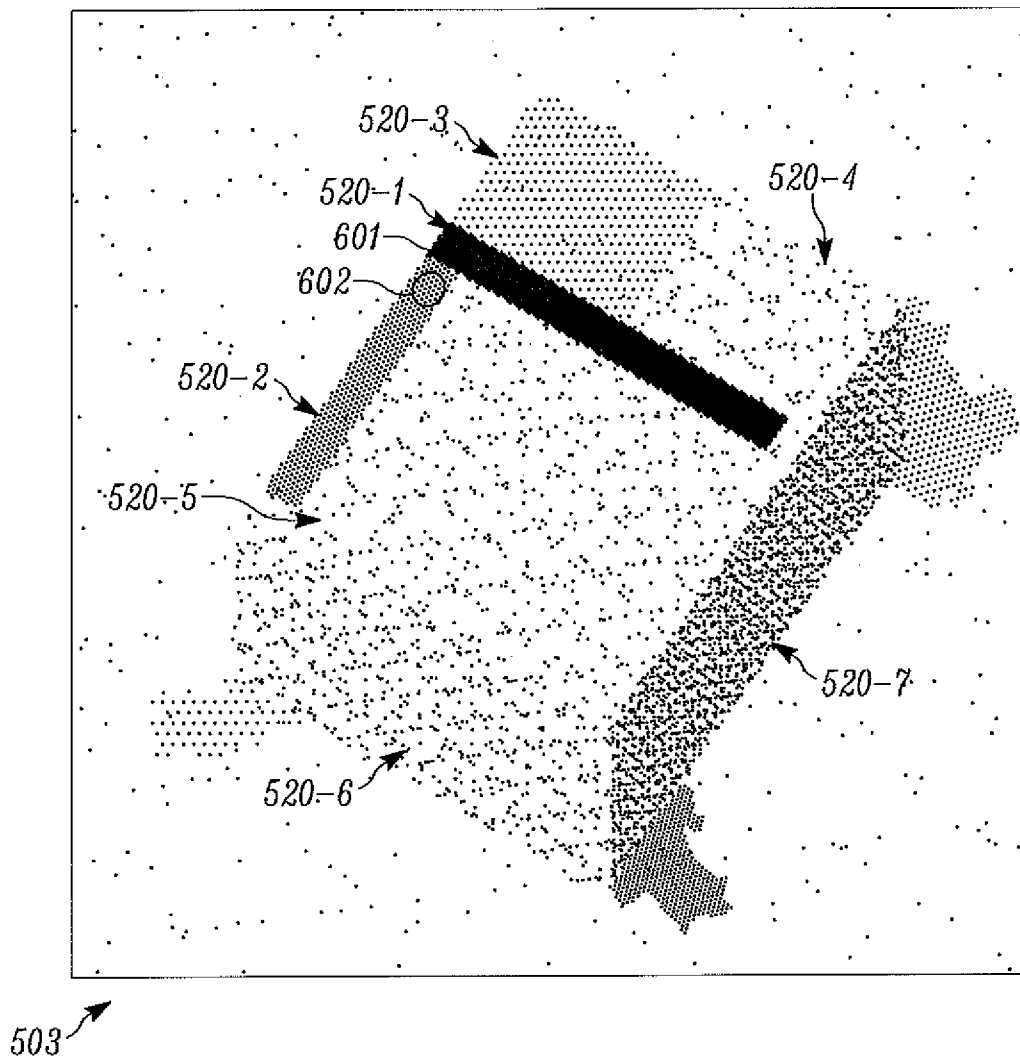


FIG. 6

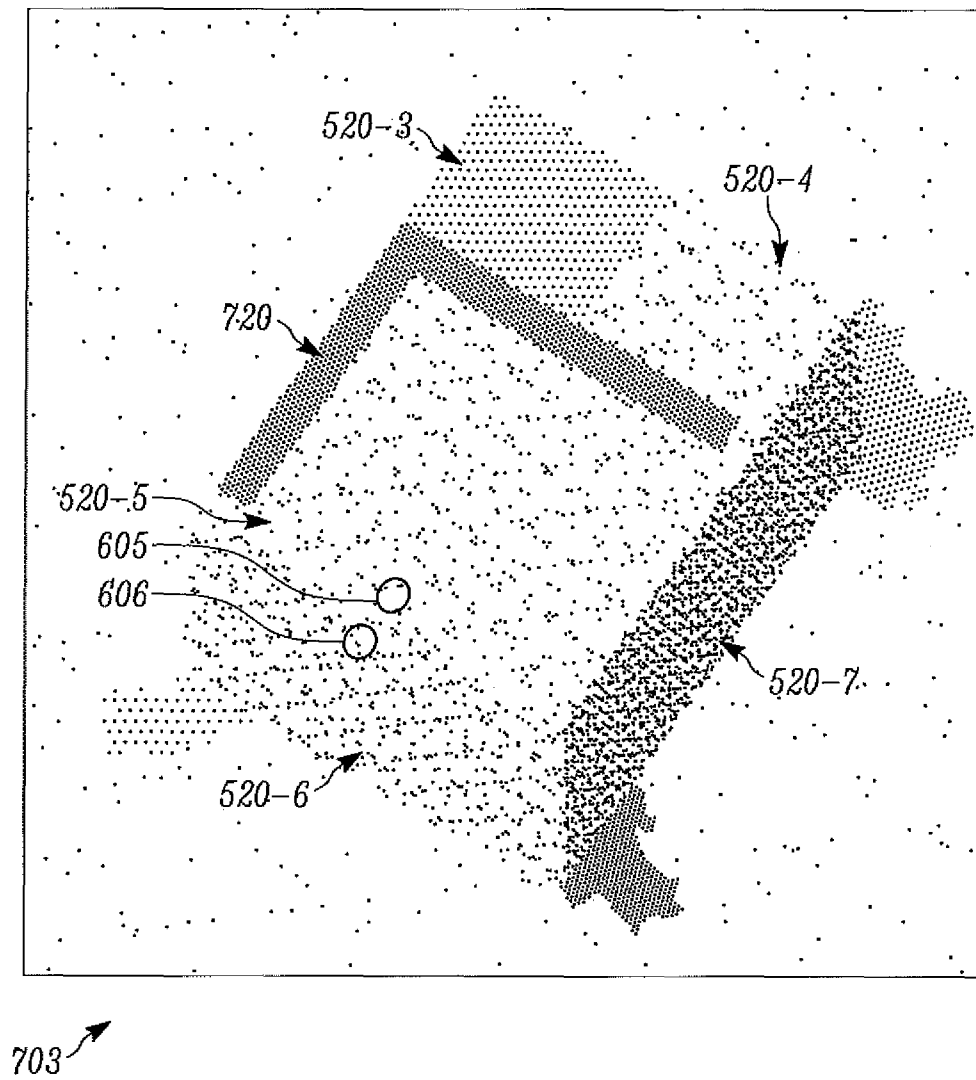


FIG. 7

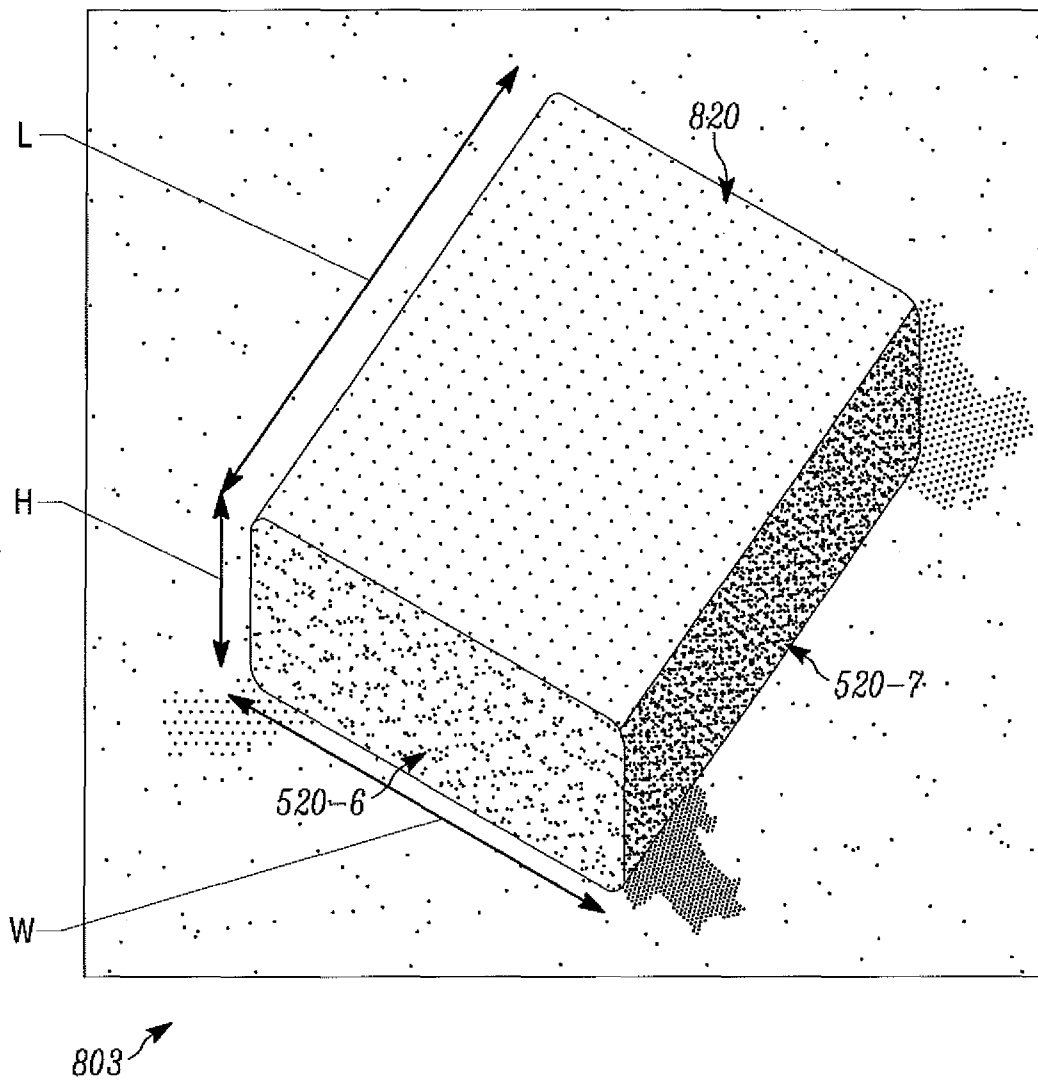


FIG. 8

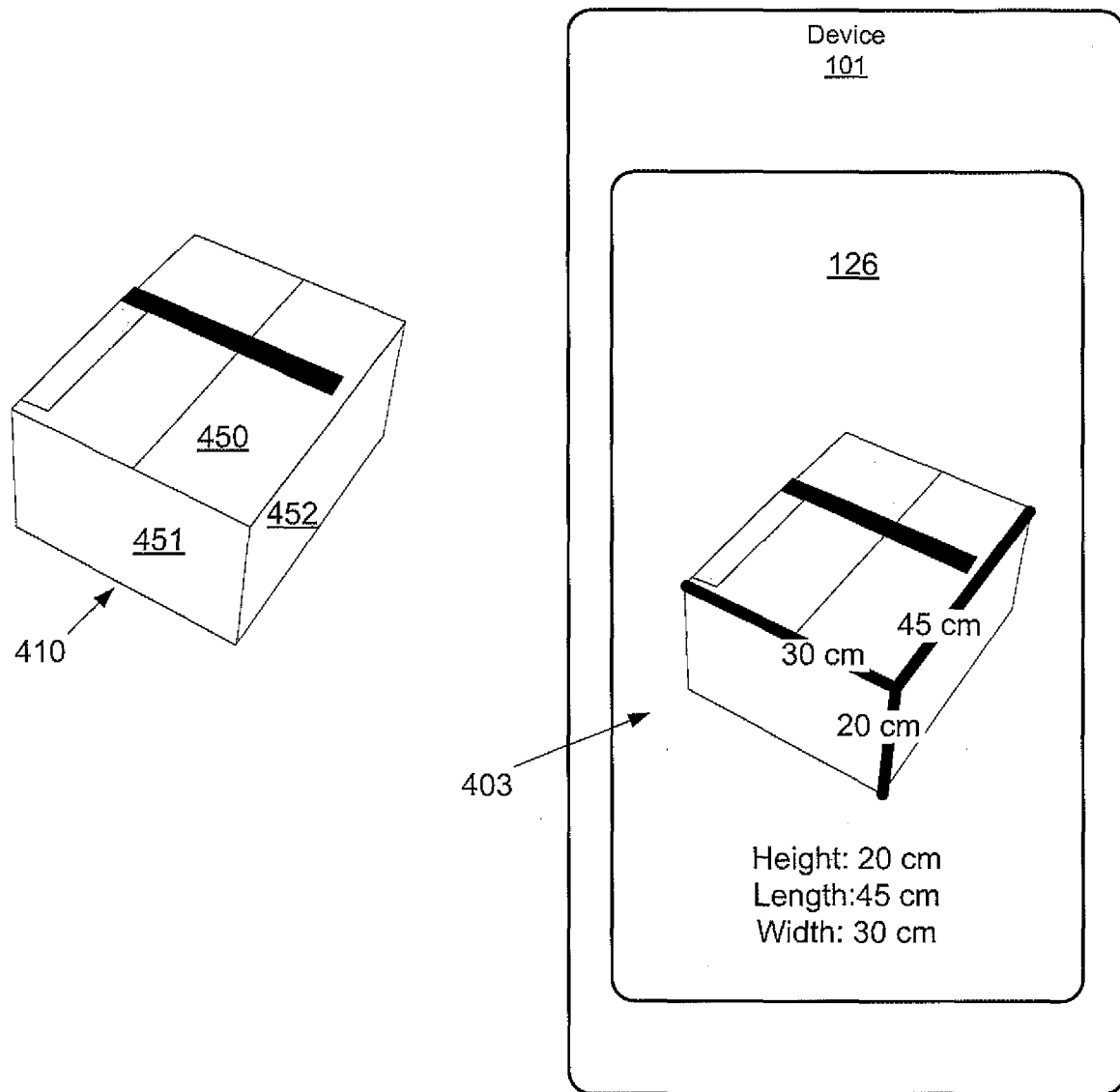


Fig. 9

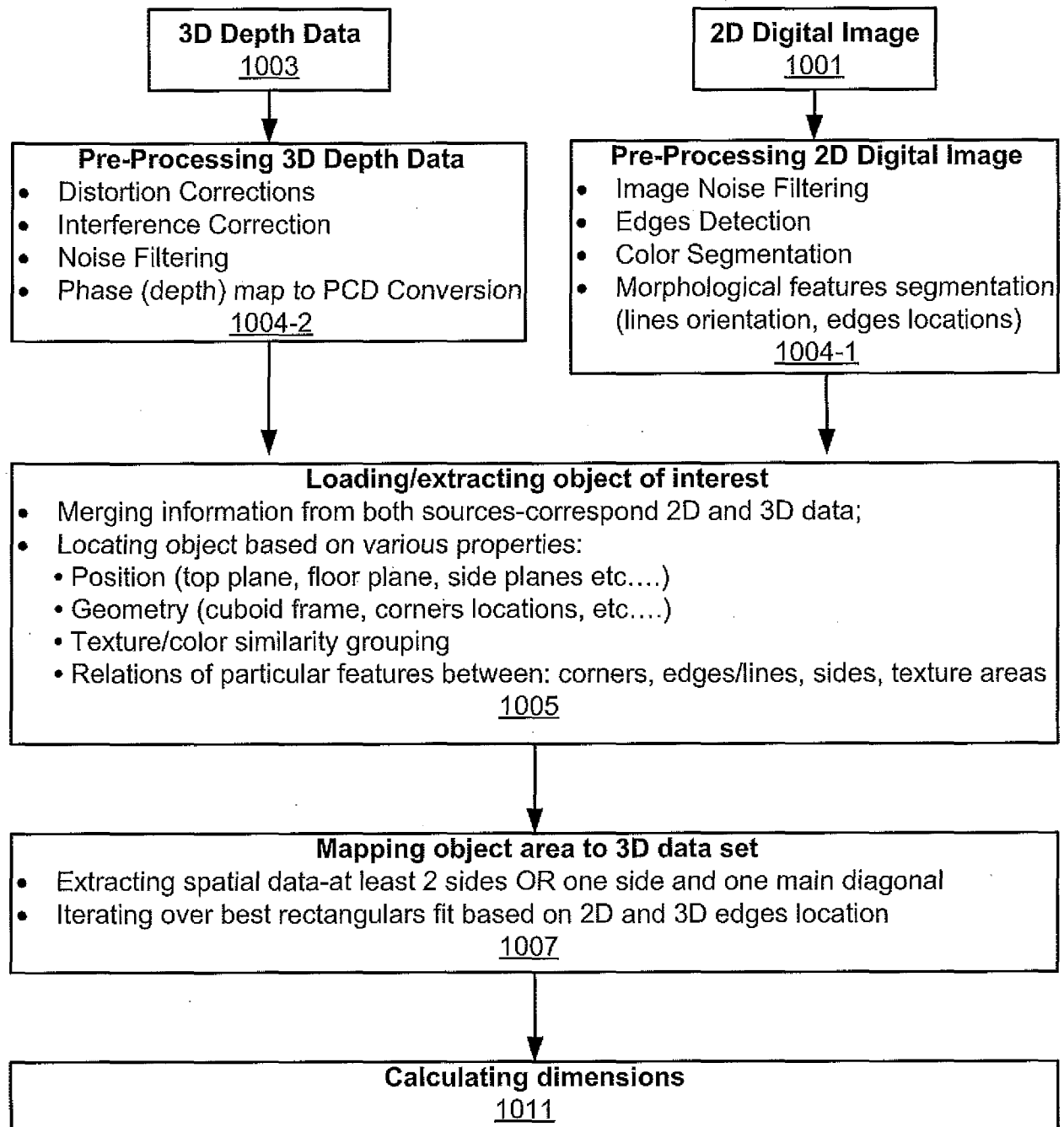


Fig. 10

1000

INTERNATIONAL SEARCH REPORT

International application No

PCT/US2016/068203

A. CLASSIFICATION OF SUBJECT MATTER

INV. G01B11/02 H04N13/02
ADD.

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

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

G01B H04N G06T G06K

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

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

EPO-Internal, WPI Data

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	US 2015/332075 A1 (BURCH REUBEN F [US]) 19 November 2015 (2015-11-19) paragraph [0030] - paragraph [0040]; figure 1 paragraph [0071] - paragraph [0095]; figures 5A, 5B, 6	1-20
X	US 2013/329013 A1 (METOIS ERIC [US] ET AL) 12 December 2013 (2013-12-12)	1-6, 10-20
Y	the whole document	1-20
	----- -/-	



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

14 March 2017

Date of mailing of the international search report

23/03/2017

Name and mailing address of the ISA/

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

Authorized officer

Ardelt, Per-Lennart

INTERNATIONAL SEARCH REPORT

International application No

PCT/US2016/068203

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	KEFEI LU ET AL: "Detecting textured objects using convex hull", MACHINE VISION AND APPLICATIONS, SPRINGER, BERLIN, DE, vol. 18, no. 2, 25 January 2007 (2007-01-25), pages 123-133, XP019487841, ISSN: 1432-1769, DOI: 10.1007/S00138-006-0060-0 the whole document -----	1-20
Y	US 6 995 762 B1 (PAVLIDIS THEO [US] ET AL) 7 February 2006 (2006-02-07) the whole document -----	1-20
Y	US 2011/142328 A1 (FRIEDHOFF RICHARD MARK [US] ET AL) 16 June 2011 (2011-06-16) the whole document -----	1-20

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No

PCT/US2016/068203

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
US 2015332075 A1	19-11-2015	CA 2949088 A1 EP 3143474 A1 US 2015332075 A1 WO 2015175964 A1	19-11-2015 22-03-2017 19-11-2015 19-11-2015
US 2013329013 A1	12-12-2013	US 2013329013 A1 WO 2013170260 A1	12-12-2013 14-11-2013
US 6995762 B1	07-02-2006	NONE	
US 2011142328 A1	16-06-2011	EP 2104907 A2 EP 2400459 A1 US 2008089576 A1 US 2011142328 A1 WO 2008051364 A2	30-09-2009 28-12-2011 17-04-2008 16-06-2011 02-05-2008